National Technical University of Athens
School of Mechanical Engineering
Sector of Industrial Management & Operational Research
Laboratory of Industrial Engineering

Generation Expansion Planning with high shares of Renewable Energy Sources: Single and Multi objective optimization based on Metamodel-assisted Evolutionary Algorithms

PhD Thesis

Constantinos Vrionis

Supervisor: Athanasios Tolis
Associate Professor NTUA

Athens, 2020

National Technical University of Athens
School of Mechanical Engineering
Sector of Industrial Management & Operational Research
Laboratory of Industrial Engineering

Generation Expansion Planning with high shares of Renewable Energy Sources: Single and Multi objective optimization based on Metamodel-assisted Evolutionary Algorithms

PhD Thesis
Constantinos Vrionis

Examination Committee:

1. Athanasios Tolis (Supervisor), Associate Professor NTUA,
School of Mechanical Engineering, (Member of the advisory Committee)
2. Nikolaos Panagiotou, Associate Professor NTUA,
School of Mechanical Engineering, (Member of the advisory Committee)
3. Ilias Tatsiopoulos, Professor NTUA,
School of Mechanical Engineering, (Member of the advisory Committee)
4. Kyriakos Giannakoglou, Professor NTUA
School of Mechanical Engineering
5. Gregoris Mentzas, Professor NTUA,
School of Electrical and Computer Engineering
6. John Anagnostopoulos, Professor NTUA,
School of Mechanical Engineering
7. Athanasios Rentizelas, Assistant Professor NTUA,
School of Mechanical Engineering

Athens, 2020

Acknowledgements

First of all, I would like to thank my supervisor, Associate Professor National Technical University of Athens (NTUA) Athanasios Tolis, for providing me the opportunity to realize this thesis under his supervision. His guidance and advices over this period I shall always remember and be deeply grateful for. Also, I would like to thank him for his patience, time and the corrections provided for the completion of this thesis, and the papers published during these years, which I also greatly appreciate.

My sincere gratitude extends to the members of the Advisory Committee, Associate Professor NTUA Nikolaos Panagiotou and Professor NTUA Ilias Tatsiopoulos, for the valuable comments and suggestions provided by closely reviewing this thesis. Also I would like to thank them for their suggestions regarding its presentation.

Also, it is an honor for me that Professor NTUA Kyriakos Giannakoglou, Professor NTUA Gregoris Mentzas, Professor NTUA John Anagnostopoulos and Assistant Professor NTUA Athanasios Rentizelas accepted to be members of the Examination Committee.

I am indebted to the other member of the research group, Vasilios Tsalavoutis, with whom I worked closely and shared most part of this journey. His determination and hard work had been exemplary. I am grateful for the constructive criticism and discussions throughout our cooperation that were needed for the completion of this thesis. I have been fortunate to have collaborated with him.

At this point I would like to gratefully acknowledge the funding source that supported this thesis. This research is co-financed by Greece and the European Union (European Social Fund-ESF) through the Operational Programme "Human Resources Development, Education and Lifelong Learning" in the context of the project "Strengthening Human Resources Research Potential via Doctorate Research" (MIS-5000432), implemented by the State Scholarships Foundation (IKY).

I would also like to thank all of my close friends and my family. The completion of this thesis would have seemed impossible without their support. Especially, I would like to thank Stephanie for her love, understanding and encouragement during these years. Also, it is hard to overstate how grateful I am to my brother, Yiannis, for his dedicated effort. His has been always beside me to motivate, help and support me. Last but not least, I would like to warmly thank my parents, Beatrice and George, for all the sacrifices they have made for me. I am blessed to have them as my parents. I dedicate this thesis to them.

Ευρωπαϊκή Ένωση
European Social Fund

Operational Programme
Human Resources Development,
Education and Lifelong Learning
Co-financed by Greece and the European Union

ΕΣΠΑ
2014-2020
ανάπτυξη - εργασία - αλληλεγγύη

Abstract

The scope of this thesis is to develop and examine single and multi objective Metamodel-assisted Evolutionary Algorithms (MAEA) for Generation Expansion Planning (GEP) models in the presence of high shares of generation by Renewable Energy Sources (RES). A GEP model may facilitate decision making in mid towards long-term energy scheduling. Commonly, a GEP model is employed to provide a road-map towards an affordable, sustainable and secure operation of a power system. These road-maps are generated in the form of scenarios for the evolution of a power system and are determined by considering both possible investments in capacity additions and the short-term operation of the power system, e.g. to determine an economic and/or environmentally optimal investment plan that is adequate to meet the growing power demand and exhibits a reliable short-term operation. Such a problem can be formulated as an optimization problem.

The pursue for emission-free power system and the increasing shares of RES, triggered by environmental considerations, have led to the introduction of various aspects to GEP models. Such an aspect is to capture economic and technical challenges related to the short-term operation of a power system with increased detail. This could be required to assess the synergy of the conventional generating fleet with the increasing installations of RES. In particular, the variability and uncertainty associated with the latter has been reported to increase the operational flexibility requirements of future generating fleets. It has also been reported that underestimating these requirements could have economic implications on the reliable and efficient short-term operation. Towards this aim, endeavours have been made to integrate, within a GEP model, a more detailed representation of the short-term operation of a power system in terms of spatial, temporal and technical detail. However, the integration of GEP model with such detail can lead to an increased computational cost. Therefore, simplifications are required.

Evolutionary Algorithms (EA) are nature-inspired algorithms which employ stochastic operators to improve a set of candidate solutions. As derivative-free algorithms, EAs can be used as direct search methods; this feature has rendered them applicable for complex optimization problems. In addition, Multi-Objective EAs (MOEA) are well established approaches for Multi-Objective Optimization (MOO). On the other hand, one main limitation is the relatively large number of function evaluations required for the algorithm to converge. This can be binding for optimization problems involving computational costly simulations. For such applications, EAs coupled with Approximating Models (AM) have been developed which are commonly referred to as MAEAs or Surrogate-Assisted EAs. The AMs replace in part the original models and provide an estimate for the adequacy of a candidate solution to reduce the computational burden.

This thesis focuses on MAEA applications for single and multi objective GEP optimization problems that include Simulation Models (SM) for the short-term operation of a power system. The most important contributions of this thesis are the following:

1. A single objective multi-period GEP approach based on MAEAs is presented. The GEP model includes a Simulation Model (SM) to provide an indicator of the cost of the short-term operation. The adopted SM is an optimization model for the short-term operation of a power system including simplifications e.g. spatial detail is not examined. However, it exhibits an increased level of technical and temporal detail w.r.t. the context of long-term planning, and it is adopted to assess on-line the operating flexibility of a candidate installed capacity. The

formulation exploits problem-specific characteristics. This is implemented by employing AMs to provide an estimate of the SM's output and reduce the number of simulations required to achieve a near-optimal solution. The AMs are Radial Basis Functions (RBF). These are built off-line and updated on-line to improve the accuracy of the achieved approximation. Both local and global AMs are built in different stages of the search. Problem specialized operators are developed to enhance the performance of the EA examined which is Differential Evolution (DE). The performance of the MAEA and the problem-specialized operators are assessed. The MAEA achieved satisfactory results based on the performed numerical experiments. Moreover, among the developed problem-specialized operators, a repair heuristic, addressing the constraint nature of the optimization problem, provided the largest improvement in the performance of the base DE algorithm. The impact of including the SM is also examined. The results indicate the importance of capturing operational flexibility requirements to adequately assess the flexibility providers considered as investment options. The metrics employed to examine the accuracy of the attained AMs indicated that a decent approximation had been achieved. Therefore, a visual analysis of the sensitivity of the operating cost towards the installed capacity of the derived near-optimal solution was carried out.

2. A multi-objective static GEP approach based on MAEAs is presented that aims at capturing cost trade-offs emerging for a MOO GEP. Operational flexibility is assessed by an adopted SM that includes technical, spatial and temporal detail. The approach is developed based on MOEA and frameworks for surrogate-assisted derivative-free optimization. Approximation models are employed to address the computational restrictions. RBF and Polynomial Regression (PR) are used as the AMs. These are updated on-line by criteria that prioritize feasibility of the planning constraints, the spatial allocation of the attained training set w.r.t. the search space, and a possible Hypervolume improvement. A local phase is also included in which gradient-based local search is implemented employing local RBF, PR and an ensemble model. The performance of the approach is examined on a MOO benchmark test suite. Numerical experiments are carried out to assess the performance optimization approach on a MOO GEP formulation neglecting the short-term operation and on five MOO GEP formulations including a SM. The latter are repeated for two different levels of temporal detail. The results attained suggest an acceptable performance of the optimization approach w.r.t. the computational restriction. Moreover, the achieved accuracy of the AMs varied among the numerical experiments. The main factors influencing the performance of the AMs are identified. An analysis of the derived cost trade-offs for each of the five formulations examined can provide a detailed evaluation of the impact of a diverse set of alternatives. This could reveal incentives required for strategic energy policy decision making. For example, based on the extreme values of the non-dominated front attained for the considered operating and investment cost functions, a 96% reduction of the investment cost could result in a nearly 40% increase of operating cost.

Decision support tools could facilitate the complex and evolving decision making process of GEP. Economic, environmental and social criteria must be considered along with aspects that are progressively identified as essential. Towards this aim, the developed EA-based approaches have been presented. Despite their heuristic nature, the results suggested that these could be promising tools to support well established state-of-the-art GEP models that could facilitate decision makers, such as investors and energy policy makers, when high shares of RES generation are considered.

Contents

Chapter 1

Introduction

1.1 Context

A transition towards power systems with low-carbon emissions can be a major factor for sustainability. Consequently, the share of energy generated by RES have been increasing in the past years to meet environmental objectives (References [1, 2, 3, 4, 5]). It is desirable that such a transition is achieved in an economic and environmental manner due to the high costs involved. These can emanate from investments in new power plants expected to operate for a number of years and, therefore, involve long-term planning. Long-term planning is a complex task since it considers several aspects that can often be conflicting. Informed decision-making can contribute towards this direction (Reference [6]).

In this context, the operational perspective of future generating fleets to optimally facilitate the integration of high levels of RES has been a research field of growing interest (References [1, 2, 3, 4, 5]). As the examined levels of RES penetration increase, their impact on short-term operation and, consequently, on optimal long-term planning has been reported to be important. This has been attributed (in part) to the fluctuating supply of RES and its interaction with the conventional capacity (References [2, 3, 4]). In particular, efficient integration of generation by RES may result in a reduction of the overall cost of short-term operation. However, high penetration levels may require adequate operating flexibility to enable meeting the demand in an efficient manner (References [7, 8]). Therefore, the adequacy of the conventional capacity is often examined also in terms of operating flexibility (References [2, 4]). This could have a diverse impact on the cost components included within long-term planning and the cost trade-offs that could emerge.

Adequately capturing the impact of increasing shares of RES generation on the short-term operation of a power system within the context of long-term planning is not straightforward (References [7, 8, 2, 4, 9]). It implies that sufficient information (e.g. technical, temporal and spatial detail) is considered regarding the short-term operation of a power system within a long-term planning model (Reference [9]). This information is required to assess the operating flexibility of a candidate generating fleet. In general, even though the importance of such detail may vary based on the examined system (Reference [10]), omitting an assessment of operating flexibility of a future candidate fleet could have an impact on optimal planning.

On the other hand, high complexity and computational restrictions arise to integrate long-term planning models with increased detail of power systems short-term operation. These derive from the requirement of combining two complex models (References [7, 8]), i.e. models for long-term and short-term planning. Therefore, the computational tractability must be also addressed along with the introduction of simplifications (References [9, 5]).

Over the years, EAs have become one of the well-established optimization techniques in a variety of fields. Applications of EA presenting satisfactory results have been also reported for optimization problems related to power systems (References [1, 11]). In general, EAs are nature-inspired derivative-free algorithms that rely on stochastic operators to improve a number of candidate solutions over a series of iterations. These have been mostly appropriate for optimization problems involving complex (e.g. non-convex and/or non-differentiable) objective and constraints functions. Moreover, EAs have been widely recognized as a promising approach for MOO problems (References [12, 13]). Despite some advantages, EAs exhibit limitations. For example, the absence of gradient information (when such is available) can render EAs less applicable for optimization problems that can be easily addressed by classical optimization methodologies. In addition, EAs can be rendered impractical for optimization problems involving computationally expensive objective and/or constraint functions due to the low number of available evaluations (References [14, 15]). Such problems can include a costly simulation model or an expensive experiment. A common approach is to enhance EAs with AMs (i.e. Metamodels or Surrogate models). These, usually, are termed in the relevant literature as MAEAs. Such approaches include AMs that are used to provide a computationally cheap indicator of the adequacy of a solution and replace in part the computational expensive model (Reference [16]). The context, scope and background of this thesis are elaborated in Chapters 2 and 3 where background on GEP and EAs are provided.

1.2 Motivation and Aims

The motivation of this thesis emanates from the growing computational intensity of long-term planning of power systems. These can be attributed to the importance of assessing the operating flexibility of a candidate future generating fleet and to a number of new aspect included in a GEP regarding the operation of a power system (References [1, 2, 3, 4, 5]). EA-based approaches for long-term planning or short-term planning have been suggested within the relevant literature. However, EA or MAEA-based approaches for long-term planning exhibiting increased detail of the short-term operation by including a SM had not been identified. This could be attributed to the computational cost associated with the latter. Therefore, the main scope of this thesis are MAEA-based approaches for single and multi objective optimization problems relevant to the context of long-term planning, specifically GEP models, including SMs for the short-term operation. Focus is directed on including a SM within the optimization process. Developing an optimal SM is beyond the scope of this thesis, therefore SMs are adopted with minor modification. These are included to assess a candidate generating fleet by accounting for dynamics emerging from short-term operation. The impact of including technical, temporal and spatial detail of the short-term operation in long-term models has been highlighted in many recent studies (References [1, 2, 3, 4, 9]). Therefore, an attempt is made to capture operating flexibility provided both by thermal (e.g. conventional) and non-thermal units. These increase the computational cost and, consequently, MAEA are examined.

The MAEAs are developed for single and multi objective optimization and are based on surrogate-assisted derivative-free optimization frameworks identified in the literature. Modifications driven by the specific application are included to enhance their efficiency.

The decision making process is highly complex since it involves several factors that must be considered (References [1, 2]). For example, environmental, such as the reduction of Green House

Gas (GHG) emissions, economic, such as profitability of the generating fleet, and/or social, such as affordable and accessible energy, objectives are usually considered. The developed approaches aim to serve as decision support tools to enhance decision making. Such an example is the assessment of incentives for increasing the penetration levels of RES by examining possible RES support schemes for which benefits or costs can be transferred to the consumers.

An enhanced assessment of a long-term plan may provide benefits to the investors, regulatory authorities and consumers. These can arise from improved signals provided for the decision making process. For example, adequate signals may influence the allocation of an investment budget among difference investment options. Regulatory authorities could be influenced to provide incentives for increasing the efficiency of the power system. An efficient and reliable operation of a power system can benefit consumers.

This thesis focuses on MAEA's applications for single and multi objective GEP optimization problems that include Simulation Models (SM) for the short-term operation of a power system. A brief overview of the most important contributions of this thesis are the following:

1. In this thesis, a multi-period GEP approach based on MAEAs is presented for enhancing optimal GEP (Reference [17]). The approach focuses on high technical and temporal detail for the context of long term planning. This had been motivated by the importance of adequately capturing the operating flexibility requirements of a future candidate generating fleet. For this task, a SM is included within the model. Due to the computational restrictions a MAEA approach had been developed. The AMs, based on RBF, are employed to provide an estimate of the adequacy of a candidate generating fleet. Therefore, computationally expensive simulations are replaced in part by computationally cheap cost indicators to reduce the number of simulations required to achieve a near-optimal solution. Both local and global AMs are built in different stages of the search. Moreover, the optimization approach exploits problem-specific characteristic. Also, problem specialized operators are developed to enhance the selected EA performance that is a DE algorithm. The results indicate that satisfactory results are observed based on a series of numerical experiments carried out. A repair heuristic, addressing the constraint nature of the optimization problem, provided the largest improvement in the performance of the base DE algorithm. In addition, the metrics employed to assess the quality of the attained AMs, in terms of accuracy, suggested that a decent approximation had been achieved. Therefore, a visual analysis of the sensitivity of the operating cost towards the installed capacity of the derived near-optimal solution is carried out. This could serve as a simple approach to assess the sensitivity of the operating cost towards the installed capacity of an attained near-optimal solution. The level of technical detail included in a SM is examined by comparing two different SMs that were employed in the approach. The results suggest that difference in the technical detail of the SMs may influence the derived investment decisions, anticipated generating mix and cost. Adequately capturing operational flexibility requirements within the context of long-term planning could provide improved signals to decision makers.

2. A multi-objective static GEP approach based on MOO-MAEAs is presented that aims at capturing cost trade-offs emerging for a MOO GEP. A SM including technical, spatial and temporal detail is adopted to assess the operational flexibility on-line. The approach is developed based on MOEA and frameworks for surrogate-assisted derivative-free optimization. The AM selected are the RBFs and the PR. These provide an estimate of the values of the

objective functions considered and are employed to address the computational restrictions. The criteria employed for updating the AM on-line prioritize feasibility of the planning constraints, the spatial allocation of the attained training set w.r.t. the search space, and possible Hypervolume improvements. A gradient-based local search is implemented employing local RBF, PR, and an ensemble model that serve as the local phase. The performance of the MOO-MAEA approach is examined on a benchmark test suite. Also, numerical experiments are carried out to assess the performance of the optimization approach on a MOO GEP formulation, neglecting the short-term operation, and on five MOO GEP formulations including the adopted SM. The results attained suggest an acceptable performance of the optimization approach w.r.t. the computational restriction. Moreover, the achieved accuracy of the AMs had varied among the numerical experiments. Main factors influencing the performance of the AMs are identified. An analysis of the derived cost trade-offs for each of the five formulations examined can provide a detailed evaluation of the attained near-optimal solutions and the impact of a diverse set of alternatives. This could reveal incentives required for strategic energy policy decision making.

1.3 Outline

This thesis consists of six Chapters. Chapter 2 and 3 provide the background for this thesis and elaborate on its scope and aims. In particular, Chapter 2 provides an introduction and the relevant literature to GEP with emphasis on long-term planning of power systems involving high levels of RES penetration. Basic characteristic of a GEP model are presented. Challenges arising due to the increasing shares of RES generation are discussed. An introduction to EAs and MAEAs is provided in Chapter 3 presenting also the relevant literature. Chapter 4 presents a single-objective multi-period GEP approach based on MAEAs. The approach includes problem-specific operators that are examined regarding their impact on the algorithm performance. The impact of including the SM and the technical detail is also examined. In Chapter 5, a static MOO GEP approach based on MOO-MAEAs is presented. The approach aims towards examining cost trade-offs arising in the presence of increased shares of generation by RES by a SM. A series of numerical experiments are conducted to examine the efficiency of the optimization approach and identified limitations are presented. Chapter 6 provides the concluding remarks of this thesis.

Chapter 2

Generation Expansion Planning and increased shares of Renewable Energy Sources

This Chapter provides a background to the short and long term optimization models that are considered in this thesis. Specifically, a brief introduction to long-term planning with a focus on GEP is provided. Then, the considerations regarding short-term planning, with a focus on the Unit Commitment Problem (UCP), are discussed. Moreover, the main challenges regarding an efficient integration of RES generation are mentioned. Lastly, the relevant literature on GEP models including aspects of the short-term operation is presented, which is the focus of the thesis. Overall, the Chapter aims to elaborate on the motivation and scope of this thesis by providing relevant background.

2.1 Brief note on long-term planning and Generation Expansion Planning

Long-term planning is very broad and different categorizations based on different criteria such as scope, focus, methodologies, and level of detail, can be made (Reference [9]). Such models are mainly intended for generating scenarios and pathways to be analyzed w.r.t. to the scope under examination. Focus is directed on long-term planning for which the scope is restricted to the electrical power sector and for which the methodology is restricted to optimization models. More specifically, this thesis is restricted to power planning models that focus on investment scenarios in generation capacity, for the electrical power sector, accounting for short-term operation in the presence of high shares of RES generation.

There are other categories with broader scopes such as energy system optimization models, energy-economy models and integrated assessment models that are not discussed in this thesis. Such models may account for other sectors of the energy system (e.g. transportation), consider the interaction of the energy system and the entire economic system, or have a large-scale energy-economic global scope. Furthermore, they may focus on scenarios for multiple regions/countries, or global scenarios, for a few decades, or a century. Interested readers are referred to Reference [9] for greater detail.

One of the main focuses of such long-term power planning optimization models is to examine a power system's ability to satisfy the anticipated demand in a long-term planning horizon (years to decades). Main considerations are the generating fleet and the transmission network which should be both analyzed (Reference [2]). These commonly are examined in terms of technical, economic and environmental efficiency, reliability and flexibility.

Specifically, models that focus on analyzing the existing generating fleet and its possible expansion, are commonly referred to as GEP models (Reference [18]). In this case, possible investments in capacity additions and/or the introduction of new generating units are examined considering economic, reliability, and environmental criteria. One of the main factors that drives these additions is, usually, the anticipated demand growth. In addition, they may be required due to the aging of the existing generating fleet which can lead to the decommissioning of old or inefficient units. Moreover, they could be driven by policy factors, e.g. a low-carbon transition motivated by the environmental factors.

Correspondingly, models with a focus on the analysis of the transmission network and its possible expansion, i.e. transmission expansion planning (TEP), are commonly referred to as TEP models (Reference [18]). In this case, the analysis focuses on the adequacy of the existing transmission network in terms of efficiency and reliability. Similarly to GEP, an expansion in the transmission network can be driven by the aging of the existing network and the anticipated demand growth. However, there are also other motivations for TEP such as to facilitate RES integration, facilitate power exchange among producers and consumers, and to create interconnections between isolated regions.

Several models include both GEP and TEP (GEP-TEP) as it is important that both must be analyzed within the context of long-term planning. It has been discussed (Reference [2]) that such models may provide alternative and beneficial expansion plans and the examination of GEP and TEP should be considered at least in parallel (even though not equally optimal).

The following Sections elaborate on GEP which is the focus of this thesis and TEP and GEP-TEP are beyond its scope. Greater detail on the aforementioned can be found in References [2, 1, 4, 18, 5].

2.1.1 Generation Expansion Planning

GEP models stand as one of the most important tools to support decision-making in power sector long-term planning (References [2, 1, 4, 6]). Moreover, recent challenges introduced to power sectors, such as new economic, technical, environmental, and regulatory issues, have gradually forced a new GEP models to emerge that deviate from the traditional GEP framework. Categorizing GEP models and approaches can be made in different ways considering different criteria. For example, Reference [1] provided a comprehensive review of the GEP problem. They reviewed and classified many studies on GEP from the perspective of five different factors, i.e. the liberalization of the electricity industry, climate change, and environmental issues, recent revolution in generating technologies, current regulatory policies, and emerging techniques in the fields of optimization and modelling.

Such factors have motivated the development of many GEP models. In the following Sections, he basic questions associated with the GEP process, the objectives and perspective of the planner, the traditional approach for the inclusion of operation and uncertainty are provided.

Generation Expansion Planning: Basic Decisions

Commonly, four basic questions are associated with a basic GEP: what, how much, where and when (References [1, 4, 18]). In particular, a GEP model should provide the answers to the type and size of the suggested capacity additions, where these should be located, and when should the investments be carried out. Each of these questions is important to optimally identify the capacity additions that adequately meet the demand growth over an examined long-term planning horizon.

- Type: Different investment options may exhibit different techno-economic characteristics. The efficiency and reliability of the system are highly dependent on the synergy of the existing and new capacity and consequently its techno-economic characteristics. For example, the techno-economic characteristics of non-dispatchable RES generating technologies differ from the ones of fossil fuel units.

- Size: Due to the different techno-economic characteristics of the investment options the optimal size of capacity additions in each such option can differ. For example, the system could benefit by larger/lower numbers of capacity additions, to the existing capacity, of a generating technology type A in comparison to B. Moreover, possible capacity addition sizes could be restricted by techno-economic characteristics.

- Location: The importance of location is, in general, associated with the network limitations and the spatial distribution of the supply and demand (Reference [18]). For example, it may be important to suggest the location of the optimal capacity additions to reduce possible transmission losses. In addition, restrictions may emerge due to possible congestion, especially for cases of high shares of generation by RES (Reference [19]). Therefore, attaining information regarding the location requires a representation of the transmission system within GEP models. Models including the aforementioned representation are referred to as network-constraint. On the contrary, when the representation of the transmission system is neglected the model is referred to as a single-node model. In this case, both supply and demand are assumed to be co-located on a single virtual node.

- Timing: A GEP model may consider a planning horizon spanning over a few years or decades and the timing of implementing the investment options may be important. Therefore, there are two main categories static models and dynamic models (Reference [18]). For a static model, it is assumed that capacity additions are ordered or constructed at a single discrete step during the beginning of the planning horizon, e.g. first year, and the final year of the planning horizon is assumed as the target year. For dynamic models, a number of discrete steps throughout the planning horizon are assumed (e.g. every five years). Within each step an investment stage and a target year are then considered. Hereafter, dynamic models shall be referred to as multi-period.

Generation Expansion Planning: Frameworks based on Industry structure

Based on the categorization made in Reference [1], there are different frameworks to approach GEP w.r.t. the liberalization of the electricity industry structure. The first is the centralized approach where a system-wide analysis is considered. It is carried out from the perspective, or under the assumption, of a central planner which is interested in the most efficient expansion plan. This case considers the traditional structure (monopoly - vertically integrated structure) of the power sector.

The second considers a market-based approach. It is implemented from the perspective of profit-oriented agents that aim to maximize their own anticipated profits by expanding their production capacity. Therefore, such agents are not responsible to meet the demand but are interested in improving their participation, aiming at higher profits, in the electricity market. This case considers a decentralized electricity industry structure (liberalized electricity industry). An intermediate case arises when the transition from a monopoly towards a liberalized market is considered. In such, the GEP problem may be addressed as a profit maximization problem, from the perspective of the independent power producers, or a cost minimizing problem, from the perspective of utilities which purchase generation from the independent power producers.

A central planner framework is adopted in this thesis. The central approach seeks for the investments in capacity additions that satisfy the defined objective and limitations in an optimal manner. Common objectives are cost minimization (e.g. total, generation, and/or investment cost) or the maximization of social welfare. Despite satisfying the demand in a reliable manner, other limitations may be also taken into consideration (e.g. environmental factor, regulatory objectives). However, under a liberalized structure, a central approach does not necessarily imply that the central agent actually implements the investments but rather provides incentives to private investors to build them (Reference [18]) .

Generation Expansion Planning: Single and Multi objective optimization frameworks

Traditionally, GEP models consider a single-objective optimization (SOO) formulation (Reference [1]). However, long-term planning could include many other objectives. For example, the environmental, operational, regulatory, and/or social aspects are also important (References [1, 4]).

In general, a least-cost GEP model could aim at identifying the optimal capacity additions by minimizing an aggregated cost function representing the system's total cost (References [1, 4]). In case more than one objectives are considered, then these can be included in a single aggregated function. Including the aforementioned within a least-cost formulation requires an aggregated weighted function for which the numerical differences must be scaled. This also could require determining weights/preferences a priori which is not always trivial when the included objectives cannot be easily expressed in monetary terms (e.g. monetary value of RES penetration) or their preferences cannot be easily quantified (e.g. reducing the operating cost and achieving regulatory goals). Moreover, the latter may include objectives that may be conflicting (e.g. an increase in the total investment cost could, possibly, reduce the operation cost). Moreover, the output of a SOO problem does not provide a set of solutions that reveal trade-offs among such terms. In these cases, where different objectives are conflicting but included within a single function, the optimization framework could be perceived as a weighted sum method. Moreover, some objectives can be imposed as a set of limitations represented as constraint functions or penalty terms. Common examples of such objectives expressed as limitations are the environmental and/or reliability factor (Reference [20]).

Alternatively, multi-objective GEP frameworks have been proposed and their popularity has been recently increased. Reference [3] attributes this increase to the growing importance of parameters such as RES, GHG emission, and energy security targets. Therefore, Multi-Criteria Decision Making (MCDM) methods have been developed to support decision-making in the presence of multiple and conflicting objectives. Commonly, these approaches are categorized to Multi-attribute decision-making and Multi-Objective Decision Making (References [4, 1]).

In multi-attribute decision-making approaches, discrete and predefined alternatives are compared based on a set of decision criteria (Reference [4]). This aims at ranking such alternatives based on the overall preferences when the criteria are considered simultaneously. Reference [21] discusses multi-attribute decision-making approaches applied to energy planning based on three main categories; namely value measurement methods, goal programming, and outranking models.

On the contrary, Multi-Objective Decision Making focuses on cases where the number of decision alternatives is large. Therefore, an optimization problem is formulated, commonly, considering the criteria as objective functions and possible limitations as constrained functions. The output of such a MOO problem is a number of possible solutions with equivalent quality w.r.t. the examined conflicting objectives. In general, the most suitable solution must be determined a posteriori based on a decision maker's preference as each may be more suitable given different weights/preferences (Reference [20]). Therefore, the aforementioned solutions can be analyzed in terms of trade-offs during decision-making.

Different GEP frameworks have focused on different objectives, such as minimization of cost, GHG emissions or risk related objectives, maximization of economic feasibility or reliability-related objectives (References [1, 3, 4]). Reference [1] identifies the analytical hierarchy process, the utility theory, the fuzzy set theory, the normal boundary intersection, the preference-order ranking and the graphical representation using trade-off curves as the most popular approaches. It also discusses the relatively low number of MOO based studies, in comparison to single-objective ones, despite the recent advances in multi-criteria decision-making.

Generation Expansion Planning: Uncertainty and risk assessment

Long-term planning is subjected to uncertainties. Therefore, investment options should be evaluated within GEP based on a series of risks posed by each candidate generation technology. Reference [2] summarized the basic risk components in the GEP process to the following categories: economic, political, regulatory, environmental, technical, social, and climate. Also they provide a relevant literature review regarding the state-of-the-art on risk assessment in GEP. Moreover, techniques employed for considering uncertainties induced by RES in the planning process are discussed in Reference [4]. In general, deterministic GEP models assume perfect information regarding such uncertainties while stochastic GEP models attempt to account for the uncertainty of the GEP problem by employing uncertainty modelling approaches (e.g. stochastic programming).

2.1.2 Generation Expansion Planning and Short-term operation

The main goal of a short-term model is to support decision-making towards the optimum scheduling of available generating units in a power system and efficiently balance supply and demand in a short-term planning horizon (one day up to two weeks). In practice, many electricity markets are structured as a Day-Ahead market and a balancing and ancillary market (Reference [18]). Decision-making is usually implemented by the market operator or an Independent System Operator (ISO). Moreover, decision-making is implemented by privately-owned companies that submit offers (supply and demand) in liberalized markets. From the ISO perspective, the main considerations are the reliable, economic and environmental operation of the power system. Reliability is considered to ensure that supply and demand in real-time operation are met. Some technical characteristics/restrictions of generating units require that the available units are committed in advance to attain an economical and efficient operation. Consequently, sufficient reserves should be available in the Day-Ahead market for the real-time operation to account for uncertain events (e.g.

transmission or generation outages, forecasting errors in demand or supply). Supply and demand must be economically met at all times including also the possible imbalances. Based on the structure of an electricity market, generation cost (partly) determines the marginal price. Therefore, the inefficient operation can impact the cost that is transferred to consumers and the profits of investors. Consequently, techno-economical characteristics must be accounted for to determine a feasible and optimal generation schedule.

Short-term operation: the merit order

Traditionally, the operational perspective, i.e. short-term operation, within the context of long-term planning must be accounted for. However, until recently, historic load patterns have been characterized by highly predictable and relatively slow time dynamics (Reference [8]). Therefore, a common assumption made was that the demand and variable costs are the main forces affecting the dispatch in different operating conditions. Specifically, detailed representation of the short-term operation, including for example technical limitations of the generating fleet (e.g. unit ramping limits) and/or operating requirements (e.g. operating reserves) had been mostly ignored. This assumption enables the employment of a simpler representation of both the generation cost and dispatch scheduling. One of the main approaches for capturing the short-term operation is based on identifying a subset of operating conditions (Reference [18]). Commonly, these are derived based on clustering and the load-duration-curves (LDC). A LDC is, in general, the aggregated and sorted in descending order electricity consumption/demand (load). An important characteristic of the LDC approach is that it maintains information of the historical data, e.g. correlation among load levels in different locations, however, chronological information is not preserved, e.g. load patterns, (Reference [18]).

Based on the aforementioned, the dispatch of an assumed available conventional generating fleet can be determined based on the merit-order defined by their variable costs and using the LDC. In particular and from an electricity-economic perspective of liberalized wholesale power markets, the market-clearing equilibrium price (supply meets demand) could be determined by the most expensive plant, among the existing ones, that is willing to produce (its variable cost determines, and equals to, the market-clearing price) under specific assumptions (Reference [18]). Based on the aforementioned, an estimation of an optimal dispatch could be determined by a merit order model. The computational efficiency of such models can provide a useful framework for examining prices and the dispatch in different market models.

There are different methods to capture short-term operation by merit-order-based models in GEP. Such, is the Screening Curve method (Reference [22]). It can provide an estimate of the total cost of thermal power plants per time period per unit of available capacity. Due to their computational efficiency, Screening Curve method has been widely employed within GEP models to provide this estimate of the least-cost option to generate electricity and identify an optimal capacity mix. For the cases where variable RES generation, is considered modifications are required. Commonly, the LDCs are replaced by the net LDC, i.e. the generation by variable RES is first subtracted from the load and then sorted. This results in the loss of chronological information which can have an impact on the attained investment decisions (Reference [23]). Moreover, the loss of chronological information can render capturing the variations of generation by RES, thermal unit operating limitations, demand-side management and/or storage capabilities challenging (Reference [9]). Therefore, efforts have been made to improve the accuracy of this estimation (References [24, 25, 26]).

Short-term operation: the Unit Commitment Problem

Another main category of models for short-term generation scheduling are models for the UCP. Such models are more technical-oriented as technical restrictions of the generating units are accounted for.

In general, a UCP model includes two sub-problems: the Unit Commitment (UC) and the Economic Dispatch (ED) problems. The former implies the definition of the operating status of each unit for each scheduling period and the latter the load allocation among the committed units. To determine the operating states and the production levels, the operating limitations of the thermal generators are considered. Such limitations arise by the units' technical and economic characteristics, e.g. the unit ramping capability, the minimum and maximum production level, the minimum on-line/off-line time constraints, up/down reserve requirement, and fuel consumption cost. Moreover, commitment decisions need to be taken in advance. Therefore, the cost for bringing a unit on-line (start-up cost) or switching it off-line (shut-down cost) is taken into account. Such cost arises due to fuel consumption during the time a unit is heating up or cooling down. In addition to efficiently meet the electricity demand, a set of other operating requirements are considered regarding the reliability of the system (e.g. operating reserves or transmission limitations). From a mathematical point of view, a UCP is a complex optimization problem that can be formulated as a Mixed Integer Linear Programming (MILP) problem.

UCP models have been also presented to capture the short-term characteristics for longer scheduling periods (up to a year). For example, Reference [27] suggested a long-term UCP model including pumped storages. Computational tractability of the MILP is tackled by decomposition methods. Reference [28] examines heuristic solutions for the long-term UCP including co-generation plants. Reference [29] suggested a linear formulation to simulate a UCP for evaluating the curtailment of RES generation and operational costs for large-scale power systems.

2.2 Long-term planning with increased detail in the representation of the short-term operation of a power system

The increasing penetration of generation from variable RES and its efficient integration within a power system has led to integrating long-term planning models with a more detailed representation of the short-term operation (Reference [2]). Variability and uncertainty of RES generation require to be addressed during short-term planning to account for sufficient operating flexibility. As a result, the value of operating flexibility must be also assessed during long-term planning to determine the optimal capacity additions when high shares of RES generation are considered. Operating flexibility refers to the ability of the power system to respond to fluctuations and meet the net load (the residual load to be met by conventional units after subtracting the contribution of intermittent energy) within an acceptable time frame by adjusting supply. These are elaborated in the following Sections and constitute one of the main motivations for this thesis. More specifically, some basic characteristics of RES generation are provided. Then, the impact of accounting for short-term operation in the context of long-term planning is discussed and a number of representative approaches are presented.

2.2.1 Characteristics and integration of generation by Renewable Energy Sources

Main characteristics of electricity production by weather-dependent (e.g. wind farms and solar photovoltaic installations) RES are the following:

- The output of RES installations is variable: This variability can be observed in different time scales. For example, seasonal fluctuations can be observed within a year, e.g. solar photovoltaic have higher output in summer months than in winter ones. Their output may fluctuate within a day, e.g. the peak in generation by solar photovoltaic is usually observed in midday while their output is zero during night hours. Besides, the output of RES may vary in large percentages within short time-scales (intraday and intra-hourly) based on shifts occurring in the weather conditions. In cases of rapid increases/decreases, the short-term generation schedule should be able to meet such shifts.
- The output of RES installations is uncertain: Uncertainty of RES generation is a consequence of their weather-dependent output. Specifically, the anticipated RES generation output is considered during the day-ahead scheduling. However, it cannot be perfectly forecasted. Therefore, the generation schedule must be able to meet cases of overproduction or underproduction of RES generation along with the other common uncertainties of generation scheduling, i.e. load error forecasting or unexpected power plant outages.
- The marginal cost of RES generation is (almost) zero: The zeros marginal cost of RES generation has a number of implications regarding the operation and profitability of the generating fleet (both for conventional and RES) and the final consumer price (Reference [30]). The negligible marginal cost of RES generation leads to efficient generation scheduling prioritizing it and conventional (thermal) generators satisfy the remaining (residual) demand and provide sufficient reserves. In a scenario of high shares of RES generation, and under current power market design conditions, implications regarding to profitability of the generating fleet could emerge (Reference [30]). For example, negative marginal prices have been observed in the electricity markets' operation.
- RES production is GHG emission-free: The necessity of tackling environmental issues (Reference [1]) has triggered the ongoing low-carbon transition of many power systems. A direction towards the reduction of GHG emissions is the increase in the share of RES generation (Reference [31]). Since the production by RES does not omit GHG emissions, and if their dispatch is prioritized in generation scheduling, their contribution towards meeting the demand can replace energy production otherwise generated by fossil units.

2.2.2 Motivation for assessing operating flexibility within long-term planning

Including short-term dynamics within long-term planning is mostly motivated by the necessity for a reliable operation of a power sector. Including higher detail concerning such dynamics may provide additional, or in some cases more accurate, insight for evaluating scenarios and options in long-term planning (References [7, 8]). Such is the case a transition towards higher shares of generation by RES is considered where the value of operating flexibility should be also assessed.

The main source of increased operating flexibility requirements is based on the variability and uncertainty of RES generation induced by their dependency on meteorological conditions; RES output follows the fluctuations that occurred within the latter. More specifically, conventional generators should be able to provide the load following deviations among the forecasted load used for unit commitment and the forecast load used in the real-time balancing market. At high RES penetration levels, load following requirements are increased, as forecasting errors occurring due to RES uncertainty and RES variability must be also accounted for. Moreover, the variability of the net load in comparison to the variability of the load can result much higher. In such cases, flexible generators could be required to operate more frequently and/or generators may be forced to more frequent cycling w.r.t. their techno-economic characteristics. This may increase operating costs (e.g. operation, maintenance, and start-up) of conventional base-load units. The aforementioned are based on Reference [32] where these are discussed in detail. In addition, Reference [33] suggests that increased wind generation levels might lead to operating flexible conventional units more frequently. The importance of highly flexible and reliable units, of adequate interconnections and storage capacity to be able to coop with uncertain and intermittent production of most RES installations, is also highlighted. Reference [34] developed a linear programming model including several operating constraints such as operating reserves and ramping ability of the units. They identified that the generation share of base-load units could be replaced by generation from intermediate units due to the penetration of wind power (attributed to its intermittency and variability), (ii) base-load units ramping restrictions are important operating limitations to be considered, (iii) the transmission system and energy storage can serve as operating flexibility providers, and (iv) peak-load technology additions could be less necessary. Reference [35] outlines steps for relying primarily on variable RES i.e. transition to zero carbon power systems with 100% RES. Specifically, 3 phases/steps towards creating power systems with sufficient flexibility to maintain a reliable and stable operation are presented. These steps are provided in the form of a comprehensive overview of policies, technical changes, and institutional systems.

2.2.3 Flexibility providers

Operating flexibility is not restricted to the technical flexibility provided by thermal units. A review on main flexibility providers is provided in Reference [36] which is a comprehensive literature study on recent flexibility mechanisms in power systems with a high penetration level of generation by variable RES. Main power system flexibility resources can be provided from both the demand and supply side. Flexibility, from the supply side, is mostly represented by the available generating fleet. For example, flexibility can be provided by (i) thermal generators, w.r.t. their short-term techno-economic characteristics and limitations, (ii) non-thermal generators, such as conventional hydro-power plants or storage facilities (discharging), which can smoothen the net-load. From the demand-side, mechanisms such as demand response, smart grids, storage facilities (charging), and possible interconnections with neighboring power systems can also contribute. Moreover, an enhancement to the transmission system may also provide flexibility to the transmission system and consequently mitigate possible congestion (Reference [2]). Lastly, the variable generation by RES output cannot be controlled as effectively as the corresponding one of the conventional thermal fleet. The main instrument for adjusting the former is through curtailment which is also restricted by weather conditions. To be specific, curtailment is the voluntary reduction of RES output to a lower level than the maximum available.

2.2.4 Limitations and simplifications

One of the main restrictions in including a detailed representation of the short-term operation in long-term planning, such as a GEP model, arises due to the computational restrictions. The primary focus towards this direction follows the inclusion of main aspects of the UCP in long-term models (Reference [9]). Both GEP and UCP models can be computationally challenging. A MILP formulation of UCP can include detailed technical constraints on a unit level, (e.g. operating levels, ramping capabilities, and minimum up/down times), system requirements on an hourly level (e.g. load and reserve requirements) and a detail network representation (e.g. transmission system). On the other hand, a long-term planning model can include techno-economic detail regarding a number of investment options, a network representation to consider the location of such investment, and a planning horizon spanning over a few years or decades. A combined GEP-UCP formulation considering the aforementioned emerges as a large scale MILP formulation that requires significant computational resources to be tackled (Reference [7]). Consequently, simplifications are introduced to the representation of short-term operation in the context of long-term planning. An overview of the aforementioned is included in Reference [9]. The simplifications are presented in terms of the level of technical, temporal, and spatial detail used to describe the electric energy system. Furthermore, endeavours have also been focused on suggesting methods to tackle the computational intractability of such models. Such approaches are reviewed in Reference [2].

2.2.5 Impact of technical, temporal and spatial detail of short-term operation on long-term planning

Capturing the impact of the intermittent output of RES generation on the short-term operation of a power system within the context of long-term planning is not straightforward (References [1, 2, 3, 4, 9]). It implies that sufficient information is considered regarding: (i) technical detail (References [7, 8, 37, 38]), e.g. technical characteristic of the operating units and system requirements, (ii) temporal detail (References [39, 40, 41, 42, 43]), e.g. chronological information on an hourly level including different operating conditions and seasonality, and (iii) spatial detail (References [44, 45]), e.g. the available transmission system to consider possible congestion. Even though the importance of neglecting RES variability can be system specific (Reference [10]), it increases with higher RES penetration levels. In general, omitting technical, temporal and/or spatial detail may have an impact on optimal planning. More specifically:

- Technical detail: Technical detail may have an impact on long-term planning as it may overestimate the operating flexibility of an examined installed capacity due to neglecting technical restrictions. For example, References [7, 8] demonstrated that neglecting flexibility requirements can significantly affect the derived generation mix and lead to sub-optimal (cost and GHG emission) generation portfolios that are infeasible to operate when high levels of RES generation are considered. Reference [37] focus on the impact of operating reserve requirements on generation capacity investments following the large-scale integration of intermittent RES generation. They revealed that neglecting such requirements and their allocation and costs may result in underestimation of RES integration costs. They also suggest innovative sizing and allocation strategies for operating reserves. Moreover, Reference [38] have demonstrated that ignoring flexibility constraints could result in high deviations of the resulting installed capacities in a case study focusing on Ireland.

- Temporal detail: In general, limiting the impact of temporal detail may result in overestimation of generation by RES and inflexible units and in an underestimation of the investments in flexible generation. This is demonstrated in a number of studies. For example, Reference [39] examined the influence of increasing temporal resolution on the optimal technology mix through a model combining long time scales of climate change mitigation and power system investments with short-term fluctuations of RES. Among their findings, an increase in flexible technologies natural gas technologies had been observed by increasing temporal resolution. Reference [40] presented an extension of the an energy planning tool to consider seasonal, daily, and hourly supply and demand dynamics. Regarding temporal detail, they highlight that lower resolution models can overestimate the optimum amount of investment in RES. Moreover, has been examined by deriving a capacity mix from a GEP model that does not include the restrictions of a UCP model and examine the dispatch decisions by employing one. For example, Reference [41] focused on quantifying and analyzing the impact of simplifications considering temporal and technical detail. They demonstrated that a suitable approach can be the appropriate selection of a set of historical days to represent an entire year. Reference [42] suggested a soft-linking methodology to verify the technical appropriateness of the energy systems developed portfolio. They highlight that an energy system model can produce reliable portfolios that however may overvalue variable RES, undervalue other flexibility providers (storage) and overestimate base-load operation. Towards this direction, Reference [43] compared different approaches to select a representative set of days. By demonstrating that increasing temporal detail may provide more robust results in the expense of a higher computational cost, they propose an optimization-based approach for selecting representative time periods and suggested indicators and metrics for the evaluation of representativeness.

- Spatial detail: Benefits may arise by spatial smoothing of RES generation due to spatial diversification. Moreover, limited representation of the transmission networks does not account for challenges arising due to possible transmission congestion. For example, Reference [44] assessed the impact of spatial resolution (spatial aggregation) on the investment planning decisions by developing a linear programming model. Their results revealed that the relative competitiveness of RES technologies may be affected and consequently lead to suboptimal investments in capacity additions. Reference [45] developed a MILP model for the optimal long-term energy planning of a power generation system. Within this context, the approach focuses on a detailed representation of spatial and technical characteristics of short-term operation. Among other findings, they identify the allocation of investment and capacity additions derived by the model aimed towards balancing the demand excess and production excess among the considered regions.

2.2.6 Representative approaches on Generation Expansion Planning models integrating dynamics of short-term operation

This Section focuses on long-term optimization models, and specifically GEP models, that focus on increased shares of RES generation. The level of detail for which short-term dynamics are accounted for differs and mostly relies on the scope and focus of each study. Due to the many different perspectives, scopes and aims of the studies and models, some categories have been excluded. Specifically, and following the categorization in Reference [3], non-optimization models such as probabilistic, simulation, life cycle assessment, cost-benefit analysis, econometric, multi-criteria, system dynamics, and modern portfolio theory models have been excluded. Moreover, computable general or partial equilibrium models have been excluded. In such models the GEP problem is part of an optimum equilibrium solution of the whole economy or the energy system respectively (Reference [3]). For greater detail on (i) integrating energy by RES in the GEP, (ii) integrating short-term variations of the power system into integrated energy system models, and (iii) the state-of-the-art on incorporating short-term dynamics on GEP decisions, the reader is kindly referred to References [2, 3, 4]. Some representative approaches are presented in the following Section and have been categorized in: (i) static GEP models including a UCP model, (ii) multi-period GEP models including a UCP model, (iii) approaches focusing on computational restrictions, (iv) approaches including flexibility metrics, (v) MOO approaches, and (vi) EA-based SOO and MOO approaches. However, some of the following approaches could be included in more than one category.

Static GEP models including a UCP model

Efforts have been made towards merged GEP-UCP models to assess operational flexibility requirements and their impact on the output of the model. For example, Reference [7] presented a MILP model for the optimal GEP problem, employing a full year representation with hourly time scale and a series of clustered unit commitment constraints utilized to handle the computational cost. The clustered unit commitment formulation considers integer variables for the UCP, in contrast to the commonly formulated binary variable UCP, leading to significant computational time reduction. Reference [8] extends the aforementioned approach and examines the impact of operational flexibility on the GEP, incorporating RES and emission reduction targets. Reference [46] proposed a GEP model that embeds a convex relaxation of a UCP as a short-term operational model that is a continuous and polynomially-solvable optimization problem. The investment decision variables are the only integer variables of the model. They identify that neglecting operational flexibility in GEP can lead to reserve shortage, load shedding, and curtailment of RES generation due to under-investment in flexible capacity. Reference [47] proposed a model to determine the optimal generating mix of a power sector by accounting for operational flexibility. The approach employs an operational model to re-evaluate the installed capacity derived by a basic GEP model and iteratively seek for an improved solution. The advantage of this approach is that operating restrictions are considered on a power plant level. Also, it demonstrates the importance of operational constraints related to thermal units on the investments and operational planning, w.r.t. the level of RES generation. Reference [48] developed a novel capacity expansion model optimizing investment decisions. The LP optimization model employs technical, economic and spatial characteristics to aggregate units and a relaxed-integer formulation of a UCP model to reduce modelling complexity. Storage technologies and policy constraints are also considered. They find that neglecting flexibility constraints would significantly underestimate the curtailment rate and costs. Moreover, they highlight the importance of lower storage cost to achieve affordable higher shares of RES penetration.

Multi-period GEP models including a UCP model

Models considering a multi-period planning horizon have been also suggested. Such models aim towards capturing additional variations in the critical parameters such as the investment costs and fuel prices. For example, Reference [49] proposed an integrated model for GEP with high shares of variable RES output that considers short-term operation by a simplified UCP formulation presented in Reference [50]. The model's applicability was demonstrated considering a 10-year planning horizon; a representative day was selected for each season. They compare the results with a traditional GEP model and reveal that assuming average operating conditions can underestimate the system costs. Moreover, they highlight the impact of RES variability on the operating conditions can lead to sub-optimal GEP decisions. Reference [51] developed a multi-period multi-regional GEP model that had been formulated as a MILP. It considers the annual constraints of the GEP problem and the short-term constraints of a UCP. The approach considers many aspects of the short-term operation and its representation exhibits a high level of technical, temporal, and spatial detail. Computational restrictions had been addressed by selecting a representative day per month over the long period to determine the optimal generation mix and energy planning details of the power system. The model aims towards deriving the optimal power production mix, capacity additions, and System Marginal Price. They find that higher RES penetration levels are correlated to higher production shares of natural gas fired units and higher levels of electricity trading which are attributed to flexibility requirements. Reference [52] developed a stochastic generation capacity expansion planning model to assess environmental policies such as renewable portfolio standards and carbon tax. Uncertainty from wind and load availability had been considered by adopting the Gaussian copula method. Moreover, they provide a comparison with a target year model.

Approaches focusing on computational restrictions

Some approaches have focused on addressing the emerging computational complexity. Different decomposition methods have been used such as Dantzig-Wolfe decomposition (Reference [53]) or Nested Bender's decomposition (Reference [54]). Reference [53] proposed a column generation approach for optimizing the multi-period GEP problem with high integration of RES. The developed GEP-UCP model had been formulated as a large scale MILP. Computational times were reduced by employing Dantzig-Wolfe decomposition and a clustering technique. The study also highlights that incorporating short-term constraints into the long-term planning horizon may provide noticeable cost reductions. Reference [54] proposed a MILP model for the long-term planning of investments in the power sector including a UCP model. Computational tractability had been addressed by modelling approximations and aggregations. Moreover, an algorithm had been proposed based on Nested Bender's Decomposition for multi-period MILP problems including acceleration techniques to improve the overall performance of the algorithm and achieve computational time reductions.

Approaches including operational flexibility metrics

A number of approaches have employed operational flexibility metrics. Reference [55] suggested the use an off-line flexibility index that estimates the individual contribution of generating units to the overall system flexibility. It considers the ramping capabilities and operating range of thermal generators. Also, a unit construction and commitment algorithm is developed to determine the optimal investments in flexible generating units. The approach aims in evaluating the flexibility level provided and investigate the role of flexibility in generation planning and market operation. Among other findings, they highlight the importance of a market design w.r.t. an efficient and profitable deployment of flexibility resources. One of the main advantages of such off-line metrics could be their computational efficiency. Towards this direction, Reference [56] formulated a metric, termed the composite flexibility metric. It considers a large number of important technical flexibility characteristics of generating units as indicators and the metric is adapted based on the whole generating fleet. It may serve as a metric for comparing different generation mixes. Moreover, the value of the metric for a unit is adapted based on the flexibility characteristics of the remaining units within the generation portfolio. Reference [57] suggested an integrated framework that includes operational flexibility assessment metrics. The developed multi-period GEP-UCP model is indented for high shares of intermittent RES. In addition, the framework includes methods to address the computational burden induced by the resulting large-scale optimization problem. Their results demonstrate that the inclusion of detailed short-term constraints within long-term planning is important for high RES penetration levels as neglecting them may result in underestimation of the required operational flexibility and the GHG emissions omitted.

Market-based approaches

Approaches employing a market-based framework have been suggested to assess the evolution of electricity prices. For example, Reference [58] proposed a mid-term market-based power systems planning model to including both GEP and TEP decisions at a yearly level and a model for the UCP on an hourly level. The formulated MILP model aims at identifying the power mix, the RES evolution, and the day-ahead prices and had been employed within their work to examine the feasibility and impact of an interconnection of the mainland power system of Greece with the autonomous power system of Crete. Reference [59] examined the impact of increasing wind share on the optimal generation mix and the profitability of the generating capacities been. They developed a GEP model including an hourly UCP model to analyze policies that support resource adequacy. They find that higher wind shares reduce average electricity prices resulting in different implications, regarding the profitability of each considered generator type, and generation expansion plans under the three examined policies.

Approaches examining storage technologies as flexibility providers

Moreover, focus has been concentrated on capturing flexibility provided by non-thermal generators. Reference [60] presented a MILP model for the optimal GEP at a future target-year. They identify an increase in the role of storage-based technologies for the examined electricity systems and that availability of different low-carbon technologies can impact the optimal capacity mix and generation patterns. Reference [61], developed a MILP long-term investment model including a continuous relaxation of the technology-clustered formulation of the short-term UCP. Moreover, energy services and frequency control provided by storage technologies are also accounted for. Their results, derived by the application on a full year, showed that integration of storage resources can facilitate RES integration and lead to cost reductions. Moreover, they demonstrate that storage technologies could reduce, to some extent, the requirements for flexible power plants and support inflexible ones.

Multi-objective optimization approaches

There are also many available MOO approaches considering RES integration for GEP and such are reviewed in References [1, 3, 4]. For example, Reference [62] presented a multi-objective model for expansion with high shares of RES. The model considered the minimization of the total cost, the maximization of generation at the peak load, and the maximization of non-hydro RES generation contribution, as conflicting objectives. Their results indicate the potential importance of solar power generation for the future Brazilian system. Reference [63] developed a multi-objective method considering also large-scale demand-side management and demand response technologies. The objectives functions defined to determine an optimal mix of the renewable system had been set as the maximization of RES contribution to the peak load, the minimization of monthly and yearly intermittence, and the minimization of production cost. Reference [64] developed a multi-period multi-objective GEP including cost, environmental and reliability objectives as the competing functions. Also, reliability is evaluated by an analytical probabilistic approach. The optimization framework is based on lexicographic optimization and the Normal Boundary Intersection method. They find that the developed optimization framework derives efficient solutions that are evenly spread on the Pareto front while ensuring that dominated solutions are not produced.

However, GEP studies that focus on capturing operating flexibility requirements, induced by RES variability, by considering a detail representation of short-term operation are relatively limited. This can be attributed to the high computational cost associated with GEP-UCP models or the increased detailed required for the representation of short-term operation. Towards this direction, Reference [65] suggested a MOO approach that considers operating flexibility as a separate objective function. It utilizes the composite flexibility metric (Reference [56]) that serves as an indicator for the operational flexibility objective function. The optimization model considers also cost and environmental objectives and a set of limitations regarding reliability. A MOEA (Reference [66]) is employed for the MOO problem. They highlight the advantages of adopting a low modelling effort approach for capturing operating flexibility and mention the computational challenges arising for the case where a highly complex model would be included. Their results suggest the importance of assessing operating flexibility and identifying potential correlations among the considered objectives. Moreover, Reference [6] presented a MILP model for a mixed hydro–thermal–wind power systems. A series of constraint SOO problems is carried out to generate solutions of the MOO problem. The objective function is set as the minimization of cost and emissions to identify extreme solutions. Additional solutions are attained by varying the restrictions imposed on the generated emissions. This had been implemented by setting the emission function as a constraint function. The different scenarios produced are analyzed and their results suggest the importance of wind and hydro capacity to meet environmental objective in a cost efficient manner.

EA based single and multi-objective approaches

Besides the approaches based on classical optimization methodologies, a number of heuristic and/or meta-heuristic techniques have been employed in an attempt to address the GEP problem. In general, classical optimization approaches exhibit an important advantage when applicable: they can guarantee global optimal solutions in a finite number of steps (w.r.t. the mathematical formulation of the optimization problem) which is important from a decision maker's point of view. On the contrary, heuristic or meta-heuristic approaches could be applicable for optimization problems that include complex, non-smooth, non-convex and/or non-differentiable objective and constraints functions. Moreover, heuristic techniques could provide satisfactory results within an acceptable time limit for computationally expensive optimization problems. In addition, meta-heuristic population-based approaches, such as EAs, are considered highly applicable for MOO problems (Reference [12]).

EA-based approaches have been suggested for the GEP problem. Reference [67] provided a review of heuristic techniques applied on the GEP problem. Moreover, heuristic and meta-heuristic approaches have been discussed within the review provided in Reference [1], and mentioned with focus on RES integration within Reference [4]. Reference [68] performed a comparative analysis of different EA approaches applied on the GEP problem.

EA based approaches to tackle the GEP problem as a SOO optimization problem have been presented. A representative direction is the employment of Genetic Algorithms (GA). For example, Reference [69] focuses on GEP for competitive markets. The developed model aimed at maximizing the expected revenues of a generation company by considering the increased risks affecting their activities in liberalized markets. Uncertainty by price volatility, reliability of generating units, demand and investment, and operation costs had been represented in the presented model by the inclusion of probability distribution functions and constraints. A GA and Monte Carlo simulation had been employed to address the combinatorial nature of the problem and to sample values from the probability distribution functions, respectively. Moreover, other EA based approaches have been employed. For example, a modified DE algorithm has been proposed in Reference [70] which focused on the impact of increasing penetration of solar power technology by a GEP model accounting for emission restrictions. They employ a DE algorithm to optimize the developed GEP model. Through a comparative analysis they observed an increase in the installed capacity of the derived optimized system when solar plants had been introduced as a possible capacity addition alternative that resulted also in a reduction of the costs considered. Reference [71] suggested a DE-based approach for the optimization of a GEP-TEP model considering thermal units. In the proposed DE method, the population is clustered using the k-means method. Their numerical experiments, on benchmark functions and the GEP-TEP model examined, reveal that the algorithm's efficiency is enhanced by the suggested modifications. Furthermore, approaches employing a Particle Swarm Optimization algorithm have been suggested. Reference [72] developed a multi-stage GEP incorporating large scale energy storage systems. A Particle Swarm Optimization algorithm had been employed to address the multi-stage mixed-integer non-linear programming model. They find that energy storage systems can contribute to cost reduction and environmental pollution for systems including variable RES and thermal generating capacity.

Reference [4] discusses that the number of MOO studies considering the GEP problem is still much lower than for single-objective ones, despite the growth of MCDM methods utilization in recent years. In some cases, EA based approaches have been employed for MOO formulations of the GEP. A motivation for the latter could be that population-based approaches can provide a set of solutions within a single run for MOO problems. For example, Reference [73] considered two different formulations for the GEP problem using the elitist NSGA-II. The first formulation considered the minimization of a total cost function and the sum of the normalized constraints violation. The second formulation considered the minimization of the investment cost and the maximization of the system's reliability. Aiming at enhancing the algorithms efficiency, a Virtual Mapping Procedure had been introduced to reduce the number of decision variables. Reference [74] further enhanced the aforementioned optimization algorithm by including a controlled elitism mechanism. The approach considered investment and outage cost as conflicting objectives. Reference [75] suggested a MOO framework for the transmission constrained GEP. The first objective is formulated as a total cost function to be minimized and the second objective to be minimized is represented by the sum of the normalized constraints violation. The elitist NSGA-II was employed. In general, the MOO framework had been suggested to address the constraint nature of the optimization problem as they demonstrate that the MOO approach had been more efficient than a constraint SOO framework for handling the constraint SOO problem.

Furthermore, hybrid approaches have been also suggested. Such an example is Reference [76] where a GEP model is developed, with a focus on thermal units, to identify the most economical investment planning considering also demand-side management programs, reliability and environmental aspects and limitations. A GA-Benders' decomposition method has been suggested to optimize the model. The approach is based on a GA and includes Benders' cuts. Their results indicate economic and environmental benefits by gas-fired power plants and an improved performance of the hybridized approach for small and medium size systems. Reference [77] employed a hybrid EA for a long-term capacity model to endogenously derive the evolution of the marginal price and examine the impact of relaxing quantitative energy objectives on the performance of the algorithm.

2.3 Discussion

Various GEP based modelling approaches with a wide range of objectives have been developed. Such approaches are categorized and thoroughly discussed within recent review studies (References [1, 2, 3, 4]). In general, such approaches differ (among other aspects) w.r.t. the inherent advantages and limitations regarding the accounted techno-economical, temporal and spatial detail considered. Moreover, the computational efficiency of such approaches may differ. GEP is known as a challenging problem due to its non-linearity, non-differentiability, high-dimensionality and discrete variables included (Reference [1]).

Therefore, modelling and optimization methodologies/approaches introduce simplifications, regarding the aforementioned basic aspects, i.e. central or market based, decisions considered (size-type-location-time), single node or network constraint, short-term representation detail, deterministic or stochastic modelling. Also, computational tractability should be addressed. Consequently, a trade-off between modelling accuracy and computational tractability should be established.

The main focus of this thesis is to include key aspects of the short-term operation of a power system within the context of long-term power planning. The examined level of detail is restricted to hourly intervals to capture the short-term dynamics and technical limitations commonly addressed within a UCP model. Within the literature, the main focus in integrating long-term models with short-term operation is also identified in employing simplified UCP models (Reference [9]). Therefore, a more detailed representation of the short-term operation of a power system including sub-hourly dynamics is beyond the scope of this thesis.

Successful applications of EAs for GEP can be identified in the literature. However, this thesis discusses MAEA based approaches for GEP models including a SM considering UCP constraints. This is motivated by the high computational cost required to perform multiple simulations which are required for an EA to converge.

On the other hand, MAEAs have been successfully applied for examining the short-term operation of a power system. For example, Reference [78] proposed an efficient MAEA-based method for solving power generating UCP with probabilistic unit outages. Moreover, Reference [79] employed a surrogate-assisted DE algorithm for the short-term combined economic and emission hydrothermal optimization utilizing a master-slave level approach. These approaches, however, focus on the short-term operation and long-term planning and investments decisions are not considered.

Moreover, hybridized EA approaches including repair heuristics have been suggested for optimizing the UCP. For example approaches which focus on a thermal UCP (References [80, 81]), Reliability UCP (Reference [82]) and a MOO formulation (Reference [83]) have provided satisfactory results. A number of customizations are developed in this thesis based on the problem characteristics for GEP. These aim towards addressing the constraint optimization problem, the discrete nature of decision variables and exploit problem characteristics. Some problem-customized operators (e.g. Virtual Mapping Procedure) for GEP have been also examined in Reference [73].

Chapter 3

Evolutionary Algorithms

Evolutionary Algorithms are nature-inspired computational methods that have been frequently employed in an attempt to tackle real-world optimization problems. In general, EAs are based on the Darwinian principle of natural selection (Reference [84]). This principle states that under specific conditions the evolution of organic beings can occur by natural selection. Their popularity lies in some interesting properties they exhibit (e.g. derivative-free) rendering them suitable for many applications. These are discussed in more detail within this Chapter. In the following Sections, some preliminaries regarding Single and Multi-objective optimization are presented. Then, the relevant literature on Evolutionary Algorithms and MAEAs are presented.

3.1 Preliminaries

In this thesis, Single and Multi-objective black-box optimization problems are considered. Moreover, some special cases of SOO, i.e. Linear Programming, Mixed-Integer Linear Programming, and Integer Linear Programming, are also considered. The aforementioned problem formulations along with some definitions are provided in the following Sections.

3.1.1 Single-objective optimization

Many engineering problems can be formulated as SOO problems. A SOO problem involving constraint functions can be formulated as follows:

$$\begin{aligned} &minimize: \quad f(\mathbf{x}) \\ &s.t.: \mathbf{G}(\mathbf{x}) \leq 0 \\ &\qquad \mathbf{x} \in \mathbb{S} \end{aligned} \tag{3.1}$$

where $\mathbf{x} = (x_1, x_2, ..., x_n)$ is the decision vector of objective/decisions variables, n is the number of decisions variables, $f(\mathbf{x})$ is the objective function, $\mathbf{G}(\mathbf{x}) = (g_1(\mathbf{x}), g_2(\mathbf{x}), ..., g_k(\mathbf{x}))$ is a vector of constraints, k is the number of constraint function, and $\mathbb{S}$ is the search space. The search space is determined by the upper and lower limits of each variable $\overline{\mathbf{x}}$ and $\underline{\mathbf{x}}$, respectively. In case a SOO problem does not involve constraint functions ($\mathbf{G}(\mathbf{x})$) then it is referred to as un-constraint.

The constraint violation, $cv(\mathbf{x})$, for a decision vector $\mathbf{x}$ can be provided by:

$$cv(\mathbf{x}) = \sum_{i=1}^{k}[\max(g_i(\mathbf{x}), 0)] \tag{3.2}$$

A solution x is called feasible when $cv(\mathbf{x}) = 0$. Correspondingly, a solution $\mathbf{x}$ is called infeasible when $cv(\mathbf{x}) > 0$. The objective of the optimization problem is to identify a feasible solution that exhibits the minimum objective function value.

Moreover, some definitions regarding the function's properties mentioned in this thesis are provided:

- **Multi-modal function**: A function is referred to as multi-modal if it has more than one optimum.
- **Black-box function**: A function is referred to as a black-box function when no specific assumption are made or no information is available regarding the function's properties, e.g. differentiable or linearity.
- **Expensive/costly function**: A function is referred to as expensive or costly when it requires significant computational resources (or money) to compute.

3.1.2 Multi-objective optimization

MOO refer to optimization problems that involve more than one conflicting objective functions. Commonly, it is expressed as:

$$\begin{aligned} &minimize \ \ \mathbf{F}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), ..., f_m(\mathbf{x})) \\ &s.t. \ \ \mathbf{G}(\mathbf{x}) \leq 0 \\ &\qquad \mathbf{x} \in \mathbb{S} \end{aligned} \tag{3.3}$$

where $\mathbf{x} = (x_1, x_2, ..., x_n)$ is the decision vector of objective/decisions variables, n is the number of decisions variables, $\mathbf{F}(\mathbf{x})$ is the objective function vector, m is the number of objective functions, $\mathbf{G}(\mathbf{x}) = (g_1(\mathbf{x}), g_2(\mathbf{x}), ..., g_k(\mathbf{x}))$ is the constraint function vector, k is the number of constraint functions and $\mathbb{S}$ is the search space. The search space is determined by the upper and lower limits of each variable $\overline{\mathbf{x}}$ and $\underline{\mathbf{x}}$, respectively. In case $m = 1$ the above expression (3.3) represents a SOO problem.

Definitions Reference [12] regarding Pareto optimality, and based on Reference [13], follow:

- Pareto dominance: A vector $\mathbf{F}^a$ dominates vector $\mathbf{F}^b$, expressed as $\mathbf{F}^a \prec \mathbf{F}^b$, if $F_i^a \leq F_i^b$, $\forall i \in \{1, 2, ..., m\}$ and $\mathbf{F}^a \neq \mathbf{F}^b$.
- Pareto optimal solution: A feasible solution vector $\mathbf{x}^* \in \mathbb{S}$ is called a Pareto optimal solution, if $\nexists y \in \mathbb{S}$ such that $\mathbf{F}(\mathbf{x}) \prec \mathbf{F}(\mathbf{x}^*)$.
- Pareto set: The set of all Pareto optimal solutions is called the Pareto set:
$\mathbb{PS} = \{\mathbf{x} \in \mathbb{S} | \nexists \ y \in \mathbb{S}, \mathbf{F}(\mathbf{y}) \prec \mathbf{F}(\mathbf{x})\}$.
- Pareto front: The Pareto front refers to the image of the Pareto set in the objective space:
$\mathbb{PF} = \{\mathbf{F}(\mathbf{x}) | \mathbf{x} \in \mathbb{PS}\}$.

3.1.3 Linear Programming, Mixed-Integer Linear Programming and Integer Linear Programming

A Linear Programming is commonly expressed as:

$$\begin{aligned} &minimize \quad f = \mathbf{C}^T\mathbf{x} \\ &s.t. \quad \mathbf{A}\mathbf{x} = \mathbf{b} \\ &\qquad \mathbf{x} \geq 0 \\ &\qquad \mathbf{x} \in \mathbb{R} \end{aligned} \tag{3.4}$$

where $\mathbf{C}$ are the parameters for the linear objective function, $\mathbf{A}$ and $\mathbf{b}$ are the parameters and the constraint bounds of the linear constraint functions, respectively. In case some of the decision variables are restricted to integers then additional constraints are included. Then, the optimization problem is referred to as Mixed-Integer Linear Programming (MILP). Correspondingly it is referred to as Integer Linear Programming when all decision variables are restricted to integers.

3.2 Single and Multi objective optimization by Evolutionary Algorithms

3.2.1 Basic characteristics and components of Evolutionary algorithms

EAs are frequently applied to optimization problems. In contrast to classical optimization methodologies, EA exhibit some differences regarding the search process. In general, EA are derivative-free population-based algorithm involving stochastic operators within the search process.

EA approaches follow a concept deriving from evolution and the principle of natural selection. Briefly, a number of individuals form a population that exists and evolves in an environment. The individuals of the population exhibit unique traits/characteristics that define their fitness and reproducing chances. These traits, or similar ones, are likely to be inherited by reproduction. The restricted resources of the environment force the population members to compete for survival. The fittest individuals survive or at least are more probable to survive if randomness is taken into account. The overall fitness of the population, over a series of generations, tends to increase due to selection pressure induced by the diverse traits within the population and the survival pressure.

The most important components of an EA are based on this evolution procedure. These components are included in most EAs as operators but may differ on their implementation. A brief description follows that does not focus on a particular EA. Consequently, the description is rather general but may provide a simple and brief note on a general EA framework. For this description, the categorization made in Reference [85], and discussed in Reference [86], is adopted*. Reference [85] summarized the basic components of EAs to five categories. These include: (i) EA-based representation, (ii) Initialization, (iii) Function evaluation, (iv) Evolution operators, and (v) EA parameters. Reference [86] provides an introduction to EA and discussed the basic components of an EA in eight steps. The eight steps consider: (i) EA-based representation, (ii) Function evaluation, (iii) Population, (iv) Mating operator, (v) Variation operators, (vi) Environmental-Selection operator, (vii) Initialization operator, and (viii) Termination condition.

*Some of the original category names have been altered for consistency.

- Individual (Variable representation and encoding): Let $\mathbf{x}$ be a decision vector. The cardinality of this decision vector represents the number of variables (different decisions) that must be made. How each variable and its possible values are represented differs based on the encoding used in an EA. More specifically, some EAs apply operators on representations of the decision vectors, or their encoding, rather than on the actual decision vectors. The word encoding is used to suggest a modelling approach employed for the decision variables within the EA. Hereafter, the EA-based representation of a solution shall be referred to as an individual and the actual solution/decision vector as a decision vector; regardless of the EA approach discussed. Likewise, the word trait is used to suggest the EA-based encoding of a decision variable. The decision vector and the corresponding individual could not always differ w.r.t. the employed representation scheme.

- Population: A population is formed by a number of individuals. Therefore, the population at a given instance is the pool of available solutions. The aim is to improve the overall fitness of the population by maintaining more fit individuals over a series of iterations. These iterations are referred to as generations.

- Fitness function: The fitness function is used to compare individuals. In the context of optimization, the fitness function provides the quality indicator for an individual which is the value of the objective function of an optimization problem or a transformation of the latter.

- Initialization operator: Commonly, an initialization operator randomly initializes the individuals to provide an initial population within the search space. In some cases, different approaches could be applied when a priori information is available.

- Mating operator: A Mating operator determines which parents are selected to reproduce offspring. In general, individuals that exhibit a high fitness value could be prioritized for mating to increase elitism. Assuming that the high fitness observed for an individual is a result of its traits, selecting such individuals as parents could produce offspring that have inherited such traits. Offspring are the EA-based representation of decision vectors generated by applying the EA operators on the parents within a generation.

- Recombination/Crossover operator: A Crossover operator combines a number of individuals from the parent population and generates the offspring population. Depending on the EA and the Crossover operator considered, the offspring may inherent a number of traits from each parent without alterations. Then, the EA relies on the Mutation operator to introduce new information.

- Mutation operator: A Mutation operator attempts to introduce small, random and unbiased changes to an offspring. One of the main reasons for introducing such small changes is to maintain a non-zero probability that an offspring can be generated at any point within the search space.

- Environmental-Selection operator: An Environmental-Selection operator determines which individuals shall form the population of the next generation among the available ones, e.g. the parent and offspring population. This operator applies selection pressure by comparing individuals, based on their fitness, and determines the most adequate ones that will be preserved.

- Termination operator/criteria: EA require termination criteria which is usually a user-defined criterion. For example, some common approaches are a predefined number of function evaluations, a time limit or a progress-related criterion.
- User defined parameters: Commonly, EAs require parameters to be determined by the user, e.g. the population size or other parameters required for the EA's operators.

The differences in the aforementioned operators distinguish the EA families of algorithms but are included in some form in most EAs. Moreover, the Mating, Crossover and Mutation operators are also referred to as the Variation operators. Moreover, the combination of Variation, Mating, and Environmental-Selection operators are also referred to as Evolution operators. In general EAs, apply the Initialization operator and the Evolution operators are then iteratively applied until the termination criteria are met.

3.2.2 Three major Evolutionary Algorithm paradigms

This Section presents a brief note on some major EA paradigms which are inspired by the principles of biological evolution. The first two, i.e. Genetic Algorithms (References [87, 88]) and Evolution Strategies (References [89, 90]), are among the earliest and most well-studied EAs approaches. The third one presented is Differential Evolution (Reference [91]) that is a more recent EA family of algorithms and has emerged as one of the most competitive ones (Reference [92]). However, some important representative EAs such as Evolution Programming (Reference [93]), Estimation of Distribution Algorithms (Reference [94]), Genetic Programming (Reference [95]) and Swarm Intelligence based approaches such as Ant Colony Optimization (Reference [96]) and Particle Swarm Optimization (Reference [97]) have not been included in this brief introduction.

Genetic Algorithms

GAs are probably the most widely-known EA approach. They mimic the genetic processes of biological organisms to evolve a population over a series of generations. However, they had been first introduced for studying adaptive systems (References [88, 98, 99, 100]). The classical GA (simple GA - SGA) (References [101, 98]) adopts a binary representation for each variable (bit strings-genes); a series of genes (expressing each decision variable) form each individual. Offspring are generated as a recombination (1-Point crossover) of selected parents (Roulette wheel) from the population and by probabilistic mutation applied on each bit independently (bit flip). Lastly, the offspring population replaces the parent population. Over the years many other factors such as elitism, real-coded representation, variation operators and self-adaptation of parameters have been introduced to the GA. Some basic characteristics of a GA are the important role of the Crossover operator, a relatively lower emphasis on the Mutation operator, and the stochastic nature of the Mating operator. Modern GAs are frequently preferred for binary search spaces and have found many applications especially for combinatorial optimization problems. Moreover, GA variants serve as the search engine for state-of-the-art multi-objective and many-objective EA-based algorithms, such as NSGA-II [66] and NSGA-III (References [102, 103]).

Evolution Strategies

Evolution strategies (ES) were introduced in References [89, 90]. The earliest ES algorithm was comprised of a single parent and a single offspring. The offspring was generated (Mutation operator) by adding a random value drawn from a Gaussian distribution with mean zero and a standard deviation σ (mutation step size). Then, the fitness value of the offspring was compared with the parent and the fittest was preserved (Environmental-Selection). Two major directions led to what is known today as the classical ES. The first regards developments regarding the adaptation of the mutation parameters i.e. the introduction of the 1/5th rule (Reference [90]), the mutative self-adaptation (Reference [90]) and the de-randomized self-adaptation (Reference [104]). The second development regards the introduction of multi-membered ES variants. The (μ+λ) and (μ, λ) notation was introduced to categorize multi-membered ES variants: μ is the number of parents, λ is the number of offspring's and the "+" and "," refer to the Environmental-Selection operator employed. The difference between the *comma* and *plus* schemes is that the first discards the parents in each generation where the second maintains them. The (μ, λ) is usually preferred since the property of forgetting solutions may assist in escaping local optima, could reduce the impact of misadapted strategy parameters and could be more suitable for tracking the moving optimum when dynamic functions are considered.

In general the classical ES initializes μ parents. The Crossover operator (Recombination) combines the available information from a number of parents to generate a single new offspring. There are many different recombination variants which differ based on the number of parents involved (global and local) and the strategy used to produce the offspring (discrete and intermediate). Specifically, local recombination implies that two parents are involved while global suggests more. Offspring generated by discrete recombination inherent each component from each parent with equal probability while intermediate recombination determines the offspring by averaging the components of the parents. This is applied for both the decision vector and strategy parameter vector and repeated to produce λ offspring. The derived strategy parameters of the offspring are then mutated, by an adaptation strategy, and then used to mutate the offspring. The population of the next generation is determined by the considered Environmental-Selection operator, i.e. the comma or plus. Some basic characteristics of ES are the importance of the Mutation operator, specifically the normally distributed mutations, and the variety in possible Crossover operators. However, the most distinct characteristic is the inclusion of the strategy parameters within the EA-based encoding representation; they are part of the individuals and evolve similarly to the decision variables. This enables the self-adaptation of the strategy parameters which are used within the Mutation Operator. Modern ES are more frequently applied on optimization problems involving real-valued decision variables. The covariance matrix adaptation ES proposed in References [104, 105, 106] which is based on improved self-adaptation of the mutation distributions is often termed as the state-of-the-art ES algorithm. For a recent comprehensible overview of ES, the interested reader is referred to Reference [107].

Differential Evolution

DE was proposed in Reference [91] and is among the most recent and popular EAs. Its relative simplicity, scalability, and robust performance have contributed to this popularity (References [108, 92]). In brevity, the basic DE algorithm randomly initializes a population. Then, for each parent an offspring is generated which is identical to the parent except for some variables which are mutated. These are determined by the Crossover operator. The Mutation operator is implemented by adding scaled differences of selected individuals to secondary parents. How to determine these selected population members and secondary parents is based on the DE mutation strategy/variant employed, i.e. the Mating operator is part of the Mutation operator. The binomial crossover is among the most frequently employed Crossover operators for DE (Reference [92]). It first uniformly selects a single variable that will be mutated. A comparison is then performed independently for each variable, to determine if the mutation is implemented for the remaining variables. This comparison is made between a parameter (the crossover rate/probability in the range [0,1]) and a randomly generated number in the range [0,1] using the uniform distribution. Lastly, the Environmental-Selection operator is employed deterministically through pairwise comparisons based on the fitness value. A main characteristic of DE is that the Mutation operator relies on scaled differences derived from the evolving population and does not require specialized adaptation of the absolute step size (References [108, 92]). Moreover, DE is relatively simple to implement and requires a low number of parameters to be determined, i.e. a scaling factor, the crossover rate, and the population size. Greater detail on the DE algorithm can be found in References [91, 92, 108, 109].

3.2.3 Single objective constraint evolutionary optimization

Constraint optimization problems (COP) can be challenging for EAs (Reference [85]). Information regarding the feasibility of a solution is required to guide the search towards both promising and feasible regions. Mechanisms aiming towards adequately addressing the aforementioned, namely constraint handling techniques (CHT), have been included within EAs. There is a number of different categorization made for CHT (References [110, 111]). Reference [112] provided a review on such CHTs. Earlier taxonomies had been simplified to: (i) Penalty functions, (ii) Decoders, (iii) Special Operators, and (iv) Separation of objective function and constraints. In addition, current and emerging CHT are also discussed.

Penalty functions transform a COP into an unconstrained one by adding the constraint violation multiplied by a penalty factor to the objective function. The most simple penalty function is known as the death penalty. It eliminates or assigns an extremely high fitness value to all infeasible solutions. This restricts extracting information from infeasible solutions. Commonly, there are three main types of penalty functions that are employed: static (e.g. Reference [113]), dynamic (e.g. Reference [114]) and adaptive (e.g. Reference [115]). Penalty functions include a penalty factor which, usually, requires to be tuned. In brevity, static penalty functions maintain a constant penalty factor. To reduce the impact of a predetermined value of a penalty factor, dynamic penalty functions include a predefined alteration of the penalty factor. Adaptive penalty functions rely on information attained during the search to adjust the penalty factor. A review on penalty functions and EA has been provided in Reference [116].

Other approaches rely on handling the constraints and objective functions separately. For example, Reference [117] suggested employing two different evaluation functions for feasible and infeasible individuals. These assigned higher fitness values to infeasible individuals, based on their total constraint violation, such that feasible solutions are prioritized. This category includes also approaches that handle the COP in two phases. For example, Reference [118] suggested an approach where in the first phase the search focuses on identifying a feasible solution and ignores the objective function values. In the second phase, the search is focused on identifying feasible individuals with improved objective function values. Moreover, many popular and state-of-the-art CHT handle the constraints and objective function separately such as the Feasibility Rules, the Stochastic Ranking, and the ϵ-constrained method and MOO-based approaches. The first adopts a set of rules for comparing individuals that prioritizes feasible solutions (References [119, 120]). Stochastic ranking (Reference [121]) introduced a user-defined probability factor to the comparison of infeasible individuals. The approach ranks individuals based on a bubble-sort-like procedure.

It aimed towards balancing objective and penalty functions stochastically to deviate from user-defined penalty factors that could lead to over or under penalization. The ϵ constrained method (Reference [122]) introduces relaxation, to the comparison among individual, for when a solution is perceived as feasible. A COP can be transformed to MOO problem. Therefore, MOO-based approaches have been also frequently employed that consider either a bi-objective or a MOO problem by Pareto-based and Non-Pareto based schemes. These are discussed in Reference [123].

EAs employing decoders search on a decoded search space which ensures that the solutions are feasible. The decoded space is attained by mapping the feasible region of the search space and facilitate the EA search (e.g. Reference [124]). However, decoders may exhibit implementation complexity and increased computational cost (Reference [112]).

A special operator used as an CHT aims towards preserving the feasibility of a feasibility solution by Variation operators or alter an infeasible to a feasible one by a repair heuristic. A representative method based on repairing infeasible solutions and including specialized operators to maintain feasibility is presented in Reference [125]. Another example is the gradient-based mutation scheme presented in Reference [122] which is employed within a DE algorithm (including also the ϵ constrained method) to repair infeasible offspring.

Approaches employing more that one CHT (i.e. hybrid or ensembles) have been also suggested. For example, Reference [126] proposed an ensemble strategy to exploit distinct characteristics of different CHT which are applied on interacting sub-populations. Reference [127] proposed an adaptive trade-off model where different CHT are employed in different stages determined by the feasibility of the evolving population.

3.2.4 Multi-objective evolutionary optimization

Many engineering problems can be formulated as MOO problems. Such are frequently encountered in practical applications. A MOO problem involves identifying the best trade-off among more than one conflicting objective, i.e. improvements in the value of one objective result in the deterioration of another. In contrast to SOO, there is no single optimal solution but rather a number of Pareto optimal solutions. The number of Pareto optimal solutions can be large or even infinite. Therefore, it is computationally impractical to identifying the entire Pareto Front. Moreover, processing a large number of Pareto optimal solutions may be impractical as well since a single best compromise solution may be of interest (Reference [102]). Therefore, a common approach is to identify an approximation of the Pareto Front.

A classical approach for MOO is to employ a weighted-sum which involves the transformation of a MOO problem to a SOO problem. The augmented function is a weighted sum of the objective functions. The main drawback to such an approach is that the output of the transformed SOO is a non-dominated solution for the MOO (Reference [128]). However not all non-dominated solutions can be always attained (e.g. the case of a concave PF). Furthermore, the weights of each function need to be determined which is, usually, problem depended.

Population-based EAs exhibit advantages which render them applicable on many SOO problems (e.g. derivative-free, simple implementation and flexible) and are highly applicable for MOO since they can provide multiple non-dominated solutions within a single run (Reference [129]). The earliest EA approach aiming towards multiple non-dominated solutions is the Vector-Evaluated Genetic Algorithm (Reference [130]). Later, Reference [101] introduced the idea of relying on dominance and niching rather than objective scores to rank the population. Early MOEA based on dominance are the Multi-Objective Genetic Algorithm (Reference [131]), the Non-dominated

Sorting Genetic Algorithm (Reference [132]) and the Niched Pareto Genetic Algorithm (Reference [133]).

Many MOEA have been proposed since. These follow different MOEAs frameworks such as Pareto-based methods, decomposition-based methods, and indicator-based methods. Some of the main representative Pareto-based methods are the NSGA-II (Reference [66]), SPEA-2 (Reference [134]) and PAES (Reference [135]). One main representative of decomposition-based methods is the MOEA/D (Reference [136]). Representative indicator-based methods are the IBEA (Reference [137]) and the HypE (Reference [138]) algorithms. There are also other categories of MOEA, such as preference-based MOEA. A representative approach is NSGA-III (Reference [102]) which is an EA for many-objective optimization and follows the NSGA-II framework. The aforementioned categorization (Reference [13]) focuses on the Environmental-Selection operator used since it is an important difference among EAs for MOO and SOO. However, modifications can be also included in other operators as well. A review on EA-based MOO is available in Reference [13].

3.2.5 Problem-Specialized operators

Two main cases can be identified when EAs are applied to real-world optimization problems. The first regards the case where an EA developed for black-box optimization is applied on the real-world problem directly, for which no a priori information is available. An EA is then applied as a direct search procedure and this is one of the advantages of EAs (Reference [129]). The second case regards optimization problems where some a priori information is available. In this case, EAs can be applied as in the previous case or an attempt to exploit the available a priori information could be made.

More specifically, even though EAs are direct search procedures there might be also value in problem-specific EAs (Reference [139]). This depends on an EAs scope and the available problem-based information (Reference [140]). In general, a problem-specific EA may be outperformed by an all-around EA in a number of problems but exhibit an improved performance on the class of problems it has been designed for. This could lead to modifications to the EA which could introduce some bias to the search. However, the effective inclusion of such bias depends on its correctness and efficient integration within the EA. In Reference [139], some options to embed explicit and implicit knowledge representation within the EA are described. Regarding explicit knowledge representation, seeding initial solutions to the population which have been identified a priori as promising (or feasible initial solutions when a constraint optimization problem is considered) could provide benefits for the search. A second option regards the hybridization of EA (these are discussed in the next Section). A third option regards the customization of the EA Variation operators. They mention as possible options the inclusion of repair mechanisms, an adequate representation to handle infeasible solutions and problem-customized Variation operator. Lastly, possible directions to handle the EA parameters are mentioned as their values can have an impact on the performance of an EA. The importance of the selection of the EA-based encoding scheme, the CHT, and the representation of the solution space are mentioned as options for implicit knowledge representation.

3.2.6 Hybrid Evolutionary Algorithms

Hybrid EAs are understood as EAs merged with at least one different approach within a single framework (Reference [86]). There is a variety of different types of hybridization, e.g. two different EAs, an EA and a gradient-based solver, an EA and a problem-specific heuristic. Reference

[141] provides a detailed categorization and discussion on such approaches. Moreover, hybrid EA approaches are also discussed in Reference [86].

Usually, hybridization aims towards an improved performance attained by the synergy of the combined approaches due to their distinct characteristics. An example is to attain an improved trade-off between the exploration and exploitation of the search space by combining global and local search techniques. The main representative of this direction is the category of Memetic Algorithms (MA) which, based on the definition provided in Reference [142], include EA-based approaches that are enhanced by the inclusion of one or more phases of local search and/or by the use of problem-specific information. In general, MAs are inspired by Darwinian principles of natural evolution and the idea of memes (Reference [143]). In the latter, a meme is defined as a unit of cultural evolution capable of local refinements (Reference [144]). Two main MA categories are the Baldwinian and Lamarckian search algorithms (Reference [145]). The former approaches commonly select an individual that undergoes local search and for which only the fitness value is updated based on the result of the local search. In contrast, a Lamarckian search algorithm implies that the individual is replaced based on the result of the local search. In general, local search can include a broad range of techniques/approaches ranging from simple permutation heuristics searching in the vicinity of a base solution to gradient-based solvers (if applicable). A taxonomy of MA approaches is provided in Reference [146].

There are many examples of hybrid EAs. Reference [147] proposed a self-adaptive DE algorithm for SOO that employed a gradient-based solver (Quasi-Newton method) as a local search method. Moreover, the aforementioned study is mostly known for introducing self-adaptation of learning strategies and parameters to the DE algorithm. Applying local search to MOO problems is not trivial. This emanates mostly from the requirement of selecting a search direction as more than one objective functions are considered, i.e. to which extent should the local search focus on improving each objective function. The earliest work addressing this issue is Reference [148] where a GA was hybridized with a local search technique. In particular, the GA variation operators were applied and the local search method was implemented. A scalar fitness with random weights was used for determining the fitness for the selection of parents and for the local search. Employing a scalar fitness with random weights for hybrid EAs for MOO has been further examined (e.g. Reference [149]) and frequently employed rendering it among the most typical directions. A review on MA is available in Reference [150].

3.3 Metamodel/Surrogate assisted Single and Multi objective optimization by Evolutionary Algorithm

Optimization problems that involve costly or expensive function evaluations, performing a large number of function evaluations may be rendered impractical or too expensive (Reference [14]). For example, the evaluation can include a time-consuming simulation model, i.e. it could take a number of minutes to hours or even days to be completed. Another example is when a costly physical experiment is involved. Towards this direction, different levels of approximation have been introduced. Reference [151] differentiates three levels of approximation:

- Problem approximation: The original problem statement is replaced by a model that resembles it and is more computationally efficient to solve.

- Function approximation: An original function (e.g. the objective function) is replaced by an alternate and explicit expression. The latter could be a model for which its output for a given input should be an accurate estimate of the output of the original function for the same input. The expression should be more computationally efficient than computing the original function.

- Evolutionary approximation: There are two types of evolutionary approximations which are EA specific (Reference [151]). These are fitness inheritance and fitness imitation. Fitness inheritance (Reference [152]) suggests that the offspring inherit the fitness from the parents used to generate them (e.g. assigning an average of the parents' fitness). The second suggests that the distance among population members can be used to estimate their fitness and usually involves clustering techniques.

3.3.1 Approximation models: Metamodels/surrogate models

Different approximation modelling techniques have been proposed to replace the expensive simulations or experiments by cheap AMs. In the following Sections, two major techniques that are considered in this thesis, i.e. Polynomial Regression (PR) and Radial Basis Functions (RBF), are presented. Other major representatives such as Gaussian Processes (Krigin), Artificial Neural Networks and Support Vector Machines have been omitted. However, interested readers are kindly referred to Reference [153] for a detailed discussion.

Polynomial Regression

Polynomial Regression, or response surface methodology (Reference [154]), approximates a response using a polynomial function which is a weighted sum of powers and products of the input. It is commonly expressed in matrix notation $\hat{f} = \tilde{\mathbf{x}}^T \beta$ where β is a vector of coefficients which defines the complexity of the model and $\tilde{\mathbf{x}}$ is a vector of the corresponding parameters (the aggregate of powers and products based on the complexity of the model). For example, when a second-order polynomial regression model including interactions among the dimensions (d) of the input vector is considered, then $\tilde{\mathbf{x}}$ is set as follows: $\tilde{\mathbf{x}} = (x_1^2, ..., x_d^2, x_1, x_2, ..., x_{d-1}, x_d, x_1, x_2, ..., x_d, 1)$. Given a number of data points (samples) the polynomial coefficients can be determined by the least-squares method or a gradient-based method.

First-order or second-order PR models are, usually, employed in the context of optimization due to limitations arising by the number of data points required w.r.t. the order of the polynomial and the dimensions of the input vector. For example, the number of data points must exceed $(d+1)(d+2)/2$ for a second-order polynomial model when the least-squares method is employed. The prediction function using a second-order PR model is as follows:

$$\hat{f}(\mathbf{x}) = a_0 + \sum_{i=1}^{d} [b_i x_i] + \sum_{j=1}^{d} \sum_{i=1}^{d} [c_{i,j} x_i x_j] \tag{3.5}$$

where $\hat{f}(\mathbf{x})$ is the estimation for point $\mathbf{x}$ and a_0, $\mathbf{b}$ and $\mathbf{c}$ are weight coefficient parameters (elements of the vector β). The polynomial can be further simplified to exclude the interactions among variables:

$$\hat{f}(\mathbf{x}) = w_0 + \sum_{i=1}^{d} [w_i x_i] + \sum_{i=d+1}^{2d} [w_i x_{i-d}^2] \tag{3.6}$$

PR models are frequently employed for EA-based optimization (References [16, 15]). The generalization property renders them appropriate for smooth functions. However, for more complex functions (e.g. multi-modal) PR models are less suitable except if smoothing the landscape is of interest (Reference [155]). More detail on PR models can be found in Reference [154].

Radial Basis Function

RBF had been introduced in Reference [156]. An RBF is a real-valued function. Its output relies on the distance (usually Euclidean norm $\|\cdot\|$) of an input point $\mathbf{x}$ to certain points $\mathbf{c}$ called the centers. The value of a function $\hat{f}(\mathbf{x})$ at point $\mathbf{x}$ is a linear combination of a set of RBFs commonly expressed as follows:

$$\hat{f}(\mathbf{x}) = \sum_{i=1}^{k} [\lambda_i \phi(\|\mathbf{x} - \mathbf{c}^{(i)}\|)] \tag{3.7}$$

where k is the number of center points, λ are unknown weight coefficients and $\phi()$ represents the kernel function.

Given k available data points ($\mathbf{x}$), to be used as centers, and their corresponding function values, $\mathbf{F} = (f_1, f_2, ..., f_k)$, a linear system of equations can be solved to determine the values of the unknown coefficients $\lambda = (\lambda_1, \lambda_2, ..., \lambda_k)$:

$$\begin{bmatrix} \phi(\|\mathbf{x}^{(1)} - \mathbf{x}^{(1)}\|) & \cdots & \phi(\|\mathbf{x}^{(1)} - \mathbf{x}^{(k)}\|) \\ \phi(\|\mathbf{x}^{(2)} - \mathbf{x}^{(1)}\|) & \cdots & \phi(\|\mathbf{x}^{(2)} - \mathbf{x}^{(k)}\|) \\ \vdots & \ddots & \vdots \\ \phi(\|\mathbf{x}^{(k)} - \mathbf{x}^{(1)}\|) & \cdots & \phi(\|\mathbf{x}^{(k)} - \mathbf{x}^{(k)}\|) \end{bmatrix} \cdot \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_k \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_k \end{bmatrix} \tag{3.8}$$

Linear splines, cubic splines, thin-plate splines, and Gaussian functions are among the most common choices for the kernel functions.

3.3.2 Infill criteria

In general, the accuracy of the AM could be improved by the inclusion of additional data points, and a balance among local and global exploitation is desired (Reference [153]). Local exploitation aims at exploiting information provided by the AM to improve its accuracy in a region of interest. Global exploitation aims to improve the global accuracy of the AM by introducing additional data points, to the dataset, that could lie in unexplored regions of the decision space (e.g. the most spatially isolated regions). In the context of EA, the term exploration is also frequently employed to express global exploitation. The criteria for selecting these additional data points are often referred to as Infill criteria. Infill criteria are discussed thoroughly in References [157, 153].

3.3.3 Evolutionary Algorithms and Model Management

Optimization problems involving expensive functions are challenging due to the limited number of function evaluations available. Focusing on EA, this poses a restriction for their successful application. More specifically, EA rely on evolving a population for a number of generations to converge. For computationally cheap optimization problems this is not binding as a large number of function evaluations can be made available. On the contrary, for computationally expensive problems the number of available function evaluations could correspond to a limited number of generations.

In this context, EA have been frequently coupled with AMs to attain acceptable results when faced with computationally expensive optimization problems. In such cases, the AM are employed to replace in part the expensive model (Reference [158]). Hereafter, the actual model shall be referred to as as the True Model (TM). How to control the interaction among the AM and the TM during the search is often termed as model management. EA-based approaches including model management are often discussed in three categories (References [14, 15, 151, 158, 159]): (i) No Evolution control, (ii) Fixed Evolution control, and (iii) Adaptive Evolution control.

In approaches that include no Evolution control, the AM would replace entirely the TM during the search. This case could be acceptable when the AM is sufficient accurate, w.r.t. the TM, as the AM is not re-examined based on information attained during the search. However, as demonstrated in Reference [160], when the accuracy is not sufficient it could result in implications for the EA-based search, since AM could include false optima that do not exist in the TM. For example, the EA may converge to the global optimum of the AM which, however, is not the global optimum, or even an optimum, of the TM (Reference [160]). In addition, the number of optima in the AM could differ from that of the TM. Therefore, Reference [158] suggested three possible directions: (i) increasing the number of data points, (ii) employing on-line sampling, and (iii) employing regularization techniques. Off-line sampling refers to generating an AM prior to the EA search while on-line sampling suggests introducing additional data points, identified during the search, and then generating an updated AM.

In the context of on-line sampling, model management usually follows either the Evolution Control framework or the Pre-selection framework. Evolution Control, Reference [158], included two main methods to determine whether the fitness of an individual is evaluated using the AM or TM:

- Controlled individuals: A number of individuals at each generation are evaluated using the TM. The AM is used for the remaining individuals.

- Controlled generations: All individuals are evaluated using the TM in some generations. The AM is used for all individuals in the remaining generations.

A third category considers a number of sub-populations that evolve based on different AM with possibly different levels of accuracy based on island models including migrations (e.g. Reference [161]). In this case, model management can be implemented on a population-based level.

The aforementioned framework introduces parameters. For example, the Controlled individuals method introduces a parameter that is the number of individuals that are evaluated using the TM in each generation. In case these parameters are user-defined, the model management strategy is referred to as Fixed Evolution control.

However, when such parameters (e.g. the number of individuals evaluated using the TM) are adapted on-line, then the model management strategy is referred to as Adaptive Evolution control. These adaptation strategies, however, are not easy to define (References [14, 15]).

The Pre-selection framework (References [162, 163]) exhibits some similarities to the Controlled individuals method. However, the Pre-selection framework suggests that offspring are pre-screened using the AM and the selected ones are evaluated using the TM. Therefore, the Environmental-Selection operator is performed only on individuals that have been evaluated using the TM.

In the context of EA assisted by AM, a variety of different Infill criteria have been employed such as random selection, best-performing individual based on the AM, distance-based criteria, clustering-based, uncertainty in approximation, and/or ensembles. A number of representative approaches are presented in Sections 3.3.7 and 3.3.8.

Moreover, AM can take part in any EA operator. For example, Reference [164] used the AM indirectly. Its employment is restricted to the Initialization, Mutation and Crossover operators of a GA and it is utilized to pre-screen offspring among a large pool of individuals.

3.3.4 Selection of an Approximating model

The selection of an AM can have an impact on the performance of the EA (Reference [165]). In general, different factors, such as number of decision variables, objectives and constraints functions, and the computational time for building the models, have to be considered when choosing an AM. Moreover, the better choice of an AM can be problem specific, i.e. some AM can be rendered more adequate than others for that specific problem. For example, Reference [166] examined through a comparative study the performance of different AM on different problems. Their results indicate that given a particular problem, different AMs may outperform others. Reference [167] provides a comparison of popular metamodeling techniques. It discusses different metrics to evaluate the accuracy of the AM and proposes the employment of a ranking preservation indicator. Among the reviewed ones, an improved performance had been observed for RBF and Support Vector Machines.

3.3.5 Global and local Approximating model

A number of different approaches have been employed to select the data points used to build the AM, given a set of available ones. In general, the selection can depend on the AM used. For example, all available data points could be used. The motivation is that each of such data points can provide important information regarding the search space which has been costly made available, i.e. using the TM. However, the computational cost of building the AM could increase, given a large number of data points.

Moreover, the AM accuracy depends on the selected data points used to build the model and their ability to provide an approximation of the entire search space. For example, a comparison of local and global models is included in Reference [167]. They identified that for problems involving a larger number of dimensions, local AM models could be more suitable. For creating the local models, different approaches have been suggested. For example, such could be based on a distance criterion that prioritizes the nearest ones to a considered point in the search space (e.g. k-nearest neighbors). Other examples prioritize the most recent entries or the best-performing entries in the dataset. It should be mentioned that local models are frequently employed when the local search is implemented (e.g. References [155, 168]).

3.3.6 Curse and Blessing of Uncertainty

In contrast to the case where the approximation error of the AM impedes the EA-based search, it is worth mentioning that it can have also a positive impact in some cases. The negative and positive impacts of this approximation error on the EA-based search are often referred to as the *curse of uncertainty* and the *blessing of uncertainty*. These are discussed in Reference [169]. The former implies that the approximation error could mislead the search (e.g. introducing false optima). An example of the latter is the case of multimodal functions where the AM could smoothen the search space, i.e. the AM could provide a function with a lower number of minima in comparison to the TM. This is examined within a generalized framework in Reference [155].

3.3.7 Representative approaches: Single Objective Optimization by Metamodel-assisted Evolutionary Algorithms

Many Metamodel-assisted EA-based approaches are available for unconstrained SOO are available (Reference [14]). For example, Reference [158] proposed the individual-based and generation-based evolution control. They examined approach is based on the covariance matrix adaptation ES with Neural Networks. Reference [162] introduced an inexact pre-evaluation phase to a GA algorithm. This pre-selection is based on an RBF network and is built for each individual independently using a distance criterion to select a subset of the available data points within the dataset. Reference [170] suggested a surrogate-assisted DE approach. The DE approach employed includes a pool of mutation and crossover strategies (ensemble). Due to the low computational cost of the AM, it is utilized to evaluate the adequacy of the mutation strategies and parameter based on the performance of offspring. The AM is built based on the evolving population, i.e. only the individuals within the population are used.

Furthermore, Metamodel-assisted EA-based approaches for constraint SOO have been also suggested. For example, Reference [123] developed an ES approach based on Fitness inheritance. The feasibility rule is used as the CHT. In contrast to earlier approaches employing fitness inheritance, two parameters are introduced: (i) the inheritance ratio and the replacement ratio. The first determines the percentage of offspring to use the fitness inheritance mechanism and the second determines the percentage of offspring with inherited fitness values to survive into the next generation. Reference [171] presented an ES based approach coupled with the stochastic ranking CHT. Two distinct AM, based on the nearest neighbor regression model, are employed for the objective function and the penalty function. Both global and local AM are examined. A comparison of three selection strategies ((i) the best individual, (ii) random selection, and (iii) maximum distance to the nearest neighbor) had been made and they identified that the first was more appropriate. The number of offspring evaluated using the TM is controlled by iteratively re-examining the parent population ranking order based on the Stochastic Ranking CHT. In case the ranking order is altered, an additional offspring is evaluated by the TM. Reference [172] focused on high-dimensional constrained optimization problems. The approach is based on Evolutionary Programming and RBF models. A distinct RBF model is employed for the objective function and each constraint function. Offspring generated by the Mutation operator of Evolutionary Programming are pre-screened based on a set of rules that prioritizes feasibility. These serve as a CHT and, in contrast to the feasibility rules, prioritize offspring with the lower number of constraint functions violated rather than sum of the constraint violation. The best offspring produced from each parent, based on the prediction of the AM, is selected to be costly evaluated.

Metamodel-assisted EA-based approaches have been proposed that include local search techniques. For example, Reference [173] developed a surrogate-assisted MA. The approach includes a global and a local phase that are implemented in each iteration. In the first, a GA is employed to generate offspring. A Gaussian Process model is used to pre-select offspring that shall undergo local search. Then, an RBF model is built by using the nearest neighbours and a gradient-based local search is implemented based on the trust-region framework. Offspring generated by this procedure are evaluated using the TM and the population is updated based on Lamarckian learning. Reference [174] proposed a hybrid EA which is a MA based on Lamarckian learning. The GA approach is enhanced by a local search strategy which employs RBF models and the trust-region framework including an SQP solver to improve individuals of the evolving GA population. Reference [168] proposed a DE based approach for SOO involving inequality constraints. The approach includes a global and local phase in which new data points enter the archive for building the models. The AMs used are (i) a generalized regression neural network for global search and (ii) a local RBF model for local search. In the global phase, candidate solutions are generated by DE Variation operators and are probabilistically selected based on the feasibility rules or an uncertainty-based metric. The selected ones undergo gradient-based local search based on local RBF models built for each of these offspring. The output of local search is used to update the evolving population.

However, there are other notable derivative-free approaches for computationally expensive global optimization that are not EA-based. For example, Reference [175] is among the earliest approaches employing an RBF model. Reference [176] introduced the Efficient Global Optimization (EGO) algorithm. It is based on the Kriging model (Gaussian Process Model) and the expected improvement metric. The latter is an indicator of the level of uncertainty of the estimation of the AM for a given point provided by the Kriging model. The EGO approach attempts to identify points that maximize the expected improvement metric. In Reference [177] an RBF-based approach is used with a stochastic sampling procedure. A number of extensions of this approach have been made available for specific types of optimization problems, e.g. COP (References [178, 179]), integer optimization problems (Reference [180]) and mixed-integer optimization problems (Reference [181]).

3.3.8 Representative approaches: Multi Objective Optimization by Metamodel-assisted Evolutionary Algorithms

This Section presents some representative MAEA for MOO problems. Some approaches have extended the EGO algorithm (Reference [176]) for MOO problems. For example, Reference [182] modified the EGO algorithm for MOO problems. The MOO problem is converted into a series of SOO problems. In each iteration, a direction based on the weighting vector is selected and a new data point is determined by maximizing the expected improvement using an EA. Reference [183] proposed a multi-objective EA based on decomposition with the Gaussian process model (MOEA/D-EGO). A number of SOO problems are formulated and a Gaussian process model is built for each one. The expected improvements for the aforementioned sub-problems are optimized at each generation by the MOEA/D algorithm. Among the generated offspring, a subset is pre-screened to be costly evaluated by a rule-based selection. A MOEA/D algorithm extended with RBF network has been also presented for computationally expensive problems (Reference [184]). It employed different kernels for the RBF networks to derive different fitness landscapes. Competitive results were attained by coupling the RBF network to the MOEA/D algorithm when compared to the corresponding ones using the Gaussian process model.

MAEA for constraint MOO have been also presented. For example, Reference [185] proposed an ES-based approach for SOO, COP and MOO. The AM is based on Gaussian random field metamodels. The study focuses on different pre-screening criteria and indicates the importance of exploiting confidence information provided by local Gaussian random field metamodels. Reference [186] employed RBF models for constrained MOO problems. An RBF model is used for each objective and constraint functions. Two different variants are examined for generating new candidate points: (i) Generating uniform random individuals within the search space, and (ii) adding

Gaussian-based perturbations to an isolated individual of the attained so-far Non-Dominated Front (NDF). The aforementioned individual is selected based on its distance (objective space) to the current NDF and the one exhibiting the largest minimum distance is prioritized. The selection of new data points is based on a weighted ranking procedure that considers distances metrics for both the decision and objective spaces and prioritizes feasible non-dominated solutions among the pool generated.

In addition, hybrid MAEA that include local search have been presented. For example, Reference [187] proposed a metamodel-assisted MA based on RBF networks. An EA-based search is first implemented for a few generations to create an initial dataset which is then updated by new entries during the search. In the remaining generations pre-selection based on inexact pre-evaluation is implemented. Within the latter, local metamodels are built for each individual which are utilized for gradient-based local search. Among the newly generated offspring, the Environmental-Selection of SPEA-2 (Reference [134]) is employed to determine which ones are costly evaluated. Reference [188] suggested a multi rule approach. It employed an RBF model and included MOEA based optimization. The introduced multi-rule approach includes selection of new data points based on random sampling, the hypervolume metric, a distance metric for the decision space, and a distance metric for the objective space. The aforementioned are applied on a pool of candidate solutions generated by a MOEA employed to solve the MOO problem. Moreover, a local search (called gap optimization problem) is also implemented by employing the MOEA to search within the vicinity of an isolated (objective space) non-dominated individual. The hypervolume metric in then employed to select a new data point from the pool generated by local search. Reference [155] proposed a generalized framework for surrogate-assisted Memetic EA. It considers an ensemble of AM, including Gaussian process, RBF and PR models, to exploit the unique characteristics of each AM and smoothing of the search landscape provided by the low-order PR models. The EA is used to evolve a population for a series of generations and create an initial dataset. Then, for each individual two independent gradient-based local searches are implemented based on the weighted sum approach using randomly generated weights. In the first, the function to be optimized is based on the ensemble of the AM and in the second a low-order PR model is considered. An archive and replacement scheme is employed to determine new entries. The generalized framework also includes a MA for SOO. Reference [189] proposed the MOPLS-N algorithm. It is a population-based approach based on RBF, Tabu Search and local search. The approach focuses on a parallel framework where each worker implements local search on a center point. Center points are selected based on rules including domination and distance criteria and a memory archive of previously points considered center points. Candidate points are generated for each center point by a local search based on Gaussian perturbations. A new point to be costly evaluated for each center point is selected by employing a hypervolume-based metric or the maximum-minimum distance (decision space) that is determined probabilistically. Alternatively, local search is skipped and an offspring generated by mutating the center point is selected.

3.4 Discussion

Many EAs and MAEAs have been suggested to address black-box SOO and MOO problems. Satisfactory results have been reported for a number of different optimization problems including SOO, COP and MOO. Moreover, hybrid approaches have also been proposed. The brief introduction presented in this Chapter aimed towards providing a background for EAs and MAEAs and is not extensive since the scope of this thesis is to examine MAEA-based approaches for GEP including SMs of the short-term operation.

Successful applications have also been reported when the GEP problem is considered (References [1, 67]). The developed MAEAs, presented in the following Chapters, are based on such frameworks and include modifications driven by the characteristics of the specific optimization problems examined. These are discussed in the following Chapters.

Chapter 4

Single objective Multi-period Generation Expansion Planning by Metamodel-assisted Evolutionary Algorithms

4.1 Motivation and Aims

This Section presents a SOO Multi-period GEP approach based on MAEAs [17]. It had been motivated by the growing computational intensity of GEP that is associated with the computational requirements for capturing short-term dynamics in the context of long-term planning. The latter is driven by the increasing shares of generation by RES. More specifically, an assessment of the operating flexibility of a candidate generating fleet is required to derive optimal expansion plans when higher RES penetration levels are considered. Moreover, the impact of technical, temporal and spatial detail has been shown in the relevant literature to be important (Section 2.2.5). These in turn have motivated the development of a series of GEP models which include increased detail of the representation of the power system's short-term operation. Along with the increase in detail, the computational intensity of GEP has risen as well.

The approach aims at providing an investment road-map by capturing the impact of the resulting installed capacity on the yearly operational cost of the system by a SM. This could be of interest if a transition towards higher shares of RES generation is examined. An attempt to include increased technical and temporal detail has been made by employing the SM for representing short-term operation and assess operational flexibility requirements. The SM serves as a cost indicator of the operating cost for a candidate installed capacity. However, spatial detail has not been included and assumptions have been made. Moreover, operating flexibility is assessed by considering also non-thermal flexibility providers such as conventional hydropower and storage installations. To address computational restrictions of the SM, unit aggregation has been implemented and representative time periods have been employed.

Many EA-based approaches have been applied on GEP models. However, EA-based or MAEA-based approaches had not been identified for GEP models embedding a SM of the short-term operation of a power system that includes UCP constraints. This could be attributed to the computational cost of the such a model. To be specific, the computational burden of EA approaches is mainly driven by the number of function evaluations required to converge to a near-optimal solution if such a SM is not included (i.e. the computational cost of the objective function is low and it may be evaluated multiple times). However, when a SM is included the computational cost is driven by the cost of each simulation (i.e. a few simulations are available as the computational cost of the objective function is high). The aforementioned had been the motivation for examining a MAEA-based approach.

The main research goal had been to develop an efficient alternative for near-optimal solutions that could be used in parallel to other mature and well-established GEP models as a decision support tool. The developed problem-customized MAEA is based on DE and surrogate-assisted frameworks identified in the literature. Problem-specific customizations are made to the formulation of the optimization problem. Moreover, problem-specific operators are examined in an attempt to improve the algorithm's efficiency. AMs are utilized to limit the amount of computationally expensive simulations required to achieve a near-optimal solution.

The approach is intended to support the decision-making process. Therefore, its applicability is examined. Satisfactory results are presented for the examined cases and identified limitation are discussed. The gain of including the SM and the corresponding impact on the resulting installed capacity, the anticipated operation cost and generation mix is examined.

4.2 Problem statement and formulation

This Section presents the problem statement, the formulation of the cost terms used to determine the objective and constraint functions of the SOO problem, and the formulation of the SM employed to provide an estimate of the short-term operation.

4.2.1 Problem statement

The following constraint SOO problem is considered:

$$\begin{aligned} &minimize \quad f(\mathbf{x}) = f^{chp}(\mathbf{x}) + \sum_{yr} f^{xp}_{yr}(\mathbf{x}) \\ &s.t. \quad \mathbf{G}(\mathbf{x}) \leq 0 \\ &\qquad\;\; \mathbf{x} \in \mathbb{S} \end{aligned} \tag{4.1}$$

where $\mathbf{x} = (x_1, x_2, ..., x_n)$ is the vector of objective/decision variables, n is the number of decision variables, $f(\mathbf{x})$ is the objective function, $\mathbf{G}(\mathbf{x}) = (g_1(\mathbf{x}), g_2(\mathbf{x}), ..., g_k(\mathbf{x}))$ is the constraint function vector, k is the number of constraint functions and $\mathbb{S}$ is the search space.

The case of a computationally expensive objective function ($f(\mathbf{x})$), computationally cheap constraint functions ($\mathbf{G}^{chp}(\mathbf{x})$), bound constraints and $\mathbf{x} \in \mathbb{Z}^n$ is considered. Furthermore, it is assumed that the minimization problem is restricted by a limited computational budget. A decision variable is a positive integer ($\mathbb{Z}^+$) which represents the number of investments in a candidate technology group. These investment decisions are considered in predefined step sizes. These steps are assumed to represent unit sizes and shall be referred to as units hereafter. In particular, when referring to units, a plant/generator or groups of plants/generators is implied.

The computationally expensive objective function ($f(\mathbf{x})$) is comprised by computationally cheap and intensive/expensive terms. Each such term can be formulated as a function of the decision vector. The sum of cost terms (e.g. investment and Fixed operational and maintenance (FO&M) cost) that do not require the output of a SM are termed as a computational cheap function ($f^{chp}(\mathbf{x})$). Correspondingly, cost terms (e.g. yearly variable cost) that require the output of a SM are termed as computationally expensive functions ($f^{xp}_{yr}(\mathbf{x})$). The consideration of a function as cheap or intensive is limited to the computational resources required to compute it (e.g. time) and not to its actual impact on the objective function value.

4.2.2 Formulation of the cost terms

Cost terms are formulated as vectors. Each element of the vector represents the corresponding cost in a year of the planning horizon. The following cost terms, which are frequently encountered in GEP models, are considered (nomenclature are described in the List of Symbols):

1. *Investment cost*: The investment cost is assumed to be paid when the unit is ordered (overnight). Therefore, the annual investment cost is formulated as the sum of the product of the annual investment decisions, the corresponding net power output per unit and the investment cost. Additionally, the salvage value of active investments at the end of the planning horizon is subtracted by the total investment cost.

$$c^{inv}_{yr} = \sum_{t}[(1 - SF_{yr,t}) \cdot P^{cap\_step}_{t} \cdot INC_t \cdot x^{inv}_{yr,t}], \forall yr \tag{4.2}$$

The salvage factor, **SF**, is computed as follows:

$$SF_{yr,t} = \begin{cases} 0, & \text{if } \; yr + CT_t + LT_t <= Yrz \\ 1, & \text{elseif } \; yr + CT_t > Yrz \\ \frac{-1+(1+DR)^{(yr+CT_t+LT_t-Yrz-1)}}{-1+(1+DR)^{(CT_t+LT_t)}}, & \text{otherwise} \end{cases} , \forall yr, t \tag{4.3}$$

Eq. 4.3 assumes that the value of a plant is equal to its investment cost and that it value remains constant during its operational lifetime (Reference [190]).

2. *Fixed operational and maintenance cost*: FO&M cost in each year is formulated as the product of each units FO&M cost and the corresponding installed capacity:

$$c^{fixed}_{yr} = \sum_{t}[FOM_t \cdot ic_{yr,t}], \; \forall yr \tag{4.4}$$

where the installed capacity is computed as the product of the number of available units and their corresponding capacity step size:

$$ic_{yr,t} = nu_{yr,t} \cdot P^{cap\_step}_{t}, \; \forall yr, t \tag{4.5}$$

The number of available units of each technology group still operational in each year considers both the construction time and operational lifetime of each investment, such that:

$$nu_{yr,t} = NU^{old}_{yr,t} + \begin{cases} \sum_{z=yr-CT_t-LT_t}^{yr-CT_t}[x^{inv}_{yr,t}], & \text{if } \; yr - CT_t > 0, yr - CT_t - LT_t \geq 1 \\ \sum_{z=1}^{yr-CT_t}[x^{inv}_{yr,t}], & \text{elseif } \; yr - CT_t > 0, yr - CT_t - LT_t < 1 \\ 0, & \text{otherwise} \end{cases} , \forall yr, t \tag{4.6}$$

3. *Variable cost*: The variable cost $\mathbf{c}^{var}$ is a vector including the sum of operating cost in each year of the planning horizon:
$$c_{yr}^{var} = \sum_{tp} [W_{tp} c_{tp}^{oc}], \forall yr \tag{4.7}$$
The operating cost is determined by a SM and the procedure is discussed in Section 4.2.3.

4.2.3 Simulation model - Problem approximation

The output of a SM is required to determine the variable cost ($\mathbf{c}^{var}$), representing the yearly operating cost. The operating cost could be estimated by a SM, e.g. a simple function or an optimization model. Moreover, given a GEP formulation and assuming a UCP optimization problem as a SM, the merged problem may be re-formulated as a large scale MILP, however, such a combined GEP and UCP model can be computationally challenging and, therefore, simplifications are introduced (References [7, 8]).

The employed simulation model

A Clustered Unit Commitment (CUC) formulation based on References [7, 8, 191] is adopted as the SM. It is a problem approximation of the traditional binary UCP which exhibits computational cost reduction and scalability properties. A CUC formulation attempts to capture operating cost and limitations of a generating fleet based on unit aggregation and considering hourly resolution; demand and production chronological profiles are preserved. Similar or identical operational characteristics are used to group units into clusters. The accepted level of similarity among units forming a cluster constitutes a trade-off between computational gain and accurate representation (Reference [191]). In general, a strict grouping criterion might lead to an improved representation of the short-term operation of a power system, however, it could also be associated with a higher computational cost. Moreover, integer variables are used to represent a unit's status. In particular, the number of units that are on-line, start-up or shut-down are represented within a cluster by integer variables, rather than binary variables that indicate the status of each unit. The aforementioned could lead to a reduction in the number of variables and constraint functions. These scale with the number of clusters rather than the number of units. For greater detail, interested readers are kindly referred to References [7, 8, 191].

The selected SM is formulated as a CUC. The MILP optimization problem represents a simplified hydrothermal UCP. It is carried out independently for examined time periods. Newly added units are assumed to be available on the beginning of each year. Additionally, RES technology groups are aggregated as a single group of units, while Hydro-Storage (HS) units and Conventional Hydro (CH) units are also represented by an aggregated unit each. An assumption is required to implement the aforementioned aggregation. Specifically, it is assumed that the generating potential of the aggregated groups (e.g. the inflows of the aggregated CH unit and the fixed capacity credits employed for RES technology groups) increases linearly w.r.t. the installed capacity which is not in general the case (Reference [48]). Moreover, variables representing RES curtailment, load shedding and reserve shortage are included within the formulation to attain a solution of the SM for a given installed capacity. This is implemented to provide a cost estimate for each examined installed capacity. Therefore, infeasibility of an installed capacity is penalized within the SM. This is required since the SM must provide an output for each set of investment decisions. The formulation of the selected SM is presented in Appendix A.

The decision variables of the SM are not included in the investment level. Therefore, the dimensionality of the latter depends only on the available technology groups and the examined investment stages. Moreover, modifications to the SM such as the addition of cost terms (e.g. emission related cost terms) or constraint functions (e.g. emission related limitations) do not affect the GEP's formulation and dimensionality. In general, other SM could, possibly, be adopted and examined.

Temporal detail

The temporal detail of the planning horizon considers a number of future years ($\mathbb{Y}$). Among the latter some or all can be selected as target years ($\mathbb{Y}^{trg} \subseteq \mathbb{Y}$). These are the years that the SM is employed to provide an estimate of the operation cost ($\mathbf{c}^{var}$). In addition, target years include a number of time periods ($\mathbb{P}$). Each such time period is comprised by a set of hourly interval ($\mathbb{H}$). More specifically, temporal detail may be tuned considering the following:

- Number of target years: The SM is employed for the years of the planning horizon that are set as target years and for the remaining ones an assumption is required. Selecting the number of target years depends on the length of the planning horizon and the computational cost of the SM. For example, the number of simulations available for each target year is expected to decrease when the number of target years increases. In addition, the computational cost of the SM can determine the number of available simulations.

- Number of time periods: The number of time periods for which the SM will be employed is also required. Setting an entire year as a time period is computationally challenging (depending on the size and formulation of the SM). Selecting an amount of representative days/weeks may be a rational alternative. In this case, the SM is employed for each time period independently under the assumption that these time periods can be simulated independently. Increasing the temporal resolution can be achieved by introducing a larger number of time periods that include a diverse set of operating conditions. However, this would also lead to an increased computational cost of the SM.

- Number of hourly intervals: Each time period is comprised by a set of consecutive hourly intervals. A larger number of hourly intervals could lead to a more accurate simulation but also lead to an increase in the computational cost of the SM.

4.2.4 Formulation of the Single-Objective optimization problem

This Section presents the objective and constraint functions.

Objective function

The objective function is set as the sum of the discounted annual investment, fixed, and variable cost.

$$TC = \sum_{yr} [DD_{yr} \cdot (c_{yr}^{inv} + c_{yr}^{fixed} + c_{yr}^{var})] \tag{4.8}$$

The yearly discount factor ($\mathbf{DD}$), based on the discount rate (DR), is computed as follows:

$$DD_{yr} = 1/(1 + DR)^{yr}, \forall yr \tag{4.9}$$

Planning constraints

The search space ($\mathbb{S}$) is set as follows:

1. *Non-negative integers variables*: The annual numbers of investments are restricted to non-negative integers:
$$\mathbf{x}^{inv} \in \mathbb{N}^{n^{inv}} \tag{4.10}$$
where n^{inv} is the dimensionality of the investment level, i.e. the product of the number of technology groups (Tz) and years of the planning horizon (Yrz).

2. *Maximum and minimum number of annual investments*: An upper bound is applied on the annual number of investments to define the search space:
$$0 \leq x^{inv}_{yr,t} \leq \overline{X}^{inv}_{yr,t},\ \forall yr, t \tag{4.11}$$
and the lower bound is set as zero. Moreover, the upper bound should consider:
   (a) A realistic estimation of the maximum annual investments ($\mathbf{X}^{max\_inv}$):
$$\overline{X}^{inv}_{yr,t} \leq X^{max\_inv}_{t}, \forall yr, t \tag{4.12}$$
   (b) No investments can be constructed after the final year of the planning horizon
$$\overline{X}^{inv}_{yr,t} = 0,\ \forall yr + CT_t > Yrz, t \tag{4.13}$$
   (c) No investments can be made for technology groups that no new units are available as investment options:
$$\overline{X}^{inv}_{yr,t^{old}} = 0,\ \forall yr, t^{old} \tag{4.14}$$
   (d) Investments can be made during years which are set as investment stages:
$$\overline{X}^{inv}_{yr,t} = \begin{cases} \overline{X}^{inv}_{yr,t}, & \text{if } yr \in \mathbb{Y}^{InvSt} \\ 0, & \text{otherwise} \end{cases}, \forall yr, t \tag{4.15}$$

3. *Planning reserve margin constraint*: The total installed capacity should be greater or equal to a predefined required minimum limit:
$$\sum_{t} [ic_{yr,t} \cdot CC_t] \geq RS_{yr},\ \forall yr \tag{4.16}$$
For the sake of simplicity, the total capacity requirement (**RS**) is set equal to the projected peak power demand increased by a reserve margin:
$$RS_{yr} = (1 + RM)PD \tag{4.17}$$

4. *Maximum installed capacity constraint*: The installed capacity of each generating technology in each year may not exceed an upper bound. The latter may be limited by an estimate of the available potential for each technology group, by restrictions posed to the installed capacity or by restrictions to the maximum generation level of a technology group:
$$ic_{yr,t} \leq \overline{IC}_{yr,t},\ \forall yr, t \tag{4.18}$$
where
$$\overline{IC}_{yr,t} = \min(\overline{IC}^{ic}_{yr,t}, (\overline{IC}^{g}_{yr,t} TD_{yr})/(8760 AV_t), IC^{ep}_{t}), \forall yr, t \tag{4.19}$$

5. *Minimum installed capacity constraint*: Correspondingly, the minimum installed capacity for each technology group in each year is limited as follows:

$$ic_{yr,t} \geq \underline{IC}_{yr,t} \ \forall yr, t \tag{4.20}$$

where

$$\underline{IC}_{yr,t} = \max(\underline{IC}^{ic}_{yr,t}, (\underline{IC}^{g}_{yr,t} TD_{yr})/(8760 AV_t), 0), \forall yr, t \tag{4.21}$$

4.3 Single-objective optimization approach

The optimization approach is based on DE and frameworks for surrogate-assisted derivative free algorithms. In addition, modifications have been introduced in the form of specialized operators. These are discussed in the following Sections.

4.3.1 Differential Evolution

The basic DE algorithm randomly initializes a population of NP individuals. The Mutation, Crossover and Environmental-selection operators are applied to improve the initial population throughout a series of generations. For each individual (target vector), an offspring (trial vector) is generated through the Mutation and Crossover operators. The Environmental-selection operator determines the trial vectors that replace the corresponding target vectors of the population.

4.3.2 Basic Differential Evolution operators

This Section presents the DE algorithm employed and its operators:

1. *Variable representation*: The following common notation is adopted for representing the i^{th} individual of the population:

$$\mathbf{x}^{gen}_i = \{x^{gen}_{i,1}, x^{gen}_{i,2}, \ldots, x^{gen}_{i,D}\} \tag{4.22}$$

where D is the number of decision variables. The decision variables of the i^{th} individual ($\mathbf{x}^{gen}_i$) are set as the investment decision variables ($\mathbf{x}^{inv}$) for which the value is not predetermined, i.e. the upper and lower bounds are not equal ($\overline{X}^{inv}_{yr,t} \neq 0$). The upper and lower bounds of the search space ($\overline{\mathbf{x}}$ and $\underline{\mathbf{x}}$) are determined in a similar manner based on the variable each element refers to. For computing the objective and constraint functions, the values of $\mathbf{x}^{gen}_i$ are remapped to determine the corresponding values of $\mathbf{x}^{inv}$. In addition, Eq. 4.6 is used as a transformation function to determine the number of installed units ($\mathbf{nu}$). Therefore, $\mathbf{nu}$ are not considered as decision variables within the individual representation. In addition, the constraint functions that determine the number of units (Eq. 4.6) are excluded since utilizing them as transformation functions can guarantee their feasibility. Consequently, the objective and constraint function are computed by using the derived values of $\mathbf{x}^{inv}$ and $\mathbf{nu}$.

2. *Initialization operator*: The initial population is randomly sampled within the search space:

$$x^1_{i,j} = \underline{x}_j + (\overline{x}_j - \underline{x}_j) \cdot U_{i,j}, \forall i, j \tag{4.23}$$

where $U_{i,j}$ denotes a randomly sampled number based on the uniform distribution, within $[0, 1]$, and drawn independently for each j^{th} variable of the i^{th} individual.

3. *Mutation operator*: The operator generates a mutant vector, $\mathbf{v}_i^{gen}$, based on scaled pairwise differences of selected individuals from the population. The DE/rand/1 mutation scheme has been adopted:

$$\mathbf{v}_i^{gen} = \mathbf{x}_{r_1^i}^{gen} + F(\mathbf{x}_{r_2^i}^{gen} - \mathbf{x}_{r_3^i}^{gen}) \tag{4.24}$$

where indices r_1^i,r_2^i and r_3^i are mutually exclusive integers randomly chosen once for each mutant vector from the range $[1, NP]$, such that $r_1^i \neq r_2^i \neq r_3^i \neq i$ and F is a user-defined parameter for scaling the difference vectors. A bound handling method is then usually applied to ensure the box constraints. The following is adopted:

$$v_{i,j}^{gen} = \begin{cases} \min\{\overline{x}_j, 2\underline{x}_j - v_{i,j}^{gen}\}, & \text{if } v_{i,j}^{gen} < \underline{x}_j \\ \max\{\underline{x}_j, 2\overline{x}_j - v_{i,j}^{gen}\}, & \text{if } v_{i,j}^{gen} > \overline{x}_j \end{cases}, \forall i, j \tag{4.25}$$

4. *Crossover operator*: A trial vector, $\mathbf{u}_i^{gen}$, is formed by the crossover operator using the corresponding target ($\mathbf{x}_i^{gen}$) and mutant ($\mathbf{v}_i^{gen}$) vectors. The binomial and exponential crossover are the two most common Crossover operators employed within DE algorithms (References [108, 92]).

   (a) *Binomial crossover*: It is performed on each of the D decision variables to determine which values are selected for the trial vector from the mutant vector as follows:

   $$u_{i,j}^{gen} = \begin{cases} v_{i,j}^{gen}, & \text{if } rand_j \leq CR \text{ } or \text{ } j = jrand \\ x_{i,j}^{gen}, & \text{else} \end{cases}, \forall j \tag{4.26}$$

   where $rand_j$ is a uniformly distributed random number in $[0, 1]$ generated anew for each j-th dimension, $jrand$ is a randomly chosen index from $[1, D]$ and the crossover rate, CR, is a user-defined parameter.

   (b) *Exponential crossover*: For implementing the exponential crossover the values of two integer are required to be determined. The first integer (n^{cr}) is randomly chosen from $[1, D]$. This integer acts as the point, in a target vector from where the exchange of components with a mutant vector initiates. The second integer, n^l, suggest the number of components that the trial vector will inherit from the mutant vector. Its value is determined stochastically. In particular, its value is initially set as $n^l = 1$. Then, a uniformly distributed random number in $[0, 1]$ is generated and compared to the crossover rate (CR). n^l is increased in steps of one (1) when the random value is lower than the value of CR and n^l does not equal D. The process is repeated until either of the aforementioned conditions no longer hold. Once the values of n^{cr} and n^l are determined the i^{th} trial vector is obtained as follows:

   $$u_{i,j}^{gen} = \begin{cases} v_{i,j}^{gen}, & \text{if } j = \langle n_{cr} \rangle_D, \langle n^{cr} + 1 \rangle_D, \ldots, \langle n^{cr} + n^L - 1 \rangle_D, \\ x_{i,j}^{gen}, & \text{else} \end{cases} \forall j \tag{4.27}$$

   where the angular brackets $\langle . \rangle_D$ denote the modulo function with modulus D.

5. *Environmental-Selection operator*: A pairwise comparison is implemented to determine the trial vectors ($\mathbf{u}_i^{gen}$) that replace the corresponding individuals ($\mathbf{x}_i^{gen}$) and form the population

of the next generation:

$$\mathbf{x}_i^{gen+1} = \begin{cases} \mathbf{u}_i^{gen}, & \text{if } \ f(\mathbf{u}_i^{gen}) \leq f(\mathbf{x}_i^{gen}) \\ \mathbf{x}_i^{gen}, & \text{else} \end{cases} \tag{4.28}$$

Eq. 4.28 assumes unconstrained minimization of the objective function. For constraint optimization problems a CHT is required. Many constraint handling techniques have been suggested and employed with DE for constraint optimization problems (References [108, 92]). Among the aforementioned, the Feasibility Rule (FR) (Reference [119]) is considered. FR is a parameter-free CHT, which treats the values of the objective and constraints functions separately. DE algorithms coupled with FR, usually, implement pairwise comparisons among individuals (trial and target vectors). The selection is based on the following rules that prioritizes feasibility of the constraint functions:

- Given two infeasible individuals, the one exhibiting the lower constraint violation is selected.
- Given a feasible individual and an infeasible one, the former is selected.
- Given two feasible individuals, the one having the lower objective function value is selected.

The constraint violation value of the i^{th} individual can be computed as follows:

$$cv_i = \sum_k [\max\{g_k(\mathbf{x}_i^{gen}, 0\}] \tag{4.29}$$

Differential Evolution for discrete optimization problems

Some modifications are usually adopted when DE is employed for optimization problems involving discrete decision variables (Reference [108]) as DE is commonly employed for real-coded variables. For example, DE variants modified by transformation functions have been applied to a number of optimization problems including discrete variables such as the UCP (e.g. References [81, 83]) and the flow shops scheduling problem (e.g Reference [192]).

The approach for handling the discrete decision variables described in Reference [193] is selected. Based on the latter, the DE operators may be applied on continuous values (floating-point/real coded representation) and a truncation could be implemented to convert the real-valued points to integers when the objective and constraint functions are required to be evaluated. Therefore, the decision vector of an individual is not altered but a modified vector is derived from the latter and used for computing its fitness. A nearest integer function is selected for this truncation to enforce the integer constraints when computing the aforementioned functions. In addition, modifications are introduced to the upper and lower bounds maintaining (a nearly) equal probability for each integer to occur. For example, for the nearest integer function ($\lfloor \cdot \rceil$) employed:

$$\overline{x'}_j = \overline{x}_j + 0.5 - e, \forall j \tag{4.30}$$

$$\underline{x'}_j = \underline{x}_j - 0.5, \forall j \tag{4.31}$$

where e is a sufficiently small number.

Restart mechanism

Restart mechanisms have been employed in DE to address the issue of stagnation (Reference [92]). During specific stages the algorithm may result in (nearly) zero difference vectors, indicating that the DE algorithm has converged, and/or failed to improve any individual for a series of generations. A simple restart mechanism is employed to reinitialize some individuals. The criterion for implementing the restart is set as a predefined number of generations (gen^{rst}) for which no improvement has occurred in any individual. If the criterion is met, all individuals, except the best population member, are reinitialized within the search space. Then, the DE algorithm proceeds to the next generation. This restart may provide larger difference vectors, however, it does not guarantee that the algorithm will escape a local optimum.

4.3.3 Examined Problem-customized operators

The DE algorithm is employed to serve as the search engine. Three simple problem-customized operators are examined in an attempt to enhance its performance. Specifically, the three operators are:

- *A Randomized Repair Heuristic*: It attempts to remap infeasible solutions within the feasible region based on problem-customized criteria. Therefore, the operator aims towards addressing challenges emerging due to the constraint nature of the optimization problem.

- *A perturbation operator*: It is based on perturbation schemes frequently encountered in combinatorial optimization and it is included to introduce diversity to the DE population by implementing perturbations in an attempt to moderate the impact of handling the discrete optimization variables.

- *A Technology-group operator*: The operator determines the EA-based representation of an individual and includes a modification to the crossover scheme in an attempt to exploit a priori knowledge of variable linkages.

The aforementioned operators are analysed in the following Sections.

Randomized Repair Heuristic

The Randomized Repair Heuristic (RRH) is presented in two Section. First, the aim and motivation of developing the repair procedure is presented and then a description of the operator follows.
Aim and Motivation: There had been two main motivations for the RRH:

- The first goal is to repair infeasible solutions generated during the search. These infeasible solutions are generated since the Variation operator cannot restrict their generation as applied in this approach. Infeasibility regarding the planning constraints is considered and a distinction is made among the aforementioned:

  - Constraints imposed on the installed capacity of each technology group (Eq. 4.20 and 4.18) are technology group specific and the repair procedure can be implemented independently on the set of variables referring to a technology group.

– Constraints imposed on the total installed capacity (Eq. 4.16) which is a function of investment decisions from technology groups that include investment stages preceding the year that the constraint function considers and for which an implemented investment can be operational by that year.

Moreover, the infeasible solutions (i.e. the input to the RRH) do not violate constraints Eq. 4.11 and Eq. 4.10 as these are ensured by applying Eq. 4.25 and the nearest integer function, respectively. Therefore, the aim of RRH is to repair constraint violation of Eq. 4.16, Eq. 4.18 and Eq. 4.20 while maintaining the feasibility of Eq. 4.10 and Eq. 4.11.

- Secondly, the RRH aims towards providing a diverse set of feasible solutions. Introducing user-defined preferences has been avoided. Therefore, the repair procedure is implemented in a stochastic manner following a set of predefined rules. Therefore, the bias introduced by the inclusion of the repair heuristic has been limited.

The RRH can be applied to repair infeasible solutions generated by the DE algorithm and by the refining strategies (presented in the Section 4.3.4).

- Within the DE algorithm, the RRH is applied on each generated infeasible individual. First, a matrix $\mathbf{x}$ is defined by remapping the rounded values of an individual (e.g. $\mathbf{x}_i^{gen}$ or $\mathbf{u}_i^{gen}$). The final output of RRH, i.e. the repaired $\mathbf{x}$, is used to alter vector $\mathbf{x}_i^{gen}$ by replacing the corresponding values. In addition, the resulting integer values of $\mathbf{x}_i^{gen}$ are stochastically set as real values:
$$x_{i,j}^{gen} = x_{i,j}^{gen} - 0.5 + U_{i,j}, \ \forall j = 1, 2, \ldots, D \tag{4.32}$$
where $U_{i,j}$ denotes a randomly sampled number based on the uniform distribution, within $[0, 1)$. This is implemented due to the adoption of a real-valued representation for the DE algorithm.

- In a similar manner, an infeasible solution (vector) generated by perturbations is remapped to a matrix form ($\mathbf{x}$) when RRH is applied to repair infeasible solutions generated by the refining strategies (Section 4.3.4). However, the resulting integer values of the RRH are not stochastically set as real values in this case.

Description of the operator: The RRH includes three parts (namely Part A, B and C). The output of each part serves as the input of the next. The input of Part A is the examined matrix $\mathbf{x}$ that does not violate Eq. 4.10 and Eq. 4.11. Its output is modified (if required) to satisfy also Eq. 4.20. The output of Part B is modified (if required) to meet also Eq. 4.18. Lastly, the output of Part C satisfies the planning constraints (Eq. 4.10, Eq. 4.11, Eq. 4.16, Eq. 4.18 and Eq. 4.20). A step-by-step description of Part A, B and C follows:

1. Part A: The first part of the RRH considers Eq. 4.20. These constraints regard the minimum limit of installed capacity of each technology group in each year of the planning horizon. Consequently, the RRH can be applied independently for each such group and the process is repeated sequentially until all lower limit constraints (Eq. 4.20) are satisfied as follows:

   (a) A forward sweep is implemented to identify if the lower limit of the installed capacity for an examined technology group (t^{trial}) is violated in any year.

   (b) If no violation is identified, skip this step. Otherwise:

i. Stochastically selected (with equal probability) an investment stage (yr^{trial_sel}) prior to the first identified violation (e.g. yr^{vio}) from a pool of possible investment stages ($\mathbb{Y}^{trial}$). The pool includes the indices of the years (yr^{trial}) for the examined technology group for which:

A. An added investment can be operational (constructed) by the considered year: $yr^{trial} + CT_{t^{trial}} \leq yr^{vio}$.

B. An added investment can remain operational by the considered year: $yr^{trial} + CT_{t^{trial}} + LT_{t^{trial}} \leq yr^{vio}$.

C. An added investment does not violate the maximum number of investments (Eq. 4.11): $x_{yr^{trial},t^{trial}} + 1 \leq \overline{X}_{yr^{trial},t^{trial}}$

ii. Determine the number of investments to be stochastically added (n^{add_st}) to the selected investment stage (yr^{trial_sel}). It is set as a randomly generated integer drawn from the range $[0, \overline{X}_{yr^{trial\_sel},t^{trial}} - x_{yr^{trial\_sel},t^{trial}}]$.

iii. Update the investment variables: $x_{yr^{trial\_sel},t^{trial}} = x_{yr^{trial\_sel},t^{trial}} + min(n^{add\_st}, \lceil (\underline{IC}_{yr^{trial\_sel},t} - IC_{yr^{trial\_sel},t^{trial}})/P^{cap\_step}_{t^{trial}} \rceil)$.

iv. Update the installed capacity (**ic**) and the number of units (**nu**).

(c) If all technology groups have been repaired then proceed to part B. Otherwise, repeat the steps (1a-1c) for the remaining technology groups.

2. Part B: The second part of the RRH considers (Eq. 4.18). These constraints regard the maximum limit of installed capacity of each technology group in each year of the planning horizon. Similarly to part A, the RRH is applied independently for each technology group and is repeated until all lower and upper limits (Eq. 4.20 and Eq. 4.18, respectively) are satisfied as follows:

(a) A forward sweep is implemented to identify if the upper limit of the installed capacity for a technology group (t^{trial}) is violated in any year.

(b) If no constraint violation is identified (Eq. 4.18), skip this step. Otherwise, remove the most recently constructed investments prior to the year where the violation occurs as follows:

i. The pool of possible investment stages ($\mathbb{Y}^{trial}$) is determined to include the indices of the years (yr^{trial}) for the examined technology group for which:

A. An investment is operational by the year the violation occurs: $yr^{trial} + CT_{t^{trial}} \leq yr^{vio}$

B. An investment remains operational by the year the violation occurs: $yr^{trial} + CT_{t^{trial}} + LT_{t^{trial}} \geq yr^{vio}$

C. At least one investment can be removed: $x_{yr^{trial},t^{trial}} > 0$

ii. Select the investment stage (yr^{trial_sel}) of the investment constructed preceding the first identified violation (e.g. yr^{vio}). In particular, set yr^{trial_sel} as the largest value of $\mathbb{Y}^{trial}$.

iii. Determine the number of investments that should be removed (n^{rem}) for the selected year (yr^{trial_sel}). It is set as the number of investments occurring in the selected investment stage: $n^{rem} = x_{yr^{trial\_sel}}$

iv. Update the investment variables:
$x_{yr^{trial\_sel},t^{trial}} = x_{yr^{trial\_sel},t^{trial}} - \min(n^{rem}, \lceil (IC_{yr^{trial\_sel},t^{trial}} - \overline{IC}_{yr^{trial\_sel},t})/P^{cap\_step}_{t^{trial}} \rceil)$.

v. Update the installed capacity (**ic**) and the number of units (**nu**).

(c) Update the lower limit (Eq. 4.20) constraint violation of the examined technology group (t^{trial}).

(d) If no constraint violation (Eq. 4.20) is identified, skip this step. Otherwise, add investments to the examined technology group in a similar manner to Part A. However, investments are added to investment stages prior to the year where the violation is identified by prioritizing the preceding ones rather than stochastically selecting them. Moreover, the number of added investments is always set as the required one to satisfy the constraint violation (if possible) rather than determining it stochastically. The steps are the following:

i. Select the possible investment stage (yr^{trial_sel}) preceding to the first identified violation (e.g. yr^{vio}). The pool of possible investment stages ($\mathbb{Y}^{trial}$) includes the indices of the years (yr^{trial}) for the examined technology group for which:

A. An added investment can be operational (constructed) by the considered year: $yr^{trial} + CT_{t^{trial}} \leq yr^{vio}$.

B. An added investment can remain operational by the considered year: $yr^{trial} + CT_{t^{trial}} + LT_{t^{trial}} \leq yr^{vio}$.

C. An added investment does not violate the maximum number of investments (Eq. 4.11):
$x_{yr^{trial},t^{trial}} + 1 \leq \overline{X}_{yr^{trial},t^{trial}}$

ii. Set yr^{trial_sel} as the largest value of $\mathbb{Y}^{trial}$.

iii. Determine the number of investments that should be added (n^{add}) from the selected investment stage (yr^{trial_sel}). It is set as the number of available remaining investments: $n^{add} = \overline{X}_{yr^{trial},t^{trial}} - x_{yr^{trial},t^{trial}}$.

iv. Update the investment variables:
$x_{yr^{trial\_sel},t^{trial}} = x_{yr^{trial}\_sel,t^{trial}} + \min(n^{add}, \lceil (\underline{IC}_{yr^{trial\_sel},t} - IC_{yr^{trial\_sel},t^{trial}})/P^{cap\_step}_{t^{trial}} \rceil)$.

v. Update the installed capacity (**ic**) and the number of units (**nu**).

vi. Repeat steps 2c-2d

(e) Update the constraint violation of the upper limit of the installed capacity for the examined technology group (t^{trial}).

(f) If no constraint violation is identified (Eq. 4.18), skip this step. Otherwise, repeat steps 2a-2f.

(g) If all technology groups have been repaired then proceed to part C. Otherwise, repeat the steps (2a-2g) for the technology group not meeting the aforesaid condition.

3. Part C: The third part of the RRH considers Eq. 4.16. These constraints regard the planning reserve margin in each year of the planning horizon. A forward sweep is implemented where identified violations of the planning reserve margin constraint are sequentially repaired. The part alters the investment decisions to serve the aforementioned constraint while ensuring that the remaining constraints are not violated. It is implemented as follows:

(a) Apply a forward sweep to identify if the planning reserve margin constraint is violated in any year.

(b) If no constraint violation is identified (Eq. 4.18), skip this step. Otherwise, stochastically (equal probability) select an investment stage prior to the first year for which a violation is identified (yr^{vio}) and add an investment to a stochastically (equal probability) selected technology group in which one can be added in the selected investment stage. This is iteratively repeated until the constraint violation of identified in the examined year (yr^{vio}) is met. The aforementioned are implemented as follows:

   i. Set the pool of possible investment stages ($\mathbb{Y}^{trial}$) to include the indices of the years (yr^{trial}) for which:

      A. $yr^{trial} - \underset{\forall t}{min}(CT_t) \leq yr^{vio}$

      B. $yr^{trial} + \underset{\forall t}{max}(CT_t + LT_t) \geq yr^{vio}$

   ii. Set the pool of possible technology groups ($\mathbb{T}^{trial}$) to include the indices of all technology groups.

   iii. Stochastically select an index yr^{trial_sel} from $\mathbb{Y}^{trial}$ with equal probability.

   iv. Remove technology group indices from $\mathbb{T}^{trial}$ that do not satisfy the following conditions:

      A. An added investment does not violate the maximum number of investments (Eq. 4.11):
      $x_{yr^{trial},t^{trial}} + 1 \leq \overline{X}_{yr^{trial},t^{trial}}$

      B. An added investment can be operational (constructed) by the considered year:
      $yr^{trial\_sel} + CT_{t^{trial}} \leq yr^{vio}$

      C. An added investment will remain operational by the considered year:
      $yr^{trial} + CT_{t^{trial}} + LT_{t^{trial}} \geq yr^{vio}$

      D. An added investment does not violate the upper limit of installed capacity:
      $ic_{yy,t} + P^{cap\_step}_{t^{trial}} \leq \overline{IC}_{yy,t^{trial}}, \ \forall yy \in [yr^{trial} + CT_{t^{trial}}, \ldots, yr^{trial} + CT_{t^{trial}} + LT_{t^{trial}}]$

   v. If $\mathbb{T}^{trial}$ is an empty set, then remove yr^{trial_sel} from $\mathbb{Y}^{trial}$. Otherwise,

      A. Stochastically select an index from $\mathbb{T}^{trial}$ with equal probability.

      B. Update the investment variables:
      $x_{yr^{trial\_sel},t^{trial}} = x_{yr^{trial}\_sel,t^{trial}} + 1.$

      C. Update the installed capacity (**ic**) and the number of units (**nu**).

      D. Compute the constraint violation (Eq. 4.18) for year yr^{vio}.

   vi. If $\mathbb{Y}^{trial}$ is not an empty set, then skip this step. Otherwise,

      A. Stochastically (equal probability) select a technology group index (t^{trial}) from a pool that includes the indices of all technology groups.

      B. Update the investment variables:
      $x_{yr,t^{trial}} = \overline{X}_{yr,t^{trial}}, \forall yr$

      C. Update the installed capacity (**ic**) and the number of units (**nu**).

      D. Apply Part B for technology group t^{trial}.

      E. Repeat steps 3a-3(b)vi.

vii. If a constraint violation (Eq. 4.18) for year yr^{vio} is not identified, then repeat steps 3a-3b. Otherwise, repeat steps 3(b)ii-3(b)vii.

(c) The updated matrix $\mathbf{x}$ of RRH is provided as the output of the procedure.

Step 3(b)vi is included for the case where a feasible solution cannot be attained due to investments placed in investment stages following the year in which the planning reserve margin constraint violation occurs. The addition of an investment would lead to a constraint violation of the upper limit of the installed capacity of a technology group and, therefore, $\mathbb{Y}^{trial}$ would result in an empty set. An illustrative example is provided: Assume the case of a single technology group, a number of investment stages and an upper limit of installed capacity enabling the inclusion of a single investment. Given the aforesaid, Part C is unable to produce a feasible solution without step 3(b)vi if the year in which the investment becomes operational follows the one where a constraint violation of the planning reserve margin occurs.

Technology-group operator

In this Section, the aim and motivation for developing the technology-group operator is presented. Then a description of the operator follows.

Aim and Motivation: Reference [194] examined the impact of unnatural dependencies between adjacent variables on the performance of the exponential crossover operator in widely used synthetic benchmarks for black-box optimization. They identified that such unnatural dependencies can be exploited by the exponential crossover operator. Arbitrary assuming such linkages, on the other hand, may result in unintended biases. They suggested the Shuffled-Exponential crossover to be coupled with the DE algorithm (instead of exponential crossover) to eliminate such unnatural dependencies, unless there is some a priori knowledge that there are dependencies between consecutive variables. Motivated by the above, an operator that attempts to exploit a priori knowledge of some dependencies is examined.

Some dependencies can be identified regarding the optimization problem considered in this Chapter, since it is not a black-box optimization problem. These dependencies, even though not strictly defined, are assumed to exist between variables referring to the same technology group. This assumption is made since the investment decision variables of a technology group form a trajectory of the installed capacity of the same technology group for the planning horizon. Each trajectory is in term bounded by constraint functions (upper and lower limit of installed capacity) which are a function of a subset of the optimization variables, i.e. decision variables of the same technology group. Assuming a feasible solution, the number of such trajectories altered (more precisely the number of decision variables from different technology groups) can affect the number of constraint violations of the installed capacity limit triggered (Eq. 4.18 and Eq. 4.20). On the other hand, optimizing each trajectory independently is not applicable, since the remaining functions (planning reserve margin constraints and the objective function) require the optimization of the decision variables of all technology groups. The operator attempts to exploit the aforesaid linkages/dependencies among some variables by (i) predefining the lexicographic order of the decision variables within the individual (decision vector) representation and (ii) restricting alterations made to the target vector by the crossover operator to a subset of the technology groups.

Description of the operator: The operator includes two modifications: (i) to the individual (decision vector) representation and (ii) to the crossover operator employed.

- Individual (decision vector) representation: The decision variables are grouped in technology blocks to focus on exploiting linkages among decision variables which refer to the same technology group. In particular, all decision variables of a technology group are set adjacent, within the corresponding block, and in ascending order, based on the investment stage that it refers to. Therefore, the i_{th} individual of the population is represented as follows:

$$\mathbf{x}_i^{gen} = \{x_{i,1}^{gen}, x_{i,2}^{gen}, \ldots, x_{i,D}^{gen}\} = \{\mathbf{x}_{i,1}^{gen\_blk}, \mathbf{x}_{i,2}^{gen\_blk}, \ldots, \mathbf{x}_{i,Tz}^{gen\_blk}\}, \tag{4.33}$$

  where

$$\mathbf{x}_{i,t}^{gen\_blk} = \{x_{i,t,1}^{gen}, x_{i,t,2}^{gen}, \ldots, x_{i,t,D_t^{blk}}^{gen}\}, \forall t \tag{4.34}$$

  where D_t^{blk} denotes the dimensionality of block t. For example, $x_{i,t,2}^{gen}$ corresponds to the decision variable of the second available investment stage of technology group t.

- Crossover operator block modification (blk): A crossover operator is implemented for a number of blocks independently for each individual rather than once for each individual. An integer, n_{blk}, is employed to stochastically determine the number of blocks for which the crossover operator will be implemented. The integer, n_{blk}, is randomly selected from the range $[1, Tz]$ with equal probability. Consequently, decision variables of some technology groups are excluded as possible values to be inherited to the trial vector (from the mutant vector) for values of integer n_{blk} lower than Tz.

The modification affects the number of variables transferred from the mutant to the trial vector. This can differ based on the selected crossover operator, the number of decision variables, the number of blocks and the dimensionality of each block. Moreover, each trajectory of a technology group exhibits an equal probability to be selected to be altered regardless of the size of the corresponding block (dimensionality). In addition, the probability of adopting variables of the same technology group from a mutant vector, while rejecting those of other groups may be increased by applying the technology-group operator. For example, if Tz technology blocks of equal dimensionality, the binomial crossover scheme and a crossover rate equal to 1 are assumed, then the probability that the trial vector adopts variables from both technology groups from the mutant vector is equal to 1. In contrast, the corresponding probability is (nearly) halved when the blk is used. However, crossover operators (usually) ensure that the trial vector adopts at least one variable from the mutant vector. Therefore, small block sizes or low values of the crossover rate have an impact on the aforementioned effect. For example, for the special case of Cr=0 (which is not commonly used in DE algorithms) at least one value would be adopted from the mutant vector for each select block and consequently for $n_{blk} > 1$ the aforesaid does not hold.

A step-by-step description for applying a selected crossover operator, including the modification, on the i$^{\text{th}}$ individual follows:

1. Create set $\mathbb{T}^{trial}$ which is a randomly shuffled permutation of set $\mathbb{T}$.

2. Set the trial vector equal to the target vector:
   $\mathbf{u}_i^{gen} = \mathbf{x}_i^{gen}$

3. Set n_{blk} as a randomly generated integer from the range $[1, Tz]$.

4. Set counter c equal to one.

5. If $c > n_{blk}$ then skip this step. Otherwise, apply the crossover operator to the selected block:
   (a) Set t^{trial} as $\mathbb{T}_c^{trial}$.
   (b) Apply the crossover operator among $\mathbf{x}_{i,t^{trial}}^{gen\_blk}$ and $\mathbf{v}_{i,t^{trial}}^{gen\_blk}$ to generate and replace the block $\mathbf{u}_{i,t^{trial}}^{gen\_blk}$ of the trial vector $\mathbf{u}_i^{gen}$. $\mathbf{v}_{i,t^{trial}}^{gen\_blk}$ is the corresponding block of the vector $\mathbf{v}_i^{gen}$ generated by the mutation operator.
   (c) Update the counter:
   $c = c + 1$
   (d) Repeat step 5.

Perturbation operator

In this Section, the aim and motivation for examining the perturbation operator (PO) is discussed. Then a description of the operator follows.
Aim and Motivation: There had been two main motivations for the inclusion of a PO:

- To improve the timing of the investments (investment stages). An example is presented that illustrates when such might be required: if an individual of the DE population $\mathbf{x} = [\{0.2, 0.3, .9, .8\}, \{0.3, 0.2, 4.9, 0\}]$ is assumed, then the individual generated by applying the rounding function is $\mathbf{x} = [\{0, 0, 1, 1\}, \{0, 0, 5, 0\}]$. In addition, if it is assumed that the optimal solution is $\mathbf{x}^* = [\{0, 1, 0, 1\}, \{0, 0, 5, 0\}]$, then it can be seen that the optimal number of investments has been identified, however, a single investment of the first block is misplaced. The optimal solution can be generated by the DE variation operators by different ways depending on the remaining individuals within the population. If it is assumed that x is set as a base vector, the difference vectors result in scaled differences within the range [0.2, 1.2) and [-1.4,-0.4) for the second and third decision variables, respectively, and the crossover operator selects both decision variables to be inherited from the mutant to the trial vector. For the remaining decision variables the difference vectors should be sufficiently low to ensure that the corresponding integer values are not altered or such alterations are rejected by the crossover operator. Moreover, these alterations could be made in steps (a number of generations):

  1. An increase in value of the third decision variable while maintaining the value of the second one (the following trial vector: $\mathbf{u} = [\{0, 1, 1, 1\}, \{0, 0, 5, 0\}]$). However, the trial vector could possibly result in a higher value of the objective function (induced by higher investment and fixed cost and an insufficient reduction of the operating cost). In this case, the generated trial vector would be rejected by the Environmental-selection operator.
  2. A decrease in value of the second decision variable while maintaining the value of the third one (a trial vector: $\mathbf{u} = [\{0, 0, 0, 1\}, \{0, 0, 5, 0\}]$). This could be an infeasible solution which violates the lower limit of installed capacity. In this case, the trial vector would undergo the repair process (RRH) and be altered stochastically to a feasible solution if the RRH is included in the DE algorithm. Otherwise, it would be rejected by the Environmental-selection operator if the Environmental-selection operator prioritizes feasible individuals over infeasible ones (e.g. FR).

The aforementioned example, attempts to illustrate that the optimization problem can require reallocation of the values within the blocks if the optimal numbers of investments has been identified.

- To generate diverse individuals for the population. A real-coded representation has been adopted for representing the integer variables. The values of the difference vectors declines as the population's diversity declines and, in some cases, the difference vectors may not be sufficiently large to alter the integer values and escape a specific region if required. This might be of importance when the population converges prematurely. A larger perturbation may generate a diverse solution from the existing ones within the population. This solution could be infeasible and in this case the RRH is applied to remap the solution within the feasible region. Due to the stochastic nature of the latter, a diverse feasible individual may be generated.

Description of the operator: The PO is applied to a number of individuals with a probability of pf_{PO} to generate a block, $\mathbf{v}^{gen\_blk}$, for the trial vector. It includes three perturbation schemes based on common schemes employed for combinatorial optimization. The first two attempt to improve investment timing. The last scheme attempts to reduce the investment cost. More specifically, for the first two schemes a block is generated that is a permutation of the elements within the blocks of the target ($\mathbf{x}_{i,t}^{gen\_blk}$) and, corresponding, base vector ($\mathbf{x}_{r_1^i,t}^{gen\_blk}$). In the first scheme, two decision variables within a technology-group block are switched similarly to Reciprocal exchange (2-exchange). The second scheme shuffles investments decisions between a randomly determined number of investment stages. The last scheme removes the investments from a randomly determined number of investment stages of the target vector ($\mathbf{x}_{i,t}^{gen\_blk}$). These are re-initialized near the lower bound of each decision variable.

A step-by-step description for applying PO on the i$^{\text{th}}$ individual follows:

1. If a random number based on the uniform distribution from the range is greater than parameter pf_{PO}, then skip this step. Otherwise:
   (a) Create set $\mathbb{T}^{trial}$ which is a randomly shuffled permutation of set $\mathbb{T}$.
   (b) Set $\mathbf{u}_i^{gen}$ equal to $\mathbf{x}_i^{gen}$.
   (c) Set n_{blk} as a randomly generated integer from the range $[1, Tz]$.
   (d) Set counter c equal to one.
   (e) If $c > n_{blk}$ the skip this step. Otherwise, apply the crossover operator to the selected block:
      i. Randomly select an integer n_{PO} from the range $[1, 3]$.
      ii. Set t as $\mathbb{T}_c^{trial}$.
      iii. If $D_t^{blk} < 2$ then set $v_{i,t}^{gen\_blk}$ equal to the block generated by using the DE operators ($\mathbf{u}_{i,t}^{gen\_blk}$). Otherwise, set $v_{i,t}^{gen\_blk}$ as $x_{i,t}^{gen\_blk}$.
      iv. If $n_{PO} \neq 1$ or $D_t^{blk} < 2$, then skip this step. Otherwise:
         A. Randomly select two exclusive elements μ and ν from $[1, 2, \ldots, D_t^{blk}]$.
         B. Set $v_{i,t,\mu}^{gen\_blk}$ as $x_{r_1^i,t,\nu}^{gen\_blk}$.
         C. Set $v_{i,t,\nu}^{gen\_blk}$ as $x_{r_1^i,t,\mu}^{gen\_blk}$.

v. If $n_{PO} \neq 2$ or $D_t^{blk} < 2$, then skip this step. Otherwise:

A. Randomly select two elements μ and ν from $[1, 2, \ldots, D_t^{blk}]$ such that $\mu < \nu$.

B. Set $\mathbb{Y}^{trial}$ as a randomly shuffled permutation of integers from $[\mu, \mu+1, \ldots, \nu]$.

C. Set $v_{i,t,yr}^{gen\_blk}$ as $x_{r_1^i,t,\mathbb{Y}_{yr}^{trial}}^{gen\_blk}, \forall yr \in [\mu, \mu+1, \ldots, \nu]$.

vi. If $n_{PO} \neq 3$ or $D_t^{blk} < 2$, then skip this step. Otherwise:

A. Randomly select two integer μ and ν from the range $[1, 2, \ldots D_t^{blk}]$ such that $\mu < \nu$.

B. Set $v_{i,t,yr}^{gen\_blk}$ as $\underline{x}_{t,yr}^{\prime gen\_blk} + U_{yr}[0,1), \forall yr \in [\mu, \mu+1, \ldots, \nu]$.

vii. Update the counter:
$c = c + 1$

viii. Repeat step 1e.

ix. Set $u_{i,t}^{gen\_blk}$ as $v_{i,t}^{gen\_blk}$.

4.3.4 Steps of the optimization approach

This Section presents the steps of the optimization approach. It follows a similar framework to the surrogate-assisted derivative free algorithms proposed in References [168, 180]. The modifications emanate from the a priori knowledge available regarding the constraint optimization problem considered. Some main differences are the following:

- A number of AM are maintained where each is built on a subset of the variables rather than a single one built for the entire set of the decision variables. This is implemented to estimate the variable cost of each considered target year based on the assumption that each cost may be computed independently. This assumption can be made since the optimization problem is not a black-box and the installed capacity of the technology groups is a function of investment decisions. The dimensionality of each AM scales with the number of technology groups that can exhibit different values of installed capacity in each target year. The number of AMs scales with the number of target years. Two possible alternatives are to select a single AM build for all investment decision variables or to select a single AM build for the installed capacity of all target years. The first scales with the number of investment stages and the number of technology groups. The second scales with the number of target years and the number of technology groups that can exhibit different values of installed capacity in each target year. In general, the performance of an AM can decrease as the number of dimensions increases (Reference [167]).

- The approach exploits characteristics of the formulation, i.e. the objective function includes computationally cheap and computationally expensive cost terms and constraint functions.

- Planning constraint functions are repaired externally by the RRH. Surrogate-assisted derivative-free algorithms are coupled with CHT for black-box constraint optimization problems since repair operators may not be easily available for such problems.

- The AMs are refined by solutions, chosen at different stages from an evolving DE population and randomly generated solutions. The trajectories generated for the installed capacity by the investment decision variables are considered within these selection processes.

The optimization approach includes two main phases, the initialization phase and the optimization phase. The initialization phase includes four main steps: (i) pre-processing a region of interest, (ii) initializing the archives, (iii) creating the initial data points for each archive and (iv) supplying initial solutions. The optimization phase includes the steps for evolving the population of the DE algorithm and the steps for applying the refining strategies. These are discussed in the following Sections.

Initialization: Pre-processing a region of interest

A region of interest is pre-processed for the provided data input. It aims towards defining a region in which the initial data points will be created. This region is defined by an estimate of the maximum ($\overline{\mathbf{NU}}$) and minimum ($\underline{\mathbf{NU}}$) limits of available units of each technology group in each year. The limits are identified based on the planning constraints. They ensure that no feasible data point exists outside these limits that respects the planning constraints. However, the limits do not define the feasible region but rather exclude a number of infeasible data points. This aims towards selecting initial data points from a smaller region and reduce the amount of costly simulations of infeasible data points that do not necessarily facilitate the search.

Furthermore, the dimensionality d_{yr} of each target year can then be determined based on the technology groups for which the installed capacity is not predetermined. For any year that the installed capacity is predetermined for all technology groups ($\overline{NU}_{yr,t} = \underline{NU}_{yr,t}, \forall t$) no AM is required. For such target years, the operational cost needs to be computed once as its value is also predetermined. Therefore, the total AMs required, N_{sur}, is at most $Yrz - \min_t\{CT_t\}$, when all years of the planning horizon are set as target years.

The procedure employed, to predefine limits that loosely consider the planning constraints and define a region of interest, is implemented as follows:

$$\overline{NU}_{yr,t} = \overline{NU}^{new}_{yr,t} + NU^{old}_{yr,t}, \forall yr, t \tag{4.35}$$

$$\underline{NU}_{yr,t} = \underline{NU}^{new}_{yr,t} + NU^{old}_{yr,t}, \forall yr, t \tag{4.36}$$

where $\underline{\mathbf{NU}}^{new}$ and $\overline{\mathbf{NU}}^{new}$ are the minimum and maximum numbers of new units possibly installed by partially considering Eq. 4.11 - 4.20. An upper limit is computed based on Eq. 4.11 and Eq. 4.18.

$$\overline{NU}^{new}_{yr,t} = \min\left(\left\lfloor \frac{\overline{IC}_{yr,t} - NU^{old}_{yr,t} \cdot P^{net}_t}{P^{net}_t} \right\rfloor, NU^{max}_{yr,t} - NU^{old}_{yr,t}\right), \forall yr, t \tag{4.37}$$

where $\mathbf{NU}^{max}$ is the maximum amount of units in a technology group in each year computed by using Eq. 4.6 and setting $\mathbf{x}^{inv}$ as $\overline{\mathbf{X}}$. Then, a backward loop is implemented to consider the operational lifetime of newly installed units and their maximum decrease.

$$\overline{NU}^{new}_{yr,t} = \begin{cases} \min(\overline{NU}^{new}_{yr,t}, \overline{NU}^{new}_{yr+1,t} + \overline{X}_{yr-CT_t-LT_t,t}), & \text{if } \ yr - CT_t - LT_t > 0 \\ \min\{\overline{NU}^{new}_{yr,t}, \overline{NU}^{new}_{yr+1,t}\}, & \text{elseif } \ yr - CT_t > 0 \\ 0, & \text{otherwise} \end{cases}, \forall yr, t \tag{4.38}$$

The lower limit is computed by considering Eq. 4.16 and Eq. 4.20. In particular, $\mathbf{NU}^{min}$ which are the minimum units required to satisfy the planning reserve margin constraint (Eq. 4.16) is determined assuming that all other possible investments are implemented:

$$NU_{yr,t}^{min} = \frac{RS_{yr} - \sum_{t' \in \mathbb{T}, t' \neq t} [\overline{NU}_{yr,t'} \cdot P_{t'}^{net} \cdot CC_{t'}]}{P_t^{net} \cdot CC_t}, \forall yr, t \tag{4.39}$$

$$NU_{yr,t}^{min} = \max(\lceil NU_{yr,t}^{min} \rceil, 0), \forall yr, t \tag{4.40}$$

Then, the minimum number of new units is computed as follows based on Eq. 4.20:

$$\underline{NU}_{yr,t}^{new} = \max\left(\left\lceil \frac{\underline{IC}_{yr,t} - NU_{yr,t}^{old} \cdot P_t^{net}}{P_t^{net}} \right\rceil, NU_{yr,t}^{min} - NU_{yr,t}^{old}, 0\right), \forall yr, t \tag{4.41}$$

For the lower limit both a backward and forward loop are implemented to consider the operational lifetime of newly installed units and their maximum possible increase and decrease:

$$\underline{NU}_{yr,t}^{new} = \begin{cases} \max\{\underline{NU}_{yr,t}^{new}, \underline{NU}_{yr+1,t}^{new} - \overline{X}_{yr-CT_t-LT_t,t}\}, & \text{if } \ yr - CT_t - LT_t > 0 \\ \max(\underline{NU}_{yr,t}^{new}, \underline{NU}_{yr+1,t}^{new} - \overline{X}_{yr-CT_t,t}), & \text{elseif } \ yr - CT_t > 0 \\ 0, & \text{otherwise} \end{cases}, \forall yr, t \tag{4.42}$$

$$\underline{NU}_{yr,t}^{new} = \begin{cases} \max\{\underline{NU}_{yr,t}^{new}, \underline{NU}_{yr-1,t}^{new} - \overline{X}_{yr-CT_t-LT_t,t}\}, & \text{if } \ yr - CT_t - LT_t > 0 \\ \max(\underline{NU}_{yr,t}^{new}, \underline{NU}_{yr-1,t}^{new} - \overline{X}_{yr-CT_t,t}), & \text{elseif } \ yr - CT_t > 0 \\ 0, & \text{otherwise} \end{cases}, \forall yr, t \tag{4.43}$$

An illustration for pre-processing a region of interest is presented in Figure 4.1 using as input data the data of the test case (Section 4.4.1). Figure 4.2 presents an illustration of the trajectories of the installed capacity attained by a set of randomly generated solutions. Eq. 4.23 has been used to generated 1000 solutions which are then rounded using the nearest-integer function. In addition, these are provided as an input to the RRH to remap them within the feasible region.

Figure 4.1: The figure presents an illustration for pre-processing a region of interest.

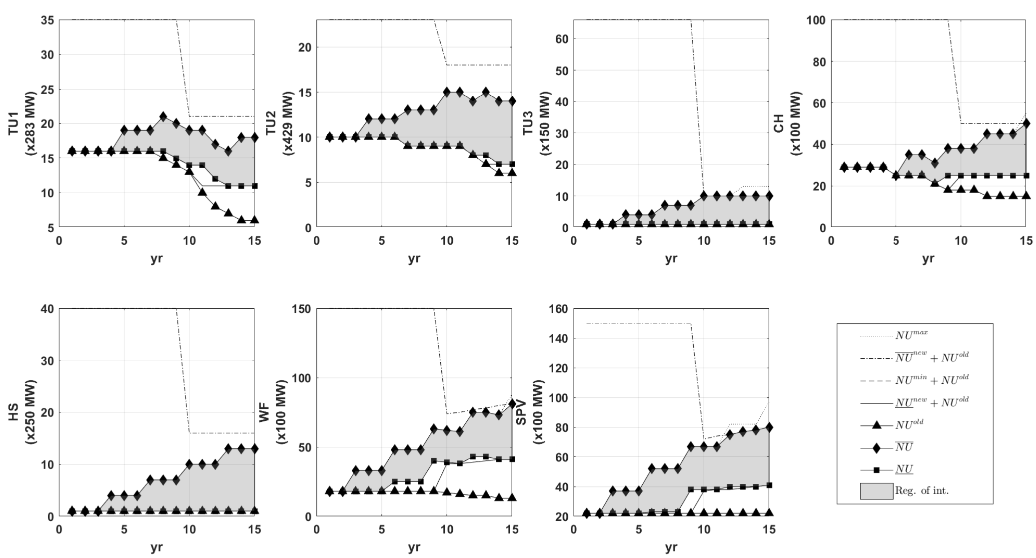
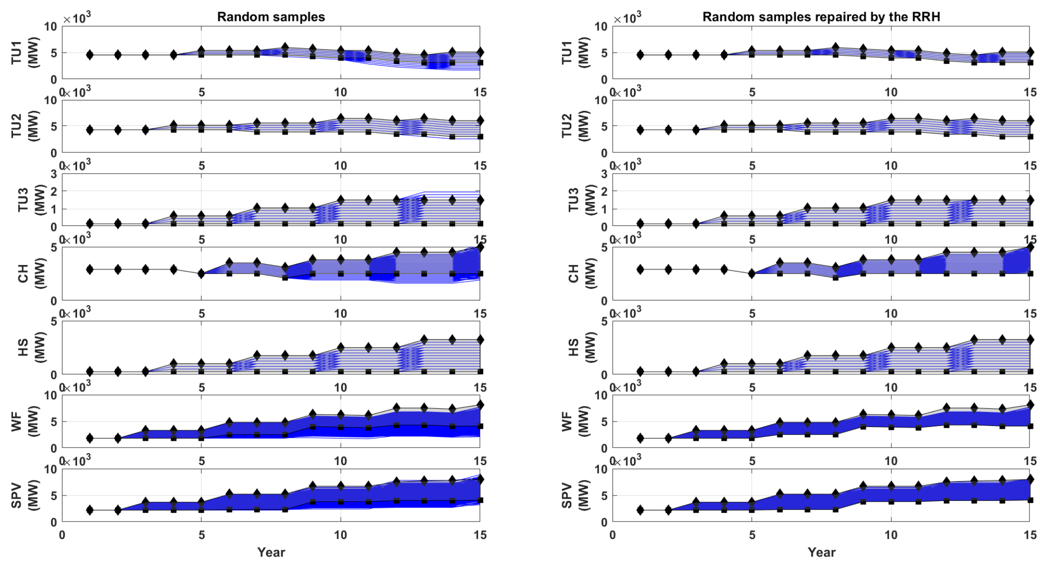


Figure 4.2: The figure presents the trajectories of the installed capacity for a set of randomly generated solutions (blue lines) in comparison to the pre-preprocessed upper (black diamonds) and lower limits (black squares) of the region of interest (grey areas). Also, the trajectories of the installed capacity for the repaired solutions attained by the RRH are provided (right column).



Initialization: Initialization of archives

A database is created that includes two types of archives (A_{yr}^{arhv} and S^{arhv}) which are initialized to an empty state. Archives A_{yr}^{arhv} are used to store all unique data points costly evaluated which are the pairs of number of installed units ($\mathbf{nu}_{yr}$) and their costly evaluated function output (c_{yr}^{var}) of a target year yr. Archive S^{arhv} is the solution archive in which unique pairs of investment decisions ($\mathbf{x}$) and their corresponding objective function value ($TC(\mathbf{x})$) are stored.

Initialization: Initial data points

The initial data points for each archive A_{yr}^{arhv} are generated in this step. The data points are generated within the region of interest ($\overline{\mathbf{NU}}$ and $\underline{\mathbf{NU}}$) using Latin hypercube sampling (Reference [195]). Then, the nearest integer function is employed to round the data points to integers (Reference [180]). These must ensure the condition regarding the minimum number of unique data points for the RBF model employed. The data points are generated anew for any archive (A_{yr}^{arhv}) for which the aforementioned condition is not met. The number of initial data points must also be met when a user-supplied minimum number of data points is included. All unique data points generated are evaluated using the SM and the values of each pair ($\mathbf{nu}_{yr}$ and c_{yr}^{var}) are stored in the corresponding archive A_{yr}^{arhv}.

The archives (A_{yr}^{arhv}) are used to build the AMs which are updated when additional data points enter the archive. The variables are assumed to be continuous for building the latter even though the prediction is computed only for integer points, as in Reference [180].

Initialization: Initial solutions

Initial solutions may improve the accuracy of the prediction near a user-supplied estimate since data points would be evaluated for an initial solution and enter the archives. Therefore, if the estimate is a decent solution that is near to the actual minimum then it could possibly reduce the costly function evaluations required to identify the specific region. Such an estimate could be the output of a different model. In case such is not available, an initial point based on the formulation of the objective function, and its cost terms, could be generated. It is assumed that an estimate is not available and a number of initial points are generated by optimization runs carried out by the DE algorithm (Section 4.3.1) and using the AMs for computing the fitness of the individuals. The selected optimization problems are based on the defined aggregated cost function and the formulation presented in Section 4.2:

1. The solution exhibiting the minimum value of the investment and fixed cost terms ($f^{chp}(\mathbf{x})$). A decent estimate could be near the aggregated minimum value these cost terms assuming that investment and fixed cost are an important cost factor to determine the optimal investment decisions. The following optimization problem is considered:

$$\begin{aligned} &minimize \;\; f^{chp}(\mathbf{x}^{inv}) = \sum_{yr}[DD_{yr} \cdot (c_{yr}^{inv} + c_{yr}^{fixed})] + \sum_{yr^{fxd}}[DD_{yr^{fxd}} c_{yr^{fxd}}^{var}] \\ &s.t.(4.11 - 4.20) \end{aligned} \tag{4.44}$$

   where yr^{fxd} denotes the indices of years for which the installed capacity is predetermined. The second term of Eq. 4.44 has been included for the sake of completeness since its value is predetermined by the input data and is a constant value.

2. The solution exhibiting the maximum value of the investment and fixed cost terms. If these cost terms have a negligible impact on the aggregated cost function, then a decent estimate could be near their aggregated maximum value assuming that the introduction of new and efficient investments triggers a reduction on the total operating cost. The following optimization problem is considered:

$$\begin{aligned} &maximize \;\; f^{chp}(\mathbf{x}^{inv}) \\ &s.t.(4.11 - 4.20) \end{aligned} \tag{4.45}$$

3. The solution exhibiting the minimum value of variable cost based on the initial prediction of the AMs. This solution could identify a region where the sum of the operating cost is low assuming that a decent approximation is available. The following optimization problem is carried out once for each different AM employed using the available data points:

$$\begin{aligned} &minimize \;\; f^{xp\_tot}(\mathbf{x}^{inv}) = \sum_{yr^{\sim}}[\widehat{f}_{yr^{\sim}}^{xp}(\mathbf{x}^{inv})] = \sum_{yr^{\sim}}[DD_{yr^{\sim}} \cdot \widehat{c}_{yr^{\sim}}^{var}] \\ &s.t.(4.11 - 4.20) \end{aligned} \tag{4.46}$$

   where $yr^{\sim}$ denotes the indices of years for which an AM is required. $\widehat{c}_{yr^{\sim}}^{var}$ is the corresponding prediction of cost term $c_{yr^{\sim}}^{var}$ attained by the AM.

These optimization problems are solved successively. For each initial solution derived, the operating costs ($\mathbf{c}^{var}$) are computed by the costly model for any vector $\mathbf{nu}_{yr}$ that has not been stored in the corresponding archive (A_{yr}^{arhv}). The value of the objective function ($TC(\mathbf{x}^{inv})$) is then computed and stored in archive S^{arhv}. A solution stored in archive S^{arhv} is compared, based on the true objective function, with the best-found solution ($\mathbf{x}^{best}$). The best-found solution is updated in case a lower function value has been achieved. These are repeated each time an initial solution is generated.

Evolve a DE population

The optimization phase follows the Initialization phase. A DE population is evolved for a predefined number of generations (nG) or function evaluation FES^{DE}. The predictions of the AMs are used to compute the fitness of the individuals. The optimization problem considered is the following:

$$\begin{aligned} &minimize \;\; \widehat{TC}(\mathbf{x}^{inv}) = f^{chp}(\mathbf{x}^{inv}) + \sum_{yr^{\sim}} [\widehat{f}^{xp}_{yr^{\sim}}(\mathbf{x}^{inv})] \\ &s.t.(4.11 - 4.20) \end{aligned} \tag{4.47}$$

The DE algorithm described in Section 4.3.1 is employed including some modifications:

1. Population initialization: The DE population is initialized by selecting up to NP top performing solutions from the archive S^{arhv}. This is implemented to supply the DE algorithm with promising solutions identified in prior steps. In case the archive includes less than NP solutions then the remaining individuals are randomly initialized within the search space (Eq. 4.23).

2. Restart mechanism: If the restart mechanism is applied, then the best individual of the population is compared to the best-found solution. In case it exhibits a lower prediction than the best-found solution it is stored in the archive (S^{arhv}). The SM is employed, the values are stored in the archives and the AMs are updated if required. The re-initialized population generated by the restart mechanism (as described in Section 4.3.2) is then used for the remaining generations.

When the termination criteria have been met, the best individual of the population is identified. It is set as the next solution to be costly evaluated. Therefore, the installed capacity is computed, new data points are costly evaluated and the total cost is computed if new data points are introduced. The archives (A_{yr}^{arhv} and S^{arhv}), the current best solution found, and the AMs are also updated if required. Th DE population is evolved once for each different AM employed.

Optimization Cycle

An Optimization Cycle consists of the sequential application of two stages: (i) refining the AMs and (ii) applying the DE algorithm. Within each optimization cycle the AMs are refined on-line, by adding new data points, and the DE algorithm is utilized to identify an improved solution based on the updated AMs:

1. The AMs are refined: Two refining strategies are employed to select solutions from a pool of candidate solutions generated by random perturbations. The first includes the Crowding Distance Metric and the second employs the prediction of local AMs as a sampling criteria.

2. The DE population is evolved: The DE algorithm is employed to evolve a population for nG generations and identify an improved solution. In this case, the AMs are built using all available data points.

These are discussed in the following Sections.

Optimization Cycle - Refining the approximation models: pool of candidate solutions

A number of surrogate-assisted derivative-free algorithms generate a large pool of candidate solutions by applying randomized perturbations on the best-found solution, or by randomly generating candidate solutions within the search space (References [177, 178, 179, 180, 181, 186]). The pool of candidate solutions is used with selection criteria to identify the next data point to be evaluated using the costly model. A similar stochastic approach is implemented to generate a pool of candidate solutions. Three different approaches are employed. The first two use the best-found solution to generate candidate solutions within its vicinity. The third generates a number of candidate solutions within the search space to ensure sufficient diversity. More specifically, these are generated as follows:

1. Candidate solutions generated by adding or removing a single investment from decision variables of the best-found solution ($\mathbf{x}^{best}$): The alteration of $\mathbf{x}^{best}$ (to add or remove an investment) is determined with equal probability. Each decision variable exhibits a probability to be altered (user-defined parameter $pf_{pert} = 0.1$).

2. Candidate solutions generated by applying the PO on the best-found solution ($\mathbf{x}^{best}$): The target vector is set as $\mathbf{x}^{best}$ and the candidate solutions of the previous step are set as base vectors to alter a number of technology blocks determined by the technology-group operator (blk).

3. Candidate solutions generated by re-initialization: A number of solutions are generated within the search space by applying Eq. 4.23.

The size of the pool of candidate solutions is determined by a user-defined parameter rNP. Constraint violations of these solutions are then repaired by applying: (i) the nearest integer function to satisfy 4.10, (ii) Eq. 4.25 to respect Eq. 4.11, and (iii) the RRH to repair solutions violating the remaining planning constraints (Eq. 4.16-4.20).

Furthermore, the RRH is used as a local search technique rather than to remap points within the feasible region. It is applied to remap the candidate solutions within the feasible vicinity of the best-found solution. More specifically, the solutions are remapped to follow similar trajectories of installed capacity to the corresponding ones of $\mathbf{x}^{best}$. This is implemented by introducing bias to the repair process by altering the upper and lower limits of the installed capacity as follows:

$$\underline{IC}^{biased}_{yr,t} = P^{net}_t \max(nu^{best}_{yr,t} - \lceil ri \cdot (\overline{NU}_{yr,t} - \underline{NU}_{yr,t}) \rceil, \underline{NU}_{yr,t}), \forall yr, t \tag{4.48}$$

$$\overline{IC}^{biased}_{yr,t} = P^{net}_t \min(nu^{best}_{yr,t} + \lceil ri \cdot (\overline{NU}_{yr,t} - \underline{NU}_{yr,t}) \rceil, \overline{NU}_{yr,t}), \forall yr, t \tag{4.49}$$

where $\overline{\mathbf{IC}}^{biased}$ and $\underline{\mathbf{IC}}^{biased}$ are biased upper and lower limits, respectively. $\mathbf{nu}^{best}$ is the number of available units generated by applying Eq. 4.6 for $\mathbf{x}^{best}$. These limits replace the initial limits of the installed capacity and are included in the RRH. It is applied to each candidate solution that derives an installed capacity violating these limits. ri is a user-defined parameter in the range $(0, 1]$ included to control the size of the sub-region of interest. A pool of values can be adopted for each refining strategy and generate different sizes for the sub region, e.g. $ri = \{1, 0.75, 0.5, 0.25\}$. The size of the pool and its values controls global and local search. In particular, the size of a pool defines the number of solutions selected by each refining strategy and the values control the maximum difference of the number of available units in each year of a candidate solution to the corresponding on of the best-found solution.

A number of candidate solutions are removed from the pool to form the final output of this step. The removed solutions are:

1. Candidate solutions that have been stored in archive S^{arhv}: a candidate solution for which the value of the Euclidean Distance towards any solution in archive S^{arhv} is lower than one is removed.

2. Candidate solutions that are not unique: a candidate solution for which the value of the Euclidean Distance towards any other candidate is lower than one. A single replicate is preserved.

In case all candidate solutions are removed by the previous step, then candidate solutions are generated anew and the RRH is implemented without including any bias. If the latter does not identify a new candidate solution, the process is repeated by generating candidate solutions within the search space by applying Eq. 4.23. The output of the aforementioned steps should be a pool of candidate solutions which are feasible w.r.t. the planning constraints and have not been stored in the archive.

Moreover, a sampling criterion could select a candidate solution that has not been stored in archive S^{arhv}. However, the resulting number of available units (**nu**) could be a combination of data points available in archives A^{arhv}_{yr}. In such a case, the candidate solution is stored in S^{arhv} and the next candidate solution, prioritized by the sampling criterion, is examined. The aforementioned steps are repeated to select a candidate solution for which at least one new data point is required to be computed by the SM (the data point is not available in archives A^{arhv}_{yr}).

Optimization Cycle - Refining the approximation models: Sampling criteria

This Section presents the sampling criteria employed within the refining strategies. The selection is made from the pool of candidate solution generated by the steps presented in Section 4.3.4 based on: (i) the Crowding Distance (CD) metric, or (ii) the minimum prediction attained by locally trained AMs.

Selection based on the Crowding Distance metric:
References [12, 119] suggested the Crowding Distance metric for MOO. Similar metrics has been also included in other surrogate-based approaches to select data points (e.g. References [196, 197]). A metric based on the CD is adopted to identify solutions that are isolated in an attempt to distribute data points spatially. The metric is based on measuring distances between archived and the generated candidate solutions. Its value is computed by comparing the newly generated candidate solutions (Section 4.3.4) with the ones in archive S^{arhv}. The candidate solution from the pool exhibiting the highest $CD(\mathbf{z})$ is selected to be costly evaluated since it is the most isolated based on the metric value. The latter is computed for each candidate solution as follows:

$$CD(\mathbf{z}) = \sum_{s}[(\|\mathbf{z} - \mathbf{z}^{(s)}\|)^2] \tag{4.50}$$

where s is the index of the archived solutions (S^{arhv}) and $\mathbf{z}$ is a vector including the normalized values of $\mathbf{x}$ for which the values are not predetermined i.e. the upper bound is greater than zero. Values are normalized based on the corresponding upper bound of each variable ($\overline{\mathbf{X}}^{inv}$).

Selection based on locally trained AMs:
A number of surrogate-assisted algorithms have employed AMs that are built by using a subset of the available data points, since the accuracy of the approximation depends on the sufficiency of the data points used to build the model. Two examples regarding surrogate-assisted DE algorithms are References [168, 170]. Reference [170] employed an AM that is built in each generation using as data points the available individuals of the evolving DE population. Reference [168] included a local phase within the DE-based MAEA. The local phase employs the interior-point method to solve an optimization problem for each individual of the DE population. An RBF model is used to replace the true function in the optimization problem of the local phase. The model is trained using a number of archived data points that are nearest to the considered individual.

The locally trained models are employed in an attempt to exclude distant data points that do not necessarily improve the approximation accuracy. The training set is determined based on the trajectory of the installed capacity generated by the best-found solution. In particular, the sub-region of interest, that had been defined to generated the pool of candidate points, is utilized to determine the number of data points that will be used to train the local AMs. This is implemented by selecting all data points that are within the aforementioned region. Consequently, the size of the region of interest (determined by parameter ri) can influence the number of available data points that will be used. However, if the minimum required number of data points for building the AM are not attained, additional data points outside the region of interest are included. The additional data points are selected by prioritizing the nearest ones (Euclidean Distance). The local AMs are built for the selected data points and are utilized to select a candidate solution, from the pool, that exhibits the minimum value of the prediction ($\widehat{TC}(\mathbf{x}^{inv})$).

Optimization Cycle - Evolve the DE population

The DE algorithm is employed to identify an improved solution when the AMs have been refined assuming that an improvement has been achieved regarding their accuracy. The procedure follows the steps described in Section 4.3.4.

4.3.5 Pool of Approximating models

The employed RBF model is described in Reference [198]. Specifically, Reference [180] is followed for this selection and for the implemented procedure since the aforementioned RBF model had been included within the surrogate-assisted model proposed for discrete global optimization problems. In addition, this selection is made since: (i) the underlying functions of the optimization problem in this case are also deterministic, (ii) the RBF model may provide an accurate prediction for the available data points, and (iii) RBFs have been reported to perform well on high dimensional problems (References [166, 167]). Moreover, it the employed RBF model has also been successfully adopted in other surrogate-assisted derivative-free algorithms (e.g. References [178, 179, 181]). The prediction function is computed as follows:

$$\hat{f}(\mathbf{x}) = \sum_{s=1}^{np} \lambda_s \phi(\|\mathbf{x} - \mathbf{x}^{(s)}\|) + \rho(\mathbf{x}) \tag{4.51}$$

where $\widehat{f}(\mathbf{x})$ is the prediction for the underlying function, $\mathbf{x}_s$ are the available data points, np is the number of available data points, $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, ..., \lambda_{np})$ are coefficients, $\phi(\cdot)$ is the kernel function and $\|\cdot\|$ is the Euclidean norm. $\rho(\mathbf{x})$ is a polynomial tail of which the order depends on the RBF kernel. A linear polynomial tail is adopted $\rho(\mathbf{x}) = \mathbf{b}^T\mathbf{x} + a$ (Reference [175]), where $\mathbf{b} = (b_1, \ldots, b_d)^T \in \mathbb{R}^d$, and $a \in \mathbb{R}$. A linear system of equations is solved in order to determine parameters $\boldsymbol{\lambda}$, $\mathbf{b}$ and a (Reference [175]):

$$\begin{bmatrix} \boldsymbol{\Phi} & \mathbf{P} \\ \mathbf{P}^T & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{\lambda} \\ \mathbf{c} \end{bmatrix} = \begin{bmatrix} \mathbf{F} \\ \mathbf{0} \end{bmatrix}, \tag{4.52}$$

where

$$\boldsymbol{\Phi} = \begin{bmatrix} \phi(\|\mathbf{x}^{(1)} - \mathbf{x}^{(1)}\|) & \phi(\|\mathbf{x}^{(1)} - \mathbf{x}^{(2)}\|) & \cdots & \phi(\|\mathbf{x}^{(1)} - \mathbf{x}^{(np)}\|) \\ \phi(\|\mathbf{x}^{(2)} - \mathbf{x}^{(1)}\|) & \phi(\|\mathbf{x}^{(2)} - \mathbf{x}^{(2)}\|) & \cdots & \phi(\|\mathbf{x}^{(2)} - \mathbf{x}^{(np)}\|) \\ \vdots & \vdots & \ddots & \vdots \\ \phi(\|\mathbf{x}^{(np)} - \mathbf{x}^{(1)}\|) & \phi(\|\mathbf{x}^{(np)} - \mathbf{x}^{(2)}\|) & \cdots & \phi(\|\mathbf{x}^{(np)} - \mathbf{x}^{(np)}\|) \end{bmatrix}, \mathbf{P} = \begin{bmatrix} \mathbf{x}^{(1)} & 1 \\ \mathbf{x}^{(2)} & 1 \\ \vdots & \vdots \\ \mathbf{x}^{(np)} & 1 \end{bmatrix}, \mathbf{c} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_d \\ a \end{bmatrix},$$

while $\mathbf{0}$ is are a null matrix/vector.

The pool of AMs includes three different kernel functions: the linear, the cubic RBF and the Thin Plate Spline (TPS) kernel functions. All three kernel functions are used since it is not easy to estimate a priori which function interpolates the underlying function more accurately.

4.4 Set-up of the Numerical Experiments

The following Sections present the considered test case, implementation notes, the numerical experiments and the employed performance metrics.

4.4.1 Test case

The test case is loosely based on the Greek Power system (Reference [199]). More specifically, the data input are modified data inspired by the aforementioned system. These are presented in Appendix C. The power system is assumed isolated and interconnections are neglected. In addition, network limitations are not considered and the examined system is assumed as a single-zone design. Therefore, it is assumed that the examined system exhibits sufficient (internal) transmission capacity.

The planning horizon is set as 15 years considering an investment stage every three years, i.e. $\mathbb{Y}^{InvSt} = \{1, 4, \ldots, 13\}$. All years within the planning horizon have been assumed as target years. However, for the first two years of the planning horizon the installed capacity is predetermined. The number of representative days has been set as one per month. For each year of the planning horizon 12 time periods have been selected (one for each month). In total 180 days are considered within the planning horizon. The weight factors $\mathbf{W}$ are set based on the number of days within each month under the assumption that the operational cost of a representative day of a month is equal to the average cost of a day within that month.

The thermal units considered are Lignite plants, Natural gas-fired plants and Natural gas-fired peak plants. For the test case these serve as base, intermediate and peak units, respectively. These shall be referred hereafter as thermal unit one (TU1), thermal unit two (TU2) and thermal unit three (TU3), respectively. The aggregated RES technology groups considered are solar PV's (SPV) and wind farms (WF). An aggregated CH unit and an aggregated HS unit are also assumed. Smaller plants, other technology groups and imports/exports are neglected. A decommissioning plan is generated based on the expected operational lifetime of each existing plant. For the sake of simplicity, old plants are assumed to have the same characteristics as new plants. The techno-economic data of the different technology groups are based on Reference [200] and have been modified for the test case.

The scenario considered assumes that the RES generation share should range within a minimum (30%) and maximum (60%) level by the 10$^{\text{th}}$ year of the planning horizon. This has been implemented by setting the lower and upper limits, for both the WF and SPV technology groups, as 15% and 30% of the total demand, respectively. The anticipated availability ($\mathbf{AV}$) of the two RES technology groups is set equal to the mean of their corresponding hourly capacity factors ($\mathbf{PP}^{res}$). Additionally, upper and lower limitations are included for the installed capacity of the remaining technology groups which are active starting from the 10$^{\text{th}}$ year of the planning horizon. These data serve as the input for an illustrative example. Therefore, the results presented are dependent on the data assumptions made and may alter, to some extent, under different input data.

4.4.2 Implementation notes

The implementation has been developed within Matlab. The available LP and MILP have been employed with default parameters. All parameters associated with the DE and MAEA algorithms are presented in Table 4.1. The Matlab implementation of Reference [180] has been also used to develop the optimization approach.

Parameters DE	Value	Parameters MAEA	Value
NP	50	rNP	3000
CR_{pool}	$[0.1, 0.2, 1]$[201]	NS^{init}	$2(d_{yr}+1)$[180]
F_{pool}	$[0.6, 0.7, \ldots, 1]$[201]	pf_{PO}	0.05
FES^{DE}	$2.5 \cdot 10^3 D$	pf_{pert}	0.1
nG	1500	ri^{CD}_{pool}	$[1, 0.75, 0.5, 0.25]$
gen^{rst}	150	ri^{LS}_{pool}	$[1, 0.75, 0.5, 0.25]$

Table 4.1: Parameter settings of the SOO MAEA.

4.4.3 Numerical experiments

In this Section, the numerical experiments conducted are presented. Specifically, the following sets have been carried out:

- The performance of the problem-customized MAEA is examined on a test case representing a modified version of a real power sector. This aims towards examining the ability of the MAEA to identify promising data points to be costly evaluated and locate a near-optimal solution when a SM is included. Two different optimization problems are examined:

  1. In the first, the SM is set as a model that does not include constraints of a UCP that are included in the CUC formulation. The aforementioned model exhibits a lower level of technical detail in comparison to the CUC model (Appendix A). The formulation is presented in Appendix B.

2. In the second, the SM is set as the CUC model presented in Appendix A.

For each optimization problem, the accuracy of the Underlying Function Approximation (UFA) achieved by the AMs is also examined. Based on the aforementioned results, the derived near-optimal solution is used to analyse the evolution of the installed capacity and the generation mix.

- The performance of the DE variant is assessed on three optimization problems. This aims towards examining the DE algorithm performance when the AMs are accurate. Moreover, the impact of the examined problem-customized operators is assessed by repeating the optimization runs for a number of variants that include combinations of the examined operators. This aims towards identifying the most competitive variants. The first optimization problem is of the form presented in Eq. 4.44. The second optimization problems is of the form presented in Eq. 4.47. A dataset attained using the MAEA, including the SM^{CUC} (Appendix A), is employed for building the AM and no additional data points are costly evaluated. This is repeated for each type of RBF kernel function. The third optimization problem exhibits the same set up (as the second) except that a dataset attained using the MAEA including the SM^{ED} (Appendix B) is used.

- The performance of the problem-customized MAEA is examined on a constraint minimization problem for which the global optimum is known. The objective function of the formulated optimization problem is set as the sum of squared deviations of the installed capacity towards a predefined targeted one. Moreover, the constraint functions of the GEP model are included.

- The near-optimal solutions derived by employing the aforementioned two SMs (SM^{ED} and SM^{CUC}) are examined. A third case is also included for which no SM has been used. In particular, the objective function considers only investment and fixed cost (the optimization problem is of the form presented in Eq. 4.44. The comparison aims towards examining the impact of the level of technical detail of the SM on the results.

These are presented in more detail in the following Sections.

Minimizing the computational cheap function using the examined DE variants

The optimization problem presented in Section 4.2 is considered. However,the operational cost term is excluded for this set of numerical experiments. Thus, the optimization problems aims towards the minimization of the total investment and fixed costs restricted by the planning constraints (as in Eq. 4.44). This optimization problem can be solved by using a MILP solver and its output can be used as a point of reference. Problem-customized DE variants are employed for solving the same optimization problem. Consequently, the results attained by the latter may be compared to the solution derived by the MILP solver to indicate the performance of the modified DE variants.

These DE variants are generated by including combinations of the examined operators (RRH, blk and PO) for three different crossover operators: the Binomial (bin), Exponential (exp) and Shuffled-Exponential (expS). Moreover, FR is used as the CHT when RRH is not included. For comparing the results of variants including RRH with the ones including FR it is important to consider that RRH heuristically repairs infeasible solutions, i.e. the computational time of RRH in comparison to the FR is higher. The number of independent runs is set to 100 for each variant.

Minimizing the prediction function attained by the approximation models for the $\mathbf{SM^{CUC}}$ using the examined DE variants

The aforementioned numerical experiment is repeated. However, the operational cost term is included in this set of numerical experiments. The optimization problem is of the form presented in Eq. 4.47. A dataset attained by an independent run of the MAEA and the $\mathrm{SM^{CUC}}$ is used for computing the parameters of the AM. Moreover, no new data point are included in the dataset during the optimization runs since the aim of this numerical experiment is to examine the problem-customized DE variants on the GEP formulation (Section 4.2) by assuming that the AMs are *perfectly* accurate. This numerical experiment is repeated for each of the three employed RBF kernel functions. Hereafter, each such set shall be referred to as SM^{CUC}*-Lin*, SM^{CUC}*-Cub* and SM^{CUC}*-TPS* based on the RBF kernel function used to construct the AM (the linear, cubic or TPS, respectively).

Minimizing the prediction function attained by the approximation models for the $\mathbf{SM^{ED}}$ using the examined DE variants

The numerical experiment set-up is similar to the one including the $\mathrm{SM^{CUC}}$. However, the dataset attained by employing the MAEA and using the $\mathrm{SM^{ED}}$ is used for computing the parameters of the AM. In this case, each set of numerical experiment shall be referred to as SM^{ED}*-Lin*, SM^{ED}*-Cub* and SM^{ED}*-TPS* based on the RBF model used to construct the AM (the linear, cubic or TPS RBF model, respectively).

Application of the MAEA to identify a targeted installed capacity

The objective function of the formulated optimization problem is set as the sum of squared deviations of the installed capacity towards a predefined targeted one. Moreover, the constraint functions of the GEP model are included. The constraint minimization problem considered is of the following form:

$$\begin{aligned} &minimize \quad F(\mathbf{x}) = \sum_{yr^{\sim}} [f^{xp}_{yr^{\sim}}(\mathbf{x})] \\ &s.t. (4.11 - 4.20) \end{aligned} \tag{4.53}$$

where $f^{xp}_{yr^{\sim}}(\mathbf{x}) = \sum_t [(nu^{targ}_{yr^{\sim},t} - nu_{yr^{\sim},t}(\mathbf{x}))^2], \forall yr^{\sim}$. The AMs are employed to estimated the functions $f^{xp}_{yr^{\sim}}(\mathbf{x})$. This numerical experiment is repeated twice. In particular, in the first set the targeted installed capacity is set as the solution derived by minimizing investment and fix cost. The optimization approach (Section 4.3) is employed for the minimization problem which is repeated 100 independent times. In the second, 100 different installed capacities are set as targeted installed capacity to examine the performance of the approach when such are located in different regions of the search space. A single independent run is performed for each targeted installed capacity. These solutions are randomly sampled and repaired by the nearest integer function and the RRH. The termination criteria for the MAEA had been set as a maximum limit of exact evaluations ($TotSim = 1000$) or a successful identification of the global optimum ($F(\mathbf{x}^{best}) = 0$). In these cases no computationally cheap functions ($f^{chp}(\mathbf{x})$) had been considered and the initial solutions are randomly generated and repaired by the nearest integer function and the RRH.

Application of the MAEA on the test case using the SM^{ED} and SM^{CUC}

The MAEA (Section 4.3) is employed to optimize the GEP model presented in Section 4.2. In the first case, the SM presented in Appendix B is employed (SM^{ED} case). In the second, the SM presented in Appendix A is used (SM^{CUC} case). The data input presented in Appendix 4.4.1 are used. A simulation is considered as computing a data point for a year of the planning horizon. The termination criteria for the MAEA has been set as a maximum number of simulations ($TotSim = 500$). The number of independent runs is set to 30 for each of the two cases.

4.4.4 Performance metrics

The metrics employed for measuring the accuracy of the UFA are presented in this Section. The accuracy is measured according to the Root Mean Square Error (RMSE), the Mean Absolute Error (MAE) and Coefficient of determination metrics (R^2) by employing a leave-one-out cross validation approach. The metrics are computed independently for each AM.

- The value of the RMSE metric is computed as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{np}[(\hat{f}(\mathbf{x_i}) - f(\mathbf{x_i}))^2]}{np}} \tag{4.54}$$

  where $f(\mathbf{x_i})$ is the true value of an objective function for $\mathbf{x_i}$, $\hat{f}(\mathbf{x_i})$ is the corresponding prediction of the AM built by excluding data point $\mathbf{x_i}$ from the dataset and np is the number of available data points within the dataset. Lower values indicate a more accurate AM and the ideal value is 0.

- The value of the MAE metric is computed as follows:

$$MAE = \frac{\sum_{i=1}^{np}[|(\hat{f}(\mathbf{x_i}) - f(\mathbf{x_i})|]}{np} \tag{4.55}$$

  Once again, lower values indicate a more accurate AM and the ideal value is 0.

- The value of the R^2 metric is computed as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{np}[(\hat{f}(\mathbf{x_i}) - f(\mathbf{x_i}))^2]}{\sum_{i=1}^{np}[(\hat{f}(\mathbf{x_i}) - f^{mean})^2]} \tag{4.56}$$

  where f^{mean} is the mean value of the available (archived) values of the objective function. For this metric, higher values are preferred which indicate a more accurate AM. The best possible value is 1.

4.5 Results

This Section presents the results of the numerical experiments, the comparison of the results based on the employed SM, the quality of the attained AM and a visual analysis of the sensitivity of the variable cost towards the attained installed capacity.

4.5.1 Results of the numerical experiments

Results of the numerical experiments using the examined DE variants

The results attained on the test cases are presented in this Section. The mean and standard deviation of the relative difference of the best-found solution in each independent run (f^{bf}) towards the value of the best-found solution over all independent runs of all variants (f^*) is presented ($\frac{f^{bf}-f^*}{f^*}$). For the case where the minimization problem does not include the prediction of the AMs, the solution attained by the MILP solver is used to determine f^*. Moreover, the Success Rate (SR) and Average number of Restarts (AR) is also presented for each variant. The SR suggests the number of independent runs that a variant identified f^*. The maximum possible number is 100 which is the number of independent runs. AR indicates the average number of restarts that have occurred in each independent run. The results are presented in Tables 4.2-4.8. Moreover, results attained by the Wilcoxon rank sum test are provided in Appendix D. These present if the attained sets of the variants including all examined operators differ in a statistically significant manner from the variants including the same crossover operator but different combinations of the examined operators.

- *Minimizing the computational cheap function using the examined DE variants*: The results indicate that the variants including the RRH are the most competitive. The SR is 100/100 for the variants including the Binomial Crossover Operator and the RRH. Variants including the Exponential or Shuffled Exponential Crossover Operator have also 100/100 SR except for the variants excluding both the blk and PO, which present a 98/100 and 99/100 value of the SR, respectively. On the contrary, the variants excluding RRH have identified the best-found solution at most once. This suggests an improved performance attained by the inclusion of RRH. The impact of the blk and PO, when RRH is included, is not clear due to the high values of SR. However, a comparison of the variants excluding the RRH suggests an improvement attained by the blk, since the mean value of variants including the blk is consistently lower in comparison to the mean of the variants excluding it. Moreover, it is observed that no restarts are triggered when RRH is excluded.

- *Minimizing the prediction function attained by the AMs for the SM^{ED} using the examined DE variants*: The results of the SM^{ED}*-Lin*, SM^{ED}*-Cub* and SM^{ED}*-TPS* cases suggest an improved performance attained by the inclusion of RRH and/or blk based on the mean values and the SR metric. In general, the most competitive variants include RRH and blk or RRH, blk and PO. The results indicate that the selection of a Crossover operator is less important when the examined operators are included. However, a comparison of the corresponding values of the AR reveals that the inclusion of PO and blk affects the number of restarts triggered for the variants including the Exponential or Shuffled Exponential Crossover Operator and the RRH.

- *Minimizing the prediction function attained by the AMs for the SM^{CUC} using the examined DE variants*: Similar observations can be made by examining the performance of the variants on the optimization problem including the AMs for the $\mathrm{SM}^{\mathrm{CUC}}$. For example, variants including RRH and blk or RRH, blk and PO had provided the highest values of SR metric.

Based on this results, the most competitive performance is attained by the DE variants including the RRH and blk and the ones including the RRH, blk and PO. Moreover, among the three examined operators the highest gain had been attained by RRH. This could suggest that exploiting domain-specific knowledge by repair heuristics within the employed DE algorithm may be prioritized. Convergence plots for the DE/rand/1/bin variants including the examined operators is provided in Figure 4.3 for the SM^{CUC}-*Cub* case. For the remaining numerical experiments the DE/rand/1/bin/RRH/blk/PO has been used.

Table 4.2: Results attained by the examined DE variants for 100 independent runs of the optimization problem including the computational cheap function.

	DE/rand/1/bin				DE/rand/1/exp				DE/rand/1/expS			
Operators	Mean	SD	SR	AR	Mean	SD	SR	AR	Mean	SD	SR	AR
/FR	5.92E-03	3.09E-03	0	0.00	3.03E-02	1.34E-02	0	0.00	2.73E-02	1.00E-02	0	0.00
/FR/blk	2.55E-03	1.66E-03	0	0.00	3.21E-03	2.05E-03	0	0.00	4.06E-03	2.12E-03	0	0.00
/FR/PO	6.99E-03	3.13E-03	1	0.00	2.87E-02	1.20E-02	0	0.00	2.57E-02	1.07E-02	0	0.00
/FR/blk/PO	2.65E-03	1.74E-03	1	0.00	3.85E-03	1.91E-03	1	0.00	4.00E-03	2.25E-03	0	0.00
/RRH	0.00E+00	0.00E+00	100	1.47	8.18E-06	6.56E-05	98	1.00	9.39E-07	9.39E-06	99	0.98
/RRH/blk	0.00E+00	0.00E+00	100	1.00	0.00E+00	0.00E+00	100	1.00	0.00E+00	0.00E+00	100	1.00
/RRH/PO	0.00E+00	0.00E+00	100	1.16	0.00E+00	0.00E+00	100	0.98	0.00E+00	0.00E+00	100	0.95
/RRH/blk/PO	0.00E+00	0.00E+00	100	1.00	0.00E+00	0.00E+00	100	1.00	0.00E+00	0.00E+00	100	1.00

Table 4.3: Results attained by the examined DE variants for 100 independent runs of the optimization problem including the prediction of the SM^{ED}-*Lin*.

	DE/rand/1/bin				DE/rand/1/exp				DE/rand/1/expS			
Operators	Mean	SD	SR	AR	Mean	SD	SR	AR	Mean	SD	SR	AR
/FR	1.58E-03	7.36E-04	2	0.00	7.63E-03	3.24E-03	0	0.00	6.53E-03	2.61E-03	0	0.00
/FR/blk	1.06E-03	5.42E-04	4	0.00	1.43E-03	6.62E-04	0	0.00	1.55E-03	6.17E-04	0	0.00
/FR/PO	1.80E-03	7.25E-04	0	0.00	7.54E-03	2.76E-03	0	0.00	6.50E-03	2.13E-03	0	0.00
/FR/blk/PO	1.25E-03	5.10E-04	1	0.00	1.63E-03	6.33E-04	0	0.00	1.68E-03	6.67E-04	0	0.00
/RRH	0.00E+00	0.00E+00	100	1.14	9.28E-06	4.09E-05	95	0.76	1.52E-05	7.00E-05	95	0.67
/RRH/blk	0.00E+00	0.00E+00	100	1.00	0.00E+00	0.00E+00	100	1.00	0.00E+00	0.00E+00	100	1.00
/RRH/PO	0.00E+00	0.00E+00	100	1.01	1.75E-06	1.75E-05	99	0.58	2.84E-05	7.82E-05	87	0.52
/RRH/blk/PO	0.00E+00	0.00E+00	100	1.00	0.00E+00	0.00E+00	100	0.99	0.00E+00	0.00E+00	100	0.99

Table 4.4: Results attained by the examined DE variants for 100 independent runs of the optimization problem including the prediction of the SM^{ED}-*Cub*.

	DE/rand/1/bin				DE/rand/1/exp				DE/rand/1/expS			
Operators	Mean	SD	SR	AR	Mean	SD	SR	AR	Mean	SD	SR	AR
/FR	1.49E-03	6.71E-04	0	0.00	7.58E-03	3.18E-03	0	0.00	6.56E-03	2.55E-03	0	0.00
/FR/blk	9.67E-04	5.64E-04	2	0.00	1.33E-03	5.49E-04	0	0.00	1.42E-03	5.97E-04	1	0.00
/FR/PO	1.68E-03	6.19E-04	1	0.00	7.26E-03	2.85E-03	0	0.00	6.44E-03	1.98E-03	0	0.00
/FR/blk/PO	1.00E-03	5.00E-04	1	0.00	1.51E-03	6.02E-04	0	0.00	1.49E-03	6.64E-04	0	0.00
/RRH	0.00E+00	0.00E+00	100	1.11	2.47E-05	6.62E-05	87	0.70	3.21E-05	1.21E-04	91	0.54
/RRH/blk	0.00E+00	0.00E+00	100	1.00	0.00E+00	0.00E+00	100	1.00	0.00E+00	0.00E+00	100	1.00
/RRH/PO	0.00E+00	0.00E+00	100	1.04	1.40E-05	5.36E-05	93	0.66	2.20E-05	8.49E-05	92	0.52
/RRH/blk/PO	0.00E+00	0.00E+00	100	1.00	0.00E+00	0.00E+00	100	0.98	0.00E+00	0.00E+00	100	1.00

Table 4.5: Results attained by the examined DE variants for 100 independent runs of the optimization problem including the prediction of the SM^{ED}-TPS.

	DE/rand/1/bin				DE/rand/1/exp				DE/rand/1/expS			
Operators	Mean	SD	SR	AR	Mean	SD	SR	AR	Mean	SD	SR	AR
/FR	1.58E-03	8.28E-04	1	0.00	7.63E-03	3.25E-03	0	0.00	6.34E-03	2.13E-03	0	0.00
/FR/blk	1.03E-03	5.30E-04	1	0.00	1.41E-03	6.24E-04	0	0.00	1.59E-03	6.83E-04	1	0.00
/FR/PO	1.84E-03	7.11E-04	0	0.00	6.94E-03	2.69E-03	0	0.00	6.69E-03	2.19E-03	0	0.00
/FR/blk/PO	1.09E-03	5.50E-04	0	0.00	1.50E-03	6.53E-04	1	0.00	1.41E-03	5.98E-04	0	0.00
/RRH	0.00E+00	0.00E+00	100	1.10	1.86E-05	7.52E-05	93	0.69	3.22E-05	1.41E-04	92	0.59
/RRH/blk	0.00E+00	0.00E+00	100	1.00	0.00E+00	0.00E+00	100	1.00	0.00E+00	0.00E+00	100	1.00
/RRH/PO	0.00E+00	0.00E+00	100	1.01	1.12E-05	5.94E-05	96	0.62	5.14E-05	1.85E-04	89	0.56
/RRH/blk/PO	0.00E+00	0.00E+00	100	1.00	0.00E+00	0.00E+00	100	1.00	0.00E+00	0.00E+00	100	1.00

Table 4.6: Results attained by the examined DE variants for 100 independent runs of the optimization problem including the prediction of the SM^{CUC}-Lin.

	DE/rand/1/bin				DE/rand/1/exp				DE/rand/1/expS			
Operators	Mean	SD	SR	AR	Mean	SD	SR	AR	Mean	SD	SR	AR
/FR	1.59E-03	5.32E-04	0	0.00	5.32E-03	2.01E-03	0	0.00	4.64E-03	1.80E-03	0	0.00
/FR/blk	1.13E-03	4.30E-04	0	0.00	1.38E-03	4.97E-04	0	0.00	1.35E-03	4.29E-04	0	0.00
/FR/PO	1.80E-03	5.60E-04	0	0.00	5.13E-03	1.80E-03	0	0.00	4.63E-03	1.63E-03	0	0.00
/FR/blk/PO	1.14E-03	4.19E-04	0	0.00	1.46E-03	4.75E-04	0	0.00	1.43E-03	5.31E-04	0	0.00
/RRH	1.87E-06	6.85E-06	93	1.00	1.95E-05	5.04E-05	70	0.35	1.22E-05	2.61E-05	65	0.34
/RRH/blk	8.01E-07	4.58E-06	97	0.99	8.01E-07	4.58E-06	97	0.93	2.67E-07	2.67E-06	99	0.90
/RRH/PO	1.87E-06	6.85E-06	93	1.00	2.59E-05	6.91E-05	70	0.27	2.33E-05	5.54E-05	66	0.25
/RRH/blk/PO	0.00E+00	0.00E+00	100	0.91	8.01E-07	4.58E-06	97	0.73	2.67E-07	2.67E-06	99	0.69

Table 4.7: Results attained by the examined DE variants for 100 independent runs of the optimization problem including the prediction of the SM^{CUC}-Cub.

	DE/rand/1/bin				DE/rand/1/exp				DE/rand/1/expS			
Operators	Mean	SD	SR	AR	Mean	SD	SR	AR	Mean	SD	SR	AR
/FR	1.55E-03	5.21E-04	0	0.00	5.60E-03	2.07E-03	0	0.00	4.56E-03	1.62E-03	0	0.00
/FR/blk	1.04E-03	3.93E-04	0	0.00	1.24E-03	3.85E-04	0	0.00	1.32E-03	4.79E-04	0	0.00
/FR/PO	1.66E-03	6.47E-04	0	0.00	5.12E-03	1.89E-03	0	0.00	4.54E-03	1.59E-03	0	0.00
/FR/blk/PO	1.17E-03	4.65E-04	0	0.00	1.32E-03	4.90E-04	0	0.00	1.37E-03	5.14E-04	0	0.00
/RRH	3.74E-06	9.32E-06	86	1.00	1.75E-05	4.14E-05	61	0.40	2.65E-05	6.09E-05	62	0.27
/RRH/blk	8.01E-07	4.58E-06	97	0.97	1.07E-06	5.26E-06	96	0.87	8.01E-07	4.58E-06	97	0.82
/RRH/PO	3.47E-06	9.03E-06	87	1.00	1.76E-05	3.73E-05	64	0.30	2.56E-05	5.43E-05	60	0.32
/RRH/blk/PO	1.07E-06	5.26E-06	96	0.91	8.01E-07	4.58E-06	97	0.69	1.87E-06	6.85E-06	93	0.60

Figure 4.3: Convergence plots of the DE/rand/1/bin variants including combinations of the examined operators (RRH, blk and PO) for the SM^{CUC}-Cub case. Each step represents 200 generations (abscissa). The ordinate depicts the relative errors ($\frac{f^{bf}-f^{*}}{f^{*}}$).

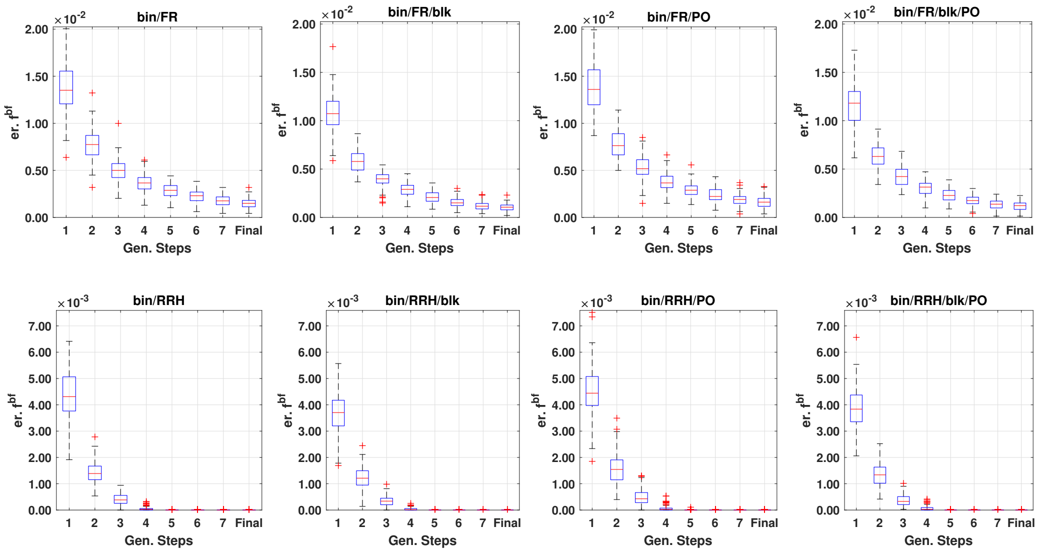


Table 4.8: Results attained by the examined DE variants for 100 independent runs of the optimization problem including the prediction of the SM^{CUC}-TPS.

	DE/rand/1/bin				DE/rand/1/exp				DE/rand/1/expS			
Operators	Mean	SD	SR	AR	Mean	SD	SR	AR	Mean	SD	SR	AR
/FR	1.54E-03	5.75E-04	0	0.00	5.36E-03	2.36E-03	0	0.00	4.65E-03	1.79E-03	0	0.00
/FR/blk	1.10E-03	4.64E-04	0	0.00	1.25E-03	4.34E-04	0	0.00	1.35E-03	4.51E-04	0	0.00
/FR/PO	1.58E-03	5.54E-04	0	0.00	5.13E-03	1.69E-03	0	0.00	4.92E-03	1.52E-03	0	0.00
/FR/blk/PO	1.21E-03	4.29E-04	0	0.00	1.33E-03	4.81E-04	0	0.00	1.46E-03	4.49E-04	0	0.00
/RRH	2.40E-06	7.68E-06	91	1.00	1.27E-05	2.39E-05	66	0.37	2.64E-05	4.92E-05	57	0.29
/RRH/blk	1.07E-06	5.26E-06	96	0.97	1.60E-06	6.38E-06	94	0.83	8.01E-07	4.58E-06	97	0.83
/RRH/PO	2.67E-06	8.05E-06	90	1.00	1.30E-05	2.95E-05	69	0.24	3.32E-05	8.02E-05	59	0.24
/RRH/blk/PO	2.67E-07	2.67E-06	99	0.89	8.01E-07	4.58E-06	97	0.72	1.07E-06	5.26E-06	96	0.59

Application of the MAEA on the test case using the SM^{ED} and SM^{CUC}

Table 4.9 present the statistical results of the computational experiments examining the application of the MAEA on the test case including the SM^{ED} and SM^{CUC}. It can be seen that the optimization approach had identified the same near-optimal solution in both examined cases. Moreover, Figure 4.4 presents the relative error ($\frac{f^{bf}-f^{*}}{f^{*}}$) of the best-found-solution in the archive in comparison to the best-found-solution among all independent runs (f^{*}).

Table 4.9: The table presents the statistical results of the computational experiments examining the application of the MAEA on the test case including the $\mathrm{SM}^{\mathrm{ED}}$ and $\mathrm{SM}^{\mathrm{CUC}}$. The distribution of the function values for the best-found solution derived for the 30 independent runs on the $\mathrm{SM}^{\mathrm{ED}}$ case are presented in Table 4.9a. Similarly, the results for the $\mathrm{SM}^{\mathrm{CUC}}$ case are presented in Table 4.9b.

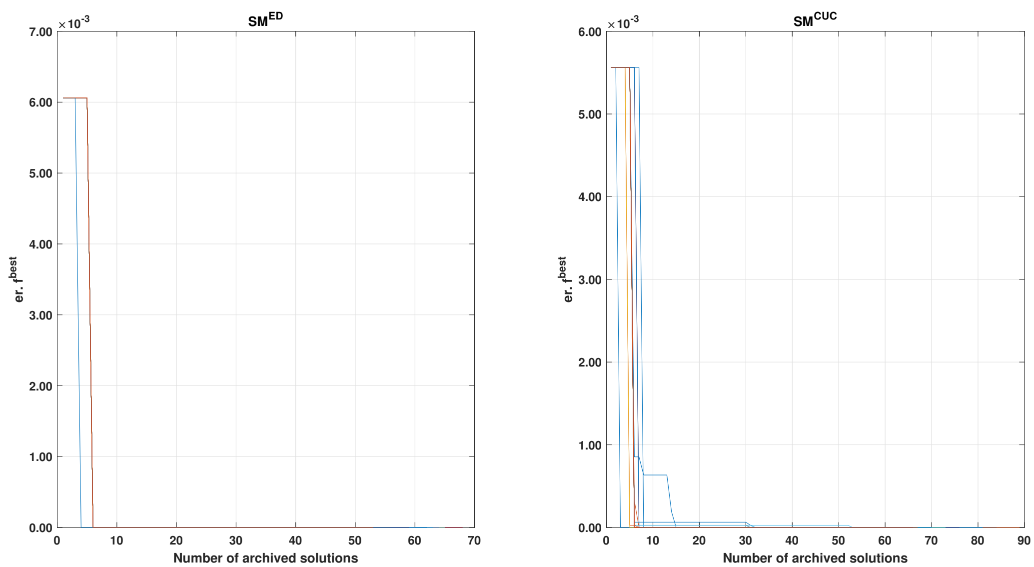
	$TC(\mathbf{x}^{best})$
Min (€)	1.647e+10
Max (€)	1.647e+10
Mean (€)	1.647e+10
SD (€)	0.00E+00
SR	30/30

(a) Statistical results for SM^{ED} case.

	$TC(\mathbf{x}^{best})$
Min (€)	1.895E+10
Max (€)	1.895E+10
Mean (€)	1.895E+10
SD (€)	0.00E+00
SR	30/30

(b) Statistical results for the $\mathrm{SM}^{\mathrm{CUC}}$.

Figure 4.4: The figure presents the progress plot of the two examined cases ($\mathrm{SM}^{\mathrm{ED}}$ and $\mathrm{SM}^{\mathrm{CUC}}$) based on the relative difference of the best archived solution ($\frac{f^{bf}-f^{*}}{f^{*}}$). The best-found solution among all independent runs is set as f^{*}.



Application of the MAEA to identify a targeted installed capacity

The results of the numerical experiments for the two cases that utilized the MAEA to identify a targeted installed capacity which had been set as the global optimum for the constraint minimization problem are presented in Table 4.10. In the case in which the solution minimizing the investment and fix cost had been set as the objective, the MAEA identified the installed capacity in all independent runs. Moreover, for the second case in which randomly generated and repaired solutions had been used to determined the targeted installed capacity the MAEA identified the installed capacity in 99/100 independent runs. The value of the best-found solution based on the entries in the archives is presented in Figure 4.5. These results indicate that the MAEA approach could consistently identify the targeted installed capacity.

Table 4.10: Table 4.10a presents the simulations utilized by the MAEA to identify the global optimum for the constraint minimization problem where a targeted installed capacity had been considered. Table 4.10b depicts the simulations utilized by the MAEA in which the randomly generated solutions are considered for setting the targeted installed capacity.

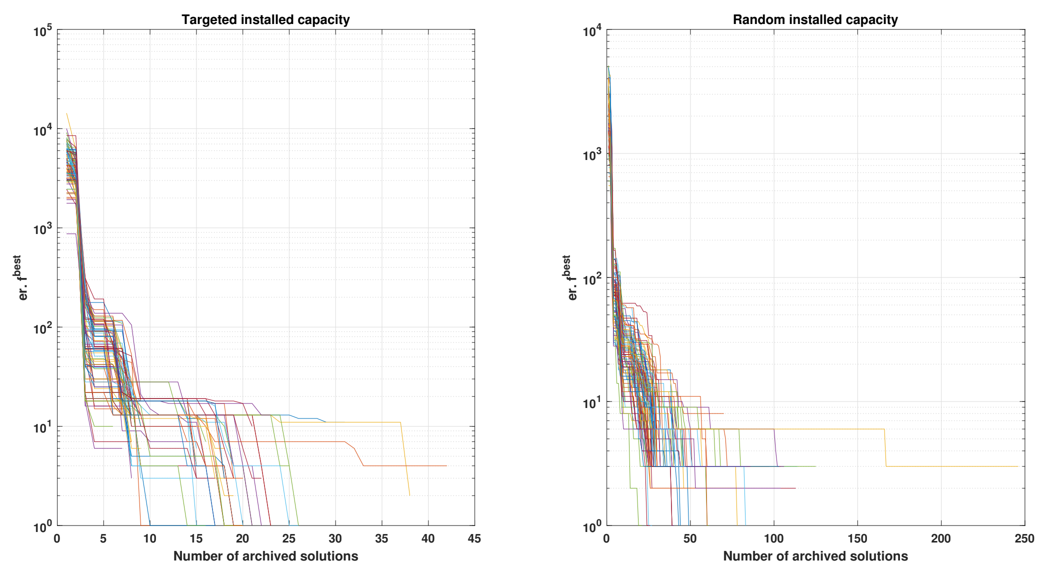
	Number of sim.
Min (Sim.)	241.00
Max (Sim.)	453.00
Mean (Sim.)	328.68
SD (Sim.)	50.90
SR	100/100

(a) Targeted installed capacity case for 100 independent runs.

	Number of sim.
Min (Sim.)	361.00
Max (Sim.)	989.00
Mean (Sim.)	502.94
SD (Sim.)	106.95
SR	99/100

(b) Randomly generated targeted installed capacity case for 100 independent runs.

Figure 4.5: The figure presents the progress plot of the two examined cases regarding the targeted installed capacity. The value of the best archived solution (f^{bf}) is depicted in logarithmic scale.



4.5.2 Comparison of the solutions attained by the simulation models

In this Section the derived near-optimal solutions attained by employing the SM$^{\mathrm{ED}}$ and SM$^{\mathrm{CUC}}$ are examined. The solution derived by the MILP solver for the formulation excluding a SM (noSM) is also presented to serve as a point of reference since the investment and FO&M cost terms had been an important cost factor to determine the near-optimal capacity additions. Each derived solution is post-processed using both available SMs. More specifically, noSM-SM$^{\mathrm{ED}}$ and noSM-

SM^{CUC} represent the solution derived by the MILP solver and post-processed by the SM^{ED} and SM^{CUC}, respectively. Similarly, SM^{CUC}-SM^{ED} and SM^{ED}-SM^{CUC} refer to the solutions derived by the SM^{ED} and SM^{CUC} and post-processed by SM^{CUC} and SM^{ED}, respectively. The comparison is carried out in terms of derived investment decisions, values of cost terms and generation mix.

Comparison of the attained investment decisions and the installed capacity

Figure 4.6 presents the derived investment decisions for the three examined cases. Moreover, the corresponding installed capacity for each solution are depicted. Some differences in the attained investment decisions are observed. For example, capacity additions in CH are made to ensure the minimum required capacity in the noSM case. In contrast, an increase is observed when the SM^{ED} or the SM^{CUC} is included. An increase is also observed in SPV capacity additions. These can be attributed to the assumed techno-economic characteristics of CH and SPV, e.g. zero generating cost and relatively higher investment cost. The inclusion of the variable cost term decreased the capacity additions in TU2 and TU3. In general, the noSM case suggests which combination of investment should be made when the operating perspective is neglected. Therefore, alterations made to the investment decisions by the SM^{ED} or the SM^{CUC} can be attributed to a sufficient variable cost reduction to render a higher investment and FO&M cost more cost-efficient.

Moreover, differences have been attained also among the SM^{ED} and the SM^{CUC} cases. For example, (i) a reduction is observed in capacity additions in TU1 and TU3, (ii) an increase in HS installed capacity, and (iii) a minor increase in SPV units are observed when the SM^{CUC} is included. These differences emerge since the aspects of the short-term operation of the power system included within the employed SMs (SM^{ED} or the SM^{CUC}) differs (Appendices B and A).

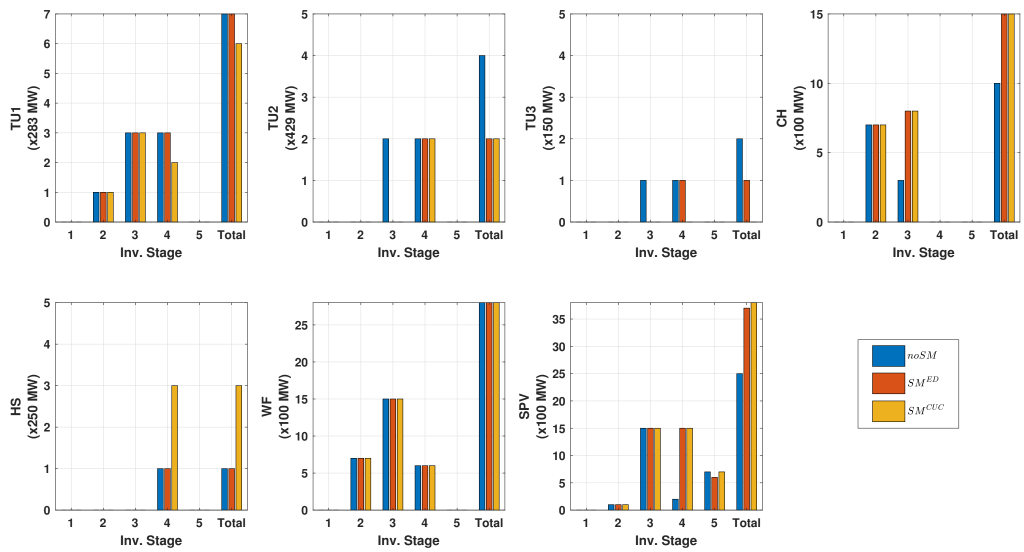
Comparison of the attained values of the cost terms

Table 4.11 and Figure 4.7 present the values of the cost terms for the examined cases during the planning horizon. The total cost values using the same SM exhibit relatively small differences which however have had an impact on the derived solutions. It can be observed that the variable cost of a near-optimal solution processed by the SM^{ED} is lower in comparison to when SM^{CUC} is used. Moreover, the derived solution by the optimization approach including SM^{CUC} exhibits a lower operating cost in comparison to the one derived by SM^{ED} when post processed by the SM^{CUC}. In addition, the corresponding comparison among the SM^{ED} and SM^{CUC}-SM^{ED} cases suggests that the former exhibits a lower value of the cost term. This can be attributed to the differences in technical detail of the SMs since different anticipated generations schedules and cost are obtained. These resulted to an alteration of the ranking of the solutions. Consequently, the near-optimal solution attained by employing the SM^{CUC} could have been identified when the SM^{ED} had been employed and rejected as less cost-efficient. More specifically, Figure 4.7 indicates that the SM^{ED} underestimates the operating cost for a candidate generating fleet rendering it more cost-efficient than others that could be less efficient when examined with a model including higher detail (SM^{CUC}).

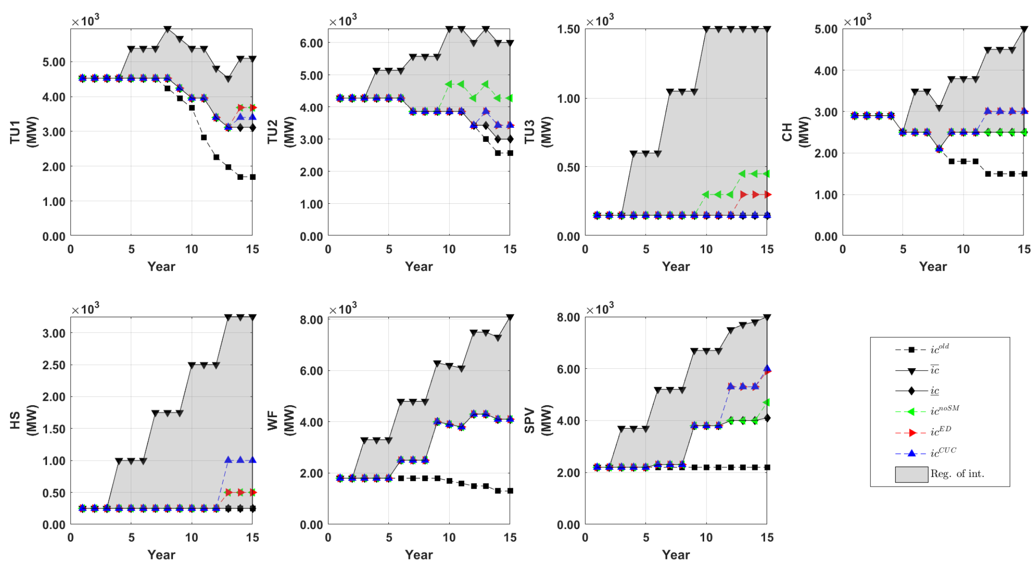
Comparison of the attained based on the derived generating mix

Differences are also observed in the derived anticipated generation mix. These are presented in Figure 4.8. In particular, the anticipated generation of each technology group is depicted for the

Figure 4.6: The figure presents the derived investment decisions and the corresponding installed capacities for the three examined cases (noSM, SM$^{\mathrm{ED}}$ and SM$^{\mathrm{CUC}}$).

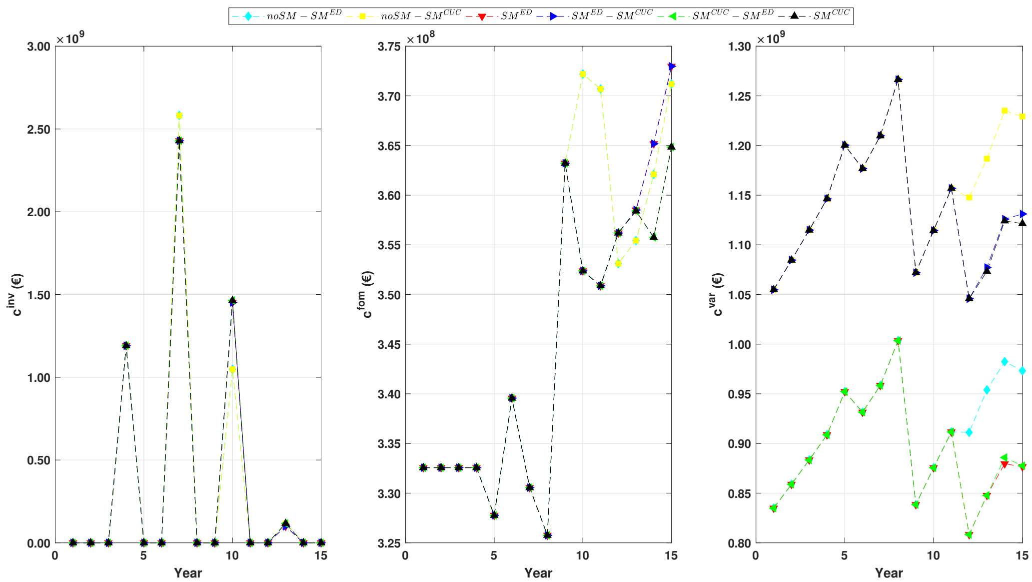


(a) Comparison of investment decisions per technology group and investment stage.

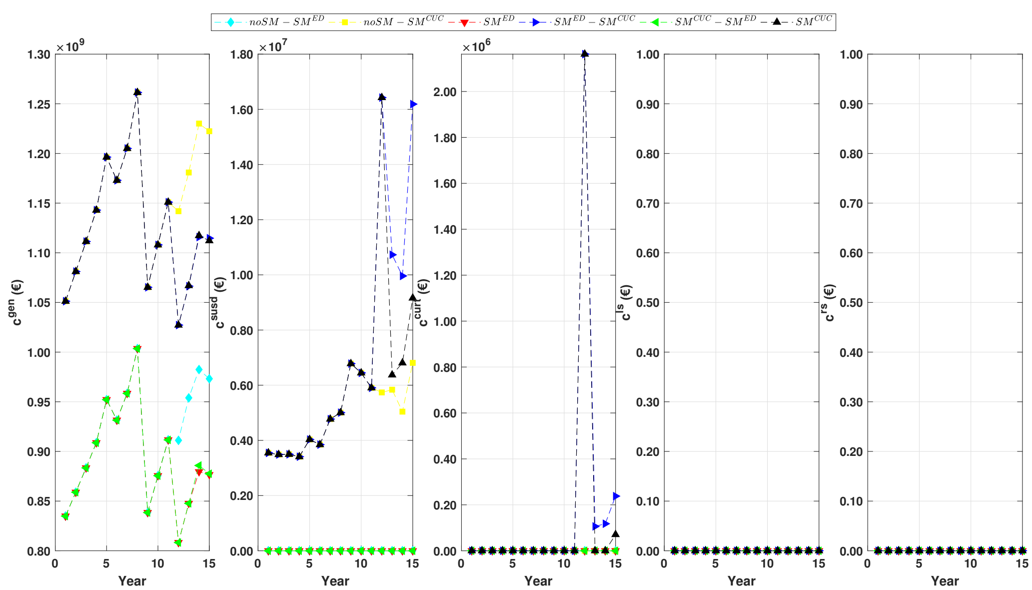


(b) Comparison of installed capacity per technology group w.r.t. the region of interest.

Figure 4.7: The figure presents the values of the cost terms based on the near-optimal solution attained for the examined cases (SM^{ED} and SM^{CUC}). The values of the cost terms derived by post-processing are also presented (noSM-SM^{ED}, noSM-SM^{CUC}, SM^{ED}-SM^{CUC} and SM^{CUC}-SM^{ED}).



(a) Values of the investment, FO&M and variable cost terms.



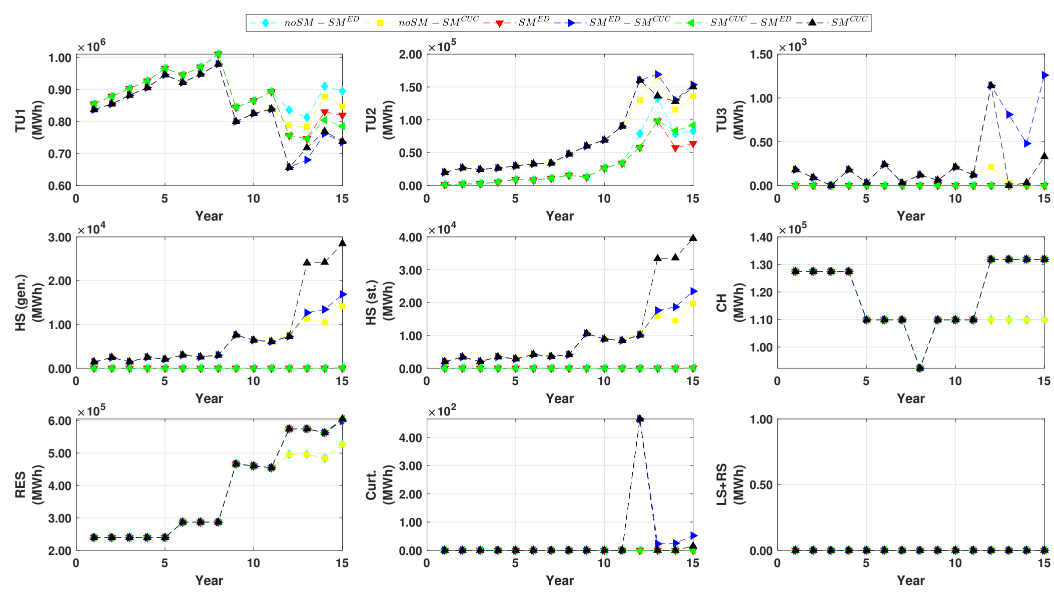
(b) Values of the cost terms of the SM.

Table 4.11: Total cost and values of cost terms for the noSM-SM^{ED}, noSM-SM^{CUC}, SM^{ED}, SM^{ED}-SM^{CUC}, SM^{CUC}-SM^{ED} and SM^{CUC} cases. The relative differences of the total cost and the aggregated discounted variable cost for each near-optimal solution are also presented.

	noSM-SM^{ED}	noSM-SM^{CUC}	SM^{ED}	SM^{ED}-SM^{CUC}	SM^{CUC}-SM^{ED}	SM^{CUC}
TC [b€]	16.565	19.055	16.465	18.950	16.476	18.950
Relative Diff.	-0.131	-	-0.131	-	-0.131	-
$\sum_{yr}[DD_{yr}c_{yr}^{inv}]$ [b€]	3.519	3.519	3.650	3.650	3.665	3.665
$\sum_{yr}[DD_{yr}c_{yr}^{fixed}]$ [b€]	3.571	3.571	3.553	3.553	3.544	3.544
$\sum_{yr}[DD_{yr}c_{yr}^{var})]$ [b€]	9.475	11.965	9.263	11.748	9.267	11.740
Relative Diff.	-0.208	-	-0.212	-	-0.211	-

examined near-optimal solution and the SMs (noSM-SM^{ED}, noSM-SM^{CUC}, SM^{ED}, SM^{CUC}-SM^{ED} and SM^{ED}-SM^{CUC} and SM^{CUC}). However, comparisons should be restricted to generation levels derived for the same near-optimal solutions (e.g. SM^{CUC}-SM^{ED} and SM^{CUC}). For example, TU1 are utilized more frequently when the SM^{ED} is used in comparison to SM^{CUC}. On the contrary, TU2 are utilized more frequently when the SM^{CUC} is used in comparison to SM^{ED}. Moreover, TU3 and HS are not utilized when SM^{ED} is used. A low level of curtailment is also observed, in comparison to the RES penetration level, for the cases employing the SM^{CUC}. In addition, CH exhibit identical generation levels due to the energy content restrictions included in both SMs.

Figure 4.8: The figure presents the total production based on the near-optimal solutions attained for the examined cases (SM^{ED} and SM^{CUC}) and the ones derived by post-processing (noSM-SM^{ED}, noSM-SM^{CUC}, SM^{ED}-SM^{CUC} and SM^{CUC}-SM^{ED}). Reserve shortage is not included in the SM^{ED} models.



The differences in the generation mix arise due to the derived generation scheduled be each SM. Figures 4.9, 4.10 and 4.11 present the generation schedules of the final year for the three near-optimal solutions computed by both SMs. The differences observed in Figures 4.9-4.11 for the generation schedules among the pairwise comparisons are consistent over all examined time periods (days) for the final year of the planning horizon. These could be attributed to the lower technical detail included in SM$^{\mathrm{ED}}$ in comparison to the SM$^{\mathrm{CUC}}$. In particular, omitting a representation of the technical restriction of a UCP in the SM resulted to underestimate the operational flexibility requirements and, therefore, flexibility providers, such as HS, are rendered less necessary. For example, the absence of ramping restriction in the SM$^{\mathrm{ED}}$ resulted to generating schedules for which thermal units are anticipated to ramp up or down in higher ratios. The operating statuses of thermal units, reserve requirements and the corresponding costs are also not captured. On the contrary, the generation schedule derived by th SM$^{\mathrm{CUC}}$ determines the statuses of thermal units by including restrictions on their operation, such as on/off statuses, minimum up/down times and ramping limitations, and the reserve requirements of the system. Consequently, the aforementioned indicate that the level of technical detail that a SM exhibits can have an impact on the signals provided by examining the anticipated generation mix.

The SM$^{\mathrm{CUC}}$ considers aspects of the short-term operation. Reference [191] examined the applicability of a CUC formulation to represent a UCP and identify that despite some limitations it may capture aspects of the short-term operation. For the context of long-term planning, it could exhibit acceptable accuracy when computational limitations are considered. However, assumptions have been made regarding the temporal, technical and spatial detail of the short-term operation. Since the SM$^{\mathrm{CUC}}$ formulation is a problem approximation of a detailed UCP, the presented results could suggest the importance of including an SM that could capture main aspects of the short-term operation of a power system within long-term multi-period planning. This is required to adequately assess thermal and non-thermal flexibility providers since including limited restriction on thermal units may underestimate operational flexibility requirements and, consequently, the utilization of flexibility providers.

4.5.3 Quality of the approximation

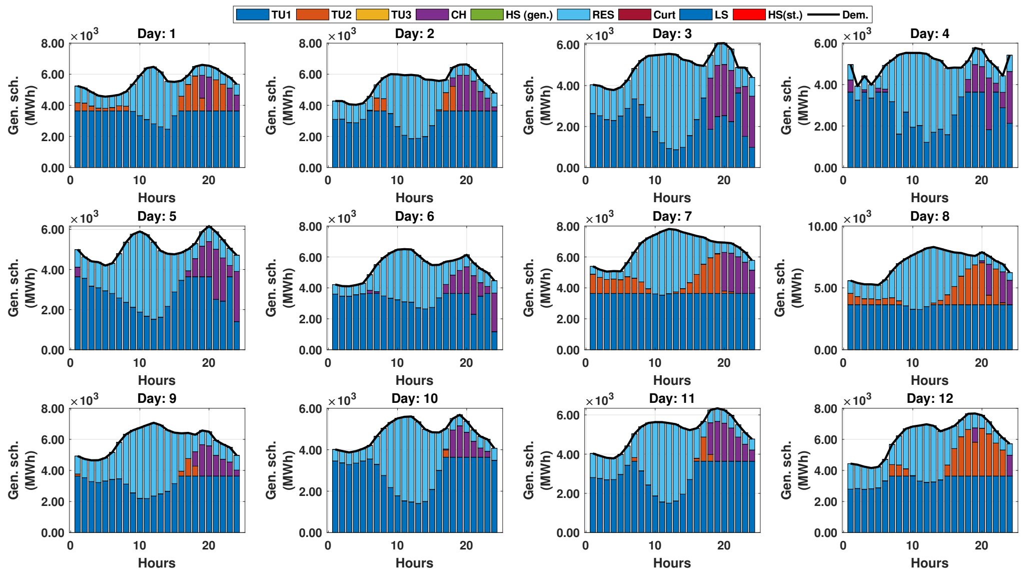
In this Section, the quality of the UFA is assessed since the approximation serves as a replacement of the computationally expensive simulations. Attaining an accurate UFA could be important as high errors might imply that it is less likely that the optimization algorithm has converged to an near-optimal solution. However, relatively large errors do not necessarily imply that the approximation is not adequate for EA-based search (Reference [14]).

In particular, the values of the performance metrics employed (Section 4.4.4) are presented for the global AMs in this Section. The derived datasets from the numerical experiments including the SM$^{\mathrm{ED}}$ and SM$^{\mathrm{CUC}}$ are examined. For assessing the accuracy of the UFA, the data points of each interdependent run are used which are available in the attained datasets and include up to 500 data points in total. However, the number of data points in each archive of each target year can differ due to the occurred allocation during the optimization run. This has an impact on the values of the metrics and should be considered. Moreover, the normalized values of the RMSE (NRMSE) and MAE (NMAE) metrics are also depicted to facilitate the comparison of the UFA for each target year.

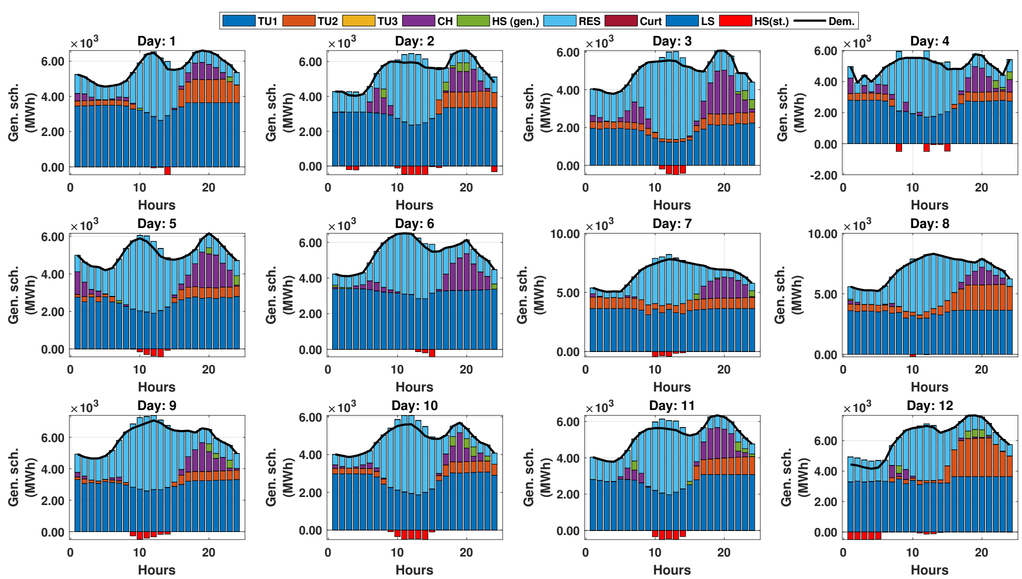
UFA accuracy for the SM^{ED} case: Figure 4.12 presents the values of the metrics for the case including the SM$^{\mathrm{ED}}$. Based on the metrics employed it can be seen that a decent UFA has been achieved.

UFA accuracy for the SM^{CUC} case: The values of the metrics for the case including the SM$^{\mathrm{CUC}}$

Figure 4.9: The figure presents the attained generation schedule for the near-optimal solution derived without including a SM and post-processed by the SM^{ED} (noSM-SM^{ED}) in comparison to the post-processed generation schedule by the SM^{CUC} (noSM-SM^{CUC}).

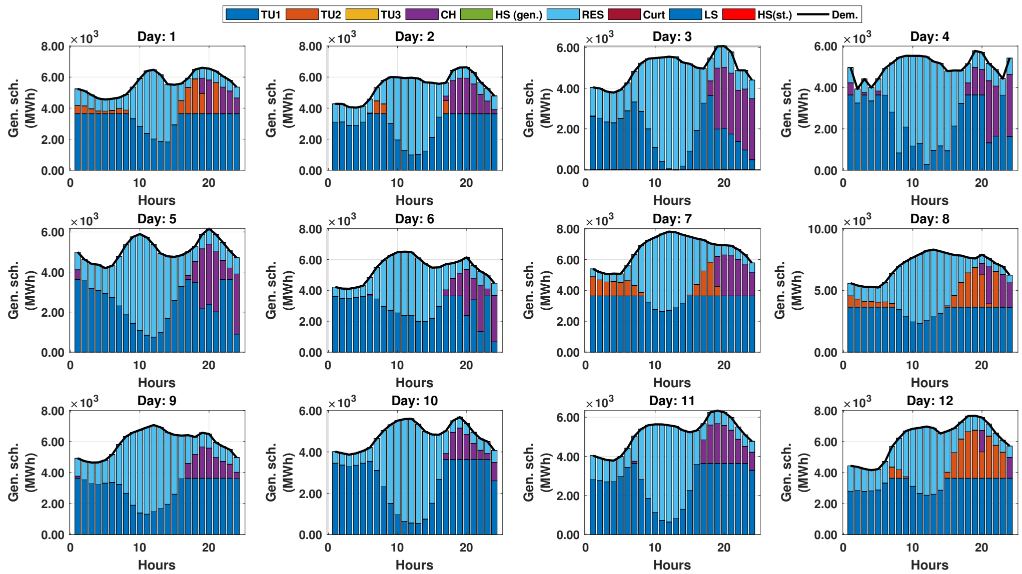


(a) Generation schedule attained by noSM-SM^{ED}.

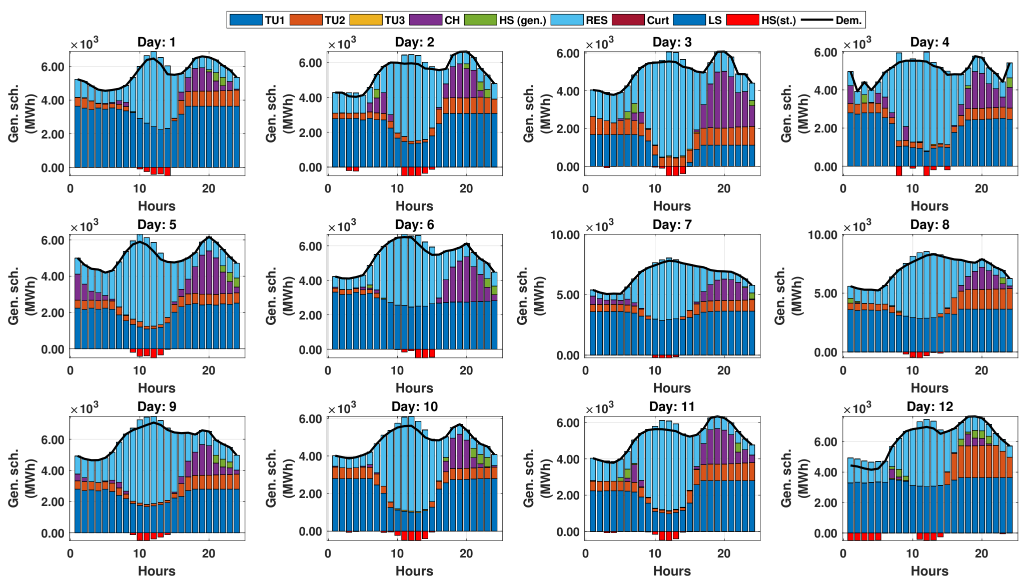


(b) Generation schedule attained by noSM-SM^{CUC}.

Figure 4.10: The figure presents the attained generation schedule for the near-optimal solution derived by SM$^{\mathrm{ED}}$ in comparison to the post-processed generation schedule by the SM$^{\mathrm{CUC}}$ (SM$^{\mathrm{ED}}$-SM$^{\mathrm{CUC}}$).

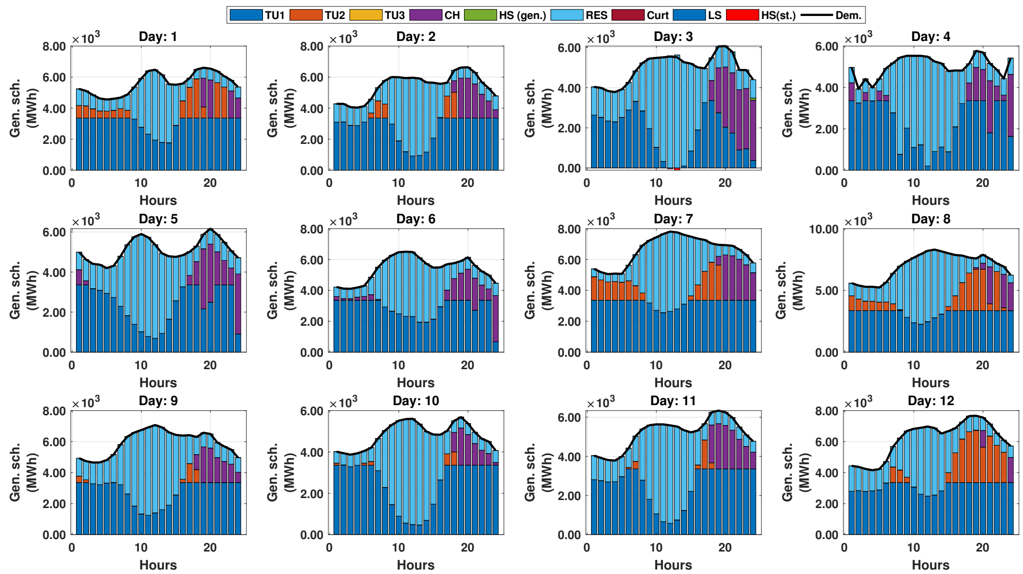


(a) Generation schedule attained by SM$^{\mathrm{ED}}$.

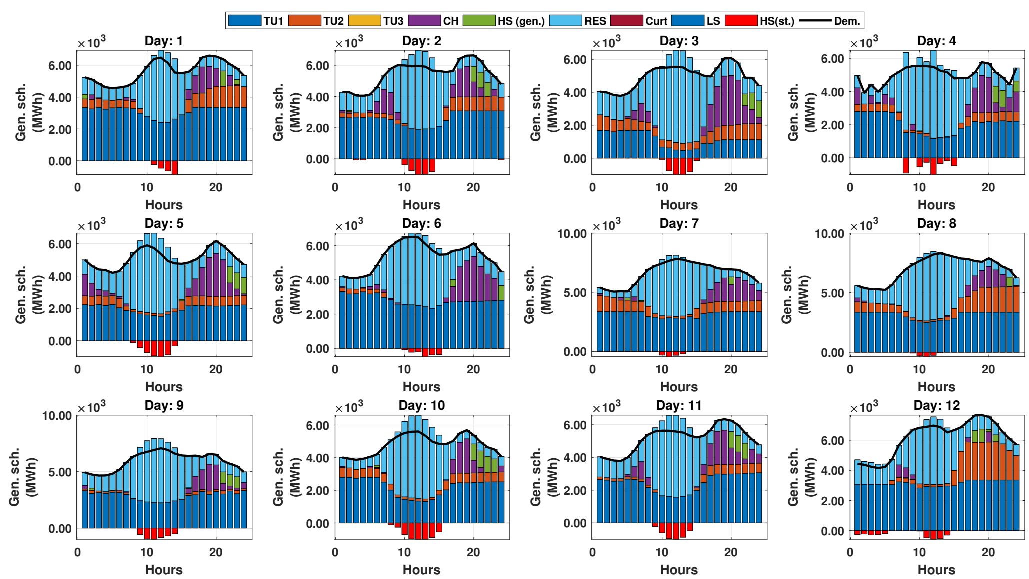


(b) Generation schedule attained by SM$^{\mathrm{ED}}$-SM$^{\mathrm{CUC}}$.

Figure 4.11: The figure presents the attained generation schedule for the near-optimal solution derived by SM$^{\mathrm{CUC}}$ in comparison to the post-processed generation schedule attained by the SM$^{\mathrm{ED}}$ (SM$^{\mathrm{CUC}}$-SM$^{\mathrm{ED}}$).



(a) Generation schedule attained by SM$^{\mathrm{CUC}}$-SM$^{\mathrm{ED}}$.



(b) Generation schedule attained by SM$^{\mathrm{CUC}}$.

are presented in Figure 4.12. The values of the metrics indicate that an acceptable UFA has been achieved.

Since the errors are not relatively high, the values of the metrics could imply that the algorithm had not converged to the same near-optimal solution due to a highly inaccurate UFA in both the $\mathrm{SM}^{\mathrm{ED}}$ and $\mathrm{SM}^{\mathrm{CUC}}$ cases. A comparison of the values of the metrics suggests that the UFA for the $\mathrm{SM}^{\mathrm{ED}}$ is more accurate than the one achieved for the $\mathrm{SM}^{\mathrm{CUC}}$ case. Moreover, the values of the metrics for the first half of the planning horizon indicate an improved UFA in comparison to the remaining ones in both cases. This could be associated to differences in the size of the region of interest, in the dimensionality of the target years and/or in the possible combination of installed capacities in each target year.

4.5.4 Sensitivity of the variable cost by a visual analysis

The output of the SOO optimization problem is a near-optimal solution. However, it may be of interest to examine the sensitivity of the cost terms towards the attained installed capacity in specific target years. This could provide information on the impact of the attained investment decisions. More specifically, investment cost is an important factor for determining an optimal expansion plan. However, it could be of interest to examine the impact of higher or lower capacity additions in specific technology groups on the operating cost.

The optimization approach includes external archives (archives A_{yr}^{arhv} and S^{arhv}) which are used to store the data points that have been costly evaluated and are used to built the AMs during the search. These archives are an output of the approach.

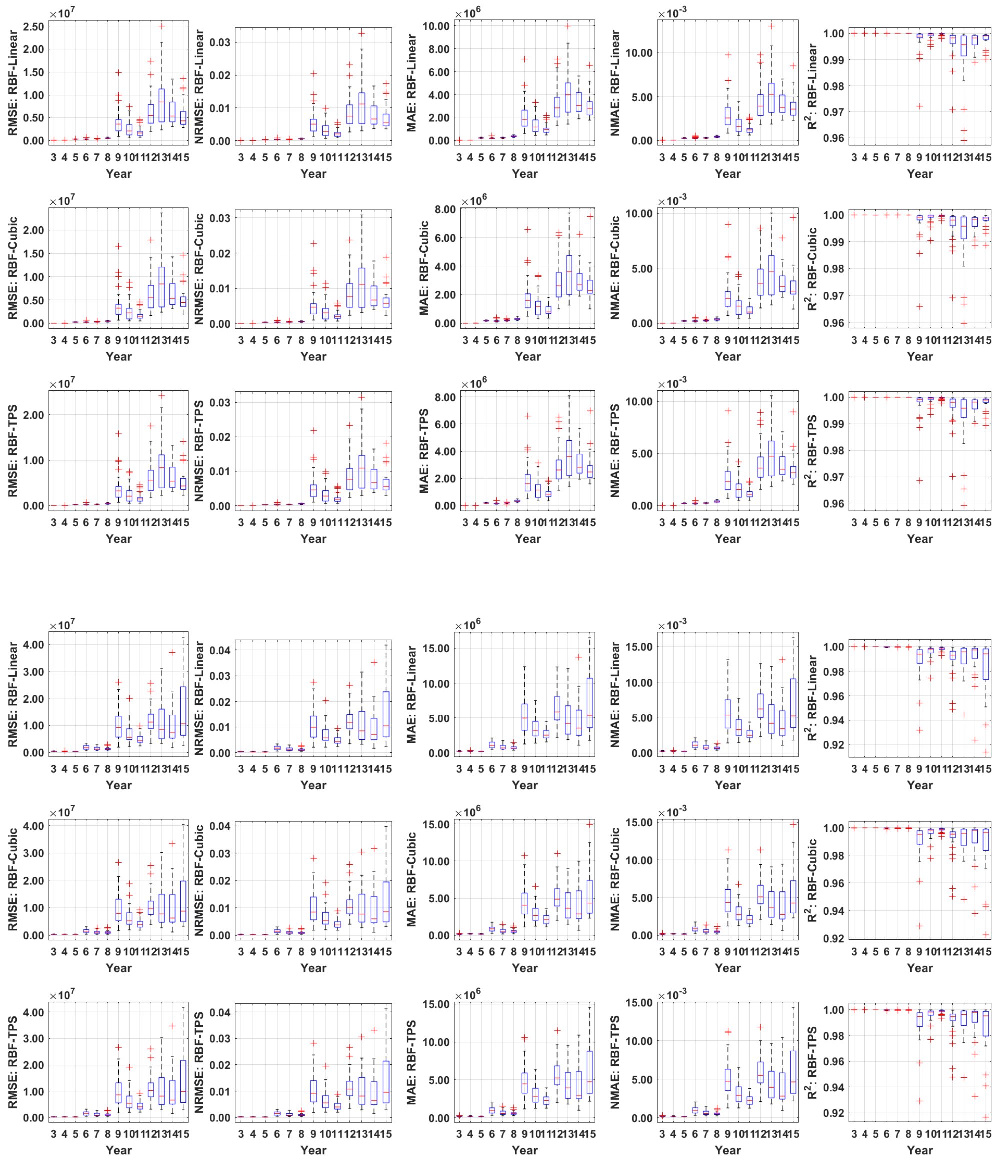
The archives are employed to provide an visual illustration of the sensitivity of the variable cost w.r.t. the installed capacity. This may provide information regarding the results of the model and could be further processed to reveal some general trends. However, the quality of such a visual illustration can be limited by the attained accuracy of the AMs since a decent approximation is required. Moreover, the quality of the approximation relies on the data points selected during an independent optimization run.

The visual illustration is attained by computing all possible combinations of pairs of technology groups within the region of interest. The installed capacity of the remaining technology groups, attained by the near-optimal solution, are remain constant.

The prediction of the variable cost for the combinations generated by the aforesaid procedure are then attained using the AMs (the linear, cubic and TPS RBF) using the dataset of an independent optimization run. Also, the corresponding variable cost is computed for these combinations by the $\mathrm{SM}^{\mathrm{CUC}}$ model to serve as a point of reference for a comparison. However, if the computational cost of the $\mathrm{SM}^{\mathrm{CUC}}$ is increased (e.g. an increase in the temporal resolution) attaining the point of reference can be rendered computational expensive. In such a case, the accuracy of the derived visualization can be based mostly on the metrics employed to assess the UFA. A number different combinations can be selected to be examined for each target year of the planning horizon. In general, higher errors at specific areas of the region of interest indicate that the quality of the approximation would benefit by the addition of a data point, in the archive, near that area.

Four examples are presented which are based on the near-optimal solution derived using the $\mathrm{SM}^{\mathrm{CUC}}$ (Figures 4.13 and 4.14). The first two present the sensitivity of the variable cost towards the installed capacity for the final year of the planning horizon w.r.t. two flexibility providers (TU3 and HS) and two RES technology groups (SPV and WF). For the remaining two examples, the SPVs-HS and CH-WF pairs are examined.

Figure 4.12: The figure presents the values of the RMSE, NRMSE, MAE, NMAE and R^2 metrics. The first, second and third row correspond to the values of the metrics for the linear, cubic and TPS RBF models for the case where SM^{ED} is employed. Correspondingly, the fourth, fifth and sixth row depict the values of the metrics for the linear, cubic and TPS RBF models for the case where SM^{CUC} is employed.



The figures indicate that some technology groups are prioritized (e.g. SPVs over WFs) due to the assumed techno-economic characteristic of the examined technology groups. In addition, it is observed that an increase in the installed capacity of the technology groups triggers variable cost reduction. Despite some observable errors in the approximation, cost reduction is in general captured by the AMs, suggesting that investment and FO&M cost had highly influenced the derived near-optimal investment schedule. More specifically:

- *TU3-HS*: A decreasing trend is observed for the variable cost of the final year of the planning horizon when the installed capacity of the HS and TU3 is increased. Assuming that both technology groups are flexibility providers, HS additions have been prioritized over TU3 additions due to the techno-economic characteristics.

- *SPV-WF*: A similar pattern for the variable cost is also observed for increased numbers of installed SPV and WF capacity. However, for the highest numbers of installed SPV and WF capacity (w.r.t. the region of interest) an increase is observed.

- *SPV-HS*: This pair includes a RES technology group and a flexibility provider. It is observed that an increase in the capacity additions of both technology groups suggests a reduction of the variable cost. On the contrary, for high levels of SPV capacity installations an increase in the variable cost is observed when the installed capacity of HS is reduced. Also, an increase in the HS installed capacity, beyond a required level, does not alter the variable cost since the excess capacity is not utilized.

- *CH-WF*: In contrast to the SPV-WF comparison, an increase in the installed units of both technology groups may trigger variable cost reduction. The generation cost of both aforementioned technology groups is assumed zero which has contributed to this result. Moreover, the output of CH units may be controlled, to a certain extent, in comparison to SPV and an efficient allocation of the available generation by CH could provide operational flexibility. This can be attributed to the limited technical restrictions considered for CH (e.g. ramping capabilities).

4.6 Discussion, limitations and future research directions

The results derived by this approach may provide some insights to decision makers regarding the challenges arising for a transition towards higher shares of RES generation. However, the approach is indented to be used in parallel to other models since there are some important limitations that should be regarded when analysing its results. In particular, limitations arise from the problem formulation. For example:

1. The formulated optimization problem does not necessarily capture the operation of a real market and its participants' behaviour. Therefore, a future direction may be to consider a wider range of market conditions.

2. A deterministic GEP model has been examined. However, long-term planning involves uncertainty deriving from various sources (e.g. fuel prices and demand growth).

Figure 4.13: The figure presents the prediction (blue circles) of the employed AMs for the installed capacity of the TU3-HS and SPV-WF pairs. The corresponding value of the cost term attained by the SM are also presented (black circles). The installed capacity for the near-optimal solution derived is marked with a red circle.

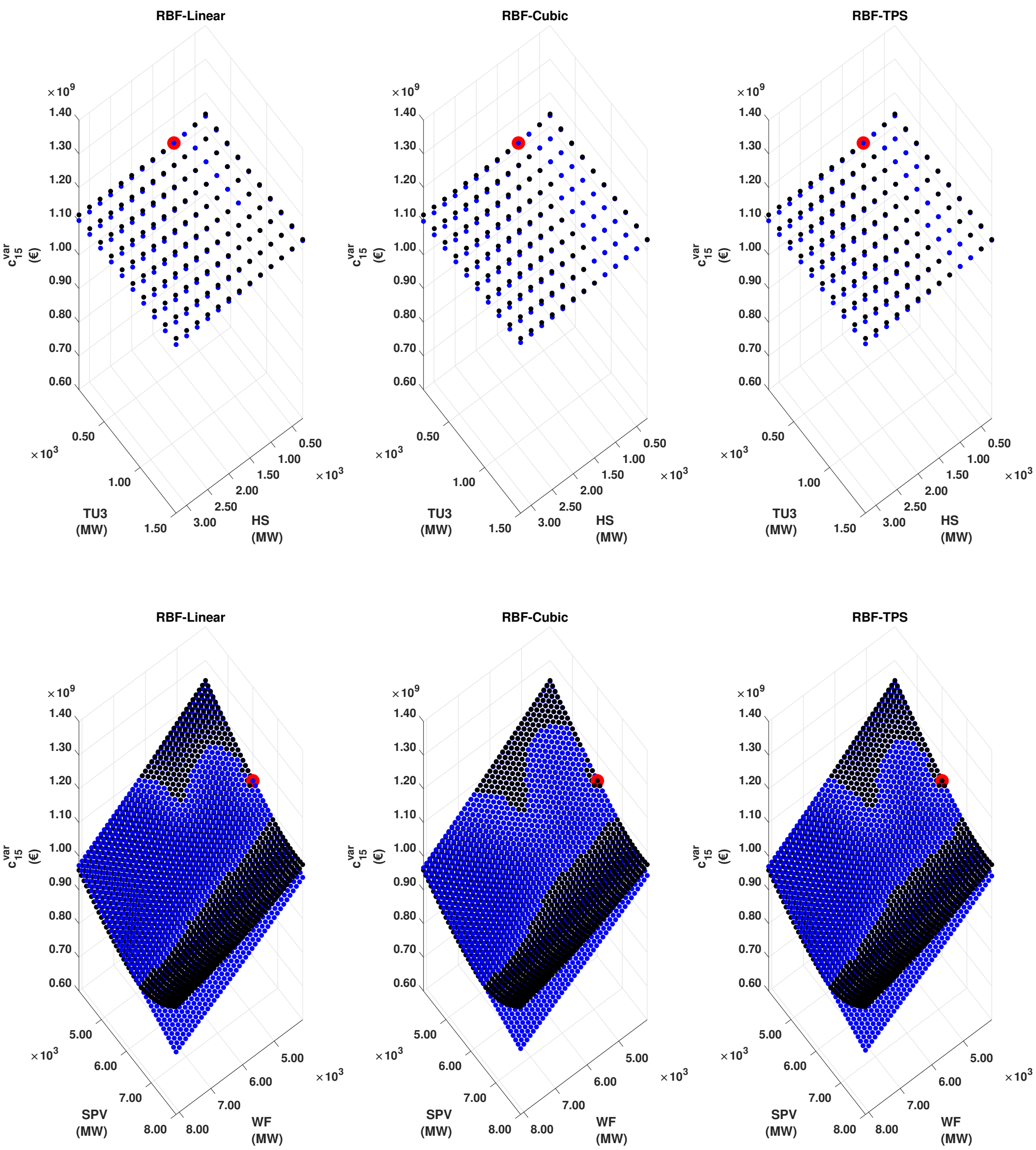
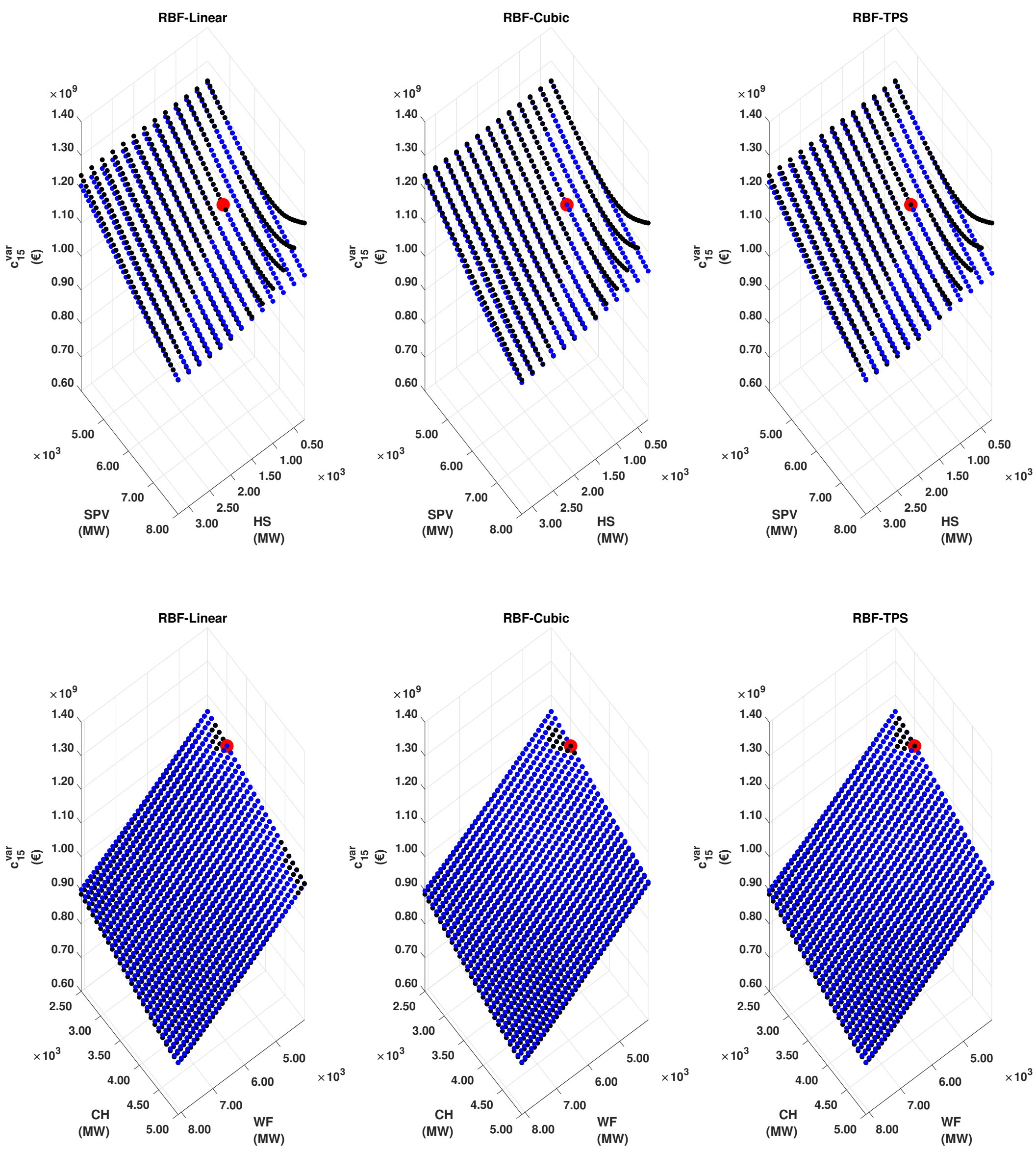


Figure 4.14: The figure presents the prediction (blue circles) of the employed AMs for the installed capacity of the SPV-HS and CH-WF pairs. The corresponding value of the cost term attained by the SM are also presented (black circles). The installed capacity for the near-optimal solution derived is marked with a red circle.



3. Assumptions have been made regarding the power system. It had been represented by a single zone and the location of new units has not been considered. Consequently, an extension of the examined formulation to evaluate also the location of new units could be of interest.

4. An expansion of the transmission grid could be also important when examining high shares of RES generation. Herein, a combined GEP and TEP formulation has not been examined and could be considered as a future direction.

Furthermore, the SM employed is based on the CUC context, which has been demonstrated to capture aspects of operational flexibility (References [7, 8, 191]). However, the inclusion of a CUC-based model as the SM exhibits also some limitations. For example:

1. Unit aggregation has been implemented which could have an impact on the representation of the short-term operation of the power system in comparison to a unit based approach (e.g. a single linear segment is employed to represent the generation cost of a group of units). However, unit aggregation can be an acceptable simplification to address the computational restrictions (Reference [202]).

2. The operation cost has been computed for a series of independent days. This may lead to underestimating the impact of the initial state of thermal units and the ability of storage units to transfer stored energy. This could be (in part) moderated by considering longer time periods and increasing the computational cost of the SM.

3. The SM aims towards capturing hourly dynamics. However, accounting for sub-hourly dynamics may lead to a detailed evaluation of the operational flexibility of an examined set of capacity additions or of a generation schedule (Reference [203]).

4. The spatial detail of the short-term operation has not been considered. For example, network-related limitations, interactions with neighbouring sectors, benefits arising from spatial smoothing of RES generation and trade flows have been neglected.

5. The maintenance scheduling has not been considered in detail.

Most of the aforementioned limitations emanate from attempts made to limit the computational cost of the SM since it is a mixed-integer optimization problem. Nevertheless, focus had not been directed on suggesting an optimal SM for a power sector (e.g. a long-term UCP), but rather, an overall optimization approach that can be modified to consider a SM.

Moreover, due to the heuristic nature of the optimization approach, the model may serve mostly as an approximation for GEP that could be adopted when a computationally expensive SM is employed. The main aspects influencing the optimality of the final solution can be categorized to the following:

1. *The efficiency of the optimization algorithm*: An attempt to examine the efficiency of the algorithm has been made by carrying out a number of numerical experiments. However, the algorithm relies on stochastic operators and heuristic mechanisms to converge and, in contrast to mathematical optimization methods, optimality guarantees can not be provided. Some main aspects that could influence the algorithm's efficiency are the following:

   (a) The size of the search space (e.g. the dimensionality of the problem and number of possible combinations).

   (b) The size of the feasible region (e.g. the number of feasible combinations).
   (c) The landscape of the search space (e.g. a highly multi-modal function).

2. *The accuracy of the UFA:* The AMs are used to provide an estimate of the TM. It is built based on selected data points which are a subset of all possible combinations. Therefore, a poor UFA could be misleading in some cases. However, error metrics can be computed to provide an indication regarding its accuracy. Some main aspects that could influence the algorithm's efficiency related to the UFA are the following:
   (a) High dimensionality (e.g. a large number of technology groups): RBF models have been employed, as the latter have been reported to be more scalable in comparison to other AMs (References [166, 167]). However, problems involving a large number of variables for which the AM is built upon may impact the number of simulation required to attain an accurate approximation (References [14, 167]).
   (b) The size of the region of interest (i.e. the region defined by the restrictions on the lower and upper limit of installed capacity in each year): A larger region of interest includes a larger number of possible combinations. More importantly, a region of interest including unrealistic cases of installed capacity could have an impact on the UFA. Evaluating such unrealistic cases, using the SM, could result in data points to enter the archive that exhibit relatively extremely high cost values. Such examples could be installed capacities that exhibit large values of load shedding and/or reserve shortage as these are associated with high penalty cost parameters. These may have an impact on the UFA accuracy. Therefore, unrealistic cases should be restricted as much as possible based on available a priori knowledge. Towards mitigating the latter, a pre-processing step had been included to define the region of interest. Nevertheless, data points exhibiting extreme values may still enter the dataset. The latter may have a relatively lower impact on the locally trained AMs as such data points might be omitted from the dataset used to train the local models. This could be the case when such points exhibit a sufficient distance from the resulting installed capacity of the best-found solution.
   (c) The fitness landscape: It may have an impact on the UFA as it is in general challenging to build an approximate model that has the same global optimum as the underlying function, especially for highly multi-modal functions (Reference [14]).

3. *The selected SM*: A SM should represent the power sector operation as accurately as possible to facilitate as a cost indicator. However, simplifications are required to reduced the computational burden to acceptable levels. Main characteristics of a SM that could affect the optimality of the final solution are the following:
   (a) The computational cost of the SM. A SM associated with a lower computational cost may enable a higher limit of costly function evaluations in comparison to a SM with a higher computational cost. Subsequently, a larger number of function evaluations enables a larger number of data points to enter the archive and, consequently, could (possibly) lead to an improved UFA.
   (b) The level of technical, temporal and spatial detail of the SM. The latter may have an impact on the output of the approach. In particular, the AMs are employed to estimate the output of the SM. Therefore, the ability of a SM to provide an accurate indication of the short-term operation of a power system is important.

(c) The ability of the SM to provide an optimal solution for a given installed capacity. Assuming that the SM is an optimization problem, it is also critical that the solver employed for this optimization problem is able to provide an optimal solution since the output of the aforementioned is used to generate the indicator. For example, a heuristic SM that provides a sub-optimal solution to be used as the indicator could underestimate the adequacy of an examined installed capacity.

Chapter 5

Multi-objective Generation Expansion Planning with increased shares of renewable energy sources by Multi-objective Metamodel-assisted Evolutionary Algorithms

5.1 Motivation and Aims

In this Chapter, a static MOO GEP model for supporting GEP is presented. It aims towards examining cost trade-offs of a computational expensive MOO GEP to facilitate decision-making. Focus is directed on including aspects of short-term operation of a power system within the MOO GEP model. More specifically, it deviates from the existing MOO GEP literature as it attempts to assess operational flexibility requirements and their impact on the considered cost terms by relying on the outcome of a short-term SM. The latter is employed to provide a decent approximation of the UCP. This, however, results in high computational cost.

Towards addressing computational tractability a MOO approach is employed. It is based on frameworks proposed for surrogate-assisted derivative-free optimization. More specifically, the MOO approach includes (i) approximation by AM (Metamodels/surrogate models), (ii) MOEA, (iii) local search and (iv) on-line sampling. It can be perceived as a Memetic Metamodel-assisted Multi-objective EA. Numerical experiments are conducted on a series of test cases to examine its efficiency, tractability and limitations. Moreover, five different economic-environmental MOO GEP variants are examined and trade-offs among the considered cost terms are analysed. The results indicate satisfactory performance of the optimization approach.

5.2 Problem statement and formulation

This Section presents the problem statement, the formulation of the cost terms used to determine the objective functions of the MOO problem and the formulation of the SM employed to provide an estimate of the short-term operation. Lastly, the objective and constraint functions of the five

MOO GEP variants are presented.

5.2.1 Problem statement

Multi-objective optimization refer to optimization problems with more than one conflicting objectives. Commonly, it is expressed as:

$$\begin{aligned} &minimize \;\; \mathbf{F}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), ..., f_m(\mathbf{x})) \\ &s.t. \;\; \mathbf{G}(\mathbf{x}) \leq 0 \\ &\quad\;\; \mathbf{x} \in \mathbb{S} \end{aligned} \tag{5.1}$$

where $\mathbf{x} = (x_1, x_2, ..., x_n)$ is the vector of objective/decision variables, n is the number of decision variables, $\mathbf{F}(\mathbf{x})$ is the objective function vector, m is the number of objective functions, $\mathbf{G}(\mathbf{x}) = (g_1(\mathbf{x}), g_2(\mathbf{x}), ..., g_k(\mathbf{x}))$ is the constraint function vector, k is the number of constraint functions and $\mathbb{S}$ is the search space.

The case of computationally expensive objective functions, computationally cheap constraint functions, bound constraints and $\mathbf{x} \in \mathbb{Z}^n$ is considered. In addition, it is assumed that the minimization problem is restricted by a limited computational budget. Moreover, a limit to the number of considered objectives functions is imposed ($m = 2$), i.e. many-objective optimization problems are not examined. In particular, GEP formulations with two computationally expensive objective functions and computationally cheap constraint functions are examined. It is assumed that a decision variable is a positive integer ($\mathbb{Z}^+$) which represents the number of capacity additions in a candidate technology group. Investments decisions, therefore, are considered in predefined step sizes.

The cost terms, objectives and constraints functions of the formulation are described in the following Sections.

5.2.2 Formulation of the cost terms

A number of different objectives have been considered to examined the corresponding trade-offs. The cost terms that are included in these objectives are formulated as follows (nomenclature is described in the List of Symbols):

- *Investment cost:* The investment cost is computed based on the capacity additions, the investment cost of each step of capacity additions and the step size:

$$c^{inv} = \sum_{\forall a,g} [x^{inv}_{a,g} IC_{a,g} P^{cap\_step}_{a,g}] \tag{5.2}$$

- *Fixed operational and maintenance cost:* FO&M is assumed as a function of the installed capacity and a cost parameter:

$$c^{fom} = \sum_{\forall a,g} [(x^{inv}_{a,g} + IU_{a,g}) FOM_{a,g} P^{cap\_step}_{a,g}] \tag{5.3}$$

- *Green policy support cost:* For the sake of simplicity, a Feed-in-Tariff (FIT) scheme with a fixed price as the green policy is assumed. In addition, RES curtailment is assumed to be fully compensated. Therefore, Green Policy Support Cost (GPSC) is computed as follows:

$$c^{gp} = \sum_{\forall a,g^{res}} [(x^{inv}_{a,g^{res}} + IU_{a,g^{res}}) P^{cap\_step}_{a,g^{res}} Pr^{res}_{a,g^{res}} \sum_{\forall h} [Av^{res}_{a,g^{res},h}]] \tag{5.4}$$

- *Generation cost:* Generation cost is assumed as a function of the fuel cost of the generating capacity:

$$c^{gen} = \sum_{\forall a,g,h} [p_{a,g,h} C^{L}_{g} + x^{su}_{a,g,h} C^{su}_{g} + x^{sd}_{a,g,h} C^{sd}_{g}] \tag{5.5}$$

- *Emission cost:* Emission cost is assumed as a linear function of the power output of each generating technology and the corresponding emission factors:

$$c^{em} = \sum_{\forall a,g,h} [p_{a,g,h} EF_{a,g,h}] C^{em} \tag{5.6}$$

- *Variable operational and maintenance cost:* The variable operational and maintenance (VO&M) cost is assumed as a function of the thermal unit outputs and cost parameters:

$$c^{vom} = \sum_{\forall a,g,h} [p_{a,g,h} VOM_{g}] \tag{5.7}$$

  VO&M cost for RES and hydro technology groups has been assumed to be negligible. Alternatively, it could be included in a similar manner.

- *RES penetration:* RES penetration is represented in monetary terms as the total RES output that contributes to satisfy the demand and is not curtailed voluntarily:

$$c^{rp} = \sum_{\forall a,h} [\overline{p}^{res}_{a,h} - \epsilon^{c}_{a,h}] C^{rp} \tag{5.8}$$

  The anticipated RES generation is computed as follows:

$$\overline{p}^{res}_{a,h} = \sum_{\forall g^{res}} [(x^{inv}_{a,g^{res}} + IU_{a,g^{res}}) P^{cap\_step}_{a,g^{res}} Av^{res}_{a,g^{res},h}], \forall a, h \tag{5.9}$$

- *Non-served Energy cost:* The NSE cost derives from the case where the installed capacity is insufficient to satisfy the demand and/or the reserves. It also includes the case where it cannot satisfy the aforementioned in an economic manner:

$$c^{nse} = \sum_{\forall a,h} [\epsilon^{r}_{a,h} C^{r} + \epsilon^{d}_{a,h} C^{d}] \tag{5.10}$$

- *Curtailment cost:* Curtailment cost is represented by the loss in value from an excess in hydro or RES generation:

$$c^{cs} = \sum_{\forall a,h} [\epsilon^{s}_{a,h} C^{s} + \epsilon^{c}_{a,h} C^{c}] \tag{5.11}$$

All cost terms have been converted to monetary terms. Within a MOEA-based approach, some of the included parameters may be unnecessary based on the objective functions defined and have been introduced mainly for consistency.

Based on the simplifications/assumptions made, a distinction can be drawn between the aforementioned cost terms. Specifically, c^{inv}, c^{fom} and c^{gp} can be computed considering only investment decisions (i.e. $\mathbf{x}^{inv}$) and neglecting short-term operation. On the contrary, cost terms c^{gen}, c^{em}, c^{vom}, c^{rp}, c^{cs} and c^{nse} could be computed based on an approximation of the short-term operation of a power system.

Let $\mathbf{v} = \{\mathbf{p},\ \mathbf{x}^{su}, \mathbf{x}^{sd}, \epsilon^{r}, \epsilon^{d}, \epsilon^{c}, \epsilon^{s}\}$ be a vector that includes all values required to compute cost terms c^{gen}, c^{em}, c^{vom}, c^{rp}, c^{cs} and c^{nse}. Then the values of $\mathbf{v}$ could be determined by a SM that considers the capacity additions fixed. The outcome of the SM can be utilized to compute the corresponding cost terms. Within the formulation, vector $\mathbf{v}$ is not considered as a set of optimization variables of the MOO problem but rather it is assumed as the output of a black-box which represents the outcome of a simulation run. The vector $\mathbf{v}$ must be computed anew for each state of capacity additions and its values are then employed to compute the corresponding cost terms. The input of the black-box is assumed as the set of capacity additions. The output of the former is a single vector $\mathbf{v}$ which represents the outcome of a simulation run.

There are some prerequisites that must be considered. First, it must be ensured that the SM provides a solution for each state of capacity additions. Therefore, a SOO SM needs to be formulated adequately to penalize infeasibility when no feasible solution exists for a given set of capacity additions. Secondly, a state of capacity additions can correspond only to a single output of the SM. If a selected SM is a SOO problem that does not exhibit a unique global optimal solution then a one-to-one relationship between capacity additions and each cost term can not be established when more than one optimal solutions of the SOO SM are considered. In many cases, non-unique global optimal solutions of a SOO SM could be restricted by its formulation or the data input. Therefore, an assumption is made. In particular, a single simulation run is carried out for each set of capacity additions and it is assumed that the outcome derived is the representative one for that state of capacity additions. Moreover, the simulations correspond to the computationally expensive part of the approach. Therefore, performing a single simulation may be acceptable as from each state of capacity additions an indicator/approximation of the short-term operation is required. In the worst case, in which the SOO problems for all states of capacity additions exhibit non-unique global optimal solutions, the indicators are based on a set of non-unique global optimal solutions attained by each SOO SM.

5.2.3 Simulation model - Problem approximation

The modelling formulation of the Fast Unit Commitment model, proposed in Reference [29], is selected as the SM for approximating a UCP. An advantage of this model is the absence of binary/integer variables within the formulation. The latter accounts for technical detail by including a relaxation of the integer constraints of the units status and modifications to the UCP constraints. In comparison to other SMs with higher computational cost, it may provide an approximation of the UCP with a relatively low computational cost. This may enable an increase in the temporal and spatial detail and also an increase in the maximum accepted number of available simulations. A representation based on continuous variables to approximate the unit statuses of the UCP within the context of long-term planning has been presented in other works (e.g. Reference [48]).

The selected SM is adopted with some modifications. The modifications are made to consider the cases where the installed capacity is inadequate to serve the demand, i.e to address infeasible installed capacities (short-term constraint violations). In addition, the cost of emissions and constraints for hydro generation are included. Moreover, a simplified fuel cost function is assumed; the minimum fuel cost of the on-line capacity is not captured, however, it can be easily included. The initial state has been assumed as an optimization variable. This assumption is made to replace a user-defined assumption regarding the initial state. Nevertheless, the impact of the latter on the results should be moderated as the number of hourly time periods increases. The formulation of the SM, including the modifications made, is presented in Appendix E and for greater detail the reader is kindly referred to Reference [29].

5.2.4 Formulation of the Multi-objective optimization problems

Let vector $\mathbf{x}$ represent the set of optimization variables converted from the matrix $\mathbf{x}^{inv}$. In particular, the matrix $\mathbf{x}^{inv}$ (representing the capacity additions in each technology group and area) is converted to the vector $\mathbf{x}$ which includes all elements of the former:

$$\mathbf{x} = \{x^{inv}_{1,1}, x^{inv}_{1,2}, \ldots, x^{inv}_{1,gz}, x^{inv}_{2,1}, x^{inv}_{2,2}, \ldots, x^{inv}_{2,gz}, x^{inv}_{az,1}, x^{inv}_{az,2}, \ldots, x^{inv}_{az,gz}\} \tag{5.12}$$

The search space ($\mathbb{S}$) is set as follows:

$$\underline{X}_{a,g} \leq x^{inv}_{a,g} \leq \overline{X}_{a,g}, \forall a, g \tag{5.13}$$

$$\mathbf{x} \in \mathbb{Z} \tag{5.14}$$

where $\underline{\mathbf{X}}$ and $\overline{\mathbf{X}}$ are the lower and upper limits of capacity additions.

Objective functions: Based on Sections 5.2.2 and 5.2.3, the five pairs of objective functions, which constitute the MOO GEP variants examined, are the following:

1. Aggregated Cost and Emission Cost (AC1-EM) variant:

   - Minimizing the sum of investment, FO&M, generation, VO&M, GPSC and curtailment cost:
   $$l_1(\mathbf{x}) = c^{inv} + c^{fom} + c^{gen} + c^{vom} + c^{gp} + c^{cs} \tag{5.15}$$
   - Minimizing emission cost:
   $$l_2(\mathbf{x}) = c^{em} \tag{5.16}$$

   The results of this variant aim in an approximation set of the PF considering economic and environmental objectives. It corresponds to a variant of the frequently examined economical and environmental MOO GEP problem. The main difference can be identified in the way cost terms are computed as is the impact of capacity additions on each cost term are assessed by the SM.

2. Aggregated Cost and RES penetration (AC2-RP) variant:

   - Minimizing the sum of investment, FO&M, generation, VO&M, emission, GPSC and curtailment cost:
   $$l_1(\mathbf{x}) = c^{inv} + c^{fom} + c^{gen} + c^{vom} + c^{em} + c^{gp} + c^{cs} \tag{5.17}$$

- Maximizing RES penetration:

$$l_2(\mathbf{x}) = (-1)c^{rp} \tag{5.18}$$

This second variant also considers an economic and environmental MOO GEP. However, the formulation focuses on examining RES penetration and curtailment. For example, emission cost reduction induced by other sources, such as generation by hydro or low-carbon emission thermal units, is not considered within the environmental objective. Therefore, the aim of the second objective is restricted towards efficiently increasing the level of RES penetration.

3. Operation Cost and GPSC (OC1-GS) variant:

- Minimizing the sum of FO&M, generation, VO&M, emission and curtailment cost:

$$l_1(\mathbf{x}) = c^{fom} + c^{gen} + c^{vom} + c^{em} + c^{cs} \tag{5.19}$$

- Minimizing GPSC:

$$l_2(\mathbf{x}) = c^{gp} \tag{5.20}$$

The results of this case aim in providing an approximation set of the Pareto front considering economical and green policy objectives. The results may assist in examining the trade-off among support cost of an energy policy and operation cost, since such costs could be transferred to the consumers.

4. Operation Cost and Investment Cost (OC2-IC) variant:

- Minimizing generation, FO&M, VO&M, emission, GPSC, and curtailment cost:

$$l_1(\mathbf{x}) = c^{gen} + c^{fom} + c^{vom} + c^{em} + c^{gp} + c^{cs} \tag{5.21}$$

- Minimizing investment cost:

$$l_2(\mathbf{x}) = c^{inv} \tag{5.22}$$

The results of this case aim in providing an approximation set of the Pareto front considering operational cost and investment cost as conflicting objectives. The results may assist in examining the effect of higher/lower investment cost and their impact on short-term operation. The aforementioned may indicate the importance of increased investment levels, when required, and suggest technology groups for which possible incentives should be provided.

5. Aggregated Cost and Aggregated Penalties (AC3-AP) variant:

- Minimizing investment, FO&M, generation, VO&M, emission and GPSC cost:

$$f_1(\mathbf{x}) = c^{inv} + c^{fom} + c^{gen} + c^{vom} + c^{em} + c^{gp} \tag{5.23}$$

- Minimizing a NSE and curtailment cost terms:

$$f_2(\mathbf{x}) = c^{nse} + c^{cs} \tag{5.24}$$

The last variant is formulated to examine the trade off between aggregated cost and the aggregated penalty cost terms. This case could be also perceived as a reformulated SOO GEP for which the penalty cost deriving from short-term operation has been set as a second objective.

Depending on the data input, the pairs of examined objectives could not always be conflicting. Moreover, NSE cost is added to the objectives as a penalty cost term in an attempt to exclude solutions exhibiting NSE from the NDF. For the first four MOO GEP variants it is included as follows:

$$f_1(\mathbf{x}) = l_1(\mathbf{x}) + w^{pen} \cdot pen(\mathbf{x}) \tag{5.25}$$

$$f_2(\mathbf{x}) = l_2(\mathbf{x}) + w^{pen} \cdot pen(\mathbf{x}) \tag{5.26}$$

where $pen(\mathbf{x}) = c^{nse}$ and $w^{pen} = 1$. The latter is a weight parameter for scaling (or excluding $w^{pen} = 0$) the penalty term.

Constraint functions: In addition, the MOO problems are subjected to the following constraint functions ($\mathbf{G}(\mathbf{x})$):

- *Reserve margin constraint:* The constraint imposes a lower limit to the total installed capacity and it is set as the peak demand, PD, increased by a reserve margin, RM.

$$(1 + RM)PD \leq \sum_{\forall a,g} [(x_{a,g}^{inv} + IU_{a,g}) P_{a,g}^{cap\_step}] \tag{5.27}$$

  The feasible region could be further limited by including an upper limit on the total installed capacity.

- *Maximum capacity additions constraint:* It is assumed that the total capacity additions are limited by an upper limit, **TCA**, as follows:

$$\sum_{\forall a} [x_{a,g}^{inv}] \leq TCA_g, \forall g \tag{5.28}$$

The total constraint violation ($cv(\mathbf{x})$) for the decision vector $\mathbf{x}$ is computed as follows:

$$cv(\mathbf{x}) = \sum_{i=1}^{k} \frac{\max(g_i(\mathbf{x}), 0)}{CV_i^{max}} \tag{5.29}$$

where CV_i^{max} is the maximum possible constraint violation of the i$^{\text{th}}$ constraint function ($g_i(\mathbf{x})$). It is included to normalize the constraint violation values within the range [0,1]. The maximum values are made available a priori by setting $x_{a,g}^{inv} = 0, \forall a, g$ and $x_{a,g}^{inv} = \overline{X}_{a,g}, \forall a, g$ for Eq. 5.27 and Eq. 5.28, respectively.

5.3 Multi-objective optimization approach

This Section presents the employed MOO approach. It is based on MOEA and surrogate-assisted derivative-free optimization. The NSGA-III algorithm (References [102, 103]) has been used as the base MOEA that is modified to by including AMs. These modifications are made based on proposed frameworks for surrogate-assisted derivative-free optimization identified in the literature (References [155, 168, 187, 188, 189, 186, 204, 205]).

To determine the steps of the approach and based on the aforementioned references, the following have been considered as important for attaining satisfactory results:

- To generate an appropriate pool of data points (PDP): Data points to be costly evaluated are selected among a pool of candidate/available ones. Therefore, the pool should include data points that are of interest for the search.

- To achieve a decent approximation of the underlying functions (UFA): The refining strategies should select the data points, to be costly evaluated, that could be most beneficial for the search, w.r.t. the exploration and exploitation of the search space. A poor approximation may be a result of low quality of data sampling or the properties of a selected AM. Moreover, a poor approximation may be attained due to the function characteristics which could render a decent UFA highly challenging. In this case the underlying functions are the output of an optimization problem (i.e. the values of the objective functions computed based on the output of the SM).

- To achieve a decent Pareto Front Approximation (PFA): Given a *perfect* UFA and an *ideal* PDP, the refining strategies should select data points to be costly evaluated that would result in a well-converged and well-distributed NDF. The latter should be a decent approximation of the PF.

For example, refining strategies that utilize the prediction function may prioritize sub-optimal data points in case the UFA is sufficiently poor to alter their ranking and if the PDP includes both *ideal* data points and sub-optimal ones. Similarly, sub-optimal data points could be selected if the *ideal* ones are not available (regardless of the UFA accuracy). Moreover, the refining strategies should be adequate to select a subset of the available data points to provide a decent PFA given an *ideal* PDP and a *perfect* UFA due to the computational restrictions.

The presented approach includes a global and local phase to benefit from (i) global and local AM, and (ii) global and local search. There is a number of approaches that employs a global and local phase to balance exploration and exploitation, global and local AM (e.g. References [155, 168, 188, 189, 204, 205]) and global and local gradient-based (e.g. References [155, 168, 204]) or non-gradient-based (e.g. Reference [206]) search. These have been discussed in more detail in Chapter 3.

For selecting sampling criteria, the frameworks of References [188, 189] are followed. In particular, the Maximum-Minimum Distance Criterion (MMDC) and a Hypervolume based metric are used within the refining strategies. Both criteria are included within the multi-rule based strategy of the GOMORS algorithm (Reference [188]) and within the MOPLS algorithm (Reference [189]) by using a probability factor. The refining strategies of the global phase are based on the multi-rule strategy of the GOMORS algorithm where a number of candidate solutions are selected by each rule. However, a sequence of steps is included to determine the rule which will be employed to select a candidate solution for each AM included within the pool of available ones. In addition, a CHT is included to address infeasible solutions. The CHT selected for prioritizing feasibility within the steps of the refining strategy is the Feasibility Rules (Reference [119]).

5.3.1 Steps of the optimization approach

Given the above, the optimization approach includes three distinct phases (i) the initialization phase, (ii) the global phase and (iii) the local phase. The last two are iteratively repeated. From here on, such an iteration will be referred to as an optimization cycle. Within the latter, a PDP is generated and selected refining strategies determine new data points to be costly evaluated in an attempt to gradually identify a NDF. These three phases are presented in the following Sections. The components and motivations of each phase are first discussed and then presented in a step-by-step manner.

Initialization Outline

Within the initialization phase the required data and the parameters are supplied. Moreover, the initial dataset (archive) is constructed. Two criteria are included for the latter. The first regards the user-supplied decimal to which all data points are rounded to. The second is the minimum distance criterion, based on the Euclidean distance, which restricts data points to enter the dataset that violate a minimum distance to any archived data point (similarly to Reference [180]).

Moreover, additional data points could be supplied based on available a priori information. Such data points may improve the approximation in regions of interest. Some initial estimates could be identified based on the considered cost terms (Section 5.2.2). For example, (i) a data point representing the minimum capacity additions could minimize the investment cost and FO&M terms, (ii) a data point representing the maximum capacity additions could present low NSE cost, and (iii) a data point representing the maximum capacity additions for thermal technology groups and the minimum for hydro and RES could present low curtailment cost and low GPSC. Alternately, initial points could be supplied by a different model if available.

Initialization step-by-step

The steps for the initialization are the following:

1. Provide the input data and parameters.
2. Generate a pool of data points by the Latin Hypercube sampling within the bound of the search space.
3. Round the data points values to a predefined decimal based on user-defined parameter dv^{Tol} (for integers the parameter is set as $dv^{Tol} = 1$) using the following equation:

$$x_i = \lfloor x_i dv^{Tol} \rceil / dv^{Tol}, \forall i = 1, 2, ..., D \tag{5.30}$$

4. Remove data points that violate minimum distance criterion:
   (a) Compute the Euclidean distances among all data points.
   (b) Remove any data point that violates a minimum distance threshold value, $dist^{min}$.
5. If the pool includes at least np^{init} data points and these ensure the minimum required rank for the matrix used to build each AM considered, then skip this step. Otherwise, repeat Steps 2-5.
6. If additional (user-supplied) data points are provided then:
   (a) Compute the Euclidean distance between the user-supplied data point and the data points of the pool.
   (b) Remove any user-supplied data point that violates the minimum distance criterion.

(c) Include the remaining ones to the pool.

7. Evaluate all data points in the pool using the expensive model.
8. Store the data points and the corresponding objective and constraint function values in the dataset.
9. Perform Environmental-Selection among the archived data points to determine the current population.
10. Set the union and evolving populations as empty sets.

Subsequently, the algorithm proceeds to the iterative phase. The latter includes the global phase and local phase which are discussed in the following Sections.

Optimization Cycle - Outline of the Global phase

The available AMs are built independently using all data points within the dataset; a maximum threshold might be required, when a relatively large number of costly evaluations are available. This threshold even though not considered in this Chapter, could be based on (i) a distance criterion, (ii) the most recent, or (iii) the best performing entries of the dataset.

Moreover, a procedure is employed that attempts to moderate the impact of outliers in the dataset, i.e. archived data points that exhibit extreme function values. The impact of extremely high values has been considered here. Simply removing such outliers may result in loss of costly information attained by a computationally expensive simulation. The procedure included generates a modified dataset which is used for building the global models. A modified dataset is created and the procedure is implemented independently for each objective function. Function values, in the modified dataset, that exceed an upper limit are set to the largest value that does not exceed this limit. The upper limit is set as the median and three standard deviations of the function values in the modified dataset. This is repeated until (i) no alterations are required or (ii) the largest value for the considered objective function in the modified dataset, which is lower than the upper limit, is equal to the minimum value of the modified dataset of the considered objective function. The median value is selected as it ensures that at least half of the dataset will not be affected. A median-based rule has been suggested in Reference [180]. Three standard deviation have been selected to limit the alterations mostly to outliers. The modified values are then used to build each global model.

The next step of the global phase is to generate an appropriate PDP. First, a Union population is generated by evolving sub-populations separately for a predefined number of generations utilizing a MOEA. When evolving the sub-populations, the AM replaces the TM for computing the objective function values. However, constraint functions are computed based on the true functions as they are considered computationally cheap. Moreover, individuals are rounded (see Eq. 5.30) only for computing their fitness.

The number of sub-populations that will evolve are determined by the number of available AMs as one is selected for each sub-population among the pool of available models. Moreover, a set of Variations operator is selected to evolve each sub-population from the pool of available ones. The latter aims towards reducing the impact of a selected one on the performance of the optimization approach. In this context, the combination of AM and Variations operator should be altered in each optimization cycle (if possible) to provide a diverse set of combinations.

The initial population of each sub-population is determined by applying the Environmental-Selection operator on a population including (if available): (i) the archived data points, (ii) the evolving population derived during the global phase of the previous optimization cycle, (iii) data points generated during the local phase of previous optimization cycle, and (iv) a population generated anew by Latin Hypercube sampling. The evolving population and data points generated during the local phase are included since in each optimization cycle only a subset of the available data points are archived. Therefore, some promising individuals identified during the previous optimization cycle are available to be costly evaluated in the proceeding one. Such individuals could survive in the evolving population for a number of optimization cycles depending on the Environmental-Selection operator employed.

The evolving population is formed by merging the attained sub-populations once the termination criteria have been met. A new set of data points is generated by Latin Hypercube sampling to ensure sufficient diversity. The evolving population is combined with the latter to form the Union population. The individuals of the latter are rounded (see Eq. 5.30). Moreover, data points within the Union population, that have been costly evaluated (or violate the minimum distance criterion to the archived data points) are removed. The Union population is then checked to ensure that at least one new data point is available. If this is not the case, the Latin Hypercube sampling is employed to a generated a new set of data points which are checked in comparison to the archived points regarding the distance criterion. The latter is repeated until at least one new data point is available (and increasing also the number of data points generated by Latin Hypercube sampling). The Union population then represents the PDP from which the refining strategies will select a subset to be costly evaluated using the TM.

The refining strategy attempts to select the most adequate data point from the PDP to be costly evaluated. Both exploration of the search space and exploitation of the AM are considered. The Expected Hypervolume Improvement (EHI) metric is included for exploitation as it may identify a data point that would improve the Hypervolume (HV) of the current NDF. It is implemented under the assumption of a sufficiently accurate UFA. The MMDC is included to assist towards the exploration of the search space by selecting data points that are in low density areas of the search space (decision or objective space).

An attempt to improve the exploration of the search space is made by including the MMDC. It could introduce new data point to the dataset that lie in unexplored regions of the objective or decision space (Reference [188]). The MMDC selects a data point, among available ones, that exhibits the maximum minimum Euclidean distance towards a second set of data points. More specifically, the minimum Euclidean distance for the i$^{\text{th}}$ candidate data point is computed as follows:

$$dist_i^{min} = \min_{\forall s}(\|\mathbf{y}_i - \mathbf{y}_s\|) \tag{5.31}$$

where if the minimum Euclidean distance is computed for the decision space, $\mathbf{y}_i$ and $\mathbf{y}_s$ are the decision vectors of the i$^{\text{th}}$ candidate data point and of the s$^{\text{th}}$ archived data point, respectively. Correspondingly, if the objective space is considered then $\mathbf{y}_i$ and $\mathbf{y}_s$ are the objective function vectors of the i$^{\text{th}}$ candidate data point and of the s$^{\text{th}}$ archived data point, respectively.

A focus on exploitation is made by the inclusion of the Hypervolume metric. It is the volume of the objective space dominated by an objective function vector and bounded by a reference vector. It is a quality indicator frequently employed to compare the performance of different MOEAs as it is strictly monotonic with regard to Pareto dominance. An estimation of the Hypervolume value can be provide by Monte Carlo simulation to address the computational cost of its exact evaluation that scales with the number of objective functions.

Therefore, the maximum EHI metric is included to identify a data point that would provide the largest improvement to the hypervolume of the current population. It is termed *expected* as it relies on the AM predictions and assumes a sufficiently accurate UFA. The EHI of the i$^{\text{th}}$ individual, for a set of non-dominated points $\mathbf{Y}$, is computed as follows:

$$EHI_i = HV(\mathbf{Y}'') - HV(\mathbf{Y}) \tag{5.32}$$

where $\mathbf{Y}''$ is a set of non-dominated points. It is a subset of a set $\mathbf{Y}'$ where $\mathbf{Y}' = \mathbf{y}_i \cup \mathbf{Y}$. The elements of $\mathbf{Y}''$ are selected from $\mathbf{Y}'$ such that: (i) $\mathbf{Y}''$ includes only non-dominated points, and (ii) the number of elements within set $\mathbf{Y}''$ does not exceed a maximum limit imposed. For example, assuming a population of NP individuals the EHI is computed to respect the maximum population size by removing elements of $\mathbf{Y}'$ if required. This is implemented to ensure that the computational cost does not increase due to the large number of data points. Moreover, it is implemented as in some cases the number of data points that should be supplied for decision-making may be restricted.

The EHI metric is computed to select a data point from the PDP for each AM. For all comparisons made within the refining strategies the objective functions' predictions are employed for both the archived and candidate data points. This is implemented as the AM might preserve the ranking of the solution despite the error in the prediction (Reference [167]). For the constraint functions the true function is used.

Moreover, the EHI metric is computed only in specific cases to limit the computational cost of the metric. In particular, the series of steps first prioritizes feasibility. Given a set of feasible data points, then diversity is considered only if a positive value of the EHI metric is not anticipated i.e. there are no data points in the Union population that dominate or are non-dominated to the NDF points of the Archived population. In this case, the selection is made from the pool of individuals that could enter the current population by using the MMDC. In the remaining cases, the EHI metric is computed for individuals that are at least non-dominated to the individuals of the Archived population and, therefore, could provide a positive value of the aforementioned metric.

More specifically, POP^{b_u}, POP^{b_a} and POP^{b_c} are determined by the Environmental-Selection operator. POP^{b_u} and POP^{b_a} represent the *best* populations derived from the Union population and Archived population, respectively. POP^{b_c} represents the data points from POP^{u} that could enter the current population if the predictions are accurate. The population size of POP^{b_c}, POP^{b_u} and POP^{b_a} is determined based on the minimum among the number of available individuals and the user-defined population size. Then, feasibility is prioritized by checking if there is a feasible data point in POP^{b_u}. When all individuals in POP^{b_u} are infeasible then the selection is made based on the minimum constraint violation. Moreover, the EHI metric is not computed if no new data point would enter the current population based on the AM predictions, i.e. POP^{b_c} is an empty set. In this case, feasibility is prioritized by removing infeasible individuals from POP^{b_u} and an individual is selected from POP^{b_u} based on the MMDC (objective space) considering POP^{a}. This is also the case when no feasible solution has entered the dataset so far. In the remaining cases, POP^{b_c} includes feasible solutions that could enter the current population, based on the predictions of the considered AM. Therefore, the feasible non-dominated individuals of POP^{b_a}, POP^{b_c} and $POP^{b\_a} \cup POP^{b\_c}$ are identified (NDF^{a}, NDF^{b_c} and $NDF^{b\_c\&a}$, respectively). Archived data points are removed from $NDF^{b\_c\&a}$ which then represents the candidate data points that are at least non-dominated to the individuals of POP^{b_a}. The EHI metric is not computed if $NDF^{b\_c\&a}$ is an empty set as a HV improvement is not possible based on the predictions of the considered

AM. The selection is then made from POP^{b_c} based on the MMDC (objective space) w.r.t. the POP^{b_a} to increase the diversity of the best archive population.

Consequently, the EHI metric is only computed for each individual of $NDF^{b\_c\&a}$. It is iteratively computed by including an individual of $NDF^{b\_c\&a}$ in NDF^{a}. Then the Environmental-Selection operator is applied on the derived combined population to ensure that it does not exceed the predefined population size limit. The NDF of the aforementioned population is then identified for which the new HV value is computed. The EHI metric for an individual is computed by subtracting the initial HV value (for NDF^{a}) from the value attained by its inclusion. Once the values of the EHI metric have been attained independently for each individual of $NDF^{b\_c\&a}$, then the one exhibiting the maximum value is selected. In case the EHI metric maximum value is zero then the selection is made from POP^{b_c} based on the MMDC (objective space) w.r.t. the POP^{b_a}.

This selection process is repeated for each available AM to select a data point for each AM. In each such iteration, a modified Archived population is utilized that includes the Archived population and the selected data points which have not been costly evaluated. In addition, the selected point is removed from the Union population so that it is excluded from the next selection process. In a similar manner, all individuals of the Union population that violate the minimum distance criterion to selected data points are also excluded.

Once a data point has been selected based on each AM, a data point is then selected based on the MMDC (decision space). Such a candidate is selected to assist towards the exploration of the search space as it may select a data point that is in a low density area of the search space (decision or objective space). The data point is selected from the Union population and the MMDC is computed w.r.t. the modified archived data points.

The global phase concludes by computing the selected data points using the true model and storing the relevant data in the archive. The termination criteria are then considered.

Optimization Cycle - The Global phase step-by-step

The steps of the global phase are presented in this Section. Four major steps comprise the global phase which are the following:

1. Build each global AM:
   (a) Create a copy, i.e. $\mathbf{F}'$, of the function values $\mathbf{F}$ which are available in the dataset.
   (b) For each of the M objective functions:
      i. Set $\overline{f}$ equal to the median and 3 standard deviations of the function values of the m^{th} objective function of $\mathbf{F}'$.
      ii. Set $\underline{f}$ equal to maximum function value, which also is lower than $\overline{f}$, of the m^{th} objective function of $\mathbf{F}'$.
      iii. If the maximum function value of the m^{th} objective function of $\mathbf{F}'$ is greater than $\overline{f}$ and $\underline{f}$ is greater than the minimum function value of the m^{th} objective function of $\mathbf{F}'$ then:
         A. Update $f'_{m,s}$ as follows: $f'_{m,s} = \min(f'_{m,s}, \underline{f}), \forall s$
         B. Repeat step 1(b)i-1(b)iii.
   (c) Use $\mathbf{F}'$ to build the AMs for each objective function m and for each available model k.

2. Generate a PDP:

   (a) For each of the K available AMs:

      i. Select the v^{th} set of Variation operators from the V available ones, where $v = \langle k + cntr \rangle_V + 1$ and $cntr$ is the optimization cycle counter.

      ii. Generate a population, POP^{lhs}, by Latin Hypercube sampling.

      iii. Create a combined population: $POP^{all} = POP^{lhs} \cup POP^{evl} \cup POP^{a} \cup POP^{ls}$

      iv. Employ the AMs to compute the objective functions predictions of POP^{all}.

      v. Employ the TM to compute the constraint violation of POP^{all}.

      vi. Determine the parent population POP^{par} by applying the Environmental-Selection operator on a combined population $POP^{lhs} \cup POP^{evl} \cup POP^{a} \cup POP^{ls}$.

      vii. Apply the v^{th} set of Variation operators to create an offspring population (POP^{off}).

      viii. Employ the AMs to compute the objective functions predictions of POP^{off}.

      ix. Employ the TM to compute the constraint violation of POP^{off}.

      x. Apply the Environmental-Selection operator on $POP^{off} \cup POP^{par}$ to update POP^{par}.

      xi. Repeat steps 2(a)vii-2(a)xi if the termination condition has not been met.

      xii. Set POP^{par} as the k^{th} sub-population (POP_k^{sp}).

   (b) Update the evolving population: $POP^{evl} = POP_1^{sp} \cup POP_2^{sp} ... \cup POP_K^{sp}$

   (c) Create a population by Latin Hypercube sampling (POP^{lhs}).

   (d) Create the Union population: $POP^{u} = POP^{evl} \cup POP^{lhs}$

   (e) Update POP^{u} by applying the rounding function (Eq. 5.30) to each individual.

   (f) Iteratively check and remove individuals from POP^{u} so that the remaining ones are unique, i.e. the minimum Euclidean distance among two individuals of POP^{u} should be greater than zero.

   (g) Remove any individuals from POP^{u} that violates the minimum distance criterion towards the Archived population.

   (h) If POP^{u} is not an empty set, the skip this step, otherwise:

      i. Replace POP^{u} by a population generated by Latin Hypercube sampling. Larger population sizes could be considered for this step.

      ii. Repeat steps 2e-2h.

3. Refine the dataset:

   (a) Set POP^{sel} as empty.

   (b) For each of the K available AMs:

      i. If POP^{u} is not empty, then skip this step. Otherwise:

         A. Go to step 4.

      ii. Employ the AMs to compute the objective functions predictions for both POP^{u} and $POP^{a} \cup POP^{sel}$.

      iii. Employ the TM to compute the constraint violation for both POP^{u} and $POP^{a} \cup POP^{sel}$.

iv. Apply the Environmental-Selection operator on population POP^u to derive POP^{b_u}.

v. Apply the Environmental-Selection operator on population $POP^a \cup POP^{sel}$ to derive POP^{b_a}.

vi. Apply the Environmental-Selection operator on a combined population $POP^c \cup POP^u$ to derive POP^{b_a}.

vii. Remove any individual from POP^{b_c} that is a member of $POP^a \cup POP^{sel}$.

viii. If POP^{b_u} includes a feasible individual, then skip this step. Otherwise:

   A. Select $\mathbf{x}^{sel}$ from POP^{b_u} based on the minimum constraint violation.

   B. Go to step 3(b)xix.

ix. If POP^{b_c} is not empty and $POP^a \cup POP^{sel}$ includes a feasible individual, then skip this step. Otherwise:

   A. Remove infeasible solutions from POP^{b_u}.

   B. Normalize each objective function value of POP^{b_u} and $POP^a \cup POP^{sel}$ based on the corresponding maximum and minimum values of each function within these populations.

   C. Select $\mathbf{x}^{sel}$ from POP^{b_u} based on the MMDC (objective space) w.r.t. $POP^a \cup POP^{sel}$.

   D. Go to step 3(b)xix.

x. Perform Non-dominating sorting on POP^{b_a} to identify NDF^a.

xi. Perform Non-dominating sorting on POP^{b_c} to identify NDF^{b_c}.

xii. Perform Non-dominating sorting on a combined population including the individuals within NDF^a and NDF^{b_c} to generate the NDF of the combined populations ($NDF^{b\_c\&a}$).

xiii. Remove any archived data points from $NDF^{b\_c\&a}$. The latter represents the candidate points that are at least non-dominated to the individuals of POP^{b_a}.

xiv. If $NDF^{b\_c\&a}$ is not empty, then skip this step. Otherwise:

   A. Set the EHI improvement to zero: $EHI = 0$.

   B. Go to step 3(b)xviiiA.

xv. Compute the HV^{old} for NDF^a as follows:

   A. Normalize and replace each objective function values of the individuals in $NDF^{b\_c\&a}$ and NDF^a based on the corresponding maximum and minimum values of each function within these populations.

   B. Set a reference point at (1.1,1.1).

   C. Compute the HV^{old} of NDF^a.

xvi. For each individuals in $NDF^{b\_c\&a}$ compute the EHI as follows:

   A. Include the selected individual of $NDF^{b\_c\&a}$ in NDF^a and form a temporary population.

   B. Apply the Environmental-Selection operator to the aforementioned temporary population to maintain a population of up to NP individuals.

   C. Perform Non-dominated sorting (NDS) on the aforementioned population to identify the NDF.

   D. Compute the HV^{new} for the individuals forming the aforementioned NDF.

E. Compute the EHI by subtracting HV^{old} from HV^{new}.

xvii. If the maximum attained EHI is not greater than zero, then skip this step. Otherwise:

A. Select $\mathbf{x}^{sel}$ from $NDF^{b\_c\&a}$ based on the maximum EHI value.

xviii. If the maximum attained EHI is greater than zero, then skip this step. Otherwise:

A. Select $\mathbf{x}^{sel}$ from POP^{b_c} based on the MMDC (objective space) w.r.t. POP^{b_a}.

xix. Include $\mathbf{x}^{sel}$ in the selected population to be evaluated (POP^{sel}).

xx. Remove individuals from POP^{u} that violate minimum distance criterion to $\mathbf{x}^{sel}$.

(c) If POP^{u} is empty, then skip this step. Otherwise:

i. Apply to all individuals of POP^{u}:

$$cv'_i = \begin{cases} cv_i, & \text{if } U(0,1) \leq pf^{MMDC} \\ 0, & \text{otherwise} \end{cases}, \forall i \tag{5.33}$$

ii. If the minimum value of $\mathbf{cv}'$ is not greater than zero, then, select $\mathbf{x}^{sel}$ from POP^{u} based on the MMDC (decision space) w.r.t. $POP^{a} \cup POP^{sel}$. Otherwise:

A. Remove all individuals from POP^{u} for which cv'_i is greater than zero.

B. Select $\mathbf{x}^{sel}$ from POP^{u} based on the MMDC (decision space) w.r.t. $POP^{a} \cup POP^{sel}$.

iii. Include $\mathbf{x}^{sel}$ in POP^{sel}.

4. Perform expensive function evaluations:

(a) If POP^{sel} is empty then, skip this step. Otherwise:

i. Perform expensive function evaluations for each individual of POP^{sel} and store the data points in the archive. The termination criteria are then considered.

Optimization Cycle - Outline of the Local phase

Locally trained models are built for selected individuals of the current NDF. The MMDC (objective space) is employed to select up to μ isolated NDF members. It is computed by normalizing the function values of the NDF members w.r.t. the current population. The user defined parameter μ is included to limit the maximum number of NDF members that will undergo local search. Such individuals are included in a population that will undergo local search, i.e. POP^{lc}.

Moreover, an additional point is included within POP^{lc}. Two adjacent NDF members that exhibit the largest Euclidean distance (objective space) among them are identified. Then, an additional point is generated along the line (decision space) connecting these two individuals. More specifically, a weighted average of these individuals is produced. The weight is randomly generated based on the uniform distribution. This is implemented under the assumption that the mapping from the decision space to the objective space is not highly non-linear. In such a case, the offspring produced could lie among the two individuals used to produce the offspring and applying local search on the aforementioned could generate a promising data point. On the other hand, if the mapping is highly non-linear then the offspring produced by the aforementioned step is a rather arbitrary point that will undergo local search.

A local AM is built independently for each individual of POP^{lc}. A model is assigned from the pool of available models to each individual of POP^{lc}. The pool consists the available models and an additional model that is the weighted sum of the predictions of the models (ensemble). For building each model, the nearest N^{lc_dp} data points (Euclidean Distance) to the considered individual are selected as in other approaches including local models (e.g. References [155, 168]). Additional data points are included when the rank of the matrix created, using the selected data points, is inadequate to built the AM. In this case, a sub-region of the search space is defined and all data points within the bounds of the region are included. This is implemented based on a parameter, RI, which is a percentage of the search space. Its value increases in user-defined steps until the condition is met, e.g. $ri = 10\%$. The parameters of the trained model are checked, as the data points may be inadequate to train the models, and additional points are included if needed by increasing the value of RI. Moreover, the AMs are built using the augmented function which is the weighted sum of the normalized function values; the function values are normalized based on the maximum and minimum corresponding values of the selected data points used to built the models.

For each such individual of POP^{lc}, local search is implemented using the locally trained models and a gradient-based non-linear solver. This is similar to the proposed local search in Reference [168] for SOO and in Reference [155] for MOO. When a MOO is considered, an augmented function is defined which is a weighted sum of the objective functions including arbitrary weights that are generated anew for each considered individual (Reference [148]). The latter is a typical approach for MA-MOEA. For surrogate-assisted MOO, the framework proposed in Reference [155] is considered. Local search is implemented based on Sequential Quadratic Programming. The non-lineal optimization problem neglects decision variables constraints regarding the step sizes (dv^{Tol}), i.e. the relaxed problem is solved even if decision variables are restricted to integers and the latter are assumed to be continuous. The locally built AM are utilized for computing the objective function value. However, the TM is employed for the constraint function since the aforementioned are computationally cheap. Moreover, the individual undergoing local search is used as the initial point for local search. The box constraints are set as the region defined by the maximum and minimum values of the data points used to built the local model. A minimum upper and lower range for each variable of the initial point is also considered, i.e. the variable step size and the minimum percentage (parameter RI) of the search space, both, restricted by the actual box-constraints.

The output of the local search ($\mathbf{x}^{lc}$) should be at least non-dominated to the individual to which local search was applied to. However, local search is implemented by ignoring the decision variable constraints regarding the step sizes. Therefore, $\mathbf{x}^{lc}$ is further processed by rounding it to the required predefined decimal. N^{lc_sr} offspring are generated by stochastically rounding $\mathbf{x}^{lc}$ and one offspring is generated by deterministically rounding $\mathbf{x}^{lc}$. For stochastically rounding $\mathbf{x}^{lc}$ the modulus, w.r.t. the decimal tolerance, is used as the probability. These are implemented, in an attempt to generated a number of offspring within the vicinity of $\mathbf{x}^{lc}$. Among these offspring the selected, as the output of the local search, offspring is determined based on the minimum prediction and minimum distance criterion and by prioritizing feasible offspring. The selected offspring then enters a population (POP^{ls}) that includes the offspring generated by local search.

Once this has been repeated for all individual in POP^{lc}, a PDP is generated by local search (POP^{ls}). Among this PDP, up to μ are selected to be computed using the TM and then stored in the archive. The selection is based on the MMDC (decision space). Furthermore, feasible solutions are prioritized.

Costly evaluating the selected individuals and storing them in the archive is the final step of the local phase. This is also the final step of the optimization cycle. Subsequently, it proceeds to the next optimization cycle if the termination criteria has not been met.

Optimization Cycle - The Local phase step-by-step

The local phase aims towards identifying improved data points within the vicinity of the best performing individuals of the archive population. This phase employs locally trained AMs for each considered individual and a gradient-based local search. The steps of the local phase are the following:

1. Apply the Environmental-Selection operator on POP^a to determine the current population (POP^{cur}). The true function values are used that are available in the dataset.

2. Apply NDS on POP^{cur} to identify the individuals forming the NDF (NDF^{cur}).

3. Set NDF^{cur} as the population that will undergo local search (POP^{lc}).

4. If there are less than μ individuals in POP^{lc}, then skip this step. Otherwise:
   (a) Normalize the objective function values of the NDF^{cur} and POP^a based on the corresponding maximum and minimum function values of the POP^a.
   (b) Select the μ individuals from POP^{lc} (objective space) that exhibit the largest minimum distance towards the Archived population.
   (c) Set POP^{lc} as the selected μ individuals.

5. If NDF^{cur} includes only one individual, then skip this step. Otherwise:
   (a) Normalize the objective function values of NDF^{cur} and POP^a based on the corresponding maximum and minimum function values of POP^a.
   (b) For each objective function m, sort the individuals from NDF^{cur} based on their normalized objective function value and identify the pair of adjacent individuals that exhibits the maximum Euclidean distance (objective space) to each other.
   (c) Identify and extract ($\mathbf{x}^a$ and $\mathbf{x}^b$) the single pair of succeeding individuals that exhibit the maximum to each other Euclidean distance in the objective space.
   (d) Generate an offspring, $\mathbf{x}^{ab}$, which is located (decision space) along the line connecting the selected pair of individuals:

$$x_j^{ab} = \kappa x_j^a + (1-\kappa)x_j^b, \forall j = 1, 2, ..., D \tag{5.34}$$

   where $\kappa = U(0,1)$ and $U(0,1)$ is a randomly generated number draw from the uniform distribution in the range [0,1].
   (e) Include offspring $\mathbf{x}^{ab}$ in POP^{lc}.

6. Associate an AM model to each individual of POP^{lc} as follows:
   (a) Generate a list of indices $K^{pool} = (1, 2, ..., K+1)$ where each index refers to one of the available AM and the last is the ensemble AM.

(b) Randomly shuffle the order of the indices in K^{pool}.

(c) For each of the i individuals in POP^{lc}, determine the index of the model that will be used (k_i) as follows:

$$k_i = 1 + \langle i \rangle_{K+1} \tag{5.35}$$

7. For each i individual of POP^{lc} build a locally trained AM:

   (a) Select the N^{lc_dp} data points that exhibit the minimum Euclidean Distance (decision space) from POP^a.

   (b) If the data points are sufficient to build the k_i model (or models for the ensemble case) then, skip this step. Otherwise:

      i. Set $RI = 0$. Parameter RI is a percentage of the search space which defines the bound of the sub-region for the i$^{\text{th}}$ individual of POP^{ls}.

      ii. Update parameter RI based on the user-defined parameter ri: $RI = RI + ri$.

      iii. Define the bounds of a sub-region based on parameter RI as follows:

$$\overline{x}_j^{sr} = \min(x_j^{(i)} + DX_j, \overline{x}_j), \forall j = 1, 2, ..., D \tag{5.36}$$

$$\underline{x}_j^{sr} = \max(x_j^{(i)} - DX_j, \underline{x}_j), \forall j = 1, 2, ..., D \tag{5.37}$$

      where $DX_j = max(RI(\overline{x}_j - \underline{x}_j), 1/dv^{Tol}), \forall j = 1, 2, ..., D$.

      iv. Include all data points within this sub-region until the condition is met.

      v. If the condition is met (sufficient data points to build the k_i AM) skip this step, otherwise repeat 7(b)ii-7(b)v by increasing the value of parameter RI in predefined steps until the condition is met.

   (c) Normalize each of the objective function values of the selected data points, for building the local model, based on their corresponding maximum and minimum function values ($\overline{f}_m^{lc}$ and $\underline{f}_m^{lc}$).

   (d) Randomly generate weight parameters (w_m) for each of the M objective functions such that $\sum_{m=1}^{M}[w_m] = 1$.

   (e) Build the k_i model from K^{pool} (or models for the ensemble case) for an aggregated function which is the weighted sum of the normalized objective functions: $f^{agg} = \sum_{m=1}^{M}[w_m(f_m - \underline{f}_m^{lc})/(\overline{f}_m^{lc} - \underline{f}_m^{lc} + \epsilon)]$.

8. For each i individual of POP^{lc} perform local search as follows:

   (a) Define the search region for the local search: It is restricted based on the variable step size, a minimum percentage (parameter ri) of the search space, the data points used to built the local model and the actual box-constraints. In particular, the local search region is defined as follows:

      i. The upper bounds for each variable are set as follows:

$$\overline{x}_j^{ls} = \max(UR_j^{min}, UR_j^{max}), \forall j = 1, 2, ..., D \tag{5.38}$$

where $\mathbf{UR}^{min}$ is the minimum upwards range for each variable, computed as follows:

$$UR_j^{min} = \min(x_j^{(i)} + 1/dv^{Tol}, \overline{x}_j), \forall j = 1, 2, ..., D \tag{5.39}$$

$\mathbf{UR}^{max}$ is the maximum upwards range for each variable. It considers the actual upper bounds of the variable, a maximum range based on a percentage of the search space (parameter ri) and the range of the data points used to build the local AM. It is computed for the j$^{\text{th}}$ decision variable as follows:

$$UR_j^{max} = \min(\overline{x}_j, x_j^{(i)} + RI(\overline{x}_j - \underline{x}_j), x_j^{(i)} + (1+sc)(\max(x_j^{(i)}, s_j^{(1)}, \ldots, s_j^{(ns_{ls})}) - x_j^{(i)}) \tag{5.40}$$

where sc is a user defined parameter (e.g. $sc = .1$) that scales the maximum difference of the decision variable value of the data points to the corresponding one of the individual. $s_j^{(n)}$ is the j$^{\text{th}}$ decision variable of the n$^{\text{th}}$ data point from the total ns_{ls} selected data points to build the local model of individual $x^{(i)}$.

ii. In a similar manner, the lower bounds for local search are computed as follows:

$$\underline{x}_j^{ls} = \min(DR_j^{min}, DR_j^{max}), \forall j = 1, 2, ..., D \tag{5.41}$$

where $\mathbf{DR}^{min}$ is the minimum downwards range for each variable, computed as follows:

$$DR_j^{min} = \max(x_j^{(i)} - 1/dv^{Tol}, \underline{x}_j), \forall j = 1, 2, ..., D \tag{5.42}$$

The maximum downwards range, DR_j^{max}, for the j$^{\text{th}}$ decision variable is computed as follows:

$$DR_j^{max} = \max(\underline{x}_j, x_j^{(i)} - RI(\overline{x}_j - \underline{x}_j), x_j^{(i)} - (1+sc)(x_j^{(i)} - \min(x_j^{(i)}, s_j^{(1)}, \ldots, s_j^{(ns_{ls})}) \tag{5.43}$$

(b) Apply local search for the relaxed SOO problem. The locally AMs are utilized for computing the objective function value and the TM is employed for the constraint function. Moreover, the individual undergoing local search is used as the initial point for local search within the region defined by $\underline{x}_j^{ls}$ and $\overline{x}_j^{ls}$.

(c) Select offspring:

i. Set $\mathbf{x}^{sel}$ as empty.

ii. Generate $\mathbf{x}^{lc\_d}$ by rounding $\mathbf{x}^{lc}$ to the required decimal (Eq. 5.30).

iii. Generate N^{lc_sr} offspring to form $\mathbf{x}^{lc\_s}$ by stochastically rounding $\mathbf{x}^{lc}$ to the required decimal. The i$^{\text{th}}$ offspring is generated as follows:

$$x_{i,j}^{lc\_s} = \begin{cases} \lfloor x_{i,j}^{lc_s} dv^{Tol} \rfloor / dv^{Tol}, & \text{if } dv^{Tol} \langle x_{i,j}^{lc_s} \rangle_{1/dv^{Tol}} \leq U[0,1] \\ \lceil x_{i,j}^{lc_s} dv^{Tol} \rceil / dv^{Tol}, & \text{otherwise} \end{cases}, \forall j = 1, 2, ..., D \tag{5.44}$$

iv. Compute the minimum distance of each offspring ($\mathbf{x}^{lc\_s}$ and $\mathbf{x}^{lc\_d}$) to POP^a.

v. Remove any offspring that violates the minimum distance criterion to POP^a.

vi. If all offspring have been removed, then skip this step. Otherwise: Compute the objective function predictions, based on the local AMs, and the constraint violation for the remaining offspring.

vii. If all offspring have been removed, then skip this step. Otherwise: If there is a feasible offspring, then set $\mathbf{x}^{sel}$ as the feasible offspring presenting the minimum prediction. Otherwise set $\mathbf{x}^{sel}$ as the offspring with the minimum constraint violation.

(d) If $\mathbf{x}^{sel}$ is empty, then skip this step. Otherwise: Store $\mathbf{x}^{sel}$ in the population attained by local search, POP^{ls}.

9. If POP^{ls} is empty, then skip this step. Otherwise:

   (a) Set POP^{sel} as an empty set.

   (b) Create a copy of POP^{ls} (i.e. POP'^{ls}).

   (c) If there are at least μ feasible individuals in POP^{ls}, then:

      i. Remove any infeasible individual from POP'^{ls}.

      ii. Select the individual from POP'^{ls} based on the MMDC w.r.t. $POP^{a} \cup POP^{sel}$.

      iii. Compute the Euclidean distance of each individual in POP'^{ls} w.r.t. $POP^{a} \cup POP^{sel}$.

      iv. Remove any individual from POP'^{ls} that violates the minimum distance criterion.

      v. If POP^{sel} includes μ entries or POP'^{ls} is an empty set, then skip this step. Otherwise, repeat steps 9(c)iii-9(c)v.

   (d) If there are less that μ feasible individuals in POP^{ls}, then:

      i. Select the individual from POP'^{ls} based on the MMDC w.r.t. $POP^{a} \cup POP^{sel}$.

      ii. Compute the distance of each individual in POP'^{ls} w.r.t. $POP^{a} \cup POP^{sel}$.

      iii. Remove any individual from POP'^{ls} that violates the minimum distance criterion.

      iv. If POP^{sel} includes μ entries or POP'^{ls} is an empty set, then skip this step. Otherwise, repeat steps 9(d)ii-9(d)iv.

10. Expensive function evaluations: The final step of the local phase is to compute the selected data points (POP^{sel}) using the expensive model and store the relevant data in the archive. The termination criteria are then considered.

5.3.2 Pool of Approximating models

Employing a pool of AM is made based on the *curse* and *blessing of uncertainty* which are thoroughly discussed in References [155, 169]. More specifically, relying on the prediction of different AM within the search may reduce the impact of the *curse of uncertainty*, i.e. inaccurate AMs may result in the search to stall or converge to false optimum, and exploit the effect of the *blessing of uncertainty*, i.e. AMs may smooth the search space and could contribute in accelerating and/or improving convergence. Therefore, the framework presented in Reference [155] is followed.

The pool of AM includes two model types. The first is a RBF model and the second is a low order PR model. The ensemble model included in the local phase is the sum (equal weight) of the prediction of these models. The AM is build based on the np available data points in the dataset. The vector $\mathbf{x}^{(s)}$ is used to suggest the s$^{\text{th}}$ archived data point. Moreover, a model is built for each objective function.

1. *PR model*: A simplified low (second) order PR model has been employed where the terms capturing the interactions among variables have been excluded. The prediction function for the m$^{\text{th}}$ objective function at point $\mathbf{x}$ is computed as follows:

$$\hat{f}_m(\mathbf{x}) = \beta_0 + \sum_{i=1}^{d}[\beta_i x_i] + \sum_{i=d+1}^{2d}[\beta_i x_{i-d}^2] \tag{5.45}$$

The least square method is used to estimate the unknown coefficients of the PR model.

$$\mathbf{y} = \mathbf{X\Theta} \tag{5.46}$$

where

$$\mathbf{X} = \begin{bmatrix} 1 & x_1^{(1)} & x_2^{(1)} & \cdots & x_d^{(1)} & (x_1^{(1)})^2 & (x_2^{(1)})^2 & \cdots & (x_d^{(1)})^2 \\ 1 & x_1^{(2)} & x_2^{(2)} & \cdots & x_d^{(2)} & (x_1^{(2)})^2 & (x_2^{(2)})^2 & \cdots & (x_d^{(2)})^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_1^{(np)} & x_2^{(np)} & \cdots & x_d^{(np)} & (x_1^{(np)})^2 & (x_2^{(np)})^2 & \cdots & (x_d^{(np)})^2 \end{bmatrix}, \mathbf{y} = \begin{bmatrix} f_m^{(1)} \\ f_m^{(2)} \\ \vdots \\ f_m^{(np)} \end{bmatrix},$$

The solution of the following system of equation determines an estimation $\hat{\mathbf{\Theta}}$ of the unknown coefficients:

$$\hat{\mathbf{\Theta}} = (\mathbf{X}^{\intercal}\mathbf{X})^{-1}\mathbf{X}^{\intercal}\mathbf{y}, \tag{5.47}$$

The aforementioned PR model requires at least $2d{+}1$ data point to determine the values of the coefficients which is also the minimum number of required linearly independent rows of matrix $\mathbf{X}$.

2. *RBF model*: The RBF model described in Section 4.52 is used with a cubic kernel function.

5.3.3 Pool of Variation operators

The pool of Variation operators includes a GA variant and a DE variant. The GA variant (Reference [207]) includes the modified tournament selection (Reference [103]), simulated binary crossover (Reference [208]) and the polynomial mutation (Reference [12]). The DE variant (Reference [209]) includes a DE/rand/1 mutation scheme and polynomial mutation. More specifically, the steps for the real-coded GA variant to generate a set of offspring are the following:

1. Apply the tournament selection to identify the parent population.
2. Apply the Simulated binary crossover based on the parent population to generate the offspring population.
3. Modify the offspring population by polynomial mutation.

Correspondingly, the steps of the DE variant are the following:

1. Apply DE/rand/1/bin to generate the trial vectors (offspring).
2. Modify the offspring by Polynomial mutation.

5.3.4 Environmental-Selection operator

The Environmental-Selection operator of NSGA-III (References [102, 103]) has been employed. NSGA-III is reference-point-based Many-objective EA that is based on the NSGA-II framework (Reference [66]). Even though the case of many-objectives is not considered, NSGA-III exhibits some interesting features which are relevant for this application: it can perform well with low population sizes, it can handle different scales of the objectives and it can include a user-defined preferred part of the Pareto-optimal front by defining representative reference points (Reference [102]). For greater detail, interested readers are kindly referred to References [102, 103].

5.4 Set-up of the Numerical Experiments

In this Section, the numerical experiments conducted are presented. Specifically, the following two sets have been carried out:

- The performance of the algorithm is examined since it is a heuristic approach. Therefore, it has been applied on a series of computationally cheap test cases.
- The applicability of the approach is examined on a computationally expensive test case. The results aim in examining trade-offs among different cost terms in MOO GEP considering flexibility constraints.

A detailed comparison of the optimization approach with other efficient surrogate-assisted derivative-free algorithms available within the literature has not been carried out for three main reasons. Firstly, it is hard to provide a fair comparison due to the inclusion of different operators (e.g. AMs or search engines employed) and different focuses (e.g. constraint or unconstrained problems). Secondly, the examined optimization approach for MOO GEP is based on such frameworks. Lastly, the scope of the numerical experiments is not to propose a new surrogate-assisted derivative-free algorithm but to examine the attained trade-offs from the application of one on a MOO GEP.

The following Sections present the considered test case, implementation notes, the numerical experiments and the employed performance metrics.

5.4.1 Test case

The test case is loosely based on the Greek Power sector [199] and serves purely as an illustrative example. The data are mostly based on data retrieved by ENTSO-e (Reference [210]) and ADMIE (Reference [199]). For the GEP level, a single target year is assumed (temporal detail). Two different areas are considered (spatial detail) to represent, to some extent, the north and south zones of the examined test case. The network limitations are assumed not binding among the two regions, however a limit has been imposed. Moreover, the test case is assumed isolated and the interaction with neighbouring power sectors has been omitted. Regarding technical detail, six candidate technologies have been assumed as possible capacity addition options. Three different thermal candidate units (TU1, TU2 and TU3), conventional hydro (CH), wind farms (WF) and solar photovoltaic (SPV) have been assumed. The latter can be included in each of the two areas. Furthermore, a distinction is made between the old and new capacity of thermal units, i.e. their technical-economical characteristics are assumed different. In addition, yearly investment cost is used for the assumed investment options which is derived based on the total investment cost. It is also assumed that up to 20 steps ($\overline{\mathbf{X}}$) of the capacity additions can be made in each area and technology group and a total of 20 steps are available (**TCA**). Moreover, other important flexibility providers, such as storage, expansions in the transmission system, demand-side products and interconnections, have been neglected. Moreover, the data used for the test case assumes zero shut-down cost. The data input are analytically presented in Appendix F.

5.4.2 Implementation notes

The approach is developed in Matlab within the Platemo platform [211] using available implementations, algorithms and operators. For local search the Sequential Quadratic Programming (SQP) solver available within Matlab is employed with a 500 FES limit. The Matlab implementation of the employed RBF and PR models (References [180, 181]) is used. The parameters of the optimization approach are presented in Table 5.1.

Parameter	Value	Parameter	Value
Max_FES	500	Max_Gen	100
μ	5	F	0.5
sc	0.1	CR	1
dv^{Tol}	$[1, 10^{-10}]$	ri	0.1
$dist^{min}$	$[1, 10^{-4}]$	N^{lc_dp}	$2(d+1)$
np^{init}	$2(d+1)$	pf^{MMDC}	0.9

Table 5.1: Parameter settings of the optimization approach. For the ZDT test cases parameters dv^{Tol} and $dist^{min}$ have been set as $dv^{Tol} = 10^{-10}$ and $dist^{min} = 10^{-4}$ while for the remaining cases their values have been set as $dv^{Tol} = 1$ and $dist^{min} = 1$.

For all independent runs a restriction of 500 FES limit of the true functions has been imposed, i.e. $Max\_FES = 500$. It should be highlighted that constraint functions (planning constraints) are not considered as a FES in all numerical experiments since we examine only the case of a MOO GEP with computationally expensive objective functions and computationally cheap constraint functions. In particular, the cost of computing the latter is negligible as within the formulation they do not require the output of the SM, i.e. they are computationally cheap functions of the capacity additions.

5.4.3 Numerical experiments

The employed optimization approach is examined on the following sets of computational experiments:

- Modified computationally cheap test suite:

  Five well-know MOO test problems which are part of the ZDT test suite (Reference [212]) are examined i.e. ZDT1, ZDT2, ZDT3, ZDT4 and ZDT6. The aforementioned are real-parameter unconstrained MOO problems that exhibit different characteristic, e.g. convex, non-convex, disconnected and/or multi-modal. However, the variables are rounded to a predefined decimal during the search, since the optimization approach considers discrete steps. In particular, variables are considered in discrete steps ($dv^{Tol} = 10^{-10}$), e.g. the range of a real parameter in $[0, 1]$ is altered to $[0, dv^{Tol}, ..., 1]$. This consequently, diverges from the original test suite.

Therefore, to make this clear the test suite should be perceived as *modified*. Moreover, the results derived for the ZDT test suite using an implementation (Reference [211]) of NSGA-III for 500, 10000 and 50000 FES are also presented as a point of reference. The performance of the two algorithms should not be directly compared due to the unfairness of the comparison; NSGA-III is not intended for computationally expensive MOO problems and the solver time is much lower.

- Computationally cheap MOO GEP test case without including a SM:

  The second case is a computationally cheap MOO GEP problem that does not require the output of a SM, i.e. short-term operation and the corresponding cost terms are omitted. Specifically, two objectives function are considered:

  1. Minimization of the sum of investment, FO&M and GPSC cost:

  $$f_1(\mathbf{x}) = c^{inv} + c^{fom} + c^{gp} \tag{5.48}$$

  2. Maximization of the anticipated RES generation:

  $$f_2(\mathbf{x}) = (-1)c^{arp} \tag{5.49}$$

     where c^{arp} is a cost term representing the anticipated RES generation (by modifying Eq. 5.8) and it is computed as follows:

  $$c^{arp} = C^{rp} \sum_{\forall a,h} [\overline{p}^{res}_{a,h}] \tag{5.50}$$

     where $C^{rp} = 1$.

  The aforementioned MOO problem is subjected to the planning constraint functions (Section 5.2.4). This test case, hereafter, will be referred to as *MOOGEP-noSM*. Since the data input may have an impact on the results, the numerical experiment is repeated for a case where the existing capacity is omitted (greenfield case). Correspondingly, it will be referred to as *MOOGEP-noSM-GF*. Despite the differences in the objective space, a main difference among the two cases regards the planning reserve margin constraint (Eq. 5.27) which can not be violated in the *MOOGEP-noSM* case. Therefore, the infeasible region, which is determined by the planning constraints, in the *MOOGEP-noSM-GF* is larger. The user defined initial points suggested in Section 5.3.1 are not supplied.

- Computationally cheap GEP with the temporal resolution set at one day: Each of the presented cases (Section 5.2.4) is examined using the optimization approach. For this numerical experiment, temporal resolution has been limited to 24h ($t^{fes} < 1sec$). The data input of the first day is used. The five presented test cases will be referred to as *MOOGEP-(XXX)-1D*, e.g. *MOOGEP-(AC1-EM)-1D*. This numerical experiment is also repeated for the greenfield case and, correspondingly, the five problems will be referred to as *MOOGEP-(XXX)-1D-GF*. The test cases (*MOOGEP-(XXX)-1D* and *MOOGEP-(XXX)-1D-GF*) also differ with regard to the feasible data points that are penalized internally, e.g. the supplied initial point representing the minimum capacity additions is feasible in the *MOOGEP-(XXX)-1D* case and infeasible in *MOOGEP-(XXX)-1D-GF*. It also exhibits an extremely high function value due to high NSE cost for the *MOOGEP-(XXX)-1D-GF* cases.

- Large scale MOO GEP considering flexibility constraints: The formulations presented in Section 5.2.4 are considered to examine the trade-offs of the examined cost terms in MOO GEP. The numerical experiments are repeated as in the corresponding computationally cheap numerical experiment. However, the temporal resolution is set at 4 weeks. Each week is computed independently and the cost terms are summed, i.e. each day/week is assumed to have an equal weight. These test cases shall be referred to *MOOGEP-(XXX)-4W* and *MOOGEP-(XXX)-4W-GF* for when the existing capacity or the greenfield case are considered, respectively.

5.4.4 Performance metrics

This Section presents the performance metrics employed to assess the PFA and UFA.

Performance metrics of PFA

The Hypervolume metric (Reference [213]) has been employed to assess the performance of the optimization approach. It is a Pareto-compliant metric. It may assess both the convergence and diversity of an attained population. It defines the volume enclosed by a population and a user-specified reference point (Reference [214]). The HV metric is computed as follows:

$$HV1 = \frac{H(NDF^{pop\text{-}f})}{H(PF^{*})} \tag{5.51}$$

where $HV1$ is the fraction of the HV computed for the NDF of the data points attained from a single run ($H(NDF^{pop\text{-}f})$) and the HV of the PF ($H(PF^{*})$). Since an estimation of the PF of each GEP test case is not available, the HV values are computed w.r.t. a NDF generated by applying NDS on the feasible data points of all 25 independent runs using the 500 FES limit. Hereafter, an NDF derived by the aforementioned procedure shall be referred to as the best-found Non-Dominated Front(NDF*); the NDF* could deviate from the true PF.

Note that function values are normalized based on the maximum and minimum values of each corresponding function value in PF^{*} (or in NDF* when required). Any population member exceeding these maximum values is removed. The reference point is set at (1.1,1.1).

Moreover, the HV metric is also employed to estimate the feasible objective space dominated by each final NDF and account for the initially supplied data points:

$$HV2 = \frac{H(NDF^{pop\text{-}f}) - H(NDF^{pop\_in})}{H(PF^{*}) - H(NDF^{pop\_in})} \tag{5.52}$$

where $H(NDF^{pop\text{-}f})$ is the HV computed for the NDF from the feasible dataset of a single run. $H(NDF^{pop\_in})$ is the HV computed for the NDF attained by the feasible data points of the initial dataset of the same run. The HV of the PF ($H(PF^{*})$) is replaced by $H(NDF^{*})$ when the former is not available. In this case, function values are normalized based on the maximum and minimum values of all attained feasible solutions within the datasets. Any population member exceeding these maximum values is removed and the reference point is set at (1.1,1.1).

This metric accounts for the progress attained regarding the dominated feasible objective space w.r.t. the initial sample. It should be highlighted that the HV is computed based on the entire dataset attained during each run and differences in the number of solutions in the resulting NDFs of the runs are not considered.

In general, higher values of $HV1$ and $HV2$ suggest an improved performance; in both cases the ideal values are 1.

Performance metrics of UFA

The metrics employed for measuring the accuracy of the UFA are presented in this Section. The accuracy is measured according to the Root Mean Square Error (RMSE) and the Coefficient of determination metrics by employing a leave-one-out cross validation approach. The metrics are computed independently for each AM and each objective function.

- The value of the RMSE metric is computed as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{np}[(\hat{f}(\mathbf{x_i}) - f(\mathbf{x_i}))^2]}{np}} \tag{5.53}$$

  where $f(\mathbf{x_i})$ is the true value of an objective function for $\mathbf{x_i}$, $\hat{f}(\mathbf{x_i})$ is the corresponding prediction of the AM built by excluding data point $\mathbf{x_i}$ from the dataset and np is the number of available data points within the dataset. Lower values indicate a more accurate AM and the ideal value is 0.

- The value of the R^2 metric is computed as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{np}[(\hat{f}(\mathbf{x_i}) - f(\mathbf{x_i}))^2]}{\sum_{i=1}^{np}[(\hat{f}(\mathbf{x_i}) - f^{mean})^2]} \tag{5.54}$$

  where f^{mean} is the mean value of the available true values of the objective function. For this metric, higher values are preferred which indicate a more accurate AM and the best possible value is 1.

5.5 Results

This Section presents the results of the numerical experiments presented in Section 5.4.

5.5.1 Performance of optimization approach on computationally cheap test cases

In the following Sections the results of the numerical experiments, described in Section 5.4.3, are presented. For each, a visual illustration of the attained NDF, the values of the HV based metrics and the performance of the UFA are presented.

Modified computationally cheap test suite (ZDT)

Figure 5.1 presents a visual comparison of non-dominated trade-offs attained by the results for the min, median and max HV1 in comparison to the true PF (it does not consider the discrete steps) of each of the five ZDT problems. Moreover, the attained datasets during all 25 optimization runs and the NDF* are also depicted. Figure 5.2 depicts the attained HV1 and HV2 values for each of the five test cases within the 500 FES limit imposed. Moreover, Figure 5.3 provides the attained results by the implementation of NSGA-III for 500, 10000 and 50000 FES to serve as a point of reference.

It can be seen that satisfactory results have been attained for the problems ZDT1, ZDT2 and ZDT3 within the limited FES considered. The optimization approach in ZDT6 had been less successful. It provided a few true PF members in each independent run. On the contrary, it had been unsuccessful in providing any true PF member within the limited FES considered for the ZDT4 problem. Each ZDT test case exhibits distinct challenges (Reference [212]). However, the corresponding ones for problems ZDT4 and ZDT6 are harder to address in this case. In particular, the problem ZDT6 exhibits a low density of solutions near the PF and the problem ZDT4 includes highly multimodal functions and a large number of local Pareto-optimal fronts; the problem ZDT4 is a challenging test case even when *perfect* information of the functions is available (Reference [66]).

In general, the deterioration in the performance of the approach could be attributed to the poor UFA attained. This might be a result of the aforementioned characteristics of the test functions (ZDT4 and ZDT6) for which attaining a decent UFA is rendered highly challenging. This can be seen in Figure 5.4 where the values of the performance metrics employed to assess the attained UFA are presented. More specifically, the values of the metrics indicate that the attained UFA is less accurate in cases ZDT4 and ZDT6 in comparison to the remaining cases. Consequently, even if the evolution search performs *ideally*, the algorithm is likely to converge to false optima, within the limited FES considered, due to a poor UFA. In addition, the case of stalling to a local Pareto-optimal front even when a *perfect* UFA has been achieved cannot be excluded since the ZDT4 test case exhibits a large number of local Pareto-optimal fronts.

Furthermore, the minimum distance criterion and the variable decimal tolerance can also exclude a number of possible solutions. An improved solution that exhibits a lower minimum distance to a data point in the archive than the threshold value cannot be costly evaluated. This, however, depends on the user-defined parameters and the location of the archived data points.

Computationally cheap GEP test cases neglecting short-term operation

Figure 5.5 presents the results for the min, median and max HV in comparison to the NDF* for each of the *MOOGEP-noSM* and *MOOGEP-noSM-GF* problems. It also presents the combined dataset attained during all 25 optimization runs and the NDF*. Moreover, the results derived by NSGA-III for 50000 FES limit are presented. In addition, it presents the corresponding HV1 and HV2 values. The values of the metrics employed to assess the UFA are presented in Figure 5.6.

The results of NSGA-III are presented to serve only as a point of reference and examine if the employment of the AMs has mislead the search. It is stressed that the comparison is not fair to any extent. For example, the NSGA-III: (i) had not been developed with a focus on computationally expensive optimization problems, (iii) exhibits a much lower solver time, (iv) does not include a gradient-based local search solver, and (v) does not include a decimal tolerance or a minimum distance criterion. More importantly, the number of solutions in the NDF attained by the dataset can exceed the population sized of NSGA-III which has an impact on the computed HV metrics. Lastly, the optimization approach could be roughly perceived as a variant of NSGA-III. Therefore, it should be made clear that the results do not undermine the performance of NSGA-III.

By comparing the attained NDFs by the NSGA-III and the optimization approach it can be observed that both provided satisfactory results. Also, it can be noticed that the AMs had not mislead the search and an acceleration w.r.t. the FES employed had been achieved. This can be seen in Figure 5.6 where the values of the RMSE and R^2 are presented. The attained values suggest that a decent UFA had been achieved. Moreover, by comparing the performance of the algorithms among the two examined cases (*MOOGEP-noSM* and *MOOGEP-noSM-GF*), the results indicate that the infeasible region defined by the planning constraint and the differences of considering the existing capacity have not heavily affected the performance of the optimization approach.

Computationally cheap GEP with limited temporal resolution (one day)

Figure 5.7 presents the combined dataset attained during all 25 optimization runs and the corresponding NDF*. Furthermore, the results for the min, median and max HV in comparison to the NDF* are presented for the *MOOGEP-(XXX)-1D* and *MOOGEP-(XXX)-1D-GF* cases, respectively. Figure 5.8 depicts the corresponding serial progress for the HV1 and HV2 values w.r.t. each NDF*. The values of the RMSE and R^2 metrics are presented in Figures 5.9 and 5.10 for each *MOOGEP-(XXX)-1D* and *MOOGEP-(XXX)-1D-GF* case, respectively.

The results for each of the five different MOO problems in each test case considered were satisfactory. Moreover, some differences can be observed among the results attained for the *MOOGEP-(XXX)-1D* and *MOOGEP-(XXX)-1D-GF* test cases. Specifically, the optimization approach performed more robustly when the existing capacity had been considered (Fig. 5.8). These differences had not been observed in the *MOOGEP-noSM* and *MOOGEP-noSM-GF* test cases and, therefore, could be induced due to the inclusion of the SM. More specifically, it could be attributed to challenges emerging by large numerical differences among the function values of the attained data points caused by high penalty costs. This can be observed by comparing the attained datasets (Fig. 5.7) and the corresponding NDFs of each test case. These numerical differences have an impact on the UFA accuracy as the metrics suggest a better UFA accuracy for the *MOOGEP-(XXX)-1D* cases in comparison to the *MOOGEP-(XXX)-1D-GF* cases (Fig. 5.9 and 5.10).

The cases differ also w.r.t. the costs derived by the existing capacity, which is not assumed as a capacity addition option in the greenfield cases. Consequently this should be considered when directly comparing the numerical values of the attained NDFs. Moreover, the results are not analyzed in terms of trade-offs since a limited temporal resolution (1 day) had been considered. Such are discussed in the following Section.

5.5.2 Examining trade-offs in MOO GEP considering flexibility constraints

This Section presents the results for the large scale test problems where the five MOO GEP variants described in Section 5.2.4 are considered. First, the performance of the optimization algorithm is assessed and the quality of the UFA in each problem is presented. Then the emerging trade-offs are analysed for each MOO instance.

Performance of the optimization approach

Figure 5.11 presents the combined dataset attained during all 25 optimization runs and the corresponding NDF*. The results for the min, median and max HV in comparison to the NDF* are presented for the *MOOGEP-(XXX)-4W* and *MOOGEP-(XXX)-4W-GF*, respectively. Figures 5.12 depicts the corresponding serial progress for the HV1 and HV2 values w.r.t. each NDF*.

The attained NDFs (Fig. 5.11) and the serial progress observed for the values of the HV metrics (Fig. 5.12) could imply an acceptable performance. This is clearer for the attained for the *MOOGEP-(XXX)-4W* cases. For some of the greenfield cases (*MOOGEP-(XXX)-4W-GF*) a larger number of FES is required as indicate by the results, e.g. *MOOGEP-AC2-RP-4W-GF* (Fig. 5.12). Figures 5.13 and 5.14 present the attained values of the metrics employed to assess the UFA.

The increase in temporal detail alters the search space (it results in a different optimization problem). This restricts the extent for which the comparison among the results derived for different temporal detail can be made. However, some observations can be made by comparing the datasets attained from the numerical experiment including one day (*MOOGEP-(XXX)-1D*) and the equivalent ones attained by setting the temporal detail to four weeks (*MOOGEP-(XXX)-4W*). In particular, it is noticed that a larger number of extreme cases are introduced in the dataset of the former. This could imply that a higher number of states that include penalties are selected to be costly evaluated. Moreover, the NDF* attained have been also affected as cost terms are derived based on a larger set of operating conditions.

A reduction in the UFA accuracy is observed based on the metrics employed. These indicate that the level of accuracy attained in the cases including lower temporal detail had not been achieved. This is consistent over all *MOOGEP-(XXX)-4W* cases. It is observed that the highest values of the metrics are attained for cases *MOOGEP-AC2-RP-4W* and *MOOGEP-AC2-RP-4W-GF*. This might suggest that the penalty term included in the objective function of the remaining cases has an impact on the UFA accuracy. More specifically, a larger set of operating conditions are examined in the case where temporal detail is increased. A number of data points evaluated by the SM could be penalized internally. This penalization adds a high cost to the objective function and, therefore, the increase in the objective function value is relatively large in comparison to the difference emerging among data points that are not penalized. Consequently, the underlying function exhibits a large increase of its value at points that are penalized.

Analysis of the derived cost term values and capacity additions

In this Section, the emerging trade-offs among the cost terms are discussed. In particular, for each *MOOGEP-(XXX)-4W* test case the evolution of these cost terms are examined along the attained NDF. The corresponding capacity additions are presented to facilitate the analysis. For each of these test case the NDF of the independent run presenting the highest value of the HV1 metric is discussed.

There are two limitations regarding the follow analysis. Firstly, it is based on the attained PFA which relies on the performance of the approach. Secondly, the implications of this analysis are bounded by assumptions made regarding the formulation of the optimization problem, e.g. the selected SM and the data input of the test case. Moreover, emerging trade-offs can be problem-specific as these are based on the data input of an examined test case.

Figures 5.15-5.19 present the values of the derived cost terms along the attained NDFs for the *MOOGEP-(XXX)-4W* test cases. Figures 5.20-5.24 present the corresponding values of the decision vectors. It can be seen that different trade-offs can emerge for each one.

- *MOOGEP-AC1-EM-4W*: Aiming towards a higher emission reduction is associated with higher investment, FO&M, generating and GPSC cost in comparison to lower aggregated cost (Fig. 5.15). Moreover, aiming towards lower emission cost results in higher RES penetration including also a low level of RES curtailment. Furthermore, a shift in the selected capacity additions of thermal units (from TU1 to TU2 in area 2) can be observed (Fig. 5.20) near the lowest levels of achieved emission cost. The latter has an impact on the trends of the cost terms. More specifically, capacity additions in TU2 are prioritized for achieving the lower emission reductions due to the lower emission factor of TU2 in comparison to TU1. Lower emission cost can be attributed to the high level of capacity additions in WF, SPV and CH. Among the aforementioned, WF are prioritized. Capacity additions tend to be allocated in both areas suggesting that the location might be important.

- *MOOGEP-AC2-RP-4W*: The results indicate that higher RES penetration levels are associated with higher investment, FO&M, GPSC and curtailment cost in comparison to lower aggregated cost (Fig. 5.16). In contrast, lower generating, VO&M and emission cost are observed for the former. This can be attributed to the increase in RES capacity additions (Fig. 5.21). Once again, WF are prioritized among the non-thermal technology groups and TU1 are prioritized among the conventional thermal units. Capacity additions are also made in TU3 for the highest achieved levels of RES penetration. Moreover, capacity additions in RES technology groups tend to be allocated in both areas suggesting that the location might be important in contrast to capacity addition in TU1. Capacity additions are not made for CH. This can be anticipated as the contribution of CH generation is not accounted for in the environmental function and, therefore, such capacity additions could be motivated merely by a reduction in the value of the aggregated cost objective function (AC2). In a similar manner, a lower number of capacity additions are made in conventional thermal units (TU1, TU2 and TU3).

- *MOOGEP-OC1-GP-4W*: The results indicate that lower operating cost (OC1) are associated with higher investment, FO&M and GPSC in comparison to lower GPSC (Fig. 5.17). However, lower GPSC are associated with higher generating, VO&M and emission cost in comparison to lower operating cost (OC1). As anticipated, lower levels of RES penetration are attained for lower GPSC. These can be attributed to the observed higher RES capacity additions (Fig. 5.22) made for achieving lower operating cost that results in higher investment cost. Consequently, the FO&M cost is also increased. CH capacity additions are prioritized along the NDF implying that such are beneficial for both objective functions. In general, capacity additions are made in TU1, RES and CH. Regarding spatial allocation TU1, CH and RES capacity additions are made in both areas.

- *MOOGEP-OC2-IC-4W*: The results indicate that lower operating cost (OC2) are associated with higher investment, FO&M and GPSC and lower generating and emission cost in comparison to lower GPSC (Fig. 5.18). In comparison, to the *MOOGEP-OC1-GP-4W* case where CH had been highly prioritized, such capacity additions are made here only for the lowest levels of operating cost (OC2) (Fig. 5.23). Moreover, a larger number of capacity additions in TU1, RES and CH is observed for lower values of operating cost (OC2) which consequently increases the value of the second objective that is comprised by the investment cost (IC).

- *MOOGEP-AC3-AP-4W*: The results indicate a single solution as optimal (Fig. 5.19). This implies the aforementioned solution could minimize both considered objective functions. Interestingly, the aggregated function prioritizes only a limited number of capacity additions in TU1 (Fig. 5.24). Consequently, the aforementioned suggests that the benefits from capacity additions in other technology groups are not sufficient to produce a solution with a lower value of the aggregated objective function. This can be also attributed to the sufficiency of the existing capacity to meet the planning reserve margin constraint.

Overall, the result indicate that a detail assessment of optimal capacity additions on a cost term level may provide additional information. Given different objective functions, different NDFs may be generated. The set of alternative solution provided as PFA can be analysed to identify the optimal one which can be determined based on the decision makers preferences.

5.5.3 Figures

This Section includes all figures mentioned within the Results Section.

Modified computationally cheap test suite (ZDT)

This Section includes four figures. Figure 5.1 presents the distribution of the data points (objective space) included within the combined datasets and the NDFs derived for the ZDT test cases. An estimation of the PF is provided for the comparison. For each such estimation 1000 points have been used except for ZDT3 which includes 1066 points. Figure 5.2 presents the values of the HV1 and HV2 metrics for the ZDT test cases. Figure 5.4 presents the values of the RMSE and R^2. Figure 5.3 presents the results derived for NSGA-III for different limits of available FES in comparison to the attained ones within the 500 FES limit.

Figure 5.1: The combined datasets (blue dots) and the corresponding NDF^* (red X mark) for each ZDT test case are presented in the upper five sub-figures. The lower five present the NDFs of a single run which exhibit the maximum (blue circles), median (cyan squares) and minimum (red diamonds) value of the HV1 metric. Moreover, an estimation of the PF is presented for comparison (black crosses). For problem ZDT4, the HV2 metric has been used as the HV1 metric is equal to zero for all 25 runs.

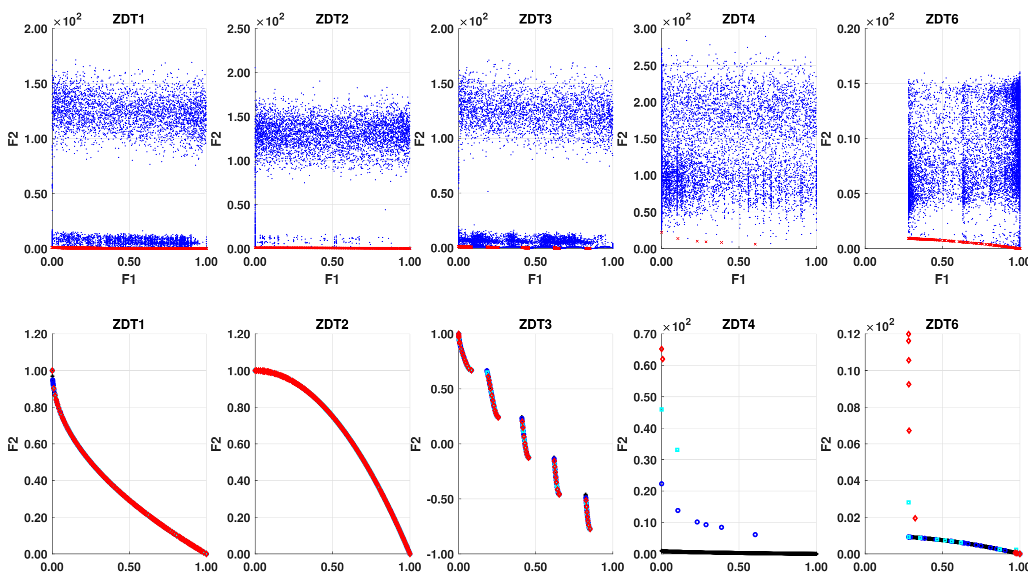


Figure 5.2: Box plots of the HV1 (left) and HV2 (right) metrics for the ZDT test cases.

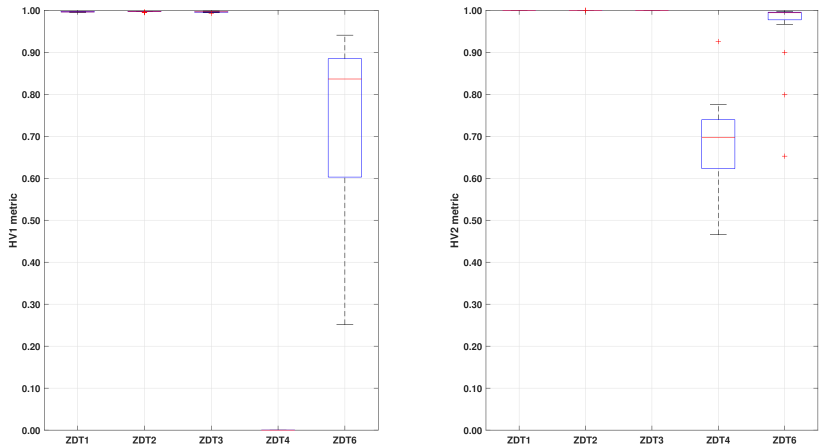


Figure 5.3: NDFs attained by NSGA-III for each ZDT test cases (blue triangles) for 500 FES (first row), 10000 FES (second row) and 50000 FES (third row). The corresponding NDFs (red circles) attained within the 500 FES limit by the optimization approach are also presented in each sub-figure. The NDFs exhibiting the median value of the HV1 metric among the 25 independent runs have been selected. For problem ZDT4, the HV2 metric has been used for this selection as the HV1 values are all equal to zero.

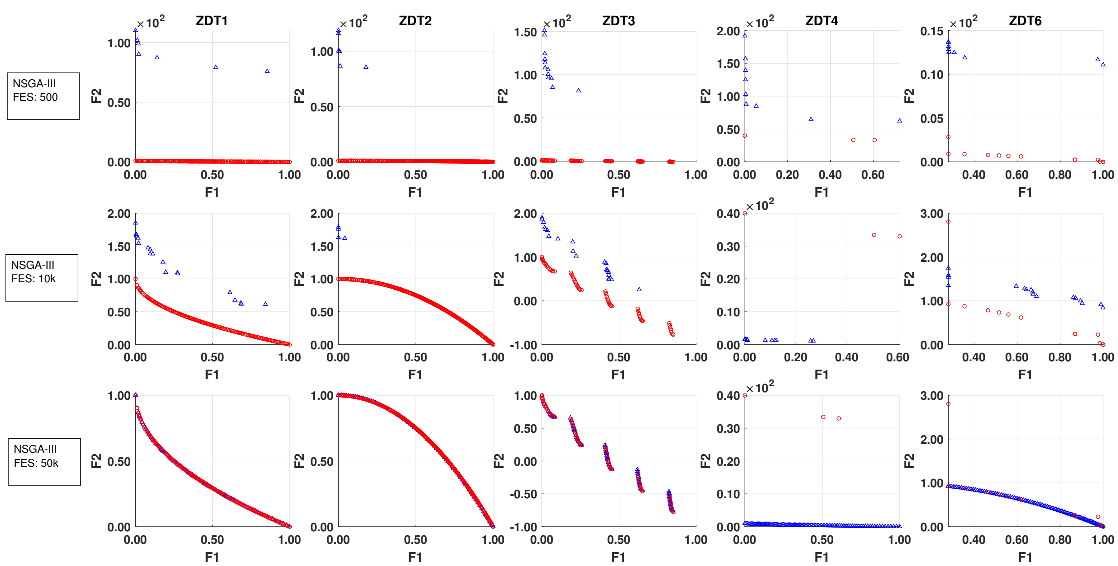
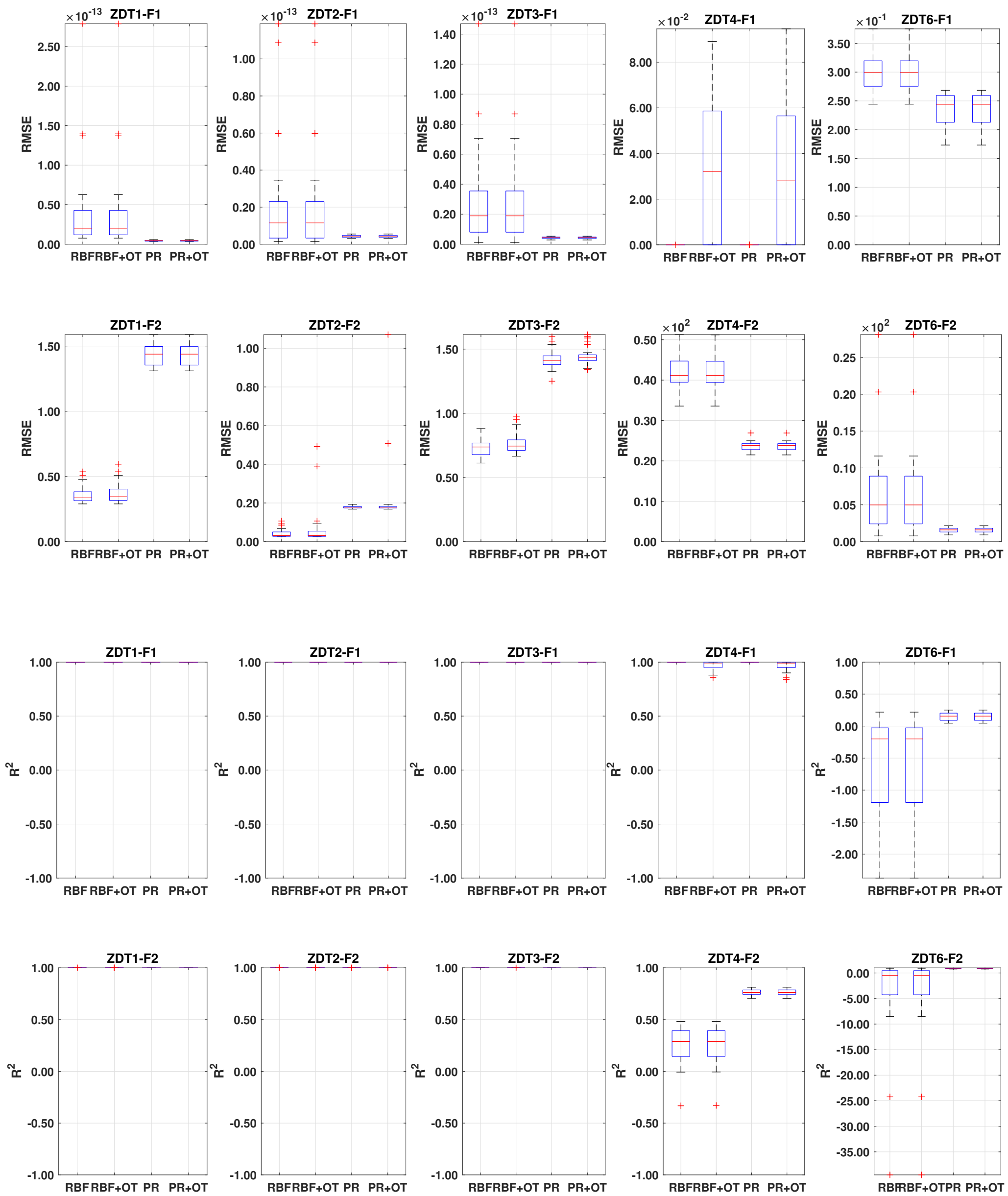


Figure 5.4: The figure presents the values of the RMSE (first and second rows) and R^2 (third and fourth rows) metrics for the ZDT test cases. The first and third row present the metric values for the first objective function. In a similar manner, the second and fourth row present the metric values for the second objective function. Each sub-figure includes four box plots presenting the values of the metrics for the RBF and PR models of each independent run without and with the inclusion of the outlier technique, namely: RBF, RBF+OT, PR and PR+OT.



Computationally cheap GEP test cases neglecting short-term operation

This Section includes Figure 5.5 which presents the combined dataset the NDFs attained for a 500 FES limit and the corresponding values of the HV1 and HV2 metrics. Moreover, a comparison is provided to the results derived by employing NSGA-III for a 50k FES limit.

Figure 5.5: Results derived for the *MOOGEP-noSM* (upper row) and *MOOGEP-noSM-GF* (lower row) test cases employing the optimization approach for a 500 FES limit and NSGA-III for a 50k FES limit. The feasible data points (blue dots) and infeasible data points (red dots) from the combined dataset and the corresponding NDF^* (black crosses) for each test case are presented in the sub-figures of the first column. The sub-figures of the second column present the NDFs of a single run which exhibit the maximum (blue circles), median (cyan squares) and minimum (red diamonds) value of the HV1 metric in comparison to the corresponding NDF^* (black crosses). In a similar manner, the corresponding results for NSGA-III (50k FES limit) are presented in the sub-figures of the third column. The box plots within the final two columns present the attained values for the HV1 (left) and HV2 (right) metrics as derived for the *MOOGEP-noSM* and *MOOGEP-noSM-GF* test cases. The numerical scales of the sub-figures corresponding to the *MOOGEP-noSM* and *MOOGEP-noSM-GF* cases differ.

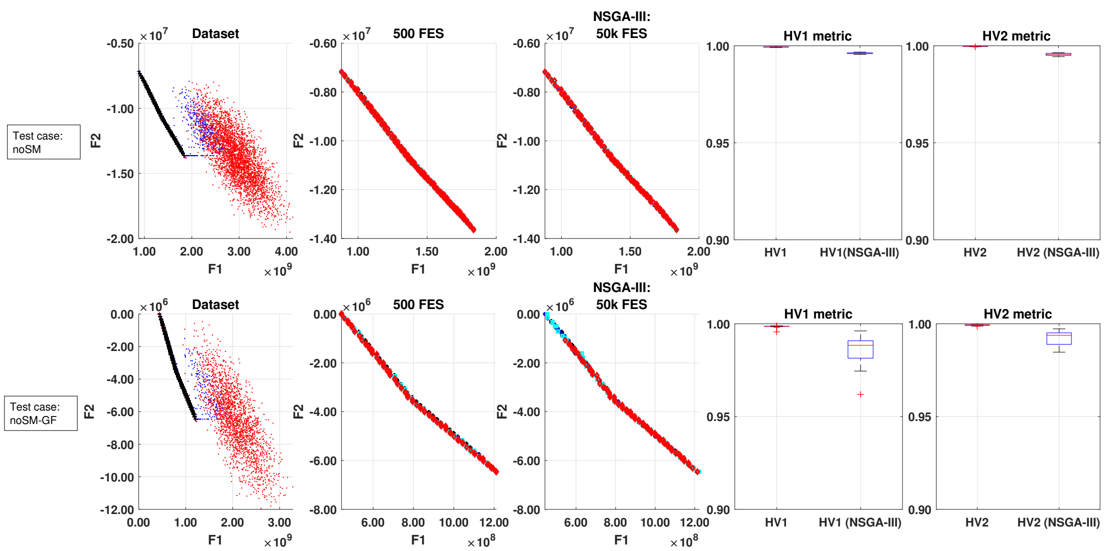
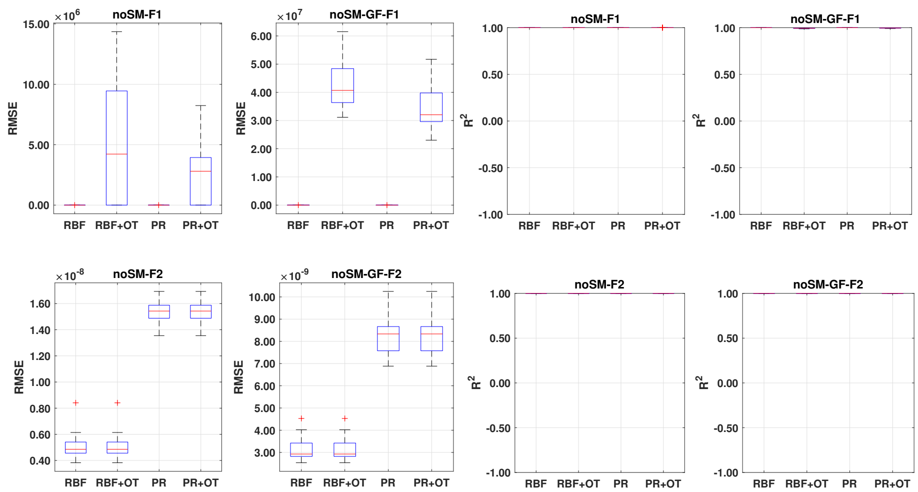


Figure 5.6: The figure presents the values of the RMSE (first and second columns) and R^2 (third and fourth columns) metrics for the *MOOGEP-noSM* and *MOOGEP-noSM-GF* test cases. The first row presents the metric values for the first objective function. In a similar manner, the second row presents the metric values for the second objective function. Each sub-figure includes four box plots presenting the values of the metrics for the RBF and PR models of each independent run without and with the inclusion of the outlier technique, namely: RBF, RBF+OT, PR and PR+OT.



Computationally cheap GEP with limited (1day) temporal detail

This Section includes four figures. Figure 5.7 presents the combined datasets, the NDFs attained for a 500 FES limit for each of the *MOOGEP-(XXX)-1D* and *MOOGEP-(XXX)-1D-GF* test cases. The corresponding values of the HV1 and HV2 metrics are presented in 5.8. Figure 5.9 presents the values of the RMSE and R^2 metrics for the *MOOGEP-(XXX)-1D* cases. In a similar manner, Figure 5.10 presents the values of the RMSE and R^2 metrics for the *MOOGEP-(XXX)-GF-1D* cases.

Figure 5.7: Results derived for the *MOOGEP-(XXX)-1D* and *MOOGEP-(XXX)-1D-GF* test cases employing the optimization approach for a 500 FES limit. The feasible data points (blue dots) and infeasible data points (red dots) from the combined dataset and the corresponding NDF^* (black crosses) for each *MOOGEP-(XXX)-1D* test case are presented in the sub-figures of the first row. The second row present the NDFs of a single run which exhibit the maximum (blue circles), median (cyan squares) and minimum (red diamonds) value of the HV1 metric in comparison to the attained NDF^* (black crosses) for each *MOOGEP-(XXX)-1D* test case. The third and fourth rows correspond to the results for the *MOOGEP-(XXX)-1D-GF* cases.

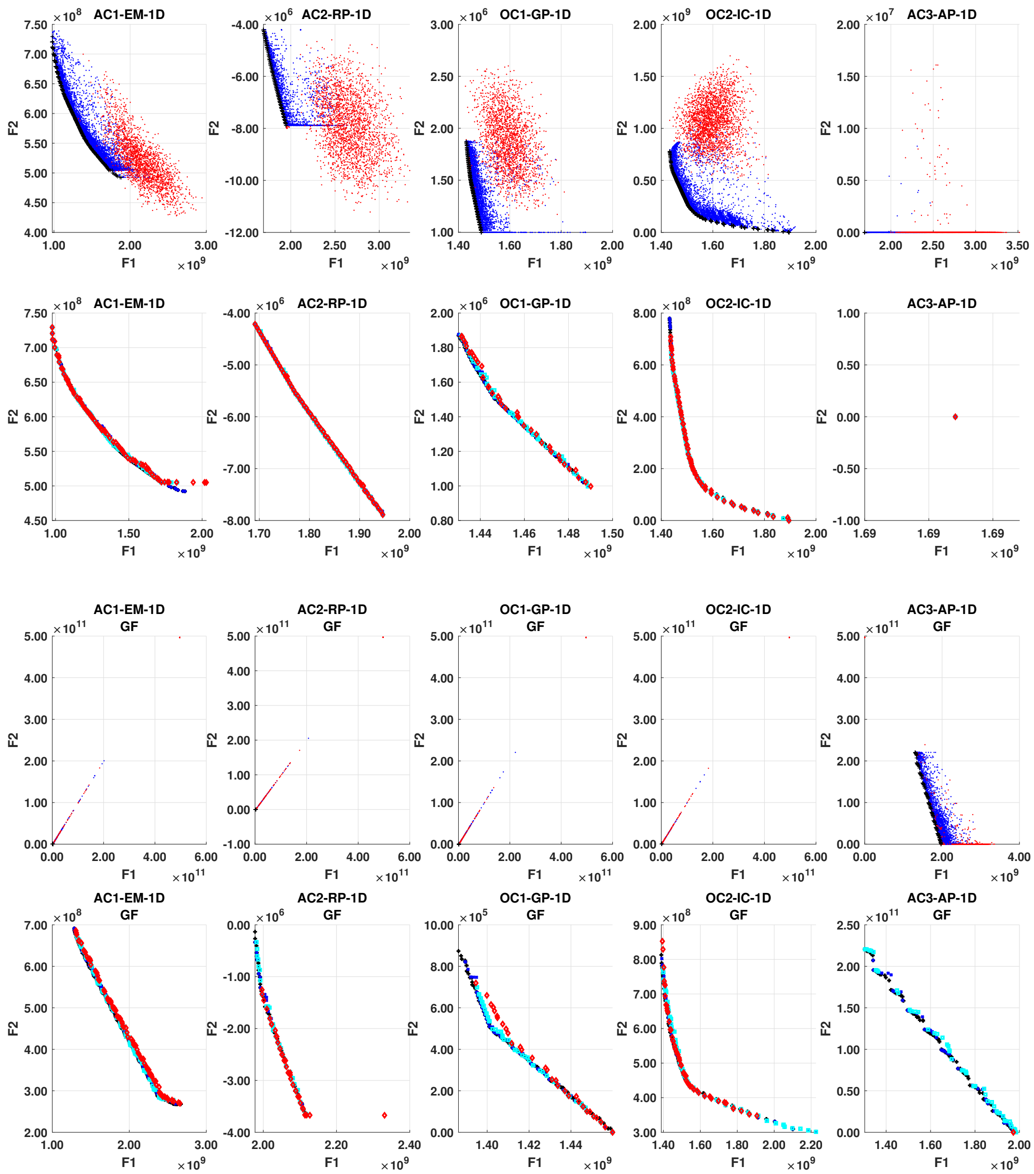


Figure 5.8: The figure presents the HV1 (first and third rows) and HV2 (second and fourth rows) values for the *MOOGEP-(XXX)-1D* (first and second rows) and *MOOGEP-(XXX)-1D-GF* (third and fourth rows) test cases. The first step of each box plot represents the value attained by the initially supplied data points of each independent run. Each of the remaining steps corresponds to the HV values attained by the datasets including up to $i * 100$ data points (FES), e.g. the third step (3) represents the HV values for the independent datasets including up to 300 data points.

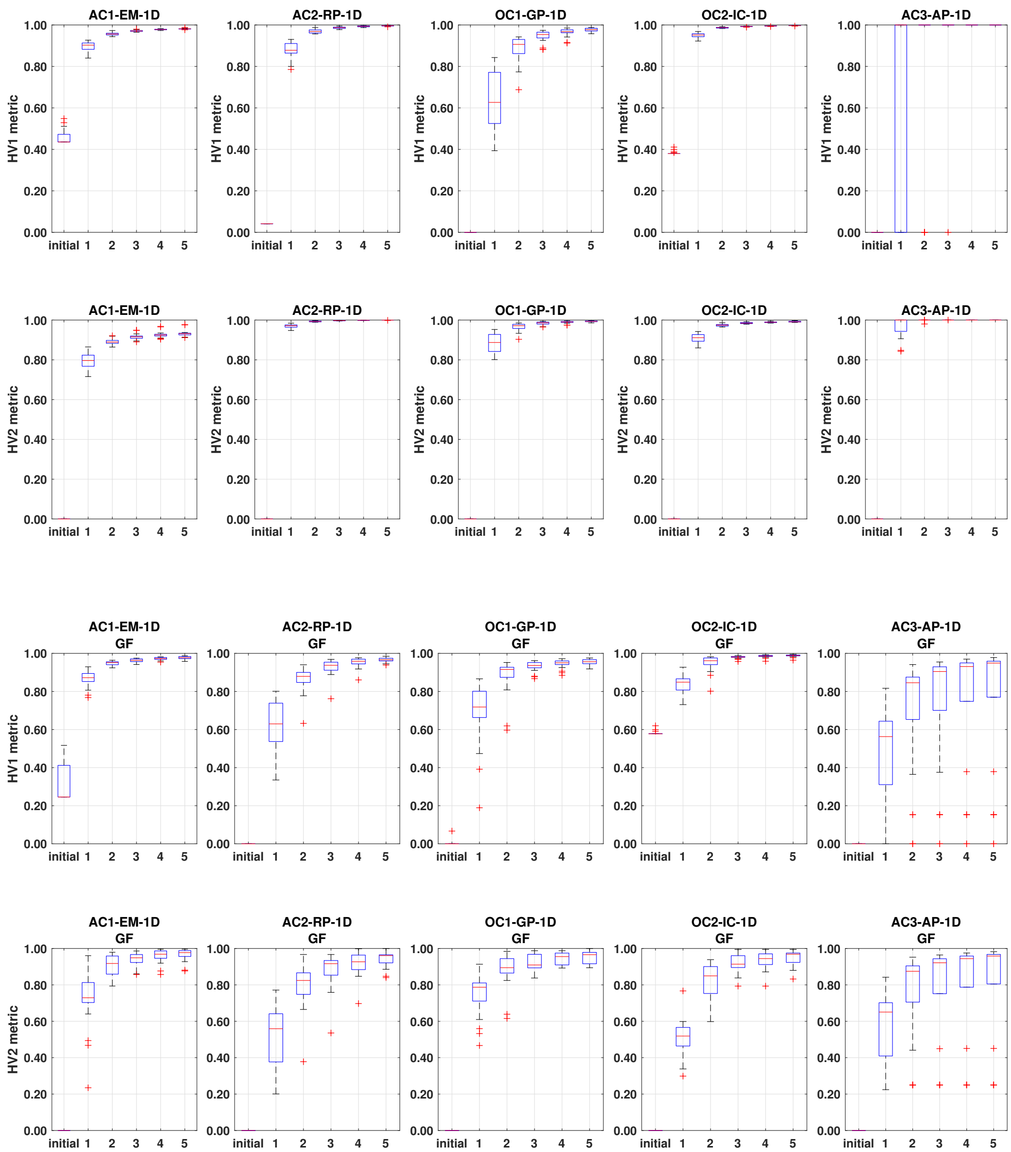


Figure 5.9: The figure presents the values of the RMSE (first and second rows) and R^2 (third and fourth rows) metrics for the *MOOGEP-(XXX)-1D* test cases. The first and third row present the metric values for the first objective function while the second and fourth row correspond to the ones of the second objective function. Each sub-figure includes four box plots presenting the values of the corresponding metrics for the RBF and PR models of each independent run without and with the inclusion of the outlier technique, namely: RBF, RBF+OT, PR and PR+OT. For the *MOOGEP-AC3-AP-1D* case 2 runs had not attained a value including a penalty term and these are excluded from computing the R^2 metric.

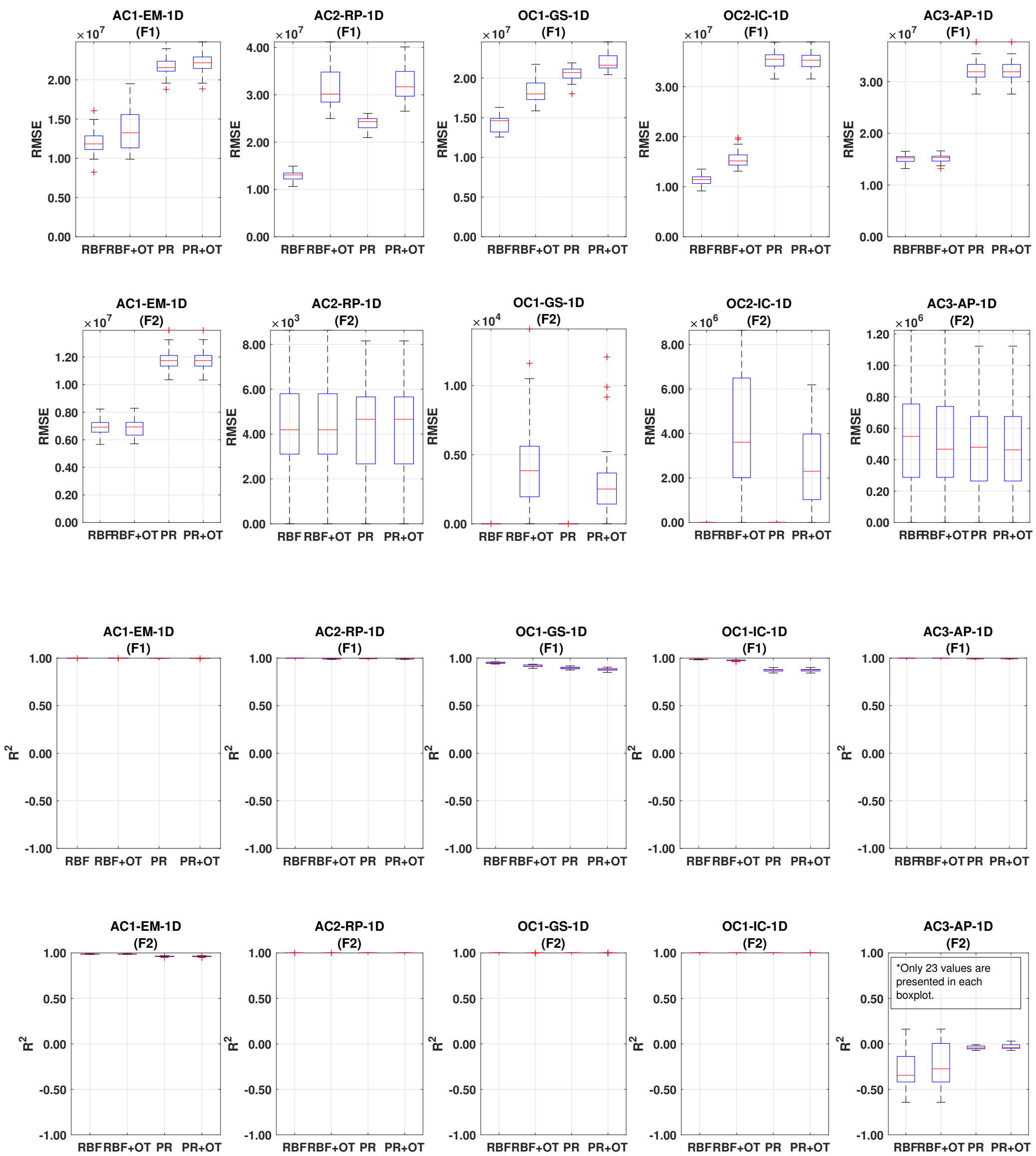
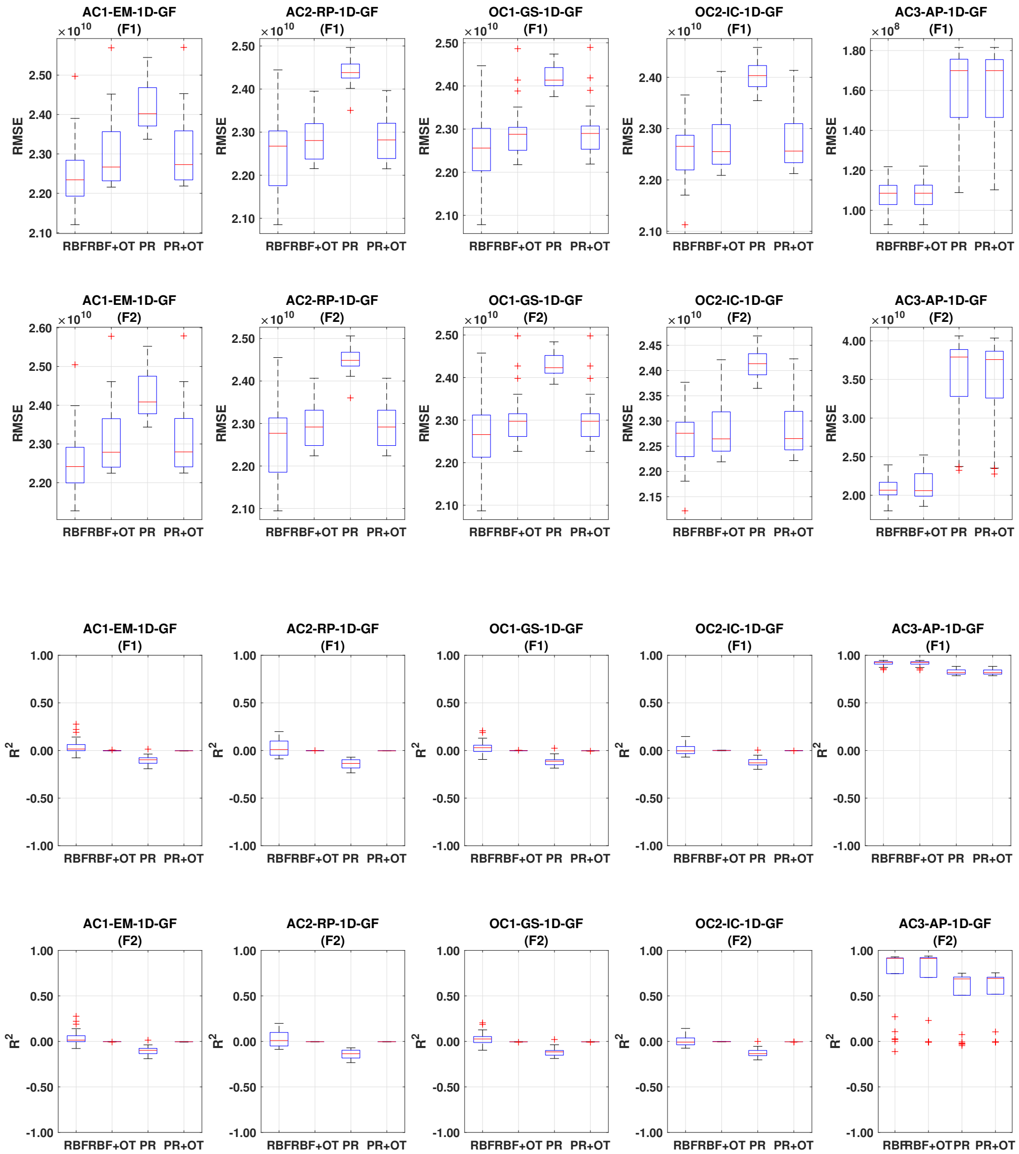


Figure 5.10: The figure presents the values of the RMSE (first and second rows) and R^2 (third and fourth rows) metrics for the *MOOGEP-(XXX)-1D-GF* test cases in a similar manner to Fig. 5.9.



Examining trade-offs in MOO GEP considering flexibility constraints

This Section includes four figures. Figure 5.11 presents the combined datasets, the NDFs attained for a 500 FES limit for each of the *MOOGEP-(XXX)-4W* and *MOOGEP-(XXX)-4W-GF* test cases. The corresponding values of the HV1 and HV2 metrics are presented in 5.12. Figures 5.13 and 5.14 present the values of both, the RMSE and R^2, metrics for the *MOOGEP-(XXX)-4W* and for the *MOOGEP-(XXX)-4W-GF* cases, respectively.

Figure 5.11: Results derived for the *MOOGEP-(XXX)-4W* and *MOOGEP-(XXX)-4W-GF* test cases employing the optimization approach for a 500 FES limit. The feasible data points (blue dots) and infeasible data points (red dots) from the combined dataset and the corresponding NDF^* (black crosses) for each *MOOGEP-(XXX)-4W* test case are presented in the sub-figures of the first row. The second row present the NDFs of a single run which exhibit the maximum (blue circles), median (cyan squares) and minimum (red diamonds) value of the HV1 metric in comparison to the attained NDF^* (black crosses) for each *MOOGEP-(XXX)-4W* test case. The third and fourth rows correspond to the results for the *MOOGEP-(XXX)-4W-GF* cases.

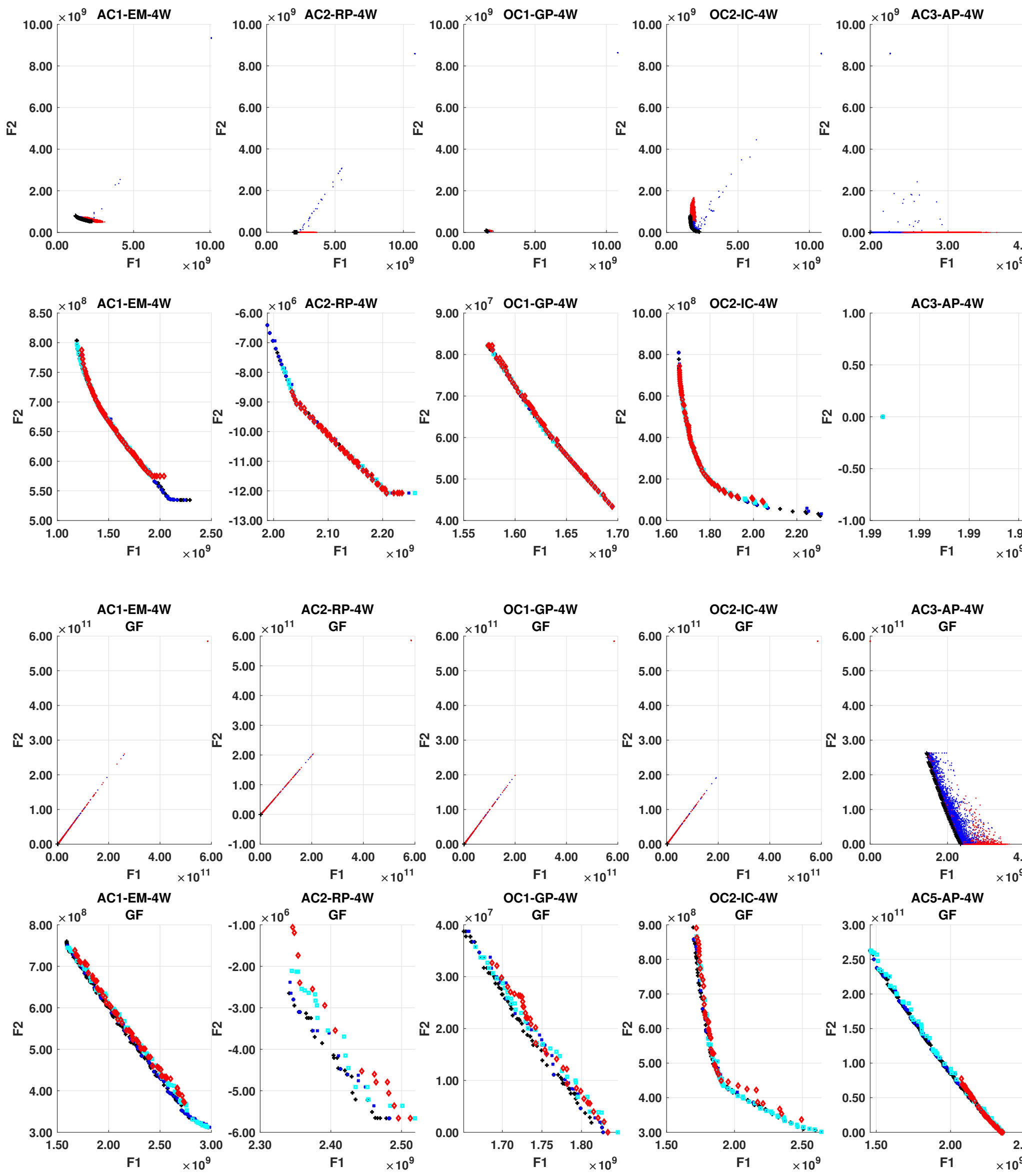


Figure 5.12: The figure presents the serial progress of the HV1 (first and third rows) and HV2 (second and fourth rows) values for the *MOOGEP-(XXX)-4W* (first and second rows) and *MOOGEP-(XXX)-4W-GF* (third and fourth rows) test cases. Each step represents an addition of 100 FES to the FES limit.

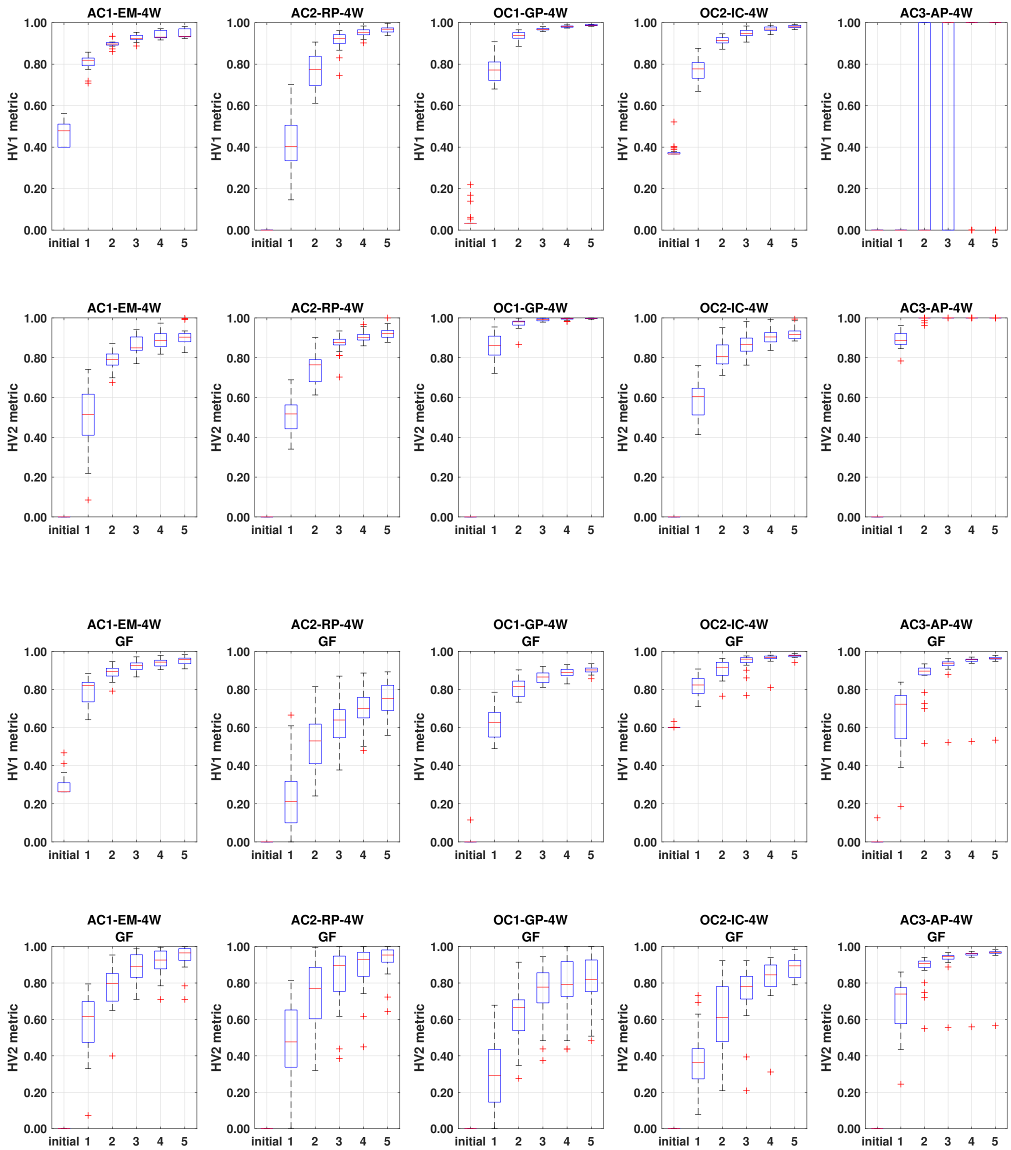


Figure 5.13: The figure presents the values of the RMSE and R^2 metrics for the *MOOGEP-(XXX)-4W* test cases.

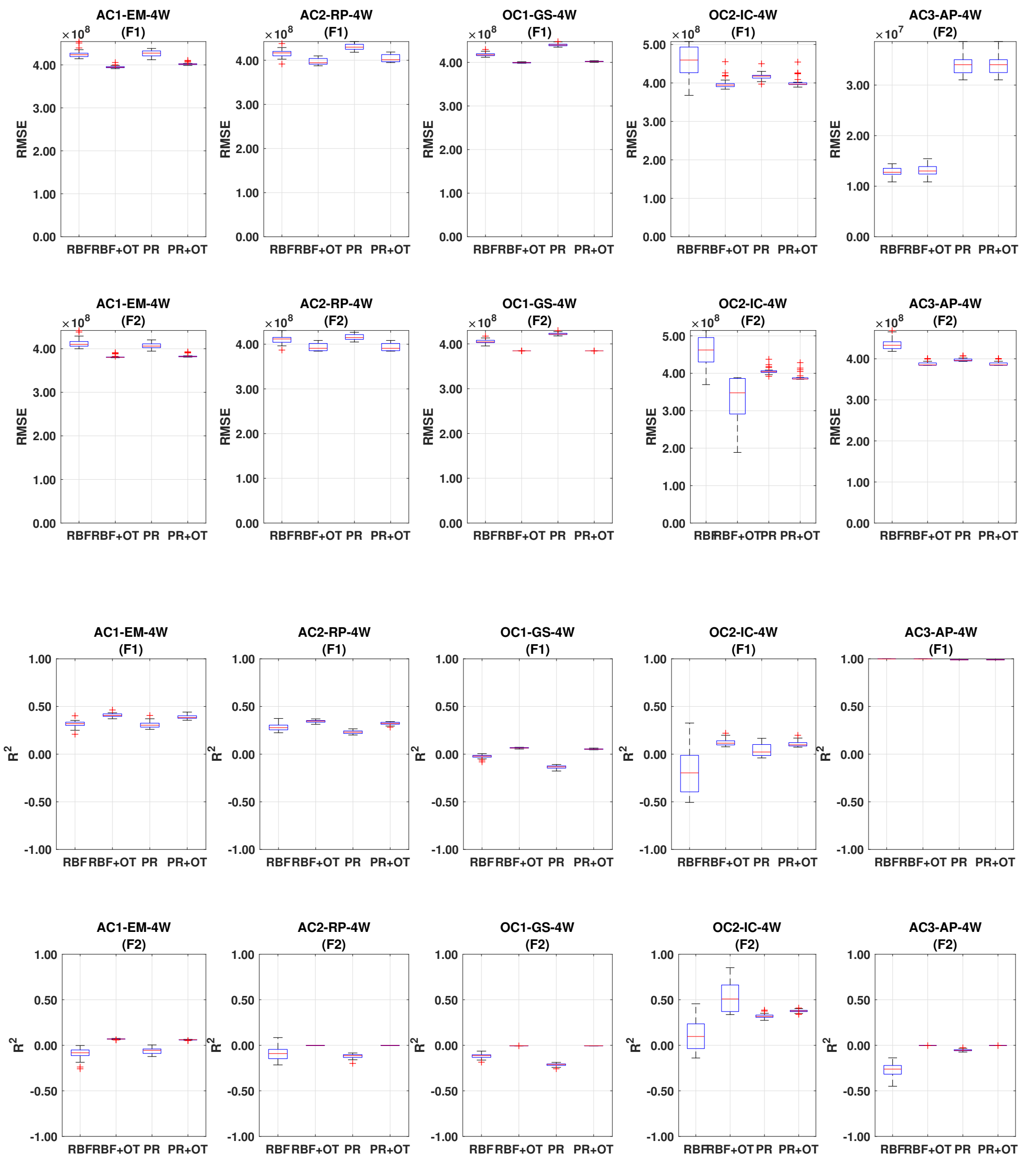
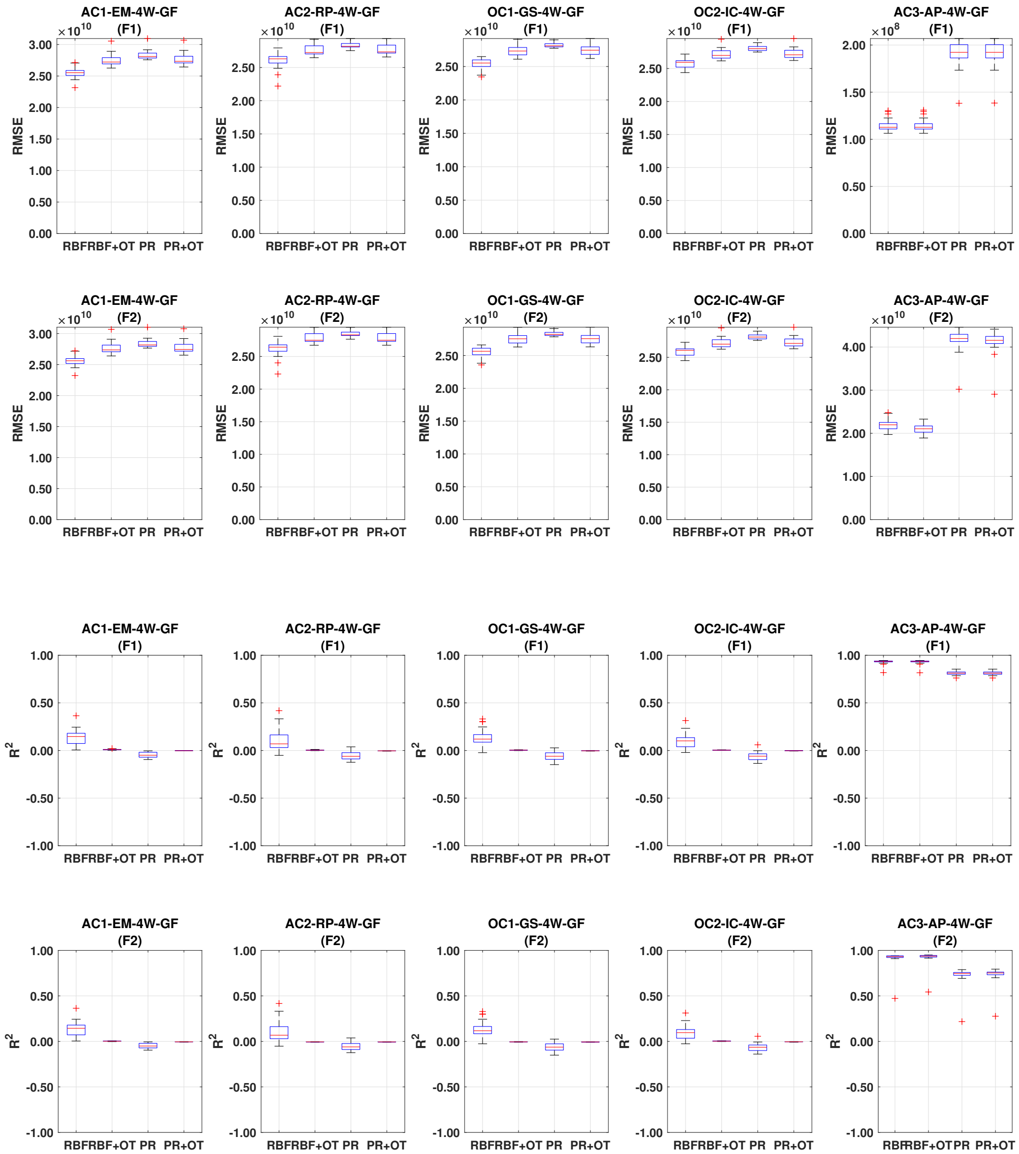


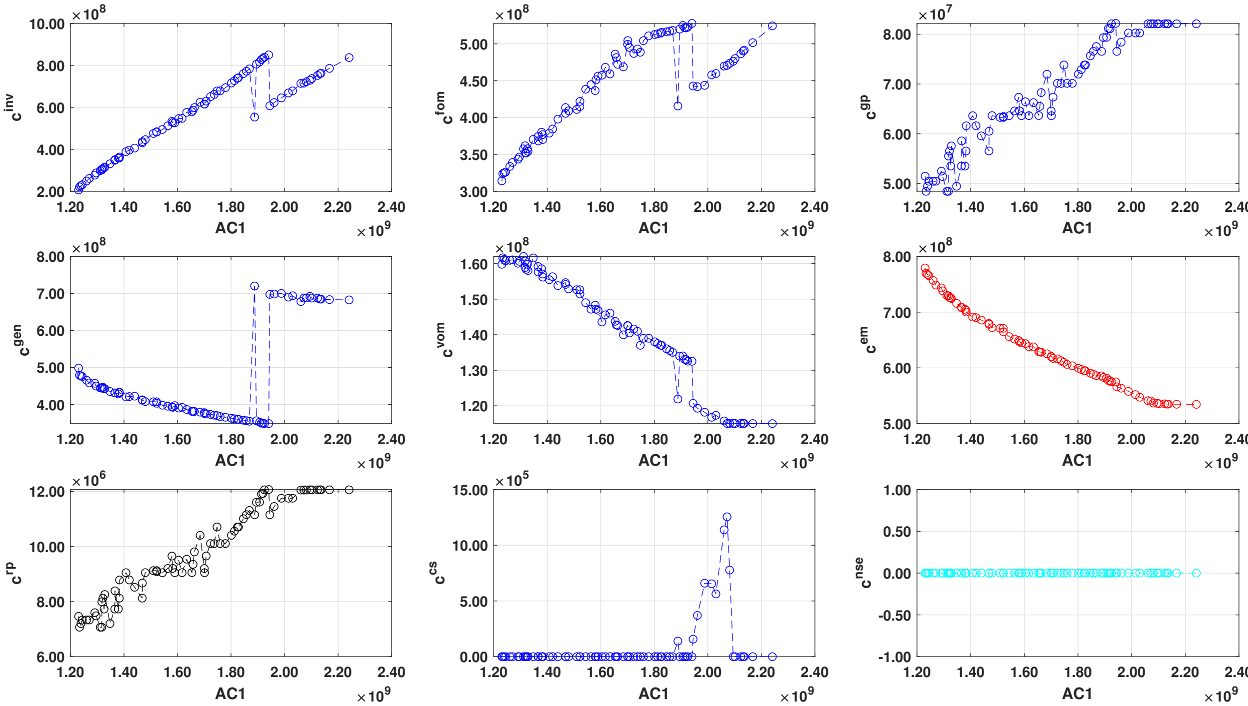
Figure 5.14: The figure presents the values attained for the RMSE and R^2 metrics for the *MOOGEP-(XXX)-4W-GF* test cases.



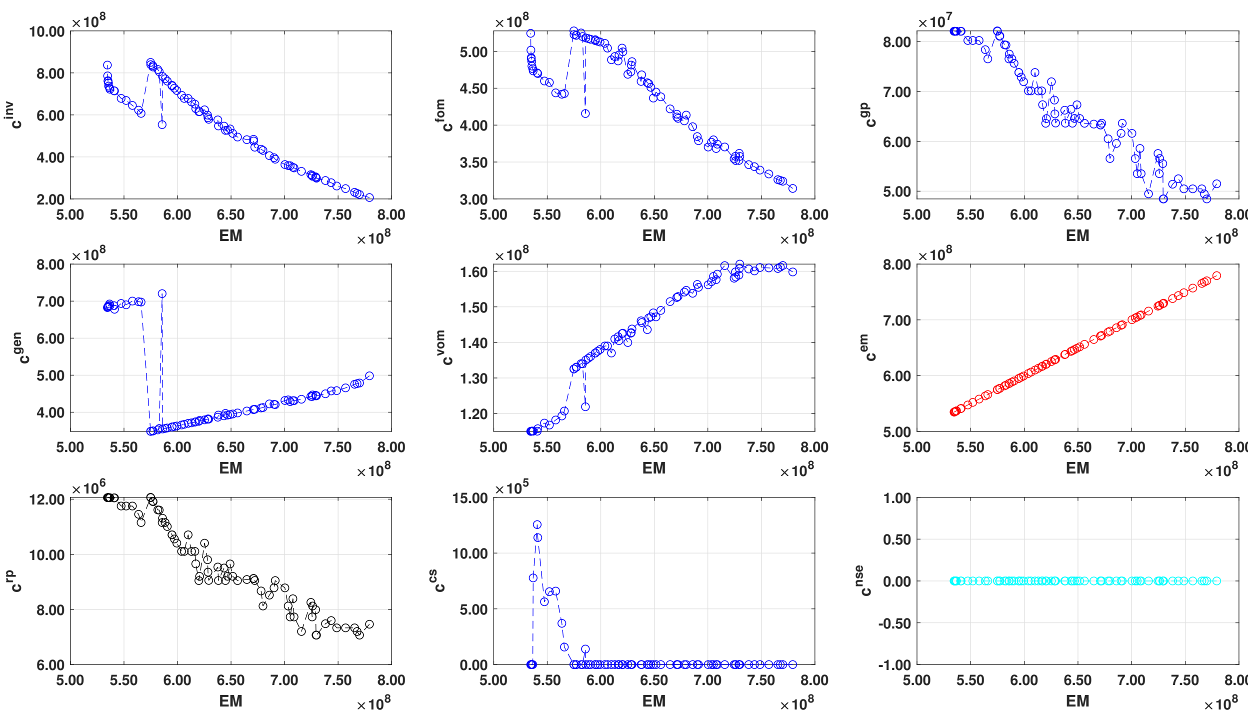
Cost terms and capacity additions for cases including the existing capacity

This Section includes ten figures. In each figure (Figures 5.15 -5.19) the evolution of the cost term values along the attained NDF for one of the five *MOOGEP-(XXX)-4W* test cases is presented. Correspondingly, Figures 5.20 -5.24 present the capacity additions along the attained NDF for each of the five *MOOGEP-(XXX)-4W* test cases.

Figure 5.15: The figure presents the cost term values along the attained NDF for the *MOOGEP-AC1-EM-4W* test case. The colour of the dots indicates that the cost term participates in the AC1 function (blue), in the EM function (red), in none (black) or both functions (cyan). All values are in monetary terms.

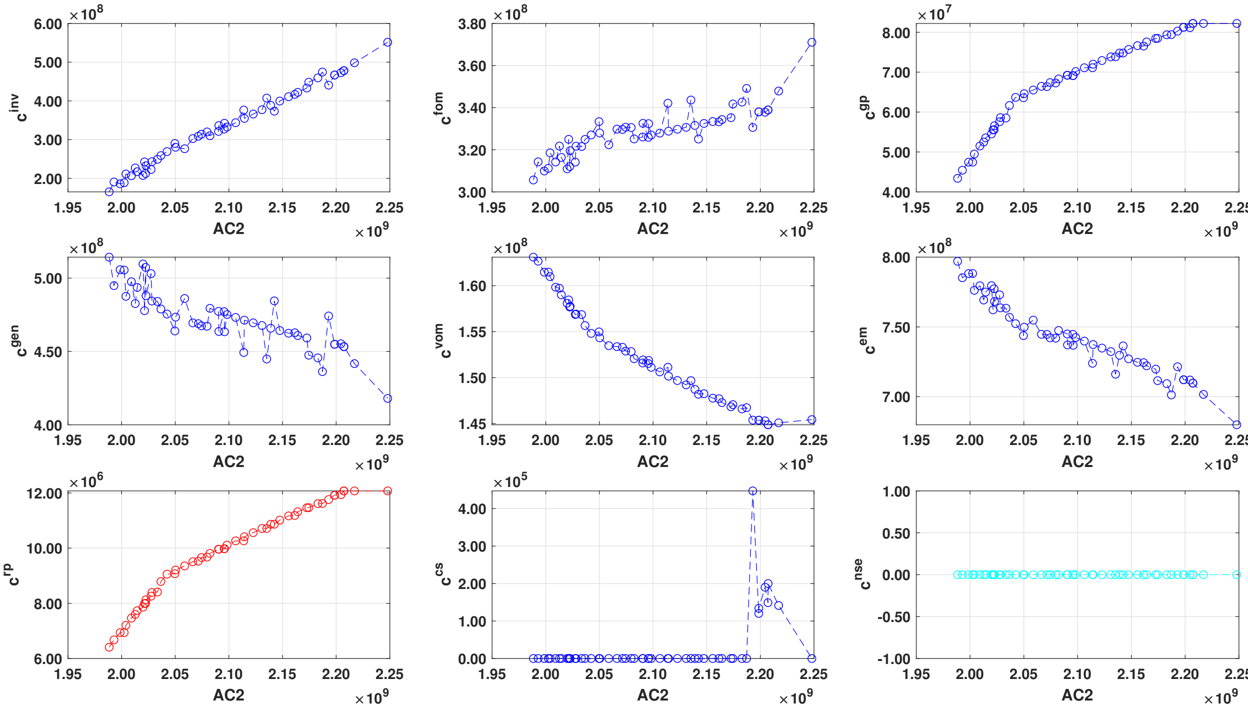


(a) Cost term values for the corresponding AC1 function values of the best attained NDF of a single independent run based on the HV1 metric.

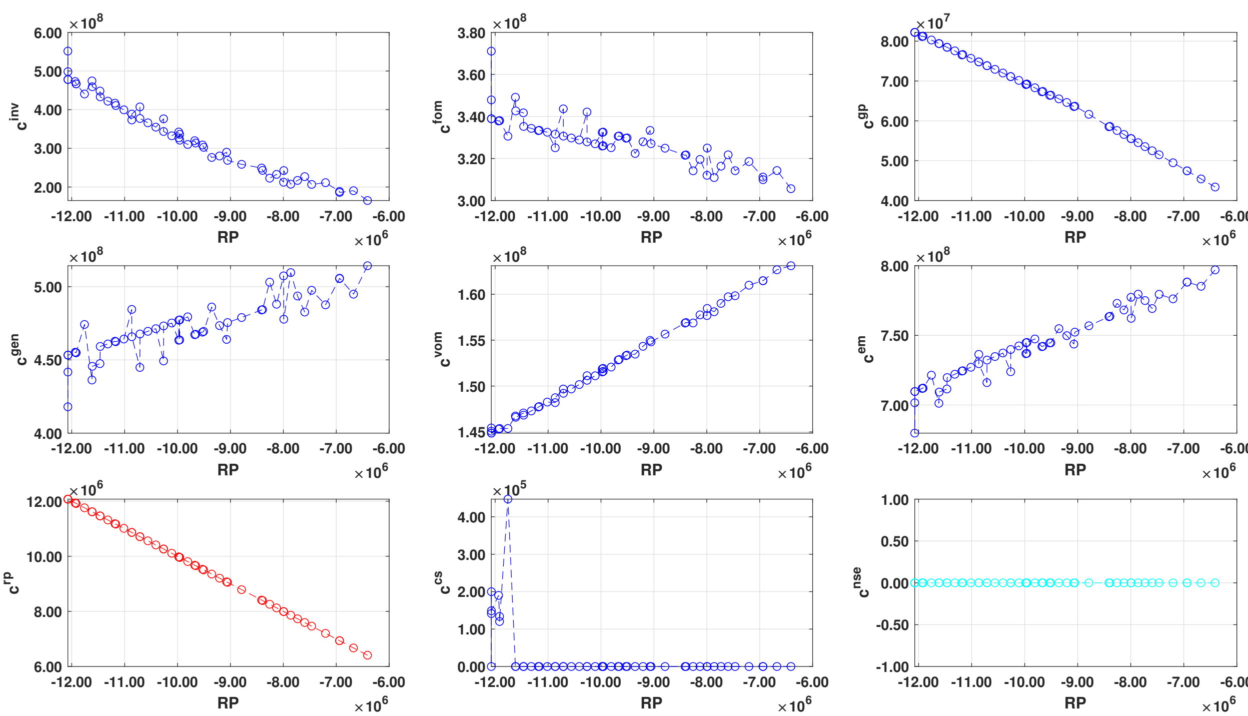


(b) Cost term values for the corresponding EM function values of the best attained NDF of a single independent run based on the HV1 metric.

Figure 5.16: The figure presents the cost term values along the attained NDF for the *MOOGEP-AC2-RP-4W* test case. The colour indicates that the cost term participates in the AC2 function (blue), in the RP function (red), in none (black) or both functions (cyan), respectively. All values are in monetary terms.

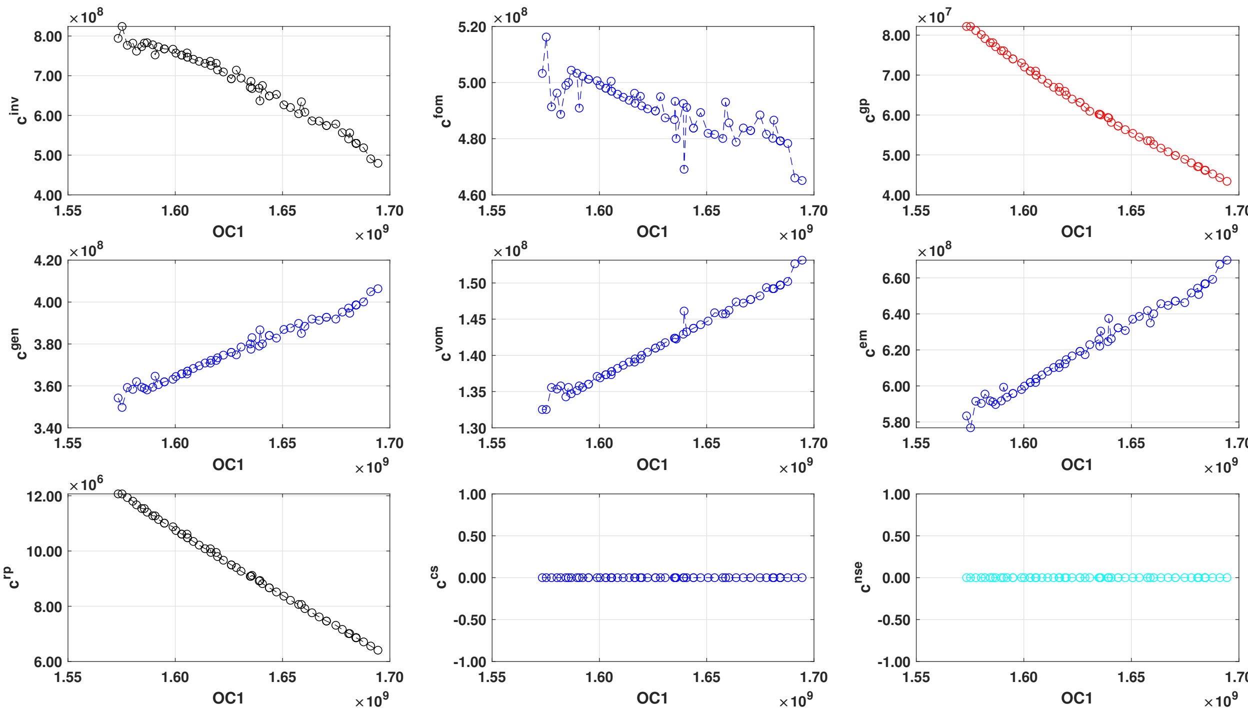


(a) Cost term values for the corresponding AC2 function values of the best attained NDF of a single independent run based on the HV1 metric.

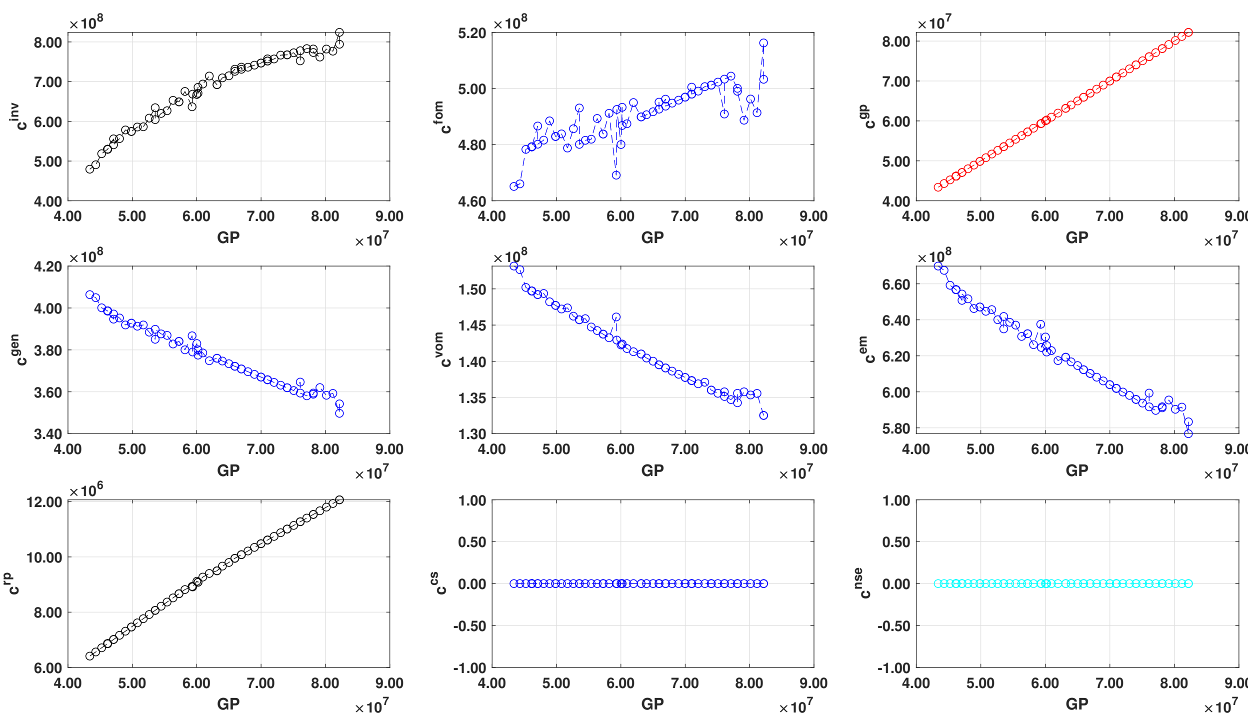


(b) Cost term values for the corresponding RP function values of the best attained NDF of a single independent run based on the HV1 metric.

Figure 5.17: The figure presents the cost term values along the attained NDF for the *MOOGEP-OC1-GP-4W* test case. The colour indicates that the cost term participates in the OC1 function (blue), in the GP function (red), in none (black) or both functions (cyan), respectively. All values are in monetary terms.

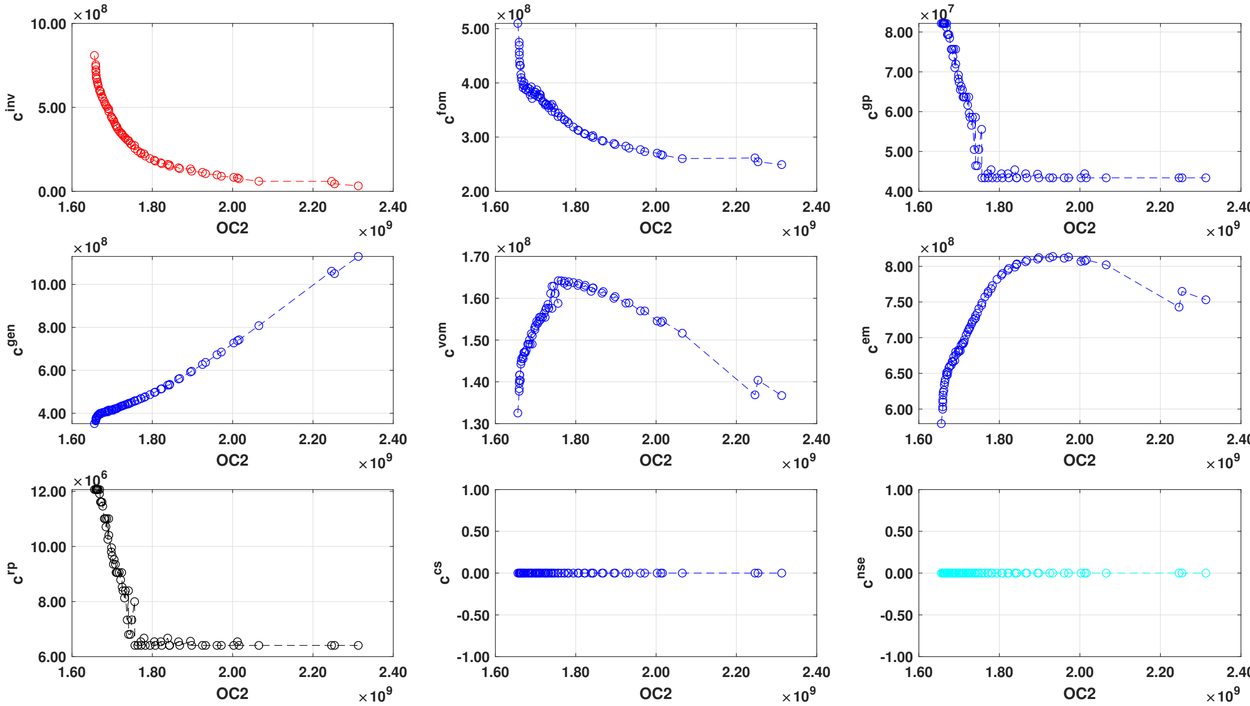


(a) Cost term values for the corresponding OC1 function values of the best attained NDF of a single independent run based on the HV1 metric.

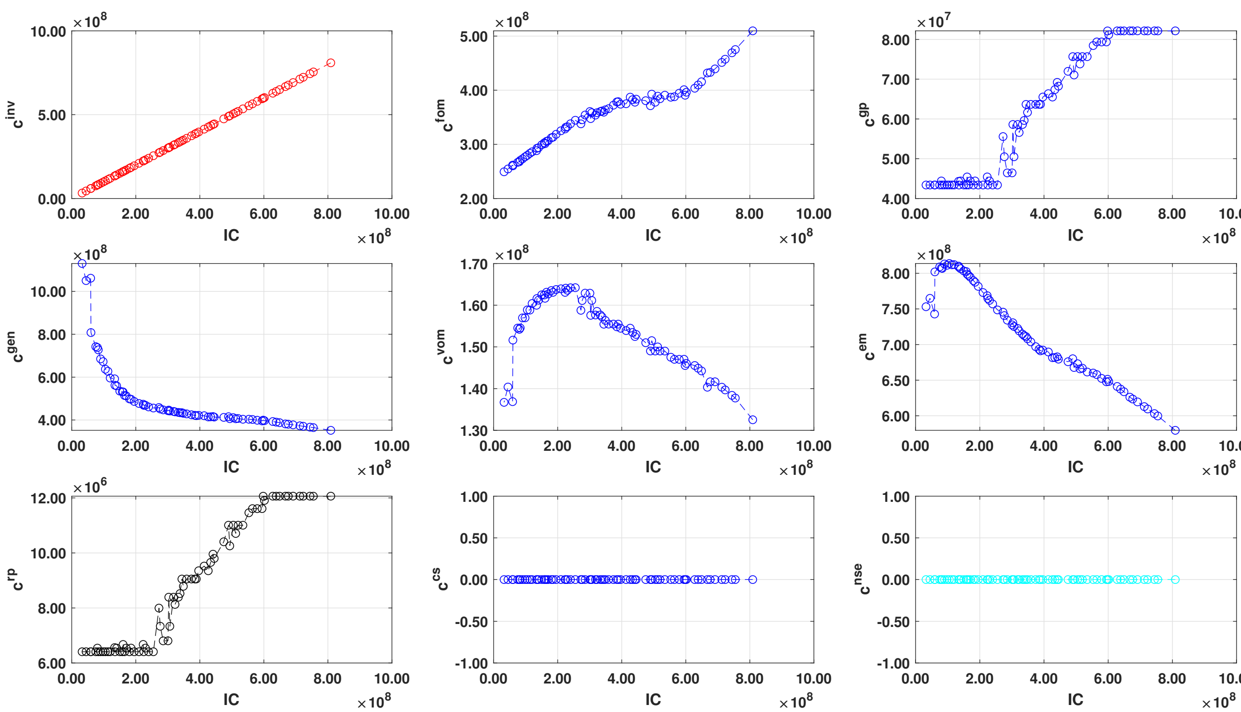


(b) Cost term values for the corresponding GP function values of the best attained NDF of a single independent run based on the HV1 metric.

Figure 5.18: The figure presents the cost term values along the attained NDF for the *MOOGEP-OC2-IC-4W* test case. The colour indicates that the cost term participates in the OC2 function (blue), in the IC function (red), in none (black) or both functions (cyan), respectively. All values are in monetary terms.

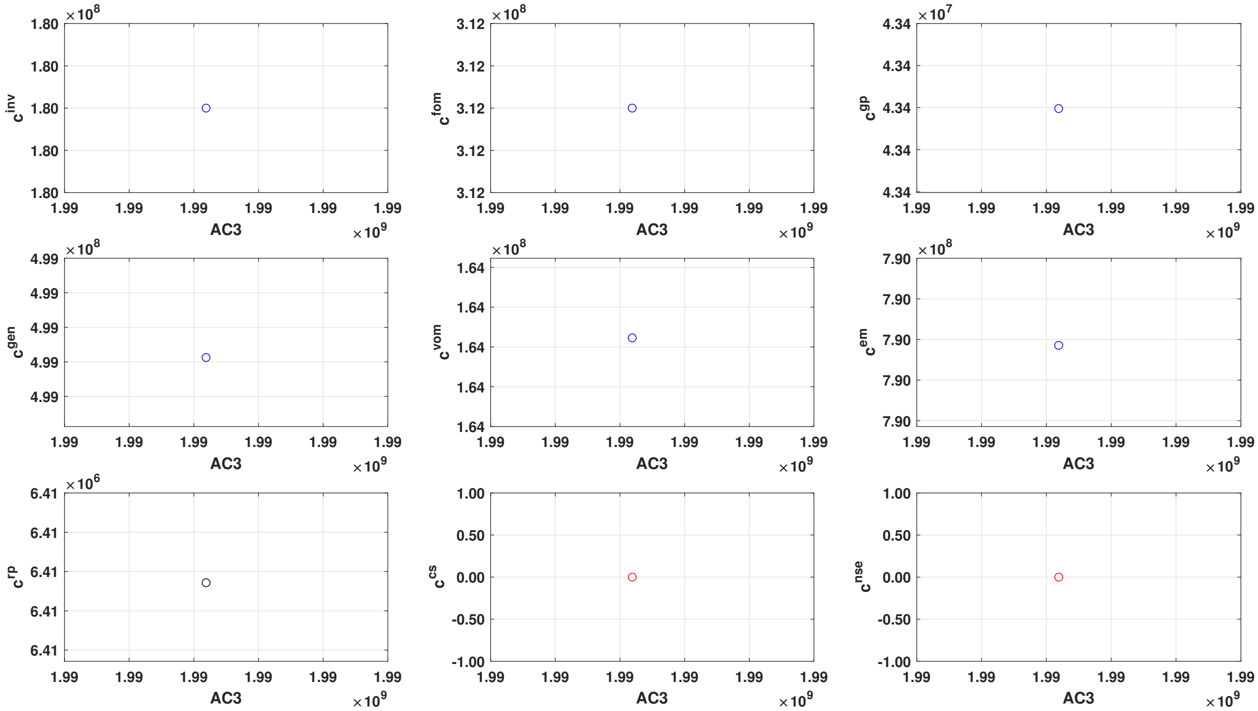


(a) Cost term values for the corresponding OC2 function values of the best attained NDF of a single independent run based on the HV1 metric.

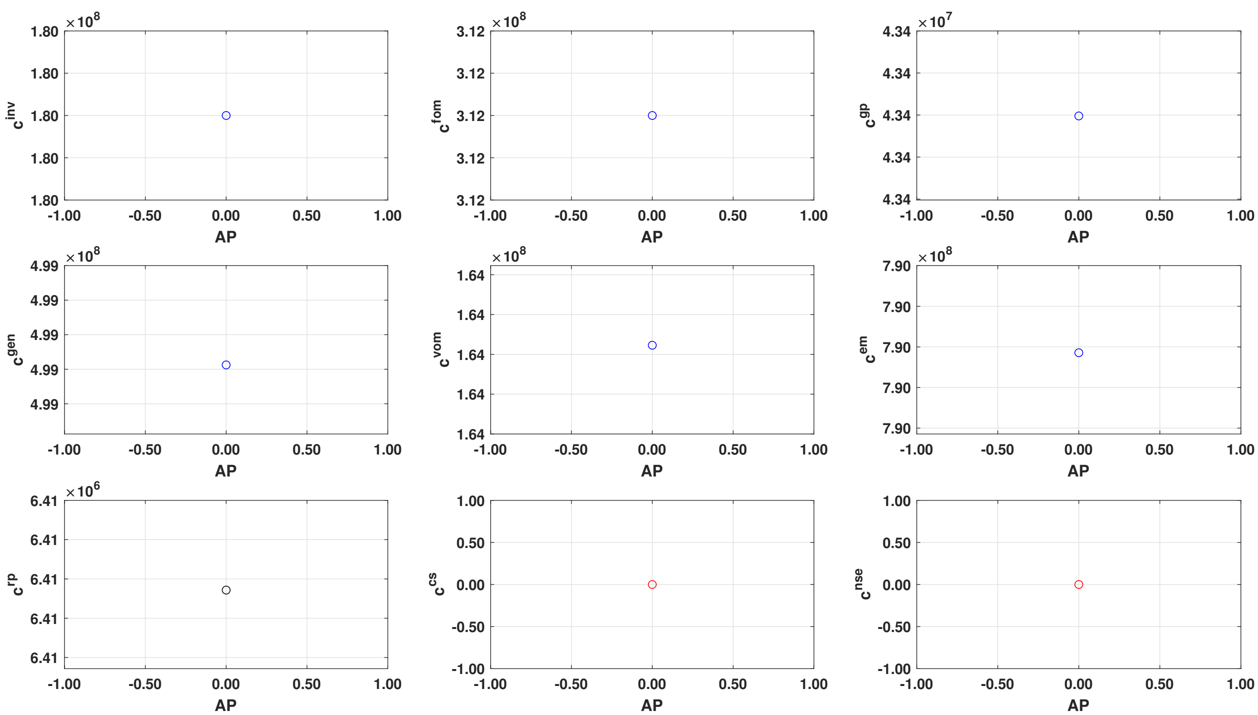


(b) Cost term values for the corresponding IC function values of the best attained NDF of a single independent run based on the HV1 metric.

Figure 5.19: The figure presents the cost term values along the attained NDF for the *MOOGEP-AC3-AP-4W* test case. The colour indicates that the cost term participates in the AC3 function (blue), in the AP function (red), in none (black) or both functions (cyan), respectively. All values are in monetary terms.



(a) Cost term values for the corresponding AC3 function values of the best attained NDF of a single independent run based on the HV1 metric.



(b) Cost term values for the corresponding AP function values of the best attained NDF of a single independent run based on the HV1 metric.

Figure 5.20: The figure presents the decision vector values along the attained NDF for the *MOOGEP-AC1-EM-4W* test case. Blue colour indicates the capacity additions made in area 1 and red is used for area 2. The total capacity additions are presented in black colour.

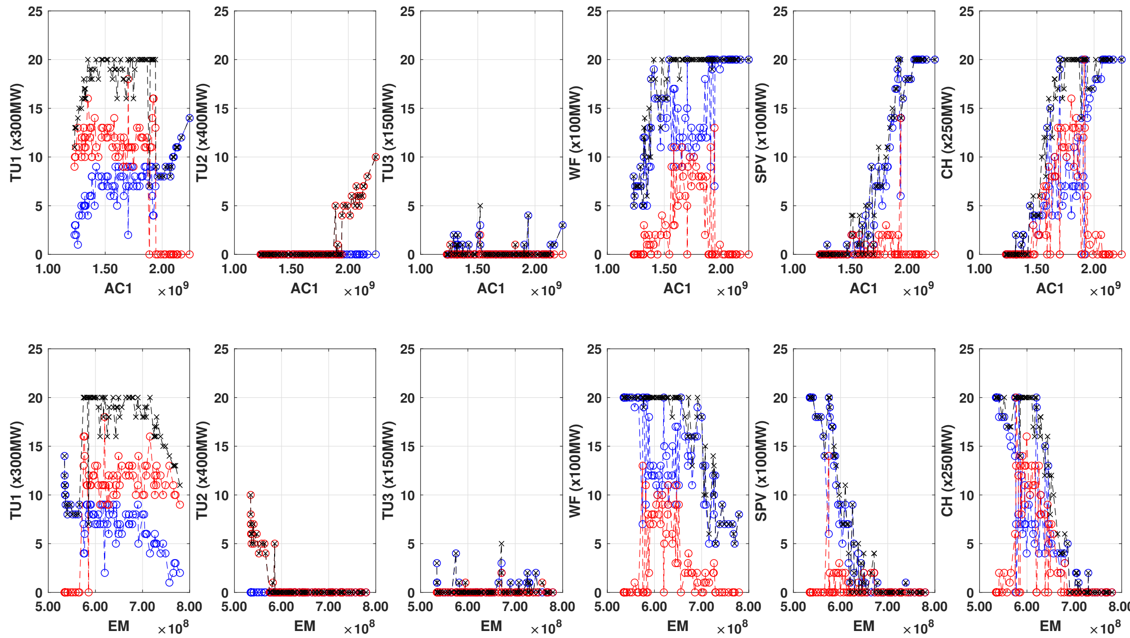


Figure 5.21: The figure presents the decision vector values along the attained NDF for the *MOOGEP-AC2-RP-4W* test case. Blue colour indicates the capacity additions made in area 1 and red is used for area 2. The total capacity additions are presented in black colour.

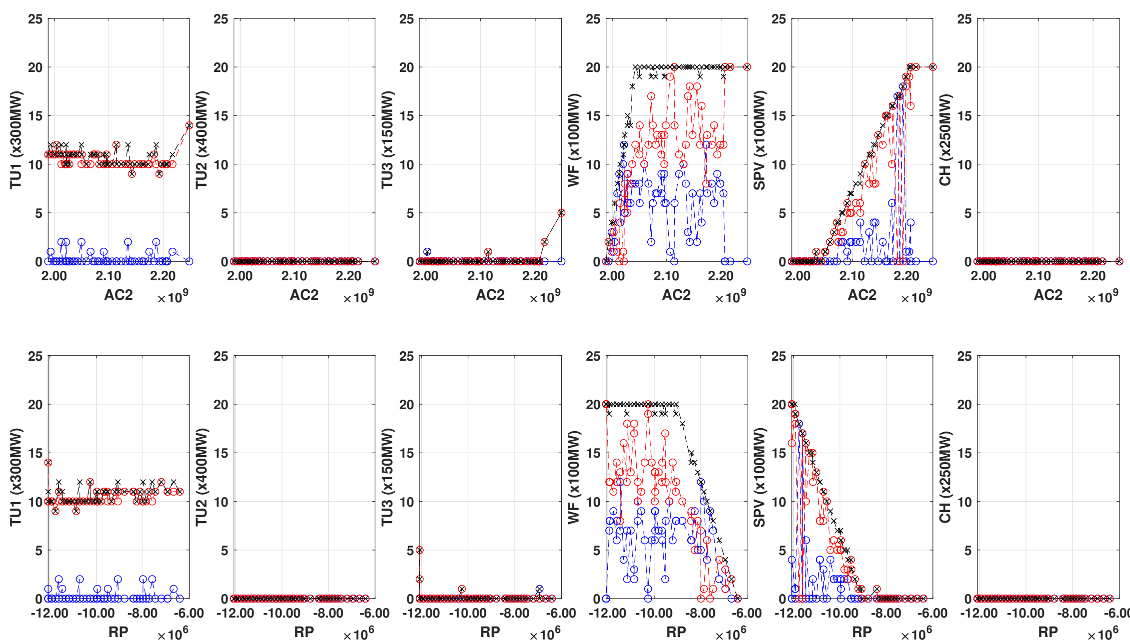


Figure 5.22: The figure presents the decision vector values along the attained NDF for the *MOOGEP-OC1-GP-4W* test case. Blue colour indicates the capacity additions made in area 1 and red is used for area 2. The total capacity additions are presented in black colour.

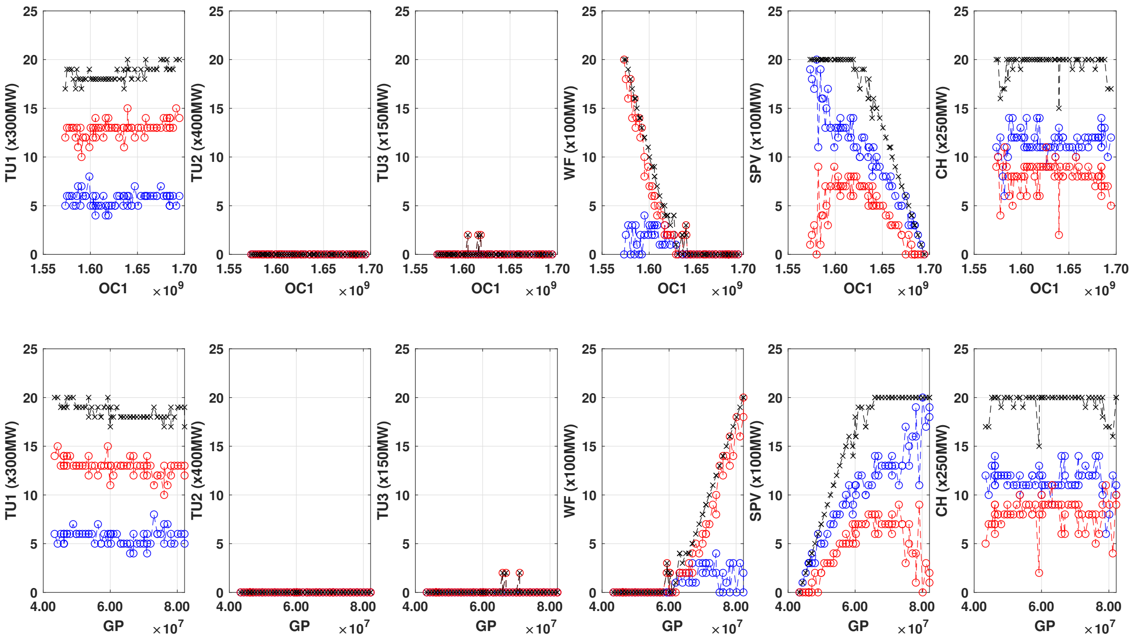


Figure 5.23: The figure presents the decision vector values along the attained NDF for the *MOOGEP-OC2-IC-4W* test case. Blue colour indicates the capacity additions made in area 1 and red is used for area 2. The total capacity additions are presented in black colour.

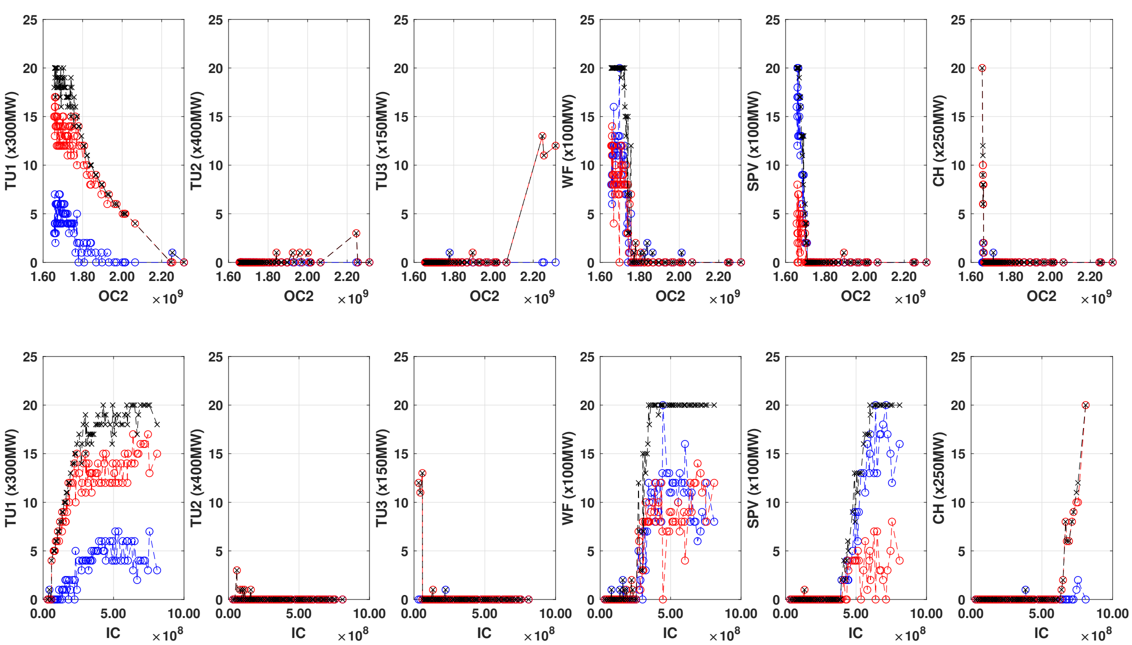
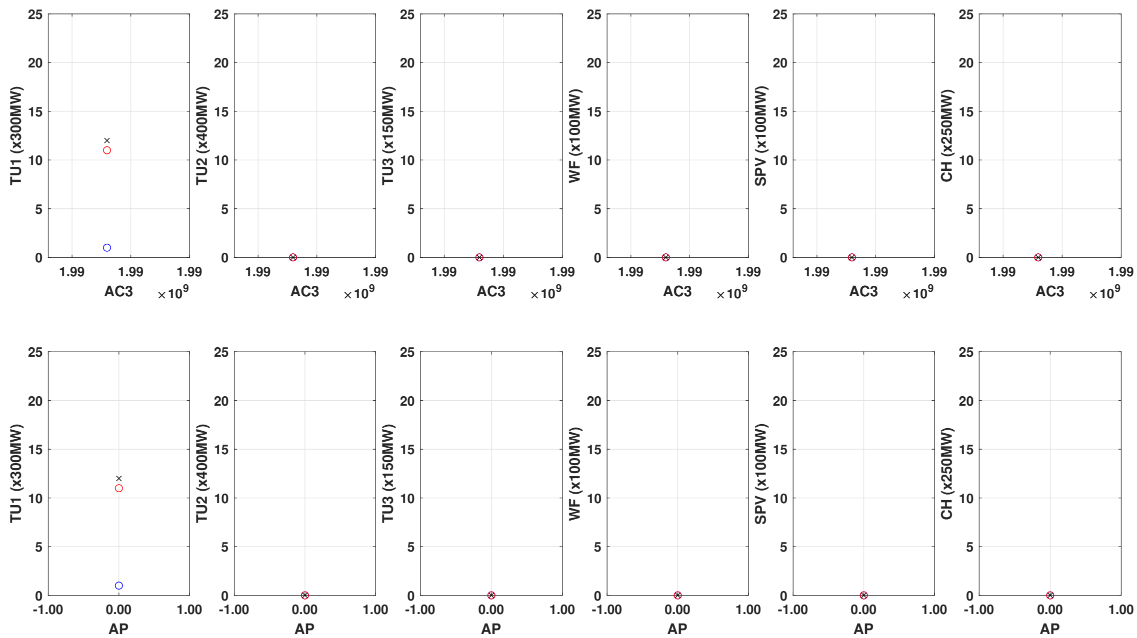


Figure 5.24: The figure presents the decision vector values along the attained NDF for the *MOOGEP-AP3-AP-4W* test case. Blue colour indicates the capacity additions made in area 1 and red is used for area 2. The total capacity additions are presented in black colour.



5.6 Discussion, limitations and future research directions

The derived results indicate that the proposed framework may constitute a promising approach for examining trade-offs within MOO GEP. The result are further discussed in this Section based on the performance of the optimization approach and the GEP model's resolution indicating also limitations and future research directions.

The performance of the approach has been examined revealing satisfactory results. Specifically, it provided a NDF which had been a decent PFA in most cases examined w.r.t. the considered computational restrictions. However, it should be considered that the framework is based on a heuristic approach when analysing its results. For example, a poor UFA impeded the algorithm from providing a decent PFA in the case of highly multimodal functions (i.e. ZDT4).

Moreover, additional limitations may arise due to the included operators. For example, high dimensionality (number of different types of candidate capacity additions and areas) may have an impact on the algorithms performance. Specifically, high dimensionality may impact the AMs and/or the EA (mating operators) performance which both have been reported to suffer from the curse of dimensionality. Therefore, larger values of available Max_FES and Max_Gen should be considered for high dimensional problems. Different ranges have been examined within the numerical experiments, i.e. 10 variables (ZDT4 and ZDT6), 12 variables (all GEP test cases) and 30 variables (ZDT1, ZDT2 and ZDT3).

The optimization approach has been examined using RBF and PR models as AMs, a DE and a GA variant as Variation operators and the Environmental-Selection operator of NSGA-III. The results suggested a satisfactory performance. However, many other AMs and corresponding operators can be identified within the literature and could, possibly, be examined.

Lastly, the inclusions of data points that present extreme values may have some impact on the performance of the algorithm. Such data points arise mostly due to the penalty term. In general, low penalty parameters may result in an augmented function for which the global optimum is located within the penalized region. On the other hand, high values of the penalty term may deteriorate the UFA accuracy within the region which is not penalized. There are some other possible directions to reduce the impact of the aforementioned: i) the penalty term can be set as an additional objective function increasing, however, the number of objective functions, ii) the penalty term could be scaled by a predefined parameter which requires tuning, iii) the search space can be limited, i.e. the $\underline{X}_{a,g}^{inv}$ and $\overline{X}_{a,g}^{inv}$, to exclude a number of unrealistic cases which requires a priori information and the inclusion of bias. The local AMs could be less sensitive to the aforementioned since in some cases extreme data points might be excluded from the dataset.

The GEP model presented considers cost terms commonly included within GEP formulations. A main advantage of this framework is that such cost terms are computed by considering short-term operation (when applicable). Therefore, operational flexibility, to some extent, is considered implicitly. In general, technical (e.g. operational flexibility constraints), temporal (e.g. a year of operation on an hourly resolution) and spatial detail (e.g. two different areas including power exchange limitations among them) have been accounted for by including a SM. The latter is based on the model presented in [29] and included for providing a decent approximation of a UCP in a relatively low computational cost. For the purposes of this study and possibly for the context of long-term planning which involves uncertainty by various other sources, it could be appropriate. However, it remains a problem approximation of a UCP model. Within this context, different SM that may provide the required output (the vector $\mathbf{v}$) for computing the objective functions can be examined within the presented framework.

The results indicate that investment and FO&M cost are an important cost factor for determining a set of optimal alternatives. For example, aggregated cost functions including the aforementioned suggest a limited number of capacity additions in comparison to the ones focusing on the short-term operation. This could imply the importance of incentives to support optimal investments in capacity additions. Overall, a larger number of capacity additions resulted to operating cost reductions (e.g. generation and emission cost) and increased levels of RES penetration. For example, the extreme values of the best non-dominated front attained (Figure 5.11) for the considered operating and investment cost functions suggest that a 96% reduction (from 0.809 b€to 0.032 b€) of the investment cost could result in a nearly 40% increase of operating cost (from 1.656 b€to 2.313 b€).

There are additional cost sources that could be examined, e.g. the cost of transmission losses or the cost of an expansion in the transmission system. Moreover, MOO GEP can consider many other objectives within a multi-period GEP formulation. Limitations arise also due to the assumptions and simplifications made when formulating the objective functions. For example, the GPSC should consider more market conditions and the impact of RES penetration on the system marginal price. For example, the determined objective function does not capture benefits arising from the reduction of operating cost. Therefore, the impact of a policy supporting RES penetration should be assessed both in terms of cost and benefits arising on short-term operation. This is shown by the attain NDF of the *MOOGEP-OC1-GP-4W* case. Furthermore, thermal units fuel and emission cost functions are in general non-linear. Moreover, the output of RES technology groups and CH has been assumed to increase linearly w.r.t. the capacity additions. Also, some data input that are characterized as stochastic have been assumed deterministic, e.g. RES generation and the demand.

Chapter 6

Concluding remarks

This Chapter summarizes the main contributions of this dissertation.

6.1 Conclusions

Chapter 4 presents a MAEA-based approach for SOO multi-period GEP. It aimed at capturing challenges considering the transition towards higher shares of generation by RES and is intended to be used in parallel to other mature and well-established GEP models. Approximation models had been employed within the MAEA-based approach to estimate the output of a SM. The latter serves as a cost indicator of the short-term operation of a power system. In general, the computational time is highly depended to the computational cost of the SM. Assuming a SM adequate to capture operational restrictions, the inclusion of the AMs aimed towards attaining a near-optimal solution by utilizing a limited number of simulations.

The first objective had been to examine the applicability of the developed approach. Therefore, a series of numerical experiments had been carried out. The results for the test case examined suggested that satisfactory results had been attained. More specifically, the targeted installed capacity had been identified within the predefined limit of simulations. In addition, the approach identified consistently a near-optimal solution for the two SM included. Moreover, the assessment of the quality of the approximation attained indicates that the estimated errors, based on the metrics employed, are not very large. These errors were not sufficient to restrict the algorithm from attaining the same near-optimal solution. Moreover, a visual analysis has been also conducted to examine the sensitivity of the operating cost towards the installed capacity revealing that the AMs had capture it in an acceptable manner. These indicated that an increase in investments in capacity additions could have triggered operating cost reductions for the final year of the planning horizon.

Consequently, the developed MAEA-based GEP model could constitute a promising approach to identify near-optimal solutions for SOO multi-period GEP including a SM of the short-term operation of a power system. In addition, employing the AMs to perform a visual analysis could be a simple approach to assess the sensitivity of the operating cost towards the installed capacity of an attained near-optimal solution assuming that a decent approximation has been achieved.

The second objective considered had been to evaluate the possible gain of including problem-customized operators (i.e. the RRH, blk and PO) based on domain-specific knowledge. These had been developed and examined as enhancements to a basic DE algorithm. A series of numerical experiments for DE variants including combinations of the aforementioned had been carried out and their impact on the performance of the algorithm had been evaluated. The results derived suggest that the variants including the RRH and blk or the RRH, blk and PO had been the most competitive. However, the highest gain had been achieved by the inclusion of the RRH.

Therefore, problem-specific operators could be relevant to enhance the performance of an EA applied on a GEP formulation. More importantly, an operator that repairs infeasible solutions generated by an EA population should be prioritized when the optimization problem involves constraint functions for which information regarding the constraint functions is available.

The third objective considered had been to assess the impact of including a SM when increased shares of generation by RES are examined. The importance of technical, temporal and spatial detail of the short-term operation represented in a long-term model had been examined in recent studies (these are discussed in Section 2.2.5). Therefore, the results attained by including different SMs had been compared. Specifically, three cases had been considered. The first, which served as a point of reference, did not include a SM and the operating cost had not been evaluated. Therefore, a near-optimal solution had been attained based on investment and FO&M cost. The second included a SM that captures the operating cost, however, it presented relatively lower technical detail (cost and limitations arising from restriction of the UCP are omitted). The third included a SM based on the CUC framework that exhibited high technical detail w.r.t. the context of long-term planning. The comparison had focused on technical detail. Temporal detail had remained constant among the two SMs, while spatial detail had been neglected. Moreover, non-thermal flexibility providers had been included as investment options in all three cases. Based on the attained results, differences had been observed in the derived near-optimal solution (investment decisions), the derived generating mix and cost of each near-optimal solution when evaluated by both SM. The differences in the generating mix suggested a decrease in the utilization of the flexibility providers when an installed capacity had been evaluated by the SM including less technical detail. This had been attributed to the omission of technical restrictions which had rendered the utilization of flexibility providers less necessary. Consequently, differences in the corresponding operating cost had emerged. These had been adequate to shift the ranking of the installed capacity perceived as more cost-efficient and had led to the differences in the derived investment decisions.

For this reason, the technical detail of the short-term operation represented in a selected SM employed to assess a candidate generating fleet within a multi-stage GEP model could be important to efficiently meet the operating flexibility requirements and examine the available flexibility providers, when such are considered as investment options.

Chapter 4 presents a MAEA-based approach for MOO static GEP. It had aimed at capturing cost trade-offs by assessing flexibility requirements and their impact on the considered cost terms. This had been implemented by including a short-term SM that had been used to provide an indicator of the short-term operation. To address computational restrictions, AMs are employed to provide an estimate of the output of the SM. The approach is developed based on frameworks for surrogate-assisted derivative-free optimization.

The first objective had been to examine the performance of the approach. This had been implemented by carrying out a series of numerical experiments. The aforementioned included a MOO test suite (modified ZDT). The results derived regarding the performance of the optimization approach had been satisfactory given the computational restrictions. More specifically, a decent PFA had been achieved for the ZDT1,ZDT2 and ZDT3 problems. However, the optimization approach had been less successfully on the ZDT4 and ZDT6 problems. This had been attributed to the poor UFA achieved. Particularly, it had been observed a decent UFA for highly multi-modal functions had not been achieving within the predetermined limit of function evaluations that had resulted to a poor PFA. In comparison to a base algorithm, used as a point of reference, an acceleration of the convergence rate had been achieved if a focus on computational restrictions is made.

Consequently, the MOO-MAEA based approach can attain an acceptable PFA within a predefined limit of function evaluations. However, for complex MOO (e.g. including a large number of local optima) achieving a decent UFA is rendered highly challenging. This can restrict the optimization approach from attaining an acceptable PFA.

Apart from the modified ZDT test suite, a series of numerical experiments had been carried out on MOO-GEP variants. Five different MOO-GEP problems had been formulated with different pairs of objective functions. The SM and the considered planning constraint functions are included in each MOO-GEP variant. Moreover, a sixth case had been examined which neglected short-term operation which corresponds to a computationally cheap MOO problem. The optimization approach had been applied on a test case loosely based on a real power sector that had considered the existing capacity and the greenfield case. The numerical experiments for the five MOO-GEP variants had been examined by including two different levels of temporal detail for the shot-term operation. In particular, in the first set a single day and in the second four weeks had been selected. Spatial and technical detail not had been altered in all cases.

The results for the two cases (i.e. existing capacity and greenfield case) regarding the performance of the optimization approach on the computationally cheap cases suggest that an acceptable PFA had been achieved. These had been comparable to a base algorithm, intended for computationally cheap optimization problems, despite the computational restrictions imposed on the optimization approach. Moreover, an acceptable UFA had been achieved based on the metrics employed to assess its quality. By comparing the results for the two cases (existing capacity and greenfield cases), it had been observed that the performance of the approach had not been considerably affected. This is shown by the metrics employed to assess the algorithms performance and the quality of the UFA.

Therefore, the MOO-MAEA based approach can attain an acceptable PFA within a predefined limit of function evaluations for a computationally cheap MOO-GEP model. However, for a computationally cheap model a MOEA could be also employed since the computational restrictions are not binding.

The results for the ten cases (i.e. existing capacity and greenfield cases for each of the five MOO-GEP variants) regarding the performance of the optimization approach on the computationally cheap cases including a SM with limited temporal detail suggest that an acceptable PFA had been achieved. Given the computational restrictions imposed on the optimization approach, the serial progress of the Hypervolume metrics suggest a progressive increase of the PFA accuracy. However, the inclusion of the SM can affect the performance of the optimization approach as a relatively lower UFA had been achieved. This is evident in the greenfield cases as for the cases including existing capacity a relatively decent UFA had been achieved. The impact on the accuracy of the UFA had been attributed to numerical differences arising due to the penalty cost term attained by the SM and included in the objective functions. In particular, the highest levels of UFA accuracy had been attained for the variant including the penalty term as one of the objective functions. Moreover, the distribution of the data point within the objective space indicates that a higher number of data points exhibit installed capacities that are penalized for the greenfield cases in comparison to the considering the existing capacity. Nevertheless, the Hypervolume metrics and the visual analysis of the PFA suggest the decrease in the accuracy of the UFA had not heavily affected the performance of the optimization approach.

Hence, the MOO-MAEA-based approach can attain an acceptable PFA within a predefined limit of function evaluations for a computationally cheap MOO-GEP model including a SM limited temporal detail. However, the inclusion of a SM can affect the accuracy of the UFA. One of the main factors identified for this effect is the CHT used to address infeasible installed capacity, since the penalty terms can exhibit relevantly high cost values.

The final ten cases (i.e. existing capacity and greenfield cases for each of the five MOO-GEP variants) are an illustration of the application of the MOO-MAEA-based approach. These had included a SM with increased temporal detail. Satisfactory results regarding the performance of the optimization approach had been attained w.r.t. the computational restrictions. In particular, the serial progress of the metrics employed for assessing the PFA indicates that the approximation of the PFA could be acceptable given the computational restriction. Also, it suggests that an increase in the available simulations could improve the PFA accuracy in some cases. Moreover, the results had also demonstrated that numerical differences arising due to the inclusion of the penalty cost term had affected the UFA accuracy. This had been observed for all cases except for the ones including the penalty term as a separate objective function. Moreover, the Hypervolume metrics and the visual analysis of the PFA suggest that the decrease in the accuracy of the UFA had not impeded the optimization approach from attaining a decent PFA w.r.t. the computational restrictions.

Therefore, the MOO-MAEA-based approach can constitute a promising approach for an analysis of emerging cost trade-offs in MOO-GEP including a SM for assessing operational flexibility. However, a trade-off among computational burden and PFA accuracy is also identified.

The second objective had been to examine the cost trade-off that had emerged. The results of the five cases considering the existing capacity had been used. The analysis had been based on the derived NDF and under the assumption that it represents a decent PFA (i.e.these cannot be termed as the PF). The results of the analysis highlight the importance of interpreting the results on a cost term level. In particular, given different objectives different conclusions can be drawn regarding the near-optimal capacity additions as anticipated. This had been attributed to the impact of each cost term on each an aggregated function. The first economic-environmental formulation (*MOOGEP-AC1-EM-4W*) suggested that a decrease in emission cost can be associated with higher aggregated cost. The second economic-environmental formulation (*MOOGEP-AC2-RP-4W*) suggested an increase of the level of RES penetration can be associated with higher aggregated cost. Both formulations indicated that investments in GHG-free capacity additions or new efficient thermal installation are required. The third formulation (*MOOGEP-OC1-GP-4W*) had considered an operational and RES policy formulation. Capacity additions are prioritized for RES and CH and thermal base-load units. The results suggested that a reduction of the operating cost requires an increase in the GPSC. This could imply that when assessing possible support schemes for RES (e.g. FIT) the benefits of an increased GPSC could trigger operating cost reductions which could have an impact on the prices observed by the consumers. The forth formulation (*MOOGEP-OC2-IC-4W*) indicated that higher investment cost can lead to operating cost reductions. Thus, highlighting the importance of incentives for capacity additions in cost efficient technology groups. The fifth formulation *MOOGEP-AC3-AP-4W* suggested that a limited number of investments (in base-load units) are required to attain a cost efficient power system. This had been attributed to the impact of specific cost terms (e.g. investment and FO&M cost) on the objective function value which had rendered less efficient more capacity additions.

Therefore, a cost term analysis based on a MOO-GEP approach could identify trade-offs and provide an assessment of capacity additions on a cost term level. Due to the different aspects and objectives that can be considered in a GEP formulation, Many-objective optimization approaches are suggested as a promising extension.

The importance of GEP models with increased technical, temporal, and spatial detail of the short-term operation of a power system for efficiently integrating high shares of RES has been stressed throughout the literature (References [1, 2, 3, 4, 9]). Given the above, this thesis focused on integrating GEP models with SMs for the short-term operation of a power system within a MAEA framework. Considering the computational restrictions and the satisfactory results for the SOO and MOO problems examined, the developed approaches could be promising to facilitate the decision making process and support well established GEP models.

The following Section provides a note on future research directions.

6.2 Future research directions

Limitations have been identified and discussed in previous Section. Based on the aforesaid, a number of future research directions that could be considered are the following:

1. The formulation exploits the modelling flexibility provided by EAs. Different SMs, for the short-term operation of a power system could be examined.

2. There are many objectives involved in the decision making process of a GEP. An extension of the MOO approach to consider a larger number of objectives (Many-objective optimization) of a Multi-stage GEP including SM of the short-term operation could be of interest.

3. The scope of this thesis had not included GEP-TEP formulations. Due to the importance of TEP, especially for the case of high shares of generation by RES, accounting for the expansion of the transmission system with increased spatial detail could be considered to capture also a wider range of market conditions and the role of electricity trading.

4. Computational efficient metrics have been suggested to assess operational flexibility of an installed capacity. It would be of interest to examine if such metrics can be exploited during the search to select new data points.

5. There are other frameworks for surrogate-assisted derivative-free optimization that have not been examined in this thesis. Also, different AMs, EA operators and refining strategies could be possibly examined.

6. Long-term and short-term planning involve uncertainty and a risk assessment could be represented in the presented approaches.

7. Extending the formulation of the GEP model to account for emerging challenges such as the integration of electric vehicles.

6.3 List of publications

The following publications have been made in peer-reviewed international journals and international conferences:

- *Articles in international journals:*
  - Vrionis, C., Tsalavoutis, V., Tolis, A. (2020). A Generation Expansion Planning model for integrating high shares of renewable energy: A Meta-Model Assisted Evolutionary Algorithm approach. Applied Energy, 259, 114085.
  - Tsalavoutis, V. A., Vrionis, C. G., Tolis, A. I. (2018). Optimizing a unit commitment problem using an evolutionary algorithm and a plurality of priority lists. Operational Research, 1-54.
  - Tsalavoutis, V. A., Vrionis, C. G., Tolis, A. I. (2017). Relaxation of quantitative energy objectives on generation expansion planning: A computational and policy study. International Transactions on Electrical Energy Systems, 27(12), e2427.
- *Articles in international conferences:*
  - Tsalavoutis, V., Vrionis, C., Tolis, A. (2019, July). A hybrid multi-objective evolutionary algorithm for economic-environmental generation scheduling. In Proceedings of the Genetic and Evolutionary Computation Conference (pp. 1338-1346).
  - Tsalavoutis, V., Vrionis, C., Tolis, A., (2019, May). Optimization of the hydrothermal generation scheduling problem using an enhanced multi-objective evolutionary algorithm. In 8th International Symposium and 30th National Conference on Operational Research (HELORS 2019).
  - Tsalavoutis, V., Vrionis, C., Tolis, A., Plataniotis, D. (2018, November). A Differential Evolution Approach for the Reliability Constrained Unit Commitment Problem. In 2018 IEEE Symposium Series on Computational Intelligence (SSCI) (pp. 2329-2339). IEEE.
  - Tsalavoutis, V., Vrionis, C., Tolis, A. (2016, December). An enhanced real coded approach for the optimization of the Unit Commitment Problem. In 2016 IEEE Symposium Series on Computational Intelligence (SSCI) (pp. 1-8). IEEE.

Appendices

Appendix A

Simulation model based on Cluster Unit Commitment

This Section presents the SM^{CUC} used to provide an indicator of the short-term operation of a power system. The CUC formulation is adopted which is based on References [7, 8, 191]. The formulation of the cost terms, objective and constraint functions are presented in the following Sections.

A.1 Cost terms - SM

In this Section the cost terms considered in the SM are presented. These are the start-up and shut-down cost, generation cost, RES curtailment cost, load shedding cost and reserve shortage cost and are computed as follows:

1. *Start-up and shut-down cost*: Start-up and shut-down cost is computed as a fixed cost added for each start-up and shut-down of a unit:

$$c^{susd} = \sum_{g}\sum_{h}[u^{su}_{g,h}SUC_g + u^{sd}_{g,h}SDC_g] \tag{A.1}$$

2. *Generation cost*: The generation cost is computed based on fixed and variable cost parameters of each group of generating units:

$$c^{gen} = \sum_{g}\sum_{h}[u^{on}_{g,h}C^{fixed}_g + p_{g,h}C^{var}_g] \tag{A.2}$$

3. *RES curtailment cost*: RES curtailment cost is set as the total curtailed RES generation penalized by a curtailment cost factor:

$$c^{curt} = PC^{curt}\sum_{h}[\epsilon^{curt}_h] \tag{A.3}$$

4. *Load shedding cost*: Load shedding is penalized by a cost factor as follows:

$$c^{ls} = PC^{ls}\sum_{h}[\epsilon^{ls}_h] \tag{A.4}$$

5. *Reserve shortage cost*: Reserve shortage is penalized by a cost factor as follows:

$$c^{rs} = PC^{rs} \sum_{h} [\epsilon_h^{rs}] \tag{A.5}$$

A.2 Objective function - SM

The objective function is set as the minimization of total operating cost. An aggregated objective function is considered which is the sum of start-up and shut-down cost, generation cost, RES curtailment cost, load shedding cost and reserve shortage cost.

$$minimize \;\; oc = c^{susd} + c^{gen} + c^{curt} + c^{ls} + c^{rs} \tag{A.6}$$

A.3 Constraint functions

The adopted formulation considers basic constraints regarding the power balance and reserves of the system, technical limitations of the thermal units, the CH aggregated unit and the HS aggregated unit.

1. *Logical constraints*: The number of on-line units is controlled by the number of units starting-up and shutting-down in previous hours:

$$u_{g,h+1}^{on} = u_{g,h}^{on} - u_{g,h}^{sd} + u_{g,h}^{su}, \forall g, h \in [1, \ldots, hz-1] \tag{A.7}$$

2. *Demand constraints*: The total generation should meet the total demand at each hour of the examined scheduling period:

$$\sum_{g} [p_{g,h}] + p_h^{ch} + p_h^{p} - p_h^{st} + \overline{P}_h^{res} - \epsilon_h^{curt} + \epsilon_h^{ls} = Dem_h, \forall h \tag{A.8}$$

where the total RES output ($\overline{\mathbf{P}}^{res}$) in each hour is computed as follows:

$$\overline{P}_h^{res} = \sum_{t^{res}} [PP_{t^{res},h}^{res} AU_{t^{res}} P_{t^{res}}^{net}], \forall h \tag{A.9}$$

3. *Reserve requirements constraints*: Reserves are assumed to be available for mitigating the impact of possible unit outages and deviations of the actual load from the anticipated one:

$$\epsilon_h^{rs} + \sum_{g} [r_{g,h}] = rr_h, \forall h \tag{A.10}$$

where reserve requirements ($\mathbf{rr}$) are assumed as a percentage of the expected demand and a percentage of the expected RES output:

$$rr_h = RR^{dem}(Dem_h - \epsilon_h^{ls}) + RR^{res}(\overline{P}_h^{res} - \epsilon_h^{curt}), \forall h \tag{A.11}$$

The maximum reserve requirements in each hour ($\overline{\mathbf{RR}}$) can be computed (Eq. A.11) by assuming zeros load shedding and RES curtailment, i.e. $\epsilon_h^{ls} = 0$ and $\epsilon_h^{curt} = 0$. These reserve requirements are a rather simplified formulation. A more detailed consideration of operating reserves has been considered in References [8, 37].

4. *Minimum up time constraints*: A committed unit must remain on-line for a minimum number of time intervals before it can be shut-down:

$$u_{g,h}^{on} \geq \sum_{hh=1}^{MUT_g} [u_{g,h-hh}^{su}], \forall g, h \tag{A.12}$$

5. *Minimum down time constraints*: A unit that has been shut-down must remain off-line for a number of time intervals before it can be brought on-line:

$$AU_g - u_{g,h}^{on} \geq \sum_{hh=1}^{MDT_g} [u_{g,h-hh}^{sd}], \forall g, h \tag{A.13}$$

6. *Ramp rate constraints*: A unit can adjust its power output at a limited rate based on its ramping capabilities. The constraints accounts also for start-up and shut-down limitations:

$$\begin{aligned} p_{g,h+1} + r_{g,h+1} - p_{g,h} \leq RU_g(u_{g,h}^{on} - u_{g,h}^{sd}) \\ + SU_g u_{g,h}^{su} - \underline{P}_g u_{g,h}^{sd} \end{aligned}, \forall h \in [1, \ldots, hz-1], g \tag{A.14}$$

$$\begin{aligned} p_{g,h} + r_{g,h} - p_{g,h+1} \leq RD_g(u_{g,h}^{on} - u_{g,h}^{sd}) \\ + SD_g u_{g,h}^{sd} - \underline{P}_g u_{g,h}^{su} \end{aligned}, \forall h \in [1, \ldots, hz-1], g \tag{A.15}$$

7. *Minimum operating point*: The lower operating point limit of the generating units is set as follows:

$$p_{g,h} \geq u_{g,h}^{on} \underline{P}_g, \forall g, h \tag{A.16}$$

8. *Maximum operating point*: The maximum power limit of the generating units, considering also the start-up and shut-down capabilities, is formulated as follows:

$$p_{g,h+1} + r_{g,h+1} \leq \overline{P}_g u_{g,h+1}^{on} - (\overline{P}_g - SU_g) u_{g,h}^{su}, \forall h \in [1, \ldots, hz-1], g \tag{A.17}$$

$$p_{g,h} + r_{g,h} \leq \overline{P}_g u_{g,h}^{on} - (\overline{P}_g - SD_g) u_{g,h}^{sd}, \forall g, h \tag{A.18}$$

9. *HS energy content constraints*: HS units are modelled by an aggregated unit. The energy content is limited by a minimum and maximum level scaled by the available HS units:

$$\sum_{t^{hs}} [AU_{t^{hs}} P_{t^{hs}}^{net} \underline{E}_{h,t^{hs}}^{hs}] \leq ec_h^{hs} \leq \sum_{t^{hs}} [AU_{t^{hs}} P_{t^{hs}}^{net} \overline{E}_{h,t^{hs}}^{hs}], \forall h \tag{A.19}$$

where the energy content in each hour is computed as follows:

$$ec_{h+1}^{hs} = ec_h^{hs} - \frac{p_{h+1}^{p}}{Eff^{p}} + p_{h+1}^{st} Eff^{st}, \forall h \in [1, \ldots, hz-1], \tag{A.20}$$

10. *CH energy content constraints*: The energy content constraints of the CH aggregated unit are set as follows:

$$ec_{h+1}^{ch} = ec_h^{ch} - \frac{p_{h+1}^{ch}}{Eff^{ch}} + \sum_{t^{ch}} [AU_{t^{ch}} In_{h,t^{ch}}^{ch} \overline{P}_{h,t^{ch}}^{ch}], \forall h \in [1, \ldots, hz-1] \tag{A.21}$$

where the energy content in each hour is set as follows:

$$\sum_{t^{ch}}[AU_{t^{ch}}P^{net}_{t^{ch}}\underline{E}^{ch}_{h,t^{ch}}] \leq ec^{ch}_h \leq \sum_{t^{ch}}[AU_{t^{ch}}P^{net}_{t^{ch}}\overline{E}^{ch}_{h,t^{ch}}], \forall h \tag{A.22}$$

For high levels of CH installed capacity, that are adequate to restrict a feasible solution, excess hydro should also be considered and penalized.

11. *Status constraints*: The non-negative integer variables, representing the status of the thermal units are restricted by the available units within each group:

$$u^{on}_{g,h}, u^{sd}_{g,h}, u^{su}_{g,h} \in \{0, 1, \ldots, AU_g\}, \forall g, h \tag{A.23}$$

12. *HS status constraints*: The HS aggregated unit may not generate and store simultaneously:

$$u^{hs}_h \in \{0, 1\}, \forall h \tag{A.24}$$

13. *Variable upper and lower bounds*:

$$0 \leq p_{g,h} \leq AU_g\overline{P}_g, \forall g, h \tag{A.25}$$

$$0 \leq r_{g,h} \leq AU_g\overline{P}_g, \forall g, h \tag{A.26}$$

$$0 \leq p^{st}_h \leq (1 - u^{hs}_h)\sum_{t^{hs}}[AU_{t^{hs}}\overline{P}^{st}_{h,t^{hs}}], \forall h \tag{A.27}$$

$$0 \leq p^{p}_h \leq u^{hs}_h\sum_{t^{hs}}[AU_{t^{hs}}\overline{P}^{p}_{h,t^{hs}}], \forall h \tag{A.28}$$

$$0 \leq p^{ch}_h \leq \sum_{t^{ch}}[AU_{t^{ch}}\overline{P}^{ch}_{h,t^{ch}}], \forall h \tag{A.29}$$

$$0 \leq \epsilon^{ls}_h \leq Dem_h, \forall h \tag{A.30}$$

$$0 \leq \epsilon^{rs}_h \leq \overline{RR}_h, \forall h \tag{A.31}$$

$$0 \leq \epsilon^{curt}_h \leq \overline{P}^{res}_h, \forall h \tag{A.32}$$

Appendix B

Simulation model based on a simplified Economic Dispatch

This Section presents the SM^{ED}. The formulation of the cost terms, objective and constraint functions are presented in the following Sections.

B.1 Cost terms

The cost terms considered in the SM^{ED} are presented in this Section. These are the generation cost, RES curtailment cost and load shedding cost and are computed as follows:

1. *Generation cost*: The generation cost is computed as the sum of variable cost of each group of generating units:

$$c^{gen} = \sum_g \sum_h [p_{g,h} AVC_g] \tag{B.1}$$

2. *RES curtailment cost*: RES curtailment cost is set as the total curtailed RES generation penalized by a curtailment cost factor:

$$c^{curt} = PC^{curt} \sum_h [\epsilon_h^{curt}] \tag{B.2}$$

3. *Load shedding cost*: Load shedding is penalized by a cost factor as follows:

$$c^{ls} = PC^{ls} \sum_h [\epsilon_h^{ls}] \tag{B.3}$$

B.2 Objective function

The objective function is set as the minimization of total operational cost which considers the sum of generation cost, RES curtailment cost and load shedding cost.

$$minimize \;\; oc = c^{gen} + c^{curt} + c^{ls} + c^{rs} \tag{B.4}$$

B.3 Constraint functions

The following basic constraints regarding the power balance and reserves of the system, technical limitations of the thermal units and the CH and HS aggregated units are included:

1. *Demand constraints*: Eq. A.8.
2. *HS energy content constraints*: Eq. A.19.
3. *CH energy content constraints*: Eq. A.21.
4. *HS status constraints*: A relaxed-integer variable is used to represent the HS aggregated unit's status. Eq. A.24 is modified to:
$$0 \leq u_h^{hs} \leq 1, \forall h \tag{B.5}$$
5. *Variable upper and lower bounds*: Eq. A.25, Eq. A.27, Eq. A.28, Eq. A.29, Eq. A.30 and Eq. A.32.

Appendix C

Data input used for the test case in Chapter 4

The data input (Section 4.4.1) are presented in this Section.

Table C.1: Technical and economic characteristics of the assumed technology groups.

	$\mathbf{P}^{cap\_step}$ [MW]	**INC** [M€/MW]	**FO&M** [€/MW]	$\mathbf{C}^{var}$ [€/MW]	$\mathbf{C}^{fixed}$ [€]	**CT** [years]	**LT** [years]	**CC** [1]	$\mathbf{X}^{max\_inv}$ [1]
TU1	283	1.550	33000	32	2100	4	40	0.80	3
TU2	429	0.850	21000	40	1320	3	25	0.60	2
TU3	150	0.486	12000	66	950	3	20	0.30	3
CH	100	1.200	12000	0	0	5	45	0.50	10
HS	250	1.175	3400	0	0	3	45	0.50	3
WF	100	1.200	15000	0	0	2	25	0.20	15
SPV	100	0.895	13000	0	0	2	25	0.25	15

Table C.2: Technical, economic and operational characteristics of thermal units.

	$\underline{\mathbf{P}}$ [MW]	$\overline{\mathbf{P}}$ [MW]	**MUT** [h]	**MDT** [h]	**SUC** [€]	**SDC** [€]	**RU** [MW]	**RD** [MW]	**SU** [MW]	**SD** [MW]
TU1	155	280	8	8	10000	5000	56	56	211	211
TU2	147	420	4	4	6000	2000	126	126	273	273
TU3	30	150	1	1	1500	500	150	150	150	150

Table C.3: Assumed restrictions on the installed capacity of each technology group. The restrictions on the installed capacity ($\underline{\mathbf{IC}}^{ic}$ and $\overline{\mathbf{IC}}^{ic}$) and on the generation levels ($\underline{\mathbf{IC}}^{g}$ and $\overline{\mathbf{IC}}^{g}$) are assumed to be active starting from the 10^{th} year of the planning horizon.

	$\mathbf{IC}^{ep}$ [GW]	$\underline{\mathbf{IC}}^{ic}$ [GW]	$\overline{\mathbf{IC}}^{ic}$ [GW]	$\underline{\mathbf{IC}}^{g}$ [1]	$\overline{\mathbf{IC}}^{g}$ [1]
TU1	10.0	3.0	6.0	-	-
TU2	10.0	3.0	8.0	-	-
TU3	10.0	0.0	1.5	-	-
CH	10.0	2.5	5.0	-	-
HS	10.0	0.0	4.0	-	-
WF	10.0	-	-	.15	.30
SPV	10.0	-	-	.15	.30

Table C.4: System-related data input.

GrR	RM	RR^{dem}	RR^{res}	PC^{curt}	PC^{ls}	PC^{rs}	DR
[%]	[1]	[1]	[1]	[€/MW]	[€/MW]	[€/MW]	[1]
2	0.1	0.05	0.05	150	10000	10000	0.05

Table C.5: Assumed data input for the HS aggregated unit.

Eff^{p}	Eff^{st}	$\underline{E}^{hs\_init}$	$\underline{E}^{hs\_final}$	$\underline{\mathbf{E}}^{hs}$	$\overline{\mathbf{E}}^{hs}$
[1]	[1]	[1]	[1]	[1]	[1]
0.9	0.8	2.5	2.5	0.5	7

Table C.6: Assumed demand profile for days one to six for the first year of the planning horizon. Demand profiles are scaled based on the growth rate for the successive years.

	D1	D2	D3	D4	D5	D6
	[MW]	[MW]	[MW]	[MW]	[MW]	[MW]
H1	3893.4	3178	3001.6	3687.6	3712.24	3133.2
H2	3796.1	3174.5	2963.8	2916.9	3436.86	3057.6
H3	3603.6	3020.5	2849.7	3276.7	3275.16	3048.5
H4	3434.2	2995.3	2809.1	2965.2	3243.8	3107.3
H5	3384.5	3066.7	2909.2	3280.2	3118.36	3192
H6	3420.2	3403.4	3166.1	3662.4	3201.66	3605.7
H7	3475.5	3952.2	3589.6	3849.3	3558.38	4085.2
H8	3620.4	4291	3886.4	4043.9	3981.74	4426.8
H9	3978.8	4460.4	4049.5	4109	4275.74	4657.8
H10	4390.4	4451.3	4074	4106.9	4381.58	4809.7
H11	4749.5	4409.3	4095	4106.9	4258.1	4834.2
H12	4810.4	4427.5	4123.7	4074.7	3985.66	4818.1
H13	4568.2	4392.5	4097.1	3904.6	3663.24	4535.3
H14	4117.4	4205.6	3914.4	3834.6	3556.42	4253.2
H15	4090.1	4186.7	3862.6	3558.8	3534.86	4066.3
H16	4146.8	4135.6	3723.3	3581.9	3595.62	4079.6
H17	4407.2	4189.5	3686.9	3576.3	3724	4230.1
H18	4823	4778.9	4074	3824.8	3933.72	4282.6
H19	4905.6	4915.4	4501.7	4285.4	4366.88	4384.1
H20	4863.6	4929.4	4512.2	4227.3	4582.48	4559.8
H21	4741.8	4681.6	4294.5	3832.5	4365.9	4171.3
H22	4501	4216.1	3627.4	3609.9	4073.86	3890.6
H23	4308.5	3924.2	3623.9	3287.9	3759.28	3709.3
H24	3978.1	3558.8	3268.3	4025.7	3502.52	3320.8

Table C.7: Assumed demand profile for days seven to twelve for the first year of the planning horizon. Demand profiles are scaled based on the growth rate for the successive years.

	D7 [MW]	D8 [MW]	D9 [MW]	D10 [MW]	D11 [MW]	D12 [MW]
H1	4004.7	4148.9	3658.9	2981.3	2997.4	3297
H2	3850	4010.3	3522.4	2925.3	2910.6	3246.6
H3	3754.8	3929.8	3455.9	2874.9	2832.9	3150
H4	3779.3	3929.8	3461.5	2921.1	2810.5	3087.7
H5	3767.4	3885.7	3572.1	3007.2	2963.8	3142.3
H6	4188.8	4196.5	3838.8	3124.8	3281.6	3504.9
H7	4692.8	4736.2	4314.8	3454.5	3739.4	4215.4
H8	5137.3	5177.2	4700.5	3739.4	4011.7	4623.5
H9	5374.6	5463.5	4892.3	3947.3	4152.4	4967.2
H10	5527.2	5688.9	5012	4095.7	4186	5071.5
H11	5687.5	5893.3	5137.3	4151	4183.9	5120.5
H12	5802.3	6094.9	5250.7	4164.3	4148.2	5179.3
H13	5764.5	6169.8	5140.1	3970.4	4104.1	5106.5
H14	5646.2	6062.7	4900.7	3703.7	3963.4	4844
H15	5559.4	5947.2	4778.9	3598	3882.2	4959.5
H16	5466.3	5843.6	4746	3591	3952.9	5100.2
H17	5399.8	5808.6	4765.6	3728.9	4208.4	5396.3
H18	5229	5688.9	4679.5	4089.4	4643.1	5681.2
H19	5164.6	5632.9	4881.1	4223.1	4708.9	5695.2
H20	5149.9	5859	4824.4	3994.2	4641.7	5610.5
H21	5119.8	5625.2	4412.8	3679.9	4442.2	5303.9
H22	4940.6	5335.4	4227.3	3446.8	4089.4	4863.6
H23	4683	5149.2	4073.3	3345.3	3787.7	4526.2
H24	4292.4	4636.1	3693.2	3026.8	3543.4	4250.4

Table C.8: Assumed SPV profile for days one to six.

	D1 [1]	D2 [1]	D3 [1]	D4 [1]	D5 [1]	D6 [1]
H1	0	0	0	0	0	0
H2	0	0	0	0	0	0
H3	0	0	0	0	0	0
H4	0	0	0	0	0	0
H5	0	0	0	0	0.0024545	0.012273
H6	0	0	0	0.12727	0.058091	0.091364
H7	0.00045455	0.0036364	0.022727	0.30727	0.22541	0.25864
H8	0.058636	0.12636	0.17591	0.65136	0.41359	0.42318
H9	0.24409	0.385	0.38091	0.49409	0.56209	0.51955
H10	0.43182	0.56636	0.55409	0.64955	0.66273	0.57864
H11	0.58864	0.67636	0.68	0.61591	0.68564	0.60455
H12	0.66455	0.72818	0.74636	0.78	0.64268	0.59409
H13	0.63545	0.71409	0.75227	0.61727	0.53264	0.59409
H14	0.53545	0.63409	0.68045	0.56045	0.39191	0.53591
H15	0.33773	0.47909	0.54636	0.53364	0.24095	0.45773
H16	0.087727	0.24545	0.34455	0.33727	0.15341	0.31818
H17	0.00090909	0.031364	0.11045	0.14818	0.081409	0.15545
H18	0	0	0.0027273	0.027727	0.017591	0.047727
H19	0	0	0	0	0	0.0059091
H20	0	0	0	0	0	0
H21	0	0	0	0	0	0
H22	0	0	0	0	0	0
H23	0	0	0	0	0	0
H24	0	0	0	0	0	0

Table C.9: Assumed SPV profile for days one to six.

	D7 [1]	D8 [1]	D9 [1]	D10 [1]	D11 [1]	D12 [1]
H1	0	0	0	0	0	0
H2	0	0	0	0	0	0
H3	0	0	0	0	0	0
H4	0	0	0	0	0	0
H5	0.0022727	0	0	0	0	0
H6	0.082273	0.067273	0.046818	0.010909	0	0
H7	0.24864	0.22273	0.19091	0.15909	0.063636	0.0086364
H8	0.42636	0.40182	0.38045	0.365	0.28318	0.14045
H9	0.585	0.56591	0.55136	0.55227	0.48091	0.31091
H10	0.69591	0.68909	0.66818	0.68182	0.62045	0.45909
H11	0.75364	0.76318	0.72455	0.74955	0.69455	0.54636
H12	0.76273	0.78636	0.73136	0.76227	0.70773	0.57773
H13	0.72318	0.765	0.68455	0.72136	0.66273	0.54636
H14	0.64455	0.69636	0.58545	0.62227	0.55409	0.43864
H15	0.53455	0.57818	0.45182	0.46909	0.37864	0.27045
H16	0.39227	0.42091	0.29318	0.27227	0.17636	0.10182
H17	0.23136	0.24409	0.14045	0.094091	0.025909	0.0036364
H18	0.096364	0.095455	0.032727	0.0036364	0	0
H19	0.017273	0.0081818	0	0	0	0
H20	0	0	0	0	0	0
H21	0	0	0	0	0	0
H22	0	0	0	0	0	0
H23	0	0	0	0	0	0
H24	0	0	0	0	0	0

Table C.10: Assumed WF profile for days one to six.

	D1	D2	D3	D4	D5	D6
H1	0.26157	0.28787	0.34268	0.18013	0.21459	0.14975
H2	0.23526	0.28286	0.35647	0.16447	0.25906	0.15727
H3	0.21772	0.286	0.3643	0.1604	0.30103	0.15602
H4	0.19579	0.28161	0.36116	0.1557	0.31136	0.15163
H5	0.18233	0.2506	0.34049	0.14349	0.30385	0.14975
H6	0.18139	0.21803	0.33485	0.13441	0.30009	0.14192
H7	0.17262	0.20206	0.33454	0.1366	0.28036	0.12877
H8	0.15946	0.18389	0.32421	0.1864	0.25624	0.12
H9	0.14537	0.17544	0.29289	0.12971	0.24246	0.12
H10	0.13315	0.17011	0.27472	0.12908	0.22022	0.12689
H11	0.13096	0.16667	0.27002	0.14725	0.20362	0.13754
H12	0.13253	0.16541	0.27096	0.14725	0.19955	0.14913
H13	0.13284	0.16573	0.27065	0.15915	0.19517	0.14506
H14	0.13691	0.16792	0.2647	0.16541	0.1911	0.13785
H15	0.14318	0.16792	0.26125	0.17011	0.18201	0.14255
H16	0.14224	0.16949	0.2553	0.1698	0.16009	0.15539
H17	0.14192	0.16416	0.25468	0.1723	0.16886	0.18076
H18	0.14662	0.16166	0.26971	0.1864	0.1626	0.17324
H19	0.16353	0.16667	0.26063	0.19517	0.17418	0.1817
H20	0.17732	0.17136	0.25374	0.20018	0.18796	0.18796
H21	0.18076	0.18076	0.2528	0.20331	0.21083	0.20018
H22	0.17105	0.19078	0.23964	0.19298	0.22116	0.21803
H23	0.16855	0.19987	0.22367	0.19454	0.20926	0.21928
H24	0.17011	0.21709	0.22336	0.19172	0.19767	0.19579

Table C.11: Assumed WF profile for days seven to twelve.

	D7 [1]	D8 [1]	D9 [1]	D10 [1]	D11 [1]	D12 [1]
H1	0.12251	0.24998	0.28098	0.13503	0.29915	0.4
H2	0.12282	0.26752	0.29383	0.13785	0.2813	0.37275
H3	0.12438	0.28286	0.3336	0.14224	0.2719	0.35302
H4	0.12501	0.28568	0.3502	0.14036	0.26282	0.32421
H5	0.12877	0.28631	0.36148	0.14349	0.24935	0.32702
H6	0.12783	0.28944	0.37087	0.14662	0.2387	0.33736
H7	0.12626	0.27378	0.35145	0.14694	0.22148	0.30635
H8	0.12783	0.27315	0.34456	0.1366	0.22336	0.29477
H9	0.13065	0.27879	0.34895	0.13221	0.21834	0.2719
H10	0.13566	0.27879	0.34143	0.12846	0.20613	0.24309
H11	0.14067	0.26595	0.32107	0.12877	0.19517	0.24497
H12	0.14506	0.24747	0.31043	0.13159	0.18389	0.24872
H13	0.15132	0.22649	0.2932	0.13315	0.19391	0.25405
H14	0.15007	0.21928	0.28819	0.13566	0.18734	0.25875
H15	0.15289	0.20488	0.28192	0.13472	0.18013	0.25687
H16	0.15727	0.1889	0.25749	0.13409	0.17074	0.23839
H17	0.15758	0.17544	0.24089	0.13221	0.16322	0.26157
H18	0.15477	0.16416	0.22837	0.13065	0.15758	0.24591
H19	0.15664	0.16573	0.22179	0.12846	0.16291	0.22523
H20	0.15477	0.16761	0.22085	0.13002	0.16291	0.20362
H21	0.1579	0.16072	0.23588	0.13253	0.15727	0.18953
H22	0.15477	0.15915	0.23588	0.13566	0.15414	0.1911
H23	0.15696	0.15351	0.22868	0.1413	0.14412	0.18922
H24	0.15383	0.15226	0.23087	0.14224	0.13566	0.18013

Table C.12: Assumed inflows profile for days one to six.

	D1 [1]	D2 [1]	D3 [1]	D4 [1]	D5 [1]	D6 [1]
H1	0.28596	0.34339	0.26198	0.2137	0.15021	0.035637
H2	0.28615	0.34335	0.26178	0.21377	0.14996	0.035577
H3	0.28635	0.34331	0.26157	0.21384	0.14971	0.035517
H4	0.28654	0.34327	0.26137	0.21392	0.14946	0.035457
H5	0.28673	0.34323	0.26117	0.21399	0.14921	0.035397
H6	0.28693	0.3432	0.26096	0.21406	0.14897	0.035337
H7	0.28712	0.34316	0.26076	0.21414	0.14872	0.035277
H8	0.28731	0.34312	0.26056	0.21421	0.14847	0.035217
H9	0.28751	0.34308	0.26035	0.21428	0.14822	0.035157
H10	0.2877	0.34304	0.26015	0.21436	0.14797	0.035097
H11	0.28789	0.343	0.25994	0.21443	0.14772	0.035037
H12	0.28809	0.34297	0.25974	0.2145	0.14748	0.034977
H13	0.28828	0.34293	0.25954	0.21457	0.14723	0.034917
H14	0.28847	0.34289	0.25933	0.21465	0.14698	0.034857
H15	0.28867	0.34285	0.25913	0.21472	0.14673	0.034797
H16	0.28886	0.34281	0.25893	0.21479	0.14648	0.034737
H17	0.28905	0.34278	0.25872	0.21487	0.14623	0.034676
H18	0.28925	0.34274	0.25852	0.21494	0.14599	0.034616
H19	0.28944	0.3427	0.25832	0.21501	0.14574	0.034556
H20	0.28963	0.34266	0.25811	0.21508	0.14549	0.034496
H21	0.28983	0.34262	0.25791	0.21516	0.14524	0.034436
H22	0.29002	0.34258	0.25771	0.21523	0.14499	0.034376
H23	0.29021	0.34255	0.2575	0.2153	0.14474	0.034316
H24	0.29041	0.34251	0.2573	0.21538	0.14449	0.034256

Table C.13: Assumed inflows profile for days seven to twelve.

	D7 [1]	D8 [1]	D9 [1]	D10 [1]	D11 [1]	D12 [1]
H1	0.0275	0.10843	0.13788	0.12435	0.16142	0.19413
H2	0.027537	0.10861	0.13778	0.12441	0.16146	0.19419
H3	0.027574	0.10879	0.13768	0.12448	0.1615	0.19424
H4	0.027612	0.10897	0.13758	0.12454	0.16154	0.19429
H5	0.027649	0.10915	0.13748	0.1246	0.16157	0.19435
H6	0.027686	0.10933	0.13738	0.12466	0.16161	0.1944
H7	0.027723	0.10951	0.13728	0.12472	0.16165	0.19445
H8	0.02776	0.10969	0.13718	0.12479	0.16169	0.19451
H9	0.027798	0.10987	0.13708	0.12485	0.16172	0.19456
H10	0.027835	0.11005	0.13698	0.12491	0.16176	0.19461
H11	0.027872	0.11023	0.13688	0.12497	0.1618	0.19467
H12	0.027909	0.11041	0.13678	0.12504	0.16184	0.19472
H13	0.027946	0.11059	0.13668	0.1251	0.16187	0.19477
H14	0.027983	0.11077	0.13658	0.12516	0.16191	0.19482
H15	0.028021	0.11095	0.13648	0.12522	0.16195	0.19488
H16	0.028058	0.11113	0.13638	0.12528	0.16199	0.19493
H17	0.028095	0.11131	0.13628	0.12535	0.16203	0.19498
H18	0.028132	0.11149	0.13618	0.12541	0.16206	0.19504
H19	0.028169	0.11167	0.13608	0.12547	0.1621	0.19509
H20	0.028207	0.11185	0.13598	0.12553	0.16214	0.19514
H21	0.028244	0.11203	0.13588	0.12559	0.16218	0.1952
H22	0.028281	0.11221	0.13578	0.12566	0.16221	0.19525
H23	0.028318	0.11239	0.13568	0.12572	0.16225	0.1953
H24	0.028355	0.11257	0.13558	0.12578	0.16229	0.19536

Table C.14: Assumed parameters for initial and final CH energy content.

	D1 [1]	D2 [1]	D3 [1]	D4 [1]	D5 [1]	D6 [1]
Initial	714.4713	844.3752	958.041	941.8031	958.041	893.0891
Final	718.6617	848.4347	957.5172	942.3443	955.9458	888.759
	D7 [1]	D8 [1]	D9 [1]	D10 [1]	D11 [1]	D12 [1]
Initial	763.1852	633.2814	617.0434	600.8054	617.0434	600.8054
Final	758.9948	632.7576	616.5021	601.3292	616.5021	604.472

Table C.15: Assumed parameters for the hourly maximum CH energy content of days one to six.

	D1 [1]	D2 [1]	D3 [1]	D4 [1]	D5 [1]	D6 [1]
H1	1071.9688	1266.8165	1437.0288	1412.7384	1436.9306	1339.363
H2	1072.2307	1267.0702	1436.9961	1412.7723	1436.7997	1339.0924
H3	1072.4926	1267.3239	1436.9634	1412.8061	1436.6687	1338.8218
H4	1072.7545	1267.5776	1436.9306	1412.8399	1436.5378	1338.5511
H5	1073.0165	1267.8313	1436.8979	1412.8737	1436.4068	1338.2805
H6	1073.2784	1268.0851	1436.8651	1412.9076	1436.2759	1338.0099
H7	1073.5403	1268.3388	1436.8324	1412.9414	1436.1449	1337.7392
H8	1073.8022	1268.5925	1436.7997	1412.9752	1436.014	1337.4686
H9	1074.0641	1268.8462	1436.7669	1413.0091	1435.883	1337.198
H10	1074.326	1269.0999	1436.7342	1413.0429	1435.7521	1336.9273
H11	1074.5879	1269.3536	1436.7015	1413.0767	1435.6211	1336.6567
H12	1074.8498	1269.6074	1436.6687	1413.1105	1435.4902	1336.3861
H13	1075.1117	1269.8611	1436.636	1413.1444	1435.3592	1336.1154
H14	1075.3736	1270.1148	1436.6032	1413.1782	1435.2283	1335.8448
H15	1075.6355	1270.3685	1436.5705	1413.212	1435.0973	1335.5742
H16	1075.8974	1270.6222	1436.5378	1413.2459	1434.9664	1335.3035
H17	1076.1593	1270.876	1436.505	1413.2797	1434.8354	1335.0329
H18	1076.4212	1271.1297	1436.4723	1413.3135	1434.7044	1334.7623
H19	1076.6831	1271.3834	1436.4396	1413.3474	1434.5735	1334.4916
H20	1076.945	1271.6371	1436.4068	1413.3812	1434.4425	1334.221
H21	1077.2069	1271.8908	1436.3741	1413.415	1434.3116	1333.9504
H22	1077.4688	1272.1446	1436.3413	1413.4488	1434.1806	1333.6797
H23	1077.7307	1272.3983	1436.3086	1413.4827	1434.0497	1333.4091
H24	1077.9926	1272.652	1436.2759	1413.5165	1433.9187	1333.1385

Table C.16: Assumed parameters for the hourly maximum CH energy content of days seven to twelve.

	D7 [1]	D8 [1]	D9 [1]	D10 [1]	D11 [1]	D12 [1]
H1	1144.516	949.8893	925.5313	901.2408	925.5313	901.4373
H2	1144.2541	949.8566	925.4974	901.2736	925.4974	901.6664
H3	1143.9922	949.8238	925.4636	901.3063	925.4636	901.8956
H4	1143.7303	949.7911	925.4298	901.3391	925.4298	902.1248
H5	1143.4683	949.7584	925.3959	901.3718	925.3959	902.3539
H6	1143.2064	949.7256	925.3621	901.4045	925.3621	902.5831
H7	1142.9445	949.6929	925.3283	901.4373	925.3283	902.8123
H8	1142.6826	949.6602	925.2944	901.47	925.2944	903.0414
H9	1142.4207	949.6274	925.2606	901.5027	925.2606	903.2706
H10	1142.1588	949.5947	925.2268	901.5355	925.2268	903.4998
H11	1141.8969	949.5619	925.193	901.5682	925.193	903.7289
H12	1141.635	949.5292	925.1591	901.601	925.1591	903.9581
H13	1141.3731	949.4965	925.1253	901.6337	925.1253	904.1873
H14	1141.1112	949.4637	925.0915	901.6664	925.0915	904.4164
H15	1140.8493	949.431	925.0576	901.6992	925.0576	904.6456
H16	1140.5874	949.3983	925.0238	901.7319	925.0238	904.8747
H17	1140.3255	949.3655	924.99	901.7646	924.99	905.1039
H18	1140.0636	949.3328	924.9562	901.7974	924.9562	905.3331
H19	1139.8017	949.3	924.9223	901.8301	924.9223	905.5622
H20	1139.5398	949.2673	924.8885	901.8629	924.8885	905.7914
H21	1139.2779	949.2346	924.8547	901.8956	924.8547	906.0206
H22	1139.016	949.2018	924.8208	901.9283	924.8208	906.2497
H23	1138.7541	949.1691	924.787	901.9611	924.787	906.4789
H24	1138.4922	949.1363	924.7532	901.9938	924.7532	906.7081

Table C.17: Assumed parameters for the hourly minimum CH energy content of days one to six.

	D1 [1]	D2 [1]	D3 [1]	D4 [1]	D5 [1]	D6 [1]
H1	357.3229	422.2722	479.0096	470.9128	478.9769	446.4543
H2	357.4102	422.3567	478.9987	470.9241	478.9332	446.3641
H3	357.4975	422.4413	478.9878	470.9354	478.8896	446.2739
H4	357.5848	422.5259	478.9769	470.9466	478.8459	446.1837
H5	357.6722	422.6104	478.966	470.9579	478.8023	446.0935
H6	357.7595	422.695	478.955	470.9692	478.7586	446.0033
H7	357.8468	422.7796	478.9441	470.9805	478.715	445.9131
H8	357.9341	422.8642	478.9332	470.9917	478.6713	445.8229
H9	358.0214	422.9487	478.9223	471.003	478.6277	445.7327
H10	358.1087	423.0333	478.9114	471.0143	478.584	445.6424
H11	358.196	423.1179	478.9005	471.0256	478.5404	445.5522
H12	358.2833	423.2025	478.8896	471.0368	478.4967	445.462
H13	358.3706	423.287	478.8787	471.0481	478.4531	445.3718
H14	358.4579	423.3716	478.8677	471.0594	478.4094	445.2816
H15	358.5452	423.4562	478.8568	471.0707	478.3658	445.1914
H16	358.6325	423.5407	478.8459	471.082	478.3221	445.1012
H17	358.7198	423.6253	478.835	471.0932	478.2785	445.011
H18	358.8071	423.7099	478.8241	471.1045	478.2348	444.9208
H19	358.8944	423.7945	478.8132	471.1158	478.1912	444.8305
H20	358.9817	423.879	478.8023	471.1271	478.1475	444.7403
H21	359.069	423.9636	478.7914	471.1383	478.1039	444.6501
H22	359.1563	424.0482	478.7804	471.1496	478.0602	444.5599
H23	359.2436	424.1328	478.7695	471.1609	478.0166	444.4697
H24	359.3309	424.2173	478.7586	471.1722	477.9729	444.3795

Table C.18: Assumed parameters for the hourly minimum CH energy content of days seven to twelve.

	D7 [1]	D8 [1]	D9 [1]	D10 [1]	D11 [1]	D12 [1]
H1	381.5053	316.6298	308.5104	300.4136	308.5104	300.4791
H2	381.418	316.6189	308.4991	300.4245	308.4991	300.5555
H3	381.3307	316.6079	308.4879	300.4354	308.4879	300.6319
H4	381.2434	316.597	308.4766	300.4464	308.4766	300.7083
H5	381.1561	316.5861	308.4653	300.4573	308.4653	300.7846
H6	381.0688	316.5752	308.454	300.4682	308.454	300.861
H7	380.9815	316.5643	308.4428	300.4791	308.4428	300.9374
H8	380.8942	316.5534	308.4315	300.49	308.4315	301.0138
H9	380.8069	316.5425	308.4202	300.5009	308.4202	301.0902
H10	380.7196	316.5316	308.4089	300.5118	308.4089	301.1666
H11	380.6323	316.5206	308.3977	300.5227	308.3977	301.243
H12	380.545	316.5097	308.3864	300.5337	308.3864	301.3194
H13	380.4577	316.4988	308.3751	300.5446	308.3751	301.3958
H14	380.3704	316.4879	308.3638	300.5555	308.3638	301.4721
H15	380.2831	316.477	308.3525	300.5664	308.3525	301.5485
H16	380.1958	316.4661	308.3413	300.5773	308.3413	301.6249
H17	380.1085	316.4552	308.33	300.5882	308.33	301.7013
H18	380.0212	316.4443	308.3187	300.5991	308.3187	301.7777
H19	379.9339	316.4333	308.3074	300.61	308.3074	301.8541
H20	379.8466	316.4224	308.2962	300.621	308.2962	301.9305
H21	379.7593	316.4115	308.2849	300.6319	308.2849	302.0069
H22	379.672	316.4006	308.2736	300.6428	308.2736	302.0832
H23	379.5847	316.3897	308.2623	300.6537	308.2623	302.1596
H24	379.4974	316.3788	308.2511	300.6646	308.2511	302.236

Table C.19: Assumed initial units of the system and decommissioning plan.

Year [1]	TU1 [1]	TU2 [1]	TU3 [1]	CH [1]	HS [1]	WF [1]	SPV [1]
1	16	10	1	29	1	18	22
2	16	10	1	29	1	18	22
3	16	10	1	29	1	18	22
4	16	10	1	29	1	18	22
5	16	10	1	25	1	18	22
6	16	10	1	25	1	18	22
7	16	9	1	25	1	18	22
8	15	9	1	21	1	18	22
9	14	9	1	18	1	18	22
10	13	9	1	18	1	17	22
11	10	9	1	18	1	16	22
12	8	8	1	15	1	15	22
13	7	7	1	15	1	15	22
14	6	6	1	15	1	13	22
15	6	6	1	15	1	13	22

Appendix D

Statistical tests for the DE variants

This Section presents the results of Wilcoxon rank sum test (the Matlab implementation has been used) for the DE variants examined. The DE variants including all examined operators are tested against the variants including combinations of the examined operators. The test result 'TRUE' indicates a rejection of the null hypothesis, and 'FALSE' indicates a failure to reject the null hypothesis at the 5% significance level.

Table D.1: The table presents the results of the Wilcoxon rank sum test for the *noSM* case.

DE/rand/1/bin							
/RRH/blk/PO vs	/FR	/FR/blk	/FR/PO	/FR/blk/PO	/RRH	/RRH/blk	/RRH/PO
Two-sided	TRUE	TRUE	TRUE	TRUE	FALSE	FALSE	FALSE
Left-tailed	TRUE	TRUE	TRUE	TRUE	FALSE	FALSE	FALSE
Right-tailed	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
p-value (Two-sided)	5.64E-39	5.63E-39	2.16E-38	2.15E-38	NaN	NaN	NaN
p-value (Left-tailed)	2.82E-39	2.82E-39	1.08E-38	1.08E-38	1.00E+00	1.00E+00	1.00E+00
p-value (Right-tailed)	1.00E+00	1.00E+00	1.00E+00	1.00E+00	1.00E+00	1.00E+00	1.00E+00
DE/rand/1/exp							
/RRH/blk/PO vs	/FR	/FR/blk	/FR/PO	/FR/blk/PO	/RRH	/RRH/blk	/RRH/PO
Two-sided	TRUE	TRUE	TRUE	TRUE	FALSE	FALSE	FALSE
Left-tailed	TRUE	TRUE	TRUE	TRUE	FALSE	FALSE	FALSE
Right-tailed	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
p-value (Two-sided)	5.64E-39	5.63E-39	5.64E-39	2.16E-38	1.58E-01	NaN	NaN
p-value (Left-tailed)	2.82E-39	2.82E-39	2.82E-39	1.08E-38	7.92E-02	1.00E+00	1.00E+00
p-value (Right-tailed)	1.00E+00	1.00E+00	1.00E+00	1.00E+00	9.23E-01	1.00E+00	1.00E+00
DE/rand/1/expS							
/RRH/blk/PO vs	/FR	/FR/blk	/FR/PO	/FR/blk/PO	/RRH	/RRH/blk	/RRH/PO
Two-sided	TRUE	TRUE	TRUE	TRUE	FALSE	FALSE	FALSE
Left-tailed	TRUE	TRUE	TRUE	TRUE	FALSE	FALSE	FALSE
Right-tailed	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
p-value (Two-sided)	5.64E-39	5.64E-39	5.64E-39	5.64E-39	3.22E-01	NaN	NaN
p-value (Left-tailed)	2.82E-39	2.82E-39	2.82E-39	2.82E-39	1.61E-01	1.00E+00	1.00E+00
p-value (Right-tailed)	1.00E+00	1.00E+00	1.00E+00	1.00E+00	8.44E-01	1.00E+00	1.00E+00

Table D.2: The table presents the results of the Wilcoxon rank sum test for the SM^{ED}-*Lin* case.

DE/rand/1/bin							
/RRH/blk/PO vs	/FR	/FR/blk	/FR/PO	/FR/blk/PO	/RRH	/RRH/blk	/RRH/PO
Two-sided	TRUE	TRUE	TRUE	TRUE	FALSE	FALSE	FALSE
Left-tailed	TRUE	TRUE	TRUE	TRUE	FALSE	FALSE	FALSE
Right-tailed	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
p-value (Two-sided)	8.14E-38	1.12E-36	5.64E-39	2.16E-38	NaN	NaN	NaN
p-value (Left-tailed)	4.07E-38	5.59E-37	2.82E-39	1.08E-38	1.00E+00	1.00E+00	1.00E+00
p-value (Right-tailed)	1.00E+00	1.00E+00	1.00E+00	1.00E+00	1.00E+00	1.00E+00	1.00E+00
DE/rand/1/exp							
/RRH/blk/PO vs	/FR	/FR/blk	/FR/PO	/FR/blk/PO	/RRH	/RRH/blk	/RRH/PO
Two-sided	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	FALSE
Left-tailed	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	FALSE
Right-tailed	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
p-value (Two-sided)	5.64E-39	5.64E-39	5.64E-39	5.64E-39	2.42E-02	NaN	3.22E-01
p-value (Left-tailed)	2.82E-39	2.82E-39	2.82E-39	2.82E-39	1.21E-02	1.00E+00	1.61E-01
p-value (Right-tailed)	1.00E+00	1.00E+00	1.00E+00	1.00E+00	9.88E-01	1.00E+00	8.44E-01
DE/rand/1/expS							
/RRH/blk/PO vs	/FR	/FR/blk	/FR/PO	/FR/blk/PO	/RRH	/RRH/blk	/RRH/PO
Two-sided	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE
Left-tailed	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE
Right-tailed	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
p-value (Two-sided)	5.64E-39	5.64E-39	5.64E-39	5.64E-39	2.42E-02	NaN	2.03E-04
p-value (Left-tailed)	2.82E-39	2.82E-39	2.82E-39	2.82E-39	1.21E-02	1.00E+00	1.02E-04
p-value (Right-tailed)	1.00E+00	1.00E+00	1.00E+00	1.00E+00	9.88E-01	1.00E+00	1.00E+00

Table D.3: The table presents the results of the Wilcoxon rank sum test for the SM^{ED}-*Cub* case.

DE/rand/1/bin							
/RRH/blk/PO vs	/FR	/FR/blk	/FR/PO	/FR/blk/PO	/RRH	/RRH/blk	/RRH/PO
Two-sided	TRUE	TRUE	TRUE	TRUE	FALSE	FALSE	FALSE
Left-tailed	TRUE	TRUE	TRUE	TRUE	FALSE	FALSE	FALSE
Right-tailed	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
p-value (Two-sided)	5.64E-39	8.14E-38	2.16E-38	2.16E-38	NaN	NaN	NaN
p-value (Left-tailed)	2.82E-39	4.07E-38	1.08E-38	1.08E-38	1.00E+00	1.00E+00	1.00E+00
p-value (Right-tailed)	1.00E+00	1.00E+00	1.00E+00	1.00E+00	1.00E+00	1.00E+00	1.00E+00
DE/rand/1/exp							
/RRH/blk/PO vs	/FR	/FR/blk	/FR/PO	/FR/blk/PO	/RRH	/RRH/blk	/RRH/PO
Two-sided	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE
Left-tailed	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE
Right-tailed	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
p-value (Two-sided)	5.64E-39	5.64E-39	5.64E-39	5.64E-39	2.03E-04	NaN	7.31E-03
p-value (Left-tailed)	2.82E-39	2.82E-39	2.82E-39	2.82E-39	1.01E-04	1.00E+00	3.65E-03
p-value (Right-tailed)	1.00E+00	1.00E+00	1.00E+00	1.00E+00	1.00E+00	1.00E+00	9.96E-01
DE/rand/1/expS							
/RRH/blk/PO vs	/FR	/FR/blk	/FR/PO	/FR/blk/PO	/RRH	/RRH/blk	/RRH/PO
Two-sided	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE
Left-tailed	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE
Right-tailed	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
p-value (Two-sided)	5.64E-39	2.16E-38	5.64E-39	5.64E-39	2.23E-03	NaN	4.04E-03
p-value (Left-tailed)	2.82E-39	1.08E-38	2.82E-39	2.82E-39	1.11E-03	1.00E+00	2.02E-03
p-value (Right-tailed)	1.00E+00	1.00E+00	1.00E+00	1.00E+00	9.99E-01	1.00E+00	9.98E-01

Table D.4: The table presents the results of the Wilcoxon rank sum test for the SM^{ED}-TPS case.

DE/rand/1/bin							
/RRH/blk/PO vs	/FR	/FR/blk	/FR/PO	/FR/blk/PO	/RRH	/RRH/blk	/RRH/PO
Two-sided	TRUE	TRUE	TRUE	TRUE	FALSE	FALSE	FALSE
Left-tailed	TRUE	TRUE	TRUE	TRUE	FALSE	FALSE	FALSE
Right-tailed	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
p-value (Two-sided)	2.16E-38	2.15E-38	5.64E-39	5.63E-39	NaN	NaN	NaN
p-value (Left-tailed)	1.08E-38	1.08E-38	2.82E-39	2.82E-39	1.00E+00	1.00E+00	1.00E+00
p-value (Right-tailed)	1.00E+00	1.00E+00	1.00E+00	1.00E+00	1.00E+00	1.00E+00	1.00E+00
DE/rand/1/exp							
/RRH/blk/PO vs	/FR	/FR/blk	/FR/PO	/FR/blk/PO	/RRH	/RRH/blk	/RRH/PO
Two-sided	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE
Left-tailed	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE
Right-tailed	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
p-value (Two-sided)	5.64E-39	5.64E-39	5.64E-39	2.16E-38	7.31E-03	NaN	4.44E-02
p-value (Left-tailed)	2.82E-39	2.82E-39	2.82E-39	1.08E-38	3.66E-03	1.00E+00	2.22E-02
p-value (Right-tailed)	1.00E+00	1.00E+00	1.00E+00	1.00E+00	9.96E-01	1.00E+00	9.78E-01
DE/rand/1/expS							
/RRH/blk/PO vs	/FR	/FR/blk	/FR/PO	/FR/blk/PO	/RRH	/RRH/blk	/RRH/PO
Two-sided	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE
Left-tailed	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE
Right-tailed	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
p-value (Two-sided)	5.64E-39	2.16E-38	5.64E-39	5.64E-39	4.04E-03	NaN	6.78E-04
p-value (Left-tailed)	2.82E-39	1.08E-38	2.82E-39	2.82E-39	2.02E-03	1.00E+00	3.39E-04
p-value (Right-tailed)	1.00E+00	1.00E+00	1.00E+00	1.00E+00	9.98E-01	1.00E+00	1.00E+00

Table D.5: The table presents the results of the Wilcoxon rank sum test for the SM^{CUC}-Lin case.

DE/rand/1/bin							
/RRH/blk/PO vs	/FR	/FR/blk	/FR/PO	/FR/blk/PO	/RRH	/RRH/blk	/RRH/PO
Two-sided	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE
Left-tailed	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
Right-tailed	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
p-value (Two-sided)	5.64E-39	5.64E-39	5.64E-39	5.64E-39	7.30E-03	8.27E-02	7.30E-03
p-value (Left-tailed)	2.82E-39	2.82E-39	2.82E-39	2.82E-39	3.65E-03	4.14E-02	3.65E-03
p-value (Right-tailed)	1.00E+00	1.00E+00	1.00E+00	1.00E+00	9.96E-01	9.60E-01	9.96E-01
DE/rand/1/exp							
/RRH/blk/PO vs	/FR	/FR/blk	/FR/PO	/FR/blk/PO	/RRH	/RRH/blk	/RRH/PO
Two-sided	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE
Left-tailed	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE
Right-tailed	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
p-value (Two-sided)	1.62E-38	1.62E-38	1.62E-38	1.62E-38	2.54E-07	1.00E+00	2.32E-07
p-value (Left-tailed)	8.11E-39	8.11E-39	8.11E-39	8.11E-39	1.27E-07	5.02E-01	1.16E-07
p-value (Right-tailed)	1.00E+00	1.00E+00	1.00E+00	1.00E+00	1.00E+00	5.02E-01	1.00E+00
DE/rand/1/expS							
/RRH/blk/PO vs	/FR	/FR/blk	/FR/PO	/FR/blk/PO	/RRH	/RRH/blk	/RRH/PO
Two-sided	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE
Left-tailed	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE
Right-tailed	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
p-value (Two-sided)	8.10E-39	8.10E-39	8.10E-39	8.10E-39	4.45E-10	1.00E+00	9.45E-10
p-value (Left-tailed)	4.05E-39	4.05E-39	4.05E-39	4.05E-39	2.22E-10	5.03E-01	4.73E-10
p-value (Right-tailed)	1.00E+00	1.00E+00	1.00E+00	1.00E+00	1.00E+00	5.03E-01	1.00E+00

Table D.6: The table presents the results of the Wilcoxon rank sum test for the SM^{CUC}-*Cub* case.

DE/rand/1/bin							
/RRH/blk/PO vs	/FR	/FR/blk	/FR/PO	/FR/blk/PO	/RRH	/RRH/blk	/RRH/PO
Two-sided	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE
Left-tailed	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE
Right-tailed	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
p-value (Two-sided)	2.26E-38	2.26E-38	2.26E-38	2.26E-38	1.38E-02	7.04E-01	2.30E-02
p-value (Left-tailed)	1.13E-38	1.13E-38	1.13E-38	1.13E-38	6.90E-03	6.51E-01	1.15E-02
p-value (Right-tailed)	1.00E+00	1.00E+00	1.00E+00	1.00E+00	9.93E-01	3.52E-01	9.89E-01
DE/rand/1/exp							
/RRH/blk/PO vs	/FR	/FR/blk	/FR/PO	/FR/blk/PO	/RRH	/RRH/blk	/RRH/PO
Two-sided	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE
Left-tailed	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE
Right-tailed	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
p-value (Two-sided)	1.62E-38	1.62E-38	1.62E-38	1.62E-38	4.43E-10	7.04E-01	3.80E-09
p-value (Left-tailed)	8.11E-39	8.11E-39	8.11E-39	8.11E-39	2.21E-10	3.52E-01	1.90E-09
p-value (Right-tailed)	1.00E+00	1.00E+00	1.00E+00	1.00E+00	1.00E+00	6.51E-01	1.00E+00
DE/rand/1/expS							
/RRH/blk/PO vs	/FR	/FR/blk	/FR/PO	/FR/blk/PO	/RRH	/RRH/blk	/RRH/PO
Two-sided	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE
Left-tailed	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE
Right-tailed	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
p-value (Two-sided)	5.79E-38	5.79E-38	5.79E-38	5.79E-38	7.83E-08	1.97E-01	2.29E-08
p-value (Left-tailed)	2.89E-38	2.89E-38	2.89E-38	2.89E-38	3.91E-08	9.03E-01	1.15E-08
p-value (Right-tailed)	1.00E+00	1.00E+00	1.00E+00	1.00E+00	1.00E+00	9.83E-02	1.00E+00

Table D.7: The table presents the results of the Wilcoxon rank sum test for the SM^{CUC}-*TPS* case.

DE/rand/1/bin							
/RRH/blk/PO vs	/FR	/FR/blk	/FR/PO	/FR/blk/PO	/RRH	/RRH/blk	/RRH/PO
Two-sided	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE
Left-tailed	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE
Right-tailed	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
p-value (Two-sided)	8.10E-39	8.10E-39	8.10E-39	8.10E-39	9.72E-03	1.77E-01	5.41E-03
p-value (Left-tailed)	4.05E-39	4.05E-39	4.05E-39	4.05E-39	4.86E-03	8.84E-02	2.71E-03
p-value (Right-tailed)	1.00E+00	1.00E+00	1.00E+00	1.00E+00	9.95E-01	9.13E-01	9.97E-01
DE/rand/1/exp							
/RRH/blk/PO vs	/FR	/FR/blk	/FR/PO	/FR/blk/PO	/RRH	/RRH/blk	/RRH/PO
Two-sided	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE
Left-tailed	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE
Right-tailed	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
p-value (Two-sided)	1.62E-38	1.62E-38	1.62E-38	1.62E-38	1.67E-08	3.09E-01	1.35E-07
p-value (Left-tailed)	8.11E-39	8.11E-39	8.11E-39	8.11E-39	8.37E-09	1.54E-01	6.73E-08
p-value (Right-tailed)	1.00E+00	1.00E+00	1.00E+00	1.00E+00	1.00E+00	8.47E-01	1.00E+00
DE/rand/1/expS							
/RRH/blk/PO vs	/FR	/FR/blk	/FR/PO	/FR/blk/PO	/RRH	/RRH/blk	/RRH/PO
Two-sided	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE
Left-tailed	TRUE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE
Right-tailed	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
p-value (Two-sided)	2.26E-38	2.26E-38	2.26E-38	2.26E-38	6.84E-11	7.04E-01	3.13E-10
p-value (Left-tailed)	1.13E-38	1.13E-38	1.13E-38	1.13E-38	3.42E-11	6.51E-01	1.57E-10
p-value (Right-tailed)	1.00E+00	1.00E+00	1.00E+00	1.00E+00	1.00E+00	3.52E-01	1.00E+00

Appendix E

Simulation model based on Fast Unit Commitment

The SM used is adopted from Reference [29]. It is formulated as a SOO problem. Some modifications to the formulation have been included which are mentioned in Section 5.2.3.

The optimization variables of the sub-problem, where $v \in \mathbb{R}^n$ and n the dimensionality of $\mathbf{v}$, are $v = \{\mathbf{x}^{on}, \mathbf{x}^{su}, \mathbf{x}^{sd}, \epsilon^d, \epsilon^r, \epsilon^c, \epsilon^s, \mathbf{p}, \mathbf{p}^H, \mathbf{p}^{Tr}, \mathbf{x}^{on(init)}, \mathbf{x}^{su(init)}, \mathbf{p}^{init}\}$.

The mathematical formulation of the SOO problem is of the following form:

$$\begin{aligned} &minimize \;\; TC(\mathbf{v}) \\ &s.t. \;\; E.4 - E.37 \\ &\qquad \mathbf{v} \in \Re^n \end{aligned} \tag{E.1}$$

The equations included within the sub-problem are the following:

- Objective Function: The objective function is set as the minimization of the total operation cost of the considered hourly time periods.

$$\begin{aligned} TC = &\sum_{\forall a,g,h} [(C_g^L + EF_g^{em} C^{em}) p_{a,g,h}] + \sum_{\forall a,g,h} [C_g^{su} x_{a,g,h}^{su} + C_g^{sd} x_{a,g,h}^{sd}] + \sum_{\forall a,h} [\epsilon_{a,h}^d] C^d + \sum_{\forall a,h} [\epsilon_{a,h}^r] C^r \\ &+ \sum_{\forall a,h} [\epsilon_{a,h}^c] C^c + \sum_{\forall a,h} [\epsilon_{a,h}^s] C^s \end{aligned} \tag{E.2}$$

The fuel and emission cost term (first sum term) could be altered to the following if sufficient data are available:

$$c^{f\&em} = \sum_{\forall a,g,h} [p_{a,g,h}(C_g^{sl} + EF_g^{sl} C^{em}) + x_{a,g,h}^{on} \underline{OL}_g (C_g^{min\_f} - C_g^{sl} + C^{em}(EF_g^{min} - EF_g^{sl\_em}))] \tag{E.3}$$

where C^{sl} and EF^{sl} are parameters based on the slope of the fuel and emission functions, respectively. C^{min_f} and EF_g^{min} are the minimum fuel cost and emission factor per unit of on-line capacity. Eqs. 5.5 and 5.6 should be updated correspondingly.

- Demand constraint: The demand constraint should ensure that supply meets demand in each hour.

$$\sum_g [p_{a,g,h}] + p^H_{a,h} + (\overline{P}^{res}_{a,h} - \epsilon^c_{a,h}) + \sum_{m\in\Psi_z} [p^{Tr}_{a,m}] = D_{a,h} - \epsilon^d_{a,h}, \forall a, h \tag{E.4}$$

- Reserve constraint: The reserve constraint ensures that sufficient capacity is on-line to moderate the impact of uncertainty.

$$\sum_g [x^{on}_{a,g,h}] + p^H_{a,h} + (\overline{P}^{res}_{a,h} - \epsilon^c_{a,h}) + \epsilon^r_{a,h} + \sum_{m\in\Psi_{zm}} [p^{Tr}_{a,zm}] \geq (1+Rf^d)(D_{a,h} - \epsilon^d_{a,h}) + Rf^{res}(\overline{P}^{res}_{a,h} - \epsilon^c_{a,h}), \forall a, h \tag{E.5}$$

- Operation status constraint:

$$x^{on}_{a,g,h} = x^{on}_{a,g,h-1} + x^{su}_{a,g,h} - x^{sd}_{a,g,h} \forall a, g, h \tag{E.6}$$

- Operation status constraint for initial hour:

$$x^{on}_{a,g,1} = x^{on(init)}_{a,g} + x^{su}_{a,g,1} - x^{sd}_{a,g,1} \forall a, g \tag{E.7}$$

- Ramp up constraints:

$$p_{a,g,h} - p_{a,g,h-1} \leq \underline{OL}_g(x^{su}_{a,g,h} - x^{sd}_{a,g,h}) + RU_g(x^{on}_{a,g,h} - x^{su}_{a,g,h} - x^{sd}_{a,g,h+1}), \forall a, g, h \tag{E.8}$$

- Ramp up constraints for first hour:

$$p_{a,g,1} - p^{init}_{a,g} \leq \underline{OL}_g(x^{su}_{a,g,1} - x^{sd}_{a,g,1}) + RU_g(x^{on}_{a,g,1} - x^{su}_{a,g,1} - x^{sd}_{a,g,2}), \forall a, g \tag{E.9}$$

- Ramp up constraints for last hour:

$$p_{a,g,hz} - p_{a,g,hz-1} \leq \underline{OL}_g(x^{su}_{a,g,hz} - x^{sd}_{a,g,hz}) + RU_g(x^{on}_{a,g,hz} - x^{su}_{a,g,hz}), \forall a, g \tag{E.10}$$

- Ramping down constraints:

$$p_{a,g,h} - p_{a,g,h-1} \geq \underline{OL}_g(x^{su}_{a,g,h} - x^{sd}_{a,g,h}) - RD_g(x^{on}_{a,g,h} - x^{su}_{a,g,h} - x^{su}_{a,g,h-1}), \forall a, g, h \tag{E.11}$$

- Ramping down constraints for first hour:

$$p_{a,g,1} - p^{init}_{a,g} \geq \underline{OL}_g(x^{su}_{a,g,1} - x^{sd}_{a,g,1}) - RD_g(x^{on}_{a,g,1} - x^{su}_{a,g,1} - x^{su(init)}_{a,g}), \forall a, g \tag{E.12}$$

- Operating limits constraints:

$$p_{a,g,h} \leq \overline{OL}_g(x^{on}_{a,g,h} - x^{su}_{a,g,h} - x^{sd}_{a,g,h+1}) + \underline{OL}_g x^{su}_{a,g,h} + \underline{OL}_g x^{sd}_{a,g,h+1}, \forall a, g, h \tag{E.13}$$

$$x^{on}_{a,g,h}\underline{OL}_g \leq p_{a,g,h} \leq x^{on}_{a,g,h}\overline{OL}_g, \forall a, g, h \tag{E.14}$$

- Operating limits constraints for last hour:

$$p_{a,g,hz} \leq \overline{OL}_g(x^{on}_{a,g,hz} - x^{su}_{a,g,hz}) + \underline{OL}_g x^{su}_{a,g,hz}, \forall a, g \tag{E.15}$$

- Operating limits constraints of initial production:

$$x_{a,g}^{on(init)}\underline{OL}_g \leq p_{a,g}^{init} \leq x_{a,g}^{on(init)}\overline{OL}_g, \forall a, g \tag{E.16}$$

- Minimum on-line time limits constraints for first hour:

$$0 \leq x_{a,g,1}^{sd} \leq x_{a,g}^{on(init)}, \forall a, g \tag{E.17}$$

- Minimum on-line time limits constraints for first MT^{on} hours except first hour:

$$0 \leq x_{a,g,h+1}^{sd} \leq x_{a,g,h}^{on} - \sum_{\tau=0}^{h-1}[x_{a,g,h-\tau}^{su}], \forall a, g, 1 \leq h \leq MT_g^{on} \tag{E.18}$$

- Minimum on-line time limits constraints for all hours except first MT^{on} hours:

$$0 \leq x_{a,g,h+1}^{sd} \leq x_{a,g,h}^{on} - \sum_{\tau=0}^{MT_g^{on}-2}[x_{a,g,h-\tau}^{su}], \forall a, g, MT_g^{on} \leq h \leq hz-1 \tag{E.19}$$

- Minimum off-line time limits constraints for first hour:

$$0 \leq x_{a,g,1}^{su} \leq \overline{P}_{a,g} - x_{a,g}^{on(init)}, \forall a, g \tag{E.20}$$

- Minimum off-line time limits constraints for first MT^{off} hours except first hour:

$$0 \leq x_{a,g,h+1}^{su} \leq \overline{P}_{a,g} - x_{a,g,h}^{on} - \sum_{\tau=0}^{h-1}[x_{a,g,h-\tau}^{sd}], \forall a, g, 1 \leq h \leq MT_g^{off} \tag{E.21}$$

- Minimum off-line time limits constraints for all hours except first MT^{off} hours:

$$0 \leq x_{a,g,h+1}^{su} \leq \overline{P}_{a,g} - x_{a,g,h}^{on} - \sum_{\tau=0}^{MT_g^{off}-2}[x_{a,g,h-\tau}^{sd}], \forall a, g, MT_g^{off} \leq h \leq hz-1 \tag{E.22}$$

- Hydro reservoir final hour constraints:

$$EC_a^{init}\overline{P}_a^H + \sum_{\tau=0}^{hz}[Inflows_{a,h}\overline{P}_a^H] - \frac{1}{Eff^H}\sum_{\tau=0}^{hz}[p_{a,h}^H] - \sum_{\tau=0}^{hz}[\epsilon_{a,h}^s] = EC_a^{final}\overline{P}_a^H, \forall a \tag{E.23}$$

- Hydro minimum and maximum reservoir limits constraints:

$$\underline{EC}_a\overline{P}_a^H \leq EC_a^{init}\overline{P}_a^H + \sum_{\tau=0}^{h}[Inflows_{a,h}\overline{P}_a^H] - \frac{1}{Eff_a^H}\sum_{\tau=0}^{h}[p_{a,h}^H] - \sum_{\tau=0}^{h}[\epsilon_{a,h}^s] \leq \overline{EC}_a\overline{P}_a^H, \forall a, h \tag{E.24}$$

- Upper and lower constraints of variables:

$$0 \leq x_{a,g,h}^{on} \leq \overline{P}_{a,g}, \forall a, g, h \tag{E.25}$$

$$0 \leq x_{a,g,h}^{su} \leq \overline{P}_{a,g}, \forall a, g, h \tag{E.26}$$

$$0 \leq x_{a,g,h}^{sd} \leq \overline{P}_{a,g}, \forall a, g, h \tag{E.27}$$

$$0 \leq x_{a,g}^{on(init)} \leq \overline{P}_{a,g}, \forall a, g \tag{E.28}$$

$$0 \leq x_{a,g}^{su(init)} \leq \overline{P}_{a,g}, \forall a, g \tag{E.29}$$

$$0 \leq \epsilon_{a,h}^{c} \leq \overline{P}_{a,g}^{res}, \forall a, h \tag{E.30}$$

$$0 \leq \epsilon_{a,h}^{s} \leq \overline{P}_{a,g}^{H}, \forall a, h \tag{E.31}$$

$$0 \leq \epsilon_{a,h}^{d} \leq D_{a,h}, \forall a, h \tag{E.32}$$

$$0 \leq \epsilon_{a,h}^{r} \leq (1 + rf^{d})D_{a,h} + rf^{res}\overline{P}_{a,h}^{res}, \forall a, h \tag{E.33}$$

$$-\overline{P}_{zm}^{Tr} \leq p_{zm}^{Tr} \leq \overline{P}_{zm}^{Tr}, zm \in \Psi_{z} \tag{E.34}$$

$$0 \leq p_{a,h}^{H} \leq \overline{P}_{a}^{H}, \forall a, h \tag{E.35}$$

$$0 \leq p_{a,g,h} \leq \overline{P}_{a,g}, \forall a, g, h \tag{E.36}$$

$$0 \leq p_{a,g,h}^{init} \leq \overline{P}_{a,g}, \forall a, g \tag{E.37}$$

Appendix F

Data input for the test case used in Chapter 5

The data input (Section 5.4.1) are presented in this Section.

Table F.1: Technical characteristics of the assumed technology groups.

	area	$\overline{P}$ [MW]	$\underline{P}$ [MW]	$\overline{\mathbf{OL}}$ [1]	$\underline{\mathbf{OL}}$ [1]	$\mathbf{RU}'$ [MW]	$\mathbf{RD}'$ [MW]	**RU** [1]	**RD** [1]	$\mathbf{MT}^{on}$ [h]	$\mathbf{MT}^{off}$ [h]
$\mathrm{TU1^{old}}$	1	300	129	1	0.43	135	135	0.45	0.45	11	11
$\mathrm{TU2^{old}}$	1	400	140	1	0.35	200	200	0.5	0.5	3	3
$\mathrm{TU3^{old}}$	1	150	45	1	0.3	150	150	1	1	1	1
TU1	1	300	129	1	0.43	135	135	0.45	0.45	8	8
TU2	1	400	140	1	0.35	200	200	0.5	0.5	2	2
TU3	1	150	45	1	0.3	150	150	1	1	1	1
SPV	1	100	0	1	0	0	0	0	0	0	0
WF	1	100	0	1	0	0	0	0	0	0	0
CH	1	250	0	1	0	0	0	0	0	0	0
$\mathrm{TU1^{old}}$	2	300	129	1	0.43	135	135	0.45	0.45	11	11
$\mathrm{TU2^{old}}$	2	400	140	1	0.35	200	200	0.5	0.5	3	3
$\mathrm{TU3^{old}}$	2	150	45	1	0.3	150	150	1	1	1	1
TU1	2	300	129	1	0.43	135	135	0.45	0.45	8	8
TU2	2	400	140	1	0.35	200	200	0.5	0.5	2	2
TU3	2	150	45	1	0.3	150	150	1	1	1	1
SPV	2	100	0	1	0	0	0	0	0	0	0
WF	2	100	0	1	0	0	0	0	0	0	0
CH	2	250	0	1	0	0	0	0	0	0	0

Table F.2: Economic characteristics of the assumed technology groups.

	area	**IC** [€/MW]	**FOM** [€/MW]	**VOM** [€/MW]	$\mathbf{C}^{su}$ [€/MW]	$\mathbf{C}^{sd}$ [€/MW]	$\mathbf{C}^{L}$ [€/MW]	$\mathbf{EF}^{em}$ [$tCO_{2\ eq}$/MW]	$\mathbf{Pr}^{res}$ [€/MW]
$TU1^{old}$	1	0	21600	3.3	86.0536	0	11.3134	1038.774	0
$TU2^{old}$	1	0	9200	1.6	44.3159	0	45.7463	427.4658	0
$TU3^{old}$	1	0	8300	1.6	52.5044	0	62.7378	586.2388	0
TU1	1	50000	21600	3.3	58.6536	0	8.608	790.3716	0
TU2	1	20000	9200	1.6	26.3059	0	37.859	353.7648	0
TU3	1	18000	8300	1.6	20.3944	0	52.2815	488.5323	0
SPV	1	52000	10700	0	0	0	0	0	80
WF	1	112000	9180	0	0	0	0	0	100
CH	1	42222.22	23800	0	0	0	0	0	0
$TU1^{old}$	2	0	21600	3.3	86.0536	0	11.3134	1038.774	0
$TU2^{old}$	2	0	9200	1.6	44.3159	0	45.7463	427.4658	0
$TU3^{old}$	2	0	8300	1.6	52.5044	0	62.7378	586.2388	0
TU1	2	50000	10700	3.3	58.6536	0	8.608	790.3716	0
TU2	2	20000	9180	1.6	26.3059	0	37.859	353.7648	0
TU3	2	18000	23800	1.6	20.3944	0	52.2815	488.5323	0
SPV	2	52000	10700	0	0	0	0	0	80
WF	2	112000	9180	0	0	0	0	0	100
CH	2	42222.22	23800	0	0	0	0	0	0

Table F.3: Restrictions on the capacity additions for the assumed technology groups.

	$\mathbf{P}^{cap\_step}$ [MW]	**TCA** [1]	$\overline{\mathbf{X}}_1$ [1]	$\overline{\mathbf{X}}_2$ [1]	$\underline{\mathbf{X}}_1$ [1]	$\underline{\mathbf{X}}_2$ [1]	$\mathbf{IU}_1$ [1]	$\mathbf{IU}_2$ [1]
$TU1^{old}$	300	0	0	0	0	0	9	4
$TU2^{old}$	400	0	0	0	0	0	2	7
$TU3^{old}$	150	0	0	0	0	0	0	1
TU1	300	20	20	20	0	0	0	0
TU2	400	20	20	20	0	0	0	0
TU3	150	20	20	20	0	0	0	0
SPV	100	20	20	20	0	0	10	10
WF	100	20	20	20	0	0	10	15
CH	250	20	20	20	0	0	6	6

Table F.4: Assumed data input for the CH capacity.

area	$\mathbf{eff}^{H}$ [1]	**Inflows** [1]	$\mathbf{EC}^{init}$ [1]	$\mathbf{EC}^{final}$ [1]	$\overline{\mathbf{EC}}$ [1]	$\underline{\mathbf{EC}}$ [1]
1	1.1111	0.1	10	10	11	9
2	1.1111	0.1	10	10	11	9

Table F.5: System-related data input.

$\mathbf{Rf}^{d}$ [1]	$\mathbf{Rf}^{res}$ [1]	RM [1]	C^{d} [€/MW]	C^{r} [€/MW]	C^{c} [€/MW]	C^{s} [€/MW]	C^{em} [€/$tCO_{2\ eq}$]	$\mathbf{p}^{Tr}$ [MW]
0.1	0.05	0.15	10000	10000	150	150	0.018	1500

Table F.6: Total demand of the first week.

	W1D1 [MW]	W1D2 [MW]	W1D3 [MW]	W1D4 [MW]	W1D5 [MW]	W1D6 [MW]	W1D7 [MW]
H1	6198.5	5671.6	6048.9	6130.3	6332.7	6425.1	6070.9
H2	5762.9	5231.6	5458.2	5621	5791.5	5822.3	5599
H3	5635.3	5240.4	5471.4	5637.5	5800.3	5761.8	5574.8
H4	5308.6	5007.2	5324	5412	5638.6	5551.7	5324
H5	4996.2	4780.6	5107.3	5228.3	5448.3	5310.8	5042.4
H6	4790.5	4810.3	5093	5212.9	5501.1	5121.6	4907.1
H7	4761.9	5119.4	5476.9	5615.5	5890.5	5109.5	4980.8
H8	4676.1	5573.7	6106.1	6230.4	6527.4	5166.7	4981.9
H9	4732.2	6109.4	6795.8	7035.6	7241.3	5440.6	5166.7
H10	5132.6	6692.4	7581.2	7851.8	7991.5	5795.9	5616.6
H11	5604.5	6965.2	7755	8036.6	8078.4	6095.1	6110.5
H12	6064.3	7199.5	7849.6	8191.7	8089.4	6543.9	6448.2
H13	6145.7	7360.1	7877.1	8341.3	8155.4	6850.8	6675.9
H14	5745.3	7386.5	7854	8250	8077.3	6668.2	6463.6
H15	5090.8	7126.9	7577.9	7791.3	7892.5	5989.5	5882.8
H16	5033.6	6986.1	7324.9	7613.1	7873.8	5943.3	5808
H17	5322.9	7114.8	7464.6	7804.5	8168.6	6246.9	6094
H18	5926.8	7546	8010.2	8419.4	8751.6	6829.9	6645.1
H19	6655	7936.5	8427.1	8969.4	9128.9	7502	7472.3
H20	6761.7	7922.2	8400.7	8939.7	9060.7	7624.1	7583.4
H21	6787	7784.7	8301.7	8787.9	8936.4	7470.1	7450.3
H22	6564.8	7482.2	7859.5	8112.5	8335.8	7174.2	7143.4
H23	6198.5	6890.4	7157.7	7365.6	7550.4	6728.7	6609.9
H24	6167.7	6653.9	6855.2	7055.4	7126.9	6531.8	6405.3

Table F.7: Total demand of the second week.

	W2D1 [MW]	W2D2 [MW]	W2D3 [MW]	W2D4 [MW]	W2D5 [MW]	W2D6 [MW]	W2D7 [MW]
H1	5549.5	5715.6	5811.3	5782.7	5687	5140.3	5380.1
H2	4955.5	5172.2	5252.5	5174.4	5064.4	4648.6	5014.9
H3	4899.4	5086.4	5149.1	5042.4	4890.6	4521	4695.9
H4	4736.6	4930.2	4930.2	4841.1	4683.8	4301	4224
H5	4647.5	4823.5	4823.5	4736.6	4448.4	4126.1	3932.5
H6	4711.3	4911.5	4896.1	4842.2	4395.6	4137.1	3847.8
H7	5204.1	5365.8	5342.7	5241.5	4375.8	4314.2	3804.9
H8	5915.8	6034.6	6067.6	5860.8	4459.4	4523.2	3768.6
H9	6806.8	6867.3	6912.4	6682.5	4912.6	5038	4103
H10	7396.4	7365.6	7471.2	7156.6	5317.4	5523.1	4556.2
H11	7367.8	7359	7445.9	7289.7	5429.6	5800.3	4867.5
H12	7381	7420.6	7477.8	7299.6	5421.9	5885	4852.1
H13	7398.6	7419.5	7455.8	7272.1	5434	5963.1	4611.2
H14	7323.8	7333.7	7483.3	7150	5415.3	5993.9	4225.1
H15	7142.3	7201.7	7378.8	6895.9	5281.1	5777.2	3824.7
H16	6988.3	6948.7	7224.8	6718.8	5291	5684.8	3803.8
H17	6920.1	6893.7	7203.9	6803.5	5347.1	5729.9	3914.9
H18	6695.7	6722.1	7048.8	6642.9	5244.8	5695.8	3853.3
H19	6584.6	6570.3	7078.5	6506.5	5286.6	5848.7	3952.3
H20	6941	6946.5	7388.7	6870.6	5528.6	6382.2	4325.2
H21	7907.9	7877.1	8032.2	7404.1	6114.9	7504.2	5320.7
H22	7711	7838.6	7856.2	7169.8	5858.6	7361.2	5553.9
H23	6974	7123.6	7141.2	6723.2	5838.8	6605.5	5285.5
H24	6452.6	6584.6	6569.2	6400.9	5711.2	5819	5066.6

Table F.8: Total demand of the third week.

	W3D1 [MW]	W3D2 [MW]	W3D3 [MW]	W3D4 [MW]	W3D5 [MW]	W3D6 [MW]	W3D7 [MW]
H1	6994.9	7368.9	7541.6	7697.8	7520.7	7284.2	7040
H2	6426.2	6782.6	6945.4	7150	6936.6	6778.2	6557.1
H3	6068.7	6446	6579.1	6724.3	6584.6	6409.7	6154.5
H4	5905.9	6244.7	6345.9	6518.6	6351.4	6162.2	5940
H5	5791.5	6157.8	6187.5	6366.8	6204	6031.3	5793.7
H6	5817.9	6130.3	6189.7	6370.1	6199.6	5935.6	5683.7
H7	5970.8	6329.4	6381.1	6514.2	6347	5852	5472.5
H8	6870.6	7134.6	7213.8	7322.7	7132.4	6275.5	5673.8
H9	7884.8	8191.7	8252.2	8377.6	8101.5	7022.4	6113.8
H10	8679	9012.3	9156.4	9258.7	8867.1	7770.4	6597.8
H11	8983.7	9288.4	9435.8	9573.3	9111.3	8153.2	7027.9
H12	9266.4	9557.9	9708.6	9823	9243.3	8450.2	7422.8
H13	9473.2	9774.6	9948.4	10044.1	9367.6	8647.1	7620.8
H14	9553.5	9810.9	10063.9	10060.6	9342.3	8609.7	7560.3
H15	9320.3	9593.1	9905.5	9790	9075	8250	7327.1
H16	8873.7	9249.9	9551.3	9509.5	8791.2	7857.3	7145.6
H17	8602	8948.5	9256.5	9282.9	8624	7724.2	7102.7
H18	8461.2	8945.2	9102.5	9311.5	8605.3	7742.9	7249
H19	8533.8	9023.3	9002.4	9276.3	8748.3	7824.3	7355.7
H20	8544.8	8993.6	8950.7	9124.5	8641.6	7801.2	7419.5
H21	8545.9	8983.7	8902.3	9016.7	8655.9	7781.4	7576.8
H22	8907.8	9130	9152	9103.6	8771.4	8241.2	8207.1
H23	8651.5	8742.8	8892.4	8747.2	8421.6	8074	8273.1
H24	8033.3	8206	8368.8	8214.8	7901.3	7604.3	7898

Table F.9: Total demand of the fourth week.

	W4D1 [MW]	W4D2 [MW]	W4D3 [MW]	W4D4 [MW]	W4D5 [MW]	W4D6 [MW]	W4D7 [MW]
H1	5232.7	5446.1	5478	5581.4	5563.8	5537.4	5416.4
H2	4804.8	5093	5073.2	5161.2	5138.1	5190.9	5058.9
H3	4628.8	4891.7	4873	4952.2	4801.5	4957.7	4771.8
H4	4543	4792.7	4761.9	4792.7	4722.3	4766.3	4621.1
H5	4514.4	4781.7	4739.9	4765.2	4774	4723.4	4554
H6	4634.3	4846.6	4842.2	4884	4901.6	4735.5	4548.5
H7	5135.9	5361.4	5327.3	5353.7	5384.5	4909.3	4524.3
H8	5930.1	6069.8	6089.6	6096.2	6076.4	5172.2	4527.6
H9	6593.4	6762.8	6760.6	6762.8	6761.7	5858.6	4941.2
H10	7243.5	7353.5	7365.6	7321.6	7207.2	6548.3	5539.6
H11	7372.2	7476.7	7445.9	7471.2	7395.3	6956.4	6053.3
H12	7551.5	7647.2	7550.4	7653.8	7558.1	7230.3	6392.1
H13	7653.8	7730.8	7636.2	7758.3	7630.7	7366.7	6554.9
H14	7667	7659.3	7590	7773.7	7607.6	7310.6	6381.1
H15	7528.4	7410.7	7401.9	7489.9	7389.8	6893.7	5843.2
H16	7060.9	7020.2	6949.8	7104.9	6991.6	6342.6	5434
H17	6699	6773.8	6758.4	6871.7	6780.4	6196.3	5362.5
H18	6754	6910.2	6749.6	6954.2	6856.3	6255.7	5580.3
H19	6777.1	7014.7	6803.5	7071.9	7015.8	6477.9	5952.1
H20	7487.7	7825.4	7626.3	7851.8	7793.5	7394.2	6785.9
H21	7958.5	8220.3	8032.2	8173	8024.5	7527.3	7004.8
H22	7558.1	7617.5	7555.9	7579	7452.5	7064.2	6794.7
H23	6759.5	6748.5	6745.2	6831	6713.3	6411.9	6257.9
H24	6067.6	6083	6167.7	6178.7	6144.6	5901.5	5746.4

Table F.10: Demand ratio for the first week in area one. The remaining demand is the corresponding one for area two (the sum of the ratios equals 1).

	W1D1 [1]	W1D2 [1]	W1D3 [1]	W1D4 [1]	W1D5 [1]	W1D6 [1]	W1D7 [1]
H1	0.32312	0.32556	0.29534	0.35962	0.35412	0.2046	0.39157
H2	0.25831	0.29827	0.29454	0.30557	0.44665	0.30874	0.39896
H3	0.46294	0.31366	0.39318	0.38824	0.37195	0.28105	0.40782
H4	0.30689	0.36517	0.34613	0.42011	0.43973	0.40291	0.37668
H5	0.33406	0.33531	0.41071	0.42112	0.30798	0.37343	0.45013
H6	0.41538	0.38936	0.40568	0.32559	0.3944	0.36362	0.30179
H7	0.37168	0.30558	0.35034	0.35887	0.345	0.29508	0.324
H8	0.33287	0.40735	0.27337	0.3598	0.37723	0.36389	0.351
H9	0.17108	0.40344	0.38848	0.27903	0.33482	0.31492	0.35174
H10	0.21153	0.39047	0.33143	0.33542	0.38002	0.45259	0.38991
H11	0.41749	0.49721	0.36128	0.34011	0.3255	0.36769	0.29907
H12	0.19825	0.27808	0.29413	0.27062	0.31303	0.39118	0.35666
H13	0.31373	0.33374	0.40445	0.39022	0.26441	0.42885	0.38573
H14	0.35315	0.38775	0.34837	0.31517	0.35971	0.3246	0.28243
H15	0.31426	0.28149	0.32237	0.30825	0.45692	0.3359	0.36124
H16	0.36025	0.43558	0.29497	0.36219	0.39198	0.34833	0.37945
H17	0.35621	0.35511	0.27279	0.33922	0.28227	0.41668	0.36469
H18	0.27552	0.36207	0.3457	0.40829	0.40361	0.29363	0.3924
H19	0.27955	0.33404	0.42458	0.4074	0.30195	0.33249	0.40601
H20	0.27914	0.33436	0.38712	0.34476	0.3438	0.36495	0.2237
H21	0.31643	0.39324	0.40308	0.31389	0.27817	0.34886	0.26723
H22	0.41037	0.3515	0.23248	0.22073	0.44804	0.3631	0.33462
H23	0.31414	0.35824	0.38078	0.38334	0.35988	0.43751	0.41286
H24	0.26849	0.31861	0.3126	0.34063	0.41039	0.36428	0.39327

Table F.11: Demand ratio for the second week in area one. The remaining demand is the corresponding one for area two (the sum of the ratios equals 1).

	W2D1 [1]	W2D2 [1]	W2D3 [1]	W2D4 [1]	W2D5 [1]	W2D6 [1]	W2D7 [1]
H1	0.2263	0.41132	0.27505	0.3374	0.28977	0.38331	0.37639
H2	0.3248	0.31035	0.34311	0.38052	0.42021	0.32608	0.36729
H3	0.39112	0.4555	0.42934	0.31195	0.33972	0.40205	0.32177
H4	0.33995	0.38997	0.40096	0.33728	0.44956	0.39103	0.39008
H5	0.40035	0.32165	0.41926	0.35764	0.34858	0.34494	0.44265
H6	0.38066	0.35007	0.30226	0.37787	0.25347	0.2958	0.34734
H7	0.38295	0.3188	0.38006	0.31615	0.30294	0.44273	0.37182
H8	0.39166	0.34368	0.40859	0.30754	0.35035	0.26001	0.336
H9	0.33109	0.31596	0.37886	0.32653	0.36691	0.39392	0.34577
H10	0.40577	0.41437	0.39182	0.36199	0.32359	0.31761	0.31354
H11	0.31664	0.3391	0.30735	0.30395	0.29716	0.29752	0.30496
H12	0.31023	0.42833	0.32613	0.39534	0.34926	0.28385	0.41628
H13	0.29813	0.31084	0.33488	0.34278	0.26926	0.37587	0.42623
H14	0.35102	0.36553	0.32921	0.28018	0.39079	0.34878	0.32488
H15	0.31905	0.31723	0.34785	0.35948	0.44989	0.35812	0.33468
H16	0.25985	0.37687	0.39744	0.39219	0.38741	0.38028	0.39696
H17	0.34735	0.33357	0.32292	0.25283	0.35997	0.33641	0.28813
H18	0.35889	0.2973	0.39106	0.36235	0.36864	0.33207	0.34835
H19	0.26138	0.44898	0.4036	0.32792	0.39237	0.36762	0.34382
H20	0.47544	0.44337	0.4037	0.39018	0.28314	0.34338	0.3363
H21	0.37283	0.44162	0.30652	0.33562	0.43082	0.21632	0.31229
H22	0.22848	0.30757	0.30095	0.32665	0.37459	0.35038	0.27378
H23	0.37357	0.32974	0.43794	0.34977	0.39847	0.31448	0.3631
H24	0.37802	0.38512	0.3574	0.3337	0.37371	0.37313	0.43653

Table F.12: Demand ratio for the third week in area one. The remaining demand is the corresponding one for area two (the sum of the ratios equals 1).

	W3D1 [1]	W3D2 [1]	W3D3 [1]	W3D4 [1]	W3D5 [1]	W3D6 [1]	W3D7 [1]
H1	0.27004	0.37381	0.30709	0.29071	0.28971	0.40617	0.37734
H2	0.36962	0.36399	0.39075	0.33881	0.40106	0.32109	0.38677
H3	0.33896	0.31487	0.35786	0.34598	0.35955	0.32948	0.3878
H4	0.37969	0.28959	0.37856	0.36655	0.35136	0.29987	0.25126
H5	0.45252	0.36238	0.34496	0.36923	0.37776	0.32546	0.36649
H6	0.34998	0.27437	0.32018	0.2971	0.39971	0.28698	0.43045
H7	0.26532	0.38625	0.31439	0.37658	0.28639	0.45024	0.30257
H8	0.36658	0.37327	0.34322	0.35236	0.31661	0.36523	0.34875
H9	0.40751	0.36935	0.39735	0.40169	0.39554	0.36816	0.36779
H10	0.39953	0.3831	0.39235	0.31981	0.40263	0.35289	0.32416
H11	0.33179	0.3065	0.33538	0.33754	0.32864	0.29664	0.30563
H12	0.37102	0.35935	0.45922	0.32792	0.38444	0.38382	0.38186
H13	0.30082	0.25742	0.40872	0.35135	0.40991	0.33738	0.34446
H14	0.36379	0.35031	0.42829	0.3871	0.36237	0.44553	0.36553
H15	0.45428	0.25669	0.3664	0.35427	0.39224	0.39577	0.37891
H16	0.37197	0.4715	0.32574	0.41421	0.30752	0.30299	0.32622
H17	0.3144	0.33191	0.23176	0.36693	0.23467	0.43717	0.31393
H18	0.29868	0.33583	0.35021	0.23464	0.43505	0.27926	0.41851
H19	0.28575	0.25324	0.40338	0.26638	0.30812	0.37421	0.29449
H20	0.29996	0.33691	0.37268	0.30661	0.32888	0.32361	0.34256
H21	0.29633	0.26104	0.32709	0.3281	0.39213	0.40881	0.3618
H22	0.2553	0.25536	0.29891	0.30867	0.30192	0.31226	0.42942
H23	0.27328	0.38102	0.37295	0.33052	0.39442	0.35007	0.36286
H24	0.34122	0.40953	0.35023	0.30341	0.44411	0.2873	0.35348

Table F.13: Demand ratio for the fourth week in area one. The remaining demand is the corresponding one for area two (the sum of the ratios equals 1).

	W4D1 [1]	W4D2 [1]	W4D3 [1]	W4D4 [1]	W4D5 [1]	W4D6 [1]	W4D7 [1]
H1	0.32247	0.32463	0.39213	0.3335	0.40069	0.27757	0.36252
H2	0.31726	0.41111	0.3312	0.42001	0.32455	0.3278	0.37344
H3	0.33566	0.34892	0.40609	0.37609	0.28677	0.42439	0.3428
H4	0.27908	0.39429	0.30876	0.29975	0.35788	0.37356	0.32329
H5	0.40595	0.35958	0.36069	0.40689	0.39493	0.35726	0.36425
H6	0.35373	0.31552	0.314	0.35766	0.4059	0.3181	0.34909
H7	0.29579	0.25203	0.39967	0.39871	0.41209	0.31436	0.4325
H8	0.35073	0.3726	0.31091	0.39898	0.33008	0.31515	0.40571
H9	0.31785	0.32334	0.36381	0.39607	0.33303	0.42235	0.36867
H10	0.43099	0.37285	0.35363	0.30884	0.38918	0.42629	0.35852
H11	0.33256	0.34553	0.36359	0.34967	0.3123	0.28272	0.40335
H12	0.29583	0.21087	0.31644	0.33129	0.40328	0.30883	0.40903
H13	0.46201	0.34093	0.25986	0.37776	0.27563	0.32939	0.31935
H14	0.36178	0.37985	0.3052	0.35623	0.45639	0.27269	0.34326
H15	0.41266	0.29989	0.34853	0.30815	0.36369	0.25765	0.28115
H16	0.40151	0.30376	0.38738	0.41437	0.31233	0.38861	0.32271
H17	0.32036	0.42608	0.3458	0.38793	0.34939	0.3479	0.40828
H18	0.29549	0.382	0.39988	0.29897	0.34755	0.32845	0.46093
H19	0.36047	0.3724	0.35437	0.26528	0.37543	0.34398	0.34304
H20	0.37596	0.36268	0.3029	0.39316	0.25174	0.33986	0.36136
H21	0.25989	0.34475	0.35437	0.32804	0.37302	0.34029	0.32227
H22	0.35431	0.37802	0.29036	0.41234	0.3151	0.29567	0.3043
H23	0.33097	0.32833	0.37491	0.31711	0.28944	0.48744	0.36001
H24	0.3674	0.32646	0.32812	0.34334	0.36699	0.35029	0.34256

Table F.14: SPV capacity factors for the first week in area one.

	W1D1 [1]	W1D2 [1]	W1D3 [1]	W1D4 [1]	W1D5 [1]	W1D6 [1]	W1D7 [1]
H1	0	0	0	0	0	0	0
H2	0	0	0	0	0	0	0
H3	0	0	0	0	0	0	0
H4	0	0	0	0	0	0	0
H5	0	0	0	0	0	0	0
H6	0	0	0	0	0	0	0
H7	0.00041321	0	0	0.00041078	0	0	0
H8	0.052888	0.043209	0.0084207	0.043263	0.019432	0.023402	0.019059
H9	0.23657	0.20083	0.032987	0.14467	0.0819	0.10523	0.12552
H10	0.37666	0.39531	0.047562	0.20406	0.13449	0.22842	0.29073
H11	0.47178	0.53128	0.062995	0.22787	0.15617	0.31001	0.45487
H12	0.58539	0.57346	0.063914	0.21188	0.18427	0.30261	0.48634
H13	0.52839	0.55275	0.046781	0.16224	0.1902	0.25796	0.43986
H14	0.4888	0.48513	0.028402	0.10433	0.17428	0.17853	0.35781
H15	0.27318	0.25039	0.041655	0.039057	0.12679	0.086689	0.24992
H16	0.081101	0.066896	0.012411	0.0094945	0.039141	0.027838	0.091093
H17	0.00077022	0.0011484	0.00042678	0	0.0012706	0.00040161	0.0018657
H18	0	0	0	0	0	0	0
H19	0	0	0	0	0	0	0
H20	0	0	0	0	0	0	0
H21	0	0	0	0	0	0	0
H22	0	0	0	0	0	0	0
H23	0	0	0	0	0	0	0
H24	0	0	0	0	0	0	0

Table F.15: SPV capacity factors for the first week in area two.

	W1D1 [1]	W1D2 [1]	W1D3 [1]	W1D4 [1]	W1D5 [1]	W1D6 [1]	W1D7 [1]
H1	0	0	0	0	0	0	0
H2	0	0	0	0	0	0	0
H3	0	0	0	0	0	0	0
H4	0	0	0	0	0	0	0
H5	0	0	0	0	0	0	0
H6	0	0	0	0	0	0	0
H7	0.00041098	0	0	0.00038796	0	0	0
H8	0.048591	0.051616	0.0081423	0.041984	0.017953	0.024124	0.020737
H9	0.22098	0.20571	0.028759	0.15259	0.073784	0.10885	0.12229
H10	0.36808	0.44191	0.051303	0.18223	0.12831	0.26273	0.29729
H11	0.50748	0.53635	0.063393	0.22974	0.15168	0.28881	0.42832
H12	0.58686	0.54344	0.063701	0.18722	0.16712	0.31098	0.45223
H13	0.57752	0.49653	0.046411	0.15533	0.16453	0.26376	0.42552
H14	0.50084	0.47121	0.030894	0.1087	0.17781	0.17864	0.37505
H15	0.29421	0.26345	0.041366	0.045642	0.12608	0.087104	0.25377
H16	0.079986	0.066346	0.01251	0.0088977	0.039058	0.027973	0.093368
H17	0.00081173	0.0011747	0.00040428	0	0.00125	0.00039634	0.0019799
H18	0	0	0	0	0	0	0
H19	0	0	0	0	0	0	0
H20	0	0	0	0	0	0	0
H21	0	0	0	0	0	0	0
H22	0	0	0	0	0	0	0
H23	0	0	0	0	0	0	0
H24	0	0	0	0	0	0	0

Table F.16: SPV capacity factors for the second week in area one.

	W2D1 [1]	W2D2 [1]	W2D3 [1]	W2D4 [1]	W2D5 [1]	W2D6 [1]	W2D7 [1]
H1	0	0	0	0	0	0	0
H2	0	0	0	0	0	0	0
H3	0	0	0	0	0	0	0
H4	0	0	0	0	0	0	0
H5	0	0	0	0	0	0	0
H6	0.10912	0.01468	0.011386	0.0222	0	0.0054802	0.023298
H7	0.26294	0.48015	0.090911	0.136	0.11358	0.16912	0.1603
H8	0.5607	0.17761	0.22425	0.3585	0.36783	0.34668	0.50753
H9	0.42956	0.39227	0.37307	0.53768	0.52194	0.63949	0.47871
H10	0.57607	0.53448	0.52281	0.50459	0.61119	0.50601	0.47942
H11	0.52114	0.47835	0.61845	0.69873	0.68278	0.74396	0.67243
H12	0.69595	0.54081	0.65916	0.68969	0.71035	0.67631	0.61647
H13	0.53577	0.60237	0.51916	0.68364	0.69974	0.713	0.49849
H14	0.50654	0.55953	0.27168	0.59418	0.59306	0.55793	0.4238
H15	0.46659	0.37896	0.38547	0.53219	0.54137	0.53369	0.34388
H16	0.31298	0.2487	0.27925	0.329	0.28508	0.35091	0.20349
H17	0.12861	0.08307	0.12171	0.1431	0.17315	0.16579	0.10963
H18	0.026989	0.025095	0.019998	0.036522	0.019358	0.034533	0.024621
H19	0	0	0	0	0	0	0
H20	0	0	0	0	0	0	0
H21	0	0	0	0	0	0	0
H22	0	0	0	0	0	0	0
H23	0	0	0	0	0	0	0
H24	0	0	0	0	0	0	0

Table F.17: SPV capacity factors for the second week in area two.

	W2D1 [1]	W2D2 [1]	W2D3 [1]	W2D4 [1]	W2D5 [1]	W2D6 [1]	W2D7 [1]
H1	0	0	0	0	0	0	0
H2	0	0	0	0	0	0	0
H3	0	0	0	0	0	0	0
H4	0	0	0	0	0	0	0
H5	0	0	0	0	0	0	0
H6	0.11003	0.015285	0.011784	0.022109	0	0.0051701	0.025925
H7	0.26136	0.48858	0.098487	0.15544	0.12838	0.15998	0.16797
H8	0.54254	0.18454	0.23282	0.32958	0.33899	0.35822	0.55073
H9	0.41005	0.44447	0.39271	0.53039	0.52834	0.69909	0.48461
H10	0.58312	0.51347	0.4772	0.51718	0.64578	0.5055	0.49784
H11	0.58504	0.48675	0.56164	0.64886	0.69892	0.68773	0.61337
H12	0.73224	0.49972	0.62561	0.69647	0.65965	0.70746	0.61883
H13	0.52896	0.57098	0.55865	0.73121	0.73256	0.72267	0.42836
H14	0.53965	0.55003	0.2909	0.6628	0.65999	0.6409	0.38006
H15	0.50952	0.41353	0.40781	0.5473	0.51102	0.52795	0.32546
H16	0.28669	0.21597	0.24708	0.3468	0.30868	0.33263	0.19796
H17	0.12881	0.082624	0.13632	0.14378	0.16965	0.17533	0.096935
H18	0.026274	0.026361	0.018536	0.037514	0.019448	0.036979	0.025451
H19	0	0	0	0	0	0	0
H20	0	0	0	0	0	0	0
H21	0	0	0	0	0	0	0
H22	0	0	0	0	0	0	0
H23	0	0	0	0	0	0	0
H24	0	0	0	0	0	0	0

Table F.18: SPV capacity factors for the third week in area one.

	W3D1 [1]	W3D2 [1]	W3D3 [1]	W3D4 [1]	W3D5 [1]	W3D6 [1]	W3D7 [1]
H1	0	0	0	0	0	0	0
H2	0	0	0	0	0	0	0
H3	0	0	0	0	0	0	0
H4	0	0	0	0	0	0	0
H5	0.0019119	0.0018506	0.0019159	0.0020948	0.0015584	0.0012398	0.0016443
H6	0.070567	0.069793	0.070398	0.076214	0.063746	0.062647	0.073619
H7	0.2272	0.21734	0.21339	0.22456	0.18763	0.19034	0.22041
H8	0.39289	0.35324	0.3554	0.37639	0.36824	0.30985	0.32041
H9	0.49181	0.50984	0.52655	0.50257	0.50182	0.44523	0.52833
H10	0.66415	0.60788	0.59721	0.54511	0.5576	0.57667	0.59445
H11	0.6943	0.69984	0.68595	0.62735	0.68227	0.62043	0.62129
H12	0.72852	0.64031	0.58998	0.68131	0.61659	0.63982	0.62601
H13	0.6421	0.63689	0.6567	0.61775	0.6019	0.66264	0.64671
H14	0.59042	0.5555	0.58359	0.59381	0.53174	0.59732	0.44639
H15	0.45946	0.41898	0.43807	0.42558	0.45991	0.47043	0.41822
H16	0.3421	0.32824	0.34563	0.34078	0.35261	0.34483	0.3286
H17	0.19973	0.1914	0.19987	0.18979	0.19305	0.20968	0.1942
H18	0.084155	0.086303	0.084562	0.094477	0.095632	0.083058	0.087028
H19	0.015961	0.014201	0.014635	0.015214	0.015557	0.014991	0.015861
H20	0	0	0	0	0	0	0
H21	0	0	0	0	0	0	0
H22	0	0	0	0	0	0	0
H23	0	0	0	0	0	0	0
H24	0	0	0	0	0	0	0

Table F.19: SPV capacity factors for the third week in area two.

	W3D1 [1]	W3D2 [1]	W3D3 [1]	W3D4 [1]	W3D5 [1]	W3D6 [1]	W3D7 [1]
H1	0	0	0	0	0	0	0
H2	0	0	0	0	0	0	0
H3	0	0	0	0	0	0	0
H4	0	0	0	0	0	0	0
H5	0.0018268	0.00208	0.0019679	0.0021007	0.0017183	0.0012314	0.0015253
H6	0.072027	0.066853	0.069583	0.067949	0.067865	0.068471	0.071709
H7	0.21398	0.2065	0.21828	0.22805	0.20332	0.19312	0.22104
H8	0.36321	0.39211	0.39351	0.38147	0.35471	0.35179	0.3542
H9	0.53803	0.54011	0.49414	0.49576	0.52242	0.43435	0.5066
H10	0.57667	0.60167	0.57784	0.57581	0.60208	0.57164	0.59167
H11	0.61842	0.63201	0.63954	0.64169	0.69998	0.63462	0.62085
H12	0.74083	0.59599	0.63777	0.61945	0.62997	0.65602	0.61212
H13	0.58092	0.61568	0.66085	0.58578	0.60464	0.5781	0.56891
H14	0.57112	0.52826	0.53515	0.57286	0.58706	0.56056	0.51548
H15	0.48892	0.44121	0.44143	0.49132	0.46754	0.45351	0.4618
H16	0.36852	0.2764	0.32195	0.34857	0.34096	0.33522	0.33512
H17	0.216	0.20832	0.19692	0.21024	0.19251	0.21222	0.19846
H18	0.083658	0.084673	0.079808	0.088167	0.093211	0.087598	0.093406
H19	0.01519	0.013941	0.01334	0.015438	0.016201	0.01481	0.014406
H20	0	0	0	0	0	0	0
H21	0	0	0	0	0	0	0
H22	0	0	0	0	0	0	0
H23	0	0	0	0	0	0	0
H24	0	0	0	0	0	0	0

Table F.20: SPV capacity factors for the fourth week in area one.

	W4D1 [1]	W4D2 [1]	W4D3 [1]	W4D4 [1]	W4D5 [1]	W4D6 [1]	W4D7 [1]
H1	0	0	0	0	0	0	0
H2	0	0	0	0	0	0	0
H3	0	0	0	0	0	0	0
H4	0	0	0	0	0	0	0
H5	0	0	0	0	0	0	0
H6	0.010457	0.0095215	0.0076058	0.00707	0.0067303	0.0049736	0.0041989
H7	0.14204	0.13556	0.12642	0.10854	0.12152	0.11737	0.10836
H8	0.30936	0.3383	0.31163	0.25418	0.2924	0.29005	0.2654
H9	0.48457	0.50549	0.42201	0.42942	0.50703	0.43363	0.42641
H10	0.58592	0.56737	0.53151	0.54258	0.54764	0.57962	0.48366
H11	0.63528	0.67617	0.53563	0.60048	0.64531	0.62961	0.61459
H12	0.67097	0.76445	0.66259	0.61645	0.6178	0.60321	0.62568
H13	0.62769	0.60555	0.48925	0.53879	0.57879	0.56642	0.56418
H14	0.54391	0.52957	0.41994	0.49751	0.59224	0.50535	0.50286
H15	0.43908	0.40468	0.29903	0.37478	0.33963	0.35655	0.33492
H16	0.26527	0.25358	0.18089	0.20928	0.21861	0.19717	0.19984
H17	0.085292	0.083992	0.059204	0.068088	0.07057	0.066844	0.070062
H18	0.0035073	0.0034375	0.0020288	0.0020836	0.0019809	0.0016215	0.00116
H19	0	0	0	0	0	0	0
H20	0	0	0	0	0	0	0
H21	0	0	0	0	0	0	0
H22	0	0	0	0	0	0	0
H23	0	0	0	0	0	0	0
H24	0	0	0	0	0	0	0

Table F.21: SPV capacity factors for the fourth week in area two.

	W4D1 [1]	W4D2 [1]	W4D3 [1]	W4D4 [1]	W4D5 [1]	W4D6 [1]	W4D7 [1]
H1	0	0	0	0	0	0	0
H2	0	0	0	0	0	0	0
H3	0	0	0	0	0	0	0
H4	0	0	0	0	0	0	0
H5	0	0	0	0	0	0	0
H6	0.010345	0.010098	0.0080917	0.0063485	0.0065007	0.0052216	0.0044252
H7	0.14817	0.13872	0.12585	0.12442	0.13488	0.12843	0.10438
H8	0.31337	0.33879	0.26146	0.29	0.29897	0.28487	0.24946
H9	0.50141	0.50723	0.44744	0.43904	0.4287	0.43012	0.39721
H10	0.68552	0.60654	0.52555	0.50911	0.56963	0.59333	0.52481
H11	0.67378	0.64854	0.59991	0.67275	0.62006	0.55621	0.51644
H12	0.74545	0.64325	0.58195	0.65908	0.6631	0.62951	0.58411
H13	0.63939	0.6667	0.50909	0.57105	0.58205	0.60034	0.53239
H14	0.58552	0.56543	0.47852	0.49146	0.5023	0.50119	0.48574
H15	0.44252	0.42818	0.32927	0.37001	0.35025	0.36246	0.37361
H16	0.27496	0.24414	0.17135	0.19574	0.2088	0.18522	0.2132
H17	0.084534	0.079501	0.06053	0.066655	0.063363	0.072793	0.06606
H18	0.0037428	0.0032734	0.0018863	0.0019716	0.002161	0.0015532	0.0012775
H19	0	0	0	0	0	0	0
H20	0	0	0	0	0	0	0
H21	0	0	0	0	0	0	0
H22	0	0	0	0	0	0	0
H23	0	0	0	0	0	0	0
H24	0	0	0	0	0	0	0

Table F.22: WF capacity factors for the first week in area one.

	W1D1 [1]	W1D2 [1]	W1D3 [1]	W1D4 [1]	W1D5 [1]	W1D6 [1]	W1D7 [1]
H1	0.21499	0.098049	0.14336	0.34805	0.5059	0.32556	0.5489
H2	0.16049	0.11753	0.17707	0.31768	0.45171	0.33908	0.52776
H3	0.16328	0.11267	0.18405	0.28936	0.4591	0.38965	0.54526
H4	0.11828	0.078685	0.1954	0.29604	0.39214	0.37815	0.45515
H5	0.098925	0.071696	0.17356	0.25032	0.3998	0.36008	0.46435
H6	0.10206	0.045963	0.2037	0.25387	0.40444	0.33482	0.35437
H7	0.09053	0.030402	0.22906	0.20016	0.40104	0.35082	0.3707
H8	0.059441	0.024689	0.22929	0.17047	0.39998	0.38337	0.28202
H9	0.045337	0.030547	0.28727	0.1447	0.3497	0.26546	0.35205
H10	0.025146	0.031675	0.33989	0.1568	0.35485	0.32774	0.38046
H11	0.019551	0.024525	0.3653	0.17107	0.36748	0.3601	0.43061
H12	0.023666	0.015302	0.42186	0.19108	0.32379	0.45336	0.41691
H13	0.022736	0.015783	0.48165	0.21027	0.36014	0.43936	0.4299
H14	0.032314	0.017854	0.45918	0.25627	0.3737	0.50349	0.44216
H15	0.04108	0.025199	0.43507	0.2928	0.32724	0.52608	0.4099
H16	0.04016	0.02946	0.44136	0.2875	0.30563	0.55073	0.39793
H17	0.038155	0.034833	0.43504	0.29674	0.27526	0.60024	0.42573
H18	0.042527	0.027671	0.43058	0.35578	0.241	0.55563	0.45545
H19	0.073034	0.028417	0.38634	0.376	0.27927	0.52146	0.41281
H20	0.090003	0.038182	0.41776	0.38829	0.28505	0.55078	0.40182
H21	0.093915	0.064271	0.39714	0.43715	0.31836	0.60411	0.42256
H22	0.077153	0.086939	0.36877	0.44838	0.31276	0.56706	0.3997
H23	0.082114	0.1133	0.39696	0.46427	0.29278	0.56403	0.37062
H24	0.074491	0.12982	0.34864	0.44962	0.28438	0.51768	0.32412

Table F.23: WF capacity factors for the first week in area two.

	W1D1 [1]	W1D2 [1]	W1D3 [1]	W1D4 [1]	W1D5 [1]	W1D6 [1]	W1D7 [1]
H1	0.2139	0.10398	0.14706	0.33159	0.46548	0.30836	0.51963
H2	0.17337	0.1217	0.18734	0.30513	0.47166	0.35812	0.53398
H3	0.15688	0.12401	0.18323	0.28757	0.39129	0.40144	0.47275
H4	0.12502	0.095196	0.18477	0.28186	0.43894	0.3819	0.48188
H5	0.097842	0.068092	0.18177	0.25726	0.40623	0.38275	0.42922
H6	0.10426	0.053124	0.22473	0.24519	0.42403	0.3709	0.42442
H7	0.08421	0.028168	0.19704	0.20893	0.391	0.3687	0.38519
H8	0.062303	0.024441	0.27291	0.17347	0.40222	0.3837	0.3377
H9	0.045729	0.030652	0.27451	0.16272	0.39925	0.27989	0.34918
H10	0.024333	0.028527	0.35595	0.15541	0.37736	0.31671	0.38798
H11	0.020895	0.021874	0.37251	0.1736	0.35779	0.36141	0.40671
H12	0.021898	0.015063	0.45406	0.18858	0.35811	0.40614	0.46201
H13	0.022116	0.017502	0.42032	0.1995	0.35097	0.45243	0.42541
H14	0.030357	0.018735	0.44477	0.24292	0.35287	0.47287	0.42958
H15	0.039932	0.025459	0.40119	0.2672	0.34948	0.56608	0.41239
H16	0.036203	0.029443	0.45261	0.26878	0.27297	0.56782	0.35043
H17	0.039486	0.034582	0.40341	0.3153	0.26811	0.54933	0.40353
H18	0.049197	0.02854	0.47041	0.33644	0.26741	0.54876	0.43816
H19	0.072265	0.030565	0.4344	0.367	0.27874	0.57897	0.44142
H20	0.093494	0.038547	0.43944	0.41072	0.30748	0.59094	0.42459
H21	0.094606	0.062648	0.40439	0.47644	0.31303	0.60577	0.3771
H22	0.080135	0.076622	0.38215	0.47574	0.34143	0.51888	0.3832
H23	0.077279	0.10692	0.3736	0.458	0.30962	0.58517	0.33013
H24	0.07977	0.12569	0.38346	0.45391	0.31847	0.57844	0.32985

Table F.24: WF capacity factors for the second week in area one.

	W2D1 [1]	W2D2 [1]	W2D3 [1]	W2D4 [1]	W2D5 [1]	W2D6 [1]	W2D7 [1]
H1	0.093119	0.090834	0.2245	0.1511	0	0.00096982	0.023257
H2	0.069846	0.066284	0.27175	0.084124	0.016512	0.0034433	0.037421
H3	0.067821	0.038353	0.27675	0.16627	0.013503	0.0073597	0.073161
H4	0.056578	0.038924	0.22576	0.17239	0.015939	0.026697	0.043389
H5	0.040919	0.030493	0.22728	0.12563	0.0014967	0.026257	0.047626
H6	0.027906	0.046041	0.33969	0.0086264	0.016227	0.022646	0.068528
H7	0.02915	0.057031	0.33796	0.14448	0.027019	0.026653	0.070298
H8	0.10379	0.048099	0.20125	0.068848	0.021999	0.02142	0.076815
H9	0.01987	0.04683	0.29199	0.053735	0.022109	0.020115	0.079028
H10	0.018614	0.025999	0.18028	0.041163	0.004883	0.017374	0.085945
H11	0.044971	0.020866	0.24739	0.03125	0.01547	0.021205	0.117
H12	0.044463	0.036057	0.23907	0.073967	0.0025005	0.027482	0.15662
H13	0.062192	0.03963	0.2033	0.058515	0.020266	0.022991	0.14256
H14	0.072224	0.047621	0.18386	0.040901	0.020999	0.022561	0.13065
H15	0.079803	0.038208	0.18581	0.031714	0.017862	0.020387	0.2071
H16	0.081187	0.037342	0.22279	0.020469	0.012633	0.01523	0.17489
H17	0.088447	0.084901	0.20595	0.0043235	0.0048457	0.0067093	0.16997
H18	0.11474	0.05434	0.24101	0.0036199	0.0032987	0.0062626	0.16261
H19	0.1135	0.089777	0.23565	0.0039875	0.0020168	0.0059217	0.15051
H20	0.1276	0.13542	0.21817	0.0048408	0.0025437	0.0081224	0.1691
H21	0.13415	0.16785	0.19831	0.0042843	0.0034783	0.012118	0.16271
H22	0.11333	0.18844	0.17022	0.0014373	0.0014553	0.011625	0.22591
H23	0.11722	0.20417	0.15464	0.00088686	0.00093708	0.018997	0.20052
H24	0.11578	0.317	0.16144	0.00045838	0.00093422	0.083661	0.19717

Table F.25: WF capacity factors for the second week in area two.

	W2D1 [1]	W2D2 [1]	W2D3 [1]	W2D4 [1]	W2D5 [1]	W2D6 [1]	W2D7 [1]
H1	0.099084	0.084505	0.24476	0.1634	0	0.00099296	0.026134
H2	0.065692	0.057928	0.26266	0.092438	0.017176	0.0033691	0.036443
H3	0.065193	0.03762	0.27942	0.17837	0.012481	0.0073776	0.078771
H4	0.06286	0.038782	0.21263	0.17806	0.015365	0.02618	0.041615
H5	0.041148	0.028821	0.2128	0.12558	0.0014177	0.026157	0.044646
H6	0.027226	0.046983	0.32056	0.008783	0.016476	0.022331	0.070794
H7	0.031919	0.051012	0.31889	0.15497	0.025502	0.02608	0.06641
H8	0.10624	0.047028	0.19075	0.069704	0.02322	0.02042	0.075196
H9	0.01746	0.045437	0.2757	0.062013	0.019207	0.017931	0.075503
H10	0.019294	0.031546	0.17911	0.043387	0.0050438	0.019558	0.08882
H11	0.044441	0.020011	0.28189	0.031352	0.01797	0.025668	0.11544
H12	0.048434	0.034429	0.2564	0.070998	0.0024634	0.027402	0.15423
H13	0.057602	0.045535	0.23362	0.060737	0.019671	0.024077	0.1444
H14	0.079495	0.043393	0.20148	0.043211	0.019439	0.021447	0.12422
H15	0.079112	0.035078	0.17511	0.029388	0.018002	0.018624	0.22567
H16	0.070467	0.036019	0.20763	0.018041	0.012516	0.016267	0.17345
H17	0.079467	0.089093	0.21997	0.004111	0.0050228	0.0070099	0.16489
H18	0.096905	0.051539	0.23591	0.0041267	0.0033463	0.0063747	0.15956
H19	0.11598	0.077027	0.21883	0.0037998	0.0020135	0.0061528	0.15955
H20	0.12585	0.15464	0.1985	0.0043265	0.0022578	0.0090889	0.14136
H21	0.1314	0.18057	0.19658	0.0044785	0.0037489	0.013104	0.17471
H22	0.10696	0.20006	0.19304	0.0013712	0.0014004	0.012442	0.23708
H23	0.12457	0.20358	0.17379	0.0009519	0.00091734	0.018293	0.20492
H24	0.1157	0.31193	0.14538	0.00047517	0.00095965	0.07857	0.18679

Table F.26: WF capacity factors for the third week in area one.

	W3D1 [1]	W3D2 [1]	W3D3 [1]	W3D4 [1]	W3D5 [1]	W3D6 [1]	W3D7 [1]
H1	0.0076309	0.061625	0.11444	0.14758	0.30675	0.14295	0.067629
H2	0.0083374	0.090674	0.113	0.18891	0.32494	0.18642	0.17637
H3	0.010893	0.11052	0.14378	0.20017	0.30059	0.18394	0.21504
H4	0.01206	0.11174	0.15498	0.26363	0.35002	0.16829	0.2002
H5	0.018509	0.13017	0.16374	0.3199	0.37459	0.17285	0.15937
H6	0.014892	0.11931	0.1517	0.31654	0.41749	0.16309	0.15865
H7	0.013642	0.10332	0.13824	0.36582	0.37014	0.16396	0.15822
H8	0.016364	0.10504	0.12826	0.39142	0.3811	0.14207	0.15113
H9	0.02206	0.12421	0.12606	0.38907	0.37392	0.13493	0.18089
H10	0.02503	0.13673	0.1352	0.43057	0.36456	0.13593	0.24537
H11	0.036788	0.13399	0.10832	0.38553	0.36436	0.11079	0.25063
H12	0.040956	0.1244	0.11583	0.38456	0.32348	0.11879	0.25971
H13	0.049306	0.10328	0.12561	0.33695	0.3229	0.11279	0.28664
H14	0.044279	0.093801	0.13647	0.34901	0.28167	0.12637	0.28502
H15	0.054167	0.10218	0.16647	0.35954	0.25035	0.11192	0.26724
H16	0.058186	0.1074	0.16226	0.34135	0.22441	0.10231	0.24983
H17	0.059319	0.09258	0.1691	0.33547	0.19319	0.09	0.24191
H18	0.056837	0.067528	0.12937	0.31235	0.1428	0.078008	0.2314
H19	0.062321	0.07934	0.12675	0.2462	0.15766	0.069429	0.2258
H20	0.056216	0.081891	0.13825	0.32438	0.16157	0.07088	0.27564
H21	0.060108	0.08247	0.14275	0.29823	0.17066	0.058904	0.29105
H22	0.057409	0.090868	0.13722	0.28618	0.15013	0.068612	0.28328
H23	0.058017	0.083889	0.1349	0.2587	0.13124	0.080318	0.31143
H24	0.052335	0.089316	0.13823	0.314	0.14535	0.07	0.29823

Table F.27: WF capacity factors for the third week in area two.

	W3D1 [1]	W3D2 [1]	W3D3 [1]	W3D4 [1]	W3D5 [1]	W3D6 [1]	W3D7 [1]
H1	0.0075595	0.062447	0.10934	0.15032	0.2918	0.14913	0.069974
H2	0.0077987	0.090232	0.11832	0.18977	0.3173	0.18042	0.17561
H3	0.011529	0.11248	0.12852	0.20017	0.32505	0.18791	0.20178
H4	0.011602	0.11146	0.16786	0.25272	0.32312	0.16974	0.21204
H5	0.016244	0.12374	0.17167	0.29363	0.38002	0.16882	0.19441
H6	0.015055	0.12115	0.16383	0.35476	0.42448	0.16341	0.15719
H7	0.01371	0.091199	0.13614	0.38749	0.39485	0.1855	0.14857
H8	0.015079	0.10897	0.13253	0.40893	0.34084	0.15246	0.15668
H9	0.018917	0.12996	0.12609	0.38914	0.38327	0.13099	0.18976
H10	0.028308	0.13532	0.12633	0.4063	0.36222	0.13725	0.22549
H11	0.032955	0.12554	0.10529	0.38258	0.3874	0.11045	0.25478
H12	0.040253	0.11288	0.10651	0.36185	0.38766	0.12053	0.27857
H13	0.054378	0.098655	0.13069	0.36181	0.29247	0.099402	0.27807
H14	0.051458	0.090142	0.13703	0.35801	0.30079	0.13489	0.28038
H15	0.055071	0.10196	0.15907	0.36205	0.2625	0.12473	0.28218
H16	0.057834	0.10698	0.16822	0.34753	0.22348	0.10779	0.24206
H17	0.060938	0.093161	0.16339	0.33713	0.16931	0.091447	0.2211
H18	0.055547	0.083521	0.12318	0.28893	0.14639	0.07619	0.20079
H19	0.055095	0.077983	0.1136	0.30188	0.1534	0.068732	0.24442
H20	0.058758	0.098942	0.1375	0.31185	0.15554	0.067971	0.27975
H21	0.064697	0.086139	0.13909	0.28805	0.16355	0.060291	0.26019
H22	0.053338	0.092628	0.13351	0.29532	0.16794	0.07136	0.2769
H23	0.059288	0.089903	0.12826	0.27045	0.13655	0.080019	0.30115
H24	0.055174	0.08484	0.13465	0.29591	0.13442	0.073952	0.3213

Table F.28: WF capacity factors for the fourth week in area one.

	W4D1 [1]	W4D2 [1]	W4D3 [1]	W4D4 [1]	W4D5 [1]	W4D6 [1]	W4D7 [1]
H1	0.054069	0.027796	0.023216	0.097975	0.04213	0.16014	0.067871
H2	0.070785	0.033057	0.020143	0.097906	0.049287	0.15213	0.073155
H3	0.073237	0.038317	0.016565	0.07138	0.055521	0.15939	0.072434
H4	0.085467	0.031063	0.014026	0.070868	0.045402	0.1617	0.066236
H5	0.091178	0.040897	0.014875	0.070333	0.04942	0.14697	0.07038
H6	0.092742	0.041413	0.018774	0.061986	0.05119	0.16695	0.064544
H7	0.07726	0.042333	0.015257	0.071066	0.036889	0.17235	0.067094
H8	0.078519	0.029866	0.0083928	0.038763	0.024341	0.14429	0.052032
H9	0.071618	0.022748	0.010159	0.028983	0.022985	0.11311	0.035382
H10	0.045716	0.017496	0.010719	0.030573	0.020703	0.12467	0.025442
H11	0.049206	0.017529	0.012086	0.045666	0.03393	0.12363	0.033649
H12	0.051004	0.020947	0.016995	0.055569	0.068005	0.10588	0.044864
H13	0.03562	0.023632	0.031624	0.078437	0.097152	0.12646	0.055642
H14	0.040022	0.026419	0.024803	0.08643	0.10991	0.11791	0.056456
H15	0.036061	0.029386	0.020815	0.070785	0.13971	0.12111	0.061948
H16	0.034078	0.027845	0.012226	0.079933	0.13203	0.13292	0.058851
H17	0.025684	0.021418	0.015968	0.062249	0.14252	0.12374	0.060382
H18	0.021174	0.019897	0.027979	0.05174	0.14434	0.1244	0.055488
H19	0.018219	0.016045	0.038722	0.043243	0.13124	0.106	0.047435
H20	0.016854	0.018222	0.06201	0.038299	0.17254	0.090824	0.050035
H21	0.015422	0.02267	0.068816	0.035893	0.17359	0.10199	0.058683
H22	0.017304	0.028016	0.080758	0.03713	0.16691	0.097504	0.035316
H23	0.022133	0.036926	0.089587	0.036703	0.14802	0.083662	0.031995
H24	0.022821	0.036989	0.09143	0.045451	0.17018	0.075567	0.031077

Table F.29: WF capacity factors for the fourth week in area two.

	W4D1 [1]	W4D2 [1]	W4D3 [1]	W4D4 [1]	W4D5 [1]	W4D6 [1]	W4D7 [1]
H1	0.053766	0.027943	0.025325	0.10318	0.035405	0.17329	0.070455
H2	0.074392	0.031537	0.020625	0.091683	0.04705	0.16361	0.068976
H3	0.080957	0.037033	0.016379	0.07819	0.054187	0.15908	0.072922
H4	0.084048	0.033586	0.015402	0.070304	0.048553	0.1443	0.066793
H5	0.090927	0.041768	0.015214	0.071118	0.049842	0.13879	0.065699
H6	0.091724	0.044639	0.017787	0.061276	0.050365	0.18283	0.063445
H7	0.077825	0.045918	0.014785	0.070878	0.043696	0.16834	0.063904
H8	0.082774	0.028265	0.0083193	0.040729	0.027446	0.14485	0.051353
H9	0.066935	0.023801	0.010015	0.032367	0.025135	0.11073	0.037223
H10	0.051977	0.01676	0.010407	0.03128	0.0232	0.12812	0.027049
H11	0.044223	0.015201	0.0115	0.044227	0.037163	0.12611	0.038299
H12	0.047062	0.020596	0.017027	0.052623	0.061417	0.11246	0.05085
H13	0.039651	0.023471	0.029495	0.075456	0.093064	0.13783	0.057066
H14	0.039027	0.029489	0.026459	0.085185	0.11161	0.12473	0.056516
H15	0.038174	0.02524	0.019355	0.085675	0.12338	0.11003	0.053418
H16	0.035442	0.026679	0.013436	0.07335	0.13444	0.13576	0.058501
H17	0.028343	0.022771	0.015912	0.058601	0.15293	0.11557	0.057105
H18	0.023029	0.018569	0.025629	0.052324	0.13896	0.12212	0.051909
H19	0.018027	0.016828	0.043029	0.041557	0.13953	0.11241	0.052138
H20	0.016326	0.01774	0.060595	0.037697	0.16368	0.092757	0.051425
H21	0.014742	0.02151	0.070097	0.036949	0.16197	0.1056	0.051045
H22	0.017445	0.028137	0.079261	0.039324	0.13741	0.1091	0.03316
H23	0.020394	0.038046	0.086013	0.03875	0.15054	0.081623	0.035567
H24	0.023813	0.03659	0.084324	0.047645	0.14276	0.078695	0.030447

List of Figures

List of Tables

Bibliography

[1] H. Sadeghi, M. Rashidinejad, A. Abdollahi, A comprehensive sequential review study through the generation expansion planning, Renewable and Sustainable Energy Reviews 67 (2017) 1369–1394.

[2] N. E. Koltsaklis, A. S. Dagoumas, State-of-the-art generation expansion planning: A review, Applied Energy 230 (2018) 563–589.

[3] A. S. Dagoumas, N. E. Koltsaklis, Review of models for integrating renewable energy in the generation expansion planning, Applied Energy 242 (2019) 1573–1587.

[4] V. Oree, S. Z. S. Hassen, P. J. Fleming, Generation expansion planning optimisation with renewable energy integration: A review, Renewable and Sustainable Energy Reviews 69 (2017) 790–803.

[5] L. Gacitua, P. Gallegos, R. Henriquez-Auba, A. Lorca, M. Negrete-Pincetic, D. Olivares, A. Valenzuela, G. Wenzel, A comprehensive review on expansion planning: Models and tools for energy policy analysis, Renewable and Sustainable Energy Reviews 98 (2018) 346–360.

[6] S. Pereira, P. Ferreira, A. I. F. Vaz, Optimization modeling to support renewables integration in power systems, Renewable and Sustainable Energy Reviews 55 (2016) 316–325.

[7] B. Palmintier, M. Webster, Impact of unit commitment constraints on generation expansion planning with renewables, in: Power and energy society general meeting, 2011 IEEE, IEEE, 2011, pp. 1–7.

[8] B. S. Palmintier, M. D. Webster, Impact of operational flexibility on electricity generation planning with renewable and carbon targets, IEEE Transactions on Sustainable Energy 7 (2) (2016) 672–684.

[9] S. Collins, J. P. Deane, K. Poncelet, E. Panos, R. C. Pietzcker, E. Delarue, B. P. O. Gallachoir, Integrating short term variations of the power system into integrated energy system models: A methodological review, Renewable and Sustainable Energy Reviews 76 (2017) 839–856.

[10] A. Shortt, J. Kiviluoma, M. O'Malley, Accommodating variability in generation planning, IEEE Transactions on Power Systems 28 (1) (2012) 158–169.

[11] R. Mallipeddi, P. N. Suganthan, Unit commitment-a survey and comparison of conventional and nature inspired algorithms, International Journal of Bio-Inspired Computation 6 (2) (2014) 71–90.

[12] K. Deb, Multi-objective optimization using evolutionary algorithms, Vol. 16, John Wiley & Sons, 2001.

[13] A. Zhou, B.-Y. Qu, H. Li, S.-Z. Zhao, P. N. Suganthan, Q. Zhang, Multiobjective evolutionary algorithms: A survey of the state of the art, Swarm and Evolutionary Computation 1 (1) (2011) 32–49.

[14] Y. Jin, Surrogate-assisted evolutionary computation: Recent advances and future challenges, Swarm and Evolutionary Computation 1 (2) (2011) 61–70.

[15] A. Díaz-Manríquez, G. Toscano, J. H. Barron-Zambrano, E. Tello-Leal, A review of surrogate assisted multiobjective evolutionary algorithms, Computational intelligence and neuroscience 2016 (2016).

[16] Y. Jin, B. Sendhoff, Fitness approximation in evolutionary computation-a survey, in: Proceedings of the 4th Annual Conference on Genetic and Evolutionary Computation, 2002, pp. 1105–1112.

[17] C. Vrionis, V. Tsalavoutis, A. Tolis, A generation expansion planning model for integrating high shares of renewable energy: A meta-model assisted evolutionary algorithm approach, Applied Energy 259 (2020) 114085.

[18] A. J. Conejo, L. Baringo, S. J. Kazempour, A. S. Siddiqui, Investment in electricity generation and transmission, Cham Zug, Switzerland: Springer International Publishing 106 (2016).

[19] L. Bird, J. Cochran, X. Wang, Wind and solar energy curtailment: Experience and practices in the united states, Tech. rep., National Renewable Energy Lab.(NREL), Golden, CO (United States) (2014).

[20] Y. Cui, Z. Geng, Q. Zhu, Y. Han, Multi-objective optimization methods and application in energy saving, Energy 125 (2017) 681–704.

[21] E. Løken, Use of multicriteria decision analysis methods for energy planning problems, Renewable and sustainable energy reviews 11 (7) (2007) 1584–1595.

[22] D. Phillips, A mathematical model for determining generation plant mix, in: Proceeding of the Third Power Systems Computation Conference, 1969.

[23] C. I. Nweke, F. Leanez, G. R. Drayton, M. Kolhe, Benefits of chronological optimization in capacity planning for electricity markets, in: 2012 IEEE International Conference on Power System Technology (POWERCON), IEEE, 2012, pp. 1–6.

[24] C. Batlle, P. Rodilla, An enhanced screening curves method for considering thermal cycling operation costs in generation expansion planning, IEEE transactions on power systems 28 (4) (2013) 3683–3691.

[25] T. Zhang, R. Baldick, T. Deetjen, Optimized generation capacity expansion using a further improved screening curve method, Electric Power Systems Research 124 (2015) 47–54.

[26] F. Ueckerdt, R. Brecha, G. Luderer, P. Sullivan, E. Schmid, N. Bauer, D. Böttger, R. Pietzcker, Representing power sector variability and the integration of variable renewables in long-term energy-economy models using residual load duration curves, Energy 90 (2015) 1799–1814.

[27] A. Franz, J. Rieck, J. Zimmermann, Fix-and-optimize procedures for solving the long-term unit commitment problem with pumped storages, Annals of Operations Research (2018) 1–25.

[28] N. H. Kjeldsen, M. Chiarandini, Heuristic solutions to the long-term unit commitment problem with cogeneration plants, Computers & Operations Research 39 (2) (2012) 269–282.

[29] X. Han, X. Chen, M. B. McElroy, S. Liao, C. P. Nielsen, J. Wen, Modeling formulation and validation for accelerated simulation and flexibility assessment on large scale power systems under higher renewable penetrations, Applied energy 237 (2019) 145–154.

[30] J. Blazquez, R. Fuentes-Bracamontes, C. A. Bollino, N. Nezamuddin, The renewable energy policy paradox, Renewable and Sustainable Energy Reviews 82 (2018) 1–5.

[31] G. Luderer, V. Krey, K. Calvin, J. Merrick, S. Mima, R. Pietzcker, J. Van Vliet, K. Wada, The role of renewable energy in climate stabilization: results from the emf27 scenarios, Climatic change 123 (3-4) (2014) 427–441.

[32] J. N. Puga, The importance of combined cycle generating plants in integrating large levels of wind power generation, The Electricity Journal 23 (7) (2010) 33–44.

[33] J. Gil, A. Caballero, A. J. Conejo, Power cycling: Ccgts: The critical link between the electricity and natural gas markets, IEEE Power and Energy Magazine 12 (6) (2014) 40–48.

[34] C. De Jonghe, E. Delarue, R. Belmans, W. D'haeseleer, Determining optimal electricity technology mix with high level of wind power penetration, Applied Energy 88 (6) (2011) 2231–2238.

[35] G. Papaefthymiou, K. Dragoon, Towards 100% renewable energy systems: Uncapping power system flexibility, Energy Policy 92 (2016) 69–82.

[36] M. Alizadeh, M. P. Moghaddam, N. Amjady, P. Siano, M. Sheikh-El-Eslami, Flexibility in future power systems with high renewable penetration: A review, Renewable and Sustainable Energy Reviews 57 (2016) 1186–1193.

[37] A. van Stiphout, K. De Vos, G. Deconinck, The impact of operating reserves on investment planning of renewable power systems, IEEE Transactions on Power Systems 32 (1) (2017) 378–388.

[38] M. Welsch, P. Deane, M. Howells, B. Ó. Gallachóir, F. Rogan, M. Bazilian, H.-H. Rogner, Incorporating flexibility requirements into long-term energy system models–a case study on high levels of renewable electricity penetration in ireland, Applied Energy 135 (2014) 600–615.

[39] S. Ludig, M. Haller, E. Schmid, N. Bauer, Fluctuating renewables in a long-term climate change mitigation strategy, Energy 36 (11) (2011) 6674–6685.

[40] A. Pina, C. Silva, P. Ferrão, Modeling hourly electricity dynamics for policy making in long-term scenarios, Energy Policy 39 (9) (2011) 4692–4702.

[41] K. Poncelet, E. Delarue, D. Six, J. Duerinck, W. D'haeseleer, Impact of the level of temporal and operational detail in energy-system planning models, Applied Energy 162 (2016) 631–643.

[42] J. Deane, A. Chiodi, M. Gargiulo, B. P. Ó. Gallachóir, Soft-linking of a power systems model to an energy systems model, Energy 42 (1) (2012) 303–312.

[43] K. Poncelet, H. Höschle, E. Delarue, A. Virag, W. D'haeseleer, Selecting representative days for capturing the implications of integrating intermittent renewables in generation expansion planning problems, IEEE Transactions on Power Systems 32 (3) (2017) 1936–1948.

[44] V. Krishnan, W. Cole, Evaluating the value of high spatial resolution in national capacity expansion models using reeds, in: 2016 IEEE Power and Energy Society General Meeting (PESGM), IEEE, 2016, pp. 1–5.

[45] N. E. Koltsaklis, A. S. Dagoumas, G. M. Kopanos, E. N. Pistikopoulos, M. C. Georgiadis, A spatial multi-period long-term energy planning model: a case study of the greek power system, Applied Energy 115 (2014) 456–482.

[46] B. Hua, R. Baldick, J. Wang, Representing operational flexibility in generation expansion planning through convex relaxation of unit commitment, IEEE Transactions on Power Systems 33 (2) (2018) 2272–2281.

[47] A. Belderbos, E. Delarue, Accounting for flexibility in power system planning with renewables, International Journal of Electrical Power & Energy Systems 71 (2015) 33–41.

[48] X. Chen, J. Lv, M. B. McElroy, X. Han, C. P. Nielsen, J. Wen, Power system capacity expansion under higher penetration of renewables considering flexibility constraints and low carbon policies, IEEE Transactions on Power Systems (2018).

[49] S. Pereira, P. Ferreira, A. I. F. Vaz, Generation expansion planning with high share of renewables of variable output, Applied Energy 190 (2017) 1275–1288.

[50] S. Pereira, P. Ferreira, A. I. F. Vaz, A simplified optimization model to short-term electricity planning, Energy 93 (2015) 2126–2135.

[51] N. E. Koltsaklis, M. C. Georgiadis, A multi-period, multi-regional generation expansion planning model incorporating unit commitment constraints, Applied energy 158 (2015) 310–331.

[52] H. Park, R. Baldick, Multi-year stochastic generation capacity expansion planning under environmental energy policy, Applied energy 183 (2016) 737–745.

[53] A. Flores-Quiroz, R. Palma-Behnke, G. Zakeri, R. Moreno, A column generation approach for solving generation expansion planning problems with high renewable energy penetration, Electric Power Systems Research 136 (2016) 232–241.

[54] C. L. Lara, D. S. Mallapragada, D. J. Papageorgiou, A. Venkatesh, I. E. Grossmann, Deterministic electric power infrastructure planning: Mixed-integer programming model and nested decomposition algorithm, European Journal of Operational Research 271 (3) (2018) 1037–1054.

[55] J. Ma, V. Silva, R. Belhomme, D. S. Kirschen, L. F. Ochoa, Evaluating and planning flexibility in sustainable power systems, in: Power and Energy Society General Meeting (PES), 2013 IEEE, IEEE, 2013, pp. 1–11.

[56] V. Oree, S. Z. S. Hassen, A composite metric for assessing flexibility available in conventional generators of power systems, Applied energy 177 (2016) 683–691.

[57] I. F. Abdin, E. Zio, An integrated framework for operational flexibility assessment in multi-period power system planning with renewable energy production, Applied energy 222 (2018) 898–914.

[58] N. E. Koltsaklis, A. S. Dagoumas, M. C. Georgiadis, G. Papaioannou, C. Dikaiakos, A mid-term, market-based power systems planning model, Applied Energy 179 (2016) 17–35.

[59] T. Levin, A. Botterud, Electricity market design for generator revenue sufficiency with increased variable generation, Energy Policy 87 (2015) 392–406.

[60] C. F. Heuberger, I. Staffell, N. Shah, N. Mac Dowell, A systems approach to quantifying the value of power generation and energy storage technologies in future electricity networks, Computers & Chemical Engineering 107 (2017) 247–256.

[61] T. Brijs, A. van Stiphout, S. Siddiqui, R. Belmans, Evaluating the role of electricity storage by considering short-term operation in long-term planning, Sustainable Energy, Grids and Networks 10 (2017) 104–117.

[62] T. Luz, P. Moura, A. de Almeida, Multi-objective power generation expansion planning with high penetration of renewables, Renewable and Sustainable Energy Reviews 81 (2018) 2637–2643.

[63] P. S. Moura, A. T. de Almeida, Multi-objective optimization of a mixed renewable system with demand-side management, Renewable and Sustainable Energy Reviews 14 (5) (2010) 1461–1468.

[64] J. Aghaei, M. A. Akbari, A. Roosta, A. Baharvandi, Multiobjective generation expansion planning considering power system adequacy, Electric power systems research 102 (2013) 8–19.

[65] V. Oree, S. Z. S. Hassen, P. J. Fleming, A multi-objective framework for long-term generation expansion planning with variable renewables, Applied Energy 253 (2019) 113589.

[66] K. Deb, A. Pratap, S. Agarwal, T. Meyarivan, A fast and elitist multiobjective genetic algorithm: Nsga-ii, IEEE transactions on evolutionary computation 6 (2) (2002) 182–197.

[67] J. Zhu, M.-y. Chow, A review of emerging techniques on generation expansion planning, IEEE Transactions on Power Systems 12 (4) (1997) 1722–1728.

[68] S. Kannan, S. M. R. Slochanal, N. P. Padhy, Application and comparison of metaheuristic techniques to generation expansion planning problem, IEEE Transactions on Power Systems 20 (1) (2005) 466–475.

[69] A. J. Pereira, J. T. Saraiva, A long term generation expansion planning model using system dynamics–case study using data from the portuguese/spanish generation system, Electric Power Systems Research 97 (2013) 41–50.

[70] K. Rajesh, A. Bhuvanesh, S. Kannan, C. Thangaraj, Least cost generation expansion planning with solar power plant using differential evolution algorithm, Renewable Energy 85 (2016) 677–686.

[71] P. Verma, K. Sanyal, D. Srinivsan, K. Swarup, Information exchange based clustered differential evolution for constrained generation-transmission expansion planning, Swarm and evolutionary computation 44 (2019) 863–875.

[72] R. Hemmati, H. Saboori, M. A. Jirdehi, Multistage generation expansion planning incorporating large scale energy storage systems and environmental pollution, Renewable Energy 97 (2016) 636–645.

[73] S. Kannan, S. Baskar, J. D. McCalley, P. Murugan, Application of NSGA-II algorithm to generation expansion planning, IEEE Transactions on Power systems 24 (1) (2009) 454–461.

[74] P. Murugan, S. Kannan, S. Baskar, NSGA-II algorithm for multi-objective generation expansion planning problem, Electric Power Systems Research 79 (4) (2009) 622–628.

[75] P. Murugan, S. Kannan, S. Baskar, Application of nsga-ii algorithm to single-objective transmission constrained generation expansion planning, IEEE Transactions on Power Systems 24 (4) (2009) 1790–1797.

[76] J. Sirikum, A. Techanitisawad, V. Kachitvichyanukul, A new efficient ga-benders' decomposition method: For power generation expansion planning with emission controls, IEEE Transactions on Power Systems 22 (3) (2007) 1092–1100.

[77] V. A. Tsalavoutis, C. G. Vrionis, A. I. Tolis, Relaxation of quantitative energy objectives on generation expansion planning: A computational and policy study, International Transactions on Electrical Energy Systems 27 (12) (2017) e2427.

[78] C. A. Georgopoulou, K. C. Giannakoglou, Metamodel-assisted evolutionary algorithms for the unit commitment problem with probabilistic outages, Applied Energy 87 (5) (2010) 1782–1792.

[79] A. Glotić, A. Zamuda, Short-term combined economic and emission hydrothermal optimization by surrogate differential evolution, Applied Energy 141 (2015) 42–56.

[80] V. Tsalavoutis, C. Vrionis, A. Tolis, An enhanced real coded approach for the optimization of the unit commitment problem, in: 2016 IEEE Symposium Series on Computational Intelligence (SSCI), IEEE, 2016, pp. 1–8.

[81] V. A. Tsalavoutis, C. G. Vrionis, A. I. Tolis, Optimizing a unit commitment problem using an evolutionary algorithm and a plurality of priority lists, Operational Research (2018) 1–54.

[82] V. Tsalavoutis, C. Vrionis, A. Tolis, D. Plataniotis, A differential evolution approach for the reliability constrained unit commitment problem, in: 2018 IEEE Symposium Series on Computational Intelligence (SSCI), IEEE, 2018, pp. 2329–2339.

[83] V. Tsalavoutis, C. Vrionis, A. Tolis, A hybrid multi-objective evolutionary algorithm for economic-environmental generation scheduling, in: Proceedings of the Genetic and Evolutionary Computation Conference, 2019, pp. 1338–1346.

[84] C. Darwin, On the origin of species london, UK: John Murray 62 (1859).

[85] Z. Michalewicz, D. B. Fogel, How to solve it: modern heuristics, Springer Science & Business Media, 2013.

[86] A. E. Eiben, J. E. Smith, et al., Introduction to evolutionary computing, Vol. 53, Springer, 2003.

[87] J. H. Holland, Genetic algorithms and the optimal allocation of trials, SIAM Journal on Computing 2 (2) (1973) 88–105.

[88] J. H. Holland, et al., Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence, MIT press, 1992.

[89] H.-P. Schwefel, Kybernetische evolution als strategie der experimentellen forschung in der stromungstechnik, Diploma thesis, Technical Univ. of Berlin (1965).

[90] I. Rechenberg, Evolution strategy: Optimization of technical systems by means of biological evolution, Fromman-Holzboog, Stuttgart 104 (1973) 15–16.

[91] R. Storn, K. Price, Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces, Journal of global optimization 11 (4) (1997) 341–359.

[92] S. Das, S. S. Mullick, P. N. Suganthan, Recent advances in differential evolution–an updated survey, Swarm and Evolutionary Computation 27 (2016) 1–30.

[93] L. J. Fogel, A. J. Owens, M. J. Walsh, Artificial intelligence through simulated evolution (1966).

[94] P. Larrañaga, J. A. Lozano, Estimation of distribution algorithms: A new tool for evolutionary computation, Vol. 2, Springer Science & Business Media, 2001.

[95] N. L. Cramer, A representation for the adaptive generation of simple sequential programs, in: Proceedings of the first international conference on genetic algorithms, 1985, pp. 183–187.

[96] M. Dorigo, Optimization, learning and natural algorithms, PhD Thesis, Politecnico di Milano (1992).

[97] J. Kennedy, R. Eberhart, Particle swarm optimization, in: Proceedings of ICNN'95-International Conference on Neural Networks, Vol. 4, IEEE, 1995, pp. 1942–1948.

[98] K. A. De Jong, Genetic algorithms are not function optimizers, in: Foundations of genetic algorithms, Vol. 2, Elsevier, 1993, pp. 5–17.

[99] H. J. Bremermann, et al., Optimization through evolution and recombination, Self-organizing systems 93 (1962) 106.

[100] A. S. Fraser, Simulation of genetic systems by automatic digital computers i. introduction, Australian Journal of Biological Sciences 10 (4) (1957) 484–491.

[101] D. E. Goldberg, Genetic algorithms in search, optimisation and machine learning, 1989, Reading, Addison, Wesley.

[102] K. Deb, H. Jain, An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part i: solving problems with box constraints, IEEE transactions on evolutionary computation 18 (4) (2013) 577–601.

[103] H. Jain, K. Deb, An evolutionary many-objective optimization algorithm using reference-point based nondominated sorting approach, part ii: handling constraints and extending to an adaptive approach, IEEE Transactions on evolutionary computation 18 (4) (2013) 602–622.

[104] N. Hansen, A. Ostermeier, A. Gawelczyk, On the adaptation of arbitrary normal mutation distributions in evolution strategies: The generating set adaptation., in: ICGA, 1995, pp. 57–64.

[105] N. Hansen, A. Ostermeier, Adapting arbitrary normal mutation distributions in evolution strategies: The covariance matrix adaptation, in: Proceedings of IEEE international conference on evolutionary computation, IEEE, 1996, pp. 312–317.

[106] A. Ostermeier, A. Gawelczyk, N. Hansen, A derandomized approach to self-adaptation of evolution strategies, Evolutionary Computation 2 (4) (1994) 369–380.

[107] N. Hansen, D. V. Arnold, A. Auger, Evolution strategies, in: Springer handbook of computational intelligence, Springer, 2015, pp. 871–898.

[108] S. Das, P. N. Suganthan, Differential evolution: A survey of the state-of-the-art, IEEE transactions on evolutionary computation 15 (1) (2010) 4–31.

[109] K. Price, R. M. Storn, J. A. Lampinen, Differential evolution: a practical approach to global optimization, Springer Science & Business Media, 2006.

[110] Z. Michalewicz, A survey of constraint handling techniques in evolutionary computation methods., Evolutionary programming 4 (1995) 135–155.

[111] C. A. C. Coello, Theoretical and numerical constraint-handling techniques used with evolutionary algorithms: a survey of the state of the art, Computer methods in applied mechanics and engineering 191 (11-12) (2002) 1245–1287.

[112] E. Mezura-Montes, C. A. C. Coello, Constraint-handling in nature-inspired numerical optimization: past, present and future, Swarm and Evolutionary Computation 1 (4) (2011) 173–194.

[113] F. Hoffmeister, J. Sprave, Problem-independent handling of constraints by use of metric penalty functions (1996).

[114] J. A. Joines, C. R. Houck, On the use of non-stationary penalty functions to solve nonlinear constrained optimization problems with ga's, in: Proceedings of the First IEEE Conference on Evolutionary Computation. IEEE World Congress on Computational Intelligence, IEEE, 1994, pp. 579–584.

[115] B. Tessema, G. G. Yen, An adaptive penalty formulation for constrained evolutionary optimization, IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans 39 (3) (2009) 565–578.

[116] H. J. Barbosa, A. C. Lemonge, H. S. Bernardino, A critical review of adaptive penalty techniques in evolutionary computation, in: Evolutionary constrained optimization, Springer, 2015, pp. 1–27.

[117] D. Powell, M. M. Skolnick, Using genetic algorithms in engineering design optimization with non-linear constraints, in: Proceedings of the 5th International conference on Genetic Algorithms, 1993, pp. 424–431.

[118] R. Hinterding, Z. Michalewicz, Your brains and my beauty: parent matching for constrained optimisation, in: 1998 IEEE International Conference on Evolutionary Computation Proceedings. IEEE World Congress on Computational Intelligence (Cat. No. 98TH8360), IEEE, 1998, pp. 810–815.

[119] K. Deb, An efficient constraint handling method for genetic algorithms, Computer methods in applied mechanics and engineering 186 (2-4) (2000) 311–338.

[120] A. I. Oyman, K. Deb, H.-G. Beyer, An alternative constraint handling method for evolution strategies, in: Proceedings of the 1999 Congress on Evolutionary Computation-CEC99 (Cat. No. 99TH8406), Vol. 1, IEEE, 1999, pp. 612–619.

[121] T. P. Runarsson, X. Yao, Stochastic ranking for constrained evolutionary optimization, IEEE Transactions on evolutionary computation 4 (3) (2000) 284–294.

[122] T. Takahama, S. Sakai, Constrained optimization by the ε constrained differential evolution with an archive and gradient-based mutation, in: IEEE congress on evolutionary computation, IEEE, 2010, pp. 1–9.

[123] E. Mezura-Montes, C. A. C. Coello, Constrained optimization via multiobjective evolutionary algorithms, in: Multiobjective problem solving from nature, Springer, 2008, pp. 53–75.

[124] S. Koziel, Z. Michalewicz, A decoder-based evolutionary algorithm for constrained parameter optimization problems, in: International Conference on Parallel Problem Solving from Nature, Springer, 1998, pp. 231–240.

[125] Z. Michalewicz, G. Nazhiyath, Genocop iii: A co-evolutionary algorithm for numerical optimization problems with nonlinear constraints, in: Proceedings of 1995 IEEE International Conference on Evolutionary Computation, Vol. 2, IEEE, 1995, pp. 647–651.

[126] R. Mallipeddi, P. N. Suganthan, Ensemble of constraint handling techniques, IEEE Transactions on Evolutionary Computation 14 (4) (2010) 561–579.

[127] Y. Wang, Z. Cai, Y. Zhou, W. Zeng, An adaptive tradeoff model for constrained evolutionary optimization, IEEE Transactions on Evolutionary Computation 12 (1) (2008) 80–92.

[128] A. Goicoechea, D. R. Hansen, L. Duckstein, Multiobjective decision analysis with engineering and business applications, Tech. rep., John Wiley & Sons (1982).

[129] K. Deb, Multi-objective optimisation using evolutionary algorithms: an introduction, in: Multi-objective evolutionary optimisation for product design and manufacturing, Springer, 2011, pp. 3–34.

[130] J. D. Schaffer, Multiple objective optimization with vector evaluated genetic algorithms, in: Proceedings of the first international conference on genetic algorithms and their applications, 1985, Lawrence Erlbaum Associates. Inc., Publishers, 1985.

[131] C. M. Fonseca, P. J. Fleming, et al., Genetic algorithms for multiobjective optimization: Formulation discussion and generalization., in: Icga, Vol. 93, Citeseer, 1993, pp. 416–423.

[132] N. Srinivas, K. Deb, Muiltiobjective optimization using nondominated sorting in genetic algorithms, Evolutionary computation 2 (3) (1994) 221–248.

[133] J. Horn, N. Nafpliotis, D. E. Goldberg, A niched pareto genetic algorithm for multiobjective optimization, in: Proceedings of the first IEEE conference on evolutionary computation. IEEE world congress on computational intelligence, Ieee, 1994, pp. 82–87.

[134] E. Zitzler, M. Laumanns, L. Thiele, Spea2: Improving the strength pareto evolutionary algorithm, TIK-report 103 (2001).

[135] J. D. Knowles, D. W. Corne, Approximating the nondominated front using the pareto archived evolution strategy, Evolutionary computation 8 (2) (2000) 149–172.

[136] Q. Zhang, H. Li, Moea/d: A multiobjective evolutionary algorithm based on decomposition, IEEE Transactions on evolutionary computation 11 (6) (2007) 712–731.

[137] E. Zitzler, S. Künzli, Indicator-based selection in multiobjective search, in: International conference on parallel problem solving from nature, Springer, 2004, pp. 832–842.

[138] J. Bader, E. Zitzler, Hype: An algorithm for fast hypervolume-based many-objective optimization, Evolutionary computation 19 (1) (2011) 45–76.

[139] P. P. Bonissone, R. Subbu, N. Eklund, T. R. Kiehl, Evolutionary algorithms+ domain knowledge= real-world evolutionary computation, IEEE Transactions on Evolutionary Computation 10 (3) (2006) 256–280.

[140] Z. Michalewicz, Genetic algorithms+ data structures= evolution programs, Springer Science & Business Media, 2013.

[141] C. Grosan, A. Abraham, Hybrid evolutionary algorithms: methodologies, architectures, and reviews, in: Hybrid evolutionary algorithms, Springer, 2007, pp. 1–17.

[142] P. Moscato, et al., On evolution, search, optimization, genetic algorithms and martial arts: Towards memetic algorithms, Caltech concurrent computation program, C3P Report 826 (1989) 1989.

[143] R. Dawkins, The selfish gene, Oxford university press, 2016.

[144] R. Meuth, M.-H. Lim, Y.-S. Ong, D. C. Wunsch, A proposition on memes and meta-memes in computing for higher-order learning, Memetic Computing 1 (2) (2009) 85–100.

[145] Y.-S. Ong, M.-H. Lim, N. Zhu, K.-W. Wong, Classification of adaptive memetic algorithms: a comparative study, IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) 36 (1) (2006) 141–152.

[146] N. Krasnogor, J. Smith, A tutorial for competent memetic algorithms: model, taxonomy, and design issues, IEEE Transactions on Evolutionary Computation 9 (5) (2005) 474–488.

[147] A. K. Qin, P. N. Suganthan, Self-adaptive differential evolution algorithm for numerical optimization, in: 2005 IEEE congress on evolutionary computation, Vol. 2, IEEE, 2005, pp. 1785–1791.

[148] H. Ishibuchi, T. Murata, Multi-objective genetic local search algorithm, in: Proceedings of IEEE international conference on evolutionary computation, IEEE, 1996, pp. 119–124.

[149] A. Jaszkiewicz, Genetic local search for multi-objective combinatorial optimization, European journal of operational research 137 (1) (2002) 50–71.

[150] J. Knowles, D. Corne, Memetic algorithms for multiobjective optimization: issues, methods and prospects, in: Recent advances in memetic algorithms, Springer, 2005, pp. 313–352.

[151] Y. Jin, A comprehensive survey of fitness approximation in evolutionary computation, Soft computing 9 (1) (2005) 3–12.

[152] R. E. Smith, B. A. Dike, S. Stegmann, Fitness inheritance in genetic algorithms, in: Proceedings of the 1995 ACM symposium on Applied computing, 1995, pp. 345–350.

[153] A. Forrester, A. Sobester, A. Keane, Engineering design via surrogate modelling: a practical guide, John Wiley & Sons, 2008.

[154] R. H. Myers, D. C. Montgomery, C. M. Anderson-Cook, Response surface methodology: process and product optimization using designed experiments, John Wiley & Sons, 2016.

[155] D. Lim, Y. Jin, Y.-S. Ong, B. Sendhoff, Generalizing surrogate-assisted evolutionary computation, IEEE Transactions on Evolutionary Computation 14 (3) (2009) 329–355.

[156] R. L. Hardy, Multiquadric equations of topography and other irregular surfaces, Journal of geophysical research 76 (8) (1971) 1905–1915.

[157] D. R. Jones, A taxonomy of global optimization methods based on response surfaces, Journal of global optimization 21 (4) (2001) 345–383.

[158] Y. Jin, M. Olhofer, B. Sendhoff, A framework for evolutionary optimization with approximate fitness functions, IEEE Transactions on evolutionary computation 6 (5) (2002) 481–494.

[159] L. Shi, K. Rasheed, A survey of fitness approximation methods applied in evolutionary algorithms, in: Computational intelligence in expensive optimization problems, Springer, 2010, pp. 3–28.

[160] Y. Jin, M. Olhofer, B. Sendhoff, On evolutionary optimization with approximate fitness functions., in: GECCO, 2000, pp. 786–793.

[161] M. Sefrioui, J. Périaux, A hierarchical genetic algorithm using multiple models for optimization, in: International Conference on Parallel Problem Solving From Nature, Springer, 2000, pp. 879–888.

[162] K. C. Giannakoglou, A. P. Giotis, M. K. Karakasis, Low-cost genetic optimization based on inexact pre-evaluations and the sensitivity analysis of design parameters, Inverse Problems in Engineering 9 (4) (2001) 389–412.

[163] M. Emmerich, A. Giotis, M. Özdemir, T. Bäck, K. Giannakoglou, Metamodel—assisted evolution strategies, in: International Conference on parallel problem solving from nature, Springer, 2002, pp. 361–370.

[164] K. Rasheed, H. Hirsh, Informed operators: Speeding up genetic-algorithm-based design optimization using reduced models, in: Proceedings of the 2nd Annual Conference on Genetic and Evolutionary Computation, 2000, pp. 628–635.

[165] D. Lim, Y.-S. Ong, Y. Jin, B. Sendhoff, A study on metamodeling techniques, ensembles, and multi-surrogates in evolutionary computation, in: Proceedings of the 9th annual conference on Genetic and evolutionary computation, 2007, pp. 1288–1295.

[166] A. Díaz-Manríquez, G. Toscano-Pulido, W. Gómez-Flores, On the selection of surrogate models in evolutionary optimization algorithms, in: Evolutionary Computation (CEC), 2011 IEEE Congress on, IEEE, 2011, pp. 2155–2162.

[167] A. Díaz-Manríquez, G. Toscano, C. A. C. Coello, Comparison of metamodeling techniques in evolutionary algorithms, Soft Computing 21 (19) (2017) 5647–5663.

[168] Y. Wang, D.-Q. Yin, S. Yang, G. Sun, Global and local surrogate-assisted differential evolution for expensive constrained optimization problems with inequality constraints, IEEE transactions on cybernetics 49 (5) (2018) 1642–1656.

[169] Y.-S. Ong, Z. Zhou, D. Lim, Curse and blessing of uncertainty in evolutionary algorithm using approximation, in: 2006 IEEE International Conference on Evolutionary Computation, IEEE, 2006, pp. 2928–2935.

[170] R. Mallipeddi, M. Lee, An evolving surrogate model-based differential evolution algorithm, Applied Soft Computing 34 (2015) 770–787.

[171] T. P. Runarsson, Constrained evolutionary optimization by approximate ranking and surrogate models, in: International Conference on Parallel Problem Solving from Nature, Springer, 2004, pp. 401–410.

[172] R. G. Regis, Evolutionary programming for high-dimensional constrained expensive black-box optimization using radial basis functions, IEEE Transactions on Evolutionary Computation 18 (3) (2013) 326–347.

[173] Z. Zhou, Y. S. Ong, P. B. Nair, A. J. Keane, K. Y. Lum, Combining global and local surrogate models to accelerate evolutionary optimization, IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) 37 (1) (2006) 66–76.

[174] Y. S. Ong, P. B. Nair, A. J. Keane, Evolutionary optimization of computationally expensive problems via surrogate modeling, AIAA journal 41 (4) (2003) 687–696.

[175] H.-M. Gutmann, A radial basis function method for global optimization, Journal of global optimization 19 (3) (2001) 201–227.

[176] D. R. Jones, M. Schonlau, W. J. Welch, Efficient global optimization of expensive black-box functions, Journal of Global optimization 13 (4) (1998) 455–492.

[177] R. G. Regis, C. A. Shoemaker, A stochastic radial basis function method for the global optimization of expensive functions, INFORMS Journal on Computing 19 (4) (2007) 497–509.

[178] R. G. Regis, C. A. Shoemaker, Constrained global optimization of expensive black box functions using radial basis functions, Journal of Global optimization 31 (1) (2005) 153–171.

[179] R. G. Regis, Constrained optimization by radial basis function interpolation for high-dimensional expensive black-box problems with infeasible initial points, Engineering Optimization 46 (2) (2014) 218–243.

[180] J. Müller, C. A. Shoemaker, R. Piché, So-i: a surrogate model algorithm for expensive nonlinear integer programming problems including global optimization applications, Journal of Global Optimization 59 (4) (2014) 865–889.

[181] J. Müller, C. A. Shoemaker, R. Piché, So-mi: A surrogate model algorithm for computationally expensive nonlinear mixed-integer black-box global optimization problems, Computers & Operations Research 40 (5) (2013) 1383–1400.

[182] J. Knowles, Parego: a hybrid algorithm with on-line landscape approximation for expensive multiobjective optimization problems, IEEE Transactions on Evolutionary Computation 10 (1) (2006) 50–66.

[183] Q. Zhang, W. Liu, E. Tsang, B. Virginas, Expensive multiobjective optimization by moea/d with gaussian process model, IEEE Transactions on Evolutionary Computation 14 (3) (2009) 456–474.

[184] S. Zapotecas Martínez, C. A. Coello Coello, Moea/d assisted by rbf networks for expensive multi-objective optimization problems, in: Proceedings of the 15th annual conference on Genetic and evolutionary computation, 2013, pp. 1405–1412.

[185] M. T. Emmerich, K. C. Giannakoglou, B. Naujoks, Single-and multiobjective evolutionary optimization assisted by gaussian random field metamodels, IEEE Transactions on Evolutionary Computation 10 (4) (2006) 421–439.

[186] R. G. Regis, Multi-objective constrained black-box optimization using radial basis function surrogates, Journal of computational science 16 (2016) 140–155.

[187] C. A. Georgopoulou, K. C. Giannakoglou, Multiobjective metamodel–assisted memetic algorithms, in: Multi-objective memetic algorithms, Springer, 2009, pp. 153–181.

[188] T. Akhtar, C. A. Shoemaker, Multi objective optimization of computationally expensive multi-modal functions with rbf surrogates and multi-rule selection, Journal of Global Optimization 64 (1) (2016) 17–32.

[189] T. Akhtar, C. A. Shoemaker, Efficient multi-objective optimization through population-based parallel surrogate search, arXiv preprint arXiv:1903.02167 (2019).

[190] R. Loulou, U. Remme, A. Kanudia, A. Lehtila, G. Goldstein, Documentation for the TIMES Model Part II, Energy technology systems analysis programme (ETSAP) (2005).

[191] J. Meus, K. Poncelet, E. Delarue, Applicability of a clustered unit commitment model in power system modeling, IEEE Transactions on Power Systems 33 (2) (2018) 2195–2204.

[192] G. Onwubolu, D. Davendra, Scheduling flow shops using differential evolution algorithm, European Journal of Operational Research 171 (2) (2006) 674–692.

[193] J. Lampinen, I. Zelinka, Mixed integer-discrete-continuous optimization by differential evolution, in: Proceedings of the 5th International Conference on Soft Computing, 1999, pp. 71–76.

[194] R. Tanabe, A. Fukunaga, Reevaluating exponential crossover in differential evolution, in: International Conference on Parallel Problem Solving from Nature, Springer, 2014, pp. 201–210.

[195] Q. Y. Kenny, W. Li, A. Sudjianto, Algorithmic construction of optimal symmetric Latin hypercube designs, Journal of statistical planning and inference 90 (1) (2000) 145–159.

[196] J. Zhang, S. Chowdhury, A. Messac, An adaptive hybrid surrogate model, Structural and Multidisciplinary Optimization 46 (2) (2012) 223–238.

[197] S. S. Garud, I. A. Karimi, M. Kraft, Smart sampling algorithm for surrogate model development, Computers & Chemical Engineering 96 (2017) 103–114.

[198] M. Powell, The theory of radial basis function approximation in 1990, Advances in Numerical Analysis II: Wavelets, Subdivision, and Radial Functions (WA Light, ed.), Oxford University Press, Oxford 105 (1992) 210.

[199] ADMIE, Independent Power Transmission Operator (2020).
URL `http://www.admie.gr/`

[200] S. Simoes, W. Nijs, P. Ruiz, A. Sgobbi, D. Radu, P. Bolat, C. Thiel, S. Peteves, et al., The JRC-EU-TIMES model SET plan energy technologies, Tech. Rep. (2013).

[201] Y. Wang, B.-C. Wang, H.-X. Li, G. G. Yen, Incorporating objective function information into the feasibility rule for constrained evolutionary optimization, IEEE Transactions on Cybernetics 46 (12) (2015) 2938–2952.

[202] M. Pavičević, K. Kavvadias, T. Pukšec, S. Quoilin, Comparison of different model formulations for modelling future power systems with high shares of renewables–the dispa-set balkans model, Applied energy 252 (2019) 113425.

[203] G. Morales-España, L. Ramírez-Elizondo, B. F. Hobbs, Hidden power system inflexibilities imposed by traditional unit commitment formulations, Applied Energy 191 (2017) 223–238.

[204] C. A. Georgopoulou, K. C. Giannakoglou, A multi-objective metamodel-assisted memetic algorithm with strength-based local refinement, Engineering optimization 41 (10) (2009) 909–923.

[205] M. K. Karakasis, K. C. Giannakoglou, On the use of metamodel-assisted, multi-objective evolutionary algorithms, Engineering Optimization 38 (8) (2006) 941–957.

[206] S. Z. Martínez, C. A. C. Coello, A memetic algorithm with non gradient-based local search assisted by a meta-model, in: International Conference on Parallel Problem Solving from Nature, Springer, 2010, pp. 576–585.

[207] K. Deb, K. Sindhya, T. Okabe, Self-adaptive simulated binary crossover for real-parameter optimization, in: Proceedings of the 9th annual conference on Genetic and evolutionary computation, 2007, pp. 1187–1194.

[208] K. Deb, R. B. Agrawal, et al., Simulated binary crossover for continuous search space, Complex systems 9 (2) (1995) 115–148.

[209] H. Li, Q. Zhang, Multiobjective optimization problems with complicated pareto sets, moea/d and nsga-ii, IEEE transactions on evolutionary computation 13 (2) (2008) 284–302.

[210] ENTSO-E, European network of transmission system operators for electricity (2013).

[211] Y. Tian, R. Cheng, X. Zhang, Y. Jin, Platemo: A matlab platform for evolutionary multi-objective optimization [educational forum], IEEE Computational Intelligence Magazine 12 (4) (2017) 73–87.

[212] E. Zitzler, K. Deb, L. Thiele, Comparison of multiobjective evolutionary algorithms: Empirical results, Evolutionary computation 8 (2) (2000) 173–195.

[213] E. Zitzler, L. Thiele, Multiobjective optimization using evolutionary algorithms—a comparative case study, in: International conference on parallel problem solving from nature, Springer, 1998, pp. 292–301.

[214] M. Emmerich, N. Beume, B. Naujoks, An emo algorithm using the hypervolume measure as selection criterion, in: International Conference on Evolutionary Multi-Criterion Optimization, Springer, 2005, pp. 62–76.

Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Μηχανολόγων Μηχανικών
Τομέας Βιομηχανικής Διοίκησης & Επιχειρησιακής Έρευνας
Εργαστήριο Οργάνωσης Παραγωγής

Σχεδιασμός επέκτασης της δυναμικότητας παραγωγής ισχύος με υψηλό μερίδιο ανανεώσιμων πηγών ενέργειας: μονο- και πολυ- κριτηριακή βελτιστοποίηση βασιζόμενη σε εξελικτικούς αλγορίθμους υποβοηθούμενους από μεταπρότυπα

Διδακτορική Διατριβή
Κωνσταντίνος Βρυώνης

Επιβλέπων: Αθανάσιος Τόλης
Αναπληρωτής Καθηγητής ΕΜΠ

Αθήνα, 2020

Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Μηχανολόγων Μηχανικών
Τομέας Βιομηχανικής Διοίκησης & Επιχειρησιακής Έρευνας
Εργαστήριο Οργάνωσης Παραγωγής

Σχεδιασμός επέκτασης της δυναμικότητας παραγωγής ισχύος με υψηλό μερίδιο ανανεώσιμων πηγών ενέργειας: μονο- και πολυ- κριτηριακή βελτιστοποίηση βασιζόμενη σε εξελικτικούς αλγορίθμους υποβοηθούμενους από μεταπρότυπα

Διδακτορική Διατριβή
Κωνσταντίνος Βρυώνης

Επταμελής Εξεταστική Επιτροπή:

1. Αθανάσιος Τόλης (Επιβλέπων), Αναπληρωτής Καθηγητής ΕΜΠ,
   Σχολή Μηχανολόγων Μηχανικών, (Μέλος Τριμελούς Συμβουλευτικής επιτροπής)
2. Νικόλαος Παναγιώτου, Αναπληρωτής Καθηγητής ΕΜΠ,
   Σχολή Μηχανολόγων Μηχανικών, (Μέλος Τριμελούς Συμβουλευτικής επιτροπής)
3. Ηλίας Τατσιόπουλος, Καθηγητής ΕΜΠ,
   Σχολή Μηχανολόγων Μηχανικών, (Μέλος Τριμελούς Συμβουλευτικής επιτροπής)
4. Κυριάκος Γιαννάκογλου, Καθηγητής ΕΜΠ
   Σχολή Μηχανολόγων Μηχανικών
5. Γρηγόρης Μέντζας, Καθηγητής ΕΜΠ,
   Σχολή Ηλεκτρολόγων Μηχανικών & Μηχανικών Υπολογιστών
6. Ιωάννης Αναγνωστόπουλος, Καθηγητής ΕΜΠ,
   Σχολή Μηχανολόγων Μηχανικών
7. Αθανάσιος Ρεντιζέλας, Επίκουρος Καθηγητής ΕΜΠ,
   Σχολή Μηχανολόγων Μηχανικών

Αθήνα, 2020

Περίληψη

Σκοπός αυτής της Διδακτορικής Διατριβής είναι να αναπτύξει και να εξετάσει τη εφαρμογή Εξελικτικών Αλγορίθμων υποβοηθούμενων από μεταπρότυπα (Metamodel-assisted Evolutionary Algorithms, MAEA) για μονο- και πολυ- κριτηριακή βελτιστοποίηση του προβλήματος επέκτασης δυναμικότητας παραγωγής ισχύος (Generation Expansion Planning, GEP) με υψηλό μερίδιο παραγωγής από Ανανεώσιμες Πηγές Ενέργειας (Renewable Energy Sources, RES). Ένα μοντέλο GEP μπορεί να συμβάλει στη διαδικασία λήψης αποφάσεων για τον μεσοπρόθεσμο και μακροπρόθεσμο προγραμματισμό παραγωγής ενέργειας. Συνήθως, αποσκοπεί στην παραγωγή σεναρίων για την οικονομικά αποδοτική, βιώσιμη και ασφαλή λειτουργία ενός συστήματος ισχύος που καθορίζονται λαμβάνοντας υπόψη τόσο τις πιθανές επενδύσεις σε προσθήκες δυναμικότητας όσο και τη βραχυπρόθεσμη λειτουργία του συστήματος ισχύος, π.χ. ο προσδιορισμός ενός οικονομικού και/ή περιβαλλοντικά βέλτιστου επενδυτικού πλάνου για την κάλυψη της αυξανόμενης ζήτησης ενέργειας και παρουσιάζει αξιόπιστη βραχυπρόθεσμη λειτουργία. Ένα τέτοιο πρόβλημα μπορεί να διατυπωθεί ως πρόβλημα βελτιστοποίησης.

Η επιδίωξη ενός συστήματος ισχύος χωρίς εκπομπές ρύπων και τα αυξανόμενα μερίδια παραγωγής των RES οδήγησαν στην εισαγωγή διαφόρων νέων παραμέτρων στα μοντέλα GEP. Μια τέτοια παράμετρος είναι ο εντοπισμός οικονομικών και τεχνικών προκλήσεων που σχετίζονται με τη βραχυπρόθεσμη λειτουργία του συστήματος ισχύος που θα μπορούσε να είναι απαραίτητη για την αξιολόγηση της συνέργειας της εγκατεστημένης ισχύος και των αυξανόμενων εγκαταστάσεων RES. Συγκεκριμένα, η παραγωγή από RES χαρακτηρίζεται από μεταβλητότητα και αβεβαιότητα. Έχει αναφερθεί στη βιβλιογραφία ότι αυτά αναμένεται να αυξήσουν τις απαιτήσεις σε λειτουργική ευελιξία και ότι η υποτίμηση τέτοιων απαιτήσεων θα μπορούσε να έχει οικονομικές επιπτώσεις στη βραχυπρόθεσμη λειτουργία του συστήματος ισχύος. Συνεπώς, έχουν γίνει προσπάθειες για την αύξηση του επιπέδου λεπτομέρειας της μοντελοποίησης της βραχυπρόθεσμης λειτουργίας ενός συστήματος ισχύος σε όρους χωρικής, χρονικής και τεχνικής λεπτομέρειας που εισάγεται σε ένα μοντέλο GEP. Ωστόσο, αυτό μπορεί να οδηγήσει σε αυξημένο υπολογιστικό κόστος, και παράλληλα, είναι απαραίτητο να εισάγονται απλοποιήσεις.

Οι Εξελικτικοί Αλγόριθμοι (EA) είναι αλγόριθμοι εμπνευσμένοι από τη φύση που εφαρμόζουν στοχαστικούς τελεστές για να βελτιώσουν ένα σύνολο υποψηφίων λύσεων. Αυτοί δεν βασίζονται στη χρήση παραγώγων και μπορούν να χρησιμοποιηθούν ως μέθοδοι άμεσης αναζήτησης για σύνθετα προβλήματα βελτιστοποίησης. Επιπλέον, EA για πολυκριτηριακή βελτιστοποίηση (Multi-Objective EA, MOEA) είναι αποδοτικές μέθοδοι για πολυκριτηριακά προβλήματα βελτιστοποίησης (Multi-Objective Optimization, MOO). Ένα βασικό τους μειονέκτημα είναι ο σχετικά μεγάλος αριθμός αξιολογήσεων που απαιτείται για τη σύγκλιση του αλγορίθμου που μπορεί να είναι απαγορευτικό για προβλήματα που περιλαμβάνουν υπολογιστικά δαπανηρές προσομοιώσεις. Για τέτοιες εφαρμογές, οι EA χρησιμοποιούνται σε συνδυασμό με μοντέλα προσέγγισης (Approximating Models, AM) τα οποία στη βιβλιογραφία αναφέρονται ως MAEA. Τα AM αντικαθιστούν εν μέρει τα αρχικά μοντέλα για να παρέχουν μια εκτίμηση της επάρκειας μιας υποψήφιας λύσης για μείωση του υπολογιστικού κόστους.

Αυτή η Διδακτορική Διατριβή επικεντρώνεται σε εφαρμογές MAEA σε GEP ενός ή πολλών στόχων που περιλαμβάνουν μοντέλο προσομοίωσης (Simulation Models, SM) για τη βραχυπρόθεσμη λειτουργία ενός συστήματος ισχύος. Οι σημαντικότερες συνεισφορές της είναι οι ακόλουθες:

1. Παρουσιάζεται μια προσέγγιση, βασισμένη σε MAEA, για GEP πολλαπλών περιόδων και ενός στόχου που περιλαμβάνει SM. Το επιλεγμένο από τη βιβλιογραφία SM είναι ένα μοντέλο βελτιστοποίησης για τη βραχυπρόθεσμη λειτουργία ενός συστήματος ισχύος που περιλαμβάνει παραδοχές,

π.χ. η χωρική λεπτομέρεια δεν εξετάζεται. Ωστόσο, παρουσιάζει σχετικά αυξημένο επίπεδο τεχνικής και χρονικής λεπτομέρειας για το πλαίσιο του μακροπρόθεσμου προγραμματισμού. Το SM εισάγεται για την αξιολόγηση της λειτουργικής ευελιξίας κατά την διαδικασίας βελτιστοποίησης του προβλήματος. Συγκεκριμένα χαρακτηριστικά του προβλήματος αξιοποιούνται για την επίλυσή του. Αυτό υλοποιείται μέσα από την χρήση των AM έτσι ώστε να παρέχουν μια εκτίμηση του SM και να μειώσουν το υπολογιστικό κόστος. Τα εξεταζόμενα AM είναι οι Συναρτήσεις Ακτινικής Βάσης (Radial Basis Function, RBF). Γίνεται χρήση τοπικών και ολικών AM που ανανεώνονται κατά τη διάρκεια της βελτιστοποίησης. Επίσης, εξετάζονται εξειδικευμένοι τελεστές που στοχεύουν στη βελτίωση της απόδοση του EA που εξετάστηκε ο οποίος είναι ο αλγόριθμος της Διαφορικής Εξέλιξης (Differential Evolution, DE). Με βάση τα αποτελέσματα των υπολογιστικών πειραμάτων, η απόδοση του MAEA κρίνεται ικανοποιητική λαμβάνοντας υπόψη τους περιορισμούς λόγω υπολογιστικού κόστος. Επιπλέον, η ευρετική τεχνική διόρθωσης των περιορισμών του προβλήματος βελτιστοποίησης παρείχε τη μεγαλύτερη βελτίωση στην απόδοση του αλγόριθμου βάσης μεταξύ των εξειδικευμένοι τελεστών που αναπτύχθηκαν. Επίσης, εξετάζεται η επίδραση της ενσωμάτωσης του SM. Τα αποτελέσματα υποδηλώνουν τη σημασία της ενσωμάτωσης τεχνικών που να εξετάζουν την λειτουργική ευελιξία για την επαρκή αξιολόγηση παρόχων ευελιξίας όταν αυτοί θεωρούνται ως επενδυτικές επιλογές. Η αξιολόγηση της ακρίβειας των AM έδειξε ότι αυτή ήταν αποδεκτή με βάση τους δείκτες που χρησιμοποιήθηκαν. Επομένως, πραγματοποιήθηκε οπτική ανάλυση της ευαισθησίας του λειτουργικού κόστους σε σχέση με την εγκατεστημένη δυναμικότητα παραγωγής ισχύος.

2. Παρουσιάζεται μια προσέγγιση, βασισμένη σε MAEA, για ένα στατικό μοντέλο GEP που στοχεύει στην ανάλυση αντικρουόμενων στόχων σε MOO GEP. Η λειτουργική ευελιξία αξιολογείται από ένα, επιλεγμένο από τη βιβλιογραφία, SM που περιλαμβάνει τεχνική, χωρική και χρονική λεπτομέρεια της βραχυπρόθεσμης λειτουργίας ενός συστήματος ισχύος. Η μεθοδολογία αναπτύσσεται με βάση τη σχετική βιβλιογραφία για τους MOEA και τη βελτιστοποίηση με μεταπρότυπα χωρίς χρήση παραγώγων. Ως AM επιλέχθηκαν τα RBF και η Πολυωνυμική Παλινδρόμηση (Polynomial Regression, PR). Αυτά ανανεώνονται με κριτήρια που δίνουν προτεραιότητα στις εφικτές λύσεις, τη χωρική κατανομή του αποθηκευμένου συνόλου λύσεων στον χώρο αναζήτησης και πιθανές βελτιώσεις του υπερόγκου. Η διαδικασία περιλαμβάνει τοπική αναζήτηση με χρήση παραγώγων, τοπικά RBF, PR και ένα μοντέλο που αποτελεί συνδυασμό αυτών. Η απόδοση του MAEA εξετάζεται σε συναρτήσεις αναφοράς για MOO προβλήματα, σε MOO GEP χωρίς SM και σε πέντε MOO GEP που συμπεριλαμβάνουν το SM. Τα τελευταία επαναλαμβάνονται για δύο διαφορετικά επίπεδα χρονικής λεπτομέρειας. Τα αποτελέσματα υποδηλώνουν αποδεκτή απόδοση του αλγορίθμου λαμβάνοντας υπόψη τους περιορισμούς λόγω υπολογιστικού κόστος. Επιπλέον, η ακρίβεια των AM διέφερε μεταξύ των προβλημάτων που εξετάστηκαν και αναφέρονται παράγοντες που την επηρέασαν. Ανάλυση του συνόλου εναλλακτικών λύσεων σε επίπεδο συντελεστών κόστους μπορεί να προσφέρει μια λεπτομερή αξιολόγηση αυτών συμβάλλοντας στον προσδιορισμό απαιτούμενων κινήτρων για τη λήψη αποφάσεων στρατηγικής ενεργειακής πολιτικής. Για παράδειγμα, με βάση τις ακραίες τιμές του μετώπου μη-κυριαρχούμενων λύσεων για τις συναρτήσεις λειτουργικού κόστους και κόστους επένδυσης που εξετάστηκαν, η μείωση κατά 96% του κόστους επένδυσης θα μπορούσε οδηγήσει σε μια αύξηση σχεδόν 40% του λειτουργικού κόστους.

Εργαλεία υποστήριξης αποφάσεων θα μπορούσαν να συνεισφέρουν στη σύνθετη και εξελισσόμενη διαδικασία λήψης αποφάσεων ενός GEP. Τα οικονομικά, περιβαλλοντικά και κοινωνικά κριτήρια πρέπει να λαμβάνονται υπόψη μαζί με άλλες πτυχές που προοδευτικά αναγνωρίζονται ως απαραίτητες. Παρά την ευρετική τους φύση, τα αποτελέσματα έδειξαν ότι οι προσεγγίσεις που αναπτύχθηκαν θα μπορούσαν να είναι υποσχόμενα εργαλεία για την υποστήριξη καθιερωμένων μοντέλων GEP ώστε να συμβάλλουν στη διαδικασία λήψης αποφάσεων όταν λαμβάνονται υπόψη υψηλά μερίδια παραγωγής από RES.

Περιεχόμενα

Κεφάλαιο 1

Εισαγωγή

1.1 Αντικείμενο

Η μετάβαση προς συστήματα ισχύος με εκπομπές χαμηλών εκπομπών διοξειδίου του άνθρακα μπορεί να είναι ένας σημαντικός παράγοντας για τη βιωσιμότητα τους. Κατά συνέπεια, το μερίδιο της ενέργειας που παράγεται από RES έχει αυξηθεί τα τελευταία χρόνια για την επίτευξη περιβαλλοντικών στόχων [1, 2, 3, 4, 5]. Είναι επιθυμητό να επιτευχθεί μια τέτοια μετάβαση με οικονομικά αποδοτικό και περιβαλλοντικά φιλικό τρόπο λόγω του υψηλού κόστους. Αυτά μπορεί να προέρχονται από επενδύσεις σε νέους σταθμούς ηλεκτροπαραγωγής που αναμένεται να λειτουργούν για αρκετά χρόνια και, συνεπώς, περιλαμβάνουν μακροπρόθεσμο σχεδιασμό. Ο μακροπρόθεσμος σχεδιασμός είναι σύνθετος, δεδομένου ότι μπορεί να εξετάζει πολλούς στόχους οι οποίοι μπορούν συχνά να είναι αντικρουόμενοι. Η ενημερωμένη λήψη αποφάσεων μπορεί να συμβάλει προς αυτήν την κατεύθυνση [6].

Σε αυτό το πλαίσιο, τα επιχειρησιακά/λειτουργικά χαρακτηριστικά του μελλοντικού εγκατεστημένου δυναμικού παραγωγής ισχύος για τη βέλτιστη ενσωμάτωση υψηλών επιπέδων RES είναι ενεργό ερευνητικό πεδίο με αυξανόμενο ενδιαφέρον [1, 2, 3, 4, 5]. Αυτό γιατί όσο αυξάνονται τα επίπεδα διείσδυσης των RES, ο αντίκτυπός τους στη βραχυπρόθεσμη λειτουργία και, κατά συνέπεια, στο βέλτιστο μακροπρόθεσμο προγραμματισμό έχει αναφερθεί να γίνεται πιο σημαντικός λόγω της μεταβαλλόμενης παραγωγής των RES και της αλληλεπίδρασή της με το συμβατικό εγκατεστημένο δυναμικό παραγωγής ισχύος [2, 3, 4]. Συγκεκριμένα, η ενσωμάτωση της παραγωγής από RES μπορεί να οδηγήσει σε μείωση του συνολικού κόστους στη βραχυπρόθεσμη λειτουργία. Ωστόσο, στο σενάριο υψηλών επιπέδων διείσδυσης μπορεί να απαιτηθεί ταυτόχρονα επαρκή λειτουργική ευελιξία για την αποδοτική ικανοποίηση της ζήτησης [7, 8]. Ως εκ τούτου, η επάρκεια της εγκατεστημένου δυναμικού εξετάζεται συχνά και ως προς τη λειτουργική του ευελιξία [2, 4].

Ο επαρκής εντοπισμός του αντίκτυπου της αύξησης των μεριδίων παραγωγής RES στη βραχυπρόθεσμη λειτουργία ενός συστήματος ισχύος στο πλαίσιο του μακροπρόθεσμου προγραμματισμού εμπεριέχει προκλήσεις [7, 8, 2, 4, 9]. Συγκεκριμένα, σε ένα μακροπρόθεσμο μοντέλο σχεδιασμού χρειάζεται η εισαγωγή επιπέδου τεχνικής, χρονικής και χωρικής λεπτομέρειας έτσι ώστε να λαμβάνει υπόψη με σχετική ακρίβεια τη βραχυπρόθεσμη λειτουργία ενός συστήματος ισχύος [9]. Αυτό απαιτείται για την αξιολόγηση της λειτουργικής ευελιξίας ενός υποψήφιου μελλοντικού δυναμικού παραγωγής ισχύος. Η σημασία του επιπέδου ανάλυσης μπορεί να ποικίλλει ανάλογα με το εξεταζόμενο σύστημα [10], όμως η παράληψη αυτής της αξιολόγησης θα μπορούσε να έχει αντίκτυπο στον βέλτιστο σχεδιασμό.

Από την άλλη πλευρά, η ενσωμάτωση λεπτομέρειας της βραχυπρόθεσμη λειτουργίας σε μοντέλα για τον μακροπρόθεσμα σχεδιασμό οδηγεί σε αυξημένη πολυπλοκότητα και υπολογιστικό κόστος. Αυτό προκύπτει από την απαίτηση για τον συνδυασμό δύο σύνθετων μοντέλων [7, 8], δηλαδή των μοντέλων για μακροπρόθεσμο και βραχυπρόθεσμο σχεδιασμό. Επομένως, το υπολογιστικό κόστος πρέπει επίσης να αντιμετωπιστεί μαζί με την ταυτόχρονη εισαγωγή παραδοχών και απλουστεύσεων [9, 5].

Οι ΕΑ έχουν γίνει μια από τις καθιερωμένες τεχνικές βελτιστοποίησης και έχουν βρει εφαρμογές σε ποικιλία προβλημάτων όπως και σε προβλήματα βελτιστοποίησης που σχετίζονται με συστήματα ισχύος [1, 11]. Οι ΕΑ είναι αλγόριθμοι εμπνευσμένοι από τη φύση και βασίζονται σε στοχαστικούς τελεστές για να βελτιώσουν μια σειρά υποψηφίων λύσεων σε βάθος επαναλήψεων. Κρίνονται κατάλληλοι επιλύτες για προβλήματα βελτιστοποίησης που περιλαμβάνουν περίπλοκες (π.χ. μη κυρτές ή/και μη παραγωγίσιμες) αντικειμενικές συναρτήσεις και συναρτήσεις περιορισμών. Επιπλέον, οι ΕΑ έχουν αναγνωριστεί ευρέως για την απόδοσή τους σε προβλήματα MOO [12, 13]. Παρ΄ όλα αυτά, οι Ε-Α παρουσιάζουν και μειονεκτήματα. Για παράδειγμα, η απώλεια της πληροφορίας για την παράγωγο, όταν αυτή είναι διαθέσιμη, μπορεί να καταστήσουν τους ΕΑ λιγότερο πρακτικούς για προβλήματα βελτιστοποίησης που μπορούν να αντιμετωπιστούν εύκολα με κλασικές μεθοδολογίες βελτιστοποίησης. Επιπλέον, η εφαρμογή ΕΑ μπορεί να καταστεί (πρακτικά) ανέφικτη για προβλήματα βελτιστοποίησης που περιλαμβάνουν υπολογιστικά ακριβές συναρτήσεις στόχου και/ή περιορισμών λόγω του χαμηλού αριθμού διαθέσιμων αξιολογήσεων [14, 15]. Τέτοια προβλήματα μπορεί να ενσωματώνουν ένα δαπανηρό μοντέλο προσομοίωσης ή ένα ακριβό πείραμα. Μια κοινή προσέγγιση είναι η υποβοήθηση ΕΑ με ΑΜ. Τέτοιες προσεγγίσεις περιλαμβάνουν ΑΜ που χρησιμοποιούνται για την παροχή ενός υπολογιστικά φτηνού δείκτη της επάρκειας μιας λύσης και αντικαταστούν εν μέρει το υπολογιστικά ακριβό μοντέλο [16].

1.2 Βασικά κίνητρα και στόχοι

Βασικό κίνητρο αυτής της Διατριβής ήταν οι αυξανόμενες υπολογιστικές απαιτήσεις του μακροπρόθεσμου προβλήματος επέκτασης δυναμικότητας παραγωγής ισχύος. Αυτές μπορούν να αποδοθούν στην σημασία της αξιολόγησης της λειτουργικής ευελιξίας ενός υποψηφίου εγκατεστημένου δυναμικού παραγωγής ισχύος αλλά και στην ενσωμάτωση νέων παραμέτρων στο GEP [1, 2, 3, 4, 5]. Εφαρμογές ΕΑ στο GEP έχουν προταθεί στη σχετική βιβλιογραφία. Ωστόσο, εφαρμογές βασισμένες σε ΕΑ ή σε ΜΑΕΑ για τον μακροπρόθεσμο σχεδιασμό που να εμπεριέχει σχετικά αυξημένη λεπτομέρεια για τη βραχυπρόθεσμη λειτουργία ενός συστήματος ισχύος μέσω μοντέλου προσομοίωσης δεν είχαν εντοπιστεί. Αυτό θα μπορούσε να αποδοθεί στο υπολογιστικό κόστος που σχετίζεται με μια τέτοια προσέγγιση. Ως εκ τούτου, το κύριο πεδίο αυτής της Διατριβής είναι οι προσεγγίσεις που βασίζονται σε ΜΑΕΑ για μονο- και πολυ- κριτηριακή βελτιστοποίηση του προβλήματος GEP με υψηλό μερίδιο παραγωγής από RES. Οι προσεγγίσεις που αναπτύσσονται εστιάζουν στη εισαγωγή ενός τέτοιου SM στη διαδικασία βελτιστοποίησης και όχι στην ανάπτυξη ενός βέλτιστου SM για τη βραχυπρόθεσμη λειτουργία ενός συστήματος ισχύος. Συνεπώς, SM που εντοπίστηκαν στη βιβλιογραφία χρησιμοποιήθηκαν με μικρές τροποποιήσεις και εισάγονται για την αξιολόγηση υποψήφιων λύσεων λαμβάνοντας υπόψη δυναμικές που προκύπτουν από τη βραχυπρόθεσμη λειτουργία του συστήματος ισχύος. Ταυτόχρονα, εξετάζεται η λειτουργική ευελιξία που μπορούν να παρέχουν τόσο θερμικές όσο και μη θερμικές μονάδες. Αυτές οι απαιτήσεις οδηγούν σε αύξηση του υπολογιστικού κόστους και, κατά συνέπεια, εξετάζονται οι ΜΑΕΑ.

Οι ΜΑΕΑ που έχουν αναπτυχθεί για προβλήματα βελτιστοποίησης ενός και πολλαπλών στόχων για το GEP βασίζονται σε πλαίσια για βελτιστοποίηση υποβοηθούμενη από μεταπρότυπα χωρίς χρήση παραγώγων που εντοπίστηκαν στη βιβλιογραφία. Ταυτόχρονα, εξετάζονται εξειδικευμένοι τελεστές που εστιάζουν στη βελτίωση της απόδοσης του ΕΑ.

Η διαδικασία λήψης αποφάσεων του GEP είναι πολύπλοκη καθώς περιλαμβάνει πολλούς παράγοντες που πρέπει να ληφθούν υπόψη [1, 2]. Μερικοί από αυτούς τους παράγοντες μπορεί να είναι περιβαλλοντικοί, όπως μείωση των εκπομπών αερίων του θερμοκηπίου (Green House Gas, GHG), οικονομικοί, όπως η κερδοφορία και βιωσιμότητα των παραγωγών ηλεκτρικής ενέργειας, ή/και κοινωνικοί, όπως οικονομικά προσιτή ενέργεια. Τα εργαλεία υποστήριξης αποφάσεων που αναπτύχθηκαν στοχεύουν να συμβάλουν στην διαδικασία λήψη αποφάσεων. Ένα τέτοιο παράδειγμα είναι η αξιολόγηση των κινήτρων για την αύξηση των επιπέδων διείσδυσης των RES εξετάζοντας πιθανούς μηχανισμούς στήριξης RES των οποίων τα οφέλη και κόστη μπορούν να μεταφερθούν στους καταναλωτές.

Η βελτιωμένη αξιολόγηση ενός μακροπρόθεσμου σχεδίου μπορεί να προσφέρει οφέλη σε επενδυτές, ρυθμιστικές αρχές και καταναλωτές. Αυτά μπορεί να προκύψουν από βελτιωμένα σήματα που θα μπορούσαν να προκύψουν από τη ανάλυση των αποτελεσμάτων για τη διαδικασία λήψης αποφάσεων, π.χ. στην κατανομή ενός διαθέσιμου επενδυτικού κεφαλαίου μεταξύ υποψήφιων επενδύσεων ή στη διαμόρφωση κίνητρων από ρυθμιστικές αρχές για προώθηση επενδύσεων στοχεύοντας προς την αύξηση της αποδοτικότητας του συστήματος ισχύος. Η αποδοτική και αξιόπιστη λειτουργία του συστήματος μπορεί να ωφελήσει τους τελικούς καταναλωτές. Οι σημαντικότερες συνεισφορές της Διδακτορικής Διατριβής συνοψίζονται πιο κάτω:

1. Παρουσιάζεται ένα μοντέλο GEP πολλαπλών περιόδων που βασίζεται σε μονοκριτηριακή βελτιστοποίηση με χρήση MAEA [17]. Το μοντέλο περιλαμβάνει σχετικά αυξημένη τεχνική και χρονική λεπτομέρειες για τους στόχους του μακροπρόθεσμου προγραμματισμού. Ένα SM συμπεριλαμβάνεται στο GEP μοντέλο λόγω της σημασίας του επαρκούς εντοπισμού των απαιτήσεων λειτουργικής ευελιξίας ενός μελλοντικού εγκατεστημένου δυναμικού παραγωγής ισχύος . Για την αντιμετώπιση των υπολογιστικών απαιτήσεων αναπτύχθηκε μια προσέγγιση βασισμένη σε MAEA. Τα AM, που βασίζονται σε RBF, χρησιμοποιούνται για να παρέχουν μια εκτίμηση της επάρκειας της λειτουργικής ευελιξίας. Επομένως, οι υπολογιστικά ακριβές προσομοιώσεις αντικαθίστανται εν μέρει από υπολογιστικά φθηνούς δείκτες κόστους για τη μείωση του αριθμού των προσομοιώσεων που απαιτούνται για την επίτευξη σχεδόν βέλτιστης λύσης. Τόσο τα τοπικά όσο και τα ολικά AM χρησιμοποιούνται σε διαφορετικά στάδια της αναζήτησης. Επιπλέον, η προσέγγιση βελτιστοποίησης εκμεταλλεύεται συγκεκριμένα χαρακτηριστικά του προβλήματος. Συγκεκριμένα, αναπτύσσονται εξειδικευμένοι τελεστές για την βελτίωση της απόδοσης του επιλεγμένου EA που ένας αλγόριθμος DE. Τα αποτελέσματα ήταν ικανοποιητικά με βάση τα υπολογιστικά πειράματα που πραγματοποιήθηκαν. Πιο συγκεκριμένα, μια ευρετική τεχνική διόρθωσης, που στοχεύει στην αντιμετώπιση προκλήσεων που εισάγονται λόγο των συναρτήσεων περιορισμών στο πρόβλημα βελτιστοποίησης, παρείχε τη μεγαλύτερη βελτίωση στην απόδοση του εξεταζόμενου αλγόριθμου βάσης. Επιπλέον, οι μετρήσεις που χρησιμοποιήθηκαν για την αξιολόγηση της ποιότητας προσέγγισης από AM, όσο αφορά την ακρίβεια, υποδηλώνουν ότι αυτή ήταν σχετικά ακριβής. Επομένως, πραγματοποιήθηκε οπτική ανάλυση της ευαισθησίας του λειτουργικού κόστους προς την εγκατεστημένη δυναμικότητα παραγωγής ισχύος για τη λύση που εντοπίστηκε. Η επίδραση του επίπεδου τεχνικής λεπτομέρειας που περιλαμβάνεται στο SM εξετάζεται μέσω μιας συγκριτικής μελέτης. Τα αποτελέσματα υποδηλώνουν ότι διαφορετικά επίπεδα τεχνικής λεπτομέρειας μπορεί να επηρεάσουν τις παραγόμενες επενδυτικές αποφάσεις, τα αναμενόμενα επίπεδα παραγωγής της εγκατεστημένης δυναμικότητας παραγωγής ισχύος και το αναμενόμενο κόστος. Αυτά τονίζουν τη σημασία της επαρκούς αξιολόγησης των απαιτήσεων λειτουργικής ευελιξίας στο πλαίσιο του μακροπρόθεσμου προγραμματισμού για την εξαγωγή βελτιωμένων σημάτων για τη λήψη αποφάσεων.

2. Παρουσιάζεται ένα στατικό μοντέλο GEP πολλών αντικειμένων στόχων που βασίζεται σε MAEAς και στοχεύει στην ανάλυση αντικρουόμενων στόχων και συντελεστών κόστους που προκύπτουν

για ένα MOO GEP. Εισάγεται ένα SM που περιλαμβάνει τεχνική, χωρική και χρονική λεπτομέρεια για την αξιολόγηση της λειτουργικής ευελιξίας κατά την αναζήτηση βέλτιστης λύσης. Ο MAEA βασίζεται σε MOEA και στη βελτιστοποίηση προβλημάτων με χρήση μεταπροτύπων και χωρίς χρήση παραγώγων. Τα AM που επιλέγονται είναι τα RBF και PR. Αυτά παρέχουν μια εκτίμηση των τιμών των αντικειμενικών συναρτήσεων που εξετάζονται και χρησιμοποιούνται για την αντιμετώπιση του υπολογιστικού κόστους. Τα κριτήρια που χρησιμοποιήθηκαν για την ανανέωση του αρχείο που χρησιμοποιούν τα AM δίνουν προτεραιότητα στη ικανοποίηση των συναρτήσεων περιορισμού του μακροπρόθεσμου σχεδιασμού, στη χωρική κατανομή του συνόλου λύσεων στο χώρο αναζήτησης που είναι αποθηκευμένες στο αρχείο και σε πιθανές βελτιώσεις του υπερόγκου. Επίσης, εφαρμόζεται τοπική αναζήτηση με χρήση παραγώγων χρησιμοποιώντας RBF, PR και ένα συνδυασμό των δυο μοντέλων. Αυτά δημιουργούνται τοπικά και χρησιμεύουν ως τοπική φάση για τοπική αναζήτηση. Η απόδοση του MAEA εξετάζεται συναρτήσεις αναφοράς. Επίσης, διεξάγονται υπολογιστικά πειράματα για την αξιολόγηση του MAEA σε MOO GEP προβλήματα. Στα πρώτα δεν εξετάζεται η βραχυπρόθεσμη λειτουργία ενώ σε πέντε παραλλαγές MOO GEP συμπεριλαμβάνεται το επιλεγμένο SM. Τα αποτελέσματα που επιτεύχθηκαν ήταν ικανοποιητικά λαμβάνοντας υπόψη τους υπολογιστικό περιορισμούς. Επιπλέον, η επιτευχθείσα ακρίβεια των AM δεν ήταν σταθερή μεταξύ των διάφορων προβλημάτων που εξετάστηκαν. Συνεπώς, αναφέρονται κύριοι παράγοντες που εντοπίστηκαν και επηρέασαν την ακρίβεια του AM. Για κάθε ένα από τα πέντε MOO GEP εξετάζονται τα αποτελέσματα σε επίπεδο συντελεστών κόστους. Τα αποτελέσματα μπορούν να παρέχουν μια λεπτομερή ανάλυση του συνόλου λύσεων σε επίπεδο συντελεστών κόστους για την ανάλυση των διαφορετικών εναλλακτικών. Αυτό θα μπορούσε να συμβάλει στο προσδιορισμών απαιτούμενων κινήτρων για τη λήψη αποφάσεων σχετικών με τη στρατηγική ενεργειακή πολιτική.

1.3 Δομή

Η δομή της Διατριβής έχει οριστεί ως εξής: Τα Κεφάλαια 2 και 3 παρουσιάζουν μια σύντομη ανασκόπηση της σχετικής βιβλιογραφίας έτσι ώστε να εξηγηθούν πιο αναλυτικά το αντικείμενο και οι στόχοι της Διατριβής. Συγκεκριμένα, στο Κεφάλαιο 2 παρουσιάζεται βιβλιογραφία σχετική με το GEP με έμφαση στο μακροπρόθεσμο σχεδιασμό συστημάτων ισχύος με υψηλό μερίδιο RES. Ταυτόχρονα παρουσιάζονται βασικά χαρακτηριστικά των μοντέλων GEP και προκλήσεις σχετικές με την αύξηση του μεριδίου των RES. Στο Κεφάλαιο 3, παρουσιάζεται βιβλιογραφία σχετική με τους EA και τους EA υποβοηθούμενους από μεταπρότυπα. Το μονοκριτηριακό μοντέλο που αναπτύχθηκε παρουσιάζεται στο Κεφάλαιο 4 ενώ το πολυκριτηριακό μοντέλο στο Κεφάλαιο 5. Στο Κεφάλαιο 6 παρουσιάζονται τα τελικά συμπεράσματα.

Κεφάλαιο 2

Προγραμματισμός επέκτασης παραγωγής και αυξημένα μερίδια παραγωγής από ανανεώσιμες πηγές ενέργειας

Αυτό το κεφάλαιο παρέχει ένα υπόβαθρο για τα βραχυπρόθεσμα και μακροπρόθεσμα μοντέλα βελτιστοποίησης. Συγκεκριμένα, παρέχεται μια σύντομη εισαγωγή στον μακροπρόθεσμο σχεδιασμό με έμφαση στο GEP. Στη συνέχεια, παρέχεται μια σημείωση για τον βραχυπρόθεσμο προγραμματισμό, με έμφαση στο Πρόβλημα Δέσμευσης Μονάδων (Unit Commitment Problem, UCP). Επιπλέον, αναφέρονται οι κύριες προκλήσεις σχετικά με την αποτελεσματική εισαγωγή της παραγωγής από RES σε ένα σύστημα ισχύος. Τέλος, παρουσιάζεται η σχετική βιβλιογραφία για την ενσωμάτωση πτυχών της βραχυπρόθεσμης λειτουργίας σε μακροπρόθεσμα μοντέλα, η οποία αποτελεί το επίκεντρο της Διατριβής. Το κεφάλαιο στοχεύει στην παρουσίαση του πεδίου αυτής της Διατριβής παρέχοντας σχετικό υπόβαθρο.

2.1 Σύντομη σημείωση για τον μακροπρόθεσμο προγραμματισμό και τον προγραμματισμό επέκτασης παραγωγής

Ο μακροπρόθεσμος προγραμματισμός είναι ευρύ πεδίο και έχουν παρουσιαστεί διαφορετικές κατηγοριοποιήσεις με βάση διαφορετικά κριτήρια [9]. Τέτοια μοντέλα προορίζονται κυρίως για τη δημιουργία σεναρίων. Η Διατριβή αυτή εστιάζει στον μακροπρόθεσμο σχεδιασμό για τον οποίο το πεδίο περιορίζεται στον τομέα της ηλεκτρικής ενέργειας και για τον οποίο η μεθοδολογία περιορίζεται σε μοντέλα βελτιστοποίησης. Πιο συγκεκριμένα, αυτή η Διατριβή περιορίζεται σε μοντέλα σχεδιασμού επέκτασης δυναμικότητας ισχύος, για τον τομέα της ηλεκτρικής ενέργειας, που εξετάζουν την βραχυπρόθεσμη λειτουργία του συστήματος ισχύος με υψηλά μερίδια παραγωγής από RES.

Ένας από τους κύριους στόχους μοντέλων μακροπρόθεσμου προγραμματισμού είναι η εξέταση της ικανότητας ενός συστήματος ισχύος να ικανοποιεί την αναμενόμενη ζήτηση σε έναν μακροπρόθεσμο ορίζοντα προγραμματισμού (έτη έως δεκαετίες). Βασικά ζητήματα που εξετάζονται είναι η δυναμικότητα παραγωγής ισχύος και το δίκτυο μεταφοράς [2]. Αυτά εξετάζονται συνήθως σε όρους τεχνικής, οικονομικής και περιβαλλοντικής αποτελεσματικότητας, αξιοπιστίας και ευελιξίας.

Συγκεκριμένα, τα μοντέλα που επικεντρώνονται στην ανάλυση της εγκατεστημένης δυναμικότητα παραγωγής ισχύος και της πιθανής επέκτασής της, αναφέρονται συνήθως ως μοντέλα GEP [18]. Σε αυτήν την περίπτωση, εξετάζονται πιθανές επενδύσεις σε δυναμικότητα και/ή η εισαγωγή νέων μονάδων παραγωγής λαμβάνοντας υπόψη οικονομικά κριτήρια, την αξιοπιστία του συστήματος και περιβαλλοντικά κριτήρια. Ένας από τους κύριους παράγοντες που οδηγεί σε αυτές τις προσθήκες είναι, συνήθως, η αναμενόμενη αύξηση της ζήτησης. Επιπλέον, μπορεί να απαιτούνται λόγω της απόσυρσης μονάδων ή από παράγοντες που σχετίζονται με ενεργειακή πολιτική.

Αντίστοιχα, τα μοντέλα με έμφαση στην ανάλυση του δικτύου μεταφοράς και την πιθανή επέκτασή του, δηλαδή τον σχεδιασμό επέκτασης του συστήματος μεταφοράς (Transmission Expansion Planning, TEP), αναφέρονται συνήθως ως μοντέλα TEP [18]. Σε αυτήν την περίπτωση, η ανάλυση επικεντρώνεται στην επάρκεια του υπάρχοντος δικτύου μεταφοράς όσον αφορά την αποτελεσματικότητα και την αξιοπιστία. Ομοίως με το GEP, η ανάγκη για επέκταση του δικτύου μεταφοράς μπορεί να πηγάζει από την μη αποδοτική λειτουργία του υπάρχοντος δικτύου και την αναμενόμενη αύξηση της ζήτησης. Ωστόσο, υπάρχουν και άλλα κίνητρα για το TEP, όπως νέες εγκαταστάσεις από RES και η δημιουργία διασυνδέσεων μεταξύ απομονωμένων περιοχών. Πολλά μοντέλα περιλαμβάνουν τόσο το GEP όσο και το TEP (GEP-TEP), καθώς είναι σημαντικό και τα δύο να αναλύονται στο πλαίσιο του μακροπρόθεσμου προγραμματισμού. Τέτοια μοντέλα (GEP-TEP) μπορεί να παρέχουν εναλλακτικά σχέδια επέκτασης και η εξέταση των GEP και TEP πρέπει να εξετάζεται τουλάχιστον παράλληλα [2]. Οι ακόλουθες ενότητες αφορούν το GEP που αποτελεί το επίκεντρο αυτής της Διατριβής.

2.1.1 Προγραμματισμός επέκτασης παραγωγής

Τα μοντέλα GEP αποτελούν ένα από τα πιο σημαντικά εργαλεία για τη στήριξη της λήψης αποφάσεων στον μακροπρόθεσμο σχεδιασμό του τομέα της ενέργειας [2, 1, 4, 6]. Επιπλέον, οι πρόσφατες προκλήσεις που παρουσιάστηκαν στους τομείς της ενέργειας, όπως νέα οικονομικά, τεχνικά, περιβαλλοντικά και ρυθμιστικά ζητήματα, ανάγκασαν τη σταδιακή δημιουργία μοντέλων που αποκλίνουν από το παραδοσιακό πλαίσιο του GEP. Η κατηγοριοποίηση των μοντέλων GEP μπορεί να γίνει με διαφορετικούς τρόπους λαμβάνοντας υπόψη διαφορετικά κριτήρια. Η Αναφορά [1] παρουσιάζει μια ταξινόμηση μελετών σχετικά με το GEP με βάση διάφορους παράγοντες όπως την απελευθέρωση της αγοράς ηλεκτρικής ενέργειας, περιβαλλοντικά ζητήματα, νέες αναδυόμενες τεχνολογίες, την ενεργειακή πολιτική και τις αναδυόμενες τεχνικές στους τομείς βελτιστοποίησης και μοντελοποίησης. Τέτοιοι παράγοντες έχουν οδηγήσει στην ανάπτυξη πολλών μοντέλων GEP.

Προγραμματισμός επέκτασης παραγωγής: Βασικές αποφάσεις

Συνήθως, τέσσερις βασικές ερωτήσεις σχετίζονται με ένα μοντέλο GEP: τι, πόσο, πού και πότε [1, 4, 18]. Συγκεκριμένα, ένα μοντέλο GEP πρέπει να παρέχει τις απαντήσεις στον τύπο και το μέγεθος των προτεινόμενων προσθηκών ικανότητας, πού πρέπει να εγκατασταθούν, και πότε πρέπει να πραγματοποιηθούν οι επενδύσεις. Κάθε μία από αυτές τις ερωτήσεις είναι σημαντική για τον βέλτιστο προσδιορισμό των επενδύσεων σε δυναμικότητα παραγωγής ισχύος που ικανοποιούν επαρκώς την αύξηση της ζήτησης σε έναν εξεταζόμενο μακροπρόθεσμο ορίζοντα προγραμματισμού.

Προγραμματισμός επέκτασης παραγωγής: Πλαίσια βασισμένα στη δομή της αγοράς ηλεκτρικής ενέργειας

Υπάρχουν διαφορετικά πλαίσια προσέγγισης του GEP με βάση την απελευθέρωση της αγορά ηλεκτρικής ενέργειας [1]. Η πρώτη είναι η κεντρική προσέγγιση όπου η ανάλυση γίνεται με βάση όλο το σύστημα.

Η δεύτερη βασίζεται στην απελευθέρωση της αγοράς ηλεκτρικής ενέργειας. Σε αυτή τη Διατριβή υιοθετείται ένα κεντρικό πλαίσιο σχεδιασμού. Η κεντρική προσέγγιση στοχεύει στον καθορισμό των επενδύσεων σε προσθήκες δυναμικότητας παραγωγής ισχύος που ικανοποιούν τον προκαθορισμένο στόχο και τους περιορισμούς με τον βέλτιστο τρόπο. Τέτοιοι στόχοι είναι η ελαχιστοποίηση του κόστους (π.χ. συνολικό, παραγωγή ή/και επενδυτικό κόστος) ή η μεγιστοποίηση της κοινωνικής πρόνοιας.

Σχεδιασμός επέκτασης παραγωγής: Πλαίσια βελτιστοποίησης ενός και πολλαπλών στόχων

Παραδοσιακά, τα μοντέλα GEP θεωρούν τη διαμόρφωση προβλήματος βελτιστοποίησης ενός στόχου (Single Objective Optimization, SOO) [1]. Ωστόσο, ο μακροπρόθεσμος προγραμματισμός περιλαμβάνει πολλούς άλλους στόχους [1, 4]. Ένα μοντέλο GEP θα μπορούσε να στοχεύει στην ελαχιστοποίηση της συνάρτησης συνολικού κόστους που αντιπροσωπεύει το συνολικό κόστος του συστήματος. Σε περίπτωση που περισσότεροι από ένας στόχοι λαμβάνονται υπόψη, τότε αυτοί μπορούν να συμπεριληφθούν σε μια ενιαία σταθμισμένη συνάρτηση στόχος. Ορισμένοι στόχοι μπορούν να επιβληθούν ως ένα σύνολο περιορισμών που αντιπροσωπεύονται ως περιορισμοί ή όροι ποινής [19].

Εναλλακτικά, έχουν προταθεί πολυ-κριτηριακές προσεγγίσεις GEP λόγω της αυξανόμενης σημασίας παραμέτρων όπως τα RES, οι εκπομπές αερίων του θερμοκηπίου και αξιοπιστίας [3]. Επομένως, έχουν εφαρμοστεί μέθοδοι λήψης αποφάσεων πολλαπλών κριτηρίων (Multi Criteria Decision Making, MCDM) για την υποστήριξη της λήψης αποφάσεων παρουσία πολλαπλών και αντικρουόμενων στόχων. Συνήθως, αυτές οι προσεγγίσεις κατηγοριοποιούνται στη λήψη αποφάσεων πολλαπλών χαρακτηριστικών και στη λήψη αποφάσεων πολλαπλών στόχων [1, 4].

Σε προσεγγίσεις λήψης αποφάσεων πολλαπλών χαρακτηριστικών, διακριτές και προκαθορισμένες εναλλακτικές λύσεις συγκρίνονται με βάση ένα σύνολο κριτηρίων απόφασης [4]. Αντίθετα, η λήψη αποφάσεων πολλαπλών στόχων εστιάζει σε περιπτώσεις όπου ο αριθμός των εναλλακτικών αποφάσεων είναι μεγάλος. Επομένως, διαμορφώνεται ένα πρόβλημα βελτιστοποίησης, λαμβάνοντας συνήθως υπόψη τα κριτήρια ως αντικειμενικές συναρτήσεις και τους πιθανούς περιορισμούς. Γενικά, η καταλληλότερη λύση πρέπει να προσδιορίζεται εκ των υστέρων βάσει της προτίμησης του υπεύθυνου λήψης αποφάσεων [19]. Διαφορετικοί στόχοι έχουν εξεταστεί σε μοντέλα GEP όπως η ελαχιστοποίηση του κόστους, των εκπομπών αερίων του θερμοκηπίου ή στόχοι που σχετίζονται με την αξιοπιστία του συστήματος [1, 3, 4].

Προγραμματισμός επέκτασης παραγωγής: Αβεβαιότητα και αξιολόγηση κινδύνου

Ο μακροπρόθεσμος προγραμματισμός υπόκειται σε αβεβαιότητες. Τα βασικά στοιχεία κινδύνου στη διαδικασία GEP έχουν κατηγοριοποιηθεί στα ακόλουθα [2]: οικονομικά, πολιτικά, ρυθμιστικά, περιβαλλοντικά, τεχνικά, κοινωνικά και κλιματικά. Τα ντετερμινιστικά μοντέλα GEP θεωρούν ακριβή πληροφορία σχετικά με αυτές τις αβεβαιότητες, ενώ τα στοχαστικά μοντέλα χρησιμοποιούν προσεγγίσεις μοντελοποίησης αβεβαιότητας (π.χ. στοχαστικός προγραμματισμός) [4].

2.1.2 Προγραμματισμός επέκτασης παραγωγής και βραχυπρόθεσμη λειτουργία

Ο κύριος στόχος ενός βραχυπρόθεσμου μοντέλου είναι να υποστηρίξει τη λήψη αποφάσεων προς τον βέλτιστο προγραμματισμό των διαθέσιμων μονάδων παραγωγής σε ένα σύστημα ισχύος για να εξισορροπήσει την προσφορά και τη ζήτηση σε ένα βραχυπρόθεσμο ορίζοντα προγραμματισμού (μία ημέρα έως δύο εβδομάδες). Με βάση τη δομή της αγοράς ηλεκτρικής ενέργειας, το κόστος παραγωγής (εν μέρει) καθορίζει την οριακή τιμή. Επομένως, η ανεπαρκής λειτουργία της μπορεί να επηρεάσει το κόστος που μεταφέρεται στους καταναλωτές και τα κέρδη των επενδυτών. Κατά συνέπεια, πρέπει να ληφθούν υπόψη τα τεχνοοικονομικά χαρακτηριστικά των μονάδων για τον καθορισμό ενός εφικτού και βέλτιστου προγράμματος παραγωγής.

Βραχυπρόθεσμη λειτουργία: η σειρά ένταξης

Παραδοσιακά, πρέπει να ληφθεί υπόψη η βραχυπρόθεσμη λειτουργία στο πλαίσιο του μακροπρόθεσμου προγραμματισμού. Ωστόσο, μέχρι πρόσφατα, τα ιστορικά μοτίβα ζήτησης χαρακτηρίζονταν από δυναμικά προβλέψιμες και σχετικά αργές χρονικές δυναμικές [8]. Συνεπώς, μια κοινή υπόθεση ήταν ότι η ζήτηση και το μεταβλητό κόστος είναι οι κύριες δυνάμεις που επηρεάζουν την κατανομή φορτίου σε διαφορετικές συνθήκες λειτουργίας. Αυτό επιτρέπει την χρήση πιο χαμηλής λεπτομέρειας όσο αφορά τη βραχυπρόθεσμη λειτουργία στο GEP. Με βάση τα προαναφερθέντα, μια βέλτιστη κατανομή φορτίου των μονάδων παραγωγής μπορεί να καθοριστεί με βάση το μεταβλητό κόστος τους (merit-order) και τις καμπύλες διάρκειας φορτίου (Load Duration Curve, LDC). Η υπολογιστική αποτελεσματικότητα τέτοιων μοντέλων μπορεί να παρέχει ένα χρήσιμο πλαίσιο για την εξέταση των τιμών και την κατανομή φορτίου σε διάφορα μοντέλα [18].

Για τις περιπτώσεις όπου εξετάζεται η παραγωγή από RES, εισάγονται τροποποιήσεις. Συνήθως, οι ΛΔ῝ς αντικαθίστανται από τις καμπύλες καθαρής διάρκειας φορτίου (Net-Load Duration Curves, NL-DC) όπου η παραγωγή από RES αφαιρείται από το φορτίο. Ένα χαρακτηριστικό τέτοιων προσεγγίσεων είναι η απώλεια χρονικής λεπτομέρειας όσο αφορά την κατανομή φορτίου που θα μπορούσε να επηρεάσει τις επενδυτικές αποφάσεις [20]. Επιπλέον, η απώλεια της χρονικής λεπτομέρειας μπορεί να καταστήσει δύσκολη την αναλυτική εξέταση της μεταβλητότητας της παραγωγής από RES, την επίδραση των τεχνικών περιορισμών των θερμικών μονάδων ή/και τις δυνατότητες αποθήκευσης [9]. Επομένως, έχουν γίνει προσπάθειες για τη βελτίωση της ακρίβειας τέτοιων προσεγγίσεων [21, 22, 23].

Βραχυπρόθεσμη λειτουργία: το πρόβλημα δέσμευσης μονάδων

Μια άλλη κύρια κατηγορία μοντέλων για το βραχυπρόθεσμο προγραμματισμό είναι μοντέλα για το UCP. Σε γενικές γραμμές, ένα μοντέλο UCP περιλαμβάνει δύο προβλήματα: το πρόβλημα δέσμευσης μονάδων και το πρόβλημα οικονομικής κατανομή φορτίου. Το πρώτο αφορά τον ορισμό της κατάστασης λειτουργίας κάθε μονάδας για κάθε περίοδο προγραμματισμού και το δεύτερο την κατανομή φορτίου μεταξύ των δεσμευμένων μονάδων. Συγκεκριμένα, ένα μοντέλο UCP λαμβάνει υπόψη τεχνικούς περιορισμούς των μονάδων, οικονομικά και τεχνικά χαρακτηριστικά των μονάδων και κριτήρια αξιοπιστίας για τον καθορισμό του προγράμματος παραγωγής και της αποδοτικής ικανοποίησης της ζήτησης. Ένα UCP είναι ένα σύνθετο πρόβλημα βελτιστοποίησης που μπορεί να διατυπωθεί ως πρόβλημα Μικτού Γραμμικού Προγραμματισμού (Mixed Integer Linear Programming, MILP). Μακροπρόθεσμα μοντέλα UCP έχουν επίσης παρουσιαστεί που εξετάζουν μεγαλύτερες περιόδους προγραμματισμού (έως ένα έτος) [24, 25, 26].

2.2 Μακροπρόθεσμος προγραμματισμός με αυξημένη λεπτομέρεια της βραχυπρόθεσμης λειτουργίας ενός συστήματος ισχύος

Η αυξανόμενη διείσδυση της παραγωγής από RES και η αποτελεσματική ενσωμάτωσή της σε ένα σύστημα ισχύος έχει οδηγήσει στην εισαγωγή αυξημένης λεπτομέρειας της βραχυπρόθεσμης λειτουργίας σε μοντέλα του μακροπρόθεσμου σχεδιασμού [2]. Αυτό γιατί η μεταβλητότητα και η αβεβαιότητα της παραγωγής RES πρέπει να ληφθούν υπόψη για να εξεταστούν οι ανάγκες σε λειτουργική ευελιξία. Η λειτουργική ευελιξία αναφέρεται στην ικανότητα του συστήματος ισχύος να ανταποκρίνεται στις διακυμάνσεις και να ικανοποιεί το καθαρό φορτίο εντός ενός αποδεκτού χρονικού πλαισίου με ρύθμιση της προσφοράς. Ως αποτέλεσμα, η αξία της λειτουργικής ευελιξίας πρέπει επίσης να εκτιμηθεί κατά τη διάρκεια του μακροπρόθεσμου προγραμματισμού όταν λαμβάνονται υπόψη υψηλά μερίδια παραγωγής από RES.

2.2.1 Βασικά χαρακτηριστικά της παραγωγής από ανανεώσιμες πηγές ενέργειας

Τα κύρια χαρακτηριστικά της παραγωγής ηλεκτρικής ενέργειας από RES που εξαρτώνται από τις καιρικές συνθήκες είναι οι ακόλουθες:

- Η παραγωγή ηλεκτρικής ενέργειας από RES είναι μεταβλητή: Αυτή η μεταβλητότητα μπορεί να παρατηρηθεί σε διαφορετικές χρονικές κλίμακες.
- Η παραγωγή ηλεκτρικής ενέργειας από RES εμπεριέχει αβεβαιότητα: Η αβεβαιότητα της παραγωγής από RES είναι συνέπεια της εξάρτησης τους από τις μετεωρολογικές συνθήκες.
- Το οριακό κόστος της παραγωγής ηλεκτρικής ενέργειας από RES είναι αμελητέο: Το αμελητέο οριακό κόστος της παραγωγής RES οδηγεί σε προτεραιότητα της παραγωγής αυτής κατά την κατανομή φορτίου [27].
- Η παραγωγή ηλεκτρικής ενέργειας από RES είναι απαλλαγμένη από εκπομπές GHG: Η αύξηση του μεριδίου παραγωγής ηλεκτρικής ενέργειας από RES μπορεί να συμβάλει στη μείωση των εκπομπών GHG [28].

2.2.2 Ανάγκη για την αξιολόγηση της λειτουργικής ευελιξίας στο μακροπρόθεσμο σχεδιασμό

Η ανάγκη ενσωμάτωσης αυξημένης λεπτομέρειας της βραχυπρόθεσμης λειτουργίας ενός συστήματος ισχύος στον μακροπρόθεσμο προγραμματισμό οφείλεται κυρίως στην ανάγκη για αξιόπιστη λειτουργία ενός συστήματος ισχύος. Αυτό μπορεί να παρέχει πρόσθετες, ή σε ορισμένες περιπτώσεις πιο ακριβείς, πληροφορίες για την αξιολόγηση σεναρίων και επιλογών στον μακροπρόθεσμο προγραμματισμό [7, 8, 29, 30]. Η αύξηση των απαιτήσεων λειτουργικής ευελιξίας οφείλεται στη μεταβλητότητα και την αβεβαιότητα της παραγωγής από RES. Πιο συγκεκριμένα [31], οι συμβατικές μονάδες θα πρέπει να μπορούν να ικανοποιούν τη ζήτηση μετά από τις αποκλίσεις μεταξύ του προβλεπόμενου φορτίου που χρησιμοποιείται για τη δέσμευση μονάδων και του φορτίου που χρησιμοποιείται στην αγορά εξισορρόπησης σε πραγματικό χρόνο. Σε υψηλά επίπεδα διείσδυσης RES τα σφάλματα πρόβλεψης που προκύπτουν λόγω της αβεβαιότητας και της μεταβλητότητας των RES πρέπει επίσης να ληφθούν υπόψη. Η μεταβλητότητα του καθαρού φορτίου σε σύγκριση με τη μεταβλητότητα του φορτίου μπορεί να είναι αυξημένη. Σε τέτοιες περιπτώσεις, οι ευέλικτες μονάδες μπορεί να λειτουργούν πιο συχνά και/ή οι μονάδες μπορεί να υποχρεωθούν σε συχνότερες εκκινήσεις. Επίσης, το μερίδιο της παραγωγής των μονάδων βάσης θα μπορούσε να αντικατασταθεί από παραγωγή από ενδιάμεσες μονάδες λόγω της διείσδυσης της αιολικής ενέργειας, οι τεχνικοί περιορισμοί ρυθμού ανάληψης και απόρριψης φορτίου των μονάδων βάσης είναι σημαντικοί περιορισμοί που πρέπει να λαμβάνονται υπόψη, το σύστημα μεταφοράς και η αποθήκευση ενέργειας μπορούν να χρησιμεύσουν ως πάροχοι λειτουργικής ευελιξίας, και η προσθήκη μονάδων αιχμής μπορεί να είναι λιγότερο απαραίτητη [32].

2.2.3 Πάροχοι λειτουργικής ευελιξίας

Η λειτουργική ευελιξία δεν περιορίζεται στην τεχνική λειτουργική ευελιξία που παρέχουν οι θερμικές μονάδες. Από την πλευρά της προσφοράς, λειτουργική ευελιξία προσφέρεται κυρίως από τις μονάδες παραγωγής ηλεκτρικής ενέργειας (θερμικές και μη-θερμικές μονάδες) [33]. Από την πλευρά της ζήτησης, λειτουργική ευελιξία μπορεί να προσφέρουν συστήματα ανταπόκρισης στη ζήτηση, έξυπνα ενεργειακά δίκτυα (Smart Energy Grid), αποθηκευτικές μονάδες και διασυνδέσεις με άλλα δίκτυα [2, 33]. Όσο αφορά τα RES, ο βασικός μηχανισμός ελέγχου της παραγωγής τους είναι η εκούσια μείωση της παραγωγής τους.

2.2.4 Υπολογιστικοί περιορισμοί και απλοποιήσεις

Ένας από τους κύριους περιορισμούς για την ενσωμάτωση μιας λεπτομερούς αναπαράστασης της βραχυπρόθεσμης λειτουργίας στον μακροπρόθεσμο σχεδιασμό, όπως ένα μοντέλο GEP, είναι οι υπολογιστικοί περιορισμοί. Μια βασική τεχνική για αυτή την ενσωμάτωση είναι η ένταξη των κύριων πτυχών του UCP σε μακροπρόθεσμα μοντέλα [9]. Και τα δύο μοντέλα GEP και UCP είναι υπολογιστικά απαιτητικά [7]. Συνεπώς σε ένα μοντέλο που αποτελεί συνδυασμό αυτών εισάγονται παράλληλα απλοποιήσεις. Οι απλοποιήσεις σχετίζονται με το επίπεδο τεχνικής, χρονικής και χωρικής λεπτομέρειας που χρησιμοποιείται για την περιγραφή του συστήματος ισχύος [9]. Επιπλέον, έχουν εξεταστεί προσεγγίσεις που εστιάζουν στην αντιμετώπιση του υπολογιστικού κόστους τέτοιων μοντέλων [2].

2.2.5 Επιπτώσεις χαμηλής τεχνικής, χρονικής και χωρικής λεπτομέρειας της βραχυπρόθεσμης λειτουργίας στον μακροπρόθεσμο σχεδιασμό

Η μελέτη της επίδρασης της μεταβαλλόμενης παραγωγής από RES στη βραχυπρόθεσμη λειτουργία ενός συστήματος ισχύος στο πλαίσιο του μακροπρόθεσμου προγραμματισμού θα πρέπει να μπορεί να λαμβάνει υπόψη: (i) τεχνική λεπτομέρεια [7, 8, 34, 35], (ii) χρονική λεπτομέρεια [36, 37, 38, 39, 40], και (iii) χωρική λεπτομέρεια [41, 42]. Η επίδραση της μεταβλητότητας της παραγωγής από RES στη βραχυπρόθεσμη λειτουργία ενός συστήματος ισχύος μπορεί να εξαρτάται από το σύστημα ισχύος υπό εξέταση [10] άλλα αυτή αυξάνεται με υψηλότερα επίπεδα διείσδυσης των RES. Γενικά, η παράλειψη τεχνικών, χρονικών και / ή χωρικών λεπτομερειών μπορεί να έχει αντίκτυπο στον βέλτιστο σχεδιασμό. Πιο συγκεκριμένα:

- Τεχνική λεπτομέρεια: Ο περιορισμός της τεχνικής λεπτομέρειας μπορεί να έχει αντίκτυπο στον μακροπρόθεσμο σχεδιασμό καθώς μπορεί να υπερεκτιμήσει την λειτουργική ευελιξία του συστήματος ισχύος λόγω της μη εξέτασης των τεχνικών περιορισμών.

- Χρονική λεπτομέρεια: Ο περιορισμός της χρονικής λεπτομέρειας μπορεί να οδηγήσει σε υπερεκτίμηση της παραγωγής από RES και από μη ευέλικτες μονάδες και σε υποτίμηση των επενδύσεων σε λειτουργικά ευέλικτες μονάδες.

- Χωρική λεπτομέρεια: Η ενσωμάτωση χωρικής λεπτομέρειας μπορεί να οδηγήσει στον εντοπισμό πιθανού οφέλους από τη εξομάλυνση της παραγωγής RES λόγω χωρικής διαφοροποίησης και προκλήσεων που σχετίζονται με πιθανή συμφόρηση του συστήματος μεταφοράς.

2.2.6 Αντιπροσωπευτικά μοντέλα επέκτασης παραγωγής που ενσωματώνουν αυξημένη λεπτομέρεια της βραχυπρόθεσμης λειτουργίας

Αυτή η ενότητα εστιάζει σε μοντέλα βελτιστοποίησης για το GEP, που εστιάζουν σε υψηλά μερίδια παραγωγής από RES. Ορισμένες αντιπροσωπευτικές προσεγγίσεις έχουν κατηγοριοποιηθεί σε: (i) στατικά μοντέλα GEP που περιλαμβάνουν μοντέλο για το UCP [7, 8, 43, 44, 45], (ii) μοντέλα GEP πολλαπλών περιόδων που περιλαμβάνουν μοντέλο για το UCP [46, 47, 48], (iii) προσεγγίσεις που εστιάζουν σε υπολογιστικούς περιορισμούς [49, 50], (iv) προσεγγίσεις που περιλαμβάνουν δείκτες λειτουργικής ευελιξίας [51, 52, 53], (v) προσεγγίσεις που εστιάζουν στην δομή της αγορά ηλεκτρικής ενέργειας [54, 55], (vi) προσεγγίσεις που εξετάζουν την επίδραση τεχνολογιών αποθήκευσης ηλεκτρικής ενέργειας [56, 57], και (vii) προσεγγίσεις MOO [6, 58, 59, 60, 61]. Εκτός από τις προσεγγίσεις που βασίζονται σε κλασικές μεθόδους βελτιστοποίησης, έχουν χρησιμοποιηθεί ευρετικές και/ή μετα-ευρετικές τεχνικές στο πρόβλημα του GEP.

Γενικά, οι κλασικές προσεγγίσεις βελτιστοποίησης παρουσιάζουν ένα σημαντικό πλεονέκτημα όταν αυτές εφαρμόζεται: μπορούν να εγγυηθούν βέλτιστες λύσεις σε έναν αριθμό βημάτων σε προβλήματα βελτιστοποίησης που μπορούν να εφαρμοστούν. Αντιθέτως, ευρετικές ή μετα-ευρετικές προσεγγίσεις θα μπορούσαν να εφαρμοστούν σε σύνθετα μονο- και πολύ- κριτητριακά προβλήματα βελτιστοποίησης. Επιπλέον, οι ευρετικές τεχνικές θα μπορούσαν να παρέχουν ικανοποιητικά αποτελέσματα εντός αποδεκτού χρονικού ορίου για υπολογιστικά δαπανηρά προβλήματα βελτιστοποίησης. Συνεπώς, έχουν αναπτυχθεί προσεγγίσεις SOO [62, 63, 64, 65, 66, 67, 68, 69] και MOO [70, 71, 72] που βασίζονται σε EA.

2.3 Συζήτηση

Έχουν αναπτυχθεί διάφορες προσεγγίσεις GEP που εξετάζουν ένα ευρύ φάσμα στόχων. Τέτοιες προσεγγίσεις διαφέρουν στα εγγενή πλεονεκτήματα και τους περιορισμούς τους σχετικά με την τεχνο-οικονομική, χρονική και χωρική λεπτομέρεια που ενσωματώνουν. Επιπλέον, το υπολογιστικό κόστος και η απόδοση τέτοιων προσεγγίσεων μπορεί να διαφέρει. Το GEP είναι γνωστό ως ένα σύνθετο πρόβλημα λόγω της μη γραμμικότητάς του, του αυξημένου αριθμού μεταβλητών και των διακριτών μεταβλητών που εμπεριέχει [1]. Επομένως εισάγονται απλουστεύσεις στην μοντελοποίηση και στο τρόπο επίλυσης για να αντιμετωπιστεί το υπολογιστικό κόστος. Κατά συνέπεια, απαιτείται η εξισορρόπηση της ακρίβειας της μοντελοποίησης και του υπολογιστικού κόστους.

Ο κύριος στόχος αυτής της Διατριβής είναι να συμπεριλάβει βασικές πτυχές της βραχυπρόθεσμης λειτουργίας ενός συστήματος ισχύος στο πλαίσιο του μακροπρόθεσμου προγραμματισμού ισχύος μέσω μοντέλου προσομοίωσης. Το εξεταζόμενο επίπεδο λεπτομέρειας περιορίζεται σε ωριαία διαστήματα για τον εντοπισμό προκλήσεων που αφορούν το βραχυπρόθεσμο προγραμματισμό και που εξετάζονται συνήθως σε ένα μοντέλο UCP [9]. Οι προσεγγίσεις που παρουσιάζονται βασίζονται σε EA και συγκεκριμένα σε MAEA για την αντιμετώπιση του υπολογιστικού κόστους. Προσεγγίσεις βασισμένες σε

ΜΑΕΑ έχουν εφαρμοστεί επιτυχώς για την εξέταση της βραχυπρόθεσμης λειτουργίας ενός συστήματος ισχύος όταν δηλαδή ο μακροπρόθεσμος προγραμματισμός δεν λαμβάνεται υπόψη [73, 74]. Επιπλέον, οι ΜΑΕΑ που παρουσιάζονται περιλαμβάνουν υβριδισμό όπως ευρετικές επιδιόρθωσης μη εφικτών λύσεων ή τοπική αναζήτηση.

Κεφάλαιο 3

Εξελικτικοί Αλγόριθμοι

Οι ΕΑ είναι υπολογιστικές μέθοδοι εμπνευσμένες από τη φύση που έχουν χρησιμοποιηθεί συχνά σε σύνθετα προβλήματα βελτιστοποίησης. Γενικά, οι ΕΑ βασίζονται στην αρχή της φυσικής επιλογής του Δαρβίνου [75]. Συγκεκριμένα χαρακτηριστικά των ΕΑ, που παρουσιάζονται στις ακόλουθες ενότητες, τους καθιστούν κατάλληλους για πολλές εφαρμογές.

3.1 Εισαγωγικά

Σε αυτή τη Διατριβή, λαμβάνονται υπόψη μονο- και πολύ-κριτηριακά προβλήματα βελτιστοποίησης μαύρου κουτιού (black-box). Επιπλέον, εξετάζονται ορισμένες ειδικές περιπτώσεις SOO, όπως Γραμμικός Προγραμματισμός, Μικτός-Ακέραιος Γραμμικός Προγραμματισμός και Ακέραιος Γραμμικός Προγραμματισμός. Αυτές οι διατυπώσεις παρουσιάζονται στο πλήρες κείμενο.

3.1.1 Βασικά χαρακτηριστικά των εξελικτικών αλγορίθμων

Οι ΕΑ βασίζονται σε χρήση πληθυσμού λύσεων και χρησιμοποιούν στοχαστικούς τελεστές στη διαδικασία αναζήτησης για να βελτιώσουν τις λύσεις αυτές.

Τα πιο σημαντικά μέρη ενός ΕΑ είναι εμπνευσμένα από την φυσική εξέλιξη και την αρχή της φυσικής επιλογής [75]. Αυτά τα στοιχεία περιλαμβάνονται στους περισσότερους ΕΑ ως τελεστές, αλλά διαφέρουν στον τρόπο υλοποίησής τους. Τα βασικά μέρη ενός ΕΑ είναι [76, 77]:

- Άτομο (αναπαράσταση και κωδικοποίηση): Ένα άτομο του πληθυσμού αναπαριστά μια υποψήφια λύση του προβλήματος δηλαδή το διάνυσμα μεταβλητών απόφασης. Ο τρόπος με τον οποίο κάθε μεταβλητή απόφασης αντιπροσωπεύεται σε ένα άτομο καθορίζεται με βάση την κωδικοποίηση που χρησιμοποιείται σε ένα ΕΑ.

- Πληθυσμός: Ο πληθυσμός είναι το σύνολο των διαθέσιμων λύσεων. Ο στόχος είναι να βελτιωθεί η συνολική καταλληλότητα του πληθυσμού μετά από μια σειρά επαναλήψεων. Αυτές οι επαναλήψεις αναφέρονται ως γενιές.

- Συνάρτηση καταλληλότητας: Η συνάρτηση καταλληλότητας χρησιμοποιείται για τη σύγκριση ατόμων και παρέχει τον δείκτη καταλληλότητας για ένα άτομο που μπορεί να είναι η τιμή της αντικειμενικής συνάρτησης ενός προβλήματος βελτιστοποίησης.

- Τελεστής αρχικοποίησης: Συνήθως, ένας τελεστής αρχικοποίησης παρέχει έναν αρχικό πληθυσμό στο χώρο αναζήτησης.
- Τελεστής επιλογής γονέων: Ένας τελεστής επιλογής γονέων καθορίζει ποιοι γονείς επιλέγονται για αναπαραγωγή απογόνων.
- Τελεστής ανασυνδυασμού: Ένας τελεστής ανασυνδυασμού συνδυάζει έναν αριθμό ατόμων από τον πληθυσμό των γονέων και δημιουργεί τον πληθυσμό των απογόνων.
- Τελεστής μετάλλαξης: Ένας τελεστής μετάλλαξης αποσκοπεί στο να εισαγάγει μικρές αλλαγές στους απογόνους. Ένας από τους κύριους λόγους για την εισαγωγή τέτοιων μικρών αλλαγών είναι η διατήρηση μη μηδενικής πιθανότητας δημιουργίας ενός απογόνου σε οποιοδήποτε σημείο εντός του χώρου αναζήτησης.
- Τελεστής περιβαλλοντικής επιλογής: Ένας τελεστής περιβαλλοντικής επιλογής καθορίζει ποιοι γονείς και απόγονοι θα σχηματίσουν τον πληθυσμό της επόμενης γενιάς.
- Κριτήρια τερματισμού: Ο ΕΑ απαιτεί κριτήρια τερματισμού που είναι συνήθως καθορίζονται από τον χρήστη.
- Παράμετροι που προσδιορίζονται από τον χρήστη: Οι ΕΑ απαιτούν τον προσδιορισμό παραμέτρων από τον χρήστη, π.χ. το μέγεθος του πληθυσμού ή άλλες παραμέτρους που απαιτούνται για τους τελεστές των ΕΑ. Αντιπροσωπευτικοί ΕΑ είναι οι Γενετικοί Αλγόριθμοι [78, 79], οι Στρατηγικές Εξέλιξης [80, 81] και η Διαφορική Εξέλιξη [82, 83].

3.1.2 Εξελικτικοί αλγόριθμοι και βελτιστοποίηση υπό περιορισμούς

Για προβλήματα βελτιστοποίησης υπό περιορισμούς (Constraint Optimization Problem, COP) εισάγονται τεχνικές χειρισμού περιορισμών (Constraint Handling Techniques, CHT) στους ΕΑ. Βασικές κατηγορίες CHT είναι οι συναρτήσεις ποινής, οι αποκωδικοποιητές, οι ειδικοί τελεστές και οι τεχνικές που χειρίζονται ξεχωριστά την αντικειμενική συνάρτηση και τους περιορισμούς, και προσεγγίσεις που χρησιμοποιούν περισσότερες από μια CHT [84, 85, 86].

3.1.3 Πολύ-κριτηριακή Βελτιστοποίηση με χρήση εξελικτικών αλγορίθμων

Πολλά προβλήματα βελτιστοποίησης μπορούν να διατυπωθούν ως προβλήματα MOO. Ένα πρόβλημα MOO περιλαμβάνει τον εντοπισμό ενός συνόλου λύσεων (το μέτωπο μη-κυριαρχούμενων λύσεων) όταν εξετάζονται περισσότεροι από έναν αντικρουόμενοι στόχοι. Σε τέτοια περίπτωση μπορεί να μην υπάρχει μία βέλτιστη λύση, αλλά ένα σύνολο βέλτιστων λύσεων Pareto.

Μια κλασική προσέγγιση για προβλήματα MOO είναι η μετατροπή ενός σε ένα πρόβλημα SOO με χρήση σταθμισμένου αθροίσματος. Η εξαγωγή μια λύσης από την επίλυση ενός SOO είναι μια μη κυριαρχούμενη λύση για το MOO [87] ωστόσο το κύριο μειονέκτημα μιας τέτοιας προσέγγισης είναι ότι η δεν μπορούν να επιτευχθούν πάντα όλες οι μη κυριαρχούμενες λύσεις. Ένα βασικό πλεονέκτημα που καθιστά τους ΕΑ ως εφαρμόσιμες μεθόδους για MOO είναι η χρήση πληθυσμού που επιτρέπει την εξαγωγή πολλαπλών μη κυριαρχούμενων λύσεων σε μία μόνο εκτέλεση [88]. Συνεπώς έχουν αναπτυχθεί πολλοί MOEA [13] όπως μέθοδοι βασισμένες στο Παρετο [89, 90, 91], μέθοδοι βασισμένες στην αποδόμηση (δεςομποσιτιον) [92] και μέθοδοι που χρησιμοποιούν δείκτες [93, 94]. Μια άλλη κατηγορία αφορά MOEA που αποσκοπούν στην εισαγωγή προτιμήσεων κατά την αναζήτηση. Μια αντιπροσωπευτική προσέγγιση είναι ο NSGA-III [95].

3.1.4 Εξειδικευμένοι τελεστές

Δύο περιπτώσεις που μπορούν να εντοπιστούν όταν οι ΕΑ εφαρμόζονται σε προβλήματα βελτιστοποίησης πραγματικού κόσμου. Η πρώτη αφορά την περίπτωση κατά την οποία ένας ΕΑ που αναπτύχθηκε για βελτιστοποίηση μαύρου κουτιού εφαρμόζεται απευθείας στο πρόβλημα του πραγματικού κόσμου. Ένας ΕΑ μπορεί να εφαρμοστεί ως διαδικασία άμεσης αναζήτησης και αυτό είναι ένα από τα πλεονεκτήματα των ΕΑ [88]. Η δεύτερη περίπτωση αφορά προβλήματα βελτιστοποίησης όπου υπάρχει διαθέσιμη πληροφορία για το πρόβλημα βελτιστοποίησης. Σε αυτήν την περίπτωση, οι ΕΑ μπορούν να εφαρμοστούν όπως στην προηγούμενη περίπτωση ή θα μπορούσε να γίνει προσπάθεια εκμετάλλευσης της διαθέσιμης πληροφορίας. Μια τεχνική για εκμετάλλευση της διαθέσιμης πληροφορίας είναι οι εξειδικευμένοι τελεστές [96, 97].

3.1.5 Υβριδικοί εξελικτικοί αλγόριθμοι

Υβριδικοί ΕΑ είναι ΕΑ που συνδυάζονται με τουλάχιστον μία διαφορετική μέθοδο [77]. Παραδείγματα υβριδοποίησης είναι ο συνδυασμός δύο διαφορετικών ΕΑς, ο συνδυασμός ενός ΕΑ και μιας κλασσικής μεθόδου βελτιστοποίησης που κάνει χρήση της παραγώγου, ή ο συνδυασμός ενός ΕΑ με εξειδικευμένους τελεστές. Συνήθως, ο υβριδισμός στοχεύει στη βελτίωση της απόδοσης του ΕΑ από τον συνδυασμό των διακριτών χαρακτηριστικών κάθε μεθόδου. Οι Μιμητικοί αλγόριθμοι (Memetic Algorithms, MA) περιλαμβάνουν προσεγγίσεις που βασίζονται σε ΕΑ στους οποίους εισάγονται τεχνικές τοπικής αναζήτησης [98].

3.2 Μονο- και πολύ-κριτηριακή βελτιστοποίηση με χρήση εξελικτικών αλγορίθμων υποβοηθούμενων από μεταπρότυπα

Για προβλήματα βελτιστοποίησης που περιλαμβάνουν υπολογιστικά δαπανηρές αντικειμενικές συναρτήσεις, η εκτέλεση μεγάλου αριθμού αξιολογήσεων μπορεί να καταστεί ανέφικτη ή υπολογιστικά δαπανηρή [14]. Σε αυτή τη περίπτωση γίνεται χρήση τεχνικών που βασίζονται σε προσεγγίσεις. Τρία επίπεδα προσέγγισης που αναφέρονται στη βιβλιογραφία είναι [99]:

- Προσέγγιση προβλήματος: Η αρχική διατύπωση του προβλήματος αντικαθίσταται από ένα μοντέλο που εισάγει απλοποιήσεις και είναι υπολογιστικά λιγότερο δαπανηρό.
- Προσέγγιση συνάρτησης: Μια αρχική συνάρτηση αντικαθίσταται από μια εναλλακτική έκφραση. Η έκφραση θα πρέπει να είναι υπολογιστικά λιγότερη δαπανηρή από την αρχική συνάρτηση.
- Εξελικτική προσέγγιση: Είναι τεχνικές που αφορούν συγκεκριμένα τους ΕΑ όπως η κληρονομιά καταλληλότητας και η απομίμηση καταλληλότητας [99].

3.2.1 Μοντέλα προσέγγισης: Μεταπρότυπα

Έχουν προταθεί διάφορες τεχνικές προσέγγισης για την αντικατάσταση των δαπανηρών προσομοιώσεων ή πειραμάτων από υπολογιστικά φθηνά ΑΜ [100]. Δύο τεχνικές που εξετάζονται σε αυτή τη Διατριβή είναι η PR και τα RBF και αυτές παρουσιάζονται αναλυτικότερα στο πλήρες κείμενο.

3.2.2 Συμπληρωματικά Κριτήρια

Η ακρίβεια της προσέγγισης ενός ΑΜ θα μπορούσε να βελτιωθεί με την εισαγωγή πρόσθετων δεδομένων που χρησιμοποιούνται για την εκπαίδευση των ΑΜ με βάση συμπληρωματικά κριτήρια [100, 101]. Γενικά επιθυμείται μια ισορροπία μεταξύ τοπικής και ολικής εκμετάλλευσης για την επιλογή επιπλέον δεδομένων [100]. Το πρώτο στοχεύει στην βελτίωση της ακρίβειας του ΑΜ σε μια περιοχή ενδιαφέροντος. Το δεύτερο στοχεύει στη βελτίωση ακρίβειας του ΑΜ σε όλο το χώρο αναζήτησης.

3.2.3 Εξελικτικοί αλγόριθμοι και διαχείριση μεταπροτύπων

Για προβλήματα βελτιστοποίησης που περιλαμβάνουν υπολογιστικά δαπανηρές προσομοιώσεις ο αριθμός προσομοιώσεις που μπορούν να γίνουν είναι περιορισμένος. Αυτό μπορεί να είναι αποτρεπτικό για την επιτυχή εφαρμογή των ΕΑ καθώς αυτοί, συνήθως, βασίζονται στην χρήση μεγάλου αριθμού αξιολογήσεων για να συγκλίνουν. Συνεπώς, οι ΕΑ έχουν συζευχθεί με ΑΜ για να επιτύχουν αποδεκτά αποτελέσματα σε υπολογιστικά δαπανηρά προβλήματα βελτιστοποίησης. Σε τέτοιες περιπτώσεις, τα μεταπρότυπα χρησιμοποιούνται για να αντικαταστήσουν εν μέρει το πραγματικό μοντέλο (True Model, TM) [102]. Τέτοιες προσεγγίσεις μπορούν να χωριστούν σε τρεις κατηγορίες [14, 15, 99, 102, 103]: (i) Χωρίς Έλεγχο Εξέλιξης, (ii) Σταθερός Έλεγχος Εξέλιξης και (iii) Προσαρμοστικός Έλεγχος Εξέλιξης.

Στο πλαίσιο της δειγματοληψίας κατά την διάρκεια της εξέλιξης του πληθυσμού ενός ΕΑ, η διαχείριση μοντέλων ακολουθεί συνήθως είτε το πλαίσιο Έλεγχος Εξέλιξης (Evolution Control) είτε το πλαίσιο Προεπιλογής (Pre-selection). Ο Έλεγχος Εξέλιξης, [102] παρέχει δύο κύριες μεθόδους για να προσδιοριστεί εάν η καταλληλότητα ενός ατόμου αξιολογείται χρησιμοποιώντας το ΑΜ ή το TM:

- Ελεγχόμενα άτομα: Ένας αριθμός ατόμων σε κάθε γενιά αξιολογείται χρησιμοποιώντας το TM. Το ΑΜ χρησιμοποιείται για τα υπόλοιπα άτομα.
- Ελεγχόμενες γενιές: Όλα τα άτομα αξιολογούνται χρησιμοποιώντας το TM σε μερικές γενιές. Το ΑΜ χρησιμοποιείται για όλα τα άτομα στις υπόλοιπες γενιές.

Μια τρίτη κατηγορία εφαρμόζει διαχείριση μοντέλων σε επίπεδο πληθυσμού για την περίπτωση όπου γίνεται χρήση πέραν του ενός υποπληθυσμού που εξελίσσονται με βάση διαφορετικό ΑΜ και πιθανώς διαφορετικά επίπεδα ακρίβειας [104]. Επιπλέον, το ΑΜ μπορεί να συμβάλει σε οποιονδήποτε τελεστή ενός ΕΑ [105]. Στο πλαίσιο Προεπιλογής [106, 107] γίνεται χρήση των ΑΜ σε κάθε γενιά για να επιλέξει απογόνους που θα αξιολογηθούν χρησιμοποιώντας το TM και η επιλογή του πληθυσμού για την επόμενη γενιά γίνεται πάντα μεταξύ ατόμων που έχουν αξιολογηθεί με βάση το TM.

3.2.4 Επιλογή μοντέλου προσέγγισης

Η επιλογή ενός ΑΜ μπορεί να έχει αντίκτυπο στην απόδοση του ΕΑ [108, 109, 110]. Παράγοντες όπως ο αριθμός των μεταβλητών απόφασης, οι συναρτήσεις στόχου, περιορισμοί και ο υπολογιστικός χρόνος για την κατασκευή των μοντέλων, πρέπει να λαμβάνονται υπόψη κατά την επιλογή ενός ΑΜ.

3.2.5 Ολικό και τοπικό μοντέλο προσέγγισης

Έχουν εξεταστεί διάφορες τεχνικές για την επιλογή των σημείων που έχουν αξιολογηθεί χρησιμοποιώντας το TM και θα χρησιμοποιηθούν για την κατασκευή του ΑΜ. Για παράδειγμα, μπορούν να χρησιμοποιηθούν όλα τα διαθέσιμα σημεία εφόσον κάθε ένα από αυτά μπορεί να παρέχει σημαντικές πληροφορίες σχετικά με το χώρο αναζήτησης [110]. Ωστόσο, το υπολογιστικό κόστος κατασκευής του ΑΜ θα μπορούσε να αυξηθεί για μεγάλο αριθμό δεδομένων. Επιπλέον, η ακρίβεια του ΑΜ εξαρτάται από τα επιλεγμένα δεδομένα που χρησιμοποιούνται για την κατασκευή του μοντέλου και την ικανότητά τους να παρέχουν μια ικανοποιητική προσέγγιση ολόκληρου του χώρου αναζήτησης [111, 112]. Συνεπώς, έχουν προταθεί ΜΑΕΑ που κάνουν χρήση τοπικών ΑΜ.

3.2.6 Κατάρα και ευλογία της αβεβαιότητας του μοντέλου προσέγγισης

Το σφάλμα προσέγγισης του ΑΜ μπορεί να αποτρέψει τον εντοπισμό βέλτιστης λύσης από ένα ΕΑ. Όμως μπορεί να έχει και θετικό αντίκτυπο σε ορισμένες περιπτώσεις. Ένα τέτοιο παράδειγμα είναι συναρτήσεις πολλών ακρότατων όπου ένα ΑΜ θα μπορούσε να οδηγήσει σε εξομάλυνση της αντικειμενικής συνάρτησης. Οι αρνητικές και θετικές επιπτώσεις αυτού του σφάλματος προσέγγισης στην αναζήτηση που βασίζεται σε ΕΑ αναφέρονται συχνά ως κατάρα της αβεβαιότητας και ευλογία της αβεβαιότητας [111, 113].

3.2.7 Αντιπροσωπευτικές προσεγγίσεις

Έχουν παρουσιαστεί αρκετές προσεγγίσεις στην βιβλιογραφία με βάση τους ΕΑ που βασίζονται σε μεταπρότυπα για SOO [14, 102, 106, 114], για SOO υπό περιορισμούς [115, 116, 117] και προσεγγίσεις που περιλαμβάνουν τεχνικές τοπικής αναζήτησης [112, 118, 119]. Επίσης, έχουν παρουσιαστεί προσεγγίσεις για υπολογιστικά δαπανηρά SOO που δεν βασίζονται σε ΕΑ [120, 121, 122, 123, 124, 125]. Αντίστοιχα, έχουν παρουσιαστεί αρκετές προσεγγίσεις στην βιβλιογραφία με βάση τους ΕΑ που βασίζονται σε μεταπρότυπα για MOO [126, 127, 128], για MOO υπό περιορισμούς [129, 130] και προσεγγίσεις που περιλαμβάνουν τεχνικές τοπικής αναζήτησης [111, 131, 132, 133]. Οι αντιπροσωπευτικές προσεγγίσεις αυτές παρουσιάζονται αναλυτικότερα στο πλήρες κείμενο της Διατριβής.

3.3 Συζήτηση

Στην βιβλιογραφία εντοπίζονται αποδοτικοί ΕΑ και ΜΑΕΑ για προβλήματα βελτιστοποίησης μαύρου κουτιού. Επιπλέον, αυτοί παρουσιάζουν ικανοποιητικά αποτελέσματα σε πολλά προβλήματα βελτιστοποίησης όπως SOO, COP και MOO. Επίσης, έχουν παρουσιαστεί στην βιβλιογραφία εφαρμογές σε προβλήματα βελτιστοποίησης που αφορούν το GEP [1, 62]. Οι ΜΑΕΑ που παρουσιάζονται στα επόμενα κεφάλαια βασίζονται στους ΕΑ και ΜΑΕΑ και εμπεριέχουν τροποποιήσεις που πηγάζουν από το πρόβλημα βελτιστοποίησης στο οποίο εφαρμόστηκαν.

Κεφάλαιο 4

Σχεδιασμός επέκτασης της δυναμικότητας παραγωγής ισχύος με χρήση εξελικτικών αλγορίθμων υποβοηθούμενων από μεταπρότυπα

4.1 Κίνητρο και στόχος

Το Κεφάλαιο αυτό παρουσιάζει ένα μοντέλο GEP πολλών χρονικών περιόδων (multi-period) βασισμένο σε EA υποβοηθούμενους από μεταπρότυπα [17]. Βασικό κίνητρο για την εξέταση ενός τέτοιου μοντέλου ήταν οι αυξανόμενες υπολογιστικές απαιτήσεις για ένα μοντέλο GEP. Αυτές σχετίζονται με την εισαγωγή SM που παρουσιάζουν σχετικά αυξημένη λεπτομέρεια της βραχυπρόθεσμης λειτουργία ενός συστήματος ισχύος, στο πλαίσιο του μακροπρόθεσμου προγραμματισμού. Η απαίτηση για αύξηση στο επίπεδο λεπτομέρειας πηγάζει από την ανάγκη εξέτασης της λειτουργικής ευελιξίας για τον καθορισμό ενός βέλτιστου πλάνου επέκτασης παραγωγής όταν λαμβάνεται υπόψη η αύξηση του μεριδίου των RES.

Συγκεκριμένα, το μοντέλο στοχεύει στο προσδιορισμό ενός σχεδίου επέκτασης παραγωγής λαμβάνοντας υπόψη τις απαιτήσεις σε λειτουργική ευελιξία μέσω ενός SM του βραχυπρόθεσμου προγραμματισμού. Το επιλεγμένο SM χαρακτηρίζεται από αυξημένη χρονική και τεχνική λεπτομέρεια. Συνεπώς, γίνονται παραδοχές όσο αφορά τη χωρική λεπτομέρεια η οποία δεν εξετάζεται. Επίσης, εξετάζονται πάροχοι λειτουργικής ευελιξίας που δεν είναι θερμικές μονάδες όπως υδροηλεκτρικές και αποθηκευτικές μονάδες. Για την αντιμετώπιση της αύξησης του υπολογιστικού κόστους του SM χρησιμοποιείται ομαδοποίηση μονάδων (unit aggregation) και χρήση αντιπροσωπευτικών χρονικών περιόδων.

Αλγόριθμοι βασισμένοι σε EA έχουν εφαρμοστεί επιτυχώς σε μοντέλα GEP. Παρ΄ όλα αυτά, τέτοιες εφαρμογές δεν εντοπίστηκαν σε GEP μοντέλα που ενσωματώνουν SM που εμπεριέχουν περιορισμούς του βραχυπρόθεσμου προγραμματισμού (UCP) παραγωγής ηλεκτρικής ενέργειας.

Ο κύριος στόχος ήταν να αναπτυχθεί μια αποτελεσματική εναλλακτική για εντοπισμό βέλτιστων, ή έστω βελτιωμένων, λύσεων που θα μπορούσε να χρησιμοποιηθεί παράλληλα με καθιερωμένα μοντέλα GEP ως εργαλείο υποστήριξης αποφάσεων. Ο τροποποιημένος MAEA που αναπτύχθηκε βασίζεται στη βελτιστοποίηση με EA υποβοηθούμενους από μεταπρότυπα και στη DE. Επίσης, εισάγονται τροποποιήσεις σε μορφή τελεστών βασιζόμενοι σε συγκεκριμένα χαρακτηριστικά του προβλήματος βελτιστοποίησης. Τα AM χρησιμοποιούνται για να μειώσουν τον αριθμό των προσομοιώσεων που απαιτούνται για τον εντοπισμό μιας βελτιωμένης λύσης.

Εξετάζεται η απόδοση του αλγορίθμου και των τροποποιήσεων που εξετάστηκαν μέσα από υπολογιστικά πειράματα. Ταυτόχρονα, αναλύεται για την επίδραση της εισαγωγής του SM με βάση τη τελική εγκατεστημένη ισχύ, τα αναμενόμενα κόστη και το μείγμα παραγωγής.

4.2 Μαθηματική διατύπωση προβλήματος

Σε αυτήν την ενότητα παρουσιάζεται η διατύπωση του προβλήματος, των συντελεστών κόστους που χρησιμοποιούνται για τον προσδιορισμό των συναρτήσεων στόχων και των περιορισμών του μονοκριτηριακού προβλήματος βελτιστοποίησης. Επίσης, παρουσιάζεται το μοντέλο προσομοίωσης που χρησιμοποιείται για την παροχή μιας εκτίμησης της βραχυπρόθεσμης λειτουργίας του συστήματος ισχύος.

4.2.1 Μαθηματική διατύπωση προβλήματος βελτιστοποίησης

Εξετάζεται το πιο κάτω μονοκριτηριακό πρόβλημα βελτιστοποίησης:

$$\begin{aligned} &minimize \;\; f(\mathbf{x}) = f^{chp}(\mathbf{x}) + \sum_{yr} f^{xp}_{yr}(\mathbf{x}) \qquad (4.1)\\ &s.t. \;\; \mathbf{G}(\mathbf{x}) \leq 0 \\ &\qquad \mathbf{x} \in \mathbb{S} \end{aligned}$$

όπου $\mathbf{x} = (x_1, x_2, ..., x_n)$ είναι το διάνυσμα μεταβλητών απόφασης/σχεδιασμού, n είναι ο αριθμός μεταβλητών απόφασης, $f(\mathbf{x})$ είναι η αντικειμενική συνάρτηση, $\mathbf{G}(\mathbf{x}) = (g_1(\mathbf{x}), g_2(\mathbf{x}), ..., g_k(\mathbf{x}))$ είναι το διάνυσμα συναρτήσεων περιορισμού, k είναι ο αριθμός συναρτήσεων περιορισμού και $\mathbb{S}$ είναι ο χώρος αναζήτησης.

Συγκεκριμένα, εξετάζεται η περίπτωση μιας συνάρτησης στόχου με αυξημένο υπολογιστικό κόστος, συναρτήσεων περιορισμού χαμηλού υπολογιστικού κόστους ($\mathbf{G}^{chp}(\mathbf{x})$), και περιορισμούς στα άνω και κάτω όρια των μεταβλητών απόφασης. Επίσης, εξετάζεται η περίπτωση ακέραιων μεταβλητών σχεδιασμού ($\mathbb{Z}^n$). Θεωρείται ότι το πρόβλημα βελτιστοποίησης περιορίζεται από έναν προκαθορισμένο άνω όριο σε υπολογιστικό κόστος.

Οι μεταβλητών απόφασης είναι θετικοί ακέραιοι ($\mathbb{Z}^+$) που αντιπροσωπεύουν τον αριθμό των επενδύσεων σε υποψήφιες τεχνολογίες παραγωγής ενέργειας. Το ύψος κάθε βήματος επένδυσης είναι προκαθορισμένο και αντιπροσωπεύει μια μονάδα παραγωγής ισχύος ή μια ομάδα μονάδων παραγωγής ισχύος.

Η αντικειμενική συνάρτησης στόχου ($f(\mathbf{x})$) αποτελείται από συντελεστές κόστους με αυξημένο ή αμελητέο υπολογιστικό κόστος. Κάθε συντελεστής κόστους μπορεί να θεωρηθεί ως συνάρτηση του διανύσματος μεταβλητών απόφασης. Το άθροισμα των συντελεστών κόστους που δεν αποτελούν μέρος του SM (κόστος επένδυσης, σταθερά λειτουργικά κόστη και κόστη συντήρησης) θεωρείται ως μια υπολογιστικά αμελητέου κόστους συνάρτηση ($f^{chp}(\mathbf{x})$). Αντίστοιχα, οι συντελεστές κόστους που υπολογίζονται με τη χρήση του SM (π.χ. ετήσια μεταβλητά κόστη) θεωρούνται ως υπολογιστικά ακριβές συναρτήσεις ($f^{xp}(\mathbf{x})$). Διευκρινίζεται ότι το υπολογιστικό κόστος μιας συνάρτησης δεν υποδηλώνει την επίδρασή της στην τιμή της συνάρτησης στόχου.

4.2.2 Συντελεστές κόστους

Οι συντελεστές κόστους ορίστηκαν σαν διανύσματα όπου κάθε στοιχείο υποδηλώνει την τιμή σε ένα έτος του υπό εξέταση χρονικού ορίζοντα. Πιο κάτω παρουσιάζονται επιγραμματικά οι συντελεστές κόστους που εξετάστηκαν:

1. *Κόστος επένδυσης*:

$$c_{yr}^{inv} = \sum_{t}[(1 - SF_{yr,t}) \cdot P_t^{cap\_step} \cdot INC_t \cdot x_{yr,t}^{inv}], \forall yr \tag{4.2}$$

όπου:

$$SF_{yr,t} = \begin{cases} 0, & \text{if } \; yr + CT_t + LT_t <= Yrz \\ 1, & \text{elseif } \; yr + CT_t > Yrz \\ \frac{-1+(1+DR)^{(yr+CT_t+LT_t-Yrz-1)}}{-1+(1+DR)^{(CT_t+LT_t)}}, & \text{otherwise} \end{cases} , \forall yr, t \tag{4.3}$$

2. *Σταθερά λειτουργικά κόστη και κόστη συντήρησης*:

$$c_{yr}^{fixed} = \sum_{t}[FOM_t \cdot ic_{yr,t}], \; \forall yr \tag{4.4}$$

3. *Μεταβλητά κόστη*:

$$c_{yr}^{var} = \sum_{tp}[W_{tp} c_{tp}^{oc}], \forall yr \tag{4.5}$$

Τα λειτουργικά κόστη υπολογίζονται με χρήση του SM το οποίο παρουσιάζεται στην ενότητα 4.2.3.

Επίσης, η εγκατεστημένη δυναμικότητα παραγωγής ισχύος και ο αριθμός αντιπροσωπευτικών μονάδων υπολογίζονται ως εξής:

$$ic_{yr,t} = nu_{yr,t} \cdot P_t^{cap\_step}, \; \forall yr, t \tag{4.6}$$

όπου

$$nu_{yr,t} = NU_{yr,t}^{old} + \begin{cases} \sum_{z=yr-CT_t-LT_t}^{yr-CT_t}[x_{yr,t}^{inv}], & \text{if } \; yr - CT_t > 0, yr - CT_t - LT_t \geq 1 \\ \sum_{z=1}^{yr-CT_t}[x_{yr,t}^{inv}], & \text{elseif } \; yr - CT_t > 0, yr - CT_t - LT_t < 1 \\ 0, & \text{otherwise} \end{cases} , \forall yr, t \tag{4.7}$$

4.2.3 Μοντέλο προσομοίωσης - Προσέγγιση προβλήματος

Το SM χρησιμοποιείται για να προσδιορίσει τα μεταβλητά κόστη σε κάθε έτος του χρονικού ορίζοντα. Εναλλακτικά, η διατύπωση ενός ενοποιημένου προβλήματος βελτιστοποίησης (GEP και UCP) ως ένα ενιαίο MILP είναι εφικτή αλλά μπορεί να είναι υπολογιστικά ακριβή [7, 8].

Για την προσομοίωση της βραχυπρόθεσμης λειτουργίας ενός συστήματος ισχύος επιλέγηκε το μοντέλο Clustered UC (CUC) [7, 8, 134]. Αυτό αποτελεί προσέγγιση του προβλήματος UCP. Εισάγει απλοποιήσεις στη διατύπωση του προβλήματος ώστε να μειώσει το υπολογιστικό κόστος της προσομοίωσης. Η βασική απλοποίηση που εισάγεται αφορά τη χρήση ακεραίων μεταβλητών για τον προσδιορισμών του επιπέδου λειτουργίας ομάδων μονάδων παραγωγής ισχύος αντί για χρήση δυαδικών μεταβλητών για καθορισμό του επιπέδου λειτουργίας κάθε μονάδας. Αυτό επιτυγχάνεται με ομαδοποίηση μονάδων με βάση τα τεχνικά τους χαρακτηριστικά. Το CUC παρά την εισαγωγή απλοποιήσεων αποτελεί ικανοποιητική προσέγγιση της βραχυπρόθεσμης λειτουργίας του συστήματος ισχύος όσο αφορά το πλαίσιο του μακροπρόθεσμου προγραμματισμού [7, 8, 134].

Το επιλεγμένο SM διατυπώθηκε με βάση το CUC. Συγκεκριμένα, το πρόβλημα βελτιστοποίησης είναι ένα MILP που αντιπροσωπεύει ένα απλοποιημένο UCP με θερμικές και υδροηλεκτρικές μονάδες παραγωγής ισχύος. Επίσης, λαμβάνεται υπόψη η παραγωγή από RES και αποθηκευτικές μονάδες βασισμένες στην αντλησιοταμίευση (Hydro-Storage, HS). Το προαναφερθέν πρόβλημα επιλύεται ανεξάρτητα για ένα προκαθορισμένο αριθμό χρονικών περιόδων για κάθε έτος για το οποίο απαιτείται η χρήση του SM. Η διατύπωση του προβλήματος βελτιστοποίησης παρουσιάζεται αναλυτικά στο εδάφιο **;;**.

Επίσης, οι μεταβλητές σχεδιασμού του SM δεν αποτελούν μεταβλητές απόφασης του μακροπρόθεσμου προγραμματισμού καθώς το SM επιλύεται ανεξάρτητα. Συνεπώς, ο αριθμός των μεταβλητών απόφασης του μακροπρόθεσμου προγραμματισμού επηρεάζεται μόνο από τον αριθμό των διάφορων επενδυτικών επιλογών και τα καθορισμένα στάδια που μπορούν να υλοποιηθούν αυτές. Επίσης, θα μπορούσαν να εξεταστούν διαφορετικά SM καθώς η επιλογή ενός SM δεν επηρεάζει άμεσα τη μαθηματική διατύπωση του προβλήματος του μακροπρόθεσμου προγραμματισμού.

Χρονική λεπτομέρεια

Η χρονική λεπτομέρεια που εισάγεται στο μοντέλο διαχωρίζεται σε τρία επίπεδα. Το πρώτο επίπεδο αφορά τα έτη στόχοι ($\mathbb{Y}^{trg} \subseteq \mathbb{Y}$) για τα οποία γίνεται χρήση του SM για προσδιορισμό του λειτουργικού κόστους. Το δεύτερο επίπεδο αφορά τον αριθμό χρονικών περιόδων ($\mathbb{P}$) εντός ενός έτους στόχος για τα οποία επιλύονται ανεξάρτητα προβλήματα βελτιστοποίησης. Το τρίτο επίπεδο αφορά τον αριθμό ωριαίων διαστημάτων ($\mathbb{H}$) που εμπεριέχει κάθε χρονική περίοδος. Αύξηση κάθε επιπέδου χρονικής λεπτομέρειας μπορεί να οδηγήσει σε αύξηση της ακρίβειας του μοντέλου αλλά και αύξηση του υπολογιστικού κόστους.

4.2.4 Διατύπωση μονοκριτηριακού προβλήματος βελτιστοποίησης

Στο εδάφιο αυτό παρουσιάζεται η συνάρτηση στόχος και οι συναρτήσεις περιορισμού του προβλήματος.

Αντικειμενική συνάρτηση

Η αντικειμενική συνάρτηση ορίστηκε ως το άθροισμα του ετήσιου κόστους επένδυσης, σταθερού κόστους και μεταβλητού κόστους λαμβάνοντας υπόψη και τον συντελεστή αναγωγής.

$$TC = \sum_{yr} [DD_{yr} \cdot (c_{yr}^{inv} + c_{yr}^{fixed} + c_{yr}^{var})] \tag{4.8}$$

όπου:

$$DD_{yr} = 1/(1 + DR)^{yr}, \forall yr \tag{4.9}$$

Συναρτήσεις περιορισμών

1. *Θετικές ακέραιες μεταβλητές*:

$$\boldsymbol{x}^{inv} \in \mathbb{N}^{n^{inv}} \tag{4.10}$$

2. *Μέγιστος και ελάχιστος αριθμός ετήσιων επενδύσεων*:

$$0 \leq x_{yr,t}^{inv} \leq \overline{X}_{yr,t}^{inv},\ \forall yr, t \tag{4.11}$$

Το άνω όριο του αριθμού των επενδύσεων λαμβάνει υπόψη:

(α΄) Μια ρεαλιστική εκτίμηση του άνω ορίου των επενδύσεων σε κάθε επενδυτικό στάδιο:

$$\overline{X}_{yr,t}^{inv} \leq X_t^{max\_inv}, \forall yr, t \tag{4.12}$$

(β΄) Το χρόνο κατασκευής των μονάδων έτσι ώστε αυτές να μπορούν να είναι λειτουργικές εντός του υπό εξέταση χρονικού ορίζοντα:

$$\overline{X}_{yr,t}^{inv} = 0,\ \forall yr + CT_t > Yrz, t \tag{4.13}$$

(γ΄) Τις μονάδες που θεωρούνται ως επενδυτικές επιλογές:

$$\overline{X}_{yr,t^{old}}^{inv} = 0,\ \forall yr, t^{old} \tag{4.14}$$

(δ΄) Τα έτη που έχουν οριστεί ως επενδυτικά στάδια:

$$\overline{X}_{yr,t}^{inv} = \begin{cases} \overline{X}_{yr,t}^{inv}, & \text{if } yr \in \mathbb{Y}^{InvSt} \\ 0, & \text{otherwise} \end{cases}, \forall yr, t \tag{4.15}$$

3. *Περιθώριο ασφαλείας σχεδιασμού*:

$$\sum_t ic_{yr,t} \cdot CC_t \geq RS_{yr},\ \forall yr \tag{4.16}$$

4. *Μέγιστη εγκατεστημένη δυναμικότητα παραγωγής ισχύος κάθε τεχνολογίας*:

$$ic_{yr,t} \leq \overline{IC}_{yr,t},\ \forall yr, t \tag{4.17}$$

όπου

$$\overline{IC}_{yr,t} = \min(\overline{IC}_{yr,t}^{ic}, (\overline{IC}_{yr,t}^{g} TD_{yr})/(8760 AV_t), IC_t^{ep}), \forall yr, t \tag{4.18}$$

5. *Ελάχιστη εγκατεστημένη δυναμικότητα παραγωγής ισχύος κάθε τεχνολογίας*:

$$ic_{yr,t} \geq \underline{IC}_{yr,t}\ \forall yr, t \tag{4.19}$$

όπου

$$\underline{IC}_{yr,t} = \max(\underline{IC}_{yr,t}^{ic}, (\underline{IC}_{yr,t}^{g} TD_{yr})/(8760 AV_t), 0), \forall yr, t \tag{4.20}$$

4.3 Προσέγγιση μονοκριτηριακής βελτιστοποίησης

Σε αυτό το εδάφιο παρουσιάζεται η μέθοδος που χρησιμοποιήθηκε για τη μονοκριτηριακή βελτιστοποίηση του προβλήματος. Αυτή βασίζεται στη DE και σε πλαίσια για βελτιστοποίηση με χρήση μεταπροτύπων χωρίς χρήση της παραγώγου. Επίσης παρουσιάζονται τροποποιήσεις στη μορφή εξειδικευμένων τελεστών.

4.3.1 Αλγόριθμος Διαφορικής Εξέλιξης

Βασικοί τελεστές του αλγορίθμου Διαφορικής Εξέλιξης

Ο βασικός αλγόριθμος DE αρχικοποιεί ένα αριθμό (NP) από υποψήφιες λύσεις στο χώρο αναζήτησης. Οι τελεστές μετάλλαξης, διασταύρωσης και επιλογής χρησιμοποιούνται για να βελτιώσουν τον πληθυσμό λύσεων σε βάθος γενιών (επαναλήψεων). Για κάθε άτομο του πληθυσμού παράγεται ένας απόγονος μέσω των τελεστών μετάλλαξης και διασταύρωσης. Ο τελεστής επιλογής καθορίζει αν ο απόγονος θα αντικαταστήσει τον γονέα στο πληθυσμό. Ως μέθοδος διαχείρισης περιορισμών εξετάστηκε η τεχνική Κανόνες Εφικτότητας (Feasibility rules) [135]. Οι προαναφερθέντες τελεστές ενός βασικού DE αλγορίθμου [136, 137] περιγράφονται αναλυτικά στο αντίστοιχο εδάφιο του πλήρους κειμένου.

Αλγόριθμος Διαφορικής Εξέλιξης για προβλήματα βελτιστοποίησης με ακέραιες μεταβλητές απόφασης

Ο βασικός αλγόριθμος DE συνήθως τροποποιείται όταν αυτός εφαρμόζεται σε προβλήματα βελτιστοποίησης που περιέχουν μεταβλητές απόφασης που περιορίζονται σε ακεραίους [136]. Παραλλαγές του αλγορίθμου DE έχουν εφαρμοστεί σε διάφορα τέτοια προβλήματα βελτιστοποίησης [138, 139, 140].

Η τεχνική που επιλέχθηκε βασίζεται στο χειρισμό των ακέραιων μεταβλητών ως συνεχείς και τη χρήση συνάρτησης στρογγυλοποίησης για τη μετατροπή τους σε ακεραίους [141]. Αυτό υλοποιείται μόνο για τη παραγωγή εισόδου για τον υπολογισμό των συναρτήσεων στόχου και περιορισμών.

Μηχανισμός επανεκκίνησης

Έχουν χρησιμοποιηθεί διάφοροι μηχανισμοί επανεκκίνησης μαζί με τη DE για την αντιμετώπιση της περίπτωσης πρόωρης σύγκλισης [137]. Χρησιμοποιήθηκε ένας τέτοιος μηχανισμός που αρχικοποιεί εκ νέου τον πληθυσμό όταν δεν υπάρξει βελτίωση σε κάποιο άτομο του πληθυσμού για ένα αριθμό επαναλήψεων (gen^{rst}). Αυτό δεν εφαρμόζεται στο καλύτερο άτομο του πληθυσμού που διατηρείται εντός του πληθυσμού. Ένας τέτοιος μηχανισμός μπορεί να εισάγει καινούργια πληροφορία στο πληθυσμό αλλά δεν μπορεί να εγγυηθεί ότι ο αλγόριθμος θα καταφέρει να διαφύγει από ένα τοπικό βέλτιστο.

4.3.2 Εξειδικευμένοι τελεστές

Ο αλγόριθμος DE χρησιμοποιήθηκε σαν αλγόριθμος βάσης που καθοδηγεί την αναζήτηση. Τρεις εξειδικευμένοι τελεστές εξετάστηκαν για να βελτιώσουν την απόδοσή του. Συγκεκριμένα:

- Ευρετική επιδιόρθωση με τυχαιότητα (*Randomized Repair Heuristic*, RRH): Ο τελεστής επιδιορθώνει λύσεις που παραβιάζουν τις συναρτήσεις περιορισμών του μακροπρόθεσμου προγραμματισμού. Η επιδιόρθωση βασίζεται σε μια αλληλουχία βημάτων που ορίστηκε με βάση χαρακτηριστικά του προβλήματος. Επίσης, εισάγεται τυχαιότητα στη διαδικασία επιδιόρθωσης ενός ατόμου έτσι ώστε να διασφαλιστεί η ανακατασκευή ποικίλων λύσεων που δεν παραβιάζουν τις συναρτήσεις περιορισμών του μακροπρόθεσμου προγραμματισμού. Συνεπώς, ο τελεστής αυτός επικεντρώνεται στην αντιμετώπιση του προβλήματος βελτιστοποίησης με έμφαση στις συναρτήσεις περιορισμών.

- Τελεστής διατάραξης (*Perturbation Operator*, PO): Τελεστές διατάραξης εφαρμόζονται συχνά σε προβλήματα διακριτής (συνδυαστικής) βελτιστοποίησης. Εξετάστηκε ένας τελεστής που βασίζεται σε μηχανισμούς που παράγουν τέτοιες διαταράξεις/μεταβολές στα διανύσματα των ατόμων του πληθυσμού. Η χρήση του στοχεύει στην δημιουργία ποικιλίας απογόνων και στην αντιμετώπιση προκλήσεων που προκύπτουν από το χειρισμό διακριτών μεταβολών.

- Τελεστής ομαδοποίησης με βάση τεχνολογικά κριτήρια (*Technology-group operator*): Ο τελεστής καθορίζει την αλληλουχία της αναπαράστασης των μεταβλητών απόφασης για ένα άτομο και εισάγει τροποποιήσεις στην εφαρμογή του τελεστή διασταύρωσης (blk). Αυτό βασίστηκε στην πιθανή ύπαρξη συνδέσεων μεταξύ μεταβλητών απόφασης. Έχει αναφερθεί ότι τέτοιες συνδέσεις όταν υπάρχουν μπορούν να επηρεάσουν την απόδοση των τελεστών διασταύρωσης [142].

4.3.3 Βήματα της μεθόδου βελτιστοποίησης

Στο εδάφιο αυτό αναφέρονται τα βήματα της μεθόδου βελτιστοποίησης που χρησιμοποιήθηκε. Αυτή βασίζεται σε αλγόριθμους [112, 124] για βελτιστοποίηση χωρίς χρήση παραγώγου και με χρήση μεταπροτύπων. Οι τροποποιήσεις που εισάγονται βασίζονται σε χαρακτηριστικά του προβλήματος. Συγκεκριμένα:

- Χρησιμοποιούνται ένας αριθμός από ΑΜ όπου κάθε ένα εκπαιδεύεται για ένα υποσύνολο των μεταβλητών αντί για ένα ΑΜ το οποίο εκπαιδεύεται με βάση το σύνολο των μεταβλητών απόφασης. Αυτό γίνεται ώστε κάθε ΑΜ να παρέχει την εκτίμηση του μεταβλητού κόστους για ένα έτος υποθέτοντας ότι το μεταβλητό κόστος κάθε έτους μπορεί να υπολογιστεί ανεξάρτητα. Αυτή η υπόθεση γίνεται εφόσον το πρόβλημα βελτιστοποίησης δεν είναι μαύρου κουτιού (black-box) και βασίζεται στο ότι ο αριθμός των εγκατεστημένων μονάδων είναι συνάρτηση των μεταβλητών απόφασης. Συνεπώς, ο αριθμός των διαστάσεων κάθε ΑΜ αυξάνει με τον αριθμό τεχνολογιών παραγωγής ισχύος που μπορούν να έχουν διαφορετικές τιμές εγκατεστημένης δυναμικότητας παραγωγής ισχύος σε κάθε έτος. Αντίστοιχα, ο αριθμός των ΑΜ που χρησιμοποιούνται αυξάνει με τον αριθμό των ετών στόχος.

- Οι συναρτήσεις περιορισμού που αφορούν το μακροπρόθεσμο προγραμματισμό επιδιορθώνονται με τη χρήση του RRH. Αλγόριθμοι για βελτιστοποίηση τύπου μαύρου κουτιού χωρίς χρήση παραγώγου και με χρήση μεταπροτύπων συνδυάζονται με τεχνικές χειρισμού περιορισμών για προβλήματα που εμπεριέχουν περιορισμούς καθώς τελεστές επιδιόρθωσης μη-εφικτών λύσεων δεν είναι πάντα διαθέσιμοι.

- Τα ΑΜ ανανεώνονται με εισαγωγή καινούργιων δεδομένων από λύσεις επιλεγμένες σε διαφορετικά στάδια από ένα πληθυσμό DE και από στοχαστικά παραγόμενες λύσεις. Η καμπύλη εξέλιξης της εγκατεστημένης δυναμικότητας παραγωγής ισχύος κάθε τεχνολογίας λαμβάνεται υπόψη σε αυτές τις διαδικασίες επιλογής.

Η μέθοδος βελτιστοποίησης αποτελείται από δυο φάσεις. Η πρώτη αφορά την αρχικοποίηση και η δεύτερη αφορά την αναζήτηση βέλτιστη λύσης του προβλήματος. Τα βήματα αυτά παρουσιάζονται αναλυτικά στο πλήρες κείμενο της Διατριβής. Τα βασικά στάδια της πρώτης φάσης είναι τα ακόλουθα:

- Προσδιορισμός περιοχής ενδιαφέροντος με βάση τους περιορισμούς του προβλήματος.

- Αρχικοποίηση των εξωτερικών αρχείων (αρχείο λύσεων και αρχεία προσομοιώσεων) όπου αποθηκεύονται εντοπισμένες λύσεις και τα αποτελέσματα των υπολογιστικά δαπανηρών προσομοιώσεων.

- Δημιουργία αρχικών σημείων για τα αρχεία προσομοιώσεων. Για τα αρχικά σημεία αυτά εφαρμόζονται υπολογιστικά δαπανηρές προσομοιώσεις. Εκπαίδευση των ΑΜ με βάση τα αρχικά σημεία αυτά.

- Εισαγωγή αρχικών λύσεων, όταν αυτές είναι διαθέσιμες, και παραγωγή αρχικών λύσεων με χρήση του DE αλγορίθμου που αναπτύχθηκε. Υπολογιστικά δαπανηρές προσομοιώσεις εφαρμόζονται για κάθε νέο συνδυασμό εγκατεστημένη ισχύς που προκύπτει από μια τέτοια λύση και το αποτέλεσμα αυτών αποθηκεύεται στα αρχεία. Στη συνέχεια τα ΑΜ ανανεώνονται με βάση τα νέα αρχεία. Επίσης, η καλύτερη λύση που έχει εντοπιστεί ορίζεται ως αυτή που εμφανίζει την ελάχιστη τιμή στο αρχείο λύσεων η οποία ανανεώνεται κάθε φορά που εισέρχεται σε αυτό μια νέα λύση.

Τα βασικά στάδια της δεύτερης φάσης είναι τα ακόλουθα:

- Εφαρμόζεται ο DE αλγόριθμος για προκαθορισμένο αριθμό γενιών στο συνολικό πρόβλημα ελαχιστοποίησης με βάση την εκτίμηση των ΑΜ. Η καλύτερη λύση που εντοπίζεται αποθηκεύεται στο αρχείο λύσεων αν δεν έχει αποθηκευτεί. Υπολογιστικά δαπανηρές προσομοιώσεις εφαρμόζονται για κάθε νέο συνδυασμό εγκατεστημένη ισχύς που προκύπτει από μια νέα λύση. Στη συνέχεια ανανεώνονται τα αρχεία, τα ΑΜ και η καλύτερη λύση που έχει εντοπιστεί. Αυτό επαναλαμβάνεται σειριακά για κάθε ένα διαθέσιμο είδος ΑΜ.

- Παραγωγή υποψήφιων λύσεων μεταβάλλοντας την καλύτερη λύση που έχει βρεθεί και παράγοντας στοχαστικά υποψήφιες λύσεις στο χώρο αναζήτησης [122, 123, 124, 125, 130, 143]. Οι υποψήφιες λύσεις στην συνέχεια επιδιορθώνονται με χρήση του RRH μεταβάλλοντας τα όρια της περιοχής ενδιαφέροντος με βάση την εγκατεστημένη ισχύ της καλύτερης λύσης που έχει εντοπιστεί. Αυτό αποσκοπεί στη χρήση του RRH ως τεχνική τοπικής αναζήτησης. Τα όρια μεταβάλλονται με βάση μια παράμετρο (ri) που καθορίζει το εύρος της περιοχής ενδιαφέροντος.

- Επιλογή υποψήφιας λύσης με βάση το κριτήριο της απόστασης πλήθους (Crowding Distance [12, 135]) από το σύνολο των υποψήφιων λύσεων που δημιουργήθηκαν. Το κριτήριο λαμβάνει υπόψη την Ευκλείδεια απόσταση των λύσεων σε σχέση με τις λύσεις αποθηκευμένες στο αρχείο λύσεων. Η υποψήφια λύση που παρουσιάζει την μέγιστη τιμή του δείκτη επιλέγεται. Στη συνέχεια εφαρμόζονται οι υπολογιστικά δαπανηρές προσομοιώσεις για κάθε νέο συνδυασμό εγκατεστημένη ισχύς που προκύπτει από την επιλεγμένη λύση, ανανεώνονται τα αρχεία και η καλύτερη λύση που έχει εντοπιστεί. Η παραγωγή υποψήφιων λύσεων και η επιλογή επαναλαμβάνεται για κάθε διαφορετική τιμή της παραμέτρου (ri).

- Η παραγωγή υποψήφιων λύσεων επαναλαμβάνεται και η επιλογή γίνεται με βάση την εκτίμηση ΑΜ που εκπαιδεύονται τοπικά. Επιλέγεται μια νέα υποψήφια λύση που παρουσιάζει την ελάχιστη τιμή της συνάρτησης με βάση την εκτίμηση των τοπικών ΑΜ. Τα σημεία που χρησιμοποιούνται για την εκπαίδευση των τοπικών ΑΜ επιλέγονται με βάση το χωρίο που ορίζεται από την περιοχή ενδιαφέροντος. Ακολούθως γίνονται οι υπολογιστικά δαπανηρές προσομοιώσεις για κάθε νέο συνδυασμό εγκατεστημένη ισχύς που προκύπτει από την επιλεγμένη λύση, ανανεώνονται τα αρχεία και η καλύτερη λύση που έχει εντοπιστεί. Αυτά επαναλαμβάνονται για κάθε διαφορετική τιμή της παραμέτρου (ri) και για κάθε διαθέσιμο είδος ΑΜ.

- Τα βήματα της δεύτερης φάσης επαναλαμβάνονται μέχρι να ικανοποιηθούν τα κριτήρια τερματισμού.

4.3.4 Επιλεγμένα μοντέλα προσέγγισης

Η επιλογή ΑΜ γίνεται με βάση την Αναφορά [124] όπου γίνεται χρήση μοντέλου RBF [144] για πρόβλημα ολικής βελτιστοποίησης με ακέραιες μεταβλητές. Συγκεκριμένα, θεωρείται ότι είναι διαθέσιμα τρία μοντέλα RBF οποία διαφέρουν ως προς τη συνάρτηση πυρήνα (kernel) τους. Τα ΑΜ αυτά παρουσιάζονται αναλυτικότερα στο πλήρες κείμενο της Διατριβής.

4.4 Υπολογιστικά πειράματα και αποτελέσματα

Το σύστημα που εξετάστηκε είναι μια απλουστευμένη μορφή του Ελληνικού συστήματος παραγωγής ισχύος. Η χρήση ορισμένων δεδομένων που βασίζονται σε αυτό το σύστημα αποσκοπεί κυρίως στην δημιουργία ενός ενδεικτικού παραδείγματος.

4.4.1 Υπολογιστικά πειράματα

Τα υπολογιστικά πειράματα που έγιναν είναι τα εξής:

- Εξέταση του MAEA με χρήση δυο διαφορετικών SM. Ο MAEA εφαρμόζεται στο πρόβλημα βελτιστοποίησης και εξετάζεται η επίδοσή του για την εισαγωγή δυο διαφορετικών SM. Το πρώτο SM παρουσιάζει χαμηλότερη τεχνική λεπτομέρεια ($\mathrm{SM}^{\mathrm{ED}}$) σε σχέση με το δεύτερο ($\mathrm{SM}^{\mathrm{CUC}}$). Επίσης τίθεται άνω όριο στον αριθμό δαπανηρών προσομοιώσεων (500). Με βάση το τελικό αρχείο προσομοιώσεων των τρεξιμάτων εξετάζεται η ακρίβεια της εκτίμησης που επιτεύχθηκε από τα ΑΜ. Τα ανεξάρτητα τρεξίματα του MAEA επαναλήφθηκαν 30 φορές.

- Εξέταση του DE αλγορίθμου και των εξειδικευμένων τελεστών. Συγκεκριμένα δημιουργήθηκαν 24 παραλλαγές του αλγορίθμου βάσης από την εισαγωγή συνδυασμών αυτών των τελεστών. Η εξέταση έγινε σε τρία προβλήματα. Το πρώτο πρόβλημα βελτιστοποίησης εμπεριέχει στην συνάρτηση κόστους μόνο το κόστος επένδυσης και τα σταθερά κόστη λειτουργίας και συντήρησης. Για τα άλλα δυο προβλήματα γίνεται χρήση των ΑΜ για εκτίμηση του μεταβλητού κόστους και επαναλαμβάνονται για κάθε διαθέσιμο ΑΜ. Συγκεκριμένα, στο δεύτερο πρόβλημα το ΑΜ εκπαιδεύεται με βάση το τελικό αρχείο προσομοιώσεων ενός τρεξίματος με χρήση του MAEA και του $\mathrm{SM}^{\mathrm{ED}}$. Στο τρίτο το ΑΜ εκπαιδεύεται με βάση το τελικό αρχείο προσομοιώσεων ενός τρεξίματος με χρήση του MAEA και του $\mathrm{SM}^{\mathrm{CUC}}$. Τα ανεξάρτητα τρεξίματα κάθε παραλλαγής του DE αλγορίθμου επαναλήφθηκαν 100 φορές.

- Εξέταση του MAEA σε πρόβλημα βελτιστοποίησης όπου η συνάρτηση στόχος ορίστηκε ως η ελαχιστοποίηση της απόκλισης από μια συγκεκριμένη εγκατεστημένης ισχύ. Δυο περιπτώσεις εξετάστηκαν. Στη πρώτη ορίστηκε ως στόχος η λύση που ελαχιστοποιεί το κόστος επένδυσης και τα σταθερά κόστη λειτουργίας και συντήρησης. Τα ανεξάρτητα τρεξίματα του MAEA επαναλήφθηκαν 100 φορές. Στη δεύτερη περίπτωση ορίστηκαν 100 διαφορετικές λύσεις ως στόχος. Το άνω όριο στον αριθμό δαπανηρών προσομοιώσεων ορίστηκε στις 1000 προσομοιώσεις.

- Ανάλυση των αποτελεσμάτων που εξήχθησαν με χρήση των δυο διαφορετικών SM και σύγκρισή τους σε σχέση με ένα σενάριο αναφοράς όπου δεν εισάγεται SM για την εξαγωγή των επενδυτικών επιλογών.

4.4.2 Αποτελέσματα

Τα αποτελέσματα των υπολογιστικών πειραμάτων που έγιναν είναι τα εξής:

- Στα δυο προβλήματα όπου εξετάστηκε ο MAEA εντοπίστηκε συστηματικά λύση (30/30) εντός του ορίου με χρήση δυο διαφορετικών SM. Επίσης η ακρίβεια της προσέγγισης των ΑΜ, με βάση τους δείκτες που χρησιμοποιήθηκαν για την εξέταση κρίθηκε ικανοποιητική. Συνεπώς, το σφάλμα της προσέγγισης των AM δεν ήταν αρκετό έτσι ώστε να εμποδίσει την σύγκλιση του αλγορίθμου στις περιπτώσεις που εξετάστηκαν.

- Τα αποτελέσματα των υπολογιστικών πειραμάτων που αφορούν την επίδοση των 24 παραλλαγών του DE αλγορίθμου και των εξειδικευμένων τελεστών υποδηλώνουν ότι οι παραλλαγές του DE αλγορίθμου που ενσωματώνουν του τελεστές RRH και blk και αυτές που και ενσωματώνουν του τελεστές RRH, blk, PO ήταν οι πιο αποδοτικές.

- Στα δυο προβλήματα όπου εξετάστηκε ο MAEA και η συνάρτηση στόχος ορίστηκε ως η ελαχιστοποίηση της απόκλισης από μια συγκεκριμένη εγκατεστημένης ισχύ η απόδοση του αλγορίθμου κρίθηκε ικανοποιητική. Στην πρώτη περίπτωση όπου ορίστηκε ως στόχος η λύση που ελαχιστοποιεί το κόστος επένδυσης και τα σταθερά κόστη λειτουργίας και συντήρησης η βέλτιστη λύση εντοπίστηκε συστηματικά (100/100). Στη δεύτερη περίπτωση όπου ορίστηκαν 100 διαφορετικές λύσεις ως στόχος η βέλτιστη λύση εντοπίστηκε συστηματικά (99/100) εντός του άνω ορίου δαπανηρών προσομοιώσεων.

- Κατά τη σύγκριση των λύσεων που εντοπίστηκαν στα δυο προβλήματα όπου εξετάστηκε ο MAEA με χρήση δυο διαφορετικών SM παρατηρήθηκαν διαφορές στις επενδυτικές επιλογές που αποδόθηκαν στη χρήση διαφορετικού SM. Αυτό οφείλεται στις διαφορές κατά την εκτίμηση του κόστους των δυο SM. Εξετάζοντας τα εκτιμώμενα επίπεδα παραγωγής κάθε τεχνολογίας παρατηρήθηκε ότι στο χαμηλότερο επίπεδο τεχνικής λεπτομέρειας του ($\mathrm{SM}^{\mathrm{ED}}$) τα επίπεδα παραγωγής παρόχων ευελιξίας ήταν μειωμένα σε σχέση με το $\mathrm{SM}^{\mathrm{CUC}}$ καθώς τα χαμηλότερα επίπεδα τεχνικής λεπτομέρειας επέτρεπαν σε μη ευέλικτες μονάδες να παρουσιάζουν μεγαλύτερα επίπεδα παραγωγής υποτιμώντας έτσι τα επίπεδα παραγωγής πιο ευέλικτων μονάδων.

- Επίσης, έγινε χρήση του τελικού αρχείου προσομοιώσεων για την εξέταση της ευαισθησίας του μεταβλητού κόστους ως προς την εγκατεστημένη ισχύ της λύσης που εντοπίστηκε με χρήση του MAEA και του $\mathrm{SM}^{\mathrm{CUC}}$. Η ευαισθησία του μεταβλητού κόστους εξετάστηκε με βάση την εκτίμηση των AM για συνδυασμούς εγκατεστημένης ισχύος τεχνολογιών. Η εκτίμηση αυτή συγκρίθηκε με τις αντίστοιχες τιμές με χρήση του TM και παρατηρήθηκε ότι η εκτίμηση παρουσίαζε ικανοποιητική ακρίβεια.

Κεφάλαιο 5

Προγραμματισμός επέκτασης παραγωγής με υψηλό μερίδιο παραγωγής από ανανεώσιμες πηγές ενέργειας: Πολυκριτηριακή βελτιστοποίηση με χρήση εξελικτικών αλγορίθμων υποβοηθούμενων από μεταπρότυπα

5.1 Κίνητρα και στόχοι

Σε αυτό το Κεφάλαιο, παρουσιάζεται μια προσέγγιση για ένα στατικό MOO GEP. Αυτή στοχεύει προς εξέταση αντικρουόμενων στόχων και συντελεστών κόστους ενός υπολογιστικού ακριβού MOO GEP για τη διευκόλυνση της λήψης αποφάσεων. Η προσέγγιση εστιάζει στη συμπερίληψη πτυχών της βραχυπρόθεσμης λειτουργίας ενός συστήματος ισχύος εντός του MOO GEP. Πιο συγκεκριμένα, αποκλίνει από την βιβλιογραφία MOO GEP καθώς οι απαιτήσεις σε λειτουργική ευελιξία και ο αντίκτυπός τους σε συντελεστές κόστους εξετάζονται μέσω ενός SM για τη βραχυπρόθεσμη λειτουργία του συστήματος ισχύος. Η ενσωμάτωση του SM όμως οδηγεί σε αύξηση του υπολογιστικού κόστους.

Για την αντιμετώπιση του υπολογιστικού κόστους αναπτύσσεται μια προσέγγιση MOO με βάση πλαίσια για βελτιστοποίηση υποβοηθούμενη από AM και χωρίς χρήση παραγώγων. Πιο συγκεκριμένα, η προσέγγιση MOO περιλαμβάνει (i) μεταπρότυπα, (ii) MOEA, (iii) τοπική αναζήτηση και (iv) ανανέωση των AM κατά την διάρκεια του τρεξίματος. Τα υπολογιστικά πειράματα περιλαμβάνουν την εξέταση της προσέγγισης MOO σε συναρτήσεις αναφοράς και σε οικονομικά-περιβαλλοντικά MOO GEP προβλήματα βελτιστοποίησης. Επιπλέον, αναλύονται οι μη κυριαρχούμενες λύσεις που εντοπίστηκαν και οι αντίστοιχοι συντελεστές κόστους. Τα αποτελέσματα υποδηλώνουν ικανοποιητική απόδοση της MOO προσέγγισης.

5.2 Μαθηματική διατύπωση προβλήματος

Αυτή η Ενότητα παρουσιάζει τη μαθηματική διατύπωση του προβλήματος βελτιστοποίησης, των συντελεστών κόστους που χρησιμοποιούνται για τον προσδιορισμό των αντικειμενικών συναρτήσεων και τη μαθηματική διατύπωση του SM που χρησιμοποιείται για την εκτίμηση της βραχυπρόθεσμης λειτουργίας του συστήματος ισχύος. Τέλος, παρουσιάζονται οι αντικειμενικές συναρτήσεις και οι περιορισμοί πέντε παραλλαγών MOO GEP.

5.2.1 Μαθηματική διατύπωση προβλήματος βελτιστοποίησης

Εξετάζεται το πιο κάτω πρόβλημα MOO:

$$\begin{aligned} & minimize \;\; \mathbf{F}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), ..., f_m(\mathbf{x})) \\ & s.t. \;\; \mathbf{G}(\mathbf{x}) \leq 0 \\ & \qquad \mathbf{x} \in \mathbb{S} \end{aligned} \tag{5.1}$$

όπου $\mathbf{x} = (x_1, x_2, ..., x_n)$ είναι το διάνυσμα μεταβλητών απόφασης, n ο αριθμός μεταβλητών απόφασης, $\mathbf{F}(\mathbf{x})$ το διάνυσμα αντικειμενικών συναρτήσεων, m ο αριθμός αντικειμενικών συναρτήσεων, $\mathbf{G}(\mathbf{x}) = (g_1(\mathbf{x}), g_2(\mathbf{x}), ..., g_k(\mathbf{x}))$ το διάνυσμα περιορισμών του προβλήματος, k ο αριθμός περιορισμών του προβλήματος και $\mathbb{S}$ ο χώρος αναζήτησης.

Εξετάζεται η περίπτωση των υπολογιστικά δαπανηρών αντικειμενικών συναρτήσεων, υπολογιστικά φθηνών περιορισμών και $\mathbf{x} \in \mathbb{Z}^n$. Συνεπώς, θεωρείται ότι το πρόβλημα ελαχιστοποίησης περιορίζεται από το υπολογιστικό κόστος των αντικειμενικών συναρτήσεων. Επιπλέον, επιβάλλεται ένα όριο στον αριθμό των αντικειμένων συναρτήσεων που θεωρούνται στόχοι ($m = 2$), δηλαδή δεν εξετάζονται προβλήματα βελτιστοποίησης πολλών αντικειμενικών στόχων. Συγκεκριμένα, εξετάζονται προβλήματα βελτιστοποίησης GEP με δύο υπολογιστικά δαπανηρές αντικειμενικές συναρτήσεις και υπολογιστικά φθηνούς περιορισμούς. Οι μεταβλητές απόφασης θεωρούνται ως θετικοί ακέραιοι ($\mathbb{Z}^+$) που αντιπροσωπεύουν τον αριθμό των επενδύσεων σε μια υποψήφια τεχνολογία σε μια περιοχή. Το μέγεθος κάθε μιας τέτοιας επένδυσης θεωρείται προκαθορισμένο.

5.2.2 Συντελεστές κόστους

Πιο κάτω παρουσιάζονται οι συντελεστές κόστους που εξετάστηκαν και χρησιμοποιήθηκαν για τη διατύπωση των αντικειμενικών συναρτήσεων:

- *Κόστος επένδυσης:*

$$c^{inv} = \sum_{\forall a,g} [x^{inv}_{a,g} IC_{a,g} P^{cap\_step}_{a,g}] \tag{5.2}$$

- *Σταθερά κόστη λειτουργίας και συντήρησης:*

$$c^{fom} = \sum_{\forall a,g} [(x^{inv}_{a,g} + IU_{a,g}) FOM_{a,g} P^{cap\_step}_{a,g}] \tag{5.3}$$

- *Κόστος μηχανισμού στήριξης RES (Green Policy Support Cost, GPSC):*

$$c^{gp} = \sum_{\forall a,g^{res}} [(x^{inv}_{a,g^{res}} + IU_{a,g^{res}}) P^{cap\_step}_{a,g^{res}} Pr^{res}_{a,g^{res}} \sum_{\forall h} [Av^{res}_{a,g^{res},h}]] \tag{5.4}$$

- *Κόστος παραγωγής:*

$$c^{gen} = \sum_{\forall a,g,h} [p_{a,g,h}C_g^L + x_{a,g,h}^{su}C_g^{su} + x_{a,g,h}^{sd}C_g^{sd}] \tag{5.5}$$

- *Κόστος εκπομπών διοξειδίου του άνθρακα:*

$$c^{em} = \sum_{\forall a,g,h} [p_{a,g,h}EF_{a,g,h}]C^{em} \tag{5.6}$$

- *Μεταβλητά κόστη λειτουργίας και συντήρησης:*

$$c^{vom} = \sum_{\forall a,g,h} [p_{a,g,h}VOM_g] \tag{5.7}$$

- *Διείσδυση RES:*

$$c^{rp} = \sum_{\forall a,h} [\overline{p}_{a,h}^{res} - \epsilon_{a,h}^{c}]C^{rp} \tag{5.8}$$

όπου:

$$\overline{p}_{a,h}^{res} = \sum_{\forall g^{res}} [(x_{a,g^{res}}^{inv} + IU_{a,g^{res}})P_{a,g^{res}}^{cap\_step}Av_{a,g^{res},h}^{res}], \forall a, h \tag{5.9}$$

- *Κόστος μη εξυπηρετούμενης ενέργειας:*

$$c^{nse} = \sum_{\forall a,h} [\epsilon_{a,h}^{r}C^{r} + \epsilon_{a,h}^{d}C^{d}] \tag{5.10}$$

- *Κόστος μείωσης παραγωγής από RES ή υδροηλεκτρικών μονάδων:*

$$c^{cs} = \sum_{\forall a,h} [\epsilon_{a,h}^{s}C^{s} + \epsilon_{a,h}^{c}C^{c}] \tag{5.11}$$

Με βάση τις παραδοχές που έγιναν, μπορεί να γίνει διάκριση μεταξύ των συντελεστών κόστους που παρουσιάστηκαν. Συγκεκριμένα, c^{inv}, c^{fom} και c^{gp} μπορούν να υπολογιστούν λαμβάνοντας υπόψη μόνο τις επενδυτικές αποφάσεις ($\mathbf{x}^{inv}$). Αντίθετα, οι συντελεστές κόστους c^{gen}, c^{em}, c^{vom}, c^{rp}, c^{cs} και c^{nse} θα μπορούσαν να υπολογιστούν με βάση μια προσομοίωση της βραχυπρόθεσμης λειτουργίας ενός συστήματος ισχύος. Για τον υπολογισμό αυτών των συντελεστών κόστους γίνεται χρήση ενός SM του οποίου η είσοδος είναι οι επενδυτικές αποφάσεις. Η έξοδος του SM είναι ένα διάνυσμα $\mathbf{v} = \{\mathbf{p}, \mathbf{x}^{su}, \mathbf{x}^{sd}, \epsilon^r, \epsilon^d, \epsilon^c, \epsilon^s\}$ το οποίο χρησιμοποιείται για να υπολογιστούν οι συντελεστές κόστους. Δηλαδή, το αποτέλεσμα του SM θεωρείται ως το αποτέλεσμα της προσομοίωσης της βραχυπρόθεσμης λειτουργίας ενός συστήματος ισχύος.

Για την εφαρμογή αυτού απαιτείται ότι το SM παρέχει ένα αποτέλεσμα της προσομοίωσης για κάθε είσοδο. Επίσης, σε περίπτωση όπου το SM είναι ένα πρόβλημα βελτιστοποίησης το οποίο παρουσιάζει πέραν της μιας βέλτιστης λύσης γίνεται η παραδοχή ότι η λύση αυτή, δηλαδή το αποτέλεσμα της προσομοίωσης, είναι μοναδική.

5.2.3 Μοντέλο προσομοίωσης και προσέγγιση του προβλήματος

Το μοντέλο Fast Unit Commitment [26] επιλέγεται ως το SM για την προσέγγιση ενός UCP. Σε αυτό εισάγονται κάποιες τροποποιήσεις που σχετίζονται με την συγκεκριμένη εφαρμογή. Ένα πλεονέκτημα αυτού του μοντέλου είναι η απουσία ακέραιων μεταβλητών. Συγκεκριμένα, το μοντέλο βελτιστοποίηση είναι πρόβλημα γραμμικού προγραμματισμού. Αντίστοιχα μοντέλα προσέγγισης, βασιζόμενα σε πρόβλημα γραμμικού προγραμματισμού, έχουν εφαρμοστεί για την εξέταση του βραχυπρόθεσμου προγραμματισμού ενός συστήματος ισχύος στο πλαίσιο του μακροπρόθεσμου προγραμματισμού [45].

5.2.4 Διατύπωση πολυκριτηριακών προβλημάτων βελτιστοποίησης

Μεταβλητές Απόφασης: Έστω $\mathbf{x}$ το διάνυσμα μεταβλητών απόφασης παραγόμενο από το μητρώο $\mathbf{x}^{inv}$ ως εξής:

$$\mathbf{x} = \{x_{1,1}^{inv}, x_{1,2}^{inv}, \ldots, x_{1,gz}^{inv}, x_{2,1}^{inv}, x_{2,2}^{inv}, \ldots, x_{2,gz}^{inv}, x_{az,1}^{inv}, x_{az,2}^{inv}, \ldots, x_{az,gz}^{inv}\} \tag{5.12}$$

όπου:

$$\underline{X}_{a,g} \leq x_{a,g}^{inv} \leq \overline{X}_{a,g}, \forall a, g \tag{5.13}$$

$$\mathbf{x} \in \mathbb{Z} \tag{5.14}$$

Αντικειμενικές συναρτήσεις: Με βάση του συντελεστές κόστους που παρουσιάστηκαν, ορίζονται οι παρακάτω πέντε συνδυασμοί αντικειμενικών συναρτήσεων που αντιπροσωπεύουν τα πέντε MOO GEP παραλλαγές που εξετάστηκαν:

1. Συνολικό κόστος και κόστος εκπομπών (AC1-EM):

   - Ελαχιστοποίηση συνολικού κόστους (AC1):

   $$l_1(\mathbf{x}) = c^{inv} + c^{fom} + c^{gen} + c^{vom} + c^{gp} + c^{cs} \tag{5.15}$$

   - Ελαχιστοποίηση κόστος εκπομπών (EM):

   $$l_2(\mathbf{x}) = c^{em} \tag{5.16}$$

2. Συνολικό κόστος και διείσδυση των RES (AC2-RP):

   - Ελαχιστοποίηση συνολικού κόστους (AC2):

   $$l_1(\mathbf{x}) = c^{inv} + c^{fom} + c^{gen} + c^{vom} + c^{em} + c^{gp} + c^{cs} \tag{5.17}$$

   - Μεγιστοποίηση διείσδυσης των RES (RP):

   $$l_2(\mathbf{x}) = (-1)c^{rp} \tag{5.18}$$

3. Λειτουργικού κόστους και GPSC (OC1-GS):

   - Ελαχιστοποίηση λειτουργικού κόστους (OC1):

   $$l_1(\mathbf{x}) = c^{fom} + c^{gen} + c^{vom} + c^{em} + c^{cs} \tag{5.19}$$

- Ελαχιστοποίηση GPSC (GS):
$$l_2(\mathbf{x}) = c^{gp} \tag{5.20}$$

4. Λειτουργικού κόστους και κόστους επένδυσης (OC2-IC):
   - Ελαχιστοποίηση λειτουργικού κόστους (OC2):
$$l_1(\mathbf{x}) = c^{gen} + c^{fom} + c^{vom} + c^{em} + c^{gp} + c^{cs} \tag{5.21}$$
   - Ελαχιστοποίηση κόστους επένδυσης (IC):
$$l_2(\mathbf{x}) = c^{inv} \tag{5.22}$$

5. Συνολικό κόστος και συνολικό κόστος ποινής (AC3-AP):
   - Ελαχιστοποίηση συνολικού κόστους (AC3):
$$f_1(\mathbf{x}) = c^{inv} + c^{fom} + c^{gen} + c^{vom} + c^{em} + c^{gp} \tag{5.23}$$
   - Ελαχιστοποίηση συνολικού κόστους ποινής (AP):
$$f_2(\mathbf{x}) = c^{nse} + c^{cs} \tag{5.24}$$

Επιπλέον, το κόστος μη εξυπηρετούμενης ενέργειας προστίθεται στους στόχους των τεσσάρων πρώτων παραλλαγών MOO GEP ως όρος κόστους ποινής σε μια προσπάθεια αποκλεισμού λύσεων που εμφανίζουν κόστος μη εξυπηρετούμενης ενέργειας από το μέτωπο μη κυριαρχούμενων λύσεων ως εξής:

$$f_1(\mathbf{x}) = l_1(\mathbf{x}) + w^{pen} \cdot pen(\mathbf{x}) \tag{5.25}$$

$$f_2(\mathbf{x}) = l_2(\mathbf{x}) + w^{pen} \cdot pen(\mathbf{x}) \tag{5.26}$$

όπου $pen(\mathbf{x}) = c^{nse}$ και $w^{pen} = 1$. Η παράμετρος w^{pen} είναι ο συντελεστής ποινής.

Περιορισμοί: Στις πέντε MOO GEP παραλλαγές εισάγονται οι ακόλουθοι περιορισμοί ($\mathbf{G}(\mathbf{x})$):

- *Περιθώριο ασφαλείας:*
$$(1 + RM)PD \leq \sum_{\forall a,g} [(x_{a,g}^{inv} + IU_{a,g})P_{a,g}^{cap\_step}] \tag{5.27}$$

- *Μέγιστος αριθμός επενδύσεων:*
$$\sum_{\forall a} [x_{a,g}^{inv}] \leq TCA_g, \forall g \tag{5.28}$$

Η συνολική παραβίαση των περιορισμών ($cv(\mathbf{x})$) υπολογίζεται ως εξής:

$$cv(\mathbf{x}) = \sum_{i=1}^{k} \frac{\max(g_i(\mathbf{x}), 0)}{CV_i^{max}} \tag{5.29}$$

όπου $\mathbf{CV}^{max}$ το διάνυσμα που περιλαμβάνει τις μέγιστες τιμές των παραβιάσεων κάθε περιορισμού.

5.3 Προσέγγιση Πολυκριτηριακής βελτιστοποίησης

Αυτή η ενότητα παρουσιάζει την προσέγγιση του MOO. Βασίζεται σε MOEA και βελτιστοποίηση με χρήση μεταπροτύπων. Ο αλγόριθμος NSGA-III [95, 145] έχει χρησιμοποιηθεί ως ο βασικός MOEA ο οποίος τροποποιείται με τη συμπερίληψη AM. Αυτές οι τροποποιήσεις γίνονται με βάση προτεινόμενα πλαίσια βελτιστοποίησης με χρήση μεταπροτύπων [111, 112, 130, 131, 132, 133, 146, 147].

Για τον προσδιορισμό των βημάτων της προσέγγισης και με βάση τις προαναφερθείσες αναφορές, τα ακόλουθα θεωρήθηκαν σημαντικά για την επίτευξη ικανοποιητικών αποτελεσμάτων:

- Η δημιουργία κατάλληλων υποψήφιων σημείων (Pool of Data Points, PDP).
- Η επίτευξη αποδεκτής προσέγγισης από τα AM (Underlying Functions Approximation, UFA).
- Η επίτευξη αποδεκτής προσέγγισης του μετώπου Pareto (Pareto Front Approximation, PFA).

Η προσέγγιση που παρουσιάζεται περιλαμβάνει ολική και τοπική φάση αναζήτησης για να επωφεληθεί από (i) ολικά και τοπικά AM [111, 112, 132, 133, 146, 147], και (ii) ολική και τοπική αναζήτηση [111, 112, 146, 148].

Δυο κριτήρια επιλογής σημείων από το PDP ενσωματώθηκαν με βάση τις Αναφορές [132, 133]. Συγκεκριμένα, επιλέχθηκε το κριτήριο μέγιστης-ελάχιστης απόστασης (Maximum-Minimum Distance Criterion, MMDC) και ένα κριτήριο βασισμένο στον υπερόγκο (Hypervolume). Ως CHT χρησιμοποιούνται τα Feasibility Rules [135].

5.3.1 Βήματα της προσέγγισης βελτιστοποίησης

Η προσέγγιση βελτιστοποίησης περιλαμβάνει τρεις διαφορετικές φάσεις (i) την αρχικοποίηση φάση, (ii) την ολική φάση και (iii) την τοπική φάση. Τα δύο τελευταία επαναλαμβάνονται σειριακά. Μια τέτοια επανάληψη θεωρείται ως ένας κύκλος βελτιστοποίησης.

Η φάση αρχικοποίηση περιλαμβάνει την εισαγωγή των απαιτούμενων δεδομένων και τον προσδιορισμό του αρχικού συνόλου σημείων (αρχείο) που υπολογίζονται με το υπολογιστικά δαπανηρό μοντέλο. Επιπλέον, πρόσθετα σημεία, σε περιοχές ενδιαφέροντος του χώρου αναζήτησης, θα μπορούσαν να εισαχθούν αν αυτά είναι διαθέσιμα.

Η φάση της ολικής αναζήτησης περιλαμβάνει την εκπαίδευση των διαθέσιμων AM, τη δημιουργία μιας PDP με χρήση MOEA και AM και την επιλογή υποψήφιων σημείων που θα υπολογιστούν με χρήση του SM. Συγκεκριμένα, πληθυσμοί εξελίσσονται ανεξάρτητα με χρήση διαφορετικών MOEA και AM. Οι τελικοί πληθυσμοί ενώνονται για τη δημιουργία της PDP. Σε αυτή εισάγονται τυχαία σημεία και σημεία που παρήχθησαν σε προηγούμενους κύκλους βελτιστοποίησης. Με τη δημιουργία της PDP εφαρμόζονται μια σειρά από κανόνες και τα κριτήρια MMDC και Hypervolume για τον προσδιορισμών των υποψήφιων σημείων που θα υπολογιστούν με χρήση του SM και θα αποθηκευτούν στο αρχείο λύσεων.

Η φάση της τοπικής αναζήτησης περιλαμβάνει την εκπαίδευση τοπικών AM για επιλεγμένες λύσεις του μετώπου μη-κυριαρχούμενων λύσεων του αρχείου και σε υποψήφιες λύσεις που παράγονται από αυτές. Σε αυτές εφαρμόζεται τοπική αναζήτηση βασιζόμενη σε τοπικά AM και βελτιστοποίηση με χρήση παραγώγων.

Οι τρεις φάσεις και τα βήματα της προσέγγισης βελτιστοποίησης που αναπτύχθηκε παρουσιάζονται αναλυτικότερα στο πλήρες κείμενο της Διατριβής.

5.3.2 Μοντέλα προσέγγισης

Ως διαθέσιμα AM θεωρούνται τα μοντέλα RBF και PR. Η χρήση διαφορετικών AM αποσκοπεί στην αντιμετώπιση της *κατάρας της αβεβαιότητας* και στην εκμετάλλευση της *ευλογίας της αβεβαιότητας* [111, 113]. Επιπλέον, στην φάση της τοπικής αναζήτησης γίνεται χρήση ενός AM που αποτελεί συνδυασμό (άθροισμα της εκτίμησης) των RBF και PR.

5.3.3 Τελεστές εξέλιξης

Γίνεται χρήση διαφορετικών EA για την εξέλιξη κάθε υπό πληθυσμού. Οι δυο διαφορετικοί συνδυασμοί τελεστών που εξετάστηκαν βασίζονται σε Γενετικούς αλγορίθμους [149] και σε DE [150].

5.3.4 Τελεστής επιλογής

Εξετάστηκε ο τελεστής επιλογής του NSGA-III [95, 145] που βασίζεται στον NSGA-II [89]. Ο NSGA-III αναπτύχθηκε για προβλήματα βελτιστοποίησης πολλών στόχων και εμπεριέχει μηχανισμό για εισαγωγή προτιμήσεων στην αναζήτηση.

5.4 Υπολογιστικά πειράματα και αποτελέσματα

Τα υπολογιστικά πειράματα που έγιναν αποσκοπούν στην εξέταση της απόδοσης του αλγορίθμου και στην εξαγωγή και ανάλυση κάθε μετώπου μη-κυριαρχούμενων λύσεων για τις πέντε MOO GEP παραλλαγές που παρουσιάστηκαν. Το σύστημα που εξετάστηκε είναι μια απλουστευμένη μορφή του Ελληνικού συστήματος παραγωγής ενέργειας με βάση τα δεδομένα εισόδου που θεωρήθηκαν [151, 152]. Η προσέγγιση υλοποιήθηκε στη Matlab χρησιμοποιώντας την πλατφόρμα Platemo [153] και την υλοποίηση των RBF και PR μοντέλων [124, 125].

5.4.1 Υπολογιστικά πειράματα

Τα υπολογιστικά πειράματα που έγιναν είναι τα ακόλουθα:

- Εξέταση της MOO προσέγγισης σε συναρτήσεις αναφοράς (ZDT test suite [154]). Στις συναρτήσεις αυτές εισάγονται τροποποιήσεις καθώς η MOO προσέγγιση αναπτύχθηκε για προβλήματα βελτιστοποίησης με ακέραιες μεταβλητές. Σε αυτές εφαρμόζεται και ο NSGA-III του οποίου τα αποτελέσματα χρησιμεύουν ως σημείο αναφοράς.

- Εξέταση της MOO προσέγγισης σε υπολογιστικά φθηνό MOO GEP πρόβλημα το οποίο δεν εμπεριέχει SM. Συγκεκριμένα:

  1. Ελαχιστοποίηση κόστους επένδυσης, σταθερού κόστους λειτουργίας και συντήρησης και GPSC:
  $$f_1(\mathbf{x}) = c^{inv} + c^{fom} + c^{gp} \tag{5.30}$$

  2. Μεγιστοποίηση της αναμενόμενης παραγωγής από RES:
  $$f_2(\mathbf{x}) = (-1)c^{arp} \tag{5.31}$$

όπου c^{arp} είναι η εκτιμώμενη παραγωγή από RES:

$$c^{arp} = C^{rp} \sum_{\forall a,h} [\overline{p}^{res}_{a,h}] \tag{5.32}$$

όπου $C^{rp} = 1$.

Το υπολογιστικό πείραμα επαναλαμβάνεται δυο φορές. Στην πρώτη λαμβάνεται υπόψη η υφιστάμενη κατάσταση του συστήματος ισχύος (*MOOGEP-noSM*) ενώ στην δεύτερη θεωρείται ότι δεν υπάρχουν εγκατεστημένες μονάδες (*MOOGEP-noSM-GF*). Σε αντίθεση με την περίπτωση που δεν υπάρχουν εγκατεστημένες μονάδες και με βάση τα δεδομένα εισόδου η υφιστάμενη κατάσταση δεν παραβιάζει τους περιορισμούς. Συνεπώς, τα δυο προβλήματα βελτιστοποίησης διαφέρουν και ως προς το μέγεθος του χώρου εφικτών λύσεων. Εφαρμόζονται 25 ανεξάρτητα τρεξίματα με όριο τους 500 υπολογισμούς των αντικειμενικών συναρτήσεων.

- Εξέταση της MOO προσέγγισης στις πέντε MOO GEP παραλλαγές με χρονική λεπτομέρεια του SM ορισμένη στη 1 ημέρα ($24h$). Συνεπώς, το υπολογιστικό κόστος δεν είναι ιδιαίτερα αυξημένο ($< 1sec$). Οι πέντε MOO GEP παραλλαγές που λαμβάνουν υπόψη την υφιστάμενη κατάσταση αναφέρονται ως *MOOGEP-(XXX)-1D* (π.χ. *MOOGEP-(AC1-EM)-1D*) ενώ οι πέντε MOO GEP παραλλαγές που δεν λαμβάνουν υπόψη την υφιστάμενη κατάσταση αναφέρονται ως *MOOGEP-(XXX)-1D-GF*. Σε αυτή τη περίπτωση οι παραλλαγές *MOOGEP-(XXX)-1D* και *MOOGEP-(XXX)-1D-GF* διαφέρουν και ως προς τον αριθμό των σημείων στο χώρο αναζήτησης που παρουσιάζουν ποινή κατά την αξιολόγησή του από το SM. Εφαρμόζονται 25 ανεξάρτητα τρεξίματα για κάθε παραλλαγή με όριο τις 500 προσομοιώσεις με χρήση του SM.

- Εξέταση της MOO προσέγγισης στις πέντε MOO GEP παραλλαγές με αυξημένη χρονική λεπτομέρεια του SM ορισμένη στις 4 βδομάδες. Το SM εφαρμόζεται ανεξάρτητα σε κάθε εβδομάδα και η άθροιση του κόστους λαμβάνει υπόψη ίση βαρύτητα στα κόστη που προκύπτουν για κάθε βδομάδα. Οι πέντε MOO GEP παραλλαγές που λαμβάνουν υπόψη την υφιστάμενη κατάσταση αναφέρονται ως *MOOGEP-(XXX)-4W* ενώ οι πέντε MOO GEP παραλλαγές που δεν λαμβάνουν υπόψη την υφιστάμενη κατάσταση αναφέρονται ως *MOOGEP-(XXX)-4W-GF*. Εφαρμόζονται 25 ανεξάρτητα τρεξίματα για κάθε παραλλαγή με όριο τις 500 προσομοιώσεις με χρήση του SM. Τα αποτελέσματα για τις *MOOGEP-(XXX)-4W* παραλλαγές αναλύονται ως προς το μέτωπο μη-κυριαρχούμενων λύσεων που εντοπίστηκε και τις τιμές των συντελεστών κόστους σε κάθε ένα από τα μέτωπα αυτά.

5.4.2 Αποτελέσματα

Τα αποτελέσματα των υπολογιστικών πειραμάτων που έγιναν είναι τα ακόλουθα:

- *Εξέταση της MOO προσέγγισης σε προβλήματα αναφοράς:* Τα αποτελέσματα για τα προβλήματα αναφοράς ZDT1, ZDT2 και ZDT3 ήταν ικανοποιητικά εντός του ορίου προσομοιώσεων που τέθηκε με βάση και τους δείκτες υπερόγκου που χρησιμοποιήθηκαν. Στα προβλήματα αναφοράς ZDT4 και ZDT6 η απόδοση του αλγορίθμου δεν ήταν εξίσου ικανοποιητική καθώς το PFA ήταν χαμηλής ακρίβειας. Αυτό αποδόθηκε σε προκλήσεις που σχετίζονται με συγκεκριμένες συναρτήσεις των προβλημάτων αναφοράς όπως ο αυξημένος αριθμός τοπικών ακρότατων (ZDT4) και η αραιή κατανομή λύσεων κοντά στο μέτωπο Pareto (ZDT6) τα οποία φαίνεται να επηρέασαν την ακρίβεια του UFA. Συγκεκριμένα, η ακρίβεια του UFA στις συναρτήσεις των προβλημάτων αναφοράς ZDT4 και ZDT6 δεν ήταν όσο υψηλή όσο στα προβλήματα ZDT1, ZDT2 και ZDT3.

- *Εξέταση της MOO προσέγγισης σε υπολογιστικά φθηνό MOO GEP πρόβλημα το οποίο δεν εμπεριέχει SM:* Τα αποτελέσματα για τα προβλήματα *MOOGEP-noSM* και *MOOGEP-noSM-GF* ήταν ικανοποιητικά με βάση και τους δείκτες υπερόγκου που χρησιμοποιήθηκαν και τους περιορισμούς σε υπολογιστικό κόστος που τέθηκαν. Η σύγκριση των αποτελεσμάτων με τα αντίστοιχα του NSGA-III υποδηλώνουν ότι η εισαγωγή των AM επιτάχυνε την σύγκλιση και δεν απέτρεψε την εξαγωγή ενός ικανοποιητικού PFA. Επιπλέον, η ακρίβεια του UFA ήταν ικανοποιητική με βάση τους δείκτες που χρησιμοποιήθηκαν. Η σύγκριση των αποτελεσμάτων για τα προβλήματα *MOOGEP-noSM* και *MOOGEP-noSM-GF* υποδηλώνει ότι η απόδοση του αλγορίθμου δεν επηρεάστηκε από τις διαφορές στο μέγεθος του χώρου των εφικτών λύσεων.

- *Εξέταση της MOO προσέγγισης στις πέντε MOO GEP παραλλαγές με χρονική λεπτομέρεια του SM ορισμένη στη 1 ημέρα (24h):* Αντίστοιχα, τα αποτελέσματα για τα προβλήματα *MOOGEP-(XXX)-1D* και *MOOGEP-(XXX)-1D-GF* ήταν ικανοποιητικά με βάση τους δείκτες υπερόγκου που χρησιμοποιήθηκαν και τους περιορισμούς σε υπολογιστικό κόστος που τέθηκαν. Σε αντίθεση με τα προβλήματα *MOOGEP-noSM* και *MOOGEP-noSM-GF*, εντοπίστηκαν διαφορές όσο αφορά την επίδοση της MOO προσέγγισης στα προβλήματα *MOOGEP-(XXX)-1D* και *MOOGEP-(XXX)-1D-GF*. Συγκεκριμένα, η ακρίβεια του UFA στις παραλλαγές *MOOGEP-(XXX)-1D-GF* φαίνεται να επηρεάστηκε από την εισαγωγή λύσεων που παρουσιάζουν ποινή με βάση το SM.

- *Εξέταση της MOO προσέγγισης στις πέντε MOO GEP παραλλαγές με αυξημένη χρονική λεπτομέρεια του SM ορισμένη στις 4 βδομάδες:* Τα αποτελέσματα για τα προβλήματα *MOOGEP-(XXX)-4W* και *MOOGEP-(XXX)-4W-GF* ήταν αποδεκτά με βάση τους περιορισμούς σε υπολογιστικό κόστος που τέθηκαν. Οι δείκτες υπερόγκου που χρησιμοποιήθηκαν υποδηλώνου σταδιακή βελτίωση του μετώπου μη κυριαρχούμενων λύσεων. Σε κάποιες παραλλαγές (π.χ. *MOOGEP-AC2-RP-4W-GF*) οι δείκτες υποδηλώνουν ότι η αύξηση των διαθέσιμων υπολογιστικά δαπανηρών προσομοιώσεων θα μπορούσε να οδηγήσει σε βελτίωση του μετώπου μη κυριαρχούμενων λύσεων. Επιπλέον, η ακρίβεια του UFA επηρεάστηκε από την εισαγωγή λύσεων που παρουσιάζουν ποινή με βάση το SM.

- *Ανάλυση των μετώπων μη κυριαρχούμενων λύσεων και των αντίστοιχων συντελεστών κόστους:* Τα αποτελέσματα του κάθε μετώπου μη κυριαρχούμενων λύσεων που εντοπίστηκαν για τα προβλήματα *MOOGEP-(XXX)-4W* αναλύονται. Μεταξύ των πέντε παραλλαγών εντοπίζονται διαφορετικές λύσεις ανάλογα με τους στόχους που τέθηκαν και εξετάζεται η επίδραση των στόχων αυτών στην εξέλιξη των τιμών των συντελεστών κόστους στο κάθε μετώπου μη κυριαρχούμενων λύσεων. Συγκεκριμένα, δεδομένων διαφορετικών συναρτήσεων στόχου, μπορούν να εξαχθούν διαφορετικά συμπεράσματα σχετικά με βάση τις προσθήκες δυναμικότητας. Αυτό είχε αποδοθεί στην επίδραση κάθε συντελεστή κόστους σε κάθε αντικειμενική συνάρτηση. Η πρώτη διατύπωση που εμπεριέχει οικονομικούς και περιβαλλοντικούς στόχους (*MOOGEP-AC1-EM-4W*) υποδηλώνει ότι η μείωση του κόστους εκπομπών μπορεί να σχετίζεται με υψηλότερο συνολικό κόστος. Η δεύτερη οικονομική-περιβαλλοντική διατύπωση (*MOOGEP-AC2-RP-4W*) υποδηλώνει ότι η αύξηση του επιπέδου διείσδυσης RES μπορεί να σχετίζεται με υψηλότερο συνολικό κόστος. Και οι δύο διατυπώσεις έδειξαν ότι απαιτούνται επενδύσεις σε προσθήκες δυναμικότητας που δεν παράγουν εκπομπές GHG ή σε νέες θερμικές εγκαταστάσεις. Η τρίτη διατύπωση (*MOOGEP-OC1-GP-4W*) είχε εξετάσει το συνολικό λειτουργικό κόστος και το κόστος ενός μηχανισμού στήριξης RES. Με βάση τα αποτελέσματα προτεραιότητα δόθηκε σε μονάδες RES, υδροηλεκτρικές μονάδες και θερμικές μονάδες βάσης. Τα αποτελέσματα υποδηλώνουν ότι η μείωση του λειτουργικού κόστους απαιτεί αύξηση του GPSC. Η τέταρτη διατύπωση (*MOOGEP-OC2-IC-4W*) υποδηλώνει ότι υψηλότερο επενδυτικό κόστος μπορεί να οδηγήσει σε μείωση του λειτουργικού κόστους. Τα αποτελέσματα που είχαν εξαχθεί για τη πέμπτη διατύπωση (*MOOGEP-AC3-AP-4W*) υποδηλώνουν ότι απαιτείται ένας περιορισμένος αριθμός επενδύσεων (σε μονάδες βάσης) για την επίτευξη ενός αποτελεσματικού οικονομικού συστήματος ισχύος. Αυτό είχε αποδοθεί στην επίδραση συγκεκριμένων συντελεστών κόστους (π.χ. κόστος επένδυσης και σταθερού κόστους λειτουργίας και συντήρησης) στην τιμή της αντικειμενικής συνάρτησης που είχε καταστήσει περεταίρω προσθήκες ως λιγότερο αποτελεσματικές.

Κεφάλαιο 6

Συμπεράσματα

Το Κεφάλαιο 4 παρουσιάζει μια μονοκριτηριακή προσέγγιση βασισμένη σε ΜΑΕΑ, για το GEP πολλαπλών περιόδων. Έχει ως στόχο την αντιμετώπιση προκλήσεων που σχετίζονται με μια μετάβαση προς υψηλότερα μερίδια παραγωγής από RES και προορίζεται να χρησιμοποιείται παράλληλα με άλλα καθιερωμένα μοντέλα GEP. Τα AM χρησιμοποιήθηκαν στο πλαίσιο των ΜΑΕΑ για να προσφέρουν μια εκτίμηση της τιμής που θα προέκυπτε από τη χρήση του SM. Το τελευταίο χρησιμεύει ως δείκτης κόστους της βραχυπρόθεσμης λειτουργίας ενός συστήματος ισχύος. Γενικά, ο υπολογιστικός χρόνος εξαρτάται σε μεγάλο βαθμό από το υπολογιστικό κόστος του SM. Υποθέτοντας ότι γίνεται χρήση ενός SM το οποίο μπορεί να εντοπίσει τις απαιτήσεις σε λειτουργική ευελιξία, η συμπερίληψη των AM στοχεύει στην επίτευξη λύσης χρησιμοποιώντας μόνο έναν περιορισμένο αριθμό προσομοιώσεων.

Ο πρώτος στόχος που τέθηκε ήταν να εξεταστεί η δυνατότητα εφαρμογής του ΜΑΕΑ που αναπτύχθηκε. Ως εκ τούτου, πραγματοποιήθηκε μια σειρά υπολογιστικών πειραμάτων. Τα αποτελέσματα που αφορούν την απόδοση του ΜΑΕΑ ήταν ικανοποιητικά. Συγκεκριμένα, η εγκατεστημένη δυναμικότητα παραγωγής ισχύος είχε εντοπιστεί, όταν αυτή είχε τεθεί σαν συνάρτηση στόχος, εντός ενός προκαθορισμένου ορίου προσομοιώσεων. Επιπλέον, η λύση που προέκυπτε για τα δύο SM που εξετάστηκαν λήφθηκε συστηματικά. Επίσης, η αξιολόγηση της ποιότητας της προσέγγισης που επιτεύχθηκε από τα AM υποδηλώνει ότι τα εκτιμώμενα σφάλματα, με βάση του δείκτες που χρησιμοποιήθηκαν, δεν ήταν πολύ υψηλά. Αυτά τα σφάλματα δεν ήταν επαρκή για να αποτρέψουν τον ΜΑΕΑ από την επίτευξη της ίδιας λύσης. Κάνοντας χρήση των AM, πραγματοποιήθηκε μια οπτική ανάλυση για την εξέταση της ευαισθησίας του λειτουργικού κόστους προς την εγκατεστημένη δυναμικότητα παραγωγής ισχύος. Αυτή έδειξε ότι η αύξηση των επενδύσεων σε προσθήκες δυναμικότητας παραγωγής ισχύος θα μπορούσε να οδηγήσει σε μείωση του λειτουργικού κόστους για το τελικό έτος του χρονικού ορίζοντα προγραμματισμού.

Κατά συνέπεια, το μοντέλο GEP με βάση τους ΜΑΕΑ θα μπορούσε να αποτελέσει μια υποσχόμενη προσέγγιση για τον εντοπισμό λύσεων για μονοκριτηριακό πρόβλημα βελτιστοποίησης πολλαπλών περιόδων GEP το οποίο συμπεριλαμβάνει ένα SM της βραχυπρόθεσμης λειτουργίας ενός συστήματος ισχύος. Επιπλέον, τα AM θα μπορούσαν να χρησιμοποιηθούν για οπτική ανάλυση της ευαισθησίας του λειτουργικού κόστους προς την εγκατεστημένη δυναμικότητα παραγωγής ισχύος μιας λύσης, αν επιτευχτεί αποδεκτή ακρίβεια της προσέγγισης από τα AM.

Ο δεύτερος στόχος που εξετάστηκε ήταν η αξιολόγηση του πιθανού κέρδος της εισαγωγής εξειδικευμένων τελεστών που αναπτύχθηκαν για το συγκεκριμένο πρόβλημα (RRH, blk και PO). Αυτοί είχαν εξεταστεί ως πιθανές βελτιώσεις για έναν βασικό αλγόριθμο DE. Πραγματοποιήθηκε μια σειρά από υπολογιστικά πειράματα σε παραλλαγές του βασικού αλγορίθμου DE, συμπεριλαμβανομένων συνδυασμών των προαναφερθέντων και αξιολογήθηκε η επίδρασή τους στην απόδοση του αλγορίθμου. Τα αποτελέσματα που προέκυψαν υποδηλώνουν ότι οι παραλλαγές που συμπεριελάμβαναν τους RRH και blk ή τους RRH, blk και PO ήταν οι πιο ανταγωνιστικές. Ωστόσο, το υψηλότερο κέρδος είχε επιτευχθεί με τη εισαγωγή του RRH.

Επομένως, εξειδικευμένοι τελεστές θα μπορούσαν να συμβάλλουν στην βελτίωση της απόδοσης ενός EA στο πρόβλημα που εξετάστηκε. Συγκεκριμένα, προτείνεται έμφαση σε τελεστές που επιδιορθώνουν ανέφικτες λύσεις, οι οποίες παράγονται από έναν πληθυσμό EA, όταν ο EA εφαρμόζεται σε προβλήματα βελτιστοποίησης υπό περιορισμούς για τα οποία υπάρχει διαθέσιμη πληροφορία σχετικά με τις συναρτήσεις περιορισμού.

Ο τρίτος στόχος που τέθηκε ήταν η εκτίμηση του αντίκτυπου της εισαγωγής ενός SM όταν εξετάζονται αυξημένα μερίδια παραγωγής από RES. Εξετάστηκαν τρεις περιπτώσεις. Η πρώτη, η οποία ορίστηκε ως σημείο αναφοράς, δεν περιλάμβανε SM. Επομένως, η λύση είχε επιτευχθεί με βάση μόνο το κόστος επένδυσης και το σταθερό κόστος λειτουργίας και συντήρησης. Η δεύτερη περίπτωση λαμβάνει υπόψη το λειτουργικό κόστος με χρήση ενός SM, ωστόσο, αυτό παρουσιάζει σχετικά χαμηλή τεχνική λεπτομέρεια (το κόστος και οι περιορισμοί που προκύπτουν από συναρτήσεις περιορισμών του UCP παραλείπονται). Η τρίτη περίπτωση περιελάμβανε ένα SM βασισμένο στο CUC, που παρουσιάζει υψηλή τεχνική λεπτομέρεια για το πλαίσιο του μακροπρόθεσμου προγραμματισμού. Συνεπώς, αυτή η σύγκριση εστιάζει στην επίδραση της ενσωμάτωσης τεχνικής λεπτομέρειας του βραχυπρόθεσμου προγραμματισμού. Η χρονική λεπτομέρεια τέθηκε σταθερή μεταξύ των δύο SM, ενώ η χωρική λεπτομέρεια δεν εξετάστηκε. Επιπλέον, μη θερμικές μονάδες που μπορούν να λειτουργήσουν και σαν πάροχοι λειτουργικής ευελιξίας είχαν συμπεριληφθεί ως επενδυτικές επιλογές και στις τρεις περιπτώσεις. Τα αποτελέσματα έδειξαν ότι η επιλογή ενός SM μπορεί να είναι επηρεάσει τις παραγόμενες λύσεις (επενδυτικές αποφάσεις), το εκτιμώμενο μείγμα παραγωγής και το συνολικό κόστος κάθε λύσης όταν αξιολογείται και από τα δύο SM. Οι διαφορές στο μείγμα παραγωγής προκύπτουν κυρίως από τη μειωμένη χρήση των παρόχων λειτουργικής ευελιξίας όταν μια εγκατεστημένη δυναμικότητα παραγωγής ισχύος είχε αξιολογηθεί από το SM το οποίο συμπεριελάμβανε χαμηλότερη τεχνική λεπτομέρεια. Αυτό είχε αποδοθεί στην παράλειψη τεχνικών περιορισμών που κατέστησαν λιγότερο απαραίτητη τη χρήση παρόχων λειτουργικής ευελιξίας.

Για το λόγο αυτό, η τεχνική λεπτομέρεια της βραχυπρόθεσμης λειτουργίας ενός συστήματος ισχύος που εισάγεται σε ένα SM για ένα μοντέλο GEP και χρησιμοποιείται για την εκτίμηση της απόδοσης μια υποψήφιας εγκατεστημένης δυναμικότητας παραγωγής ισχύος, μπορεί να είναι σημαντική για να εντοπιστούν αποτελεσματικά οι απαιτούμενες ανάγκες σε λειτουργική ευελιξία και να αξιολογηθούν κατάλληλα οι πάροχοι λειτουργικής ευελιξίας, όταν αυτοί θεωρούνται ως επενδυτικές επιλογές.

Το Κεφάλαιο 5 παρουσιάζει ένα στατικό GEP που βασίζεται σε MAEA για MOO. Είχε ως στόχο την ανάλυση αντικρουόμενων στόχων και συντελεστών κόστους που προκύπτουν για ένα MOO GEP. Οι τιμές των συντελεστών κόστους υπολογίζονταν με βάση ένα SM έτσι ώστε να αξιολογηθούν οι ανάγκες σε λειτουργική ευελιξία και να μελετηθεί το αντίκτυπό τους στους συντελεστές κόστους. Το SM είχε χρησιμοποιηθεί για να παρέχει έναν δείκτη της βραχυπρόθεσμης λειτουργίας του συστήματος ισχύος. AM χρησιμοποιούνται για να παρέχουν μια εκτίμηση του αποτελέσματος ενός SM για την μείωση του υπολογιστικού κόστους. Ο MAEA αναπτύχθηκε με βάση πλαίσια για βελτιστοποίηση με χρήση μεταπροτύπων και χωρίς χρήση παραγώγου.

Ο πρώτος στόχος ήταν να εξεταστεί η απόδοση του MAEA. Αυτό είχε υλοποιηθεί πραγματοποιώντας υπολογιστικά πειράματα. Τα προαναφερθέντα περιελάμβαναν συναρτήσεις αναφοράς MOO (χρησιμοποιήθηκαν τα προβλήματα ZDT με μικρές τροποποιήσεις). Τα αποτελέσματα που προέκυψαν σχετικά με την απόδοση του MAEA, λαμβάνοντας υπόψη τους υπολογιστικούς περιορισμούς, ήταν ικανοποιητικά. Συγκεκριμένα, είχε επιτευχθεί αξιοπρεπές PFA για τα προβλήματα ZDT1, ZDT2 και ZDT3. Ωστόσο, τα αποτελέσματα ήταν λιγότερο ικανοποιητικά στα προβλήματα ZDT4 και ZDT6. Αυτό αποδόθηκε στο χαμηλό UFA που επιτεύχθηκε. Συγκεκριμένα, είχε παρατηρηθεί ότι η ακρίβεια του UFA, εντός του προκαθορισμένου ορίου αξιολογήσεων της συνάρτησης, ήταν χαμηλή σε συναρτήσεις με μεγάλο αριθμό τοπικών ακροτάτων (multi-modal) που είχε οδηγήσει σε χαμηλής ακρίβειας PFA. Σε σύγκριση με τον αλγόριθμο βάσης, που χρησιμοποιήθηκε ως σημείο αναφοράς, είχε επιτευχθεί επιτάχυνση του ρυθμού σύγκλισης εάν η σύγκριση εστιαστεί στο υπολογιστικό κόστος των συναρτήσεων.

Κατά συνέπεια, ο MAEA που αναπτύχθηκε μπορεί να επιτύχει ένα αποδεκτό PFA εντός ενός προκαθορισμένου ορίου αξιολογήσεων συνάρτησης. Ωστόσο, για σύνθετα MOO (π.χ. συναρτήσεων που εμπεριέχουν ένα μεγάλου αριθμό τοπικών ακροτάτων) η επίτευξη αποδεκτού UFA καθίσταται εξαιρετικά δύσκολη. Αυτό μπορεί να αποτρέψει την επίτευξη ενός αποδεκτού PFA.

Εκτός από τα προβλήματα που βασίζονταν στις συναρτήσεις αναφοράς, πραγματοποιήθηκε μια σειρά υπολογιστικών πειραμάτων σε παραλλαγές MOO-GEP. Πέντε διαφορετικά προβλήματα MOO-GEP είχαν διατυπωθεί με διαφορετικά ζεύγη αντικειμενικών συναρτήσεων. Το SM και οι συναρτήσεις περιορισμού του μακροχρόνιου προγραμματισμού περιλαμβάνονται σε κάθε παραλλαγή αυτών των MOO-GEP. Επιπλέον, εξετάστηκε μια έκτη περίπτωση που αμελούσε τη βραχυπρόθεσμη λειτουργία του συστήματος ισχύος και αντιστοιχεί σε ένα υπολογιστικά φθηνό MOO πρόβλημα. Η προσέγγιση βελτιστοποίησης είχε εφαρμοστεί σε μια απλοποιημένη μορφή ενός πραγματικού τομέα ισχύος, που είχε λάβει υπόψη την εγκατεστημένη δυναμικότητα παραγωγής ισχύος και σε μια δεύτερη όπου έγινε η θεώρηση μηδενικής εγκατεστημένη δυναμικότητα παραγωγής ισχύος (greenfield case). Τα υπολογιστικά πειράματα για τις πέντε παραλλαγές MOO-GEP είχαν εξεταστεί για δύο διαφορετικά επίπεδα χρονικής λεπτομέρειας για τη βραχυπρόθεσμη λειτουργία του συστήματος ισχύος. Συγκεκριμένα, είχαν επιλεγεί (i) μια αντιπροσωπευτική ημέρα και (ii) τέσσερις αντιπροσωπευτικές εβδομάδες. Η χωρική και τεχνική λεπτομέρεια είχε παραμένει σταθερή σε όλες τις περιπτώσεις.

Τα αποτελέσματα για τις δύο περιπτώσεις (υφιστάμενη και μηδενικής εγκατεστημένη δυναμικότητα παραγωγής ισχύος) σχετικά με την απόδοση της μεθόδου βελτιστοποίησης στις υπολογιστικά φθηνές περιπτώσεις υποδηλώνουν ότι είχε επιτευχθεί ένα αποδεκτό PFA. Αυτά ήταν συγκρίσιμα με έναν αλγόριθμο βάσης, που έχει αναπτυχτεί για προβλήματα βελτιστοποίησης με υπολογιστικά φθηνές συναρτήσεις, παρά τους υπολογιστικούς περιορισμούς που επιβλήθηκαν. Επιπλέον, επιτεύχθηκε ένα αποδεκτό UFA με βάση τους δείκτες που χρησιμοποιήθηκαν για την αξιολόγηση της ακρίβειάς του. Συγκρίνοντας τα αποτελέσματα για τις δύο περιπτώσεις (υφιστάμενη και μηδενικής εγκατεστημένη δυναμικότητα παραγωγής ισχύος), παρατηρήθηκε ότι η απόδοση του αλγορίθμου δεν επηρεάστηκε σημαντικά με βάση τους δείκτες που είχαν χρησιμοποιηθεί για την αξιολόγηση της απόδοσης των αλγορίθμων.

Ως εκ τούτου, η μέθοδος βασισμένη σε MAEA μπορεί να επιτύχει ένα αποδεκτό PFA εντός ενός προκαθορισμένου ορίου αξιολογήσεων της συνάρτησης για ένα υπολογιστικά φθηνό μοντέλο MOO-GEP. Ωστόσο, για ένα υπολογιστικά φθηνό μοντέλο GEP θα μπορούσε επίσης να χρησιμοποιηθεί ένας MOEA, καθώς οι υπολογιστικές ανάγκες δεν είναι δεσμευτικές.

Τα αποτελέσματα για τις δέκα περιπτώσεις (υφιστάμενη και μηδενικής εγκατεστημένη δυναμικότητα παραγωγής ισχύος) υποδηλώνουν ότι είχε επιτευχθεί ένα αποδεκτό PFA σχετικά με την απόδοση της βελτιστοποίησης με χρήση του MAEA στις υπολογιστικά φθηνές περιπτώσεις, συμπεριλαμβανομένου ενός SM με περιορισμένη χρονική λεπτομέρεια. Δεδομένων των υπολογιστικών περιορισμών που επιβλήθηκαν, η σειριακή πρόοδος των δεικτών βασισμένων στον υπερόγκο υποδηλώνουν μια προοδευτική αύξηση της ακρίβειας του PFA. Ωστόσο, η συμπερίληψη του SM μπορεί να επηρεάσει την απόδοση του αλγορίθμου. Αυτό παρατηρήθηκε στις περιπτώσεις μηδενικής εγκατεστημένης δυναμικότητας παραγωγής ισχύος καθώς είχε επιτευχθεί σχετικά χαμηλότερης ακρίβειας UFA, ενώ για τις περιπτώσεις συμπεριλαμβανομένης της υφιστάμενης δυναμικότητας αυτή ήταν αποδεκτή.

Συνεπώς, ο αλγόριμθος MAEA για MOO που εξετάστηκε μπορεί να επιτύχει ένα αποδεκτό PFA μέσα σε ένα προκαθορισμένο όριο αξιολογήσεων των συναρτήσεων στόχου για ένα υπολογιστικά φθηνό μοντέλο MOO-GEP, το οποίο περιλαμβάνει ένα SM με περιορισμένη χρονική λεπτομέρεια. Ωστόσο, η συμπερίληψη ενός SM μπορεί να επηρεάσει την ακρίβεια του UFA. Ένας από τους κύριους παράγοντες που εντοπίστηκαν είναι η τεχνική χειρισμού των περιορισμών του SM (συνάρτηση Ποινής) που χρησιμοποιείται για την αντιμετώπιση της περίπτωση ανεπαρκούς εγκατεστημένης δυναμικότητας παραγωγής ισχύος, δεδομένου ότι οι όροι ποινής μπορούν να παρουσιάζουν σχετικά υψηλές τιμές κόστους.

Οι τελικές δέκα περιπτώσεις (υφιστάμενη και μηδενικής εγκατεστημένη δυναμικότητα παραγωγής ισχύος) αποτελούν ένα πιο ολοκληρωμένο παράδειγμα εφαρμογής του MAEA καθώς το SM παρουσιάζει αυξημένη χρονική λεπτομέρεια. Τα αποτελέσματα σχετικά με την απόδοση του MAEA είχαν κριθεί ικανοποιητικά. Συγκεκριμένα, η σειριακή πρόοδος των δεικτών που είχαν χρησιμοποιηθεί για την αξιολόγηση του PFA υποδηλώνουν ότι η απόδοση θα μπορούσε να είναι αποδεκτή δεδομένων των υπολογιστικών περιορισμών. Επίσης, οι δείκτες υποδηλώνουν ότι αύξηση του αριθμού των διαθέσιμων προσομοιώσεων (υπολογιστικό κόστος) θα μπορούσε να είχε οδηγήσει σε βελτίωση της ακρίβεια του PFA σε ορισμένες από τις περιπτώσεις. Επίσης, τα αποτελέσματα είχαν δείξει ότι οι αριθμητικές διαφορές που προέκυψαν λόγω της συμπερίληψης του συντελεστή κόστους ποινής είχε επηρεάσει την ακρίβεια του UFA. Αυτό είχε παρατηρηθεί για όλες τις περιπτώσεις που εξετάστηκαν εκτός από αυτές που περιλαμβάνουν την ποινή ως ξεχωριστή αντικειμενική συνάρτηση. Επιπλέον, οι δείκτες που βασίζονται στον υπερόγκο και η οπτική ανάλυση του PFA υποδηλώνουν ότι η μείωση της ακρίβειας του UFA δεν είχε αποτρέψει την επίτευξη ενός αποδεκτού PFA, με βάση και τους υπολογιστικούς περιορισμούς.

Συνεπώς, ο MAEA μπορεί να αποτελεί υποσχόμενη προσέγγιση για την ανάλυση των αντικρουόμενων στόχων και συντελεστών κόστους σε MOO-GEP, που συμπεριλαμβάνουν SM της βραχυπρόθεσμης λειτουργίας ενός συστήματος ισχύος για την αξιολόγηση της λειτουργικής ευελιξίας. Ωστόσο, αναφέρεται ότι το χαμηλό υπολογιστικό κόστος και η ακρίβεια του PFA μπορούν επίσης να είναι αντικρουόμενα.

Τα αποτελέσματα των πέντε περιπτώσεων που λάμβαναν υπόψη την υφιστάμενη εγκατεστημένη δυναμικότητα παραγωγής ισχύος χρησιμοποιήθηκαν για την εξέταση των συντελεστών κόστους που είχαν προκύψει για τους συνδυασμούς συναρτήσεων στόχων που είχαν οριστεί. Η ανάλυση βασίστηκε στο παραγόμενο το μέτωπο μη κυριαρχούμενων λύσεων υποθέτοντας ότι αυτό αποτελεί μια αποδεκτή PFA. Τα αποτελέσματα της ανάλυσης αφορούν την επίδραση των συντελεστών κόστους στις συναρτήσεις στόχου που θεωρήθηκαν και στο μέτωπο μη κυριαρχούμενων λύσεων που επιτεύχθηκε

Επομένως, η ανάλυση της επίδρασης των συντελεστών κόστους των αντικρουόμενων συναρτήσεων στόχου ενός MOO-GEP θα μπορούσε να παρέχει μια λεπτομερή αξιολόγηση των προσθηκών δυναμικότητας σε επίπεδο συντελεστών κόστους.

Bibliography

[1] H. Sadeghi, M. Rashidinejad, A. Abdollahi, A comprehensive sequential review study through the generation expansion planning, Renewable and Sustainable Energy Reviews 67 (2017) 1369–1394.

[2] N. E. Koltsaklis, A. S. Dagoumas, State-of-the-art generation expansion planning: A review, Applied Energy 230 (2018) 563–589.

[3] A. S. Dagoumas, N. E. Koltsaklis, Review of models for integrating renewable energy in the generation expansion planning, Applied Energy 242 (2019) 1573–1587.

[4] V. Oree, S. Z. S. Hassen, P. J. Fleming, Generation expansion planning optimisation with renewable energy integration: A review, Renewable and Sustainable Energy Reviews 69 (2017) 790–803.

[5] L. Gacitua, P. Gallegos, R. Henriquez-Auba, A. Lorca, M. Negrete-Pincetic, D. Olivares, A. Valenzuela, G. Wenzel, A comprehensive review on expansion planning: Models and tools for energy policy analysis, Renewable and Sustainable Energy Reviews 98 (2018) 346–360.

[6] S. Pereira, P. Ferreira, A. I. F. Vaz, Optimization modeling to support renewables integration in power systems, Renewable and Sustainable Energy Reviews 55 (2016) 316–325.

[7] B. Palmintier, M. Webster, Impact of unit commitment constraints on generation expansion planning with renewables, in: Power and energy society general meeting, 2011 IEEE, IEEE, 2011, pp. 1–7.

[8] B. S. Palmintier, M. D. Webster, Impact of operational flexibility on electricity generation planning with renewable and carbon targets, IEEE Transactions on Sustainable Energy 7 (2) (2016) 672–684.

[9] S. Collins, J. P. Deane, K. Poncelet, E. Panos, R. C. Pietzcker, E. Delarue, B. P. O. Gallachoir, Integrating short term variations of the power system into integrated energy system models: A methodological review, Renewable and Sustainable Energy Reviews 76 (2017) 839–856.

[10] A. Shortt, J. Kiviluoma, M. O'Malley, Accommodating variability in generation planning, IEEE Transactions on Power Systems 28 (1) (2012) 158–169.

[11] R. Mallipeddi, P. N. Suganthan, Unit commitment-a survey and comparison of conventional and nature inspired algorithms, International Journal of Bio-Inspired Computation 6 (2) (2014) 71–90.

[12] K. Deb, Multi-objective optimization using evolutionary algorithms, Vol. 16, John Wiley & Sons, 2001.

[13] A. Zhou, B.-Y. Qu, H. Li, S.-Z. Zhao, P. N. Suganthan, Q. Zhang, Multiobjective evolutionary algorithms: A survey of the state of the art, Swarm and Evolutionary Computation 1 (1) (2011) 32–49.

[14] Y. Jin, Surrogate-assisted evolutionary computation: Recent advances and future challenges, Swarm and Evolutionary Computation 1 (2) (2011) 61–70.

[15] A. Díaz-Manríquez, G. Toscano, J. H. Barron-Zambrano, E. Tello-Leal, A review of surrogate assisted multiobjective evolutionary algorithms, Computational intelligence and neuroscience 2016 (2016).

[16] Y. Jin, B. Sendhoff, Fitness approximation in evolutionary computation-a survey, in: Proceedings of the 4th Annual Conference on Genetic and Evolutionary Computation, 2002, pp. 1105–1112.

[17] C. Vrionis, V. Tsalavoutis, A. Tolis, A generation expansion planning model for integrating high shares of renewable energy: A meta-model assisted evolutionary algorithm approach, Applied Energy 259 (2020) 114085.

[18] A. J. Conejo, L. Baringo, S. J. Kazempour, A. S. Siddiqui, Investment in electricity generation and transmission, Cham Zug, Switzerland: Springer International Publishing 106 (2016).

[19] Y. Cui, Z. Geng, Q. Zhu, Y. Han, Multi-objective optimization methods and application in energy saving, Energy 125 (2017) 681–704.

[20] C. I. Nweke, F. Leanez, G. R. Drayton, M. Kolhe, Benefits of chronological optimization in capacity planning for electricity markets, in: 2012 IEEE International Conference on Power System Technology (POWERCON), IEEE, 2012, pp. 1–6.

[21] C. Batlle, P. Rodilla, An enhanced screening curves method for considering thermal cycling operation costs in generation expansion planning, IEEE transactions on power systems 28 (4) (2013) 3683–3691.

[22] T. Zhang, R. Baldick, T. Deetjen, Optimized generation capacity expansion using a further improved screening curve method, Electric Power Systems Research 124 (2015) 47–54.

[23] F. Ueckerdt, R. Brecha, G. Luderer, P. Sullivan, E. Schmid, N. Bauer, D. Böttger, R. Pietzcker, Representing power sector variability and the integration of variable renewables in long-term energy-economy models using residual load duration curves, Energy 90 (2015) 1799–1814.

[24] A. Franz, J. Rieck, J. Zimmermann, Fix-and-optimize procedures for solving the long-term unit commitment problem with pumped storages, Annals of Operations Research (2018) 1–25.

[25] N. H. Kjeldsen, M. Chiarandini, Heuristic solutions to the long-term unit commitment problem with cogeneration plants, Computers & Operations Research 39 (2) (2012) 269–282.

[26] X. Han, X. Chen, M. B. McElroy, S. Liao, C. P. Nielsen, J. Wen, Modeling formulation and validation for accelerated simulation and flexibility assessment on large scale power systems under higher renewable penetrations, Applied energy 237 (2019) 145–154.

[27] J. Blazquez, R. Fuentes-Bracamontes, C. A. Bollino, N. Nezamuddin, The renewable energy policy paradox, Renewable and Sustainable Energy Reviews 82 (2018) 1–5.

[28] G. Luderer, V. Krey, K. Calvin, J. Merrick, S. Mima, R. Pietzcker, J. Van Vliet, K. Wada, The role of renewable energy in climate stabilization: results from the emf27 scenarios, Climatic change 123 (3-4) (2014) 427–441.

[29] C. De Jonghe, E. Delarue, R. Belmans, W. D'haeseleer, Determining optimal electricity technology mix with high level of wind power penetration, Applied Energy 88 (6) (2011) 2231–2238.

[30] G. Papaefthymiou, K. Dragoon, Towards 100% renewable energy systems: Uncapping power system flexibility, Energy Policy 92 (2016) 69–82.

[31] J. N. Puga, The importance of combined cycle generating plants in integrating large levels of wind power generation, The Electricity Journal 23 (7) (2010) 33–44.

[32] J. Gil, A. Caballero, A. J. Conejo, Power cycling: Ccgts: The critical link between the electricity and natural gas markets, IEEE Power and Energy Magazine 12 (6) (2014) 40–48.

[33] M. Alizadeh, M. P. Moghaddam, N. Amjady, P. Siano, M. Sheikh-El-Eslami, Flexibility in future power systems with high renewable penetration: A review, Renewable and Sustainable Energy Reviews 57 (2016) 1186–1193.

[34] A. van Stiphout, K. De Vos, G. Deconinck, The impact of operating reserves on investment planning of renewable power systems, IEEE Transactions on Power Systems 32 (1) (2017) 378–388.

[35] M. Welsch, P. Deane, M. Howells, B. Ó. Gallachóir, F. Rogan, M. Bazilian, H.-H. Rogner, Incorporating flexibility requirements into long-term energy system models–a case study on high levels of renewable electricity penetration in ireland, Applied Energy 135 (2014) 600–615.

[36] S. Ludig, M. Haller, E. Schmid, N. Bauer, Fluctuating renewables in a long-term climate change mitigation strategy, Energy 36 (11) (2011) 6674–6685.

[37] A. Pina, C. Silva, P. Ferrão, Modeling hourly electricity dynamics for policy making in long-term scenarios, Energy Policy 39 (9) (2011) 4692–4702.

[38] K. Poncelet, E. Delarue, D. Six, J. Duerinck, W. D'haeseleer, Impact of the level of temporal and operational detail in energy-system planning models, Applied Energy 162 (2016) 631–643.

[39] J. Deane, A. Chiodi, M. Gargiulo, B. P. Ó. Gallachóir, Soft-linking of a power systems model to an energy systems model, Energy 42 (1) (2012) 303–312.

[40] K. Poncelet, H. Höschle, E. Delarue, A. Virag, W. D'haeseleer, Selecting representative days for capturing the implications of integrating intermittent renewables in generation expansion planning problems, IEEE Transactions on Power Systems 32 (3) (2017) 1936–1948.

[41] V. Krishnan, W. Cole, Evaluating the value of high spatial resolution in national capacity expansion models using reeds, in: 2016 IEEE Power and Energy Society General Meeting (PESGM), IEEE, 2016, pp. 1–5.

[42] N. E. Koltsaklis, A. S. Dagoumas, G. M. Kopanos, E. N. Pistikopoulos, M. C. Georgiadis, A spatial multi-period long-term energy planning model: a case study of the greek power system, Applied Energy 115 (2014) 456–482.

[43] B. Hua, R. Baldick, J. Wang, Representing operational flexibility in generation expansion planning through convex relaxation of unit commitment, IEEE Transactions on Power Systems 33 (2) (2018) 2272–2281.

[44] A. Belderbos, E. Delarue, Accounting for flexibility in power system planning with renewables, International Journal of Electrical Power & Energy Systems 71 (2015) 33–41.

[45] X. Chen, J. Lv, M. B. McElroy, X. Han, C. P. Nielsen, J. Wen, Power system capacity expansion under higher penetration of renewables considering flexibility constraints and low carbon policies, IEEE Transactions on Power Systems (2018).

[46] S. Pereira, P. Ferreira, A. I. F. Vaz, Generation expansion planning with high share of renewables of variable output, Applied Energy 190 (2017) 1275–1288.

[47] N. E. Koltsaklis, M. C. Georgiadis, A multi-period, multi-regional generation expansion planning model incorporating unit commitment constraints, Applied energy 158 (2015) 310–331.

[48] H. Park, R. Baldick, Multi-year stochastic generation capacity expansion planning under environmental energy policy, Applied energy 183 (2016) 737–745.

[49] A. Flores-Quiroz, R. Palma-Behnke, G. Zakeri, R. Moreno, A column generation approach for solving generation expansion planning problems with high renewable energy penetration, Electric Power Systems Research 136 (2016) 232–241.

[50] C. L. Lara, D. S. Mallapragada, D. J. Papageorgiou, A. Venkatesh, I. E. Grossmann, Deterministic electric power infrastructure planning: Mixed-integer programming model and nested decomposition algorithm, European Journal of Operational Research 271 (3) (2018) 1037–1054.

[51] J. Ma, V. Silva, R. Belhomme, D. S. Kirschen, L. F. Ochoa, Evaluating and planning flexibility in sustainable power systems, in: Power and Energy Society General Meeting (PES), 2013 IEEE, IEEE, 2013, pp. 1–11.

[52] V. Oree, S. Z. S. Hassen, A composite metric for assessing flexibility available in conventional generators of power systems, Applied energy 177 (2016) 683–691.

[53] I. F. Abdin, E. Zio, An integrated framework for operational flexibility assessment in multi-period power system planning with renewable energy production, Applied energy 222 (2018) 898–914.

[54] N. E. Koltsaklis, A. S. Dagoumas, M. C. Georgiadis, G. Papaioannou, C. Dikaiakos, A mid-term, market-based power systems planning model, Applied Energy 179 (2016) 17–35.

[55] T. Levin, A. Botterud, Electricity market design for generator revenue sufficiency with increased variable generation, Energy Policy 87 (2015) 392–406.

[56] C. F. Heuberger, I. Staffell, N. Shah, N. Mac Dowell, A systems approach to quantifying the value of power generation and energy storage technologies in future electricity networks, Computers & Chemical Engineering 107 (2017) 247–256.

[57] T. Brijs, A. van Stiphout, S. Siddiqui, R. Belmans, Evaluating the role of electricity storage by considering short-term operation in long-term planning, Sustainable Energy, Grids and Networks 10 (2017) 104–117.

[58] T. Luz, P. Moura, A. de Almeida, Multi-objective power generation expansion planning with high penetration of renewables, Renewable and Sustainable Energy Reviews 81 (2018) 2637–2643.

[59] P. S. Moura, A. T. de Almeida, Multi-objective optimization of a mixed renewable system with demand-side management, Renewable and Sustainable Energy Reviews 14 (5) (2010) 1461–1468.

[60] J. Aghaei, M. A. Akbari, A. Roosta, A. Baharvandi, Multiobjective generation expansion planning considering power system adequacy, Electric power systems research 102 (2013) 8–19.

[61] V. Oree, S. Z. S. Hassen, P. J. Fleming, A multi-objective framework for long-term generation expansion planning with variable renewables, Applied Energy 253 (2019) 113589.

[62] J. Zhu, M.-y. Chow, A review of emerging techniques on generation expansion planning, IEEE Transactions on Power Systems 12 (4) (1997) 1722–1728.

[63] S. Kannan, S. M. R. Slochanal, N. P. Padhy, Application and comparison of metaheuristic techniques to generation expansion planning problem, IEEE Transactions on Power Systems 20 (1) (2005) 466–475.

[64] A. J. Pereira, J. T. Saraiva, A long term generation expansion planning model using system dynamics–case study using data from the portuguese/spanish generation system, Electric Power Systems Research 97 (2013) 41–50.

[65] K. Rajesh, A. Bhuvanesh, S. Kannan, C. Thangaraj, Least cost generation expansion planning with solar power plant using differential evolution algorithm, Renewable Energy 85 (2016) 677–686.

[66] P. Verma, K. Sanyal, D. Srinivsan, K. Swarup, Information exchange based clustered differential evolution for constrained generation-transmission expansion planning, Swarm and evolutionary computation 44 (2019) 863–875.

[67] R. Hemmati, H. Saboori, M. A. Jirdehi, Multistage generation expansion planning incorporating large scale energy storage systems and environmental pollution, Renewable Energy 97 (2016) 636–645.

[68] V. A. Tsalavoutis, C. G. Vrionis, A. I. Tolis, Relaxation of quantitative energy objectives on generation expansion planning: A computational and policy study, International Transactions on Electrical Energy Systems 27 (12) (2017) e2427.

[69] J. Sirikum, A. Techanitisawad, V. Kachitvichyanukul, A new efficient ga-benders' decomposition method: For power generation expansion planning with emission controls, IEEE Transactions on Power Systems 22 (3) (2007) 1092–1100.

[70] S. Kannan, S. Baskar, J. D. McCalley, P. Murugan, Application of NSGA-II algorithm to generation expansion planning, IEEE Transactions on Power systems 24 (1) (2009) 454–461.

[71] P. Murugan, S. Kannan, S. Baskar, NSGA-II algorithm for multi-objective generation expansion planning problem, Electric Power Systems Research 79 (4) (2009) 622–628.

[72] P. Murugan, S. Kannan, S. Baskar, Application of nsga-ii algorithm to single-objective transmission constrained generation expansion planning, IEEE Transactions on Power Systems 24 (4) (2009) 1790–1797.

[73] C. A. Georgopoulou, K. C. Giannakoglou, Metamodel-assisted evolutionary algorithms for the unit commitment problem with probabilistic outages, Applied Energy 87 (5) (2010) 1782–1792.

[74] A. Glotić, A. Zamuda, Short-term combined economic and emission hydrothermal optimization by surrogate differential evolution, Applied Energy 141 (2015) 42–56.

[75] C. Darwin, On the origin of species london, UK: John Murray 62 (1859).

[76] Z. Michalewicz, D. B. Fogel, How to solve it: modern heuristics, Springer Science & Business Media, 2013.

[77] A. E. Eiben, J. E. Smith, et al., Introduction to evolutionary computing, Vol. 53, Springer, 2003.

[78] J. H. Holland, Genetic algorithms and the optimal allocation of trials, SIAM Journal on Computing 2 (2) (1973) 88–105.

[79] J. H. Holland, et al., Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence, MIT press, 1992.

[80] H.-P. Schwefel, Kybernetische evolution als strategie der experimentellen forschung in der stromungstechnik, Diploma thesis, Technical Univ. of Berlin (1965).

[81] I. Rechenberg, Evolution strategy: Optimization of technical systems by means of biological evolution, Fromman-Holzboog, Stuttgart 104 (1973) 15–16.

[82] R. Storn, K. Price, Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces, Journal of global optimization 11 (4) (1997) 341–359.

[83] K. Price, R. M. Storn, J. A. Lampinen, Differential evolution: a practical approach to global optimization, Springer Science & Business Media, 2006.

[84] Z. Michalewicz, A survey of constraint handling techniques in evolutionary computation methods., Evolutionary programming 4 (1995) 135–155.

[85] C. A. C. Coello, Theoretical and numerical constraint-handling techniques used with evolutionary algorithms: a survey of the state of the art, Computer methods in applied mechanics and engineering 191 (11-12) (2002) 1245–1287.

[86] E. Mezura-Montes, C. A. C. Coello, Constraint-handling in nature-inspired numerical optimization: past, present and future, Swarm and Evolutionary Computation 1 (4) (2011) 173–194.

[87] A. Goicoechea, D. R. Hansen, L. Duckstein, Multiobjective decision analysis with engineering and business applications, Tech. rep., John Wiley & Sons (1982).

[88] K. Deb, Multi-objective optimisation using evolutionary algorithms: an introduction, in: Multi-objective evolutionary optimisation for product design and manufacturing, Springer, 2011, pp. 3–34.

[89] K. Deb, A. Pratap, S. Agarwal, T. Meyarivan, A fast and elitist multiobjective genetic algorithm: Nsga-ii, IEEE transactions on evolutionary computation 6 (2) (2002) 182–197.

[90] E. Zitzler, M. Laumanns, L. Thiele, Spea2: Improving the strength pareto evolutionary algorithm, TIK-report 103 (2001).

[91] J. D. Knowles, D. W. Corne, Approximating the nondominated front using the pareto archived evolution strategy, Evolutionary computation 8 (2) (2000) 149–172.

[92] Q. Zhang, H. Li, Moea/d: A multiobjective evolutionary algorithm based on decomposition, IEEE Transactions on evolutionary computation 11 (6) (2007) 712–731.

[93] E. Zitzler, S. Künzli, Indicator-based selection in multiobjective search, in: International conference on parallel problem solving from nature, Springer, 2004, pp. 832–842.

[94] J. Bader, E. Zitzler, Hype: An algorithm for fast hypervolume-based many-objective optimization, Evolutionary computation 19 (1) (2011) 45–76.

[95] K. Deb, H. Jain, An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part i: solving problems with box constraints, IEEE transactions on evolutionary computation 18 (4) (2013) 577–601.

[96] P. P. Bonissone, R. Subbu, N. Eklund, T. R. Kiehl, Evolutionary algorithms+ domain knowledge= real-world evolutionary computation, IEEE Transactions on Evolutionary Computation 10 (3) (2006) 256–280.

[97] Z. Michalewicz, Genetic algorithms+ data structures= evolution programs, Springer Science & Business Media, 2013.

[98] P. Moscato, et al., On evolution, search, optimization, genetic algorithms and martial arts: Towards memetic algorithms, Caltech concurrent computation program, C3P Report 826 (1989) 1989.

[99] Y. Jin, A comprehensive survey of fitness approximation in evolutionary computation, Soft computing 9 (1) (2005) 3–12.

[100] A. Forrester, A. Sobester, A. Keane, Engineering design via surrogate modelling: a practical guide, John Wiley & Sons, 2008.

[101] D. R. Jones, A taxonomy of global optimization methods based on response surfaces, Journal of global optimization 21 (4) (2001) 345–383.

[102] Y. Jin, M. Olhofer, B. Sendhoff, A framework for evolutionary optimization with approximate fitness functions, IEEE Transactions on evolutionary computation 6 (5) (2002) 481–494.

[103] L. Shi, K. Rasheed, A survey of fitness approximation methods applied in evolutionary algorithms, in: Computational intelligence in expensive optimization problems, Springer, 2010, pp. 3–28.

[104] M. Sefrioui, J. Périaux, A hierarchical genetic algorithm using multiple models for optimization, in: International Conference on Parallel Problem Solving From Nature, Springer, 2000, pp. 879–888.

[105] K. Rasheed, H. Hirsh, Informed operators: Speeding up genetic-algorithm-based design optimization using reduced models, in: Proceedings of the 2nd Annual Conference on Genetic and Evolutionary Computation, 2000, pp. 628–635.

[106] K. C. Giannakoglou, A. P. Giotis, M. K. Karakasis, Low-cost genetic optimization based on inexact pre-evaluations and the sensitivity analysis of design parameters, Inverse Problems in Engineering 9 (4) (2001) 389–412.

[107] M. Emmerich, A. Giotis, M. Özdemir, T. Bäck, K. Giannakoglou, Metamodel—assisted evolution strategies, in: International Conference on parallel problem solving from nature, Springer, 2002, pp. 361–370.

[108] D. Lim, Y.-S. Ong, Y. Jin, B. Sendhoff, A study on metamodeling techniques, ensembles, and multi-surrogates in evolutionary computation, in: Proceedings of the 9th annual conference on Genetic and evolutionary computation, 2007, pp. 1288–1295.

[109] A. Díaz-Manríquez, G. Toscano-Pulido, W. Gómez-Flores, On the selection of surrogate models in evolutionary optimization algorithms, in: Evolutionary Computation (CEC), 2011 IEEE Congress on, IEEE, 2011, pp. 2155–2162.

[110] A. Díaz-Manríquez, G. Toscano, C. A. C. Coello, Comparison of metamodeling techniques in evolutionary algorithms, Soft Computing 21 (19) (2017) 5647–5663.

[111] D. Lim, Y. Jin, Y.-S. Ong, B. Sendhoff, Generalizing surrogate-assisted evolutionary computation, IEEE Transactions on Evolutionary Computation 14 (3) (2009) 329–355.

[112] Y. Wang, D.-Q. Yin, S. Yang, G. Sun, Global and local surrogate-assisted differential evolution for expensive constrained optimization problems with inequality constraints, IEEE transactions on cybernetics 49 (5) (2018) 1642–1656.

[113] Y.-S. Ong, Z. Zhou, D. Lim, Curse and blessing of uncertainty in evolutionary algorithm using approximation, in: 2006 IEEE International Conference on Evolutionary Computation, IEEE, 2006, pp. 2928–2935.

[114] R. Mallipeddi, M. Lee, An evolving surrogate model-based differential evolution algorithm, Applied Soft Computing 34 (2015) 770–787.

[115] E. Mezura-Montes, C. A. C. Coello, Constrained optimization via multiobjective evolutionary algorithms, in: Multiobjective problem solving from nature, Springer, 2008, pp. 53–75.

[116] T. P. Runarsson, Constrained evolutionary optimization by approximate ranking and surrogate models, in: International Conference on Parallel Problem Solving from Nature, Springer, 2004, pp. 401–410.

[117] R. G. Regis, Evolutionary programming for high-dimensional constrained expensive black-box optimization using radial basis functions, IEEE Transactions on Evolutionary Computation 18 (3) (2013) 326–347.

[118] Z. Zhou, Y. S. Ong, P. B. Nair, A. J. Keane, K. Y. Lum, Combining global and local surrogate models to accelerate evolutionary optimization, IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) 37 (1) (2006) 66–76.

[119] Y. S. Ong, P. B. Nair, A. J. Keane, Evolutionary optimization of computationally expensive problems via surrogate modeling, AIAA journal 41 (4) (2003) 687–696.

[120] H.-M. Gutmann, A radial basis function method for global optimization, Journal of global optimization 19 (3) (2001) 201–227.

[121] D. R. Jones, M. Schonlau, W. J. Welch, Efficient global optimization of expensive black-box functions, Journal of Global optimization 13 (4) (1998) 455–492.

[122] R. G. Regis, C. A. Shoemaker, A stochastic radial basis function method for the global optimization of expensive functions, INFORMS Journal on Computing 19 (4) (2007) 497–509.

[123] R. G. Regis, C. A. Shoemaker, Constrained global optimization of expensive black box functions using radial basis functions, Journal of Global optimization 31 (1) (2005) 153–171.

[124] J. Müller, C. A. Shoemaker, R. Piché, So-i: a surrogate model algorithm for expensive nonlinear integer programming problems including global optimization applications, Journal of Global Optimization 59 (4) (2014) 865–889.

[125] J. Müller, C. A. Shoemaker, R. Piché, So-mi: A surrogate model algorithm for computationally expensive nonlinear mixed-integer black-box global optimization problems, Computers & Operations Research 40 (5) (2013) 1383–1400.

[126] J. Knowles, Parego: a hybrid algorithm with on-line landscape approximation for expensive multiobjective optimization problems, IEEE Transactions on Evolutionary Computation 10 (1) (2006) 50–66.

[127] Q. Zhang, W. Liu, E. Tsang, B. Virginas, Expensive multiobjective optimization by moea/d with gaussian process model, IEEE Transactions on Evolutionary Computation 14 (3) (2009) 456–474.

[128] S. Zapotecas Martínez, C. A. Coello Coello, Moea/d assisted by rbf networks for expensive multi-objective optimization problems, in: Proceedings of the 15th annual conference on Genetic and evolutionary computation, 2013, pp. 1405–1412.

[129] M. T. Emmerich, K. C. Giannakoglou, B. Naujoks, Single-and multiobjective evolutionary optimization assisted by gaussian random field metamodels, IEEE Transactions on Evolutionary Computation 10 (4) (2006) 421–439.

[130] R. G. Regis, Multi-objective constrained black-box optimization using radial basis function surrogates, Journal of computational science 16 (2016) 140–155.

[131] C. A. Georgopoulou, K. C. Giannakoglou, Multiobjective metamodel–assisted memetic algorithms, in: Multi-objective memetic algorithms, Springer, 2009, pp. 153–181.

[132] T. Akhtar, C. A. Shoemaker, Multi objective optimization of computationally expensive multi-modal functions with rbf surrogates and multi-rule selection, Journal of Global Optimization 64 (1) (2016) 17–32.

[133] T. Akhtar, C. A. Shoemaker, Efficient multi-objective optimization through population-based parallel surrogate search, arXiv preprint arXiv:1903.02167 (2019).

[134] J. Meus, K. Poncelet, E. Delarue, Applicability of a clustered unit commitment model in power system modeling, IEEE Transactions on Power Systems 33 (2) (2018) 2195–2204.

[135] K. Deb, An efficient constraint handling method for genetic algorithms, Computer methods in applied mechanics and engineering 186 (2-4) (2000) 311–338.

[136] S. Das, P. N. Suganthan, Differential evolution: A survey of the state-of-the-art, IEEE transactions on evolutionary computation 15 (1) (2010) 4–31.

[137] S. Das, S. S. Mullick, P. N. Suganthan, Recent advances in differential evolution–an updated survey, Swarm and Evolutionary Computation 27 (2016) 1–30.

[138] V. A. Tsalavoutis, C. G. Vrionis, A. I. Tolis, Optimizing a unit commitment problem using an evolutionary algorithm and a plurality of priority lists, Operational Research (2018) 1–54.

[139] V. Tsalavoutis, C. Vrionis, A. Tolis, A hybrid multi-objective evolutionary algorithm for economic-environmental generation scheduling, in: Proceedings of the Genetic and Evolutionary Computation Conference, 2019, pp. 1338–1346.

[140] G. Onwubolu, D. Davendra, Scheduling flow shops using differential evolution algorithm, European Journal of Operational Research 171 (2) (2006) 674–692.

[141] J. Lampinen, I. Zelinka, Mixed integer-discrete-continuous optimization by differential evolution, in: Proceedings of the 5th International Conference on Soft Computing, 1999, pp. 71–76.

[142] R. Tanabe, A. Fukunaga, Reevaluating exponential crossover in differential evolution, in: International Conference on Parallel Problem Solving from Nature, Springer, 2014, pp. 201–210.

[143] R. G. Regis, Constrained optimization by radial basis function interpolation for high-dimensional expensive black-box problems with infeasible initial points, Engineering Optimization 46 (2) (2014) 218–243.

[144] M. Powell, The theory of radial basis function approximation in 1990, Advances in Numerical Analysis II: Wavelets, Subdivision, and Radial Functions (WA Light, ed.), Oxford University Press, Oxford 105 (1992) 210.

[145] H. Jain, K. Deb, An evolutionary many-objective optimization algorithm using reference-point based nondominated sorting approach, part ii: handling constraints and extending to an adaptive approach, IEEE Transactions on evolutionary computation 18 (4) (2013) 602–622.

[146] C. A. Georgopoulou, K. C. Giannakoglou, A multi-objective metamodel-assisted memetic algorithm with strength-based local refinement, Engineering optimization 41 (10) (2009) 909–923.

[147] M. K. Karakasis, K. C. Giannakoglou, On the use of metamodel-assisted, multi-objective evolutionary algorithms, Engineering Optimization 38 (8) (2006) 941–957.

[148] S. Z. Martínez, C. A. C. Coello, A memetic algorithm with non gradient-based local search assisted by a meta-model, in: International Conference on Parallel Problem Solving from Nature, Springer, 2010, pp. 576–585.

[149] K. Deb, K. Sindhya, T. Okabe, Self-adaptive simulated binary crossover for real-parameter optimization, in: Proceedings of the 9th annual conference on Genetic and evolutionary computation, 2007, pp. 1187–1194.

[150] H. Li, Q. Zhang, Multiobjective optimization problems with complicated pareto sets, moea/d and nsga-ii, IEEE transactions on evolutionary computation 13 (2) (2008) 284–302.

[151] ENTSO-E, European network of transmission system operators for electricity (2013).

[152] ADMIE, Independent Power Transmission Operator (2020).
URL `http://www.admie.gr/`

[153] Y. Tian, R. Cheng, X. Zhang, Y. Jin, Platemo: A matlab platform for evolutionary multi-objective optimization [educational forum], IEEE Computational Intelligence Magazine 12 (4) (2017) 73–87.

[154] E. Zitzler, K. Deb, L. Thiele, Comparison of multiobjective evolutionary algorithms: Empirical results, Evolutionary computation 8 (2) (2000) 173–195.