



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

Διατμηματικό Πρόγραμμα Μεταπτυχιακών Σπουδών
«Παραγωγή και Διαχείριση Ενέργειας»

ΜΕΤΑΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

«Ανάπτυξη τεχνητών νευρωνικών δικτύων για πρόβλεψη της κατανάλωσης ηλεκτρικής ενέργειας»

Του Μεταπτυχιακού Φοιτητή

Μέξη Φίλιππου Δημήτριου

Επιβλέπων

Χάρης Δούκας, Αναπληρωτής Καθηγητής, Ηλεκτρολόγων
Μηχανικών & Μηχανικών Υπολογιστών

Αθήνα, Σεπτέμβριος 2019

Η παρούσα σελίδα παρέμεινε σκοπίμως λευκή.

ΜΕΤΑΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ:**«Ανάπτυξη τεχνητών νευρωνικών δικτύων για πρόβλεψη της κατανάλωσης ηλεκτρικής ενέργειας»****ΦΟΙΤΗΤΗΣ:****Μέξης Φίλιππος Δημήτριος****ΕΠΙΒΛΕΠΩΝ:****Χάρης Δούκας, Αναπληρωτής Καθηγητής,
Ηλεκτρολόγων Μηχανικών & Μηχανικών
Υπολογιστών****ΑΚΑΔΗΜΑΪΚΟ ΕΤΟΣ:****2018-2019****Σύνοψη**

Οι διακυμάνσεις και οι μεταβολές της ζήτησης ηλεκτρικής ενέργειας είναι κύριο χαρακτηριστικό των διασυνδεδεμένων συστημάτων, αστικών και μη, και επηρεάζουν σε τεράστιο βαθμό τον προγραμματισμό των εταιριών παροχής ενέργειας καθώς και την χάραξη πολιτικής τόσο σε εθνικό όσο και σε ευρωπαϊκό επίπεδο. Οι μεταβολές στην ζήτηση εξαρτώνται από φυσικούς, κοινωνικούς, οικονομικούς, τεχνικούς και άλλους παράγοντες, με τον καθένα από αυτούς να επηρεάζει το φορτίο με διαφορετικό βαθμό, βαρύτητα και περιοδικότητα. Στο πλαίσιο της παρούσας πτυχιακής εργασίας, θα γίνει μελέτη και ανάλυση με την χρήση προηγμένων υπολογιστικών εργαλείων, μη γραμμικού προγραμματισμού, με σκοπό την ανάπτυξη εργαλείου πρόβλεψης των μεταβολών της ενεργειακής ζήτησης. Το εργαλείο θα υλοποιηθεί μέσω της διαδικασίας εκμάθησης μηχανής, αξιοποιώντας ποικίλα στατιστικά δεδομένα, όπως π.χ. δημογραφικά και μετεωρολογικά στοιχεία, δείκτες ανάπτυξης ενώ παράλληλα θα μελετηθούν περαιτέρω στατιστικά στοιχεία τα οποία ενδέχεται να συμβάλλουν στην διαμόρφωση της ενεργειακής ζήτησης με όχι και τόσο εμφανή τρόπο. Το εργαλείο θα είναι σε θέση να υπολογίζει και να προβλέπει με όσο το δυνατόν καλύτερη ακρίβεια την ζήτηση ηλεκτρικής ενέργειας σε εθνικό επίπεδο με μακροπρόθεσμο ορίζοντα. Η πρότερη γνώση της μελλοντικής ζήτησης, στο βαθμό που είναι εφικτό, διευκολύνει την αποδοτικότερη διαχείριση των ενεργειακών πόρων και, κυρίως, την ανάπτυξη κατάλληλων σχεδίων δράσης με στόχο την εξοικονόμηση ενέργειας και κατ' επέκταση τη μείωση των εκπομπών CO₂ στην ατμόσφαιρα επιτυγχάνοντας έτσι τους στόχους που έχει θέσει η ΕΕ.

Περίληψη

Η παρούσα εργασία αποτελεί μια ερευνητική προσέγγιση στο πεδίο της πρόβλεψης κατανάλωσης ηλεκτρικής ενέργειας με μακροπρόθεσμο ορίζοντα με την χρήση νευρωνικών δικτύων. Η μακροπρόθεσμη πρόβλεψη έχει απασχολήσει αρκετά την επιστημονική κοινότητα για τεχνικούς, οικονομικούς, επενδυτικούς και τις τελευταίες δεκαετίες για περιβαλλοντικούς λόγους. Είναι απαραίτητο σε ένα σύγχρονο σύστημα ισχύος να έχουμε μια εκτίμηση της τάσης ανόδου και των αυξομειώσεων που πρόκειται να συμβούν στο φορτίο τα επόμενα χρόνια. Οι μέχρι πρότινος προσεγγίσεις για την πρόβλεψη του μελλοντικού φορτίου βασίζονται είτε σε στατιστικά μοντέλα, είτε σε υπολογιστικές προσεγγίσεις τεχνητής ευφυίας, είτε σε κλασσικές αναλυτικές μεθόδους. Με βάση τις μελλοντικές εκτιμήσεις ηλεκτρικού φορτίου προγραμματίζονται οι επενδυτικές στρατηγικές για την ανάπτυξη υποδομών, έτσι ώστε το δίκτυο μεταφοράς και διανομής να παραμένει επαρκές και αποδοτικό σε βάθος χρόνου, αλλά και να κατασκευάζονται έγκαιρα καινούριες μονάδες παραγωγής ισχύος. Η πρότερη γνώση της ζήτησης παίζει, προφανώς, σημαντικό ρόλο και στην διείσδυση των Ανανεώσιμων Πηγών Ενέργειας στο δίκτυο, καθώς και τον σωστό προγραμματισμό της προσφοράς ισχύος των συμβατικών μονάδων ισχύος έτσι ώστε να εξασφαλίζεται η ευστάθεια και η επάρκεια του δικτύου. Τέλος, οι στρατηγικές μείωσης των εκπομπών CO₂ που βρίσκονται σε εξέλιξη στην επικράτεια αλλά και σε Ευρωπαϊκό και παγκόσμιο επίπεδο, απαιτούν την γνώση του μελλοντικού φορτίου, έτσι ώστε να μπορούν να πραγματοποιήσουν μελλοντικά σχέδια, όχι μόνο σε ποιοτικό αλλά και σε ποσοτικό επίπεδο.

Στα πλαίσια της παρούσας εργασίας πραγματοποιείται πρόβλεψη μέσω προηγμένων προγραμματιστικών εργαλείων και συγκεκριμένα με την χρήση της μεθόδου των τεχνητών νευρωνικών δικτύων. Τα τεχνητά νευρωνικά δίκτυα είναι μια ευρέως διαδεδομένη μέθοδος εκμάθησης μηχανής, κατά την οποία ένα σύστημα που προσομοιάζει τον ανθρώπινο εγκέφαλο, δομημένο με νευρώνες είναι σε θέση να εκπαιδευτεί με γνωστά δεδομένα του προβλήματος και να παράγει τις επιθυμητές εξόδους.

Εφαρμογές νευρωνικών δικτύων έχουν οι περισσότερες σύγχρονες γλώσσες προγραμματισμού, στην συγκεκριμένη έρευνα χρησιμοποιήθηκε το λογισμικό της Matlab. Η Matlab θεωρείται γλώσσα προγραμματισμού με πληθώρα έτοιμων πακέτων εντολών για εφαρμογές τεχνητών νευρωνικών δικτύων. Χρησιμοποιείται ευρύτατα στον επιστημονικό κόσμο καθώς έχει βασιστεί και εξελιχθεί με προχωρημένο μαθηματικό υπόβαθρο.

Για την πραγματοποίηση της έρευνας ήταν απαραίτητο να βρεθούν ιστορικά δεδομένα κατανάλωσης ενέργειας για την προς μελέτη περιοχή. Έτσι, η μελέτη προσαρμόστηκε στα διαθέσιμα καταγεγραμμένα δεδομένα κατανάλωσης ηλεκτρικής ενέργειας, τα οποία διατίθενται από διεθνείς οργανισμούς, ελεύθερα στο διαδίκτυο. Η Ελληνική Στατιστική Αρχή παρέχει δεδομένα σε επίπεδο πόλης-νομού, τα οποία, όμως, δεν είχαν μεγάλο χρονικό ορίζοντα ούτε ικανοποιητική συχνότητα για την ανάπτυξη των νευρωνικών δικτύων και την πραγματοποίηση της πρόβλεψης. Έτσι, η μελέτη περατώθηκε με δεδομένα σε εθνικό επίπεδο.

Επιπλέον, για τις ανάγκες της έρευνας αναζητήθηκαν κλιματολογικά στοιχεία, κυρίως καταγραφές θερμοκρασιών, σε μηνιαία βάση για όλη την επικράτεια. Τα δεδομένα αυτά βρέθηκαν από την Ελληνική Μετεωρολογική Υπηρεσία και χρειάστηκαν ιδιαίτερη επεξεργασία για να μετατραπούν σε χρονοσειρές κατάλληλες για χρήση σε εφαρμογές νευρωνικών δικτύων. Επιπλέον, χρησιμοποιήθηκαν οικονομικά και δημογραφικά δεδομένα, όπως το ΑΕΠ της χώρας και ο πληθυσμός έτσι ώστε να δοκιμαστούν συνδυαστικά και να ερευνηθεί το κατά πόσο συμβάλλουν σε μία μοντελοποίηση νευρωνικών δικτύων.

Αναλόγως το είδος της εφαρμογής και των δεδομένων του προβλήματος η Matlab παρέχει διαφορετικά είδη νευρωνικών δικτύων, τα οποία δύναται να παραμετροποιηθούν περαιτέρω. Για την συγκεκριμένη εφαρμογή χρησιμοποιήθηκαν αυτοπαλινδρομικά νευρωνικά δίκτυα πρόβλεψης με ή χωρίς εξωτερική βοηθητική χρονοσειρά. Έγινε συνδυασμός των διαθέσιμων δεδομένων και πραγματοποιήθηκαν δοκιμές με διάφορες παραμέτρους. Για αυτοματοποίηση της διαδικασίας και την γενίκευση του προβλήματος γράφτηκε κώδικας ο οποίος φορτώνει τα δεδομένα, δημιουργεί το δίκτυο, αρχικοποιεί τις παραμέτρους, εκπαιδεύει το δίκτυο, εκτυπώνει τα αποτελέσματα της εκπαίδευσης σε διαγράμματα και πραγματοποιεί την πρόβλεψη εξάγοντας το αντίστοιχο γράφημα.

Η ποιότητα της εκπαίδευσης αξιολογείται από διάφορες παραμέτρους όπως το μέσο τετραγωνικό σφάλμα, ο συντελεστής προσδιορισμού R^2 των εξαγόμενων δεδομένων σε σχέση με τα πραγματικά καθώς και τα διαγράμματα αυτοσυσχετίσης και κατανομής σφαλμάτων. Με βάση τα αποτελέσματα που προκύπτουν από κάθε εκπαίδευση κρίνεται αν το δίκτυο είναι κατάλληλο να πραγματοποιήσει πρόβλεψη ή όχι. Κάθε φορά που το δίκτυο εκπαιδεύεται μπορεί να δώσει διαφορετικά αποτελέσματα, καθώς ο αλγόριθμος εκπαίδευσης ακολουθεί μια επαναληπτική διαδικασία η οποία σε κάθε δοκιμή συγκλίνει σε διαφορετικά αποτελέσματα. Η εκπαίδευση ακολουθεί τον αλγόριθμο της αναστροφής μετάδοσης λάθους, όπου στο τέλος κάθε κύκλου εκπαίδευσης αξιολογείται το μέσο τετραγωνικό σφάλμα και αναπροσαρμόζονται τα συναπτικά βάρη.

Εάν η εκπαίδευση κριθεί ικανοποιητική τότε το δίκτυο προχωρά στην πρόβλεψη, όπου δέχεται αρχικές τιμές της χρονοσειράς και στη συνέχεια παράγει μόνο του τα επόμενα βήματα. Η πρόβλεψη κρίνεται και σχολιάζεται με βάση την τάση, τις διακυμάνσεις και τις ακραίες τιμές της προβλέψεων. Η παραγόμενη χρονοσειρά θα πρέπει να ακολουθεί τις βασικές αρχές της καμπύλης φορτίου, να παρουσιάζει αυξητική τάση και να μην εμφανίζει πολύ μεγάλες ακραίες τιμές, έτσι ώστε η πρόβλεψη να θεωρηθεί έγκυρη.

Ο χρονικός ορίζοντας της πρόβλεψης, στα πλαίσια της έρευνας, δεν είναι αυστηρά ορισμένος. Η πρόβλεψη είναι μακροπρόθεσμη, αλλά με βάση τα αποτελέσματα των εξόδων των τεχνητών νευρωνικών δικτύων αξιολογείται το εύρος εμπιστοσύνης της πρόβλεψης. Στην βιβλιογραφία, οι προβλέψεις που μπορεί να εξάγει ένα καλά εκπαιδευμένο δίκτυο μπορούν να φτάσουν και τον αριθμό των ήδη γνωστών τιμών. Στην πράξη αυτό εξαρτάται από την φύση των δεδομένων της χρονοσειράς, την απόδοση κατά την εκπαίδευση και την σωστή παραμετροποίηση του δικτύου. Σε κάποιες περιπτώσεις, λόγω των διαφορών τύπων χρονοσειρών που χρησιμοποιήθηκαν, ήταν εφικτό να δοκιμαστεί το δίκτυο και να προβλέψει τιμές για τις οποίες υπάρχουν πραγματικά δεδομένα. Στην περίπτωση αυτή, το δίκτυο πρόβλεψε με μεγάλη ακρίβεια τα επόμενα βήματα της χρονοσειράς, δηλαδή την κατανάλωση ηλεκτρικής ενέργειας.

Ολοκληρώνοντας, έγινε συνολική αξιολόγηση των αποτελεσμάτων και των παραγόντων που επηρεάζουν την εκπαίδευση των τεχνητών νευρωνικών δικτύων και την πρόβλεψη που είναι σε θέση να παράγουν.

Λέξεις κλειδιά:

Πρόβλεψη ηλεκτρικής ενέργειας, καμπύλη φορτίου, τεχνητά νευρωνικά δίκτυα, εκμάθηση μηχανής, ανάλυση χρονοσειρών

POST-GRADUATE THESIS: «Development of Artificial Neural Networks for predicting electricity consumption. »

STUDENT: Filippos Dimitrios Mexis

SUPERVISOR: Haris Doukas, Associate Professor, Division of Industrial Electric Devices and Decision Systems, School of Electrical and Computer Engineering

ACADEMIC YEAR: 2018-2019

Abstract

Forecasting electricity consumption plays a vital role in keeping the electric power system up and running. Power plants need to have short term forecasting of the upcoming load in order to manage efficiently the generators to provide the power needed instantly, as the electricity cannot be saved in the electrical grid. Moreover, long-term estimation of the load can help European Union member states and European Commission to set energy efficiency goals and develop clean energy policies, by deploying renewable sources, finance energy efficiency investments and control carbon dioxide emissions. Electricity load fluctuations depend on several environmental, social, financial and technical factors in a nonlinear relation that make the load forecasting multivariate and complex. This study aims to predict future energy load for Greece, using non-linear state of the art computational tools, such as artificial neural networks. The latter have been designed in MATLAB® environment, where Neural Autoregressive Time Series Prediction Networks have been deployed. In order to complete the neural network training process, Greece's historic electricity consumption time series have been used. Similarly, in order to examine the correlation and impact of environmental, economic and demographic parameters to the electricity load, Neural Autoregressive Networks with exogenous input have been exploited. The calculated predictions by the trained neural networks have been examined on their accuracy, reliability and their forecast timeframe. In addition, an analytic comparison has been realised between the forecasts of this research and the prediction of the Greek Independent Power Transmission Operation.

Summary

The following study aims to predict energy consumption for Greece exploiting machine-learning algorithms. The study has been concluded by the supervision of the Decision Support Systems Laboratory School of Electrical and Computer Engineering of the National Technical University of Athens. The importance of the electricity consumption forecasts relies on the fact that the latter is crucial for the electric grid, for the power plants efficiency, and for environmental, decision making and policy deployment reasons, depending on the timeframe of the prediction. Insight of the short-term future load permits power plants to equilibrate the electricity production with demand at any time. This ensures the safety of the grid and of the consumers and the efficient managing of the plants schedule. Additionally, long-term forecasts play vital role in planning grid and infrastructure development and energy efficiency projects. European Commission and the United Nations has set long-term goals and targets that have to be accomplished by the European Union member states. Among others, countries have to deliver energy efficiency projects and develop clean energy policies, by deploying renewable sources, finance energy efficiency investments and control carbon dioxide emissions.

Variations is a common characteristic of the long-term electricity load time series. Several environmental, social, financial and technical factors affect the electricity demand in a nonlinear way that makes the load forecasting multivariate and complex. The scope of this research is the long-term prediction of the energy load of Greece by implementing machine-learning tools that exploit country's historical electricity consumption, as well as the aforementioned variables that alter the electricity load. The historical consumption was provided global energy agencies and financial institutions and the auxiliary variables time series were provided from Greek authorities. The data were pre-processed in various software such as Microsoft Excel and finally inserted in MATLAB® environment, where Artificial Neural Networks have been deployed and trained to produce the desideratum forecast.

MATLAB® is a software environment for engineers and scientists that is suitable for iterative analysis and design processes with a programming language that expresses matrix and array mathematics directly. Among others, in

MATLAB® can be found parameterized neural networks that are suitable for time series generalization and prediction. A neural network is a series of algorithms that endeavours to recognize underlying relationships in a set of data through a process that mimics the way the human brain operates. Neural networks can adapt to changing input; so the network generates the best possible result without needing to redesign the output criteria. The concept of neural networks, which has its roots in artificial intelligence, is swiftly gaining popularity in the development of trading systems. In this research, two kinds of neural networks have been deployed. These are the Neural Autoregressive Networks (NAR) and the Neural Autoregressive Networks with exogenous input (NARX). The NAR networks are fitted to be trained only on the previous steps of the time series that is predicted, without any other auxiliary data. On the contrary, the NARX networks are applied to predict a time series, using the original time series as well as an auxiliary one of an independent variable that affects the prediction. The NARX neural networks avail the analysis of the examination of the correlation and impact of environmental, economic and demographic parameters to the electricity load.

In respect of the data collection, various time series had to be acquired in order to develop the several scenarios of the analysis. The essential data were the historic electricity consumption, which has been provided from the International Energy Agency and the World Bank. The first is a monthly time series of electricity consumption ranged from January 2000 to January 2019 resulting to 229 - time steps. The World Bank time series is an annually energy consumption record since 1960 expanding up to 2015, providing only 44 - time steps. The World Bank's electricity consumption data are expressed in kWh per capita, combining the population information into the primary time series. Both amount of data is considered insufficient to properly train a neural network, although there are no available open data of a wider range nor of a denser frequency. Additionally, complementary time series have been included to form the training scenarios, in order to enhance neural networks performance. These are the mean temperature and GDP, which have been acquired from Greece's Statistical Authority and the Hellenic National Meteorological Service, respectively. Concerning the scenarios that the auxiliary time series have been implemented, the electricity load data had to be trimmed and reprocessed as a result of the different time step and range of the aforementioned time series. The latter has caused reduction of the information inserted into the neural network system and cut down its performance. Additionally, the temperature data were not provided in editable form, so they had to be converted manually and subsequently calculate the country's mean temperature and the corresponding time series.

The study includes five forecast scenarios of trained neural networks, in which one different data combination have been deployed. The first two scenarios are simple NAR neural networks, using only the two distinct electricity consumption time series, and the following three combine electricity consumption with temperature values, month index values and country's GDP in NARX neural networks. After the training process, to produce forecast in an NARX neural network, future values of the auxiliary data is needed. That dictates the necessity to produce temperature and GDP forecast using a NAR neural network, as an input to the NARX scenario. The fact that the electricity forecast of the complex NARX scenarios is based on precedent forecasts of the supplementary data dwindles the forecasting accuracy and increases thereof uncertainty.

The MATLAB® training process of the neural network chosen for this study is based on the Levenberg-Marquardt backpropagation algorithm. The latter updates the network weight and biases values similar to the quasi-Newton methods, and it was designed to approach second-order training speed without having to compute the Hessian matrix. This algorithm appears to be the fastest method for training moderate-sized feedforward neural networks and it also has an efficient implementation in MATLAB® software, because the solution of the matrix equation is a built-in function, so its attributes become even more pronounced in a MATLAB® environment.

The generated predictions and the trained neural networks have been evaluated on their accuracy, reliability and their forecast timeframe. After the training process, MATLAB® provides graphs and metrics that facilitate the assessment of the produced neural network and forecast. The particular indicators are the coefficient of determination (R^2) between the real data and the estimations of the neural network in the training process, the error's autocorrelation and frequency graphs as well as the mean squared error graph during the training epochs. Furthermore, the forecasts time series have been evaluated over their similarity to the real characteristics of an electricity consumption data set. Additionally, in some scenarios the predicted values have been compared to the real electricity consumption, as the ultimate data recorded in the time series were already a few years old.

To generalize the forecast accuracy assessment, an analytic comparison have been concluded between the forecasts of this research and the prediction of the Greek Independent Power Transmission Operation. Comparing the different forecast trends and characteristics, conclusions have been carried out about the correlation of the original time series time-step, range and the effect of complementary data to the predictions.

Key words:

Electricity consumption forecast, load profile, artificial neural networks, machine learning, time series analysis

Πρόλογος

Η παρούσα μεταπτυχιακή εργασία αποτελεί ερευνητική προσέγγιση στο επιστημονικό πεδίο των προβλέψεων με την χρήση εργαλείων τεχνητής νοημοσύνης. Η πρόβλεψη αφορά την κατανάλωση ηλεκτρικής ενέργειας σε εθνικό επίπεδο, για την Ελλάδα. Η πραγματοποίηση προβλέψεων αποτελεί βασικό εργαλείο για την χάραξη πολιτικών εξοικονόμησης ενέργειας και την εκτίμηση των μελλοντικών προβλημάτων για την επιτυχή εφαρμογή των πολιτικών αυτών. Η Ευρωπαϊκή Ένωση έχει ορίσει αυστηρούς στόχους για το 2030 και χρηματοδοτεί δράσεις έτσι ώστε να αυξηθεί η ενεργειακή αυτονομία και να μειωθεί η εξάρτηση από τα ορυκτά καύσιμα μέχρι το έτος 2050. Στην προσπάθεια αυτή, η συμβολή των τεχνολογιών της σύγχρονης εποχής, όπως τα νευρωνικά δίκτυα και η εκμάθηση μηχανής, είναι πλέον σε θέση να συμβάλλουν δραστικά και να επιταχύνουν την πρόοδο και την αποτελεσματικότητα των δράσεων.

Την επίβλεψη της εργασίας ανέλαβε ο Αναπληρωτής Καθηγητής του Ε.Μ.Π. κ. Χάρης Δούκας, τον οποίο ευχαριστώ θερμά, ενώ συνέβαλαν καταλυτικά οι συνεργάτες του Εργαστηρίου Συστημάτων Αποφάσεων και Διοίκησης του Τομέα Ηλεκτρικών Βιομηχανικών Διατάξεων και Συστημάτων Αποφάσεων της σχολής Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών με ιδιαίτερη μνεία στην κ. Ήρα Νεοφύτου. Τέλος, θα ήθελα να ευχαριστήσω την διοίκηση του Διατμηματικού Προγράμματος Μεταπτυχιακών Σπουδών «Παραγωγή & Διαχείριση Ενέργειας», διότι το πρόγραμμα αυτό αποτέλεσε ορόσημο στην ακαδημαϊκή μου πορεία και την εκπόνηση της μεταπτυχιακής εργασίας.

Περιεχόμενα

	Εισαγωγή	13
	Σκοπός και μεθοδολογία	13
	Σκοπός.....	13
1.	Γενική μεθοδολογική προσέγγιση και Φάσεις υλοποίησης εργασίας.....	14
1.1	Στατιστικές Μέθοδοι Πρόβλεψης Ηλεκτρικού Φορτίου	17
	Εισαγωγή.....	17
2.	Παράγοντες που επηρεάζουν την ζήτηση ηλεκτρικής ενέργειας.....	17
2.1	Μέθοδοι πρόβλεψης.....	18
2.2	Μέθοδοι πρόβλεψης.....	18
2.3	Τεχνητά Νευρωνικά Δίκτυα.....	21
3.	Γενικά.....	21
3.1	Χρήση Νευρωνικών Δικτύων σε προβλέψεις	24
3.2	Χρήση Νευρωνικών Δικτύων σε προβλέψεις	24
3.3	Τεχνητά Νευρωνικά δίκτυα στην Matlab	24
	Non Linear Autoregressive Neural Network (NAR).....	25
4.	Non Linear Autoregressive with External (Exogenous) Input (NARX)	25
4.1	Μεθοδολογία.....	27
	Παράμετροι δικτύου	27
	Διαστασιολόγηση Δικτύου.....	27
	Υστερήσεις εισόδου (Input Delays)	27
4.2	Λειτουργία ανοιχτού βρόχου (Open Loop).....	27
	Εκπαίδευση	28
	Λειτουργία κλειστού βρόχου (Closed Loop)	28
	Σχεδιασμός Σεναρίων	29

	Σενάριο 1.....	29
	Σενάριο 2.....	30
	Σενάριο 3.....	30
	Σενάριο 4.....	31
	Σενάριο 5.....	32
	Λήψη & Επεξεργασία Δεδομένων.....	32
	Δεδομένα Κατανάλωσης Ενέργειας	32
4.3	Κλιματικά Δεδομένα.....	34
	Δημογραφικά Δεδομένα	35
	Οικονομικός Δείκτης ΑΕΠ	35
	Προετοιμασία δεδομένων	36
	Κανονικοποίηση δεδομένων	37
5.	Αποτελέσματα.....	39
5.1	Σενάριο 1	40
5.2	Σενάριο 2.....	47
5.3	Σενάριο 3.....	52
5.4	Πρόβλεψη Θερμοκρασίας.....	52
5.5	Πρόβλεψη Κατανάλωσης Ενέργειας.....	57
	Σενάριο 4.....	61
5.6	Σενάριο 5.....	65
6.	Πρόβλεψη ΑΕΠ.....	65
	Πρόβλεψη Κατανάλωσης Ενέργειας.....	68
	Γενικά συμπεράσματα.....	73
	Προτάσεις για μελλοντική έρευνα	76

Παράρτημα - Κώδικες Matlab Live Script.....	79
Σενάριο 1.....	79
Σενάριο 2.....	82
Σενάριο 3.....	84
Πρόβλεψη NAR θερμοκρασίας.....	84
Πρόβλεψη NARX Κατανάλωσης Ηλεκτρικής Ενέργειας.....	86
Σενάριο 4.....	87
Σενάριο 5.....	89
Πρόβλεψη NAR ΑΕΠ.....	89
Πρόβλεψη NARX Κατανάλωσης Ηλεκτρικής Ενέργειας.....	90

Ευρετήριο Εικόνων

Εικόνα 3-1: Μοντέλο τεχνητού νευρώνα (Πηγή: I. Βλαχαβάς [12]).....	21
Εικόνα 3-2 Νευρωνικό Δικτύου Πρόσθιας Τροφοδότησης (Πηγή: H. Hamedmoghadam, [15]).....	22
Εικόνα 3-3: Σχηματική απεικόνιση αλγορίθμου αναστροφής λάθους (Πηγή: I. Βλαχάβας et. Al [12])	23
Εικόνα 4-1: Δομή νευρωνικού δικτύου τύπου NAR ανοιχτού βρόχου	28
Εικόνα 4-2: Εργαλείο εκπαίδευσης νευρωνικού δικτύου σε περιβάλλον MATLAB	28
Εικόνα 4-3: Δομή νευρωνικού δικτύου τύπου NAR κλειστού βρόχου	29
Εικόνα 4-4: Δομή NAR νευρωνικού δικτύου σεναρίου 1	29
Εικόνα 4-5: Δομή NAR νευρωνικού δικτύου σεναρίου 2	30
Εικόνα 4-6: Δομή NAR νευρωνικού δικτύου σεναρίου 3	31
Εικόνα 4-7: Δομή NARX νευρωνικού δικτύου σεναρίου 3.....	31
Εικόνα 4-8 Δομή NARX νευρωνικού δικτύου σεναρίου 4.....	31
Εικόνα 4-9: Δομή NARX νευρωνικού δικτύου σεναρίου 5.....	32

Ευρετήριο Γραφημάτων

Γράφημα 2-1: Παγκόσμια αύξηση μέσης θερμοκρασίας την περίοδο 1880-2020. Πηγή: NASA's Goddard Institute for Space Studies (GISS).....	18
Γράφημα 4-1: Μηνιαία κατανάλωση ηλεκτρικής ενέργειας (Πηγή: International Energy Agency)	33
Γράφημα 4-2: Ετήσια κατά κεφαλήν κατανάλωση ηλεκτρικής ενέργειας (Πηγή: World Bank)	34
Γράφημα 4-3: Μέση μηνιαία θερμοκρασία Ελλάδος (Πηγή: EMY).....	35
Γράφημα 4-4: Ακαθάριστο Εθνικό Προϊόν (Πηγή: ΕΛΣΤΑΤ	36
Γράφημα 5-1: Σενάριο 1, Μέσο τετραγωνικό σφάλμα κατά την εκπαίδευση	41
Γράφημα 5-2: Σενάριο 1,Παλινδρόμηση πραγματικών – εκτιμώμενων τιμών.....	43
Γράφημα 5-3: Σενάριο 1, Διασπορά πραγματικών- εκτιμώμενων τιμών	44
Γράφημα 5-4: Σενάριο 1, Αυτοσυσχέτιση σφαλμάτων	45
Γράφημα 5-5: Σενάριο 1, Κατανομή σφαλμάτων	45
Γράφημα 5-6: Σενάριο 1: Πρόβλεψη κατανάλωσης ηλεκτρικής ενέργειας.....	46
Γράφημα 5-7: Σενάριο 2, Μέσο τετραγωνικό σφάλμα κατά την εκπαίδευση	47
Γράφημα 5-8: Σενάριο 2, Παλινδρόμηση πραγματικών- εκτιμώμενων τιμών	48
Γράφημα 5-9: Σενάριο 2, Διασπορά πραγματικών- εκτιμώμενων τιμών	49
Γράφημα 5-10: Σενάριο 2, Αυτοσυσχέτιση σφαλμάτων	49
Γράφημα 5-11: Σενάριο 2, Κατανομή σφαλμάτων	50
Γράφημα 5-12 : Σενάριο 2, Πρόβλεψη κατανάλωσης ηλεκτρικής ενέργειας.....	51
Γράφημα 5-13: Σενάριο 2, Σύγκριση πραγματικών – προβλεπόμενων τιμών 2015-2018	52
Γράφημα 5-14: Σενάριο 3a,Μέσο τετραγωνικό σφάλμα κατά την εκπαίδευση	53
Γράφημα 5-15: Σενάριο 3a, Παλινδρόμηση πραγματικών – εκτιμώμενων τιμών	53
Γράφημα 5-16: Σενάριο 3a, Διασπορά πραγματικών- εκτιμώμενων τιμών.....	54
Γράφημα 5-17: Σενάριο 3a, Αυτοσυσχέτιση σφαλμάτων.....	55
Γράφημα 5-18: Σενάριο 3a, Κατανομή σφαλμάτων.....	55
Γράφημα 5-19: Σενάριο 3a, Πρόβλεψη θερμοκρασίας.....	56
Γράφημα 5-20: Σενάριο 3b ,Μέσο τετραγωνικό σφάλμα κατά την εκπαίδευση	57

Γράφημα 5-21: Σενάριο 3b, Παλινδρόμηση πραγματικών- εκτιμώμενων τιμών	58
Γράφημα 5-22: Σενάριο 3b, Διασπορά πραγματικών- εκτιμώμενων τιμών	59
Γράφημα 5-23: Σενάριο 3b, Αυτοσυσχέτιση σφαλμάτων	59
Γράφημα 5-24: Σενάριο 3b, Κατανομή σφαλμάτων.....	60
Γράφημα 5-25: Σενάριο 3b, Πρόβλεψη κατανάλωσης ηλεκτρικής ενέργειας.....	61
Γράφημα 5-26: Σενάριο 4, Μέσο τετραγωνικό σφάλμα κατά την εκπαίδευση	61
Γράφημα 5-27: Σενάριο 4, Παλινδρόμηση πραγματικών – εκτιμώμενων τιμών.....	62
Γράφημα 5-28: Σενάριο 4, Διασπορά πραγματικών- εκτιμώμενων τιμών	63
Γράφημα 5-29: Σενάριο 4, Αυτοσυσχέτιση σφαλμάτων	63
Γράφημα 5-30: Σενάριο 4, Κατανομή σφαλμάτων	64
Γράφημα 5-31: Σενάριο 4, Πρόβλεψη κατανάλωσης ηλεκτρικής ενέργειας.....	64
Γράφημα 5-32: Σενάριο 5a, Μέσο τετραγωνικό σφάλμα κατά την εκπαίδευση	65
Γράφημα 5-33: Σενάριο 5a, Παλινδρόμηση πραγματικών – εκτιμώμενων τιμών	66
Γράφημα 5-34: : Σενάριο 5a, Διασπορά πραγματικών- εκτιμώμενων τιμών	66
Γράφημα 5-35: Σενάριο 5a, Αυτοσυσχέτιση σφαλμάτων.....	67
Γράφημα 5-36: Σενάριο 5a, Κατανομή σφαλμάτων.....	67
Γράφημα 5-37: Σενάριο 5a, Πρόβλεψη Ακαθάριστου Εγχώριου Προϊόντος	68
Γράφημα 5-38: Σενάριο 5b, Μέσο τετραγωνικό σφάλμα κατά την εκπαίδευση	69
Γράφημα 5-39: Σενάριο 5b, Παλινδρόμηση πραγματικών – εκτιμώμενων τιμών.....	69
Γράφημα 5-40: Σενάριο 5b, Διασπορά πραγματικών- εκτιμώμενων τιμών	70
Γράφημα 5-41: Σενάριο 5b, Αυτοσυσχέτιση σφαλμάτων	70
Γράφημα 5-42: Σενάριο 5b, Κατανομή σφαλμάτων.....	71
Γράφημα 5-43: Σενάριο 5b, Πρόβλεψη κατανάλωσης ηλεκτρικής ενέργειας.....	71
Γράφημα 5-44 Σύγκριση προβλέψεων διαφορετικών σεναρίων.....	75

Εισαγωγή

Σκοπός και μεθοδολογία

1. Σκοπός

- 1.1 Σκοπός της εργασίας είναι η πρόβλεψη της κατανάλωσης ηλεκτρικής ενέργειας με την χρήση τεχνητών νευρωνικών δικτύων. Ένα νευρωνικό δίκτυο αποτελείται από ένα σύστημα τεχνητών νευρώνων, αντιγράφοντας την δομή των πραγματικών νευρών ενός βιολογικού εγκεφάλου. Οι νευρώνες επεξεργάζονται τα δεδομένα των πληροφοριών ενώ ταυτοχρόνως επικοινωνούν μεταξύ τους. Η παράλληλη επικοινωνία διέπεται από λογικές συνθήκες και κανόνες οι οποίοι στο σύνολό τους συνθέτουν μια τεχνητή νοημοσύνη που είναι ικανή να εκπαιδευτεί και να παράγει νέα πληροφορία με βάση ένα προϋπάρχον σύνολο δεδομένων. Έτσι, η αξιοποίηση ενός νευρωνικού δικτύου για την πραγματοποίηση πρόβλεψης φορτίου ηλεκτρικής ενέργειας βασίζεται στην εκπαίδευση επάνω σε παλαιότερες καταγραφές κατανάλωσης, έτσι ώστε να παράγει τις μελλοντικές τιμές αυτής. Η παρούσα μελέτη έγινε σε εθνικό επίπεδο (κατανάλωση ηλεκτρικής ενέργειας για όλη την χώρα) με την εφαρμογή αυτοπαλινδρομικών νευρωνικών δικτύων, ενώ παράλληλα έγιναν δοκιμές με συνδυασμούς δεδομένων όπως η θερμοκρασία περιβάλλοντος, δημογραφικοί και κοινωνικοί δείκτες, οι οποίοι είναι γνωστό ότι επηρεάζουν το φορτίο. Με αυτό τον τρόπο εξετάστηκε το κατά πόσο συμβάλλουν ή επηρεάζουν την εκπαίδευση και την παραγόμενη πρόβλεψη των τεχνητών νευρωνικών δικτύων.

Η πρόβλεψη του μελλοντικού φορτίου αποτελεί σημαντικό εργαλείο στα χέρια των παραγωγών ηλεκτρικής ενέργειας, του διαχειριστή του συστήματος μεταφοράς και διανομής αυτής, του κράτους, των φορέων που έχουν ως ρόλο την χάραξη πολιτικών εξοικονόμησης ενέργειας καθώς και ερευνητών που ασχολούνται με το φαινόμενο της ενεργειακής φτώχειας, των περιβαλλοντικών προβλημάτων που συνδέονται με την χρήση ορυκτών καυσίμων και την ανάπτυξη των ανανεώσιμων πηγών ενέργειας. Τα τελευταία χρόνια έχει γίνει παγκοσμίως αισθητό το γεγονός πως οι ενεργειακοί πόροι βρίσκονται σε έλλειψη, ενώ παράλληλα οι ενεργειακές απαιτήσεις καλπάζουν. Παρακολουθώντας τις συνεχώς αυξανόμενες ενεργειακές απαιτήσεις που επιτάσσει η αστικοποίηση και ο σύγχρονος τρόπος ζωής, τα περιβαλλοντικά προβλήματα διογκώνονται, με τρανό παράδειγμα εκείνο της κλιματικής αλλαγής. Οι μεταβολές στο κλίμα, έχουν γίνει ιδιαίτερα αισθητές, με την εμφάνιση ακραίων κλιματικών φαινομένων, την συνεχόμενη αύξηση περιόδων παρατεταμένης ξηρασίας, ασταμάτητων πυρκαγιών και πλημμυρικών φαινομένων.

Στο πλαίσιο της ανεξάρτησης από τα ορυκτά καύσιμα και της ενεργειακής αυτονομίας του ευρωπαϊκού χώρου η Ευρωπαϊκή Ένωση βρίσκεται συνεχώς σε επιφυλακή για την προώθηση υφιστάμενων μέτρων και την χάραξη νέων πολιτικών για την εξοικονόμηση ενέργειας, την ανεξάρτηση από τα συμβατικά ρυπογόνα καύσιμα, την ενεργειακή αυτάρκεια και την αντιμετώπιση της κλιματικής αλλαγής. Για την εκτίμηση του όγκου, της έντασης και την παραμετροποίηση των δράσεων στα νέα δεδομένα, είναι απαραίτητη η γνώση του μελλοντικού φορτίου. Οι απαιτήσεις σε ηλεκτρική ενέργεια επηρεάζονται από πληθώρα φυσικών, περιβαλλοντικών, οικονομικών, τεχνικών και κοινωνικών παραγόντων, οι οποίοι δρουν συνεργιστικά και μη γραμμικά, καθιστώντας το πρόβλημα υπολογισμού του μελλοντικού φορτίου σύνθετο και πολυπαραμετρικό. Η ανάπτυξη της τεχνητής ευφυίας και η ραγδαία αύξηση της υπολογιστικής ισχύος έχουν δημιουργήσει υπολογιστικά εργαλεία τα διαθέτουν την ικανότητα εκμάθησης και ανάλυσης μεγάλου όγκου δεδομένων. Έτσι, είναι πλέον εφικτό να είμαστε σε θέση να παράγουμε ακριβείς προβλέψεις, αξιοποιώντας την υπάρχουσα καταγεγραμμένη πληροφορία, η οποία θα ήταν αδύνατο να επεξεργαστεί με συμβατικά μέσα.

Γενική μεθοδολογική προσέγγιση και Φάσεις υλοποίησης εργασίας

Η εργασία έχει πειραματικό χαρακτήρα, στα πλαίσια του οποίου έγινε αξιολόγηση της απόδοσης των τεχνητών νευρωνικών δικτύων σε διαφορετικής προέλευσης δεδομένα κατανάλωσης ηλεκτρικής ενέργειας. Το πείραμα προσεγγίστηκε προγραμματιστικά σε περιβάλλον Matlab, στο οποίο δοκιμάστηκαν διαφορετικού τύπου δεδομένα, χωρισμένα σε σενάρια με την εφαρμογή ποικίλων παραμέτρων. Μετά την εκπαίδευση των δικτύων πραγματοποιήθηκαν προβλέψεις τα αποτελέσματα των οποίων σχολιάστηκαν, ως προς την εγκυρότητα τους ενώ παράλληλα έγινε ερμηνεία των προβλέψεων και της συσχέτισής τους με τις περιβαλλοντικές, κοινωνικές και οικονομικές παραμέτρους.

Έπειτα, εξάγεται η πρόβλεψη με ορίζοντα 2n επί των προηγούμενων καταγεγραμμένων βημάτων των χρονοσειρών που χρησιμοποιήθηκαν για την εκπαίδευση του νευρωνικού δικτύου του κάθε σεναρίου. Οι προβλέψεις αξιολογούνται ως προς την εγκυρότητά τους, συγκρίνονται μεταξύ τους και με την πρόβλεψη της μελέτης επάρκειας ισχύος του ΑΔΜΗΕ και εξάγονται συμπεράσματα για την ποιότητα τους και την ικανότητα των νευρωνικών δικτύων του κάθε σεναρίου στην εκμάθησή των χαρακτηριστικών των χρονοσειρών κατανάλωσης ηλεκτρικής ενέργειας.

Παρακάτω ακολουθούν οι φάσεις υλοποίησης της εργασίας:

1^η Φάση: Ανάθεση Διπλωματικής Εργασίας

Συζητήθηκε η δυνατότητα και ο τρόπος πραγματοποίησης της μεταπτυχιακής εργασίας, με γνώμονα τα διαθέσιμα ιστορικά δεδομένα, το πλαίσιο αναζήτησης περαιτέρω πληροφοριών, την σκοπιμότητα του αντικειμένου της εργασίας, την προγραμματιστική υλοποίηση και τον χρονικό ορίζοντα εκπόνησής της.

2^η Φάση: Αναζήτηση και συγκέντρωση απαιτούμενων δεδομένων

Έγινε αναζήτηση διαδικτυακά, ενώ παράλληλα στάλθηκαν επιστολές στους αρμόδιους φορείς για την χορήγηση των απαιτούμενων δεδομένων για την υλοποίηση των τεχνητών νευρωνικών δικτύων. Τα δεδομένα περιλαμβάνουν ιστορικές καταγραφές κατανάλωσης ηλεκτρικής ενέργειας, κλιματικά, δημογραφικά καθώς και οικονομικά δεδομένα της χώρας.

3^η Φάση: Προετοιμασία των δεδομένων για την είσοδο τους στο προγραμματιστικό περιβάλλον.

Καθώς τα δεδομένα προέρχονταν από διαφορετικούς φορείς, παρουσίαζαν ποικιλομορφία στην συχνότητα, στον τύπο του αρχείου δεδομένων, στη μορφοποίηση κ.α. Έτσι, ήταν αναγκαίο να γίνει περαιτέρω επεξεργασία έτσι ώστε να είναι συμβατά με το προγραμματιστικό περιβάλλον της Matlab.

4^η Φάση: Επιλογή κατάλληλων νευρωνικών δικτύων από το πακέτο της Matlab για την πραγματοποίηση της πρόβλεψης.

Η Matlab παρέχει μια πληθώρα παραμετροποιημένων νευρωνικών δικτύων για διάφορες εφαρμογές, Έτσι, ήταν αναγκαίο να μελετηθούν οι τύποι νευρωνικών δικτύων οι οποίοι ήταν καταλληλότεροι για την συγκεκριμένη εφαρμογή.

5^η Φάση: Συγγραφή κώδικα και υλοποίηση σεναρίων.

Έγινε συγγραφή πρωτότυπου κώδικα σε περιβάλλον Matlab έτσι ώστε να εισέρχονται τα απαραίτητα δεδομένα για το κάθε σενάριο στο σύστημα, να εκτελούνται οι εντολές δημιουργίας και εκπαίδευσης του νευρωνικού δικτύου, να εκτυπώνονται τα γραφήματα των αποτελεσμάτων σε αρχείο για την μετέπειτα παρουσίασή τους και, τέλος, να πραγματοποιείται η πρόβλεψη. Ο κώδικας είναι γραμμένος έτσι, ώστε να είναι εύκολη η παραμετροποίηση των δικτύων για τις ανάγκες του κάθε σεναρίου.

6^η Φάση: Σχολιασμός και σύγκριση των αποτελεσμάτων και των προβλέψεων. Έγινε εκτενής ανάλυση των αποτελεσμάτων της εκπαίδευσης των νευρωνικών δικτύων για κάθε σενάριο, ενώ στην συνέχεια σχολιάστηκαν οι παραγόμενες προβλέψεις και έγινε σύγκριση μεταξύ τους, καθώς και με την πρόβλεψη της μελέτης επάρκειας φορτίου του ΑΔΜΗΕ.

Στατιστικές Μέθοδοι Πρόβλεψης Ηλεκτρικού Φορτίου

Εισαγωγή

2.

2.1 Η πρόβλεψη της κατανάλωσης ηλεκτρικής ενέργειας αποτελεί απαραίτητη προϋπόθεση για την αποτελεσματική κάλυψη της ζήτησης αλλά και την εύρυθμη λειτουργία του δικτύου μεταφοράς και διανομής. Καθώς η ηλεκτρική ενέργεια δεν είναι εφικτό να αποθηκευτεί ως έχει, η παραγωγή της θα πρέπει να είναι ταυτόσημη με την ζήτησή της, σε κάθε στιγμή. Σε αντίθετη περίπτωση κινδυνεύει η ευστάθεια του δικτύου, οι μονάδες παραγωγής και τα επιμέρους τμήματα αυτών, αλλά και οι συνδεδεμένες ηλεκτρικές συσκευές των καταναλωτών. Η βύθιση τάσης του δικτύου και η διακοπή της ηλεκτροδότησης είναι απαγορευτική όχι μόνο για τεχνικούς, αλλά και για κοινωνικούς και οικονομικούς λόγους. Η αγορά της ηλεκτρικής ενέργειας στηρίζει την βιομηχανία, τις αστικές και τις κρατικές δομές, όπως τα νοσοκομεία, τις τηλεπικοινωνιακές εγκαταστάσεις, ιδρύματα και άλλες μονάδες για τις οποίες είναι καταστροφικό να μείνουν χωρίς ηλεκτροδότηση, έστω και για μικρό χρονικό διάστημα. Παράλληλα, η σωστή πρόβλεψη της ζήτησης είναι αναγκαία για τον βέλτιστο οικονομικό σχεδιασμό των μονάδων ηλεκτροπαραγωγής και την ρύθμιση της αγοράς ενέργειας με το μικρότερο δυνατό κόστος.

Υπάρχει πληθώρα μεθόδων πρόβλεψης, οι οποίες βασίζονται σε κλασσικές μαθηματικές μεθόδους όπως η απλή παλινδρόμηση, η μη γραμμική παλινδρόμηση, η μέθοδος των κινητών μέσων όρων κ.α. [1] Στις επόμενες ενότητες θα γίνει μια σύντομη αναφορά αυτών, αλλά θα εξηγηθούν και οι λόγοι για τους οποίους είναι χρήσιμη η έρευνα και η μελέτη νέων μεθόδων, οι οποίες, με την διεύρυνση της υπολογιστής ισχύος, είναι ικανές να προσαρμόζονται άμεσα στις αλλαγές των παραγόντων που επηρεάζουν την ζήτηση και εκτιμούν την μελλοντικές τιμές της σε μακροπρόθεσμο ή βραχυπρόθεσμο ορίζοντα σε πραγματικό χρόνο. [2]

2.2

Παράγοντες που επηρεάζουν την ζήτηση ηλεκτρικής ενέργειας

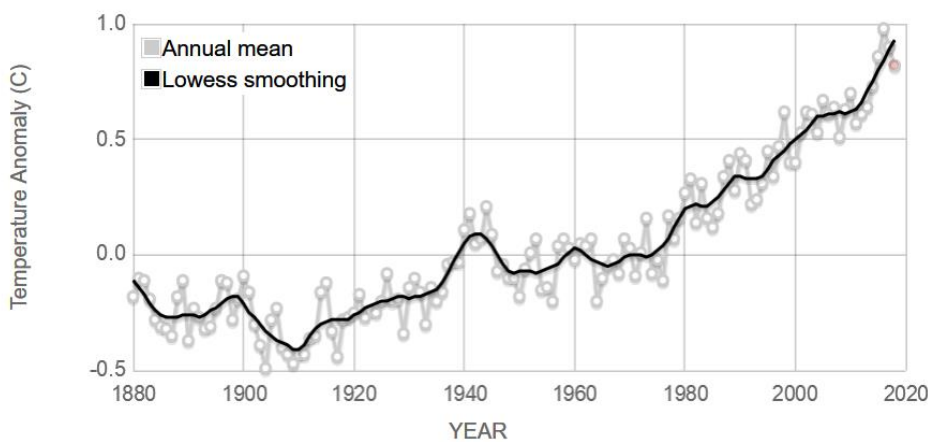
Οι προβλέψεις ζήτησης ηλεκτρικής ενέργειας, όπως αναλύεται παραπάνω, είναι απαραίτητες για την σωστή και ευσταθή λειτουργία του συστήματος μεταφοράς και διανομής ηλεκτρικής ενέργειας. Χωρίζονται σε διάφορες κατηγορίες, η κάθε μία από αυτές επιτελεί ξεχωριστό σκοπό. Όταν η πρόβλεψη είναι της τάξεως των μερικών δευτερολέπτων έως και μερικών λεπτών τότε μιλάμε για πολύ βραχυπρόθεσμη πρόβλεψη. Αντίστοιχα, η πρόβλεψη μερικών ωρών έως και μία εβδομάδα θεωρείται βραχυπρόθεσμη, ενώ η μεσοπρόθεσμη αποτελεί την εκτίμηση μερικών μηνών έως 5 ετών. Τέλος, η μακροπρόθεσμη κυμαίνεται από 10 έως και 25 έτη. [3] Στα πλαίσια της παρούσας εργασίας θα γίνει προσπάθεια πρόβλεψης σε μεσοπρόθεσμο και μακροπρόθεσμο ορίζοντα.

Με βάση την μελέτη επάρκειας του ΑΔΜΗΕ 2017-2020 για τον ελλαδικό χώρο, οι κύριοι παράγοντες που επιδρούν στη διαμόρφωση της ζήτησης ηλεκτρικής ενέργειας στη χώρα σε μεσομακροπρόθεσμη βάση είναι:

- Οι οικονομικές συνθήκες της χώρας, με βασικό δείκτη μέτρησης το ΑΕΠ.
- Οι αλλαγές στις καταναλωτικές συνήθειες (κλιματισμός, χρήση ηλεκτρισμού στις μεταφορές, χρήση υπολογιστών, χρήση λαμπτήρων LED κ.λπ.) λόγω βελτίωσης βιοτικού επιπέδου, αλλά και η βελτίωση των συνθηκών διαβίωσης συγκεκριμένων πληθυσμιακών ομάδων (π.χ. οικονομικοί μετανάστες).

- Η γενικότερη κατάσταση του ενεργειακού τομέα και της αγοράς ηλεκτρισμού (επίπεδο τιμών ηλεκτρικής ενέργειας, ανταγωνισμός με Φυσικό Αέριο κ.λπ.).
- Ειδικές συνθήκες (π.χ. υλοποίηση έργων Κοινοτικού Πλαισίου Στήριξης).
- Πληθυσμιακή εξέλιξη. Διάφορα μέτρα εξειδίκευσης πολιτικών, όπως εξοικονόμηση ενέργειας, περιβαλλοντικοί περιορισμοί, κ.λπ. [4]

Στα παραπάνω, είναι απαραίτητο να προστεθεί η θερμοκρασιακή μεταβολή. Αν και οι μεταβολές στην θερμοκρασία επηρεάζουν κυρίως βραχυ- μεσοπρόθεσμα το φορτίο, λόγω των φυσιολογικών εποχικών διακυμάνσεων, η θερμοκρασία τις τελευταίες δεκαετίες έχει ανοδική πορεία, λόγω της κλιματικής αλλαγής. Αυτό έχει ως αποτέλεσμα την μεταβολή των βραχυπρόθεσμων προτύπων των θερμοκρασιακών αυξομειώσεων, καθώς και την γενικότερη αυξανούσα ζήτηση σε ηλεκτρική ενέργεια, όσο η θερμοκρασία ανεβαίνει.



Γράφημα 2-1: Παγκόσμια αύξηση μέσης θερμοκρασίας την περίοδο 1880-2020. Πηγή: NASA's Goddard Institute for Space Studies (GISS).

2.3

Μέθοδοι πρόβλεψης

Οι συγκεντρωμένες καταγραφές κατανάλωσης ηλεκτρικής ενέργειας σε σταθερή βάση αποτελούν χρονοσειρές οι οποίες παρουσιάζουν συγκεκριμένα χαρακτηριστικά και διέπονται από κανόνες οι οποίοι έχουν αναλυθεί από ερευνητές με σκοπό την πρόβλεψη των μελλοντικών τιμών τους. Στην βιβλιογραφία υπάρχει πληθώρα κλασικών στατιστικών μεθόδων ανάλυσης χρονοσειρών οι οποίες έχουν εφαρμοστεί ευρέως και αποτελούν λύσεις για την πρόβλεψη του μελλοντικού φορτίου.

Τα βασικά χαρακτηριστικά των χρονοσειρών είναι [5]:

- Στασιμότητα: Μία χρονοσειρά παρουσιάζει διακυμάνσεις οι οποίες δεν διαφοροποιούνται σε βάθος χρόνου. Τέτοιες διαφοροποιήσεις είναι οι τάσεις, η περιοδικότητα και η εποχικότητα.
- Αιτιοκρατία και στοχαστικότητα: Μια μεταβλητή που καταγράφεται συστηματικά και συνθέτει μια χρονοσειρά συνήθως επηρεάζεται από πληθώρα εξωγενών τυχαίων παραγόντων, με αποτέλεσμα οι διακυμάνσεις της να μην ακολουθούν πάντα κάποια συγκεκριμένη αιτία και να παρουσιάζεται θόρυβος εντός της χρονοσειράς, ο οποίος δίνει το χαρακτηριστικό της στοχαστικότητας.
- Γραμμικότητα και μη- γραμμικότητα: Οι ανεξάρτητες μεταβλητές που συμβάλλουν στις διακυμάνσεις και τις αυξομειώσεις της μελετώμενης μεταβλητής ενδέχεται να έχουν

γραμμική ή μη γραμμική συσχέτιση με αυτήν. Έτσι η πορεία της μεταβλητής στο χρόνο μπορεί να είναι γραμμικό σύστημα των παραπάνω μεταβλητών ή μη.

Στις παρακάτω παραγράφους ακολουθεί μια συνοπτική περιγραφή των κλασικών μεθόδων ανάλυσης και πρόβλεψης χρονοσειρών κατανάλωσης ηλεκτρικής ενέργειας:

Ανάλυση παλινδρόμησης

Αποτελεί μια διαδεδομένη και απλή μέθοδο πρόβλεψης. Η μέθοδος μαθηματικοποιεί την σχέση της μελλοντικής κατανάλωσης με βάση τις προηγούμενες καταναλώσεις, τις καιρικές συνθήκες, τον τύπο των καταναλωτών αλλά και των ημερών κατανάλωσης.[2] Έτσι προκύπτει η παρακάτω μαθηματική σχέση, η οποία κατηγοριοποιεί σε δύο συνιστώσες. Η πρώτη είναι η κανονική ζήτηση και η δεύτερη είναι μια γραμμική συσχέτιση μεταξύ των παραγόντων που επηρεάζουν την κατανάλωση.

$$L(t) = L_n(t) + \sum_{\text{Εξίσωση 1.2}} a_i x_i(t) + e(t)$$

Όπου $L_n(t)$ η κανονική κατανάλωση για την χρονική στιγμή t , a_i ο εκτιμώμενος συντελεστής αργής μεταβολής φορτίου, $x(i)$ οι ανεξάρτητοι παράγοντες που επηρεάζουν το φορτίο (π.χ. τα καιρικά φαινόμενα), $e(t)$ ο θόρυβος και n ο αριθμός των παρατηρήσεων για τον προσδιορισμό των συντελεστών.

Πολλαπλή παλινδρόμηση

Είναι η πλέον εφαρμοσμένη μέθοδος, η οποία βασίζεται στην απλή παλινδρόμηση, αλλά λαμβάνει υπ' όψιν πληθώρα παραγόντων οι οποίοι αποδεδειγμένα επηρεάζουν το φορτίο [6] όπως οι μετεωρολογικές συνθήκες, η αύξηση του ΑΕΠ, οι τιμές της ηλεκτρικής ενέργειας, η πληθυσμιακή αύξηση. Η μέθοδος βασίζεται στην παλινδρόμηση των ελαχίστων τετραγώνων και το μοντέλο πραγματοποιεί την πρόβλεψη ως εξής:

$$Y(t) = V(t) a(t) + e_t \quad [7]$$

Εξίσωση 1.3

Όπου t , ο χρόνος δειγματοληψίας, $Y(t)$ το ολικό μετρούμενο φορτίο, $V(t)$ το διάνυσμα που περιέχει τις προσαρμοζόμενες μεταβλητές όπως η χρονική περίοδος, η θερμοκρασία, η ταχύτητα του ανέμου, η υγρασία, ο τύπος της ημέρας (καθημερινή ή αργία) κ.α. , $a(t)$ το διάνυσμα των συντελεστών της παλινδρόμησης και e_t το σφάλμα του μοντέλο για την χρονική στιγμή t .

Εκθετική εξομάλυνση

Το μοντέλο εκθετικής εξομάλυνσης είναι εμπειρική μέθοδος εξομάλυνσης χρονοσειρών. Μοντελοποιούν προηγούμενες τιμές φορτίου μέσω μιας συνάρτησης προσαρμογής και με βάση αυτήν κάνουν πρόβλεψη των επόμενων τιμών. Κάποιες προσεγγίσεις έχουν το πλεονέκτημα ότι είναι ικανές να αναλύσουν εποχιακές χρονοσειρές απευθείας, χωρίς προ-επεξεργασία των δεδομένων. Συνήθως χρησιμοποιούνται σε συνδυασμό με άλλες μεθόδους πρόβλεψης. Η μελέτη των Moghram and Rahman προβλέπει το φορτίο με εξίσωση της μορφής:

$$y(t) = \beta(t)T f(t) + e(t) \quad [8]$$

Εξίσωση 1.4

Όπου $f(t)$ το διάνυσμα της συνάρτησης προσαρμογής, $\beta(t)$ ο συντελεστής του διανύσματος, $e(t)$ το σφάλμα λόγω του θορύβου και T ο τελεστής μεταφοράς.

Αυτοπαλινδρομικά μοντέλα (AutoRegressive Models)

Τα αυτοπαλινδρομικά μοντέλα είναι στοχαστικές προσεγγίσεις για την μοντελοποίηση και πρόβλεψη χρονοσειρών. Είναι αρκετά αποτελεσματικά γιατί στην για την προσέγγιση αναλύουν εσωτερικά χαρακτηριστικά της χρονοσειράς όπως η αυτοσυσχέτιση των τιμών της, η τάση και η εποχική διακύμανση.

Το μοντέλο των [9], θεωρώντας ότι το μελλοντικό φορτίο είναι ένας γραμμικός συνδυασμός προηγούμενων τιμών, μοντελοποιεί την χρονοσειρά κατανάλωσης ηλεκτρικής ενέργειας ως εξής:

$$Lk = \sum_{i=1}^m a_i k \cdot i + e_k$$

Εξίσωση 1.5

όπου Lk είναι το προβλεπόμενο φορτίο για τον χρόνο k (min), e_k είναι η στοχαστική μεταβολή του k και a_i για $i=1 \dots m$ είναι οι άγνωστοι συντελεστές.

Αυτοπαλινδρομικά μοντέλα κινητού μέσου όρου (Moving Average Autoregressive Models)

Αποτελούν διεύρυνση των απλών αυτοπαλινδρομικών μοντέλων, με περισσότερους όρους (έναν αυτοπαλινδρομικό και έναν για τον κινητό μέσο όρο) και πιο ευέλικτη προσαρμογή στην μελετώμενη χρονοσειρά. Η εκτίμηση του φορτίου της χρονοσειράς $y(t)$ γίνεται γραμμικά με βάση προηγούμενες περιόδους $[(y(t-1), y(t-2), \dots)]$ της χρονοσειράς και με $[a(t), a(t-1)]$ προηγούμενες περιόδους του λευκού θορύβου. Η μαθηματική σχέση ενός μοντέλου ARMA (AutoRegressive Moving Average) για (p,q) γράφεται ως εξής:

$$y(t) = \phi_1 y(t-1) + \dots + \phi_p y(t-p) + a(t) - \phi_1 a(t-1) - \dots - \phi_q a(t-q)$$

Εξίσωση 1.6

Διάφορες προσεγγίσεις μοντέλων, όπως αυτή των [10] αναλύσει την χρονοσειρά του φορτίου σε επιμέρους στοχαστικά ή ντετερμινιστικά μέρη, εφαρμόζοντας το μοντέλο ARMA για τα πρώτα.

Ολοκληρωμένα αυτοπαλινδρομικά μοντέλα κινητού μέσου όρου (AutoRegressive Integrated Moving Average- ARIMA)

Στην μοντελοποίηση των ολοκληρωμένων αυτοπαλινδρομικών μοντέλων κινητού μέσου όρου πραγματοποιείται διαφόριση της χρονοσειράς, έτσι ώστε να σταθεροποιηθεί η δυναμική μεταβολή λόγω της εποχικότητας. Το μοντέλο παίρνει την μορφή:

$$\phi(B)\nabla^u P(t) = \theta(B)a(t)$$

όπου $P(t)$ η χρονοσειρά του φορτίου, $a(t)$ η ακολουθία λευκού θορύβου, $\phi(B)=1-\phi_1 B-\phi_2 B^2-\dots-\phi_p B^p$, το πολυώνυμο AR παραμέτρων, $\theta(B)=1-\theta_1 B-\theta_2 B^2-\dots-\theta_q B^q$, το πολυώνυμο MA παραμέτρων, B είναι ο τελεστής της προς τα πίσω ολίσθησης, ϕ_i, θ_i είναι σταθερές και $\nabla = 1-B$ είναι ο τελεστής προς τα πίσω διαφοράς

Τεχνητά Νευρωνικά Δίκτυα

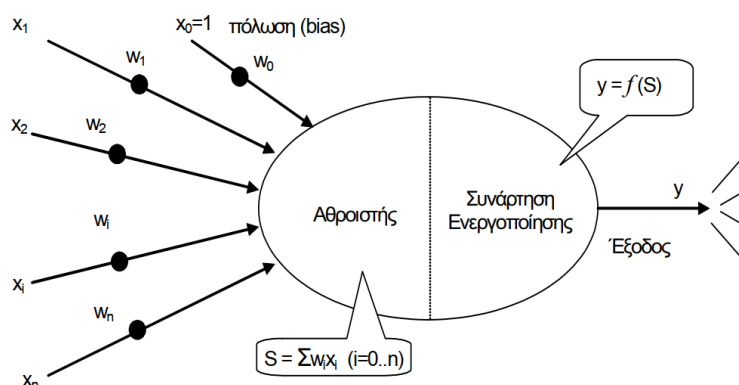
Γενικά

- Τα νευρωνικά δίκτυα ξεκίνησαν να αναπτύσσονται την δεκαετία του 1940 με πρωτοπόρους τον Warren McCulloch, Walter Pitts και D.O. Hebb [11]. Είναι εμπνευσμένα από την λειτουργία των βιολογικών νευρώνων στους εγκεφάλους βιολογικών οργανισμών, όπως αυτόν του ανθρώπου. Ένα δίκτυο αποτελείται από κόμβους οι οποίοι είναι τεχνητοί νευρώνες (nodes ή neurons) που μεταφέρουν πληροφορία μεταξύ τους (κυρίως προς μία κατεύθυνση). Η μεταφορά των δεδομένων ανάμεσα στους νευρώνες ορίζεται από την συνάρτηση ενεργοποίησης ή οποία είναι μία μη γραμμική, συνήθως, σιγμοειδής συνάρτηση της μορφής:

$$\Phi(S) = \frac{1}{1 + e^{-a \cdot S}}$$

Εξίσωση 3-1 (Πηγή: I. Βλαχαβάς [12])

Ο κάθε νευρώνας του δικτύου δέχεται τα δεδομένα όλων των προηγούμενων νευρώνων, τροποποιημένα αναλόγως το συναπτικό βάρος (synaptic weight) της σύνδεσης του κάθε νευρώνα. Στην συνέχεια αθροίζονται όλα τα δεδομένα στον νευρώνα και με βάση την συνάρτηση ενεργοποίησης, η πληροφορία μεταδίδεται στα επόμενα επίπεδα. Παρακάτω ακολουθεί μια γραμμική απεικόνιση νευρώνα ενός τεχνητού νευρωνικού δικτύου:



[12]

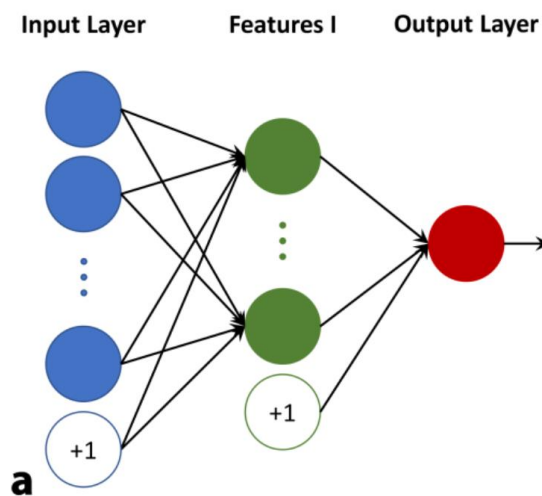
Εικόνα 3-1: Μοντέλο τεχνητού νευρώνα (Πηγή: I. Βλαχαβάς [12])

Η έξοδος του κάθε νευρώνα μπορεί να είναι 0 ή 1 αναλόγως εάν το άθροισμα $\sum x_i w_i$ είναι μεγαλύτερο ή μικρότερο από την τιμή κατωφλιού. Όπως και τα συναπτικά βάρη, έτσι και η τιμή κατωφλιού είναι πραγματικός αριθμός που αποτελεί παράμετρο του νευρώνα. Η αλγεβρική απεικόνιση έχει ως εξής:

$$\text{έξοδος} = \begin{cases} 0 & \text{εάν } \sum x_i w_i \leq \text{τιμή κατωφλιού} \\ 1 & \text{εάν } \sum x_i w_i > \text{τιμή κατωφλιού} \end{cases}$$

Εξίσωση 2.1 (Πηγή: M. Nielsen [13])

Ο τρόπος σύνδεσης και μετάδοσης της πληροφορίας μεταξύ των νευρώνων διαφέρει αναλόγως τον τύπο του δικτύου. Οι δύο βασικοί τύποι είναι τα δίκτυα πρόσθιας τροφοδότησης και τα επαναληπτικά δίκτυα. Η ειδοποιός διαφορά μεταξύ τους είναι ότι στα δίκτυα πρόσθιας τροφοδότησης η πληροφορία μεταδίδεται προς μία κατεύθυνση στους νευρώνες, ενώ τα επαναληπτικά δίκτυα έχουν συνδέσεις ανάδρασης από την έξοδο προς την είσοδο. [14]



[15][15][15][15]

Εικόνα 3-2 Νευρωνικό Δίκτυο Πρόσθιας Τροφοδότησης (Πηγή: H. Hamedmoghadam, [15])

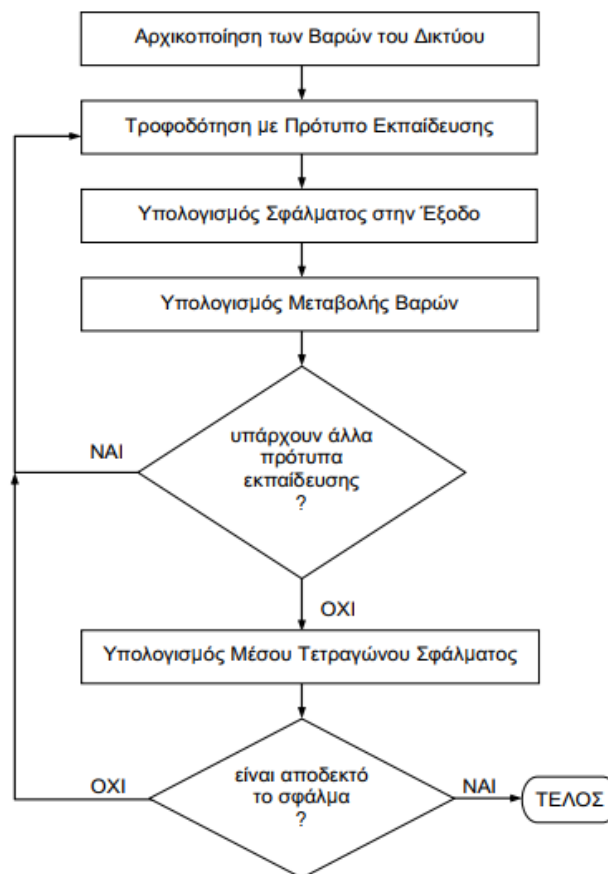
Η διαδικασία που καθιστά τα νευρωνικά δίκτυα λειτουργικά και χρήσιμα για πληθώρα εφαρμογών και μοντελοποιήσεων είναι η εκπαίδευση τους. Η διαδικασία της εκπαίδευσης τροποποιεί τα συναπτικά βάρη έτσι ώστε οι έξοδοι του δικτύου να ανταποκρίνονται στις ανάγκες του εκάστοτε προβλήματος με βάση τα ήδη υπάρχοντα δεδομένα που το περιγράφουν. Η εκπαίδευση γίνεται με βάση αλγόριθμους οι οποίοι αξιολογούν τις εξόδους του δικτύου σε κάθε κύκλο εκπαίδευσης (epoch) με σκοπό την ελαχιστοποίηση του σφάλματος.

Υπάρχει πληθώρα αλγορίθμων εκπαίδευσης ο κάθε ένας από τους οποίους χρησιμοποιεί διαφορετική μεθοδολογία για την εκπαίδευση του δικτύου, τον διαχωρισμό των δεδομένων και την βελτιστοποίηση του νευρωνικού δικτύου. Η επιλογή του καταλληλότερου αλγορίθμου εξαρτάται από την φύση των δεδομένων του προβλήματος και την κρίση του μελετητή. Οι πιο γνωστοί και ευρέως διαδεδομένοι αλγόριθμοι εκπαίδευσης είναι:

- ο κανόνας δέλτα
- ο αλγόριθμος ανάστροφης μετάδοσης λάθους (back propagation)
- η ανταγωνιστική
- η τυχαία μάθηση [12]

Στα πλαίσια της παρούσας εργασίας έγιναν δοκιμές εκπαίδευσης δικτύων με διαφορετικούς αλγόριθμους που είναι διαθέσιμοι στο πακέτο της Matlab, αλλά ο αποδοτικότερος είναι ο αλγόριθμος Levenberg-Marquardt.

Ο συγκεκριμένος αλγόριθμος είναι ταχύς, διότι παρακάμπτει τον υπολογισμό του Εσσιανού πίνακα, δηλαδή του πίνακα των δεύτερων παραγώγων της συνάρτησης που προκύπτει κατά την εκπαίδευση. Επίσης είναι αξιόπιστος, καθώς βασίζεται στον γενικευμένο κανόνα Δέλτα, αλλά σε αντίθεση με τον τελευταίο, καθορίζει το ποσοστό του συνολικού σφάλματος που αντιστοιχεί σε κάθε νευρώνα, ακόμη και αυτών που ανήκουν στα κρυφά επίπεδα. Η αναπροσαρμογή των βαρών γίνεται από το επίπεδο εξόδου προ το επίπεδο εισόδου, και αυτό το χαρακτηριστικό είναι που ονομάζεται ανάστροφο πέρασμα (backward pass) ή ανάστροφη μετάδοση (back propagation). [16] Η βελτιστοποίηση του σφάλματος (ελαχιστοποίησή του) πραγματοποιείται με την διαδικασία βελτιστοποίησης επικλινούς καθόδου (gradient descent optimization procedure) η οποία ελαχιστοποιεί το μέσο τετραγωνικό σφάλμα μεταξύ της εξόδου του δικτύου και της επιθυμητής εξόδου για τα p διανύσματα εκπαίδευσης. [13]



Εικόνα 3-3: Σχηματική απεικόνιση αλγορίθμου ανάστροφης λάθους (Πηγή: I. Βλαχάβας et. Al [12])

Χρήση Νευρωνικών Δικτύων σε προβλέψεις

Τα νευρωνικά δίκτυα, λόγω της δομής και της φύσης τους εκπαιδεύονται και προσαρμόζονται με μεγάλη ευκολία σε μη γραμμικές εφαρμογές. Έτσι αποτελούν ευρύ πεδίο μελέτης για μοντελοποιήσεις και προβλέψεις οι οποίες δεν έχουν γραμμικό χαρακτήρα, λόγω των πολλών μεταβλητών από τις οποίες επηρεάζονται και της στοχαστικότητας των δεδομένων. Τέτοια περίπτωση αποτελεί και η πρόβλεψη της καμπύλης φορτίου, η οποία, όπως έχει αναλυθεί σε προηγούμενα κεφάλαια, διαμορφώνεται από πληθώρα παραγόντων, οι οποίοι δεν είναι εφικτό να μοντελοποιηθούν με απόλυτη ακρίβεια μέσω ντετερμιστικών μαθηματικών μοντέλων.

Ισχυρό πλεονέκτημα των νευρωνικών δικτύων είναι ότι μπορούν με ευκολία να αναγνωρίσουν τις μη γραμμικές συσχετίσεις και αλληλεπιδράσεις μεταξύ των εισόδων και των εξόδων ενός συστήματος. Η ικανότητα τους να μαθαίνουν και να συγκρατούν πληροφορίες σχετικά με την μη γραμμικότητα ενός συστήματος έχει προσελκύσει το ενδιαφέρον αρκετών ερευνητών. Π.χ. ένα radial basis function νευρωνικό δίκτυο έχει χρησιμοποιηθεί ως ελεγκτής για συστήματα ελέγχου ροής (UPFC) με αποτέλεσμα την βελτίωση της σταθερότητας του συστήματος. [17]

Τα πιο συνηθισμένα συστήματα πρόβλεψης βασίζονται σε νευρωνικά δίκτυα πρόσθιας τροφοδότησης (feed forward) εκπαιδευόμενα με μέθοδο ανάστροφης μετάδοσης λάθους (back propagation). [9] Στην μελέτη των Djukanovic et. Al [18] χρησιμοποιήθηκε αλγόριθμος επιβλεπόμενης και μη μάθησης για την εκπαίδευση του δικτύου στην συσχέτιση του καταναλισκόμενου φορτίου και της θερμοκρασίας περιβάλλοντος για δεδομένη εποχή, τύπο και ώρα της ημέρας. Με την χρήση του προαναφερόμενου αλγορίθμου έγινε πρόβλεψη ωριαίου φορτίου με ορίζοντα 24^{ωv} ωρών. Ο Παπαλεξόπουλος [3] σχεδίασε και εφάρμοσε ένα τεχνητό νευρωνικό δίκτυο για το κέντρο ελέγχου ενέργειας την εταιρία «Pacific Gas and Electric Company». Ιδιαίτερη έμφαση δόθηκε στην σωστή μοντελοποίηση των ιδιοτήτων της καμπύλης φορτίου, όπως οι αργίες, τα θερμά ρεύματα αέρα, οι απρόβλεπτες εκδηλώσεις κακοκαιρίας κ.α. Οι Srinivasan et. Al χρησιμοποίησαν ANN (Artificial Neural Network) ανάστροφης μετάδοσης λάθους για πρόβλεψη καμπύλης φορτίου και επισήμαναν την υπεροχή του μοντέλου σε σχέση με τις παραδοσιακές μεθόδους [19]. Ο Liu et Al. [9] σύγκρινε ένα οικονομετρικό μοντέλο και ένα τεχνητό νευρωνικό δίκτυο, μέσω μελέτη περίπτωσης επάνω στην κατανάλωση ενέργειας στην Σιγκαπούρη αλλά έδειξε πως ένα πλήρως εκπαιδευμένο δίκτυο με καλή απόδοσης προσαρμογής ίσως να μην πραγματοποιεί σωστή πρόβλεψη μελλοντικών τιμών.

Τεχνητά Νευρωνικά δίκτυα στην Matlab

Χρησιμοποιώντας εντολές σε κώδικα Matlab είναι εφικτή η δημιουργία αλγορίθμων οι οποίοι επεξεργάζονται τα δεδομένα εισόδου (χρονοσειρές), εκπαιδεύουν το δίκτυο και τυπώνουν αυτόματα τα αποτελέσματα της εκπαίδευσης. Στην συνέχεια τροποποιούν το δίκτυο, έτσι ώστε να είναι ικανό να παράγει πρόβλεψη, και, τέλος, πραγματοποιούν και εξάγουν τα αποτελέσματα της πρόβλεψης.

Η βασική μεθοδολογία σχεδιασμού ενός τεχνητού νευρωνικού δικτύου στην Matlab είναι [20]:

1. Προσπέλαση και προετοιμασία δεδομένων
2. Δημιουργία Νευρωνικού Δικτύου
3. Παραμετροποίηση εισόδων – εξόδων του δικτύου
4. Παραμετροποίηση των παραμέτρων του δικτύου
5. Εκπαίδευση του δικτύου
6. Αξιολόγηση των αποτελεσμάτων

7. Ενσωμάτωση του δικτύου σε γραμμή παραγωγής

Οι εντολές που χρησιμοποιήθηκαν για την συγκεκριμένη μελέτη περιλαμβάνονται στο πακέτο Deep Learning Toolbox της Matlab και παρέχουν πληθώρα διαφορετικών παραμετροποιημένων τεχνητών νευρωνικών δικτύων για διάφορες εφαρμογές μεταξύ αυτών η προσαρμογή συνάρτησης, η ταυτοποίηση μοτίβων, η πρόβλεψη χρονοσειράς κ.α. Παρακάτω αναλύονται τα είδη που χρησιμοποιήθηκαν:

Non Linear Autoregressive Neural Network (NAR)

Εκπαιδεύει το νευρωνικό δίκτυο και προβλέπει την χρονοσειρά κατανάλωσης ενέργειας $y(t)$ με δεδομένα d προηγούμενα βήματα της ίδιας χρονοσειράς, κοινώς του διαθέσιμου ιστορικού κατανάλωσης ενέργειας για την μελετώμενη πόλη.

Μαθηματική έκφραση μη αυτοπαλινδρομικού μοντέλου

$$y(t) = f(y(t-1), y(t-2), \dots, y(t-p)) + e(t)$$

Εξίσωση 3-2: Μαθηματική έκφραση μοντέλου NAR

Όπου y η μελετώμενη μεταβλητή. Οι τιμές της μεταβλητής y εξαρτώνται από τις προηγούμενες p τιμές της ίδιας μεταβλητής. Ο όρος e_i είναι η ποσοτικοποίηση του σφάλματος (θόρυβος).

Αυτή η μέθοδος μπορεί να λειτουργήσει ενδεικτικά, για την αξιολόγηση της εγκυρότητας των αποτελεσμάτων της πρώτης μεθόδου. Δηλαδή να δούμε πόσο καλύτερα εκπαιδεύεται και προβλέπει την κατανάλωση ενέργειας ένα νευρωνικό δίκτυο, όταν του δώσουμε επιπλέον δεδομένα που είναι επιστημονικά αποδεδειγμένο ότι επηρεάζουν την μελετώμενη μεταβλητή (κατανάλωση ενέργειας).

Non Linear Autoregressive with External (Exogenous) Input (NARX)

Πρόκειται για εκπαίδευση ενός νευρωνικού δικτύου για την πρόβλεψη μια χρονοσειράς $y(t)$ με δεδομένα ένα πλήθος d προηγούμενων βημάτων της ίδιας χρονοσειράς και ένα αντίστοιχο πλήθος βημάτων μια τρίτης (εξωτερικής) χρονοσειράς $x(t)$ η οποία αποτελείται από όμοιο αριθμό βημάτων (τιμών) μίας ανεξάρτητης μεταβλητής η οποία θα μπορούσε να βελτιώνει την ακρίβεια των προβλέψεων. Στην περίπτωση μας η χρονοσειρά $y(t)$ αποτελεί την καταγεγραμμένη ζήτηση ηλεκτρικής ενέργειας (σε Kwh ή Mwh) και η χρονοσειρά $x(t)$ είναι ή οι τιμές του ΑΕΠ, ή η θερμοκρασία περιβάλλοντος, ή ο πληθυσμός της πόλης που μελετάται.

Για την εφαρμογή τέτοιου τύπου δικτύου, όπως συμπεραίνεται παραπάνω, θα πρέπει να είναι γνωστές οι τιμές της χρονοσειράς $x(t)$ για τις οποίες θα γίνουν οι μελλοντικές προβλέψεις.

Τα μη γραμμικά αυτοπαλινδρομικά μοντέλα με εξωγενή είσοδο (NARX) εμπεριέχουν και έναν όρο σφάλματος, ο οποίος σχετίζεται με το γεγονός ότι η γνώση των λοιπών όρων δεν σημαίνει πως η επερχόμενη τιμή της χρονοσειράς θα προβλεφθεί με ακρίβεια.

Η μαθηματική περιγραφή ενός δικτύου NARX μπορεί να γραφεί ως εξής:

$$y_t = F(y_{t-1}, y_{t-2}, y_{t-3}, \dots, u_t, u_{t-1}, u_{t-2}, u_{t-3}, \dots) + \varepsilon_t$$

Εξίσωση 3-3: Μαθηματική περιγραφή μοντέλου NARX

Όπου y η μελετώμενη μεταβλητή, και u η εξωγενής ανεξάρτητη μεταβλητή. Στην παραπάνω έκφραση φαίνεται πως η πληροφορία για την τιμή του u βοηθά στην πρόβλεψη της y , όπως βοηθούν και οι προηγούμενες τιμές της y . Ο όρος του σφάλματος είναι ο ε_t (θόρυβος). [21]

Είναι η καταλληλότερη μέθοδος για την συγκεκριμένη μελέτη, καθώς εκτός των άλλων, μελετάμε την επίδραση εξωτερικών παραγόντων, όπως αναφέρεται παραπάνω, οι οποίοι επηρεάζουν την κατανάλωση ενέργειας.

Μεθοδολογία

- Στο παρόν κεφάλαιο αναλύεται διεξοδικά η μεθοδολογία που ακολουθήθηκε για τον σχεδιασμό, την υλοποίηση, την προετοιμασία των δεδομένων για την εισαγωγή στα νευρωνικά δίκτυα, την εκπαίδευση των δικτύων και την πραγματοποίηση της πρόβλεψης. Επεξηγούνται οι δομές των νευρικών δικτύων που χρησιμοποιήθηκαν, καθώς και οι διαφορετικές λειτουργίες τους κατά την
4. διαδικασία της εκπαίδευσης και της πρόβλεψης. Έπειτα, παρουσιάζονται και επεξηγούνται τα πέντε διαφορετικά σενάρια βάσει των οποίων πραγματοποιήθηκαν οι προβλέψεις, ενώ αναφέρονται οι χρονοσειρές που συνδυάστηκαν για το καθένα.

Η μεθοδολογία παρουσιάζεται σε μορφή κειμένου, ενώ στο Παράρτημα στο τέλος του παρόντος τεύχους παρατίθεται ο αντίστοιχος κώδικας μέσω του οποίου εκτελούνται οι εντολές για την υλοποίηση, εκπαίδευση, εισαγωγή των δεδομένων, αλλαγή της λειτουργίας των δικτύων και την πραγματοποίηση της πρόβλεψης. Επίσης, γίνεται επισκόπηση των δεδομένων που θα χρησιμοποιηθούν στην εκπαίδευση καθώς και των χαρακτηριστικών που παρουσιάζουν οι χρονοσειρές αυτών.

Παράμετροι δικτύου

4.1 Διαστασιολόγηση Δικτύου

Κατά την διαστασιολόγηση του ενός νευρωνικού δικτύου επιδιώκουμε να βρούμε το δίκτυο εκείνο που για τα δεδομένα εισόδου αποδίδει ικανοποιητικά με το μικρότερο αριθμό νευρώνων. Πρόκειται για ένα πρόβλημα βελτιστοποίησης, του οποίου η περιπλοκότητα αυξάνει όσο αυξάνει το πλήθος των δεδομένων ενώ, επίσης, εξαρτάται από τον τύπο και την διακύμανση των δεδομένων της χρονοσειράς.

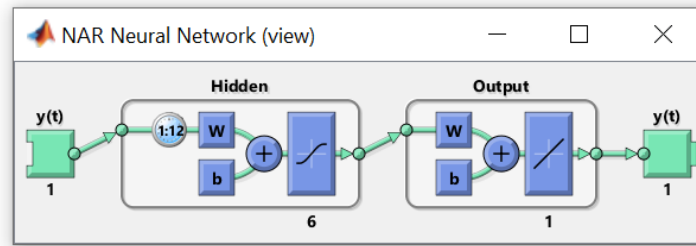
Υστερήσεις εισόδου (Input Delays)

Όταν το νευρωνικό δίκτυο βρίσκεται σε ανοιχτή μορφή, η εκπαίδευσή του γίνεται με την χρήση της χρονοσειράς για την οποία θέλουμε να προβλέψουμε τα μελλοντικά βήματα. Η ανατροφοδότηση γίνεται με τα προηγούμενα γνωστά βήματα εφαρμόζοντας μία υστέρηση. Η υστέρηση εξαρτάται από τα εσωτερικά χαρακτηριστικά της χρονοσειράς, όπως η αυτοσυσχετίσεις που παρουσιάζει ή η αν τα δεδομένα παρουσιάζουν εποχικότητα. Οι υστερήσεις στο δίνουν την δυνατότητα να αναγνωρίζει μοτίβα τα οποία υπάρχουν και έτσι να εκπαιδεύεται αρτιότερα επάνω στην χρονοσειρά.

Λειτουργία ανοιχτού βρόχου (Open Loop)

Για την πραγματοποίηση της εκπαίδευσης είναι απαραίτητο το δίκτυο να βρίσκεται σε λειτουργία ανοιχτού βρόχου. Πρακτικά αυτό σημαίνει ότι το δίκτυο κατά την εκπαίδευση δέχεται τα γνωστά δεδομένα (προηγούμενα βήματα της χρονοσειράς) ως είσοδο και με βάση αυτά προβλέπει τις επόμενες τιμές, οι οποίες είναι και αυτές γνωστές και με βάση την απόκλιση των προβλεπόμενων και των πραγματικών τιμών, κατά την εκπαίδευση, τροποποιεί αναλόγως τα συναπτικά βάρη των νευρώνων.

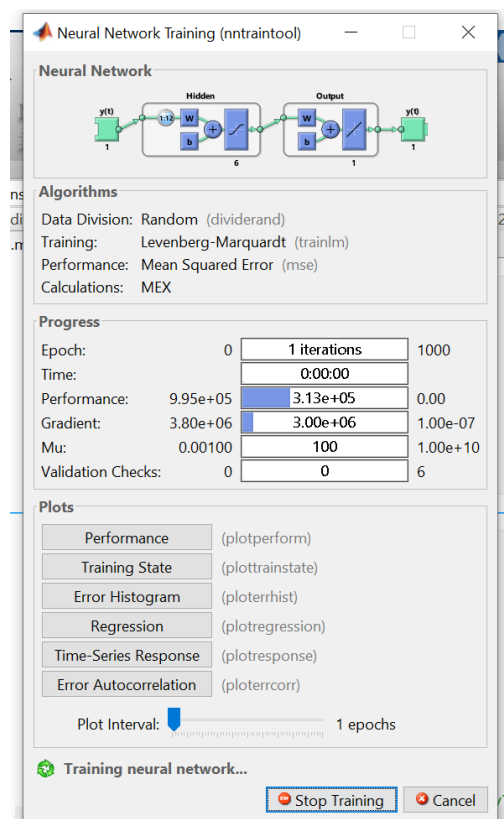
Στην παρακάτω εικόνα φαίνεται η δομή ενός δικτύου σε λειτουργία ανοιχτού βρόχου:



Εικόνα 4-1: Δομή νευρωνικού δικτύου τύπου NAR ανοιχτού βρόχου

Εκπαίδευση

Η εκπαίδευση πραγματοποιήθηκε με τον αλγόριθμο Levenberg-Marquardt μέσω του εργαλείου εκπαίδευσης «Training Tool». Με βάση τον αλγόριθμο η εκπαίδευση σταματά όταν το σφάλμα των δεδομένων εκπαίδευσης μειωθεί αισθητά και σταθεροποιηθεί, αλλά ταυτόχρονα δεν αυξηθεί το σφάλμα που προκύπτει από τα δεδομένα αξιολόγησης.



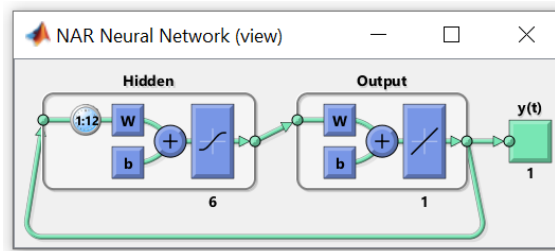
Εικόνα 4-2: Εργαλείο εκπαίδευσης νευρωνικού δικτύου σε περιβάλλον MATLAB

Λειτουργία κλειστού βρόχου (Closed Loop)

Για την πρόβλεψη πολλαπλών βημάτων της χρονοσειράς, εφόσον έχει ολοκληρωθεί η δημιουργία και η εκπαίδευση του νευρωνικού δικτύου, θα πρέπει να γίνει μετατροπή σε λειτουργία κλειστού βρόχου. Κατά την λειτουργία του κλειστού βρόχου το εκπαιδευμένο δίκτυο προβλέπει τα επόμενα βήματα της χρονοσειράς, χρησιμοποιώντας ως είσοδο την ανατροφοδότηση των προηγούμενων

προβλέψεων. Η υπολογιστική διαδικασία ξεκινά με είσοδο αρχικών τιμών, οι οποίες εξαρτώνται από την καθυστέρηση με την οποία έχει σχεδιαστεί το δίκτυο.

Στην παρακάτω εικόνα απεικονίζεται η δομή ενός δικτύου σε λειτουργία κλειστού βρόχου:

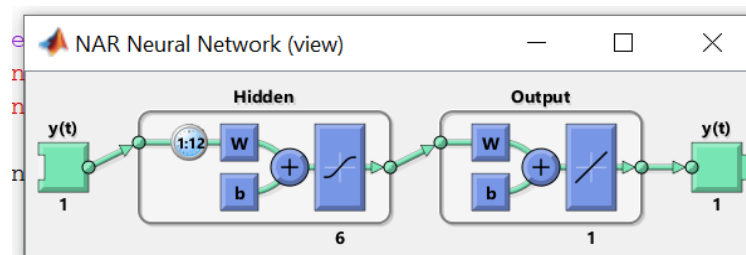


Εικόνα 4-3: Δομή νευρωνικού δικτύου τύπου NAR κλειστού βρόχου

Σχεδιασμός Σεναρίων

- 4.2 Λαμβάνοντας υπόψιν τα διαθέσιμα δεδομένα και τον στόχο της έρευνας δημιουργήθηκαν τύποι σεναρίων με βάση τους οποίους έγινε σχεδιασμός των αντίστοιχων νευρωνικών δικτύων. Σε κάθε σενάριο γράφτηκε κώδικας σε μορφή Live Script στο περιβάλλον της Matlab, ο οποίος τρέχει και αυτόματα εισάγει τις μεταβλητές στο Workspace, προετοιμάζει τα δεδομένα ανάλογα με τον τύπο του σεναρίου έτσι ώστε να γίνει σωστά η εκπαίδευση του δικτύου και εκπαιδεύει το δίκτυο. Μετά την ολοκλήρωση της εκπαίδευσης υπολογίζει την απόδοση του δικτύου, καθώς και τα σχετικά διαγράμματα τα οποία αξιολογούν την εκπαίδευση. Για τον σχεδιασμό των διαγραμμάτων απαιτείται η κατασκευή των αξόνων, και ο υπολογισμός των απαραίτητων μεταβλητών όπως είναι η διαφορά των εξαγόμενων τιμών από το δίκτυο σε σχέση με τις πραγματικές, τα σφάλματα και η κατανομή τους, αλλά και η αυτοσυσχέτιση που προκύπτει μεταξύ τους. Στην συνέχεια ο κώδικας μετατρέπει το ήδη εκπαιδευμένο δίκτυο σε μορφή κλειστού βρόχου, πραγματοποιεί την πρόβλεψη και εκτυπώνει το αντίστοιχο γράφημα. Οι κώδικες για το κάθε σενάριο παρατίθεται στο Παράρτημα στο τέλος του τεύχους της παρούσας εργασίας Σενάριο 1

Δίκτυο τύπου NAR με χρονοσειρά εισόδου την μηνιαία κατανάλωση ηλεκτρικής ενέργειας για την περίοδο από 01/2000 έως και 01/2019.



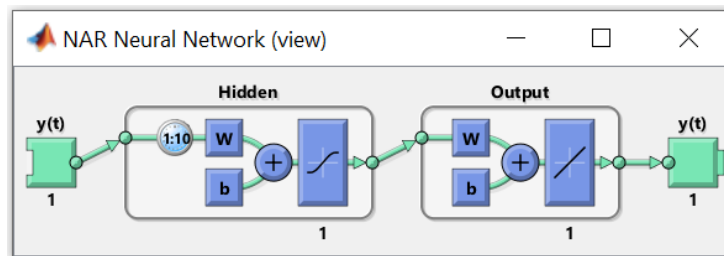
Εικόνα 4-4: Δομή NAR νευρωνικού δικτύου σεναρίου 1

Στην παραπάνω εικόνα απεικονίζεται η βασική δομή του δικτύου. Μετά από δοκιμές για διαφορετικό πλήθος νευρώνων στο κρυφό στρώμα, ο βέλτιστος αριθμός βρέθηκε να είναι οι 6 νευρώνες. Με αυτή την δομή το δίκτυο αποδίδει ικανοποιητικά κατά την εκπαίδευση, ενώ ταυτοχρόνως αποφεύγεται η υπερπροσαρμογή. Επίσης, η πρόβλεψη για την κατανάλωση ενέργειας κυμαίνεται σε ανεκτά, λογικά και αναμενόμενα επίπεδα.

Οι υστερήσεις του δικτύου επιλέχθηκε να είναι 12 καθώς η χρονοσειρά έχει μηνιαίο βήμα, οπότε αναμένεται να έχει ετήσια περιοδικότητα (12 μήνες).

Σενάριο 2

Στο σενάριο 2 χρησιμοποιήθηκε νευρωνικό δίκτυο πρόβλεψης χρονοσειράς τύπου NAR. Η εκπαίδευση έγινε με δεδομένα εισόδου την ετήσια κατανάλωση ενέργειας από το 1960 έως και το 2014.



Εικόνα 4-5: Δομή NAR νευρωνικού δικτύου σεναρίου 2

Όπως φαίνεται παραπάνω το δίκτυο έχει μόνο 1 νευρώνα στο κρυφό στρώμα. Η επιλογή έγινε μετά από δοκιμές με γνώμονα το γεγονός ότι είναι αναμενόμενο το δίκτυο να αποδίδει καλύτερα με μικρό αριθμό νευρώνων. Αυτό συμβαίνει διότι το πλήθος το δεδομένων είναι αρκετά μικρό (55 βήματα) οπότε μεγάλος αριθμός νευρώνων οδηγεί σε υπερπροσαρμογή του δικτύου.

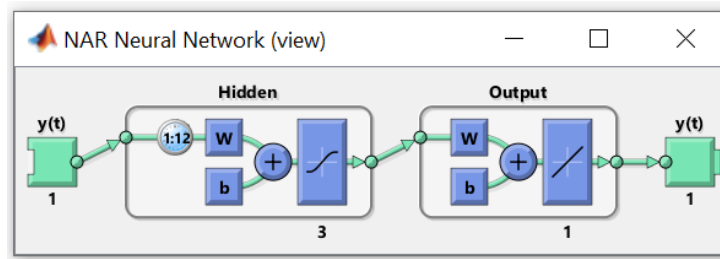
Για τον προσδιορισμό του βέλτιστου αριθμού υστερήσεων του δικτύου, σε αντίθεση με το Σενάριο 1, εδώ δεν έχουμε προφανή περιοδικότητα της χρονοσειράς. Εξετάζοντας την χρονοσειρά παρατηρούμε ότι δεν υπάρχει αυτοσυσχέτιση. Όπως φαίνεται γράφημα η κατανάλωση ενέργειας αυξάνει από το 1960 μέχρι και το 2008. Τα επόμενα έτη παρατηρείται μείωση με αποτέλεσμα το 2010 η κατανάλωση να βρίσκεται στα επίπεδα του 2006. Προφανώς και η κατανάλωση ηλεκτρικής ενέργειας είναι άμεσα εξαρτημένη από κοινωνικοπολιτικούς παράγοντες, όπως π.χ. η οικονομική κρίση που ξεκίνησε το 2009. Μέχρι και το 2014 που τελειώνει η χρονοσειρά, η κατανάλωση αυξομειώνεται διατηρώντας όμως φθίνουσα τάση με τελικά επίπεδα καταναλώσεις όμοιες με του 2003. Μετά από διάφορες δοκιμές, ο αριθμός υστερήσεων για τον οποίο το δίκτυο αποδίδει τα βέλτιστα είναι οι 10.

Σενάριο 3

Νευρωνικό δίκτυο τύπου NARX με εισόδους την κατανάλωση ενέργειας και την μέση θερμοκρασία της χώρας με μηνιαίο βήμα και εύρος από τον Ιούλιο του 2009 έως και τον Ιανουάριο του 2019. Αν και υπάρχουν διαθέσιμα δεδομένα μηνιαίας κατανάλωσης ενέργειας από τον Ιανουάριο του 2000, στο συγκεκριμένο σενάριο το εύρος περιορίζεται έτσι ώστε να ταυτίζεται με το εύρος των διαθέσιμων δεδομένων θερμοκρασίας.

Καθώς ένα δίκτυο NARX απαιτεί για την πρόβλεψη της κύριας χρονοσειράς τις αντίστοιχες μελλοντικές τιμές της βοηθητικής χρονοσειράς, θα πρέπει να γίνει πρόβλεψη της θερμοκρασίας. Για αυτό τον σκοπό κατασκευάστηκε ένα δίκτυο NAR, όπως στα Σενάρια 1 και 2 που στην προκειμένη περίπτωση θα προβλέπει τις μελλοντικές τιμές της θερμοκρασίας.

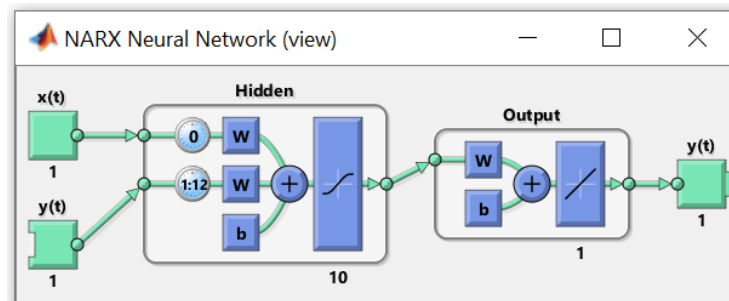
Για την πρόβλεψη της θερμοκρασίας χρησιμοποιήθηκαν τα διαθέσιμα δεδομένα θερμοκρασίας από τον Ιούλιο του 2009 έως και τον Ιανουάριο του 2019, δηλαδή μια χρονοσειρά 115 βημάτων. Η δομή του δικτύου NAR παρουσιάζεται παρακάτω:



Εικόνα 4-6: Δομή NAR νευρωνικού δικτύου σεναρίου 3

Όπως φαίνεται, οι υστερήσεις του δικτύου ορίστηκαν 12, καθώς η χρονοσειρά είναι μηνιαία και προβλέπεται να υπάρχει ετήσια περιοδικότητα. Μετά από δοκιμές, ο βέλτιστος αριθμός νευρώνων για την συγκεκριμένη χρονοσειρά ορίστηκε 3.

Στη συνέχεια σχεδιάστηκε το δίκτυο NARX με το οποίο θα γίνει η πρόβλεψη της κατανάλωσης ηλεκτρικής ενέργειας, με την βοήθεια της χρονοσειράς θερμοκρασιών. Η δομή του δικτύου είναι ως εξής, με 10 νευρώνες και υστέρηση 12:

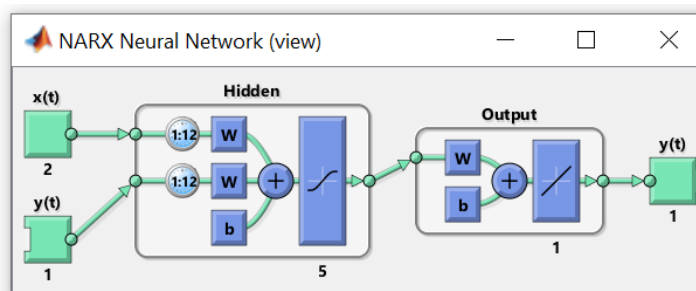


Εικόνα 4-7: Δομή NARX νευρωνικού δικτύου σεναρίου 3

Σενάριο 4

Σε αυτή την περίπτωση έγινε δοκιμή να χρησιμοποιούν ως είσοδοι σε μοντέλο NARX η κατανάλωση ηλεκτρικής ενέργειας και ως βοηθητική χρονοσειρά ο δείκτης του μήνα (π.χ. 01 για τον Ιανουάριο). Η βασική σκέψη πίσω από το σενάριο είναι ότι ο δείκτης του μήνα τροφοδοτεί το δίκτυο με μία επιπλέον πληροφορία η οποία είναι καθοριστική για την κατανάλωση της ηλεκτρικής ενέργειας. Έτσι επικυρώνεται η συσχέτιση μεταξύ ίδιων μηνών στο βάθος του χρόνου, με μια επιπλέον πληροφορία εκτός από την περιοδικότητα της χρονοσειράς που εισάγεται μέσω των υστερήσεων.

Το δίκτυο έχει την παρακάτω δομή, με υστέρηση 12 και 5 νευρώνες σε ένα κρυφό στρώμα:

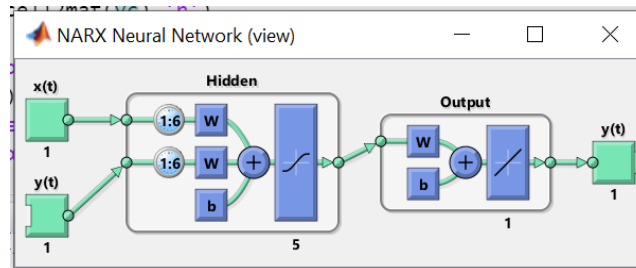


Εικόνα 4-8 Δομή NARX νευρωνικού δικτύου σεναρίου 4

Σενάριο 5

Στο Σενάριο 5 θα χρησιμοποιηθεί ως βοηθητική χρονοσειρά το ΑΕΠ της χώρας, έτσι ώστε να μελετηθεί εάν και κατά πόσο επηρεάζει ο δείκτης αυτός την πρόβλεψη του φορτίου, στα πλαίσια της μελέτης σε τεχνητό νευρωνικό δίκτυο. Η πρόβλεψη γίνεται σε δύο στάδια, όπως και στα προηγούμενα μοντέλα NARX. Αρχικά πραγματοποιείται πρόβλεψη του ΑΕΠ, με την εκπαίδευση νευρωνικού δικτύου τύπου NAR, και στην συνέχεια η πρόβλεψη χρησιμοποιείται για την πρόβλεψη της κατανάλωσης σε ηλεκτρική ενέργεια.

Το μοντέλο NAR για την πρόβλεψη του ΑΕΠ, μετά από δοκιμές σχεδιάστηκε έτσι ώστε να έχει υστέρηση 6 βημάτων και 5 νευρώνες στο κρυφό στρώμα όπως φαίνεται στην εικόνα 4-9.



Εικόνα 4-9: Δομή NARX νευρωνικού δικτύου σεναρίου 5

Το μοντέλο NARX που εκτελεί την τελική πρόβλεψη χρησιμοποιώντας την πρόβλεψη ΑΕΠ του μοντέλου NAR, μετά από δοκιμές κρίθηκε καταλληλότερο να είναι όμοιο του NAR, δηλαδή με 6 υστερήσεις και 5 νευρώνες.

4.3 Λήψη & Επεξεργασία Δεδομένων

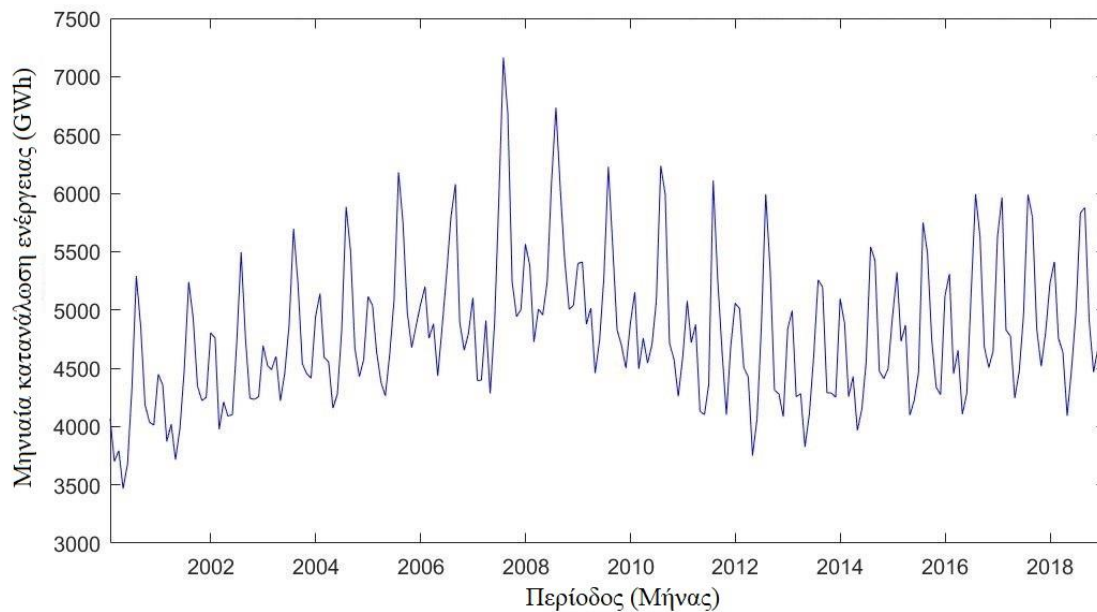
Τα δεδομένα τα οποία συνθέτουν τις χρονοσειρές για την εκπαίδευση των νευρωνικών δικτύων ποικίλλουν σε ποσότητα, εύρος καταγραφής ενώ αποτελούνται εκτός από τις ιστορικές καταγραφές κατανάλωσης ηλεκτρικής ενέργειας, και από τις βοηθητικές χρονοσειρές. Ανάμεσα σε αυτές είναι ο πληθυσμός της χώρας (ο οποίος εμπεριέχεται ως πληροφορία σε μία από τις δύο βασικές χρονοσειρές ηλεκτρικού φορτίου), η μέση θερμοκρασία της χώρας, καθώς και το Ακαθάριστο Εθνικό Προϊόν.

Η σύνθεση κάποιων από τις παρακάτω χρονοσειρές, όπως αυτή της μέση θερμοκρασίας της χώρας, έγινε στα πλαίσια της παρούσας μελέτης καθώς η πρωτογενής πληροφορία παρεχόταν σε αναφορές με μη επεξεργάσιμη μορφή.

Δεδομένα Κατανάλωσης Ενέργειας

Τα δεδομένα κατανάλωσης ηλεκτρικής ενέργειας που χρησιμοποιήθηκαν στην παρούσα μελέτη προέρχονται από δύο διαφορετικές πηγές.

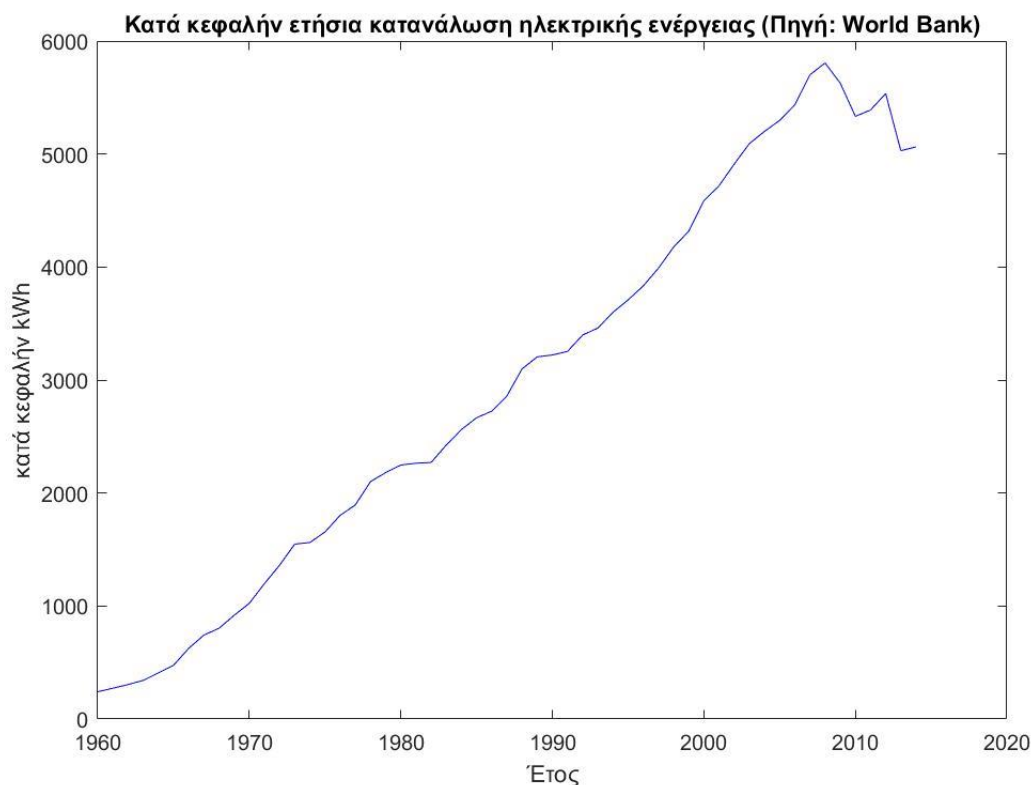
Η πρώτη χρονοσειρά είναι από την Eurostat και αποτελεί μηνιαία συνολική κατανάλωση ηλεκτρικής ενέργειας σε επίπεδο χώρας από τον Ιανουάριο του 2000 έως και τον Ιανουάριο του 2019. Η χρονοσειρά παρουσιάζεται στο παρακάτω γράφημα:



Γράφημα 4-1: Μηνιαία κατανάλωση ηλεκτρικής ενέργειας (Πηγή: International Energy Agency)

Τα χαρακτηριστικά της ετήσιας καμπύλης φορτίου αποδίδονται στο γράφημα, με τις αντίστοιχες αιχμές, την περιοδικότητα αλλά και την μακροχρόνια αυξομείωση των ενεργειακών απαιτήσεων. Σε όλα τα έτη παρατηρείται μια περιοδικότητα, όπου οι τιμές κατανάλωσης του Ιανουαρίου κυμαίνονται σε μέσα επίπεδα, ακολουθώντας πτωτική πορεία όσο προχωρούμε στην άνοιξη. Τον Απρίλιο σημειώνεται η χαμηλότερη καταγεγραμμένη τιμή ενέργειας του έτους, χαρακτηριστικό που επαναλαμβάνεται στα περισσότερα έτη της συγκεκριμένης χρονοσειράς. Τους επόμενους μήνες, με την αύξηση της θερμοκρασίας και την έναρξη του καλοκαιριού, η κατανάλωση αυξάνεται κατά πολύ, σημειώνοντας την αιχμή του έτος τον Ιούλιο και τον Αύγουστο. Από εκεί και στο εξής, μέχρι και τον Δεκέμβριο το φορτίο μειώνεται σταδιακά. Αναλόγως την χρονιά, παρουσιάζονται αυξομειώσεις και διαφοροποιήσεις την περίοδο που οριοθετείται από τον Αύγουστο μέχρι και τον Απρίλιο. Τα χαρακτηριστικά αυτής της περιόδου είναι ότι το φορτίο παραμένει σε μέσα επίπεδα, με τοπικό μέγιστο συνήθως τον Ιανουάριο και τον Δεκέμβριο και χαμηλότερες καταναλώσεις τον Φεβρουάριο. Η τάση της χρονοσειράς είναι αυξητική, από την έναρξη των καταγραφών έως και το έτος 2007. Στη συνέχεια παρουσιάζεται πτωτική πορεία, με μικρές αυξομειώσεις.

Η δεύτερη χρονοσειρά αποτελεί την κατά κεφαλήν ετήσια κατανάλωση ηλεκτρικής ενέργειας σε επίπεδο χώρας από το 1960 έως και το 2014. Τα δεδομένα παρέχονται από τον Οργανισμό World Bank. Η χρονοσειρά παρουσιάζεται παρακάτω:



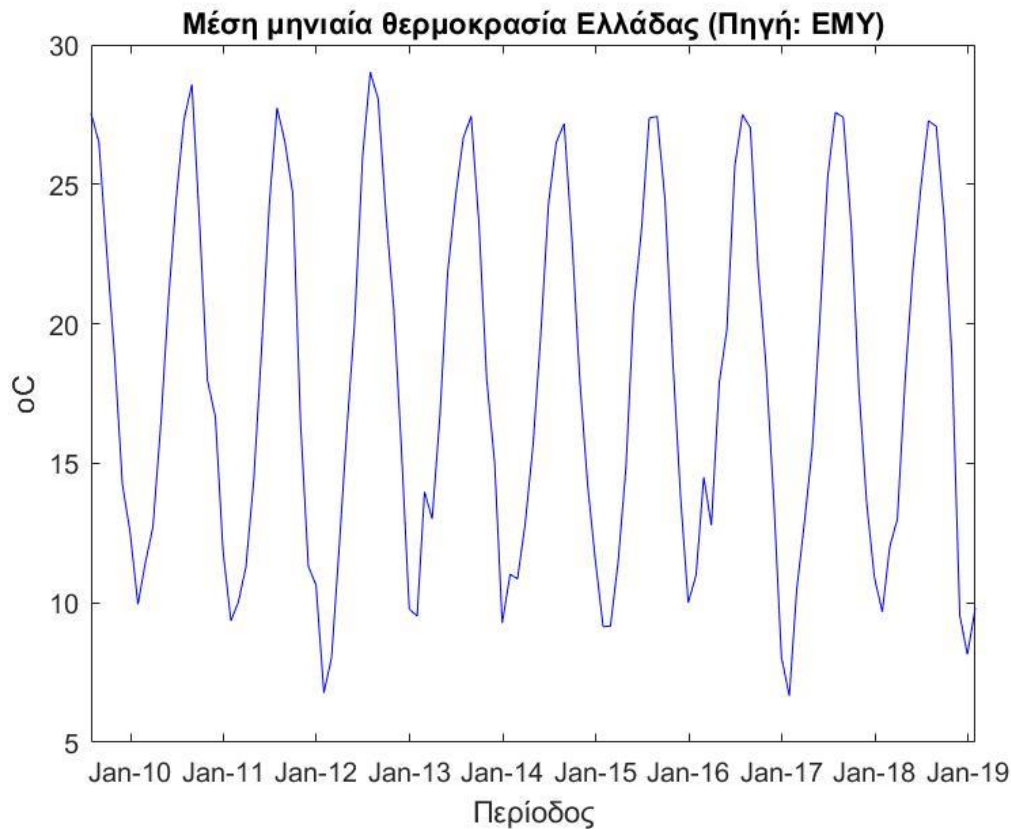
Γράφημα 4-2: Ετήσια κατά κεφαλήν κατανάλωση ηλεκτρικής ενέργειας (Πηγή: World Bank)

Ο μέσος ετήσιος ρυθμός αύξησης της συνολικής καθαρής ζήτησης κατά τη δεκαετία 2000- 2010 ήταν 2.17%, παρουσιάζοντας σημαντική μείωση σε σχέση με τις περασμένες δεκαετίες. Κατά την περίοδο 2000-2007 ο μέσος ετήσιος ρυθμός αύξησης της συνολικής καθαρής ζήτησης ηλεκτρικής ενέργειας ήταν 3.39%. Το 2008, απαρχή της οικονομικής κρίσης, η συνολική καθαρή ζήτηση στο Σύστημα ανήλθε στις 5805 kWh κατά κεφαλήν που αποτελεί ιστορικό μέγιστο, παρουσιάζοντας αύξηση 1.11% σε σχέση με το 2007. Το 2009 χαρακτηρίστηκε από σημαντική μείωση της συνολικής καθαρής ζήτησης στο Σύστημα, κατά 5.01% έναντι του 2008, η οποία οφείλεται στην αξιοσημείωτη μείωση των βιομηχανικών φορτίων κατά 20.19% σε σχέση με το 2008. Από το 2010 και μετά η συνολική καθαρή ζήτηση ηλεκτρικής ενέργειας στο ΕΣΜΗΕ υποχωρεί ελαφρώς κάθε χρόνο, με εξαίρεση το 2013 όπου παρατηρήθηκε μείωση κατά 3.7% έναντι του 2012. [4]

Κλιματικά Δεδομένα

Τα ιστορικά κλιματικά δεδομένα δεν υπήρχαν άμεσα διαθέσιμα ως τελικό προϊόν. Τα δεδομένα που υπήρξαν διαθέσιμα ήταν τα κλιματολογικά δελτία της ΕΜΥ τα οποία εκδίδονται κάθε μήνα σε αρχείο pdf. Το δελτίο περιέχει περιγραφή του κλίματος για τον μήνα για διάφορες περιοχές. Εκτός των άλλων περιέχεται και ένας πίνακας με την μέση και την μέγιστη τιμή της θερμοκρασίας, την βροχόπτωση και κάποια άλλα μετεωρολογικά δεδομένα για 55 σταθμούς της ΕΜΥ σε όλη την επικράτεια.

Για την χρήση των δεδομένων και τον υπολογισμό της μέσης μηνιαίας θερμοκρασίας της χώρας χρειάστηκε επεξεργασία. Ο τυπωμένος πίνακας σε pdf με τις μηνιαίες τιμές περάστηκε σε αρχείο Excel για τα 140 μετεωρολογικά δελτία από τον Ιούλιο του 2009 έως και τον Ιανουάριο του 2019. Στη συνέχεια, σε περιβάλλον Excel, πλέον, υπολογίστηκε η μέση τιμή της θερμοκρασίας στην χώρα, για κάθε μήνα. Η χρονοσειρά έχει ως εξής:



Γράφημα 4-3: Μέση μηνιαία θερμοκρασία Ελλάδας (Πηγή: ΕΜΥ)

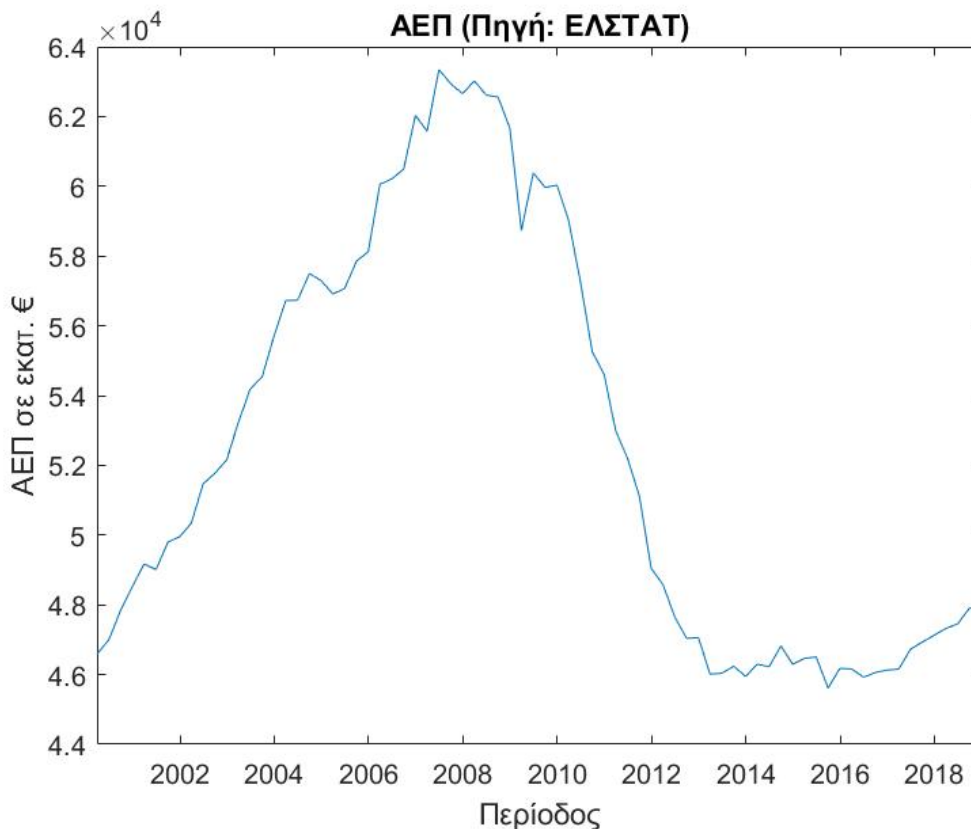
Δημογραφικά Δεδομένα

Τα δεδομένα για τον ελληνικό πληθυσμό προέρχονται από την απογραφή της ΕΛΣΤΑΤ το 2001 και το 2011. Η φύση των δεδομένων δεν αποτελεί χρονοσειρά καθώς δεν υπάρχει κάποιο μικρότερο σταθερό βήμα, πέραν της 10ετίας όπου γίνεται απογραφή. Μια, όχι και τόσο αξιόπιστη προσέγγιση, θα ήταν να γίνει γραμμική παρεμβολή ανάμεσα στις δύο γνωστές καταγραφές.

Παρόλα αυτά, η χρονοσειρά της ετήσιας κατά κεφαλήν κατανάλωσης ηλεκτρικής ενέργειας (γράφημα 4-2) από τον οργανισμό World Bank περιέχει, εμμέσως, πληροφορία για τον πληθυσμό. Το γεγονός ότι η κατανάλωση ενέργειας έχει υπολογιστεί κατά κεφαλήν παρέχει στο νευρωνικό δίκτυο μία επιπλέον πληροφορία, την οποία δεν είναι εφικτό να την παρέχουμε με διαφορετικό τρόπο στο δίκτυο λόγω έλλειψης δεδομένων. Επίσης, τα δεδομένα για τον πληθυσμό που χρησιμοποίησε ο οργανισμός World Bank για να υπολογίσει την κατά κεφαλήν κατανάλωση ηλεκτρικής ενέργειας σε ετήσια βάση, πιθανότατα να προέρχονται από επιστημονική έρευνα στην οποία έχει εφαρμοστεί εμπειριστατωμένο πληθυσμιακό μοντέλο. Έτσι, θεωρείται, ότι η πληθυσμιακή εκτίμηση είναι έγκυρη.

Οικονομικός Δείκτης ΑΕΠ

Το Ακαθάριστο Εθνικό Προϊόν παρέχεται από την ΕΛΣΤΑΤ και υπολογίζεται για κάθε τρίμηνο. Έτσι αποτελεί χρονοσειρά από το 1^ο τρίμηνο του 2000 έως και το 4^ο τρίμηνο του 2018. Τα δεδομένα έχουν ως εξής:



Γράφημα 4-4: Ακαθάριστο Εθνικό Προϊόν (Πηγή: ΕΛΣΤΑΤ)

Προετοιμασία δεδομένων

Για την εκπαίδευση του δικτύου προετοιμάζονται τα δεδομένα της χρονοσειράς. Η προετοιμασία περιλαμβάνει τον διαχωρισμό των δεδομένων σε τρεις κατηγορίες: εκείνα της εκπαίδευσης, του ελέγχου και της αξιολόγησης. Τα δεδομένα της εκπαίδευσης χρησιμοποιούνται κατά την υπολογιστική διαδικασία στην οποία προσδιορίζονται τα συναπτικά βάρη των νευρώνων, έτσι ώστε οι τιμές της εξόδου του δικτύου να είναι όσο το δυνατόν πιο κοντά στις τιμές της χρονοσειράς και να επιτευχθεί επαρκής εκπαίδευση του δικτύου.

Τα υποσύνολο των δεδομένων αξιολόγησης χρησιμοποιείται για την αξιολόγηση της εκπαίδευσης κατά τη διάρκειά της. Αυτό επιτυγχάνεται μέσω της παρακολούθησης του σφάλματος που προκύπτει μεταξύ των δεδομένων αξιολόγησης και των αποτελεσμάτων που εξάγονται από το σύστημα. Το σφάλμα παρακολουθείται σε κάθε στάδιο της διαδικασίας και είναι ιδιαίτερα χρήσιμο, καθώς βοηθά στην αποφυγή της υπερπροσαρμογής. Συνήθως το σφάλμα των δεδομένων αξιολόγησης μικραίνει καθώς προχωράνε οι κύκλοι της εκπαίδευσης (epochs), όπως ακριβώς και το σφάλμα των δεδομένων εκπαίδευσης. Όταν όμως το δίκτυο αρχίζει να υπερπροσαρμόζεται στα δεδομένα, το σφάλμα των δεδομένων αξιολόγησης ξεκινά να αυξάνει. Έτσι δίνεται εντολή, μέσω του αλγορίθμου εκπαίδευσης, να σταματήσει η διαδικασία και να αποφευχθεί η υπερπροσαρμογή του νευρωνικού δικτύου. Η εκπαίδευση ουσιαστικά σταματά όταν το σφάλμα της αξιολόγησης σταθεροποιηθεί, και οι ιδιότητες του δικτύου (πολώσεις και βάρη) που αποθηκεύονται ως τελικές είναι εκείνες που δίνουν την χαμηλότερη τιμή σφάλματος εκπαίδευσης. [20]

Τέλος, το τρίτο υποσύνολο αποτελεί το υποσύνολο ελέγχου (test). Αυτό δεν λαμβάνει μέρος στην διαδικασία της εκπαίδευσης, είναι ανεξάρτητο από τα υπόλοιπα δεδομένα, έχει όμως παρόμοια κατανομή πιθανότητας. Αποτελεί ακόμη μια δικλείδα ασφαλείας που εξασφαλίζει την σωστή εκπαίδευση και την αποφυγή υπερπροσαρμογής, καθώς εάν ένα δίκτυο ανταποκρίνεται ικανοποιητικά τόσο στα δεδομένα της εκπαίδευσης όσο και στα δεδομένα ελέγχου, τότε η διαδικασία

της εκπαίδευσης έχει ολοκληρωθεί με επιτυχία. Σε περίπτωση υπερμοντελοποίησης (overfitting) ένα δίκτυο θα είχε πολύ καλή απόδοση στα δεδομένα εκπαίδευσης, ενώ θα παρουσίαζε μεγάλη απόκλιση στα δεδομένα ελέγχου. [22]

Κανονικοποίηση δεδομένων

Η κανονικοποίηση των δεδομένων είναι απαραίτητη πριν την είσοδό τους στην υπολογιστική διαδικασία της εκπαίδευσης του νευρωνικού δικτύου. Η διαδικασία δύναται να γίνει είτε χειροκίνητα, είτε αυτόματα. Αυτό εξαρτάται από το προγραμματιστικό περιβάλλον στο οποίο υλοποιείται το δίκτυο ή από την φύση των δεδομένων και τις ιδιαιτερότητες του προβλήματος.

Σε περιπτώσεις που γίνεται κανονικοποίηση σε δεδομένα για τα οποία πρόκειται να γίνει μελλοντική πρόβλεψη θα πρέπει να ληφθεί υπόψιν εκτός από το εύρος των υπαρχόντων τιμών και το αναμενόμενο εύρος των προβλεπόμενων. Σε αντίθετη περίπτωση τα αποτελέσματα της πρόβλεψης θα βγουν εκτός του ορίου της κανονικοποίησης με αποτέλεσμα την λανθασμένη ή μη πρόβλεψη.

Στα πλαίσια της εργασίας η κανονικοποίηση γίνεται αυτόματα, καθώς υπάρχει μέσα στον κώδικα της Matlab για την κατασκευή νευρωνικού δικτύου πρόβλεψης τύπου NAR και NARX.

Αποτελέσματα

- Σε αυτό το κεφάλαιο παρουσιάζονται και αναλύονται τα αποτελέσματα που προέκυψαν για τα 5 διαφορετικά σενάρια κατά την ολοκλήρωση της εκπαίδευσης. Η εκμάθησή του τεχνητού νευρωνικού δικτύου ολοκληρώνεται με την εκτύπωση γραφημάτων και τον υπολογισμό μεταβλητών, όπως το μέσο τετραγωνικό σφάλμα και ο συντελεστής προσδιορισμού (ή συσχέτισης- R^2). Οι εξαγόμενες
5. μεταβλητές καθώς και η μορφή των γραφημάτων είναι απαραίτητες για την αξιολόγηση της διαδικασίας εκπαίδευσης και της επιτυχούς γενίκευσης του νευρωνικού δικτύου στα δεδομένα, για αυτό παρουσιάζονται και αναλύονται παρακάτω. Στην περίπτωση που τα γραφήματα και οι τιμές των μεταβλητών είναι ικανοποιητικά, τότε το δίκτυο θεωρείται πως έχει την δυνατότητα να παράγει ασφαλή πρόβλεψη για τα επόμενα βήματα της μελετώμενης χρονοσειράς.

Τα αποτελέσματα που παρουσιάζονται στις παρακάτω παραγράφους είναι τα βέλτιστα που έχουν προκύψει από αρκετές δοκιμές, διότι κάθε εκπαίδευση του δικτύου εμπεριέχει έναν παράγοντα τυχαιότητας και δίνει διαφορετικά αποτελέσματα. Ο λόγος που κάθε κύκλος εκπαίδευσης διαφοροποιείται από τον προηγούμενο απαντάται στο γεγονός ότι η εκπαίδευση αποτελεί έναν αλγόριθμο επαναληπτικής διαδικασίας, της οποίας οι τελικές παράμετροι εξαρτώνται από τις αρχικές τιμές συναπτικών βαρών, οι οποίες συνήθως δίνονται τυχαία από το σύστημα. Έτσι, στα πλαίσια της μελέτης, ο κώδικας του κάθε σεναρίου έχει δοκιμαστεί αρκετές φορές, με κάθε μία να δίνει διαφορετικής ποιότητας και γενίκευσης νευρωνικό δίκτυο. Στην εργασία παρουσιάζονται τα αποτελέσματα των δικτύων τα οποία θεωρούνται τα πιο αξιόπιστα, για λόγους που αναλύονται στις παρακάτω παραγράφους, ενώ δεν κρίθηκε σκόπιμο να παρουσιαστούν τα αποτελέσματα τα οποία έχουν απορριφθεί, διότι δεν έχουν κάποια ερευνητική αξία, λόγω του τυχαίου χαρακτήρα τους.

Για τα 5 διαφορετικά σενάρια έχουν υλοποιηθεί 7 νευρωνικά δίκτυα, καθώς τα 2 σενάρια τύπου NARX περιέχουν από ένα επιπλέον νευρωνικό δίκτυο τύπου NAR για την πρόβλεψη της βοηθητικής χρονοσειράς, όπως περιγράφεται αναλυτικότερα στα παραπάνω κεφάλαια. Οι τύποι των εξαγόμενων γραφημάτων που παρουσιάζονται στις παρακάτω παραγράφους ως αποτελέσματα των 7 νευρωνικών δικτύων είναι:

- Μέσο τετραγωνικό σφάλμα
- Γραμμική παλινδρόμηση μεταξύ πραγματικών και εκτιμώμενων τιμών
- Γράφημα διασποράς πραγματικών & εκτιμώμενων τιμών
- Αυτοσυσχέτιση των σφαλμάτων εκπαίδευσης
- Κατανομή σφαλμάτων
- Πρόβλεψη

Στις παραγράφους 5.1 έως 6 παρουσιάζονται τα εν λόγω γραφήματα και σχολιάζεται η μορφή και η κατανομή τους με γνώμονα την ποιότητα εκπαίδευσης του δικτύου, για το κάθε σενάριο. Μεταξύ των σεναρίων, τα ίδιου τύπου γραφήματα παρουσιάζουν ομοιότητα, καθώς έχουν επιλεγεί σκόπιμα έτσι ώστε τα νευρωνικά δίκτυα να έχουν καλή προσαρμογή αλλά και γενίκευση στα δεδομένα. Έτσι, τα εκπαιδευμένα τεχνητά νευρωνικά δίκτυα του κάθε σεναρίου έχουν μεγάλη πιθανότητα να παράγουν αξιόπιστη και ασφαλή πρόβλεψη για τις μελλοντικές τιμές της χρονοσειράς επάνω στην οποία έχουν εκπαιδευτεί.

Η παραγόμενη πρόβλεψη, επίσης, παρουσιάζεται και αναλύεται στις παραγράφους 5.1 έως 6. Σχολιάζονται τα χαρακτηριστικά της χρονοσειράς, όπως η περιοδικότητα, οι τάσεις και η στοχαστικότητα και εξετάζεται εάν τα παραπάνω χαρακτηριστικά αντικατοπτρίζουν μία πραγματική χρονοσειρά καταγραφών κατανάλωσης ηλεκτρικής ενέργειας.

Σενάριο 1

Το γράφημα μέσου τετραγωνικού σφάλματος παραθέτει την τιμή της ομόνυμης μεταβλητής συναρτήσεως των κύκλων εκπαίδευσης που πραγματοποιούνται έως ότου τα κριτήρια τερματισμού του αλγόριθμου της διαδικασίας εκπληρωθούν. Το γράφημα περιέχει τις γραμμές 3 διαφορετικών τετραγωνικών σφαλμάτων, μία για κάθε κατηγορία δεδομένων. Η χρησιμότητα και ο ρόλος του κάθε 5.1 πακέτου δεδομένων από τα παραπάνω περιγράφονται στην παράγραφο 4.3.

Συνοπτικά, οι τρεις κατηγορίες είναι:

1. Δεδομένα εκπαίδευσης
2. Δεδομένα ελέγχου (test)
3. Δεδομένα επαλήθευσης

Η μορφή του γραφήματος που προκύπτει, προϋποθέτει τον αναλυτή για την ποιότητα της εκπαίδευσης του νευρωνικού δικτύου. Βασικό ρόλο παίζουν η συσχέτιση, η εξέλιξη και η διαφοροποίηση ή μη των σφαλμάτων των 3 παραπάνω πακέτων δεδομένων καθώς προχωρούν οι κύκλοι της εκπαίδευσης.

Βασικό χαρακτηριστικό του γραφήματος είναι η επισήμανση του καλύτερου κύκλου εκπαίδευσης. Όπως φαίνεται και στο γράφημα 5.1, ενώ έχουν πραγματοποιηθεί 9 κύκλοι, το σύστημα όρισε ως βέλτιστο τον 3^ο κύκλο εκπαίδευσης. Πρακτικά, αυτό σημαίνει πως οι τελικές τιμές για τα συναπτικά βάρη και οι λοιπές παράμετροι του δικτύου κατά την λήξη της διαδικασίας εκμάθησης θα είναι αυτές του 3^{ου} κύκλου.

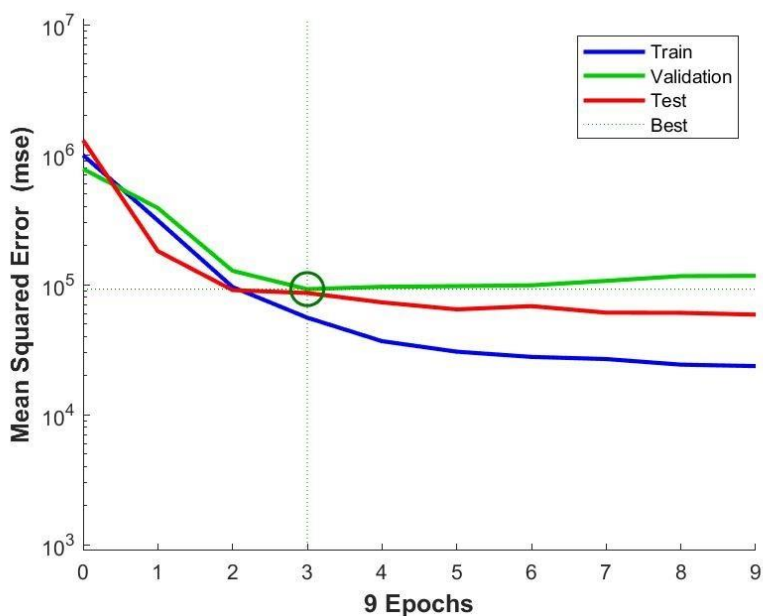
Αρχικά, η σταθεροποίηση των σφαλμάτων, όσο προχωρούν οι κύκλοι της εκπαίδευσης δηλώνει ότι η επαναληπτική διαδικασία του αλγορίθμου συγκλίνει γύρω από συγκεκριμένες παραμέτρους του δικτύου και το σφάλμα έχει σταθεροποιηθεί ή φθίνει ελάχιστα. Το γεγονός αυτό είναι ένα αρχικό δείγμα ότι η εκπαίδευση πρέπει να σταματήσει. Όσο αναφορά το σφάλμα των δεδομένων επαλήθευσης, θα πρέπει οι τιμές του να ακολουθούν το σφάλμα των τιμών της εκπαίδευσης. Συνήθως όμως, αναλόγως την φύση και τα χαρακτηριστικά της χρονοσειράς με την οποία εκπαιδεύεται το δίκτυο, το σφάλμα των δεδομένων επαλήθευσης αρχίζει και αυξάνει, ενώ το σφάλμα των δεδομένων εκπαίδευσης μειώνεται, διότι γίνεται υπερπροσαρμογή του δικτύου.

Η υπερπροσαρμογή είναι ένα ανεπιθύμητο φαινόμενο, καθώς σημαίνει ότι το δίκτυο έχει προσαρμοστεί έτσι ώστε να αντιγράφει πιστά την χρονοσειρά εισόδου. Το αποτέλεσμα είναι ότι δεν θα μπορέσει να παράγει σωστές προβλέψεις για τα επόμενα βήματα της χρονοσειράς, απλώς θα αναπαράγει τα προηγούμενα βήματα. Εκεί βρίσκεται το νόημα της παρακολούθησης των σφαλμάτων εκπαίδευσης και επαλήθευσης κατά την διάρκεια της διαδικασίας. Το γράφημα 5-1 απεικονίζει την πορεία των σφαλμάτων κατά την διαδικασία της εκπαίδευσης και με βάση τις παραπάνω επεξηγήσεις κρίνεται εάν το δίκτυο έχει εκπαιδευτεί καταλλήλως.

Για το σενάριο 1 η εκπαίδευση σταμάτησε μετά από 9 κύκλους αφού το σφάλμα της εκπαίδευσης, των δεδομένων επαλήθευσης και των δεδομένων ελέγχου άρχισαν να σταθεροποιούνται. Το δίκτυο όμως δεν προσδιορίστηκε με τις παραμέτρους του 9^{ου} κύκλου, καθώς από εκείνο το σημείο και έπειτα τα σφάλμα της επαλήθευσης άρχισε να αυξάνει. Η καλύτερη απόδοση ορίστηκε από τον αλγόριθμο στον 3^ο κύκλο, όπου, όπως παρουσιάζεται στο γράφημα 5-1 το σφάλμα των δεδομένων αξιολόγησης παρουσιάζει ολικό ελάχιστο, ενώ το σφάλμα εκπαίδευσης έχει μειωθεί αισθητά σε σχέση με τους πρώτους κύκλους εκπαίδευσης. Με αυτό τον τρόπο αποφεύγεται η υπερπροσαρμογή του δικτύου.

Αξίζει να σημειωθεί ότι η μέση τετραγωνική απόκλιση, οι τιμές της οποίας χρησιμοποιούνται για την αξιολόγηση της απόδοσης του δικτύου παρουσιάζει μεγάλες τιμές. Σε γενικές γραμμές, αυτό που προσπαθούμε να επιτύχουμε κατά την εκπαίδευση του δικτύου είναι να μειωθεί όσο το δυνατόν περισσότερο το σφάλμα. Έτσι, θα μπορούσε κανείς να υποθέσει πως το σφάλμα θα πρέπει να

κυμαίνεται σε οριακά μηδενικές τιμές. Η θεώρηση αυτή δεν είναι λάθος, απλά είναι ιδανική θεώρηση που συνήθως δεν συμβαίνει σε πραγματικά δεδομένα. Στην πραγματικότητα μικρό σφάλμα είναι εκείνο του οποίου η απόκλιση, για την φύση και την στοχαστικότητα των δεδομένων που μελετώνται, βρίσκεται μέσα στα όρια εμπιστοσύνης ή δεν επηρεάζουν αισθητά το εξαγόμενο προϊόν. Το γεγονός ότι το σφάλμα στην περίπτωση μας είναι της τάξεως του 10^6 δεν σημαίνει πως το δίκτυο δεν έχει καλή απόδοση, καθώς η μέση τετραγωνική απόκλιση ορίζεται με τετραγωνικές μονάδες της μελετώμενης χρονοσειράς, η οποία στην περίπτωση μας είναι της τάξεως των GWh. Τέλος, η απόδοση του δικτύου δεν κρίνεται μόνο με βάση το μέσο τετραγωνικό σφάλμα, αλλά και από πληθώρα άλλων παραγόντων που περιγράφονται στην συνέχεια.



Γράφημα 5-1: Σενάριο 1, Μέσο τετραγωνικό σφάλμα κατά την εκπαίδευση

Παρακάτω, στο γράφημα 5-2 απεικονίζεται η παλινδρόμηση που προκύπτει ανάμεσα στα πραγματικά δεδομένα που εισέρχονται στο νευρωνικό δίκτυο κατά την εκπαίδευση και τα αντίστοιχα εκτιμώμενα που προκύπτουν από το ήδη εκπαιδευμένο δίκτυο μετά την ολοκλήρωση της διαδικασίας. Τα εκτιμώμενα δεδομένα δεν είναι η χρονοσειρά πρόβλεψης, αλλά το εξαγόμενο προϊόν του νευρωνικού δικτύου σε λειτουργία ανοιχτού βρόχου, όπως περιγράφεται αναλυτικά στην παράγραφο 4.1. Τα βήματα της εκτιμώμενης χρονοσειράς αντιστοιχούν στα βήματα της πραγματικής καταγεγραμμένης χρονοσειράς, και με βάση την απόκλιση αυτών των δύο εκτιμώνται τα σφάλματα του δικτύου. Μια επιτυχημένη εκπαίδευση θα πρέπει να παραμετροποιεί το δίκτυο με τέτοιο τρόπο έτσι ώστε τα εξαγόμενα δεδομένα να συγκλίνουν όσο το δυνατόν περισσότερο με τα πραγματικά, υπό κάποιες προϋποθέσεις οι οποίες αναλύονται παραπάνω στην παρούσα παράγραφο, έτσι ώστε να μην έχει προκύψει υπερπροσαρμογή.

Το γράφημα παλινδρόμησης που εξάγεται από το προγραμματιστικό εργαλείο της εκπαίδευσης συμπεριλαμβάνει 4 διαφορετικά διαγράμματα. Το κάθε ένα από αυτά απεικονίζει την παλινδρόμηση που προκύπτει από τα 3 διαφορετικά σύνολα δεδομένων που χρησιμοποιούνται για την εκπαίδευση και την αξιολόγησή της. Συνοπτικά, έχουν ως εξής:

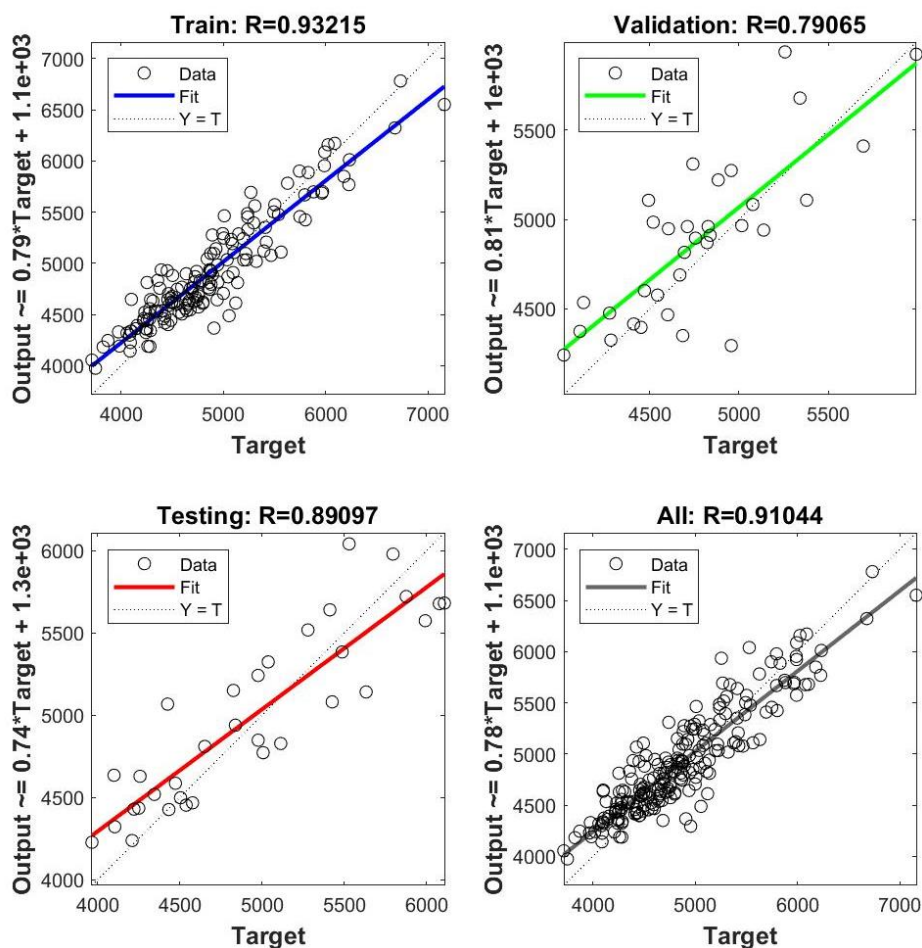
- Παλινδρόμηση συνόλου δεδομένων εκπαίδευσης
- Παλινδρόμηση συνόλου δεδομένων ελέγχου
- Παλινδρόμηση συνόλου δεδομένων επαλήθευσης

Οι 4 διαφορετικές παλινδρομήσεις θα πρέπει να συγκλίνουν όσο το δυνατόν περισσότερο ως προς τον συντελεστή προσδιορισμού R^2 , έτσι ώστε η συνολική απόδοση του νευρωνικού δικτύου να είναι ικανοποιητική. Οι διαφορές μεταξύ των συντελεστών προσδιορισμού για το κάθε υποσύνολο δεδομένων είναι αναμενόμενες και θεμιτές, καθώς αντικατοπτρίζουν καλή γενίκευση του δικτύου. Σε περίπτωση που υπάρχουν μεγάλες αποκλίσεις, τότε είναι βέβαιο πως το δίκτυο δεν έχει εκπαιδευτεί σωστά.

- Εάν το υποσύνολο εκπαίδευσης παρουσιάζει πολύ μικρή συσχέτιση με τα πραγματικά δεδομένα, τότε το νευρωνικό δίκτυο δεν έχει προσαρμοστεί καθόλου ικανοποιητικά στα δεδομένα της χρονοσειράς. Αντίστοιχα, και οι συντελεστές προσδιορισμού του υποσυνόλου ελέγχου και επαλήθευσης θα κυμαίνονται σε χαμηλές τιμές.
- Σε περίπτωση που το υποσύνολο εκπαίδευσης έχει καλό συντελεστή προσδιορισμού σε σχέση με τα πραγματικά δεδομένα, ενώ για το υποσύνολο ελέγχου ή επαλήθευσης δεν ισχύει το ίδιο, τότε η εκπαίδευση δεν έχει πραγματοποιήσει σωστή γενίκευση για την χρονοσειρά. Πιθανότατα το δίκτυο να έχει υπερπροσαρμοστεί στα δεδομένα της εκπαίδευσης και να μην είναι σε θέση να προβλέψει σωστά τιμές οι οποίες δεν βρίσκονται στο σύνολο αυτό.

Συνήθως, οι συντελεστές προσδιορισμού για τα 3 σύνολα δεδομένων έχουν μικρή απόκλιση μεταξύ τους, οπότε στο σύνθετο γράφημα των παλινδρομήσεων συμπεριλαμβάνεται ένα συγκεντρωτικό διάγραμμα στο οποίο παρουσιάζεται η συνολική παλινδρόμηση της χρονοσειράς. Με γνώμονα το συγκεντρωτικό διάγραμμα και τον συνολικό συντελεστή προσδιορισμού κρίνεται η απόδοση της εκπαίδευσης. Οι μεμονωμένες παλινδρομήσεις για το κάθε υποσύνολο απεικονίζονται, ωστόσο, ως δικλείδα ασφαλείας μαζί με το γράφημα σφαλμάτων (5-1) σε περίπτωση που ο αλγόριθμος της εκπαίδευσης δεν ολοκληρωθεί με επιτυχία και δεν εξάγει τα επιθυμητά αποτελέσματα.

Για το σενάριο 1, από το γράφημα 5-2 συμπεραίνουμε πως η εκπαίδευση είναι ικανοποιητική, καθώς ο συντελεστής προσδιορισμού για όλες τις παλινδρομήσεις συγκεντρωτικά υπολογίζεται σε υψηλά επίπεδα ($R^2=0,91$). Ωστόσο, οι συντελεστές των δεδομένων αξιολόγησης και ελέγχου είναι ελαφρώς ασθενέστεροι, χωρίς όμως να υποδεικνύουν σοβαρό πρόβλημα στην γενίκευση του νευρωνικού δικτύου.



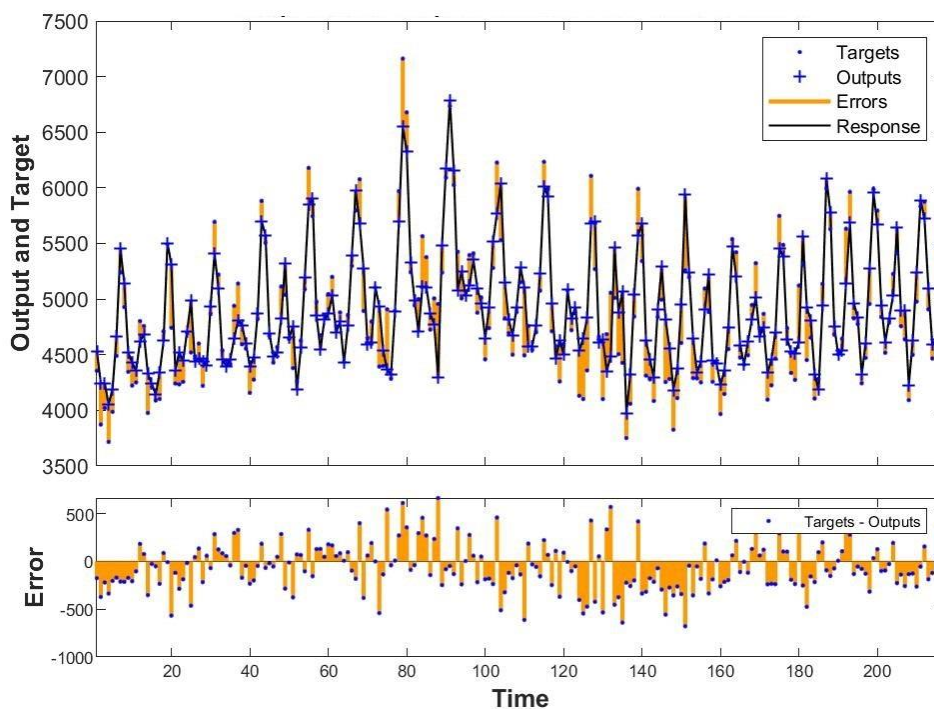
Γράφημα 5-2: Σενάριο 1, Παλινδρόμηση πραγματικών – εκτιμώμενων τιμών

Επίσης πολύ σημαντικό για την αξιολόγηση του αποτελέσματος της εκπαίδευσης είναι και η οπτική απεικόνιση της διασποράς της πραγματικής χρονοσειράς με την εκτιμώμενη (Γράφημα 5-3). Ως επί το πλείστον, οι πραγματικές χρονοσειρές εμπεριέχουν θόρυβο, λανθασμένες τιμές, καταγραφές τυχαίων συμβάντων και άλλες αλλοιώσεις στα δεδομένα. Το νευρωνικό δίκτυο θα πρέπει να είναι σε θέση να αναγνωρίσει τα πραγματικά χαρακτηριστικά της χρονοσειράς και να μην λαμβάνει υπόψιν τις τυχαίες αποκλίσεις. Το γράφημα διασποράς βοηθάει τον μελετητή να παρατηρήσει την σχέση των δύο αυτών χρονοσειρών και συμπεράνει εάν το νευρωνικό δίκτυο έχει καταφέρει να αφομοιώσει χαρακτηριστικά-κλειδιά της πραγματικής χρονοσειράς, τα οποία καθορίζουν την σωστή γενίκευση του νευρωνικού δικτύου.

Τέτοια χαρακτηριστικά είναι οι τάσεις, η εποχικότητα, η στασιμότητα κ.α. Έχοντας γνώση για τα χαρακτηριστικά της μελετώμενης χρονοσειράς, με την επισκόπηση του γραφήματος διασποράς γίνεται αντιληπτό αν η εκτιμώμενη χρονοσειρά ακολουθεί σωστά αυτές τις τάσεις, ενώ δεν επηρεάζεται από τον θόρυβο και τα τυχαία συμβάντα που υπάρχουν στην καταγεγραμμένα δεδομένα. Π.χ. στο τρέχον σενάριο βασικά χαρακτηριστικά της μηνιαίας κατανάλωσης χρονοσειράς κατανάλωσης ηλεκτρικής ενέργειας είναι η 12μηνη εποχικότητα, τα μέγιστα και τα ελάχιστα κάθε χρόνου, η αυξητική τάση κ.α. που περιγράφονται αναλυτικά στην παράγραφο 4.3.

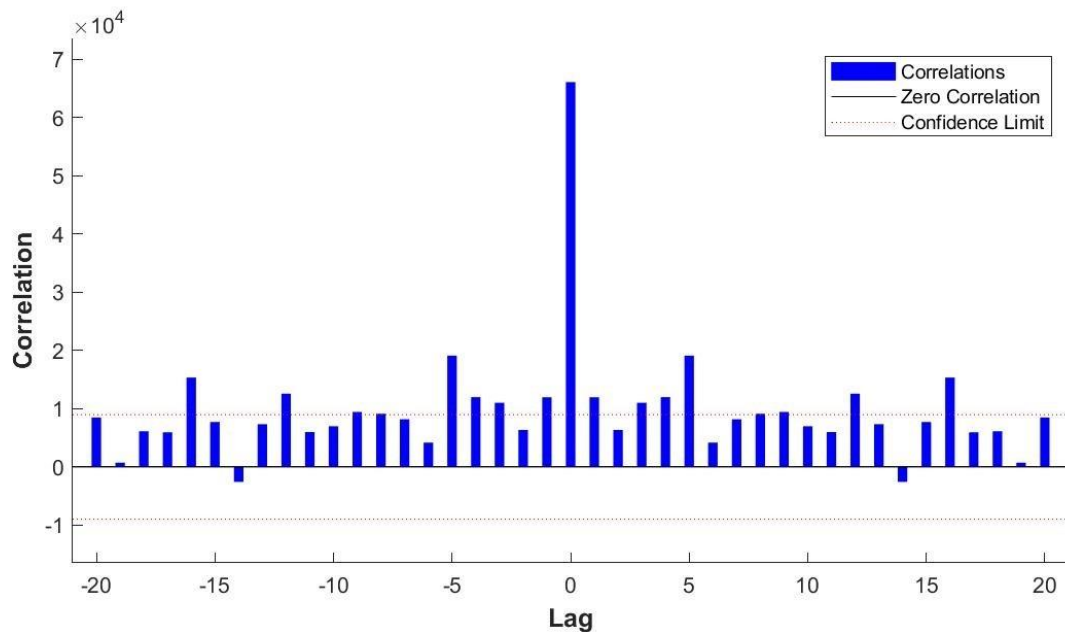
Συνοπτικά, το δίκτυο φαίνεται να έχει προσαρμοστεί αρκετά καλά στα πραγματικά δεδομένα, καθώς παρατηρείται (Γράφημα 5-3) πώς η ετήσια περιοδικότητα ακολουθείται πλήρως στην εξαγόμενη από το δίκτυο χρονοσειρά. Τα περισσότερα και μεγαλύτερα σφάλματα είναι συγκεντρωμένα την περίοδο 2008 με 2012, όπου η τάση της χρονοσειράς αλλάζει απότομα και από αύξουσα γίνεται φθίνουσα καθώς και την περίοδο 2012-2014 όπου η τάση γίνεται αύξουσα πάλι. Συγκεκριμένα, φαίνεται πως

το δίκτυο προσπαθεί να μην δημιουργεί απότομες αλλαγές τάσεις στην χρονοσειρά, προσπαθώντας να κρατήσει ομαλή αύξουσα τάση, όπως ακριβώς συμβαίνει στο μεγαλύτερο και αρχικό κομμάτι (2000-2008). Η συμπεριφορά αυτή είναι θεμιτή, καθώς αποδεικνύει πως το δίκτυο αντιλαμβάνεται τα σημαντικά, μακροχρόνια χαρακτηριστικά της χρονοσειράς (όπως π.χ. η αυξητική τάση, εποχικότητα) και επηρεάζεται όσο το δυνατόν λιγότερο από την στοχαστικότητα της (π.χ. απότομη πτώση κατανάλωσης λόγω οικονομικής κρίσης, μεμονωμένη μη συνηθισμένη αύξηση ή πτώση κατανάλωσης κάποιου μήνα).



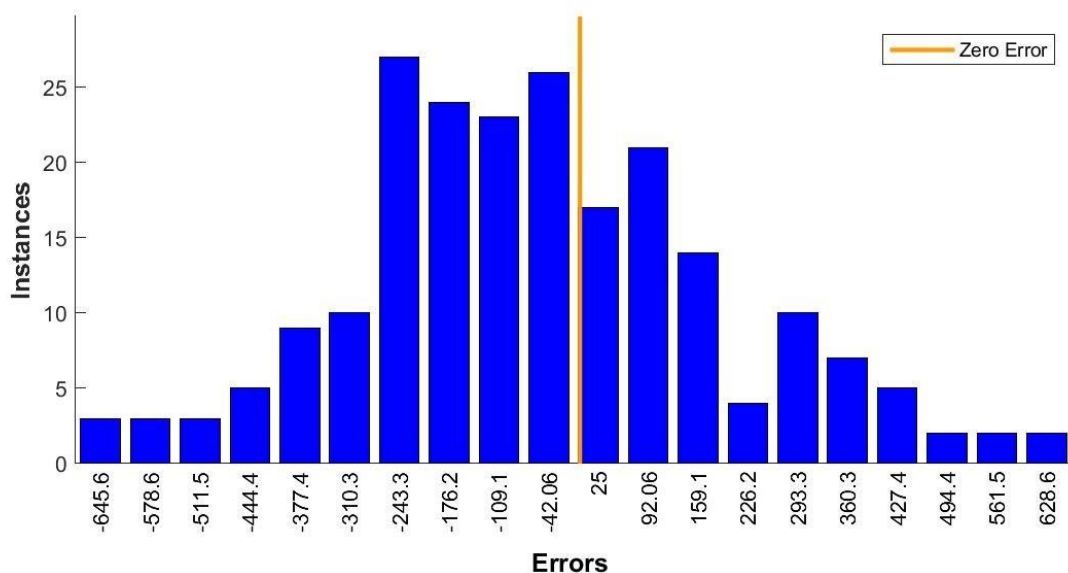
Γράφημα 5-3: Σενάριο 1, Διασπορά πραγματικών- εκτιμώμενων τιμών

Ακόμη ένας παράγοντας που παίζει σημαντικό ρόλο στα αποτελέσματα του δικτύου είναι η αυτοσυσχέτιση των σφαλμάτων. Τα σφάλματα που προκύπτουν θα πρέπει να είναι τυχαία και να μην επαναλαμβάνονται περιοδικά σε επόμενα βήματα της χρονοσειράς. Αν αυτό δεν συμβαίνει, τότε εξάγεται το συμπέρασμα πως το δίκτυο δεν έχει προσαρμοστεί σωστά στην περιοδικότητα της χρονοσειράς, ή αναπαράγει μοτίβα τα οποία ενδέχεται να έχουν προκληθεί από θόρυβο ή από επανειλημμένες τυχαίες αποκλίσεις. Το βέλτιστο είναι τα σφάλματα να παρουσιάζουν μηδενική ή ελάχιστη αυτοσυσχέτιση. Στην πραγματικότητα πάντα θα υπάρχουν κάποιες αυτοσυσχετίσεις σφαλμάτων, αλλά είναι θεμιτό να βρίσκονται μέσα στο διάστημα εμπιστοσύνης 5% (Γράφημα 5-4). Για την περίπτωση του σεναρίου 1, παρατηρείται πως υπάρχει πληθώρα επαναλαμβανόμενων σφαλμάτων, ενώ αρκετά από αυτά ξεπερνούν το διάστημα εμπιστοσύνης. Αυτό μπορεί να συμβαίνει διότι το νευρωνικό δίκτυο έχει αντιληφθεί μοτίβα τα οποία συνδέονται με την εποχικότητα της κατανάλωσης και επαναλαμβάνονται σε μεγάλο κομμάτι των καταγραφών. Το δίκτυο τείνει να επαναλαμβάνει τα μοτίβα αυτά σε όλη το χρονικό εύρος της χρονοσειράς, ενώ στην πραγματικότητα τα χαρακτηριστικά τους (π.χ. αιχμές, διάρκεια, περιοδικότητα) μεταβάλλονται λόγω απρόβλεπτων παραγόντων, όπως η έναρξη της οικονομικής κρίσης του 2008. Οι σημαντικότερες επαναλήψεις σφαλμάτων παρατηρούνται με περιοδικότητα 5, 12 και 16, δηλαδή επαναλαμβάνονται σε περιόδους που είτε αλλάζουν οι κλιματικές συνθήκες (5μηνο- 6 μήνο) και το δίκτυο δεν προσαρμόζεται άμεσα στην αλλαγή ή αλλάζει το έτος, οπότε το δίκτυο επαναλαμβάνει τα μοτίβα του προηγούμενου ενώ μπορεί αυτό να μην συμβαίνει στα καταγεγραμμένα δεδομένα.



Γράφημα 5-4: Σενάριο 1, Αυτοσυσχέτιση σφαλμάτων

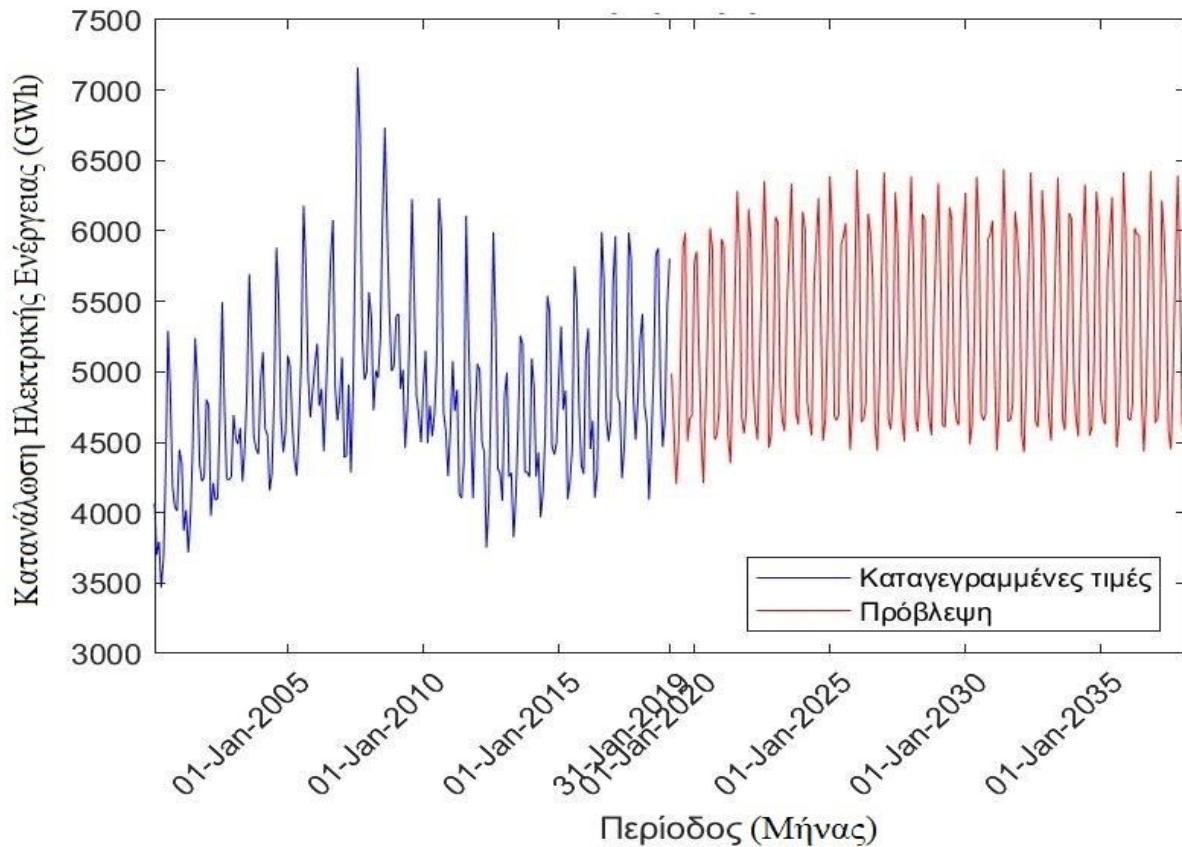
Στη συνέχεια ακολουθεί το ιστόγραμμα κατανομής σφαλμάτων, δηλαδή το πόσες φορές έχουν εμφανιστεί ίδια σφάλματα στο εξαγόμενο προϊόν του νευρωνικού δικτύου. Το ιστόγραμμα των σφαλμάτων είναι θεμιτό να παρουσιάζει κανονική κατανομή, με μικρή τυπική απόκλιση και μέση τιμή 0. Το μεγαλύτερο πλήθος των σφαλμάτων, με άλλα λόγια, θα πρέπει να κυμαίνεται σε χαμηλά επίπεδα τα οποία θα πρέπει να προσεγγίζουν το 0. Σε περίπτωση που έχουμε μεγάλη συχνότητα σφαλμάτων με υψηλές τιμές, τότε η απόκριση του δικτύου χάνει σε αξιοπιστία και αυξάνεται η πιθανότητα να μην προχωρήσει σε ορθή πρόβλεψη. Στο τρέχον σενάριο η κατανομή των σφαλμάτων αποκλίνει αρκετά από την κανονική κατανομή, και παρουσιάζεται μεγάλη συχνότητα σφαλμάτων σε τιμές διάφορες του μηδενός.



Γράφημα 5-5: Σενάριο 1, Κατανομή σφαλμάτων

Μετά την ολοκλήρωση της διαδικασίας εκπαίδευσης και της αξιολόγησης των αποτελεσμάτων, το νευρωνικό δίκτυο μετατρέπεται σε λειτουργία κλειστού βρόχου, κατά την οποία τροφοδοτείται από

αρχικές τιμές και εκείνο παράγει τα επόμενα βήματα. Οι αρχικές τιμές, στην παρούσα μελέτη, είναι οι τελευταίες τιμές της καταγεγραμμένης χρονοσειράς, έτσι ώστε το δίκτυο να παράγει τα επόμενα βήματα, δηλαδή την πρόβλεψη της κατανάλωσης ηλεκτρικής ενέργειας. Οι αρχικές τιμές που δίνονται στο δίκτυο είναι ίσες σε πλήθος με τις υστερήσεις δικτύου που έχουν οριστεί κατά τον σχεδιασμό του. Αναλυτικότερα, στην παράγραφο 4-1 εξηγείται το πως λειτουργούν οι υστερήσεις του δικτύου σε λειτουργία ανοιχτού και κλειστού βρόχου.



Γράφημα 5-6: Σενάριο 1: Πρόβλεψη κατανάλωσης ηλεκτρικής ενέργειας

Ανακεφαλαιώνοντας, η εκπαίδευση του δικτύου στο τρέχον σενάριο φαίνεται ικανοποιητική, καθώς η συνολική προσαρμογή του δικτύου στα δεδομένα της χρονοσειράς δίνει συντελεστή προσδιορισμού $R^2=0,91$. Αυτό θα μπορούσε να υποδεικνύει και υπερπροσαρμογή του δικτύου, αλλά στα διαγράμματα που συσχετίζουν τις πραγματικές με τις εκτιμώμενες τιμές (Γράφημα 5-3) παρατηρείται ότι η έξοδος του δικτύου στο πέρας της εκπαίδευσης έχει καλή προσαρμογή στα πραγματικά δεδομένα μεν, αλλά δεν τα ακολουθεί πιστά και δεν επηρεάζεται από τυχαίες μεταβολές. Παρατηρείται ότι το δίκτυο προβλέπει σωστά τις διακυμάνσεις και την τάση της χρονοσειράς, χωρίς όμως να ακολουθεί τυχαίες ακραίες τιμές ή μεμονωμένα συμβάντα. Η εκπαίδευση, δηλαδή, έχει σταματήσει έγκαιρα πριν γίνει υπερπροσαρμογή.

Η πρόβλεψη θεωρείται έμπιστη, διότι η τάση των προβλεπόμενων τιμών είναι αυξητική, οι διακυμάνσεις είναι παρόμοιες με τις παρελθοντικές, ενώ δεν παρατηρείται κάποια υπερβολικά υψηλή ή χαμηλή τιμή. Το συγκεκριμένο σενάριο, όμως, δεν λαμβάνει υπόψιν αστάθμητους παράγοντες, όπως η θερμοκρασία ή τα οικονομικά και δημογραφικά μεγέθη. Αυτό είναι φανερό και στις προβλεπόμενες τιμές, διότι σε σχέση με τα παρελθοντικά δεδομένα παρουσιάζουν πολύ πιο ομαλές τάσεις ενώ έχουν σταθερές αυξομειώσεις, οι οποίες όμως είναι παρόμοιες με αυτές των τελευταίων 4 ετών της υπάρχουσας χρονοσειράς. Η πρόβλεψη, σε γενικές γραμμές παρουσιάζει μεγάλη ομοιογένεια στην ετήσια εποχικότητα, πράγμα που θα μπορούσε να συμβεί σε πραγματικές συνθήκες. Αυτό που δεν είναι τόσο πιθανό είναι να επαναλαμβάνονται πιστά οι μεταβολές στο ετήσιο

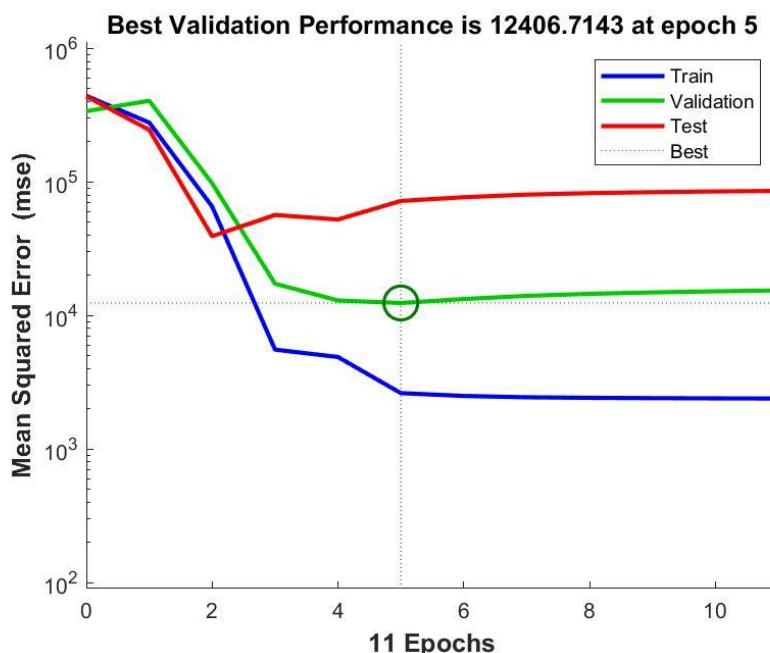
φορτίο για πολλά συναπτά έτη. Καθώς αλλάζουν οι καταναλωτικές και κοινωνικές συμπεριφορές, εξελίσσονται και διαφοροποιούνται οι ενεργειακές ανάγκες των νοικοκυριών μαζί με την εξέλιξη της τεχνολογίας και των οικουσκευών. Έτσι η κατανομή του φορτίου εντός του έτους είναι πολύ πιθανό να αλλάζει ανά κάποια χρόνια, συναρτήσει περιβαλλοντικών, κοινωνικών και οικονομικών παραγόντων.

Σενάριο 2

5.2 Στο Σενάριο 2, χρησιμοποιείται για την εκπαίδευση του δικτύου η χρονοσειρά ετήσιων καταγραφών κατανάλωσης ηλεκτρικής ενέργειας από το 1960 έως και το 2014. Τα δεδομένα, όπως περιγράφεται στην παράγραφο 4-3, παρέχονται από την Παγκόσμια Τράπεζα η οποία τηρεί ελεύθερες βάσεις δεδομένων για ποικίλους οικονομικούς, κοινωνικούς, περιβαλλοντικές μεταβλητές και αναπτυξιακούς δείκτες.

Το πρώτο γράφημα των αποτελεσμάτων είναι η πορεία του μέσου τετραγωνικού σφάλματος κατά την πάροδο των κύκλων εκπαίδευσης. Παρατηρείται πως το σφάλμα των δεδομένων ελέγχου και επαλήθευσης είναι αρκετά μεγαλύτερο από το σφάλμα των δεδομένων εκπαίδευσης. Το γεγονός αυτό είναι αρνητικό για την διαδικασία της εκπαίδευσης, αλλά εάν δεν σχολιαστούν και τα υπόλοιπα αποτελέσματα, όπως τα διαγράμματα κατανομής σφαλμάτων και η διασπορά των πραγματικών-εκτιμώμενων τιμών, δεν είναι δυνατό να εξαχθεί ολοκληρωμένο συμπέρασμα.

Παρόλα αυτά, το νευρωνικό δίκτυο ολοκληρώθηκε με τις παραμέτρους που προέκυψαν από τον 5^ο κύκλο εκπαίδευσης. Σε εκείνο το σημείο, όπως φαίνεται στο Γράφημα 5-7, το σφάλμα της επαλήθευσης παρουσιάζει ολικό ελάχιστο, ενώ ταυτόχρονα το σφάλμα της εκπαίδευσης λαμβάνει οριακά την χαμηλότερη τιμή του. Με αυτό τον τρόπο αποφεύγεται η υπερπροσαρμογή του δικτύου. Το σφάλμα των δεδομένων ελέγχου υπερβαίνει αρκετά τα δύο υπόλοιπα, πράγμα που μπορεί να δείχνει κακή γενίκευση του δικτύου στα δεδομένα της χρονοσειράς.

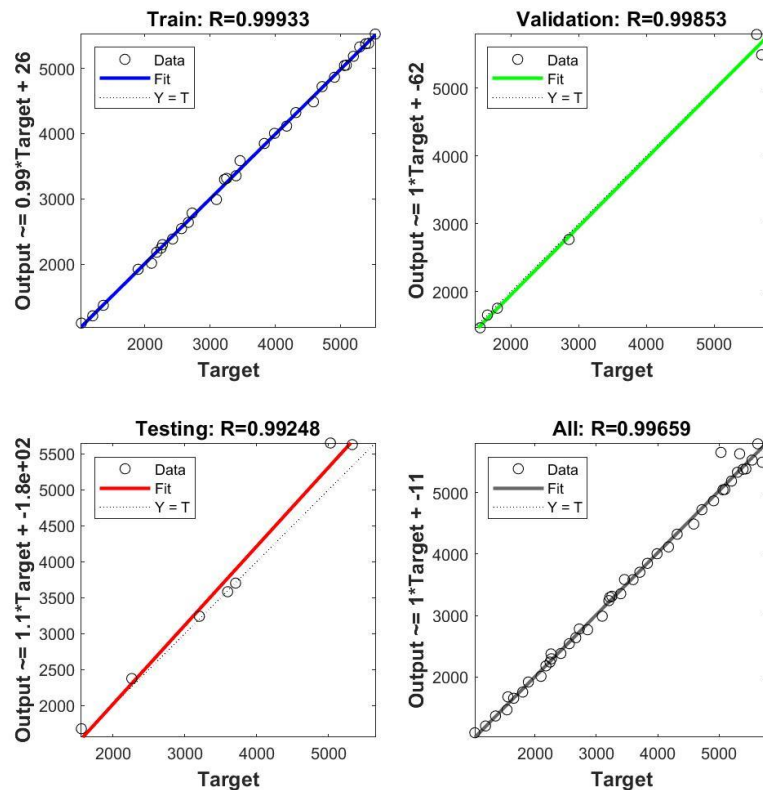


Γράφημα 5-7: Σενάριο 2, Μέσο τετραγωνικό σφάλμα κατά την εκπαίδευση

Στη συνέχεια ακολουθεί το σύνθετο γράφημα των παλινδρομήσεων. Ενώ, παραπάνω προκύπτει πως τα σφάλματα των δεδομένων ελέγχου και επαλήθευσης είναι αρκετά μεγαλύτερα από αυτά της εκπαίδευσης, στο παρακάτω γράφημα παλινδρόμησης φαίνεται πως η απόδοση του δικτύου είναι

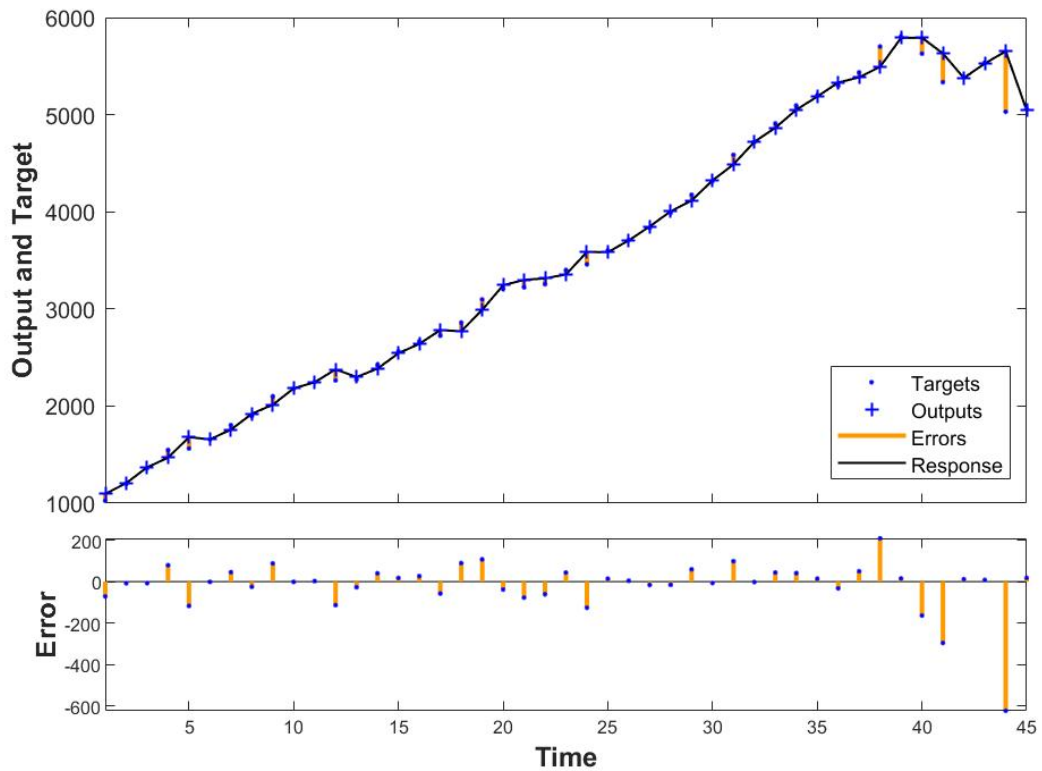
άκρως ικανοποιητική σε όλα τα σύνολα δεδομένων. Οι συντελεστές προσδιορισμού για τα 3 διαφορετικά σύνολα, καθώς και ο συνολικός συντελεστής διαφέρουν ελάχιστα μεταξύ τους. Παράλληλα και οι 4 συντελεστές είναι αρκετά υψηλοί, της τάξεως του $R^2=0,997$.

Με βάση τα παραπάνω η απόδοση του δικτύου κρίνεται πολύ υψηλή, αν και υπάρχει πιθανότητα να έχει γίνει υπερπροσαρμογή με αποτέλεσμα η γενίκευση του δικτύου να μην είναι καλή και να μην είναι σε θέση να εξάγει ασφαλή πρόβλεψη. Στη συνέχεια, τα γραφήματα κατανομής και αυτοσυσχέτισης σφαλμάτων, αλλά και η ίδια η πρόβλεψη, θα διαμορφώσουν καλύτερα την εικόνα σχετικά με την ποιότητα του εκπαιδευμένου δικτύου.



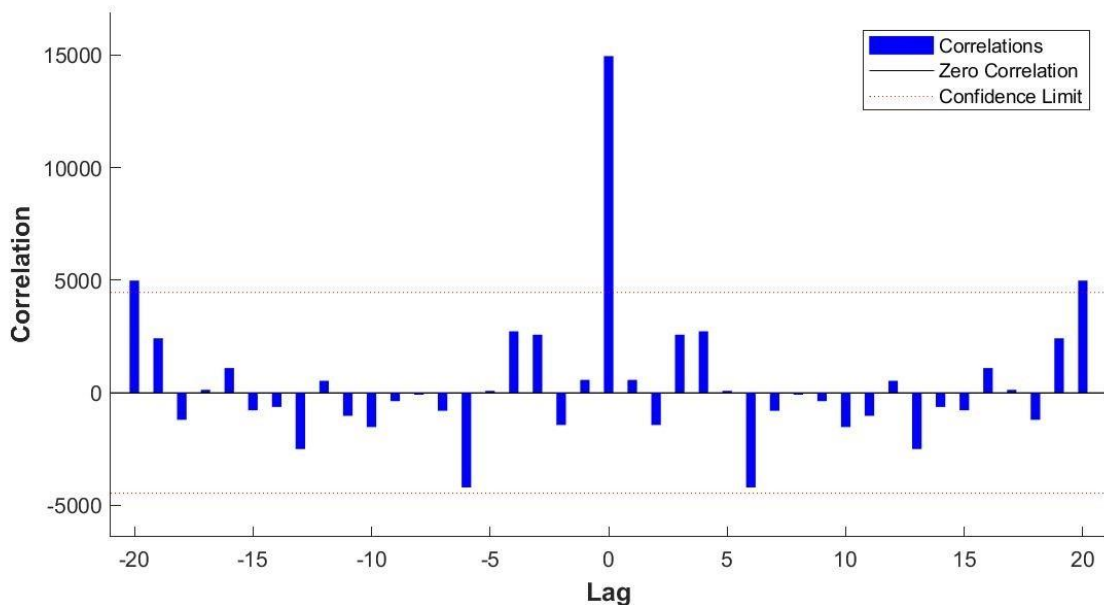
Γράφημα 5-8: Σενάριο 2, Παλινδρόμηση πραγματικών- εκτιμώμενων τιμών

Ακολουθεί το γράφημα διασποράς των πραγματικών καταγραφών της χρονοσειράς και των αντίστοιχων εκτιμώμενων από το δίκτυο. Όπως φαίνεται, το δίκτυο ακολουθεί με μεγάλη ακρίβεια την ιστορική χρονοσειρά κατανάλωσης ηλεκτρικής ενέργειας. Σφάλματα υπάρχουν, αλλά περιορίζονται κυρίως στα τελευταία βήματα της χρονοσειράς, όπου η κατανάλωση ενέργειας μεταβλήθηκε ραγδαία και άναρχα λόγω της οικονομικής κρίσης. Παρόλα αυτά, παρατηρείται ότι το νευρωνικό δίκτυο αφομοίωσε την πτωτική τάση των τελευταίων ετών, χωρίς όμως να την ακολουθεί πιστά. Αυτό δείχνει καλή γενίκευση και αποφυγή υπερπροσαρμογής.



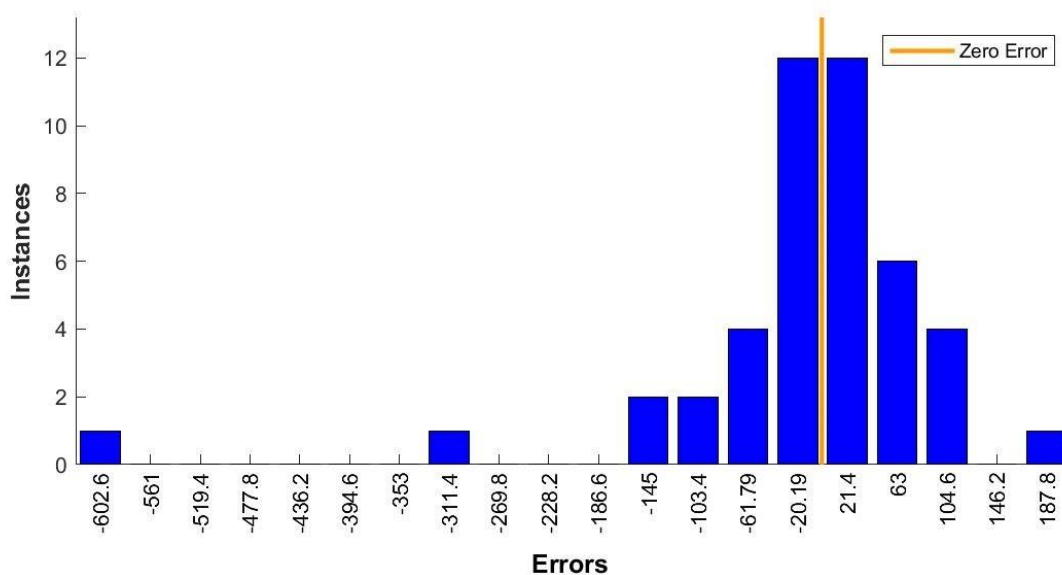
Γράφημα 5-9: Σενάριο 2, Διασπορά πραγματικών- εκτιμώμενων τιμών

Μετάπειτα παρατίθεται το γράφημα αυτοσυσχέτισης σφαλμάτων, όπου παρατηρείται ότι οι αυτοσυσχετίσεις κυμαίνονται σε πολύ μικρά επίπεδα, εντός του διαστήματος εμπιστοσύνης. Οριακά εκτός του διαστήματος εμφανίζονται σφάλματα με περιοδικότητα 20 ετών, που αποτελούν και τα μεγαλύτερα σε τιμή. Αμέσως μετά ακολουθεί σφάλμα με περιοδικότητα 6.



Γράφημα 5-10: Σενάριο 2, Αυτοσυσχέτιση σφαλμάτων

Επιβεβαιώνοντας την καλή απόδοση του δικτύου, στο Γράφημα 5-11 παρατηρείται πως ο μεγαλύτερος αριθμός σφαλμάτων περιορίζεται σε μικρές τιμές, ενώ η κατανομή τους είναι κανονική με πολύ μικρή διασπορά.



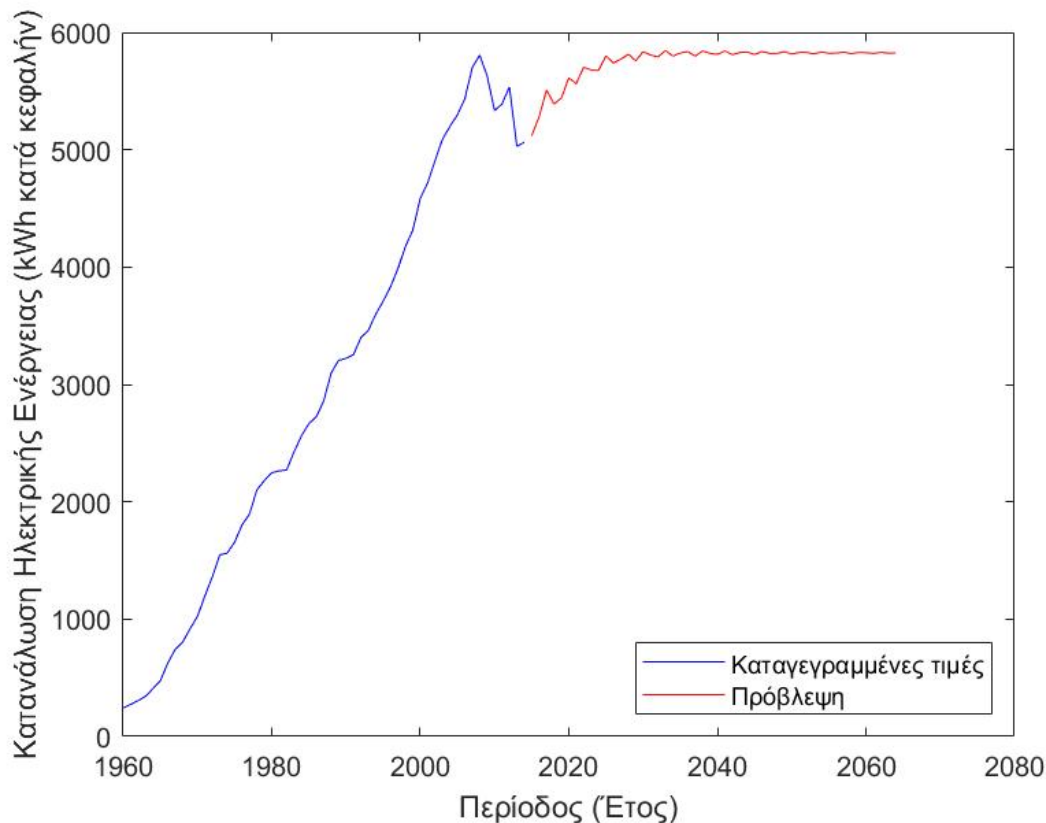
Γράφημα 5-11: Σενάριο 2, Κατανομή σφαλμάτων

Είναι φανερό πως το Σενάριο 2 παρουσιάζει πολύ καλύτερα αποτελέσματα εκπαίδευσης σε σχέση με το Σενάριο 1. Συνοπτικά, ο συντελεστής προσδιορισμού R^2 είναι υπερβολικά υψηλός, παράγοντας που προϋποθέτει πολύ καλή προσαρμογή της εκπαίδευσης στα δεδομένα αλλά εμπεριέχει τον κίνδυνο της υπερπροσαρμογής του δικτύου. Παρόλα αυτά, στο Γράφημα 5-7 φαίνεται πως η εκπαίδευση έχει ολοκληρωθεί σε σημείο που το σφάλμα αξιολόγησης (Validation Error) παραμένει στο χαμηλότερο δυνατό επίπεδο. Έτσι, θα γίνει δοκιμή του δικτύου σε λειτουργία κλειστού βρόχου, για να πραγματοποιηθεί πρόβλεψη.

Με βάση το διάγραμμα 5-12, παρακάτω, το δίκτυο έχει προβλέψει τα επόμενα βήματα της χρονοσειράς με τέτοιο τρόπο ώστε η κατά κεφαλήν κατανάλωση ενέργειας να συνεχίσει να αυξάνει, γεγονός άκρως ρεαλιστικό. Το δίκτυο δεν επηρεάστηκε αρνητικά από το ακραίο συμβάν απότομης μείωσης της κατανάλωσης ενέργειας την περίοδο της οικονομικής κρίσης, έπειτα από 40 έτη συνεχούς αύξησης. Αντιθέτως, αφομοίωσε τις μεταβολές αυτές και προσάρμοσε τα μελλοντικά βήματα έτσι ώστε να ακολουθούν αυξητική τάση, αλλά με διακυμάνσεις παρόμοιες με αυτές της πτωτικής τάσης των τελευταίων ετών.

Η πρόβλεψη μπορεί να θεωρηθεί αξιόπιστη μέχρι και το έτος 2030 κατά προσέγγιση, καθώς στη συνέχεια παρατηρείται σταθεροποίηση των προβλεπόμενων τιμών γύρω από μια περιοχή. Αυτό είναι αναμενόμενο για προβλέψεις με τόσο μακρύ ορίζοντα και δεν σημαίνει ότι το δίκτυο δεν έχει καλή απόκριση στις προηγούμενες προβλεπόμενες τιμές τους.

Ένα χαρακτηριστικό που θα μπορούσε να μην αντικατοπτρίζει την μελλοντική πραγματικότητα είναι το γεγονός ότι η πρόβλεψη δεν ξεπερνά σε τιμές εκείνες της μέγιστης παρελθοντικής τιμής του 2008. Φαίνεται πως η μελλοντική χρονοσειρά αυξάνει, αλλά τείνει να σταθεροποιείται σε τιμές παρόμοιες με εκείνες του 2008 χωρίς όμως να τις ξεπερνά. Σε επόμενες παραγράφους της παρούσας μελέτης (Παρ. 5-6) πραγματοποιείται σύγκριση των προβλέψεων από κάθε Σενάριο, καθώς και με τις προβλέψεις που προκύπτουν από την μελέτη επάρκειας του ΑΔΜΗΕ. Έτσι, είναι εφικτό να σχολιαστούν μεθοδικότερα τα αποτελέσματα και να γίνει σύγκριση με περισσότερα κριτήρια.



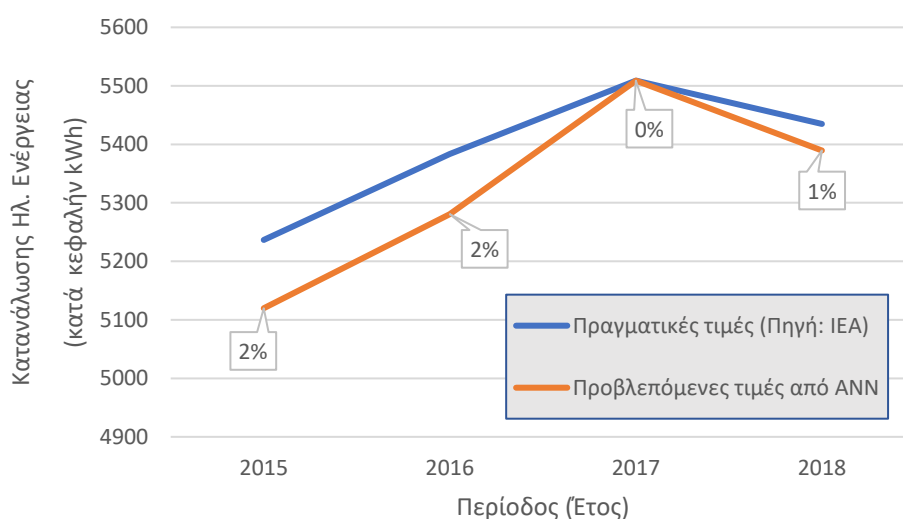
Γράφημα 5-12 : Σενάριο 2, Πρόβλεψη κατανάλωσης ηλεκτρικής ενέργειας

Ιδιαίτερο ενδιαφέρον για την αξιολόγηση της πρόβλεψης του νευρωνικού δικτύου του παρόντος Σεναρίου έχει να συγκρίνουμε τις προβλεπόμενες τιμές που προκύπτουν για τα έτη 2015, 2016, 2017 και 2018 με τις αντίστοιχες πραγματικές. Για το δίκτυο οι τιμές των προαναφερθέντων ετών αποτελούν προβλέψεις, καθώς στα δεδομένα της αρχικής χρονοσειράς της Παγκόσμιας Τράπεζας δεν συμπεριλαμβάνονται (Παράγραφος 4-3, Γράφημα 4-2), οπότε δεν χρησιμοποιήθηκαν για την εκπαίδευση του δικτύου. Υπάρχουν, όμως, διαθέσιμες πραγματικές τιμές από την χρονοσειρά του International Energy Agency (Παράγραφος 4-3 – Γράφημα 4-1) που χρησιμοποιήθηκε στο Σενάριο 1 (Παράγραφος 5-1).

Για να γίνει σύγκριση των προβλεπόμενων τιμών του σεναρίου για τα έτη 2015-2018 σε σχέση με τις πραγματικές καταναλώσεις υπολογίστηκαν οι αντίστοιχες τιμές κατανάλωσης αθροίζοντας τις μηνιαίες τιμές που παρέχονται από το International Energy Agency. Στη συνέχεια, με βάση τις απογραφές της ΕΛΣΤΑΤ για τον πληθυσμό της Ελλάδας (Παράγραφος 4-3) υπολογίστηκαν οι κατά κεφαλήν καταναλώσεις.

Όπως φαίνεται στο Γράφημα 5-13, οι προβλεπόμενες τιμές από το τεχνητό νευρωνικό δίκτυο για τα έτη 2015-2017 είναι πάρα πολύ κοντά στις πραγματικές τιμές. Το νευρωνικό δίκτυο κατάφερε να προβλέψει την κατανάλωση ενέργειας με απόκλιση της τάξης 0% έως 2%. Η απόκλιση για κάθε έτος ξεχωριστά παρουσιάζεται στο εν λόγω γράφημα.

Έτσι, συμπεραίνεται ότι οι αυξομειώσεις που παρουσιάζει η έξοδος του δικτύου για την πρόβλεψη των μελλοντικών ετών είναι έγκυρη, καθώς κατάφερε να προβλέψει με ακρίβεια και σωστό ρυθμό την άνοδο της κατανάλωσης μέχρι το έτος 2017, ενώ ταυτόχρονα προέβλεψε την πτώση της κατανάλωσης, επίσης με σωστό ρυθμό, για το έτος 2018. Το παραπάνω αποτέλεσμα θεωρείται σημαντικό και δεν μπορεί να θεωρηθεί τυχαίο, καθώς ήδη από την εκπαίδευση το δίκτυο παρουσίασε πολύ καλή προσαρμογή και γενίκευση στα δεδομένα.



Γράφημα 5-13: Σενάριο 2, Σύγκριση πραγματικών – προβλεπόμενων τιμών 2015-2018

Σενάριο 3

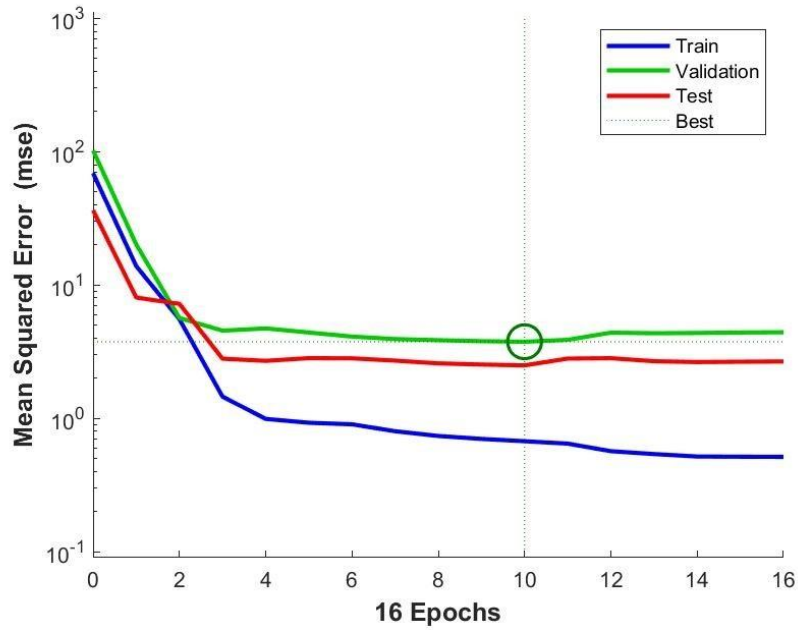
5.3

Το Σενάριο 3 περιλαμβάνει 2 νευρωνικά δίκτυα. Το πρώτο χρησιμοποιήθηκε για την πρόβλεψη της θερμοκρασίας, αξιοποιώντας ιστορικές καταγραφές, ενώ το δεύτερο χρησιμοποίησε την πρόβλεψη αυτή για να προβλέψει, με την σειρά του, την μελλοντική κατανάλωση ηλεκτρικής ενέργειας. Το πρώτο νευρωνικό δίκτυο είναι τύπου NAR, ενώ το δεύτερο είναι τύπου NARX, όπως επεξηγείται στην παράγραφο 4.2

Πρόβλεψη Θερμοκρασίας

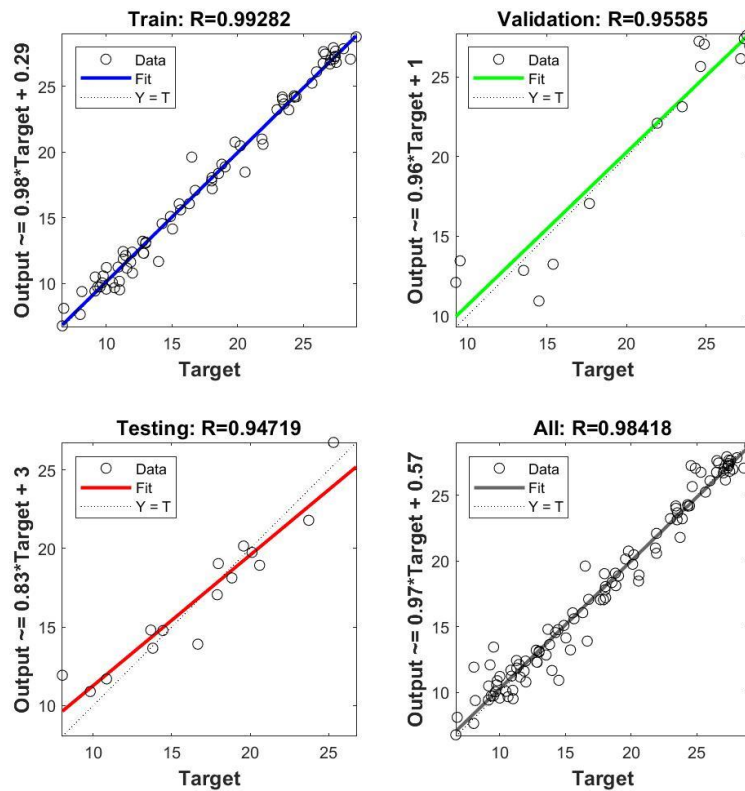
Κατά την ολοκλήρωση της εκπαίδευσης του νευρωνικού δικτύου τύπου NAR παρουσιάζεται, αρχικά, το γράφημα της προόδου των σφαλμάτων. Όπως φαίνεται, ο βέλτιστος και επιλεγόμενος κύκλος εκπαίδευσης για την τελική παραμετροποίηση του δικτύου είναι ο 10^{ος}, κατά τον οποίο, όπως και σε όλα τα υπόλοιπα σενάρια, το σφάλμα των δεδομένων επαλήθευσης λαμβάνει την χαμηλότερη του τιμή. Παράλληλα, το σφάλμα της εκπαίδευσης παραμένει χαμηλό.

Παρόμοιο χαρακτηριστικό με το Σενάριο 2 είναι το γεγονός ότι τα σφάλμα των δεδομένων ελέγχου και επαλήθευσης είναι μεγαλύτερα από εκείνο των δεδομένων εκπαίδευσης. Στο παρόν σενάριο η απόκλιση των πρώτων με τα δεύτερα είναι αισθητά μικρότερη. Στα υπόλοιπα σενάρια ο επιλεγόμενος κύκλος εκπαίδευσης ήταν μικρότερος, ο 3^{ος} για το Σενάριο 1 και ο 2^{ος} για το Σενάριο 2. Στο παρόν Σενάριο χρειάστηκαν 10 κύκλοι εκπαίδευσης για να παρουσιάσει το σφάλμα επαλήθευσης ολικό ελάχιστο.



Γράφημα 5-14: Σενάριο 3α, Μέσο τετραγωνικό σφάλμα κατά την εκπαίδευση

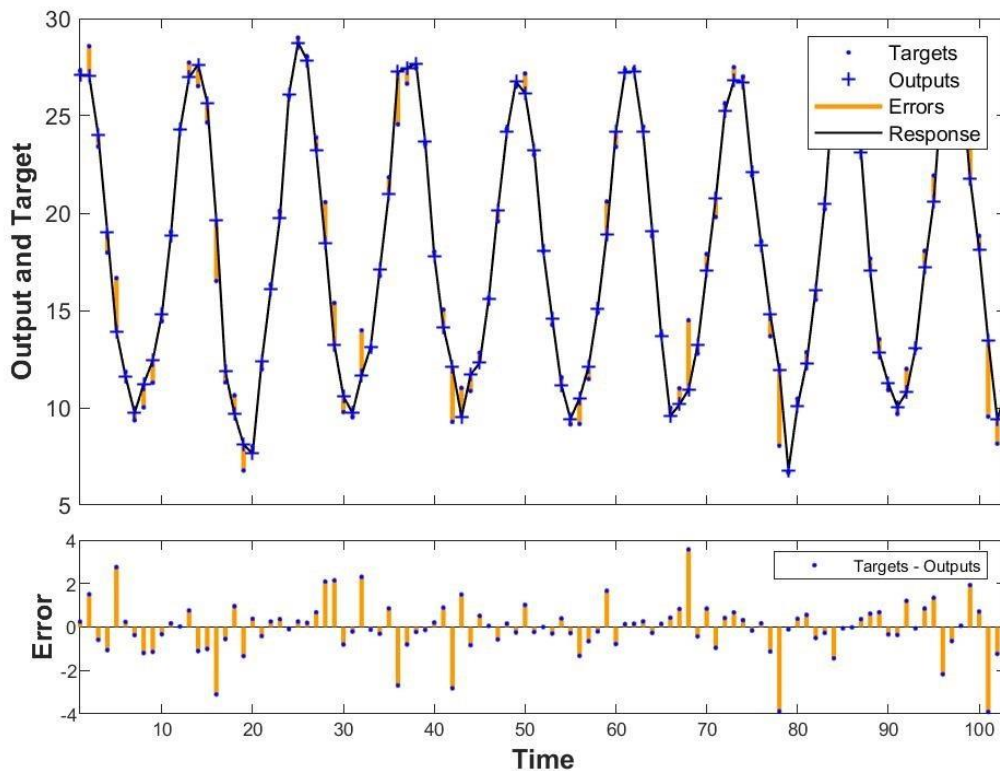
Στη συνέχεια ακολουθεί το γράφημα των παλινδρομήσεων, στο οποίο φαίνεται η καλή προσαρμογή του δικτύου στα δεδομένα. Και οι 3 συντελεστές προσδιορισμού για τα 3 διαφορετικά υποσύνολα δεδομένων είναι αρκετά υψηλοί, με χαμηλότερο εκείνο των δεδομένων ελέγχου, κάτι που είναι αναμενόμενο. Ο συνολικός συντελεστής προσδιορισμού ανέρχεται στο $R^2=0,984$.



Γράφημα 5-15: Σενάριο 3α, Παλινδρόμηση πραγματικών – εκτιμώμενων τιμών

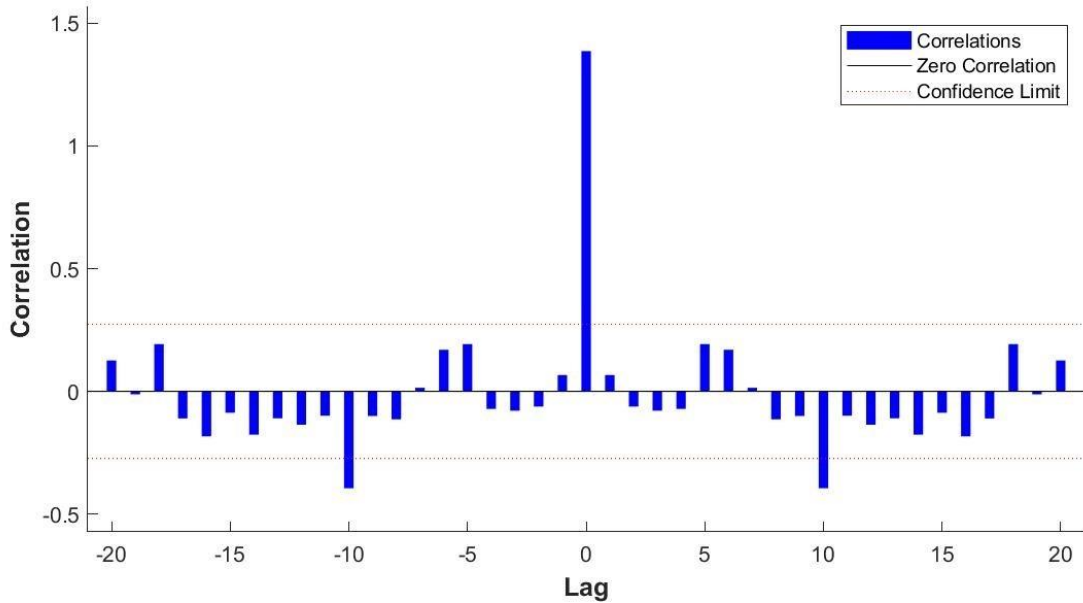
Το γράφημα διασποράς των εκτιμώμενων με τις πραγματικές τιμές (5-16) επιβεβαιώνει την καλή προσαρμογή του δικτύου και δεν αφήνει ιδιαίτερα περιθώρια για υπερπροσαρμογή του δικτύου. Αυτό εξηγείται καθώς στο γράφημα παρατηρείται πως το εξαγόμενη χρονοσειρά ακολουθεί την πραγματική, αλλά υπάρχουν σφάλματα τα οποία εντοπίζονται σε καταγραφές οι οποίες είναι τυχαίες μεταβολές των κανονικών διακυμάνσεων της χρονοσειράς και οφείλονται στην στοχαστικότητα της μεταβλητής. Σε περίπτωση υπερπροσαρμογής τα σφάλματα θα ήταν ελάχιστα και η εξαγόμενη χρονοσειρά θα ήταν πιστή αντιγραφή της πραγματικής.

Ένας από τους παράγοντες που οδηγούν στην καλή προσαρμογή του δικτύου στην χρονοσειρά είναι ότι οι θερμοκρασιακές μεταβολές ακολουθούν σε βάθος χρόνου, πολύ αυστηρές διακυμάνσεις εντός του έτους, ενώ οι τυχαίες αποκλίσεις από την συνήθη περιοδικότητα συμβαίνουν αλλά αποτελούν εξαιρέσεις.



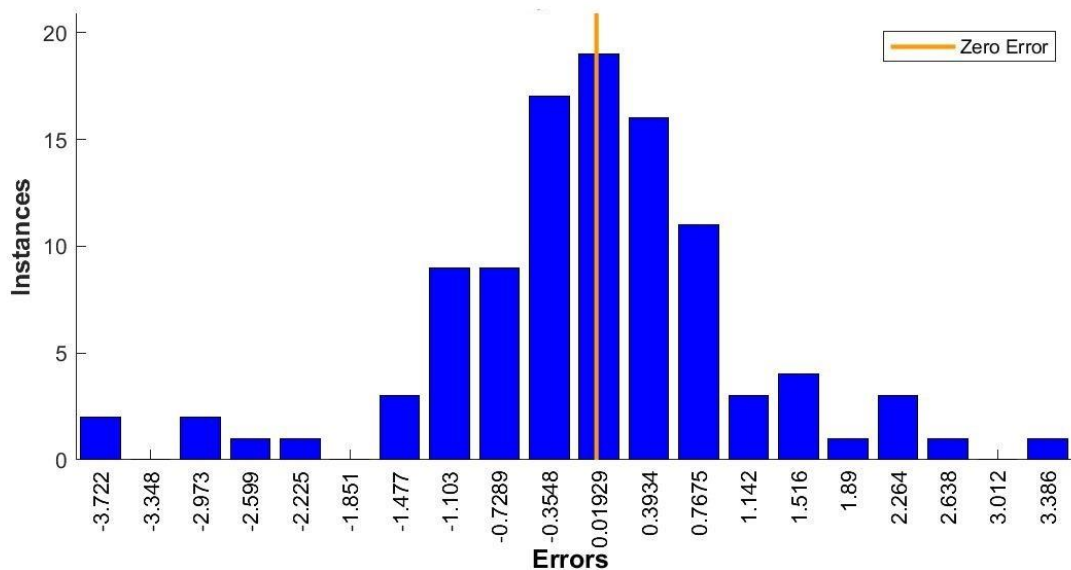
Γράφημα 5-16: Σενάριο 3α, Διασπορά πραγματικών- εκτιμώμενων τιμών

Στο επόμενο γράφημα ακολουθεί η αυτοσυσχέτιση των σφαλμάτων. Όπως εξηγείται και στην παράγραφο 5.1 η αυτοσυσχέτιση είναι πολύ σημαντικό να είναι μικρή, διότι αποδεικνύει ότι η έξοδος του δικτύου παρουσιάζει μόνο τυχαία σφάλματα και δεν υπάρχει κάποια παράμετρος η οποία να έχει εσφαλμένες τιμές κατά την ολοκλήρωση της εκπαίδευσης που να εξάγει συστηματικό σφάλμα. Η αυτοσυσχέτιση στο τρέχον Σενάριο είναι ελάχιστη, ενώ όλα τα επαναλαμβανόμενα σφάλματα είναι πολύ μικρά, εντός του ορίου εμπιστοσύνης. Μόνο μία τιμή σφάλματος γύρω στο 0,7 ελάχιστα το όριο, το οποίο έχει περιοδικότητα 10 και δεν μπορεί να συσχετιστεί με κάποιο παράγοντα, οπότε θεωρείται τυχαίο συμβάν.



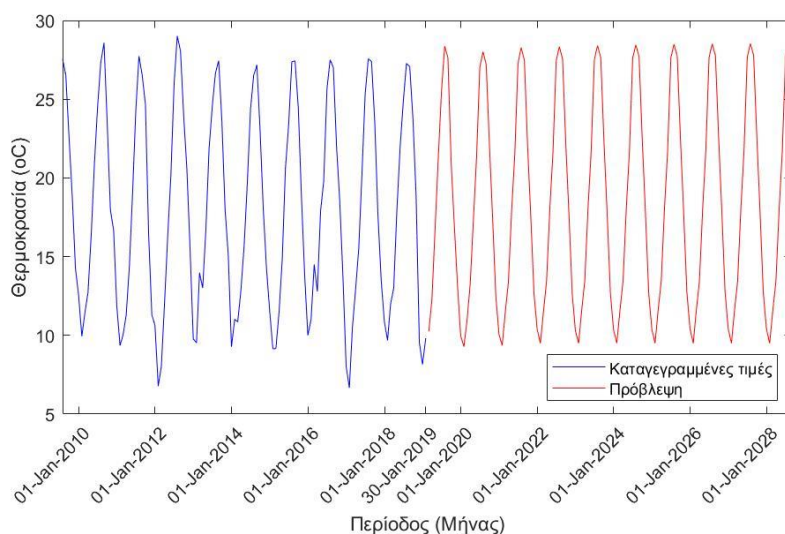
Γράφημα 5-17: Σενάριο 3α, Αυτοσυσχέτιση σφαλμάτων

Όπως και το παραπάνω γράφημα των αυτοσυσχετίσεων των σφαλμάτων, έτσι και το γράφημα 5-18 της κατανομής των σφαλμάτων αποδεικνύει την καλή απόδοση του δικτύου. Όπως φαίνεται στο εν λόγω γράφημα η κατανομή των σφαλμάτων συγκλίνει κοντά στο μηδέν, ενώ έχει αρκετά χαρακτηριστικά κανονικής κατανομής. Διαφέρει σε κάποια σημεία και δεν είναι απολύτως συμμετρική. Επίσης, υπάρχουν κάποιες μεγαλύτερες τιμές σφαλμάτων οι οποίες έχουν λίγο μεγαλύτερη συχνότητα από ότι θα έπρεπε, καθώς βρίσκονται στις ακραίες θέσεις της κατανομής. Τα παραπάνω δεν μας προβληματίζουν, διότι είναι αναμενόμενο να υπάρχουν σφάλματα σε σχέση με την πραγματική χρονοσειρά, εφόσον υπάρχει καλή γενίκευση στο δίκτυο και δεν έχει γίνει υπερπροσαρμογή. Ο σκοπός είναι αποτελέσματα να έχουν καλή απόδοση και όχι να έχουν απόλυτη μορφή τα γραφήματα, κάτι που θα ήταν ουτοπικό.



Γράφημα 5-18: Σενάριο 3α, Κατανομή σφαλμάτων

Εφόσον έχει ολοκληρωθεί η διαδικασία της εκπαίδευσης και έχει κριθεί πως ικανοποιεί τα κριτήρια για την καλή γενίκευση του δικτύου, τότε προχωρούμε στην πρόβλεψη με λειτουργία κλειστού βρόχου. Η πρόβλεψη των μελλοντικών βημάτων της χρονοσειράς παρουσιάζεται στο διάγραμμα 5-19 και φαίνεται πως το δίκτυο παράγει αρκετά καλό αποτέλεσμα.

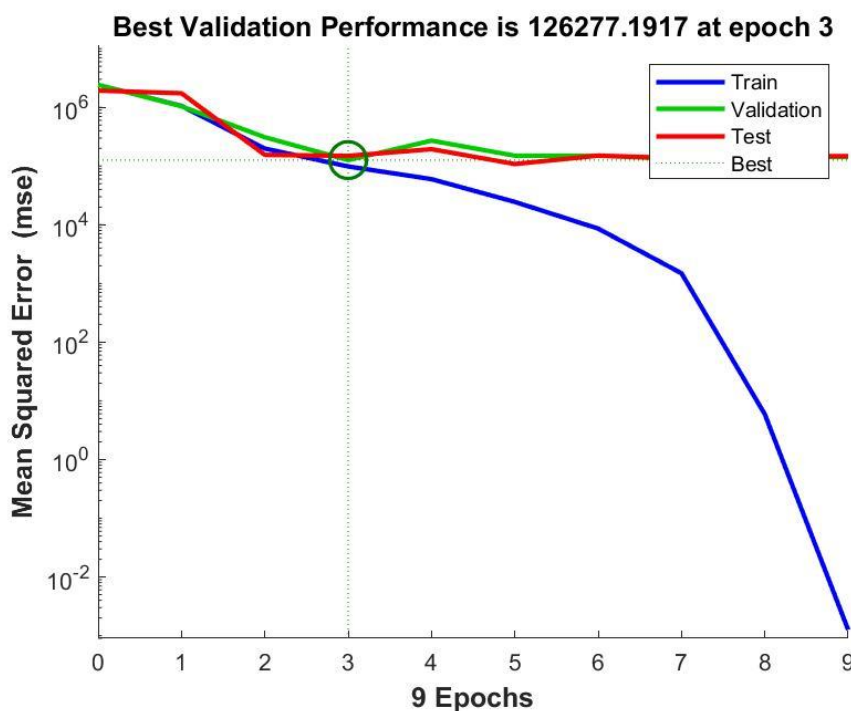


Γράφημα 5-19: Σενάριο 3α, Πρόβλεψη θερμοκρασίας

Μετά από δοκιμές και αξιολόγηση των διαφορετικών προβλέψεων που προκύπτουν σε κάθε εκπαίδευση του νευρωνικού δικτύου, κρίθηκε σκόπιμο να χρησιμοποιηθεί η παραπάνω πρόβλεψη θερμοκρασίας, η οποία αποτελεί το χειρίστο σενάριο για την αύξηση της θερμοκρασίας του πλανήτη. Η μέση θερμοκρασία των καταγεγραμμένων δεδομένων είναι $17,95^{\circ}\text{C}$ ενώ η μέση τιμή της χρονοσειράς των προβλέψεων είναι $18,5^{\circ}\text{C}$, γεγονός που φανερώνει αύξηση της θερμοκρασίας κατά $1,5^{\circ}\text{C}$ περίπου. Παράλληλα με τον μέσο όρο, όπως φαίνεται στο γράφημα 5-19, αυξάνονται οι μέγιστες και οι ελάχιστες τιμές θερμοκρασιών εντός του έτους, δημιουργώντας έτσι μεγάλες διακυμάνσεις. Η αύξηση της μέσης τιμής στο σενάριο αυτό είναι υπερβολική και έχει άμεσες επιπτώσεις στο κλίμα και κατ' επέκταση στην κατανάλωση ενέργειας. Η πρόβλεψη αυτή χρησιμοποιήθηκε έτσι ώστε να εξεταστεί η συμπεριφορά του νευρωνικού δικτύου σε δοκιμές με ακραία φαινόμενα. Συν τοις άλλοις, με αυτό τον τρόπο γίνεται διερεύνηση της επίδρασης της αύξησης της θερμοκρασίας στην κατανάλωση ηλεκτρικής ενέργειας, όπως την προσομοιώνει ένα νευρωνικό δίκτυο.

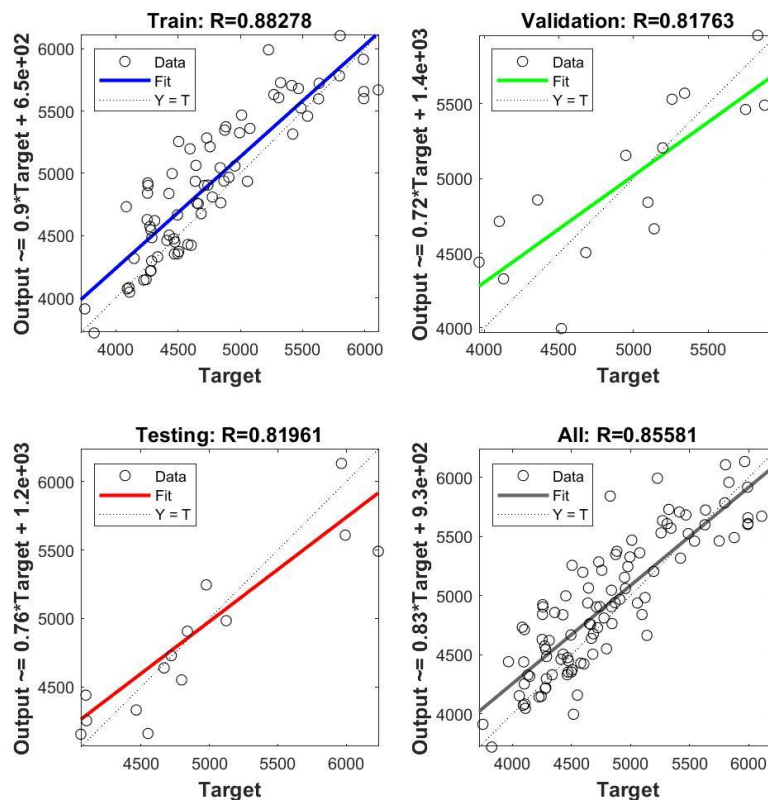
Πρόβλεψη Κατανάλωσης Ενέργειας

Μετά την ολοκλήρωση της πρόβλεψης της βοηθητικής χρονοσειράς στο δίκτυο NAR, τα υπάρχοντα καταγεγραμμένα δεδομένα κατανάλωσης ενέργειας και θερμοκρασίας αξιοποιούνται σε δίκτυο τύπου NARX για την πρόβλεψη της κατανάλωσης ενέργειας. Ακολουθεί το διάγραμμα εξέλιξης του μέσου τετραγωνικού σφάλματος (Γράφημα 5-20). Ο βέλτιστος κύκλος εκπαίδευσης επιτυγχάνεται σύντομα, ύστερα από 3 επαναλήψεις, ενώ εκείνο το σημείο παρουσιάζεται η ελάχιστη τιμή του σφάλματος ελέγχου και επαλήθευσης, όπως και στα υπόλοιπα Σενάρια. Μετά από κάποιους κύκλους φαίνεται πως το σφάλμα του συνόλου δεδομένων εκπαίδευσης φθίνει απότομα και σε πολύ χαμηλές τιμές. Σε αυτά τα επίπεδα, είναι σίγουρο πως πρόκειται για υπερδιαστασιολόγηση και αποτυχία εκπαίδευσης του δικτύου, οπότε απορρίπτονται εξ' ορισμού. Ακόμη και το γεγονός ότι το σφάλμα επαλήθευσης και ελέγχου σε μετέπειτα κύκλους από το βέλτιστο (π.χ. κύκλος 5) κυμαίνεται στα ίδια χαμηλά επίπεδα με τον κύκλο 3 δεν είναι ικανό για να οριστεί ο κύκλος 5 ως βέλτιστος, καθώς η απότομη φθίνουσα πορεία του σφάλματος εκπαίδευσης δείχνει την αποτυχία του δικτύου. Για τον βέλτιστο κύκλο εκπαίδευσης που επιλέχθηκε από τον αλγόριθμο, θεωρούμε πως το δίκτυο λειτουργεί ικανοποιητικά, μένει μόνο να ελεγχθούν και τα γραφήματα των υπόλοιπων παραγόντων που ακολουθούν στη συνέχεια.



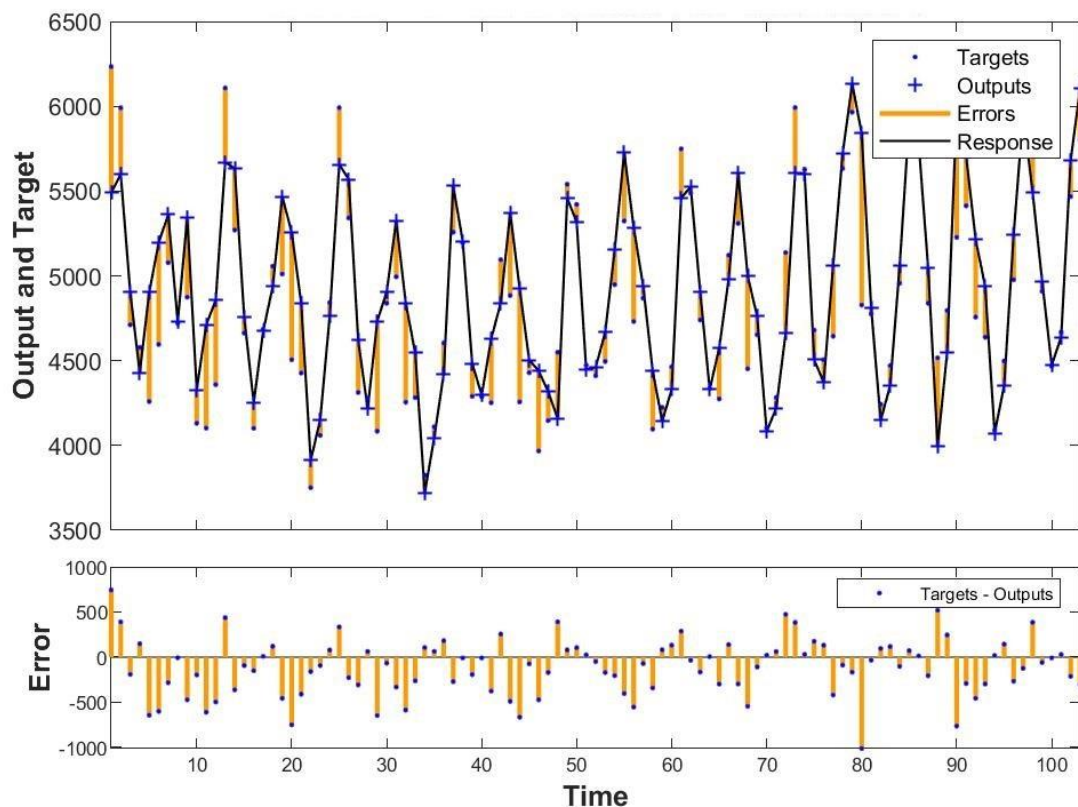
Γράφημα 5-20: Σενάριο 3b, Μέσο τετραγωνικό σφάλμα κατά την εκπαίδευση

Ακολουθεί το γράφημα των παλινδρομήσεων (Γράφημα 5-21). Η προσαρμογή του δικτύου είναι ικανοποιητική. Αν συγκριθεί, όμως, με το Σενάριο 1 (Παράγραφος 5.1) για το οποίο έχει χρησιμοποιηθεί η ίδια χρονοσειρά κατανάλωσης ηλεκτρικής, χωρίς την συμβολή της θερμοκρασίας, η απόδοση είναι μικρότερη και η εκπαίδευση δίνει μικρότερες τιμές συσχέτισης και συντελεστή προσδιορισμού ($R^2=0,856$ για το τρέχον σενάριο και $R^2=0,91$ για το Σενάριο 1). Με βάση την πολυπλοκότητα του τρέχοντος Σεναρίου σε σχέση με το Σενάριο 1 και δεδομένου ότι δίνονται περισσότερα δεδομένα στο σύστημα, που επηρεάζουν την κύρια χρονοσειρά, η αρχική θεώρηση είναι πως θα έπρεπε να εξάγονται καλύτερα αποτελέσματα.



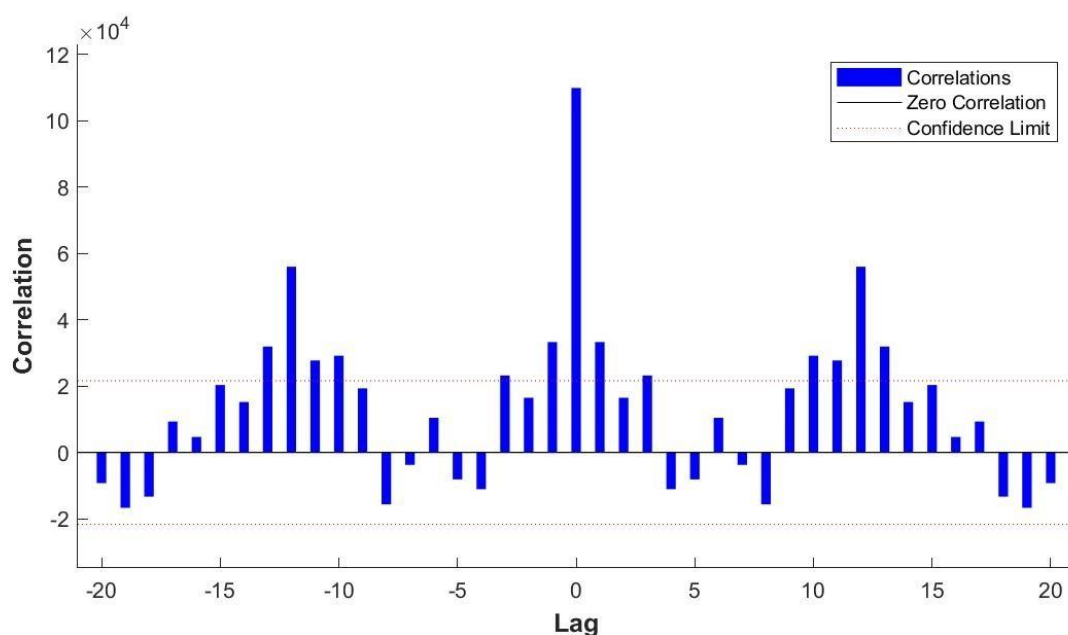
Γράφημα 5-21: Σενάριο 3b, Παλινδρόμηση πραγματικών- εκτιμώμενων τιμών

Στο γράφημα διασποράς των πραγματικών και εκτιμώμενων τιμών του δικτύου φαίνεται πως υπάρχουν σφάλματα σε όλο το εύρος της χρονοσειράς. Η εκτίμηση του νευρωνικού δικτύου αποκλίνει σε διάφορα σημεία από τις πραγματικές διακυμάνσεις. Τα σφάλματα εντοπίζονται κυρίως στις αιχμές και στα ελάχιστα του φορτίου, για τα οποία το δίκτυο εκτιμά υψηλότερες και χαμηλότερες τιμές αντίστοιχα. Στις μεταβατικές περιόδους, από την ελάχιστη κατανάλωση έως την αιχμή, για κάθε εξαμηνιαία περίοδο το δίκτυο ακολουθεί ικανοποιητικά τα πραγματικά δεδομένα. Με βάση τα παραπάνω, η απόδοση του δικτύου δεν φαίνεται να είναι εξίσου ικανοποιητική σε σχέση με το Σενάριο 1. Ένας λόγος που μπορεί να συμβαίνει αυτό είναι πως για την υλοποίηση του τρέχοντος Σεναρίου χρησιμοποιήθηκε μικρότερο εύρος καταγεγραμμένων τιμών ηλεκτρικής ενέργειας, έτσι ώστε να αντιστοιχεί στην χρονοσειρά των θερμοκρασιακών μεταβολών. Όσο λιγότερα είναι τα διαθέσιμα δεδομένα, τόσο πιο δύσκολη είναι η προσαρμογή του δικτύου στην χρονοσειρά, κάτι που είναι βασικό χαρακτηριστικό των τεχνητών νευρωνικών δικτύων.



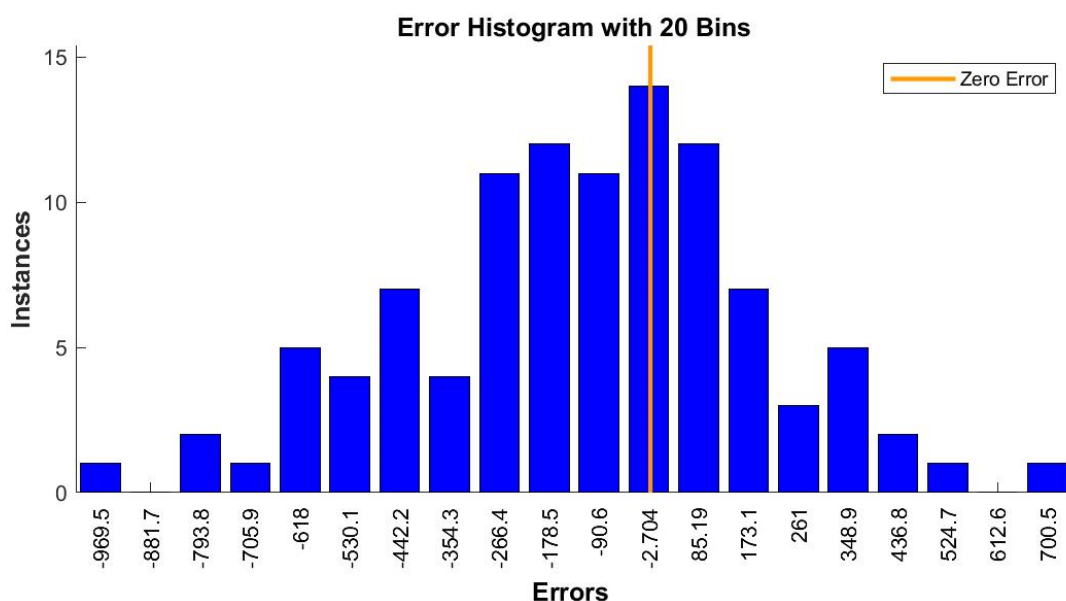
Γράφημα 5-22: Σενάριο 3b, Διασπορά πραγματικών- εκτιμώμενων τιμών

Το γράφημα αυτοσυσχέτισης των σφαλμάτων του δικτύου δεν αποκλίνει από τα παραπάνω συμπεράσματα, καθώς παρατηρείται μεγάλο εύρος σφαλμάτων τα οποία είναι εκτός του ορίου εμπιστοσύνης. Το μεγαλύτερο σφάλμα έχει ετήσια περιοδικότητα (12 μήνες), ενώ οι υπόλοιπες αυτοσυσχετίσεις φθίνουν γύρω από την τιμή αυτού.



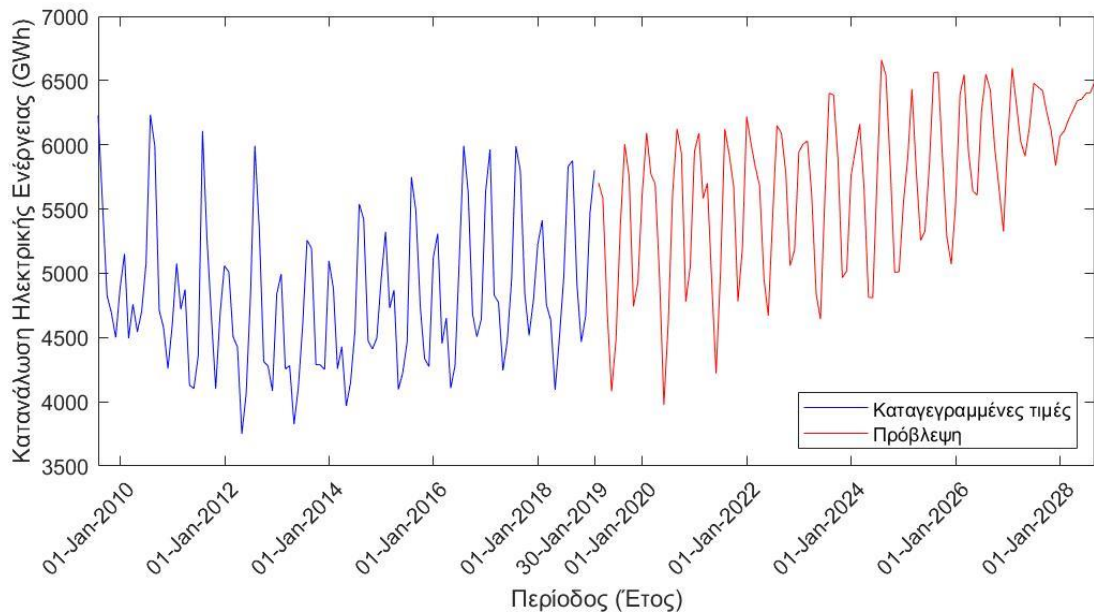
Γράφημα 5-23: Σενάριο 3b, Αυτοσυσχέτιση σφαλμάτων

Αντίστοιχα, η κατανομή των σφαλμάτων αποκλίνει αρκετά από την κανονική κατανομή, ενώ δεν είναι συμμετρική. Όπως φαίνεται στο γράφημα υπάρχει πληθώρα υψηλών σφαλμάτων που εμφανίζονται με μεγάλη συχνότητα, ενώ η τιμή του πιο συχνού σφάλματος που αποτελεί την κορυφή της κατανομής δεν συγκλίνει στο μηδέν (-2,704).



Γράφημα 5-24: Σενάριο 3b, Κατανομή σφαλμάτων

Αν και η αξιολόγηση των αποτελεσμάτων της εκπαίδευσης συγκλίνει πως το δίκτυο δεν έχει ικανοποιητική απόδοση εκπαίδευσης, η πρόβλεψη που προκύπτει και απεικονίζεται στο γράφημα 5-25 έχει όλα τα χαρακτηριστικά μιας πραγματικής χρονοσειράς κατανάλωσης ηλεκτρικής ενέργειας. Αν εξαιρέσουμε τα τελευταία βήματα της πρόβλεψης από το 2027 και ύστερα, η υπόλοιπη χρονοσειρά παρουσιάζει αυξητική τάση, ενώ οι ετήσιες διακυμάνσεις είναι παρόμοιες με τις πραγματικές και καθόλου ομοιογενείς μεταξύ τους όσο αναφορά τις μικροδιακυμάνσεις. Όπως αναλύεται στην παράγραφο 5.6 όπου συγκρίνονται όλα τα σενάρια, η πρόβλεψη του τρέχοντος σεναρίου είναι εκείνη που δίνει την μεγαλύτερη αύξηση σε κατανάλωση ενέργειας τα επόμενα έτη. Η παρατήρηση αυτή είναι ενδιαφέρουσα καθώς η πρόβλεψη θερμοκρασίας που χρησιμοποιήθηκε για την πρόβλεψη της κατανάλωσης ενέργειας παρουσιάζει μεγάλες αυξήσεις και μεγάλες διακυμάνσεις για τα επόμενα έτη. Έτσι η μεταβολή των θερμοκρασιακών συμπεριφορών συνδέονται άμεσα με την κατανάλωση ενέργειας και την επηρεάζουν σε μεγάλο βαθμό. Με βάση το Γράφημα 5-25, η πρόβλεψη δείχνει πως αυξάνονται οι διακυμάνσεις (αιχμές φορτίου και ελάχιστα) όπως αυξάνονται και οι διακυμάνσεις στην πρόβλεψη θερμοκρασίας που χρησιμοποιήθηκε. Αν και μέχρι το έτος 2027 το δίκτυο φαίνεται να εξάγει δομημένα αποτελέσματα, θα θεωρηθεί ασφαλής πρόβλεψη μέχρι και το έτος 2025, καθώς στην συνέχεια οι διακυμάνσεις μειώνονται αισθητά.



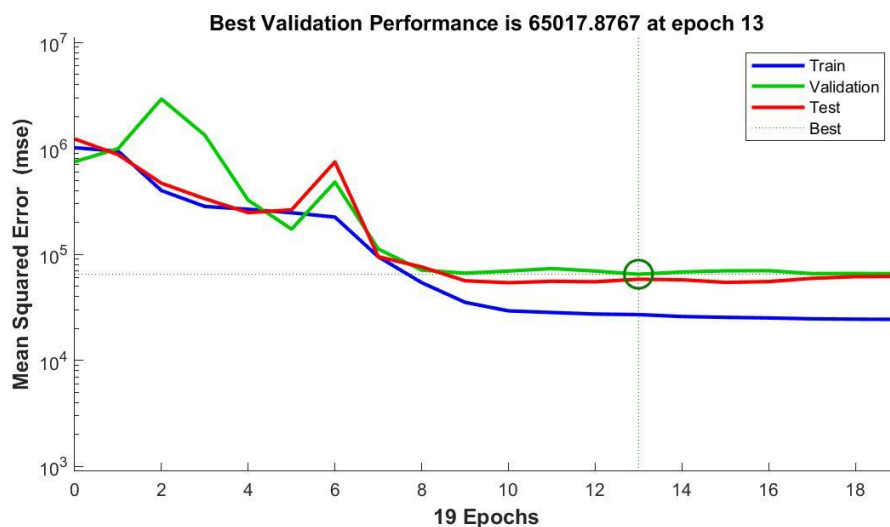
Γράφημα 5-25: Σενάριο 3b, Πρόβλεψη κατανάλωσης ηλεκτρικής ενέργειας

Σενάριο 4

5.4

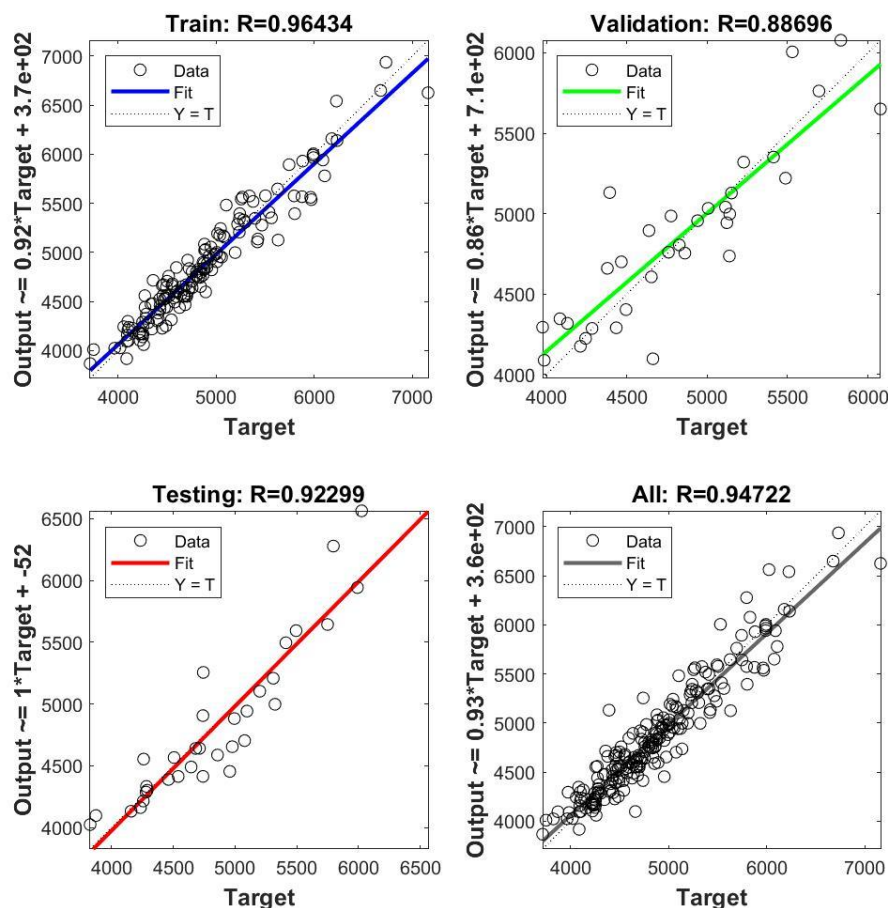
Για την εκπαίδευση του Σεναρίου 4 χρησιμοποιείται η μηνιαία χρονοσειρά κατανάλωσης ενέργειας, όπως στο Σενάριο 1, με την διαφορά ότι εισάγονται στο δίκτυο βοηθητικά δεδομένα τα οποία περιέχουν τον δείκτη του μήνα για κάθε βήμα της βασικής χρονοσειράς.

Όπως φαίνεται στο γράφημα 5-26, τα σφάλματα του δικτύου σταθεροποιούνται μετά τον 8^ο κύκλο εκπαίδευσης, ενώ βέλτιστος κύκλος επιλέγεται ο 13^{ος}, καθώς σε εκείνο το σημείο το μέσο τετραγωνικό σφάλμα των συνόλων ελέγχου και επαλήθευσης σταθεροποιούνται στην ελάχιστη τιμή τους. Παράλληλα, στον 13^ο κύκλο το μέσο τετραγωνικό σφάλμα της εκπαίδευσης παραμένει σε χαμηλά επίπεδα. Όπως συμβαίνει συνήθως, και παρατηρείται και στα υπόλοιπα Σενάρια, τα σφάλματα επαλήθευσης και ελέγχου είναι μεγαλύτερα από εκείνα της εκπαίδευσης. Η διαφορά μεταξύ των δύο αυτών σφαλμάτων με το σφάλμα της εκπαίδευσης είναι μικρότερη σε σχέση με το Σενάριο 2 και 4, κάτι που είναι επιθυμητό και δείχνει καλή απόδοση του δικτύου.



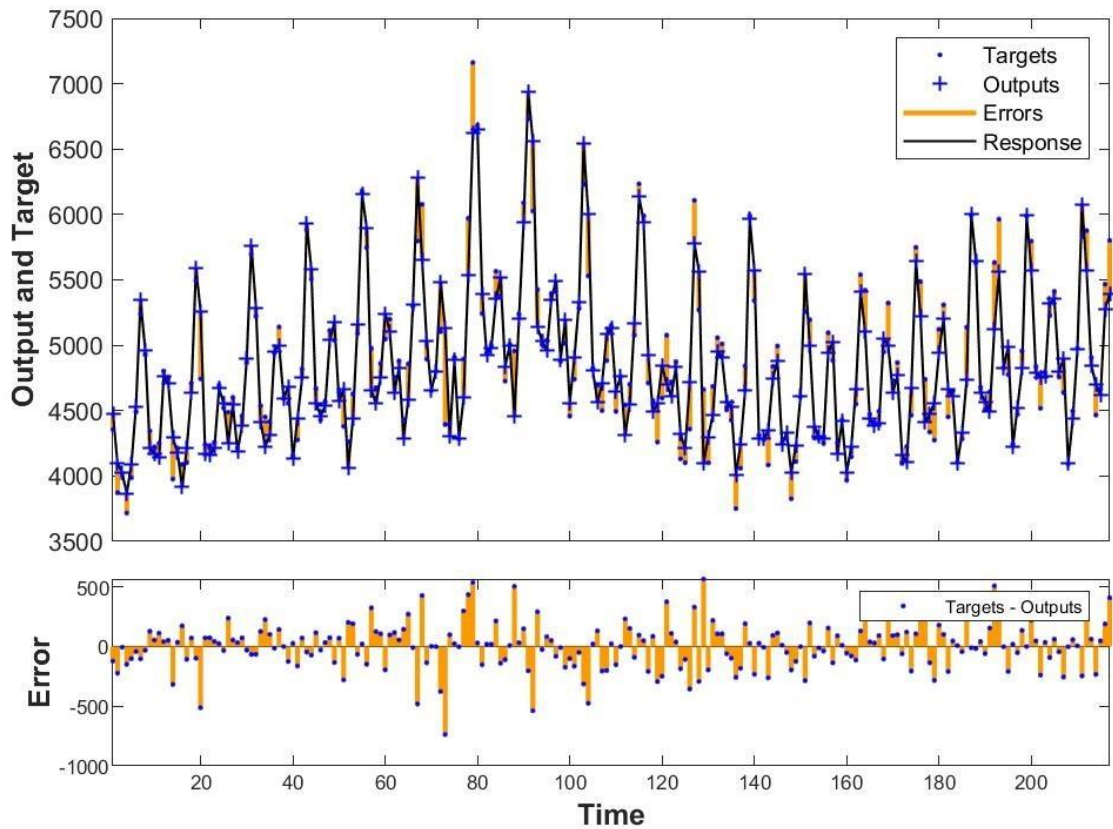
Γράφημα 5-26: Σενάριο 4, Μέσο τετραγωνικό σφάλμα κατά την εκπαίδευση

Κλασικά, ακολουθεί το γράφημα παλινδρομήσεων στο οποίο βλέπουμε πως ο γενικός συντελεστής προσδιορισμού είναι υψηλός ($R^2=0,947$), υψηλότερος από αυτόν του Σεναρίου 1. Αυτό δείχνει πως η βοηθητική χρονοσειρά με τον δείκτη του μήνα βοήθησε όντως το δίκτυο να αυξήσει την ικανότητα του στην ανάλυση και αφομοίωση της μηνιαίας χρονοσειράς κατανάλωσης ηλεκτρικής ενέργειας. Χαρακτηριστικά, οι συντελεστές προσδιορισμού για τα σύνολα δεδομένων ελέγχου και επαλήθευσης προφανώς χαμηλότεροι από εκείνον της εκπαίδευσης, αλλά με μικρότερη απόκλιση από εκείνον σε σχέση με τα υπόλοιπα σενάρια. Επίσης, είναι υψηλότεροι από αυτούς του Σεναρίου 1 με τιμές $R_{val}^2=0,886$, $R_{test}^2=0,922$ και $R_{val}^2=0,790$, $R_{test}^2=0,890$ αντίστοιχα.



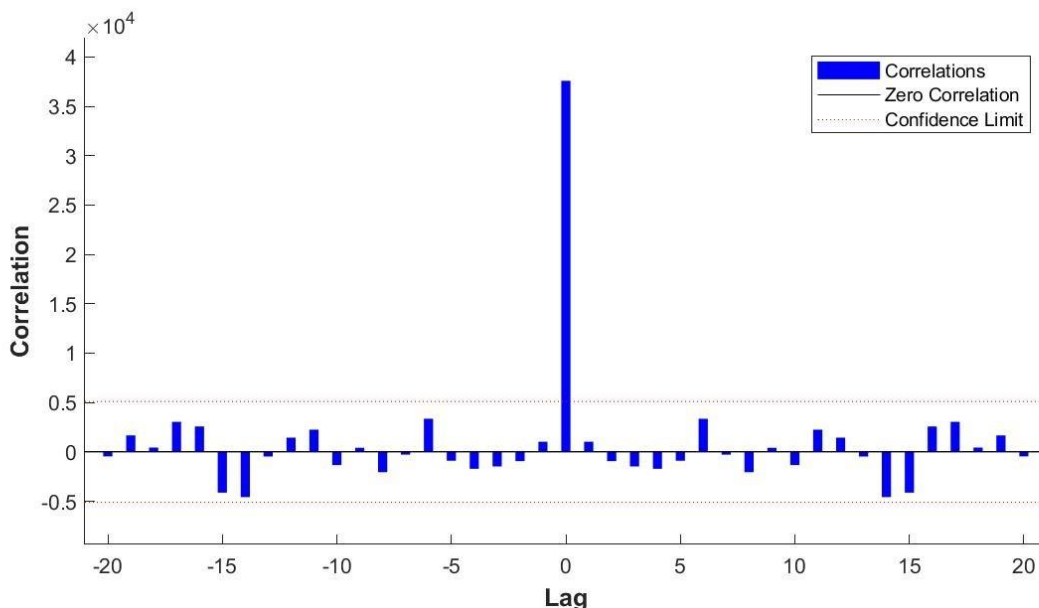
Γράφημα 5-27: Σενάριο 4, Παλινδρόμηση πραγματικών – εκτιμώμενων τιμών

Την καλή συσχέτιση στα δεδομένα παρουσιάζει και το παρακάτω γράφημα (5-28) όπου φαίνεται πως το δίκτυο έχει ελάχιστα σφάλματα στις αιχμές και τα ελάχιστα φορτία, ενώ ακολουθεί σωστά τις τάσεις της χρονοσειράς. Οι περισσότερες αποκλίσεις εμφανίζονται στις μεταβατικές περιόδους των ετήσιων διακυμάνσεων, όπου συμβαίνουν αρκετές μικρές αυξομειώσεις πολλές από τις οποίες οφείλονται στην στοχαστικότητα της χρονοσειράς, οπότε είναι θεμιτό το δίκτυο να μην τις ακολουθεί πιστά. Έτσι, επιβεβαιώνεται η καλή γενίκευση του τεχνητού νευρωνικού δικτύου στα δεδομένα, ενώ αποκλείεται η υπερπροσαρμογή και η πιστή αντιγραφή της χρονοσειράς στα μελλοντικά βήματα.



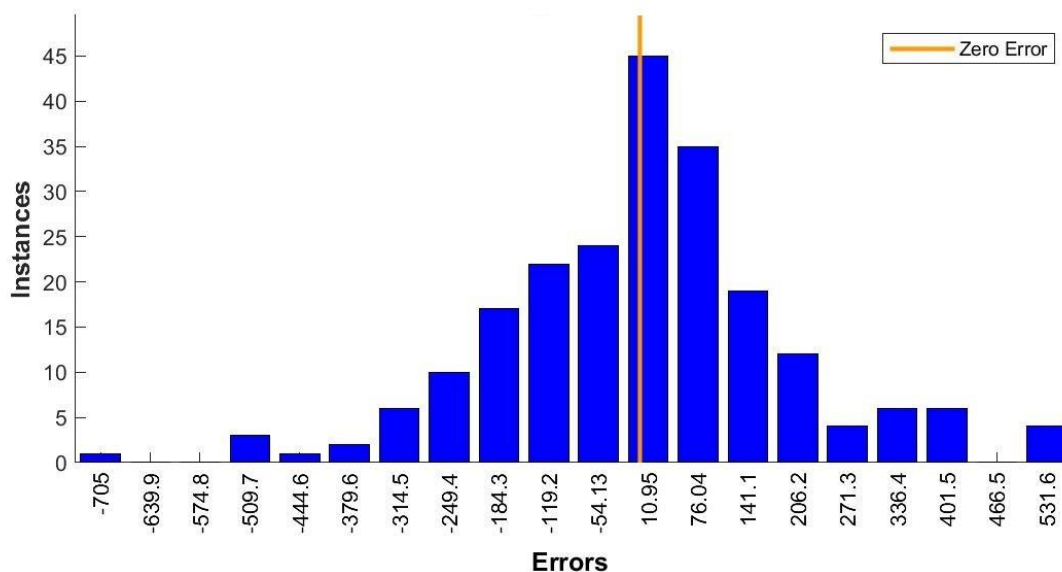
Γράφημα 5-28: Σενάριο 4, Διασπορά πραγματικών- εκτιμώμενων τιμών

Συνεχίζοντας την ανάλυση των αποτελεσμάτων, στο γράφημα 5-29 παρατηρείται πολύ μικρή αυτοσυσχέτιση μεταξύ των σφαλμάτων. Το σύνολό του βρίσκεται εντός του ορίου εμπιστοσύνης, ενώ το μεγαλύτερο σφάλμα είναι ταυτόσημο με εκείνο της εκπαίδευσης και δεν παρουσιάζει καμία περιοδικότητα (Lag=0).



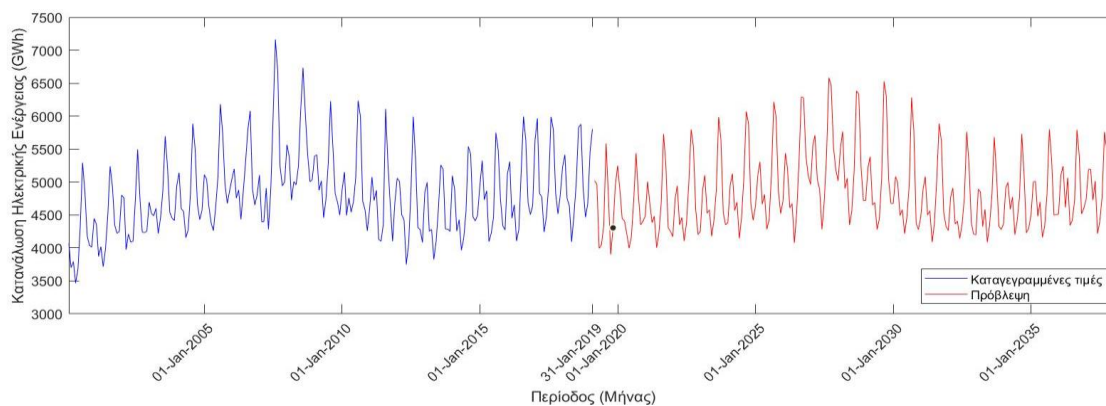
Γράφημα 5-29: Σενάριο 4, Αυτοσυσχέτιση σφαλμάτων

Αντίστοιχα, η κατανομή των σφαλμάτων είναι πολύ κοντά στην κανονική κατανομή, αν και η ράβδος των περισσότερων συχνών σφαλμάτων δεν είναι σχεδόν μηδενική αλλά λαμβάνει τιμή 10,95. Παρ' όλα αυτά, η εκπαίδευση στο σύνολο της φαίνεται αρκετά ικανοποιητική, και αναμένεται και η πρόβλεψη να είναι αντίστοιχη.



Γράφημα 5-30: Σενάριο 4, Κατανομή σφαλμάτων

Συνολικά, η απόδοση της εκπαίδευσης σε αυτό το σενάριο είναι ικανοποιητική και αναμενόμενη, καθώς η τιμή του μήνα προσδίδει στο δίκτυο επιπρόσθετη ποιοτική πληροφορία. Παρ' όλα αυτά, η πρόβλεψη που ακολουθεί στο γράφημα 5-31 δεν είναι ρεαλιστική, καθώς είναι φανερό πως οι προβλεπόμενες τιμές είναι επανάληψη των προηγούμενων τιμών της χρονοσειράς. Ακόμη, φαίνεται πως επαναλαμβάνονται οι μεταβολές στις τάσεις της χρονοσειράς των προηγούμενων ετών, ενώ παράλληλα τάση της πρόβλεψης είναι αυξητική μέχρι το έτος 2028, ενώ στην συνέχεια φθίνει και μειώνονται οι διακυμάνσεις. Κοινό χαρακτηριστικό με τα υπόλοιπα σενάρια, εκτός του Σεναρίου 3, είναι πως η μελλοντική κατανάλωση αυξάνει, αλλά δεν ξεπερνά το μέγιστο του 2008. Η πρόβλεψη μπορεί να θεωρηθεί ασφαλής μέχρι και το έτος 2028, όπου από εκεί και ύστερα η χρονοσειρά αρχίζει να έχει φθίνουσα τάση.



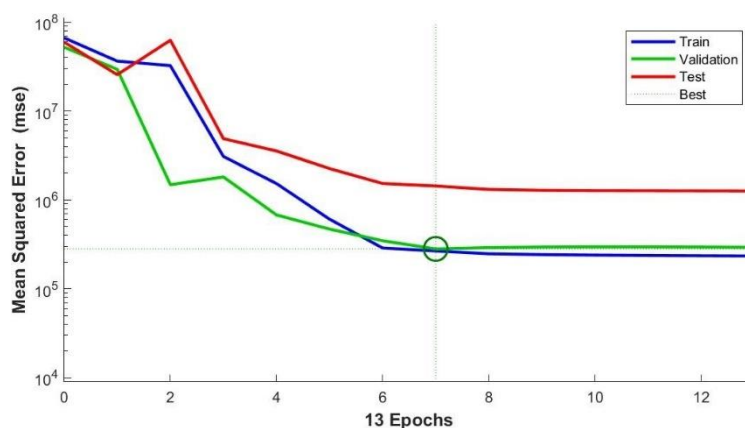
Γράφημα 5-31: Σενάριο 4, Πρόβλεψη κατανάλωσης ηλεκτρικής ενέργειας

Σενάριο 5

Πρόβλεψη ΑΕΠ

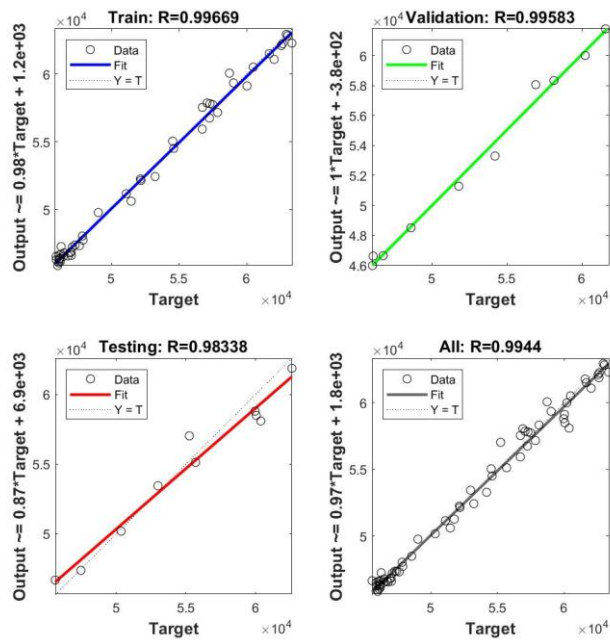
5.5 Για την υλοποίηση του Σεναρίου πραγματοποιείται αρχικά πρόβλεψη του ΑΕΠ της χώρας με την χρήση τεχνητού νευρωνικού δικτύου τύπου NAR. Η πρόβλεψη θα χρησιμοποιηθεί ως είσοδος στο σύνθετο νευρωνικό δίκτυο NARX στη συνέχεια. Τα γραφήματα που προκύπτουν από τα αποτελέσματα της εκπαίδευσης αλλά και η πρόβλεψη του ΑΕΠ παρουσιάζονται και σχολιάζονται παρακάτω.

Η πορεία των σφαλμάτων κατά την διάρκεια της εκπαίδευσης έχει φθίνουσα τάση, όπως αναμένεται, ενώ χαρακτηριστική είναι η σύγκλιση του μέσου τετραγωνικού σφάλματος του υποσυνόλου των δεδομένων εκπαίδευσης με εκείνα της επαλήθευσης. Το μέσο τετραγωνικό σφάλμα των δεδομένων ελέγχου κυμαίνεται σε υψηλότερα επίπεδα. Ο βέλτιστος κύκλος εκπαίδευσης ορίζεται ο 7^{ος}, καθώς και τα τρία σφάλματα λαμβάνουν τις χαμηλότερες τιμές τους. Με βάση το γράφημα 5-32, η πιθανότητα υπερπροσαρμογής είναι μικρή, καθώς όπως παρουσιάζεται το μέσο τυπικό σφάλμα του συνόλου δεδομένων αξιολόγησης ακολουθεί φθίνουσα πορεία, χωρίς να αυξάνει.



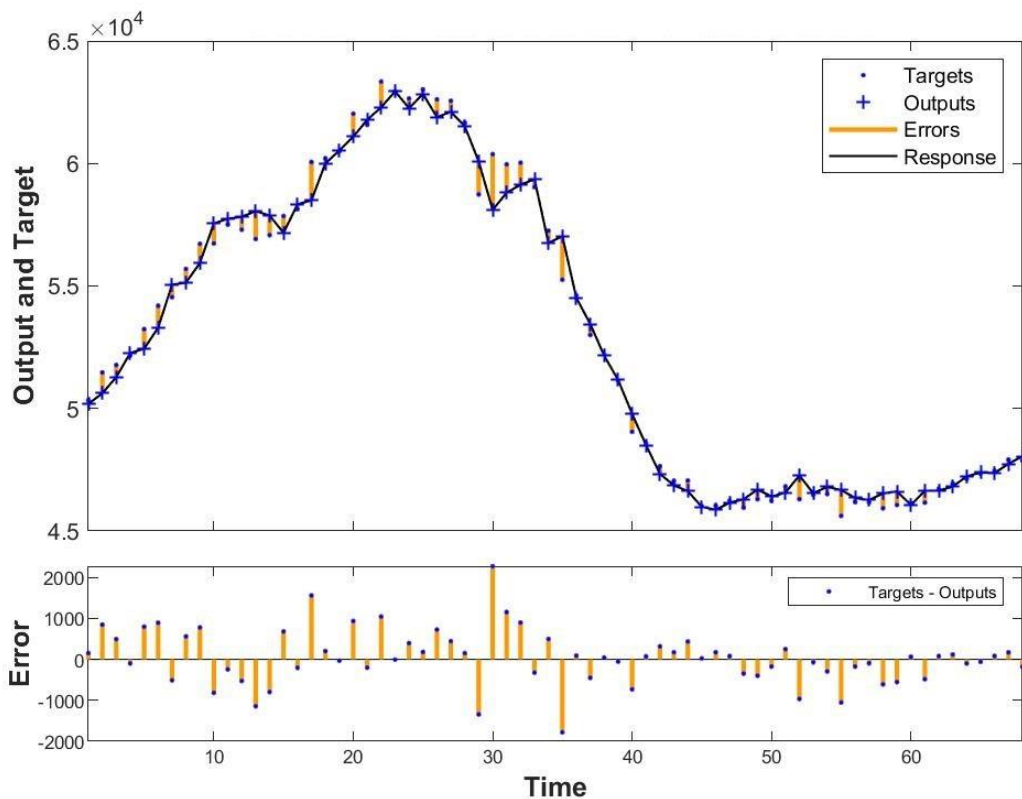
Γράφημα 5-32: Σενάριο 5a, Μέσο τετραγωνικό σφάλμα κατά την εκπαίδευση

Όπως φαίνεται, στο γράφημα 5-33 η διαδικασία της εκπαίδευσης ολοκληρώνεται επιτυχώς με αρκετά υψηλό συντελεστή προσδιορισμού $R^2 = 0,9944$, ενώ αντίστοιχα υψηλοί είναι και οι επιμέρους συντελεστές για τα δεδομένα ελέγχου, εκπαίδευσης και επαλήθευσης. Ο συντελεστής προσδιορισμού είναι σχεδόν 1, πράγμα που δείχνει καλή προσαρμογή στα δεδομένα, αλλά εμπεριέχει και τον κίνδυνο της υπερδιαστασιολόγησης του δικτύου και, κατά συνέπεια, της κακής γενίκευσης επί της τριμηνιαίας χρονοσειράς του ΑΕΠ.



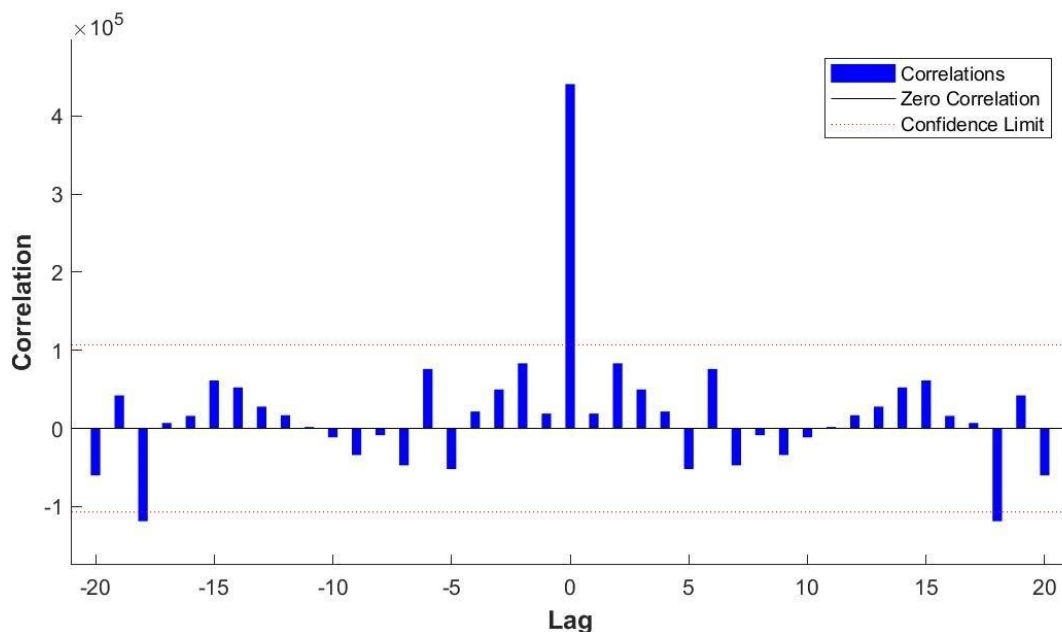
Γράφημα 5-33: Σενάριο 5a, Παλινδρόμηση πραγματικών – εκτιμώμενων τιμών

Στα γράφημα 5-34 παρατηρείται καλή προσαρμογή των εξαγόμενων τιμών του νευρωνικού δικτύου σε σχέση με τις πραγματικές, ενώ τα περισσότερα σφάλματα απαντώνται στα σημεία τα οποία υπάρχει απότομη μεταβολή στις τιμές της χρονοσειράς.



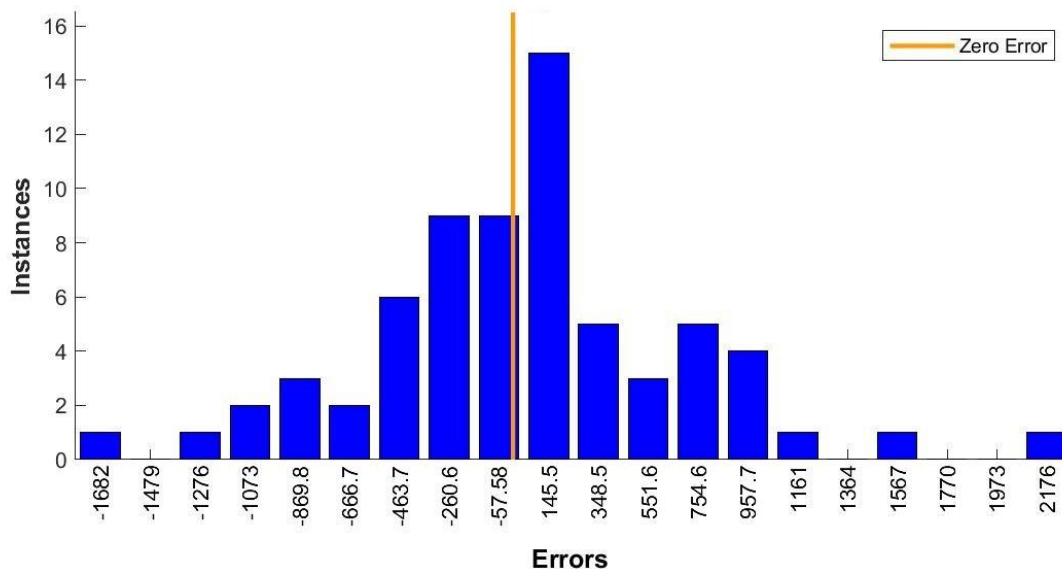
Γράφημα 5-34: : Σενάριο 5a, Διασπορά πραγματικών- εκτιμώμενων τιμών

Η αυτοσυσχέτιση των σφαλμάτων βρίσκεται εντός του ορίου εμπιστοσύνης, με κάποιες μικρές εξαιρέσεις, οι οποίες ξεπερνούν οριακά το όριο και εμφανίζονται με περιοδικότητα 18 βημάτων.



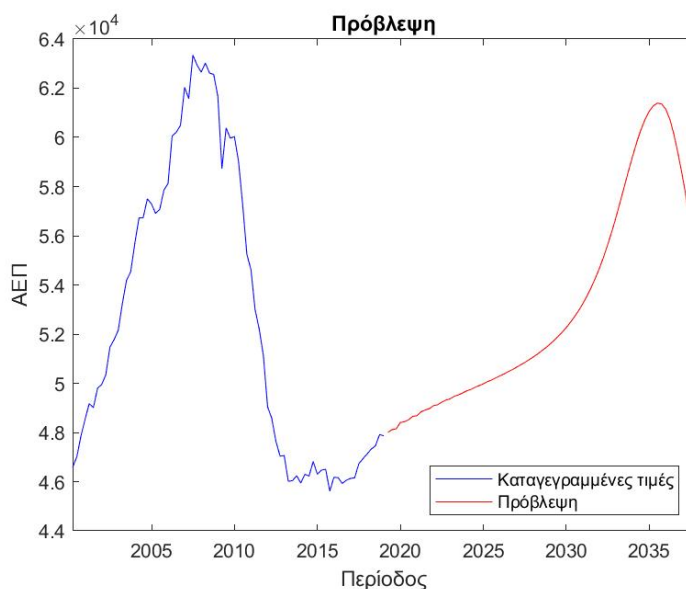
Γράφημα 5-35: Σενάριο 5α, Αυτοσυσχέτιση σφαλμάτων

Η κατανομή των σφαλμάτων δεν είναι συμμετρική, ενώ τα περισσότερα σφάλματα απαντώνται σε μεγάλη απόλυτη τιμή ίση με |57,58|. Η απόδοση του δικτύου είναι αμφίβολη, καθώς παρατηρούνται διάφορες αποκλίσεις στα αποτελέσματα. Στη συνέχεια το δίκτυο θα μετατραπεί σε λειτουργία κλειστού βρόχου και θα πραγματοποιηθεί η πρόβλεψη.



Γράφημα 5-36: Σενάριο 5α, Κατανομή σφαλμάτων

Η πρόβλεψη του δείκτη ΑΕΠ ξεκινάει από την 1/1/2019 και μπορούμε να θεωρήσουμε ότι φτάνει μέχρι και την 31/12/2024 όπου από εκεί και στο εξής η καμπύλη της πρόβλεψης γίνεται εκθετική, ομαλή και δεν φαίνεται να ανταποκρίνεται σε πραγματικές διακυμάνσεις. Είναι αναμενόμενο ο ορίζοντας της πρόβλεψης να είναι μικρότερος, καθώς τα καταγεγραμμένα δεδομένα δεν είναι επαρκή σε αριθμό για την εκπαίδευση του δικτύου, γεγονός που μειώνει την προσαρμοστικότητα του δικτύου για την μελετώμενη χρονοσειρά.

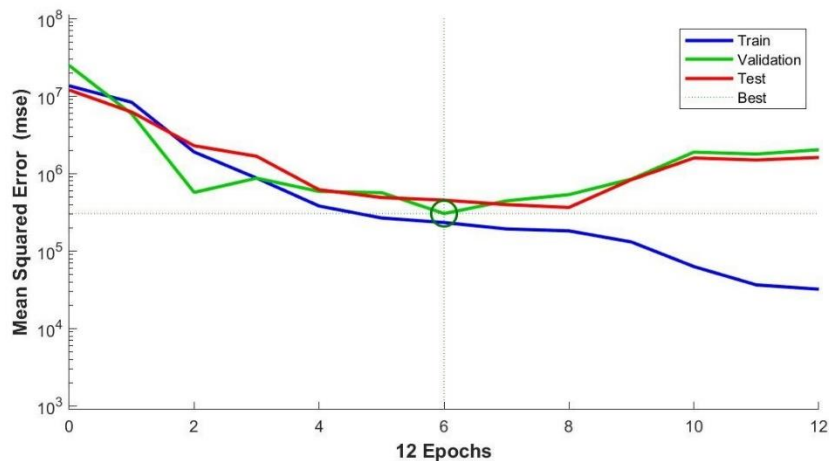


Γράφημα 5-37: Σενάριο 5α, Πρόβλεψη Ακαθάριστου Εγχώριου Προϊόντος

Η πρόβλεψη, όπως φαίνεται στο γράφημα 5-37, αποδεικνύει πως το δίκτυο δεν ανταποκρίνεται σωστά στα χαρακτηριστικά της χρονοσειράς. Βέβαια, παίζει μεγάλο ρόλο η μη περιοδικότητα και οι μεγάλες απρόβλεπτες μεταβολές του ΑΕΠ την περίοδο 2008-2018, οι οποίες οφείλονται στην οικονομική κρίση της τελευταίας δεκαετίας. Η πρόβλεψη μέχρι και το 2024 θα θεωρηθεί επαρκής για να μελετηθεί η επίδραση της στο επόμενο βήμα, που είναι η μοντελοποίηση και η πρόβλεψη της κατανάλωσης ηλεκτρικής ενέργειας με νευρωνικό δίκτυο τύπου NARX.

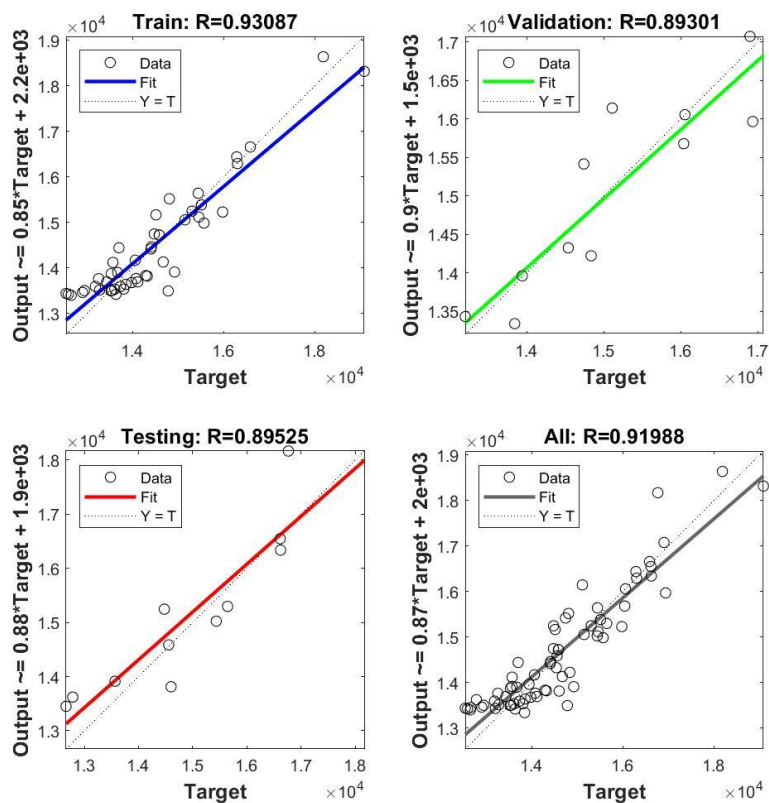
Πρόβλεψη Κατανάλωσης Ενέργειας

Στο γράφημα 5-38 φαίνεται η επαρκής εκπαίδευση του δικτύου, η οποία τερματίστηκε στην 6^η επανάληψη, έτσι ώστε να μην προκληθούν φαινόμενα υπερπροσαρμογής. Τα σφάλμα των τριών συνόλων (εκπαίδευσης, ελέγχου και επαλήθευσης) συγκλίνουν αρκετά, στο σημείο του βέλτιστου κύκλου εκπαίδευσης, με χαμηλότερο το σφάλμα των δεδομένων εκπαίδευσης, όπως είναι αναμενόμενο.



Γράφημα 5-38: Σενάριο 5b, Μέσο τετραγωνικό σφάλμα κατά την εκπαίδευση

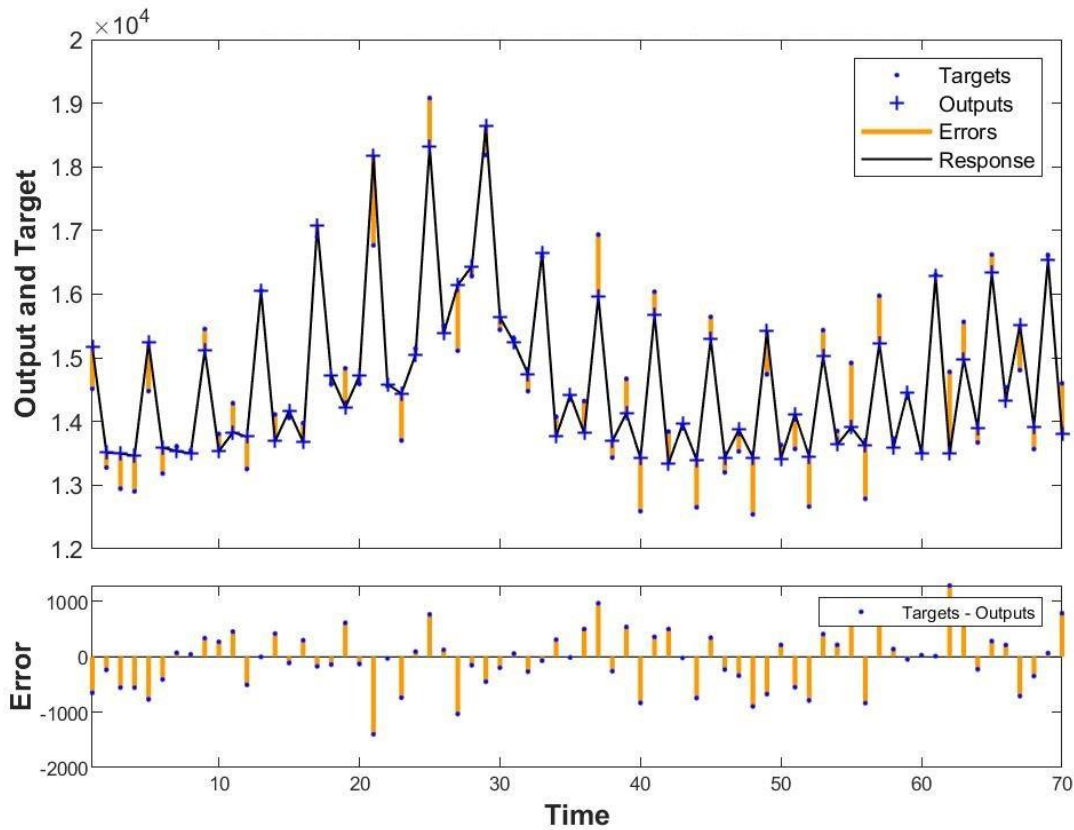
Η συσχέτιση των εξαγομένων δεδομένων με τα πραγματικά, μέσω της παλινδρόμησης, είναι πολύ καλή, όπως φαίνεται στο γράφημα 5-39, με ολικό συντελεστή προσδιορισμού $R^2 = 0,92$. Οι συντελεστές προσδιορισμού για τα επιμέρους σύνολα δεδομένων κυμαίνονται σε επίσης υψηλά επίπεδα.



Γράφημα 5-39: Σενάριο 5b, Παλινδρόμηση πραγματικών – εκτιμώμενων τιμών

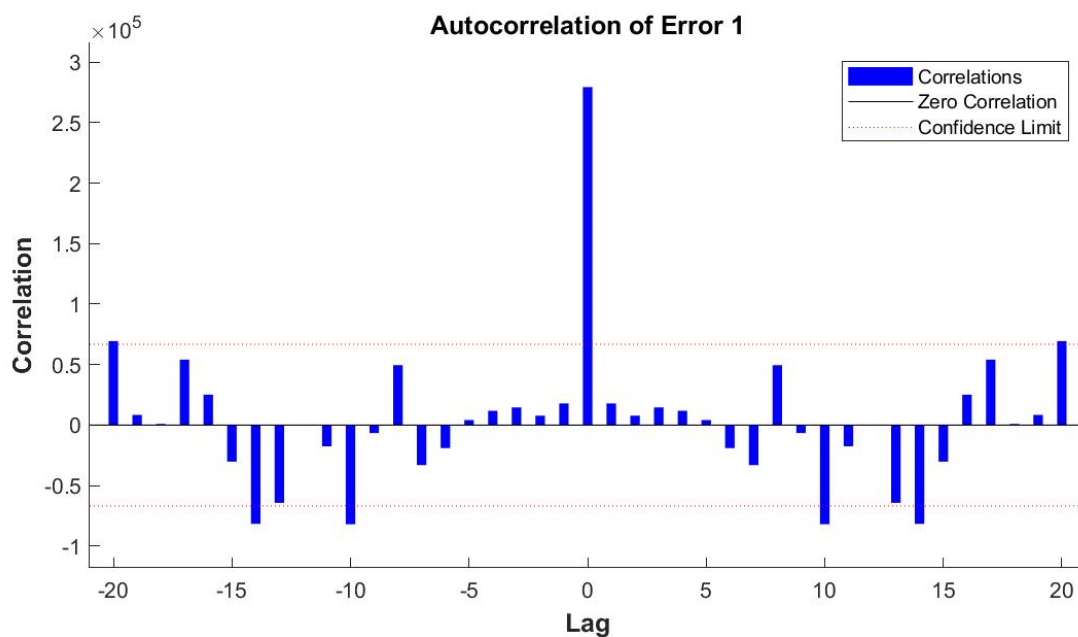
Στο γράφημα διασποράς που ακολουθεί (5-40) φαίνεται πως, αν και το δίκτυο παρουσιάζει υψηλό συντελεστή προσδιορισμού παλινδρόμησης, δεν είναι ικανό να εκτιμήσει σωστά τις αιχμές και τα

ελάχιστα της κατανάλωσης. Παρατηρείται πως υπάρχουν αρκετά σφάλματα σε όλο το εύρος της χρονοσειράς, με κυριότερα εκείνα που βρίσκονται στα ακραία σημεία.



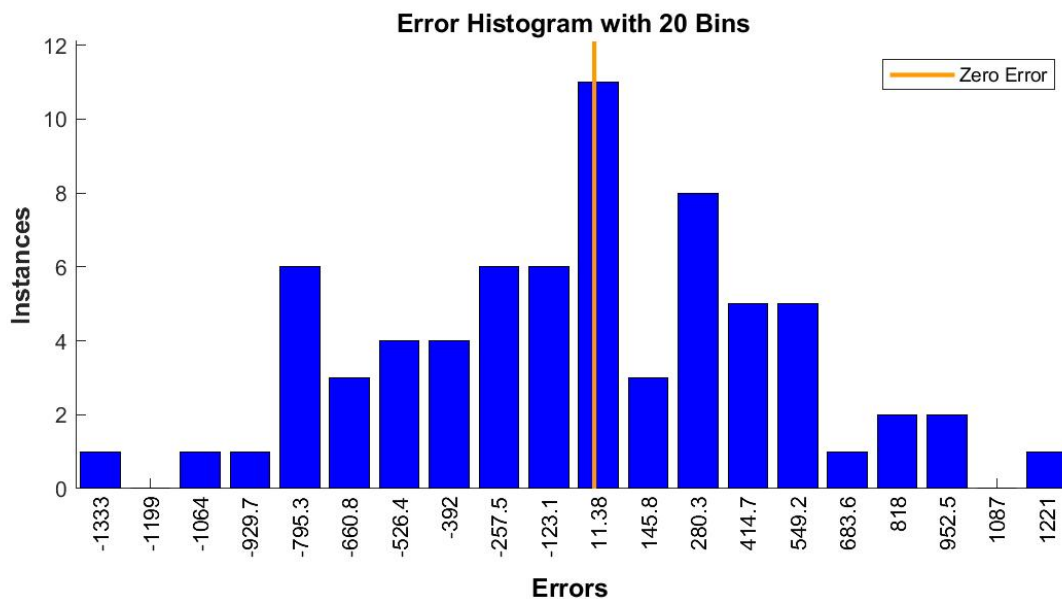
Γράφημα 5-40: Σενάριο 5b, Διασπορά πραγματικών- εκτιμώμενων τιμών

Όσο αναφορά την αυτοσυσχέτιση, υπάρχουν αρκετά επαναλαμβανόμενα σφάλματα τα οποία ξεπερνούν το όριο εμπιστοσύνης και έχουν περιοδικότητα 10 και 16 βήματα (τρίμηνα).



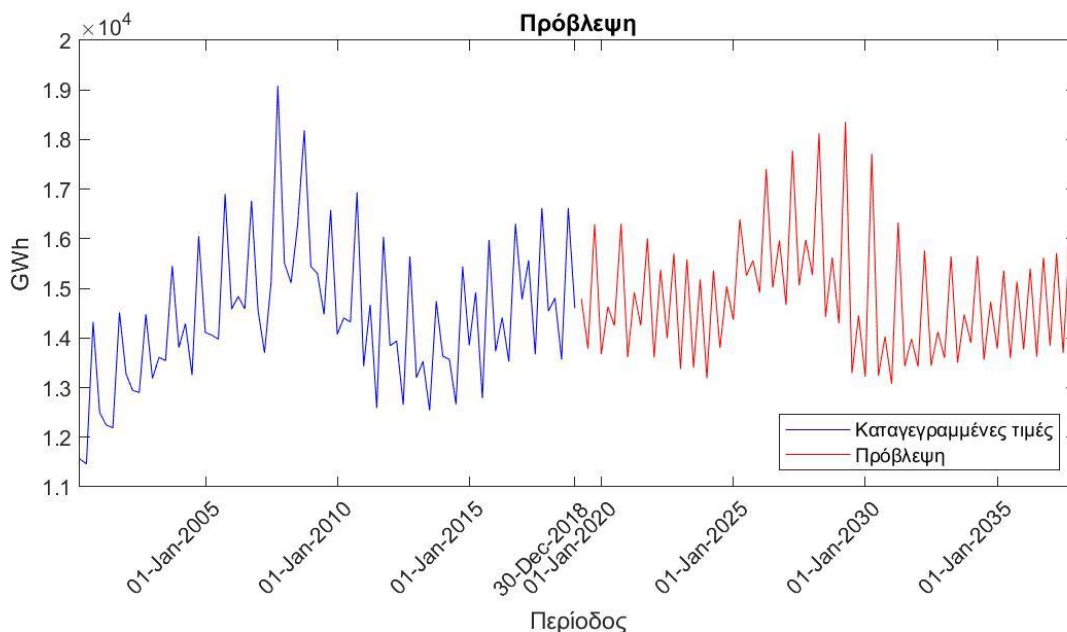
Γράφημα 5-41: Σενάριο 5b, Αυτοσυσχέτιση σφαλμάτων

Η κατανομή των σφαλμάτων, επίσης, δεν είναι ικανοποιητική καθώς αποκλίνει αρκετά από την κανονική, ενώ απαντώνται αρκετά σφάλματα σε διάφορες υψηλές τιμές. Επίσης, το συχνότερο σφάλμα λαμβάνει τιμές μεγαλύτερες του μηδενός (11,38).



Γράφημα 5-42: Σενάριο 5b, Κατανομή σφαλμάτων

Η πρόβλεψη δεν φαίνεται να είναι ικανοποιητική. Μέχρι και το έτος 2024 κυμαίνεται σε αποδεκτές τιμές, αν και ακολουθεί πτωτική πορεία. Χρειάζεται περαιτέρω εξέταση και σύγκριση με πραγματικές τιμές, εάν υπάρχουν, ή με τα αποτελέσματα των υπόλοιπων σεναρίων έτσι ώστε να επαληθευτεί η εγκυρότητα της πρόβλεψης. Το χρονικό εύρος για το οποίο η χρονοσειρά πρόβλεψης κρίνεται ρεαλιστική, ταυτίζεται με το χρονικό εύρος στο οποίο η πρόβλεψη του ΑΕΠ ήταν ρεαλιστική επίσης, γεγονός που είναι αναμενόμενο καθώς η βοηθητική χρονοσειρά (ΑΕΠ) επηρεάζει άμεσα την κύρια, σε ένα νευρωνικό δίκτυο τύπου NARX.



Γράφημα 5-43: Σενάριο 5b, Πρόβλεψη κατανάλωσης ηλεκτρικής ενέργειας

Σχολιασμός Αποτελεσμάτων & Προτάσεις για Μελλοντική Έρευνα

6. Στο τελευταίο κεφάλαιο σχολιάζονται οι προβλέψεις που πραγματοποιήθηκαν στα προηγούμενα βήματα της παρούσας μελέτης. Αναλύεται η επίδραση των εξωγενών παραγόντων, για τα σενάρια που εφαρμόστηκε νευρωνικό δίκτυο τύπου NARX, ενώ αξιολογείται η εγκυρότητα της πρόβλεψης με γνώμονα τις πραγματικές διακυμάνσεις, τάσεις και χαρακτηριστικά που θα παρουσίαζε μια πραγματική καμπύλη φορτίου. Συν τοις άλλοις, γίνεται σύγκριση μεταξύ των προβλέψεων της μελέτης με την πρόβλεψη της μελέτης επάρκειας φορτίου του Ανεξάρτητου Διαχειριστή Μεταφοράς Ηλεκτρικής Ενέργειας. Επιπλέον, εντοπίζονται τα αδύναμα σημεία της παρούσας μελέτης και με βάση αυτά, αναπτύσσονται προτάσεις για μελλοντική έρευνα και βελτιστοποίηση των νευρωνικών δικτύων αλλά και των προβλέψεων αντίστοιχα.

Γενικά συμπεράσματα & Σύγκριση προβλέψεων σεναρίων

- 6.1 Παρακάτω ακολουθεί σύγκριση σεναρίων, συγκρίνοντας τους συντελεστές προσδιορισμού R^2 και τον ορίζοντα πρόβλεψης που θεωρείται έμπιστος. Όπως παρατηρείται, και σχολιάζεται σε παραπάνω παραγράφους, όλα τα σενάρια παρουσιάζουν αρκετά υψηλούς συντελεστές προσδιορισμού κατά την εκπαίδευσή τους, ενώ ο εκτιμώμενος ορίζοντας πρόβλεψης δεν διαφέρει δραστικά μεταξύ των σεναρίων, αλλά φαίνεται να επηρεάζεται με τον αριθμό των βημάτων των χρονοσειρών εκπαίδευσης.

A/A	Χρησιμοποιούμενα Δεδομένα	Συντελεστής R	Εκτιμώμενος ορίζοντας πρόβλεψης
Σενάριο 1	Μηνιαία κατανάλωση από 01/2000 έως και 01/2019	0,910	Ιαν 2030
Σενάριο 2	Ετήσια κατανάλωση 1960-2014	0,997	2030
Σενάριο 3	Μηνιαία κατανάλωση & θερμοκρασία (περίοδο)	0,856	2025
Σενάριο 4	Μηνιαία κατανάλωση & δείκτης μήνα	0,947	Ιαν 2030
Σενάριο 5	Μηνιαία κατανάλωση & ΑΕΠ	0,912	Ιαν 2024

Πίνακας 6-1: Συνοπτική περιγραφή σεναρίων και ορίζοντα πρόβλεψης

Στη συνέχεια ακολουθεί συμπεριληπτικό γράφημα όλων των προβλέψεων, έτσι ώστε να γίνει εποπτική ανασκόπηση των αποκλίσεων της πρόβλεψης μεταξύ των σεναρίων. Για την κατασκευή του διαγράμματος έγινε επεξεργασία των εξαγόμενων προβλέψεων έτσι ώστε να έχουν το ίδιο βήμα, καθώς μερικές από αυτές ήταν μηνιαίες και άλλες ετήσιες, αναλόγως την χρονοσειρά που χρησιμοποιήθηκε για την κατασκευή του μοντέλου. Παράλληλα, καθώς η ετήσια χρονοσειρά είναι

καταγεγραμμένη ως κατά κεφαλήν κατανάλωση ηλεκτρικής ενέργειας (kWh/κατά κεφαλήν) , έγινε μετατροπή σε GWh.

Με σημείο αναφοράς την μελέτη επάρκειας ισχύος του ΑΔΜΗΕ 2017-20127, όλα τα σενάρια της παρούσας μελέτης προβλέπουν αρκετά μεγαλύτερες απαιτήσεις σε ηλεκτρική ενέργεια.

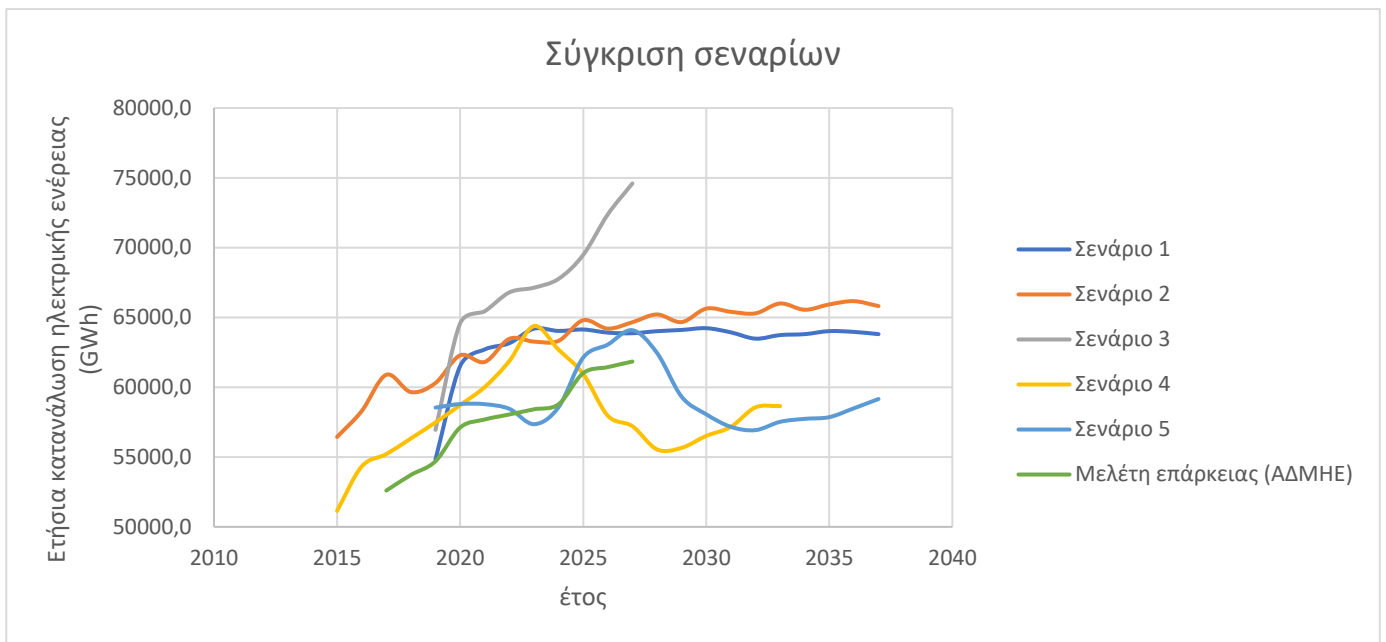
Στον παρακάτω πίνακα παρουσιάζεται η μέση απόκλιση σε σχέση με την μελέτη επάρκειας, για κάθε σενάριο.

Σενάριο 1	6,30%
Σενάριο 2	7,45%
Σενάριο 3	14,24%
Σενάριο 4	-1,91%
Σενάριο 5	2,05%

Πίνακας 6-2: Απόκλιση Σεναρίων από την μελέτη επάρκειας ΑΔΜΗΕ.

Με βάση την απόκλιση, κοντινότερα στην μελέτη επάρκειας είναι το Σενάριο 4, στο οποίο χρησιμοποιήθηκε η μηνιαία χρονοσειρά κατανάλωσης ενέργειας από την Eurostat, ενώ παράλληλα εισήχθη ο δείκτης κάθε μήνα για τον εμπλουτισμό της διαδικασίας εκπαίδευσης με περισσότερα δεδομένα. Ενώ η απόκλιση είναι μικρή, η καμπύλη της πρόβλεψης δεν ακολουθεί παρόμοια πορεία με αυτήν της μελέτης επάρκειας. Οι δύο καμπύλες μεταξύ τους διαφοροποιούνται αρκετά, με σημαντικότερη διαφοροποίηση την φθίνουσα πορεία του Σεναρίου 4 από το 2023 και έπειτα, σε αντίθεση με την μελέτη επάρκειας που ακολουθεί αύξουσα πορεία σε όλο τον ορίζοντα πρόβλεψης. Η πρόβλεψη του Σεναρίου 4 δεν μπορεί να θεωρηθεί αξιόπιστη διότι είναι σχετικά απίθανο η κατανάλωση ηλεκτρικής ενέργειας να φθίνει σε μακροπρόθεσμο ορίζοντα. Μια τέτοια περίπτωση θα πραγματοποιούνταν σε ακραίες συνθήκες οικονομικής κρίσης, ενεργειακής φτώχειας ή άλλου καταστροφικού συμβάντος, όπως συνέβη με την οικονομική κρίση του 2008.

Το σενάριο 5, επίσης έχει μικρή απόκλιση από την μελέτη επάρκειας, ενώ ταυτόχρονα οι δύο καμπύλες παρουσιάζουν μια ομοιότητα μεταξύ τους. Ενώ η πρόβλεψη του σεναρίου 5 για το 2020 βρίσκεται σε υψηλότερα επίπεδα από την μελέτη επάρκειας, καθώς ο αυξάνει ο ορίζοντας πρόβλεψης, οι καμπύλες συγκλίνουν μέχρι το έτος 2022 και στη συνέχεια αποκλίνουν, δημιουργώντας ένα τοπικό ελάχιστο στα έτη 2023-2024. Παρόμοια πορεία ακολουθούν και στην συνέχεια, συνεχίζοντας σχεδόν ταυτόσημα έως ότου αποκλίνουν στο έτος 2027 κατά 4% , όπου τελειώνει η μελέτη επάρκειας ισχύος.



Γράφημα 6-1 Σύγκριση προβλέψεων διαφορετικών σεναρίων

Τα σεσάρια 1 και 2 έχουν μικρή απόκλιση της τάξεως του 10% μεταξύ τους και ακολουθούν παρόμοια πορεία στο μεγαλύτερο εύρος της πρόβλεψης, από το 2020 μέχρι και το 2037. Στα δύο αυτά σεσάρια, αν και έχουν παρόμοιες προβλέψεις, χρησιμοποιήθηκαν διαφορετικές ιστορικές χρονοσειρές κατανάλωσης ηλεκτρικής ενέργειας, κατά την εκπαίδευσή τους. Στο 1 χρησιμοποιήθηκε η μηνιαία χρονοσειρά της Eurostat και στο 2 η ετήσια χρονοσειρά του World Bank. Τα σεσάρια αυτά παράγουν μία πιο σταθερή πρόβλεψη, με μικρές διακυμάνσεις και αύξουσα καμπύλη. Κοινό τους χαρακτηριστικό είναι ότι είναι τα μόνα σεσάρια νευρωνικών δικτύων τύπου NAR στα οποία η πρόβλεψη βασίστηκε μόνο στις παλαιότερες καταγραφές κατανάλωσης. Αν και κοντινά μεταξύ τους, απέχουν κατά μέσο όρο $\approx 7\%$ παραπάνω από την πρόβλεψη του ΑΔΜΗΕ.

Ως συμπέρασμα της ομοιότητας των παραπάνω δύο σεναρίων και της διαφοροποίησης τους από τα υπόλοιπα, εξάγεται ότι οι βοηθητικές χρονοσειρές για τις οποίες εκπαιδεύτηκαν δίκτυα τύπου NARX επιδρούν αισθητά στην παραγόμενη πρόβλεψη. Έτσι, η εγκυρότητα της πρόβλεψης επηρεάζεται ισχυρά από την πρότερη πρόβλεψη που απαιτείται για την βοηθητική χρονοσειρά, αυξάνοντας ενδεχομένως το τελικό σφάλμα. Με αυτόν τον τρόπο, όμως, είναι ευκολότερη η ανάλυση της επίδρασης του κάθε εξωγενή παράγοντα και η εξαγωγή συμπερασμάτων. Π.χ. το γεγονός ότι το σεσάριο 5 είναι εκείνο που κυμαίνεται πιο κοντά στα αποτελέσματα της μελέτης του ΑΔΜΗΕ μας δείχνει ότι η επίδραση του ΑΕΠ στο νευρωνικό δίκτυο και στην τελική πρόβλεψη είναι καταλυτική, καθώς το μοντέλο για την μελέτη επάρκειας επίσης λαμβάνει υπόψιν οικονομικά στοιχεία για την ενέργεια.

Σε αυτό το σημείο αξίζει να αναφερθεί ξανά η επιτυχία του Σεναρίου 2 κατά την επαλήθευση των προβλέψεων του με πραγματικές τιμές και όχι με άλλη μελέτη, που αναλύθηκε στην παράγραφο 5.5. Η επαλήθευση ήταν εφικτή καθώς οι καταγραφές της χρονοσειράς από το World Bank τελειώνουν το 2014, όποτε τα έτη 2015-2018 αποτέλεσαν πρόβλεψη για το νευρωνικό δίκτυο. Η απόκλιση μεταξύ των πραγματικών και των προβλεπόμενων τιμών ήταν ελάχιστη, ενώ οι καμπύλες παρουσιάζουν ακριβώς τις ίδιες διακυμάνσεις. Με αυτό το δεδομένο συμπεραίνουμε ότι η πρόβλεψη για το 2017 και 2018 του ΑΔΜΗΕ, απείχε αρκετά από την πραγματική κατανάλωση η οποία εκτιμήθηκε επιτυχώς από το νευρωνικό δίκτυο του σεναρίου 2.

Το σεσάριο με την μεγαλύτερη αύξηση κατανάλωσης ηλεκτρικής ενέργειας για τα επόμενα έτη είναι το σεσάριο 3, στο οποίο, όπως περιγράφεται στην παράγραφο 5.3 χρησιμοποιείται δυστοπική

πρόβλεψη θερμοκρασίας με μεγάλη αύξηση για τα επόμενα χρόνια. Όπως παρουσιάζεται στο γράφημα 5-44 η αύξηση της κατανάλωσης (ή των απαιτήσεων) σε ηλεκτρική ενέργεια αυξάνονται κατά πολύ σε περίπτωση μεγάλων θερμοκρασιακών μεταβολών. Με βάση την πρόβλεψη η κατανάλωση θα αυξηθεί κατά 22% από το 2019 μέχρι το 2025, ενώ θα συνεχίσει να αυξάνεται. Η πρόβλεψη δείχνει καλή απόκριση του νευρωνικού δικτύου στον εξωγενή παράγοντα της θερμοκρασίας, χωρίς όμως να αποτελεί ρεαλιστικό σενάριο.

Προτάσεις για μελλοντική έρευνα

6.2 Τα αποτελέσματα της μελέτης θεωρείται ικανοποιητικό, καθώς ο σκοπός είναι να διερευνηθεί το κατά πόσο τα τεχνητά νευρωνικά δίκτυα είναι σε θέση να επιτύχουν μακροπρόθεσμη κατανάλωση ηλεκτρικής ενέργειας. Στο πλαίσιο, βέβαια, που αυτό είναι εφικτό με βάση τα διαθέσιμα καταγεγραμμένα ιστορικά δεδομένα για τον ελληνικό χώρο. Έτσι, πραγματοποιήθηκε η μελέτη και σχολιάστηκαν τα αποτελέσματα για κάθε έναν από τους διαφορετικούς τύπους δεδομένων και καταγεγραμμένων μετρήσεων που χρησιμοποιήθηκαν για την εκπαίδευση των Τεχνητών Νευρωνικών Δικτύων και την πραγματοποίηση της πρόβλεψης.

Τα συμπεράσματα των αποτελεσμάτων και των εξαγόμενων προβλέψεων αναλύονται στην παράγραφο 6, ενώ ακολουθεί ανάλυση και προτάσεις για περαιτέρω έρευνα, ανάπτυξη και βελτίωση των αποτελεσμάτων. Αρχικά θα αναλυθεί για ποιόν λόγο, οι εκπαιδευτές νευρωνικών δικτύων με την χρήση τρίτης χρονοσειράς παρουσίασαν χαμηλότερη ικανότητα προσαρμογής και γενίκευσης, σε σχέση με τα 2 σενάρια στα οποία χρησιμοποιήθηκε μόνο η χρονοσειρά κατανάλωσης ηλεκτρικής ενέργειας.

Στις περιπτώσεις που χρησιμοποιείται πρόβλεψη τρίτων δεδομένων, όπως η θερμοκρασία, το ΑΕΠ ή ο πληθυσμός της χώρας, μια λύση θα ήταν να πραγματοποιείται πρόβλεψη των βοηθητικών χρονοσειρών με μοντέλα που εξειδικεύονται στα αντίστοιχα δεδομένα. Στην παρούσα εργασία, έγινε πρόβλεψη της θερμοκρασίας με την χρήση παρόμοιου νευρωνικού δικτύου με αυτό που χρησιμοποιήθηκε για την εκπαίδευση της κύριας χρονοσειράς. Προφανώς έγιναν δοκιμές με διαφορετικές παραμέτρους στα δευτερεύοντα δεδομένα για την καλύτερη γενίκευση του δικτύου, αλλά δεν παύει το εξαγόμενο αποτέλεσμα, το οποίο χρησιμοποιείται ως είσοδος στο κύριο νευρωνικό να αποτελεί πρόβλεψη, γεγονός που εμπεριέχει μεγάλη πιθανότητα σφαλμάτων και αποκλίσεων. Κύριο συμπέρασμα από την συμπεριφορά των νευρωνικών τύπου NARX σε αυτές τις δοκιμές είναι ότι ίσως να μην είναι κατάλληλα για αυτού του είδους την εφαρμογή. Π.χ. μια τυπική χρήση νευρωνικών τύπου NARX είναι η πρόβλεψη της θέσης ενός ηλεκτρομαγνήτη γνωρίζοντας τις προηγούμενες θέσεις του (σε μορφή χρονοσειράς για την εκπαίδευση του νευρωνικού), και εισάγοντας ως ανεξάρτητη μεταβλητή το ρεύμα ελέγχου που εισάγεται στον ηλεκτρομαγνήτη. Έτσι, το νευρωνικό προβλέπει τις επόμενες θέσεις που θα λάβει ο αιωρούμενος μαγνήτης. Στην περίπτωση αυτή η πρόβλεψη είναι βραχυπρόθεσμη, ενώ η ανεξάρτητη μεταβλητή αποτελεί πραγματική μέτρηση μιας παραμέτρου που επηρεάζει το μέγεθος που μας ενδιαφέρει. Στα πλαίσια της παρούσας εργασίας, η πρόβλεψη ήταν μεσο- μακροπρόθεσμη, ενώ η ανεξάρτητη μεταβλητή που χρησιμοποιήθηκε ως είσοδος ήταν ήδη ένα εξαγόμενο αποτέλεσμα.

Το παραπάνω συμπέρασμα ίσως αποτελέσει τροχοπέδη στην χρήση νευρωνικών δικτύων με τον συγκεκριμένο τρόπο, αλλά παράλληλα αποτελεί εμπνευση για διαφορετικού τύπου μελέτες. Π.χ. θα μπορούσε να γίνει εφαρμογή νευρωνικών δικτύων για την δημιουργία σεναρίων για διαφορετικές προβλέψεις θερμοκρασιών. Δηλαδή να σχεδιαστούν σενάρια μεταβολής της θερμοκρασίας στην μελετώμενη περιοχή, τα οποία να αντικατοπτρίζουν την μεταβολή του μικροκλίματος και του θερμοκρασιακού εύρους λόγω της κλιματικής αλλαγής. Τα αποτελέσματα αυτών θα εισαχθούν σε νευρωνικό δίκτυο, παρόμοιο με αυτό που χρησιμοποιήθηκε στην παρούσα εργασία, έτσι ώστε να μελετηθεί η μεταβολή και οι επιπτώσεις στην κατανάλωση ηλεκτρικής ενέργειας για κάθε διαφορετικό σενάριο μεταβολής θερμοκρασίας.

Ακόμη μια ιδέα η οποία θα μπορούσε να βελτιώσει τα αποτελέσματα των νευρωνικών δικτύων, και να δώσει μια πιο ολοκληρωμένο μοντέλο των απαιτήσεων σε ενέργεια βασιζόμενων σε εξωτερικούς παράγοντες θα ήταν να χρησιμοποιηθούν σε κατάλληλη δομή νευρωνικών δικτύων όλες οι βοηθητικές χρονοσειρές εξωγενών παραγόντων μαζί. Έτσι οι εξωγενείς παράγοντες θα δρούσαν συνεργιστικά εντός του δικτύου, λαμβάνοντας ταυτόχρονα υπόψιν και τα ιστορικά στοιχεία κατανάλωσης και θα έδιναν μια πιο ρεαλιστική πρόβλεψη. Προϋπόθεση σε αυτή την περίπτωση είναι η ασφαλής πρόβλεψη των εξωγενών παραγόντων, διαδικασία η οποία είναι αρκετά επιρρεπής σε σφάλματα.

Με βάση παλαιότερες μελέτες, η μεγαλύτερη κατηγοριοποίηση των δεδομένων βοηθά τα νευρωνικά δίκτυα να προσαρμοστούν καλύτερα στα προς επεξεργασία δεδομένα. Εάν τα διαθέσιμα δεδομένα είχαν μικρότερο βήμα και εξίσου μεγάλο εύρος καταγραφής, τότε θα μπορούσαν να παρέχουν περισσότερη πληροφορία. Π.χ. εάν κατείχαμε ημερήσια χρονοσειρά κατανάλωσης ηλεκτρικής ενέργειας θα μπορούσε να γίνει διαχωρισμός καθημερινών με αργίες κ.ο.κ. Σε αυτή την περίπτωση θα έπρεπε να ερευνηθεί η ικανότητα πρόβλεψης σε μακροπρόθεσμο ορίζοντα. Οι περισσότερες έως τώρα μελέτες εφαρμόζουν νευρωνικά δίκτυα για την πρόβλεψη σε βραχυπρόθεσμο και πολύ βραχυπρόθεσμο ορίζοντα. στα πλαίσια της παρούσας έγινε μελέτη για τον αν είναι ένα δίκτυο ικανή να παράγει μεσοπρόθεσμη και μακροπρόθεσμη πρόβλεψη με βάση ιστορικές καταγραφές.

Η έρευνα θα μπορούσε να συνεχιστεί χρησιμοποιώντας νέες μετρήσεις, οι οποίες είναι ικανές να επαληθεύσουν τα μοντέλα και να δώσουν μια κατευθυντήρια γραμμή για την παραμετροποίηση τους και την βελτίωση των αποτελεσμάτων. Αυτό θα μπορούσε να πραγματοποιηθεί και με υπάρχουσες μετρήσεις, αλλά κρίθηκε προτιμότερο να χρησιμοποιηθούν όσο το δυνατόν περισσότερα δεδομένα κατά την εκπαίδευση, καθώς ήδη το πλήθος ήταν μικρό. Στην περίπτωση του Σεναρίου 2, εφόσον υπήρχαν διαθέσιμα πραγματικά δεδομένα από άλλη πηγή, έγινε σύγκριση για την αξιολόγηση του μοντέλου. Όπως αναλύεται και στην παράγραφο 5.2., τα αποτελέσματα ήταν ιδιαίτερος ικανοποιητικά.

Τα δίκτυα τα οποία σχεδιάστηκαν στα πλαίσια της εργασίας μπορούν να χρησιμοποιηθούν σε μελλοντικές μελέτες σεναρίων με υποθέσεις ακραίων μεταβολών των εξωγενών παραγόντων (θερμοκρασία, πληθυσμός, ΑΕΠ). Έτσι είναι εφικτό να μελετηθεί η επίδραση της αύξησης της θερμοκρασίας ή η συσχέτιση των οικονομικών μιας χώρας στις μελλοντικές ενεργειακές απαιτήσεις σε ηλεκτρισμό και κατά πόσο μπορεί π.χ. η αύξηση του πληθυσμού να τις μεταβάλλει.

Παράρτημα - Κώδικες Matlab Live Script

Σενάριο 1

```
%NAR TEST 1
%εισαγωγή μεταβλητών
clear all
load ('Tvariable.mat','T')
%εκπαίδευση νευρωνικού δικτύου
net = narnet(1:12,6,'open','trainlm');
net=init(net);
[Xs,Xi,Ai,Ts] = preparets(net,{}, {},T);
[net,tr] = train(net,Xs,Ts,Xi,Ai);
[y,Xf,Af]=net(Xs,Xi,Ai)
%view(net);
perf = perform(net,Ts,y)
rmse=sqrt(perf)
%δημιουργία αξόνων χρόνου
timeb= datetime(1999,12,31)+calmonths(1:numel(cell2mat(T)))
timedelay= datetime(1999,12,31)+calmonths(1:(length(Xi)))
timey=datetime(1999,12,31)+calmonths((length(Xi)))+calmonths(1:((length(T)-length(Xi))))
% γράφημα εκπαίδευσης
set(gcf,'Visible','on')
plot(timeb,cell2mat(T),'b')
xtickformat('MMM-yy')
hold on
plot (timedelay,cell2mat(Xi),'y')
plot (timey,cell2mat(y),'r')
hold off
xlabel('Περίοδος')
ylabel("GWh")
title('Εκτιμώμενες τιμές κατά την εκπαίδευση')
legend({'Πραγματικές τιμές','Υστέρηση δικτύου','Εκτιμώμενες τιμές'},,'Location','southeast')
saveas(gcf,'ekpaideysi.jpg');
```

```

%δημιουργία διαγραμμάτων αποτελεσμάτων εκπαίδευσης
plotperform(tr)
saveas(gcf,'performance.jpg');
fig=plottrainstate(tr)
saveas(fig,'plottrainstate.jpg');
trOut = y(tr.trainInd);
vOut = y(tr.valInd);
tsOut = y(tr.testInd);
trTarg = Ts(tr.trainInd);
vTarg = Ts(tr.valInd);
tsTarg = Ts(tr.testInd);
fig=plotregression(trTarg, trOut, 'Train', vTarg, vOut, 'Validation',
tsTarg, tsOut, 'Testing',Ts,y,'All')
saveas(fig,'Regression.jpg');
fig=plotresponse(Ts,y);
saveas(fig,'TimeSeries_response.jpg');
E = gsubtract(Ts,y);
fig=ploterrcorr(E)
saveas(fig,'Error_autocorrelation.jpg');
e=cell2mat(Ts)-cell2mat(y)
fig=ploterrhist(e)
saveas(fig,'error_histogram.jpg');
%μετατροπή σε κλειστό βρόχο και πρόβλεψη
[netc,Xic,Aic] = closeloop(net,Xf,Af);
yc = netc(cell(0,229),Xic,Aic)
% γράφημα πρόβλεψης
timeyc=timeb(length(timeb))+calmonths(1:length(yc))
plot(timeb,cell2mat(T),'b')
hold on
plot(timeyc,cell2mat(yc),'r')
hold off
xlabel('Περίοδος')
ylabel("GWh")
title('Πρόβλεψη')
legend({'Καταγεγραμμένες τιμές','Πρόβλεψη'},,'Location','southeast')
set(gcf,'Visible','on')

```



```
set(gca, 'XTick', sort([timeb(length(timeb)), get(gca, 'XTick')]));  
xtickangle(45)  
saveas(gcf, 'provlepsi.jpg');
```

Σενάριο 2

```
%NAR TEST 2
load ('matlab.mat','T')

delay=10
net = narnet(1:delay,1,'open',['trainlm']);
net=init(net)
[Xs,Xi,Ai,Ts] = preparets(net,{}, {},T);
[net,tr] = train(net,Xs,Ts,Xi,Ai);
[y,Xf,Af]=net(Xs,Xi,Ai);
%view(net)
perf = perform(net,Ts,y)
rmse=sqrt(perf)
Timedelay=[1960:(1960+numel(Xi)-1)]
TimeT=[1960:(1960+numel(T)-.1)]
Timey=[(1960+numel(Xi)):(numel(y)+1960+numel(Xi)-1)]

set(gcf,'Visible','on')
plot(TimeT,cell2mat(T),'b')
hold on
plot (Timedelay,cell2mat(Xi),'y')
plot(Timey,cell2mat(y),'r')
hold off
xlabel('Έτος')
ylabel("κατά κεφαλήν kWh")
title('Εκτιμώμενες τιμές κατά την εκπαίδευση')
legend({'Πραγματικές τιμές','Υστέρηση δικτύου','Εκτιμώμενες τιμές'},,'Location','southeast')
set(gca, 'XTick', sort([2014, get(gca, 'XTick')]));
xtickangle(45)
saveas(gcf,'ekpaideysi.jpg');

plotperform(tr)
saveas(gcf,'performance.jpg');
fig=plottrainstate(tr)
saveas(fig,'plottrainstate.jpg');
trOut = y(tr.trainInd);
vOut = y(tr.valInd);
tsOut = y(tr.testInd);
trTarg = Ts(tr.trainInd);
vTarg = Ts(tr.valInd);
tsTarg = Ts(tr.testInd);
fig=plotregression(trTarg, trOut, 'Train', vTarg, vOut, 'Validation',
tsTarg, tsOut, 'Testing',Ts,y,'All')
saveas(fig,'Regression.jpg');
fig=plotresponse(Ts,y);
saveas(fig,'TimeSeries_response.jpg');
E = gsubtract(Ts,y);
fig=ploterrcorr(E)
saveas(fig,'Error_autocorrelation.jpg');
e=cell2mat(Ts)-cell2mat(y)
fig=ploterrhist(e)
saveas(fig,'error_histogram.jpg');

[netc,Xic,Aic] = closeloop(net,Xf,Af);
yc = netc(cell(0,50),Xic,Aic)
```

```
Timeyc=[(TimeT(1,length(T))+1):(TimeT(1,length(T))+length(yc))]  
fig= plot(TimeT,cell2mat(T),'b')  
hold on  
fig=plot(Timeyc,cell2mat(yc),'r')  
hold off  
xlabel('Έτος')  
ylabel("κατά κεφαλήν kWh")  
title('Πρόβλεψη')  
legend({'Καταγεγραμμένες τιμές','Πρόβλεψη'},,'Location','southeast')  
set(gcf,'Visible','on')  
saveas(fig,'provlepsi.jpg');
```

Σενάριο 3

Πρόβλεψη NAR θερμοκρασίας

```
%NARX TEST 2 NAR TEMP
%εισαγωγή μεταβλητών
clear all
load ('T.mat','T')

%εκπαίδευση νευρωνικού δικτύου
net = narnet(1:12,3,'open','trainlm');
net=init(net);
[Xs,Xi,Ai,Ts] = preparets(net,{}, {},T);
[net,tr] = train(net,Xs,Ts,Xi,Ai);
[y,Xf,Af]=net(Xs,Xi,Ai)
%view(net);
perf = perform(net,Ts,y)
rmse=sqrt(perf)

%δημιουργία αξόνων χρόνου
timeb= datetime(2009,06,30)+calmonths(1:numel(cell2mat(T)))
timedelay= datetime(2009,06,30)+calmonths(1:(length(Xi)))
timey=datetime(2009,06,30)+calmonths((length(Xi))+calmonths(1:(length(T)-length(Xi))))

% γράφημα εκπαίδευσης
set(gcf,'Visible','on')
plot(timeb,cell2mat(T),'b')
xtickformat('MMM-yy')
hold on
plot (timedelay,cell2mat(Xi),'y')
plot (timey,cell2mat(y),'r')
hold off
xlabel('Περίοδος')
ylabel("oC")
title('Εκτιμώμενες τιμές κατά την εκπαίδευση')
legend({'Πραγματικές τιμές','Υστέρηση δικτύου','Εκτιμώμενες τιμές'},,'Location','southeast')
saveas(gcf,'ekpaideysi.jpg');

%δημιουργία διαγραμμάτων αποτελεσμάτων εκπαίδευσης
plotperform(tr)
saveas(gcf,'performance.jpg');
fig=plottrainstate(tr)
saveas(fig,'plottrainstate.jpg');
trOut = y(tr.trainInd);
vOut = y(tr.valInd);
tsOut = y(tr.testInd);
trTarg = Ts(tr.trainInd);
vTarg = Ts(tr.valInd);
tsTarg = Ts(tr.testInd);
fig=plotregression(trTarg, trOut, 'Train', vTarg, vOut, 'Validation',
tsTarg, tsOut, 'Testing',Ts,y,'All')
saveas(fig,'Regression.jpg');
fig=plotresponse(Ts,y);
saveas(fig,'TimeSeries_response.jpg');
```

```

E = gsubtract(Ts,y);
fig=ploterrcorr(E)
saveas(fig, 'Error_autocorrelation.jpg');
e=cell2mat(Ts)-cell2mat(y)
fig=ploterrhist(e)
saveas(fig, 'error_histogram.jpg');

%μετατροπή σε κλειστό βρόχο και πρόβλεψη
[netc,Xic,Aic] = closeloop(net,Xf,Af);
temp_pred = netc(cell(0,115),Xic,Aic)

% γράφημα πρόβλεψης
timetemp_pred=timeb(length(timeb))+calmonths(1:length(temp_pred))
plot(timeb,cell2mat(T), 'b')
hold on
plot(timetemp_pred,cell2mat(temp_pred), 'r')
hold off
xlabel('Περίοδος')
ylabel("GWh")
title('Πρόβλεψη')
legend({'Καταγεγραμμένες τιμές', 'Πρόβλεψη'}, 'Location', 'southeast')
set(gcf, 'Visible', 'on')
set(gca, 'XTick', sort([timeb(length(timeb)), get(gca, 'XTick')]));
xtickangle(45)
saveas(gcf, 'provlepsi.jpg');

```

Πρόβλεψη NARX Κατανάλωσης Ηλεκτρικής Ενέργειας

```
%NARX TEST 2
%εισαγωγή μεταβλητών
load ('matlab.mat','elec','temp')

%εκπαίδευση νευρωνικού δικτύου
delay=12
net = narxnet(0,1:delay,10,'open','trainlm');
net=init(net)
[Xs,Xi,Ai,Ts] = preparets(net,temp,{},elec);
[net,tr] = train(net,Xs,Ts,Xi,Ai);
[y,Xf,Af]=net(Xs,Xi,Ai);
perf = perform(net,Ts,y)
rmse=sqrt(perf)

%δημιουργία αξόνων χρόνου
timeb= datetime(2009,06,30)+calmonths(1:numel(cell2mat(elec)))
timedelay= datetime(2009,06,30)+calmonths(1:(length(Xi)))
timey=datetime(2009,06,30)+calmonths((length(Xi))+calmonths(1:(length(elec)-length(Xi))))

% γράφημα εκπαίδευσης
set(gcf,'Visible','on')
plot(timeb,cell2mat(elec),'b')
xtickformat('MMM-yy')
hold on
plot (timedelay,cell2mat(Xi(2,:)),'y')
plot (timey,cell2mat(y),'r')
hold off
xlabel('Περίοδος')
ylabel("GWh")
title('Εκτιμώμενες τιμές κατά την εκπαίδευση')
legend({'Πραγματικές τιμές','Υστέρηση δικτύου','Εκτιμώμενες τιμές'},'Location','southeast')
saveas(gcf,'ekpaideysi.jpg');

%δημιουργία διαγραμμάτων αποτελεσμάτων εκπαίδευσης
plotperform(tr)
saveas(gcf,'performance.jpg');
fig=plottrainstate(tr)
saveas(fig,'plottrainstate.jpg');
trOut = y(tr.trainInd);
vOut = y(tr.valInd);
tsOut = y(tr.testInd);
trTarg = Ts(tr.trainInd);
vTarg = Ts(tr.valInd);
tsTarg = Ts(tr.testInd);
fig=plotregression(trTarg, trOut, 'Train', vTarg, vOut, 'Validation', tsTarg, tsOut, 'Testing',Ts,y,'All')
saveas(fig,'Regression.jpg');
fig=plotresponse(Ts,y);
saveas(fig,'TimeSeries_response.jpg');
E = gsubtract(Ts,y);
fig=ploterrcorr(E)
saveas(fig,'Error_autocorrelation.jpg');
e=cell2mat(Ts)-cell2mat(y)
```

```

fig=ploterrhist(e)
saveas(fig, 'error_histogram.jpg');

[netc,Xic,Aic] = closeloop(net,Xf,Af);
yc=netc(temp_pred,Xic,Aic)

% γράφημα πρόβλεψης
timeyc=timeb(length(timeb))+calmonths(1:length(yc))
plot(timeb,cell2mat(elec),'b')
hold on
plot(timeyc,cell2mat(yc),'r')
hold off
xlabel('Περίοδος')
ylabel("GWh")
title('Πρόβλεψη')
legend({'Καταγεγραμμένες τιμές','Πρόβλεψη'},'Location','southeast')
set(gcf,'Visible','on')
set(gca,'XTick',sort([timeb(length(timeb)), get(gca,'XTick')]));
xtickangle(45)
saveas(gcf,'provlepsi.jpg');

```

Σενάριο 4

```

%NARX TEST 2
%εισαγωγή μεταβλητών
load ('αρχικα_δεδομενα.mat','elec')
%autoc = nncorr(T,T,10)

%συγχώνευση του χρονικού άξονα στα δεδομένα εισόδου
X2= [01 02 03 04 05 06 07 08 09 10 11 12 01 02 03 04 05 06 07 08 09 10 11 12
01 02 03 04 05 06 07 08 09 10 11 12 01 02 03 04 05 06 07 08 09 10 11 12 01
02 03 04 05 06 07 08 09 10 11 12 01 02 03 04 05 06 07 08 09 10 11 12 01 02
03 04 05 06 07 08 09 10 11 12 01 02 03 04 05 06 07 08 09 10 11 12 01 02 03
04 05 06 07 08 09 10 11 12 01 02 03 04 05 06 07 08 09 10 11 12 01 02 03 04
05 06 07 08 09 10 11 12 01 02 03 04 05 06 07 08 09 10 11 12 01 02 03 04 05
06 07 08 09 10 11 12 01 02 03 04 05 06 07 08 09 10 11 12 01 02 03 04 05 06
07 08 09 10 11 12 01 02 03 04 05 06 07 08 09 10 11 12 01 02 03 04 05 06 07
08 09 10 11 12 01 02 03 04 05 06 07 08 09 10 11 12 01 02 03 04 05 06 07 08
09 10 11 12 01]

for i=1:length(elec)
X{i}=[X2(1,i);cell2mat(elec(1,i))];
end

%εκπαίδευση νευρωνικού δικτύου
delay=12
net = narxnet(1:delay,1:delay, 5, 'open', 'trainlm');
%net.numInputs=2
net=init(net)
[Xs,Xi,Ai,Ts] = preparets(net,X,{},elec);
[net,tr] = train(net,Xs,Ts,Xi,Ai);
[y,Xf,Af]=net(Xs,Xi,Ai);
%view(net)
perf = perform(net,Ts,y)
rmse=sqrt(perf)

```

```

%δημιουργία αξόνων χρόνου
timeb= datetime(1999,12,31)+calmonths(1:numel(cell2mat(elec)))
timedelay= datetime(1999,12,31)+calmonths(1:(length(Xi)))
timey=datetime(1999,12,31)+calmonths((length(Xi))+calmonths(1:((length(elec)
)-length(Xi))))

% γράφημα εκπαίδευσης
set(gcf, 'Visible', 'on')
plot(timeb,cell2mat(elec), 'b')
xtickformat('MMM-yy')
hold on
plot (timedelay,cell2mat(Xi(2,:)), 'y')
plot (timey,cell2mat(y), 'r')
hold off
xlabel('Περίοδος')
ylabel("GWh")
title('Εκτιμώμενες τιμές κατά την εκπαίδευση')
legend({'Πραγματικές τιμές', 'Υστέρηση δικτύου', 'Εκτιμώμενες
τιμές', }, 'Location', 'southeast')
saveas(gcf, 'ekpaideysi.jpg');

%δημιουργία διαγραμμάτων αποτελεσμάτων εκπαίδευσης
plotperform(tr)
saveas(gcf, 'performance.jpg');
fig=plottrainstate(tr)
saveas(fig, 'plottrainstate.jpg');
trOut = y(tr.trainInd);
vOut = y(tr.valInd);
tsOut = y(tr.testInd);
trTarg = Ts(tr.trainInd);
vTarg = Ts(tr.valInd);
tsTarg = Ts(tr.testInd);
fig=plotregression(trTarg, trOut, 'Train', vTarg, vOut, 'Validation',
tsTarg, tsOut, 'Testing', Ts, y, 'All')
saveas(fig, 'Regression.jpg');
fig=plotresponse(Ts, y);
saveas(fig, 'TimeSeries_response.jpg');
E = gsubtract(Ts, y);
fig=ploterrcorr(E)
saveas(fig, 'Error_autocorrelation.jpg');
e=cell2mat(Ts)-cell2mat(y)
fig=ploterrhist(e)
saveas(fig, 'error_histogram.jpg');

[netc, Xic, Aic] = closeloop(net, Xf, Af);
yc=netc(X, Xic, Aic)

% γράφημα πρόβλεψης
timeyc=timeb(length(timeb))+calmonths(1:length(yc))
plot(timeb,cell2mat(elec), 'b')
hold on
plot(timeyc,cell2mat(yc), 'r')
hold off
xlabel('Περίοδος')
ylabel("GWh")
title('Πρόβλεψη')
legend({'Καταγεγραμμένες τιμές', 'Πρόβλεψη', }, 'Location', 'southeast')

```



```

set(gcf, 'Visible', 'on')
set(gca, 'XTick', sort([timeb(length(timeb)), get(gca, 'XTick')]));
xtickangle(45)
saveas(gcf, 'provlepsi.jpg');

```

Σενάριο 5

Πρόβλεψη NAR AEP

```

%NARX TEST 5 NAR aep
%εισαγωγή μεταβλητών
clear all
load ('cell.mat', 'cella')
%εκπαίδευση νευρωνικού δικτύου
net = narnet(1:8,6, 'open', 'trainlm');
net=init(net);
[Xs,Xi,Ai,Ts] = preparets(net, {}, {}, cella);
[net,tr] = train(net,Xs,Ts,Xi,Ai);
[y,Xf,Af]=net(Xs,Xi,Ai)
%view(net);
perf = perform(net,Ts,y)
rmse=sqrt(perf)
%δημιουργία αξόνων χρόνου
timeb= datetime(2000,03,30):calmonths(3):datetime(2018,12,31)
timedelay= datetime(2000,03,30):calmonths(3):datetime(2002,02,31)
timey=datetime(2002,02,31):calmonths(3):datetime(2018,12,31)
% διάγραμμα εκπαίδευσης
set(gcf, 'Visible', 'on')
plot(timeb, cell2mat(cella), 'b')
%xtickformat('MMM-yy')
hold on
plot (timedelay, cell2mat(Xi), 'y')
plot (timey, cell2mat(y), 'r')
hold off
xlabel('Περίοδος')
ylabel("AEP")
title('Εκτιμώμενες τιμές κατά την εκπαίδευση')
legend({'Πραγματικές τιμές', 'Υστέρηση δικτύου', 'Εκτιμώμενες τιμές'}, 'Location', 'southeast')
saveas(gcf, 'ekpaideysi.jpg');
%δημιουργία διαγραμμάτων αποτελεσμάτων εκπαίδευσης
plotperform(tr)
saveas(gcf, 'performance.jpg');
fig=plottrainstate(tr)
saveas(fig, 'plottrainstate.jpg');
trOut = y(tr.trainInd);
vOut = y(tr.valInd);
tsOut = y(tr.testInd);
trTarg = Ts(tr.trainInd);

```

```

vTarg = Ts(tr.valInd);
tsTarg = Ts(tr.testInd);
fig=plotregression(trTarg, trOut, 'Train', vTarg, vOut, 'Validation',
tsTarg, tsOut, 'Testing',Ts,y,'All')
saveas(fig, 'Regression.jpg');
fig=plotresponse(Ts,y);
saveas(fig, 'TimeSeries_response.jpg');
E = gsubtract(Ts,y);
fig=ploterrcorr(E)
saveas(fig, 'Error_autocorrelation.jpg');
e=cell2mat(Ts)-cell2mat(y)
fig=ploterrhist(e)
saveas(fig, 'error_histogram.jpg');

%μετατροπή σε κλειστό βρόχο και πρόβλεψη
[netc,Xic,Aic] = closeloop(net,Xf,Af);
AEP_pred = netc(cell(0,75),Xic,Aic)

%διάγραμμα πρόβλεψης
timetemp_pred=datetime(2019,03,31):calmonths(3):datetime(2037,09,31);
plot(timeb,cell2mat(cella),'b')
hold on
plot(timetemp_pred,cell2mat(AEP_pred),'r')
hold off
xlabel('Περίοδος (Τρίμηνο)')
ylabel("Ακαθάριστο Εθνικό Προϊόν (€)")
legend({'Καταγεγραμμένες τιμές', 'Πρόβλεψη'}, 'Location', 'southeast')
set(gcf, 'Visible', 'on')
saveas(gcf, 'provlepsi.jpg');

```

Πρόβλεψη NARX Κατανάλωσης Ηλεκτρικής Ενέργειας

```

%NARX TEST 2
%εισαγωγή μεταβλητών
load ('αρχικα_δεδομενα.mat','elec')
load ('AEP_pred','AEP_pred')
elec=table2cell(elec)
%autoc = nncorr(T,T,10)
X=table2cell(aep)

%εκπαίδευση νευρωνικού δικτύου
delay=6
net = narxnet(1:delay,1:delay, 5, 'open', 'trainlm');
%net.numInputs=2
net=init(net)
[Xs,Xi,Ai,Ts] = preparets(net,X,{},elec);
[net,tr] = train(net,Xs,Ts,Xi,Ai);
[y,Xf,Af]=net(Xs,Xi,Ai);
%view(net)
perf = perform(net,Ts,y)
rmse=sqrt(perf)

%δημιουργία αξόνων χρόνου
timeb= datetime(2000,03,30):calmonths(3):datetime(2018,12,31)
timedelay= datetime(2000,03,30):calmonths(3):datetime(2001,06,31)
timey=datetime(2001,09,30):calmonths(3):datetime(2018,12,31)

% διάγραμμα εκπαίδευσης

```

```

set(gcf, 'Visible', 'on')
plot(timeb, cell2mat(elec), 'b')
hold on
plot (timedelay, cell2mat(Xi(2, :)), 'y')
plot (timey, cell2mat(y), 'r')
hold off
xlabel('Περίοδος')
ylabel("GWh")
title('Εκτιμώμενες τιμές κατά την εκπαίδευση')
legend({'Πραγματικές τιμές', 'Υστέρηση δικτύου', 'Εκτιμώμενες τιμές'}, 'Location', 'southeast')
saveas(gcf, 'ekpaideysi.jpg');

%δημιουργία διαγραμμάτων αποτελεσμάτων εκπαίδευσης
plotperform(tr)
saveas(gcf, 'performance.jpg');
fig=plottrainstate(tr)
saveas(fig, 'plottrainstate.jpg');
trOut = y(tr.trainInd);
vOut = y(tr.valInd);
tsOut = y(tr.testInd);
trTarg = Ts(tr.trainInd);
vTarg = Ts(tr.valInd);
tsTarg = Ts(tr.testInd);
fig=plotregression(trTarg, trOut, 'Train', vTarg, vOut, 'Validation',
tsTarg, tsOut, 'Testing', Ts, y, 'All')
saveas(fig, 'Regression.jpg');
fig=plotresponse(Ts, y);
saveas(fig, 'TimeSeries_response.jpg');
E = gsubtract(Ts, y);
fig=ploterrcorr(E)
saveas(fig, 'Error_autocorrelation.jpg');
e=cell2mat(Ts)-cell2mat(y)
fig=ploterrhist(e)
saveas(fig, 'error_histogram.jpg');

[netc, Xic, Aic] = closeloop(net, Xf, Af);
yc=netc(X, Xic, Aic)

%διάγραμμα πρόβλεψης
timeyc=datetime(2019, 03, 31):calmonths(3):datetime(2037, 12, 31);
plot(timeb, cell2mat(elec), 'b')
hold on
plot(timeyc, cell2mat(yc), 'r')
hold off
xlabel('Περίοδος (Τρίμηνο)')
ylabel("Κατανάλωση Ηλεκτρικής Ενέργειας (GWh)")
legend({'Καταγεγραμμένες τιμές', 'Πρόβλεψη'}, 'Location', 'southeast')
set(gcf, 'Visible', 'on')
set(gca, 'XTick', sort([timeb(length(timeb)), get(gca, 'XTick')]));
xtickangle(45)
saveas(gcf, 'provlepsi.jpg');

```


ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1] K. B. Debnath and M. Mourshed, "Forecasting methods in energy planning models," *Renew. Sustain. Energy Rev.*, vol. 88, no. August 2016, pp. 297–325, 2018.
- [2] A. Kumar Singh, S. Khatoun, M. Muazzam, and D. K. Chaturvedi, "Network and Complex Systems An Overview of Electricity Demand Forecasting Techniques," vol. 3, no. 2, pp. 38–48, 2013.
- [3] A. Papalexopoulos, S. Hao, and T. M. Peng, "Implementation of a neural network based load forecasting model for the EMS," *IEEE Trans. Power Syst.*, 1994.
- [4] ΑΔΜΗΕ, "Μελέτη Επάρκειας Ισχύος 2017-2020," Athens, 2017.
- [5] M. I. Ullah, "Basic Statistics and Data Analysis," *WEN Themes*, 2013. [Online]. Available: <http://itfeature.com/time-series-analysis-and-forecasting/time-series-analysis-forecasting>. [Accessed: 15-Aug-2019].
- [6] J. A. M. Joe H. Chow, Felix F. Wu, *Applied mathematics for restructured electric power systems*. 2001.
- [7] G. A. N. Mbamalu and M. E. El-Hawary, *Load Forecasting Via Suboptimal Seasonal Autoregressive Models And Iteratively Reweighted Least Squares Estimation*. IEEE Transaction on Power System, 1992.
- [8] I. Rahman and S. Moghram, *Analysis and evaluation of five short-term load forecasting techniques*, vol. 4, no. 4. IEEE Transaction on Power System, 1989.
- [9] K. Liu *et al.*, "Comparison Of Very Short-Term Load Forecasting," *IEEE Trans. Power Syst.*, vol. 11, pp. 877–882, 1996.
- [10] E. H. Barakat, J. M. Al-Qassim, and A. Al-Rashed, "New Model For Peak Demand Forecasting Applied To Highly Complex Load Characteristics Of A Fast Developing Area," *IEE Proc. - C*, vol. 139, pp. 136–149, 1992.
- [11] McCulloch, Warren, and W. Pitts, "A Logical Calculus of Ideas Immanent in Nervous Activity," *Bull. Math. Biophys.*, p. 1943.
- [12] I. Βλαχαβάς, Π. Κεφαλάς, Β. Ν., Φ. Κόκκορας, and Η. Σακελλαρίου, *Τεχνητή Νοημοσύνη*. Εκδόσεις Πανεπιστημίου Μακεδονίας, 2011.
- [13] M. Nielsen, *Neural networks and deep learning*. Determination Press, 2015.
- [14] Α. Λύκας, *Τεχνητά Νευρωνικά Δίκτυα*. Πανεπιστήμιο Ιωαννίνων, 2008.
- [15] H. Hamedmoghadam, N. Joorabloo, and M. Jalili, "Australia's long-term electricity demand forecasting using deep neural networks," 2018.
- [16] OSarena, "Εισαγωγή στα Νευρωνικά Δίκτυα," 2013. [Online]. Available: <http://osarena.net/logismiko/applications/neural-networks-ti-einai.html>.
- [17] M. . Irving and C. N. Macqueen, "Robust algorithm for Load Estimation in Distribution Networks," no. IEEE Transactions on Power Delivery, 1994.
- [18] M. Djukanovic, B. Babic, O. J. Sobajic, and Y. H. Pao, "24-Hour Load Forecasting," *IEE Proc. D C*, vol. 140, pp. 331–318, 1993.
- [19] D. Srinivasan and M. A. Lee, "Survey of Hybrid Fuzzy Neural Approaches to Electric Load Forecasting," *Proc. IEEE Int. Conf. Syst. Man Cybern.*, 1995.
- [20] Mathworks, "Global Optimization Toolbox: User's Guide (r2018b)," 2018. .

- [21] S. A. Billing, *Nonlinear System Identification: NARMAX Methods in the Time, Frequency, and Spatio-Temporal Domains*. Wiley, 2013.
- [22] B. D. Ripley, “Pattern Recognition and Neural Networks,” *Cambridge Univ. Press*, p. 354.