Hurst-Kolmogorov dynamics in hydroclimatic processes and in the microscale of turbulence;

PhD thesis

by Panayiotis Dimitriadis

Supervisory board:

Demetris Koutsoyiannis (Supervisor; NTUA)

Panos Papanicolaou (Co-supervisor; NTUA)

Christian Onof (Imperial College of London)

Examination board:

Athanasios Loukas (University of Thessaly)

Anastasios Stamou (NTUA)

Theodoros Karakasidis (University of Thessaly)

Andreas Langousis (University of Patras)

National Technical University of Athens

Athens, May 2017

Chaos was the law of nature; Order was the dream of man.

(Henry Adams, 1918)

In this workshop we are venturing into a smoky area of science
where nobody knows what the real truth is.
Such fields are always dominated by the compensation phenomenon:
supreme self-confidence takes the place of rational arguments.

(ET. Jaynes, 1990)

That new data that we insist on analyzing
in terms of old ideas
(that is, old models which are not questioned)
cannot lead us out of the old ideas.

(ET. Jaynes, 1996)

Everything should be as simple as it can be,
but not simpler

(quote attributed to A. Einstein in 1933).

When you have combined experimentation,
mathematically and physically based justification,
time-series analysis of billions of observations,
new parsimonious ideas applied to old and new data,
and some things seem to be puzzled out , then
you may have put some order into the chaos in Nature
but also, Nature has put some chaos into the order in you.

Abstract

The high complexity and uncertainty of atmospheric dynamics has been long identified through the observation and analysis of hydrometeorological processes such as temperature, humidity, atmospheric wind, precipitation, atmospheric pressure, river discharges etc. Particularly, all these processes seem to exhibit high unpredictability due to the clustering of events, a behaviour first identified by H.E. Hurst in 1951 while working at the River Nile, although its mathematical description is attributed to A. N. Kolmogorov who developed it while studying turbulence in 1940. To give credits to both scientists this behaviour and dynamics is called Hurst-Kolmogorov (HK). In order to properly study the clustering of events as well as the stochastic behaviour of hydrometeorological processes in general we would require numerous of measurements in annual scale. Unfortunately, large lengths of high quality annual data are hardly available in observations of hydrometeorological processes. However, the microscopic processes driving and generating the hydrometeorological ones are governed by turbulent state. By studying turbulent phenomena in situ we may be able to understand certain aspects of the related macroscopic processes in field. Certain strong advantages of studying microscopic turbulent processes in situ is the recording of very long time series, the high resolution of records and the controlled environment of the laboratory. The analysis of these time series offers the opportunity of better comprehending, control and comparison of the two scientific methods through the deterministic and stochastic approach.

In this thesis, we develop the stochastic framework for the empirical as well as theoretical estimation of the marginal characteristic and dependence structure of a process. Also, we develop and apply explicit and implicit algorithms for stochastic synthesis of mathematical processes as well as stochastic prediction of physical processes. Moreover, we discuss and suggest a definition for turbulent processes through the Hurst parameter and the drop of variance with scale based on experiments held at the laboratory. Additionally, we propose a stochastic model for the behaviour of a process from the micro to the macro scale that results from the maximization of entropy. Finally, we apply this model to other microscale turbulent processes as well as to temperature, precipitation, humidity, atmospheric pressure, river discharges and wind time series from thousands of stations around the globe and several billions of data.

A summary of the major innovations of the thesis are: (a) the further development, and extensive application to numerous processes, of the classical second-order stochastic framework including innovative approaches to account for discretization effects and statistical bias; (b) the further development of stochastic generation schemes such as the Sum of Autoregressive (SAR) models, e.g. AR(1) or ARMA(1,1), the Symmetric-Moving-Average (SMA) scheme in multiple dimensions (that can generate any process second-order dependence structure, marginal distribution and certain aspects of the intermittency behaviour) and an implicit and explicit cyclo-stationary (CSAR and CSMA) schemes for simulating the periodicities of a process such as seasonal and diurnal; and (c) the introduction and application of an extended HK stochastic model (with an identical expression of marginal distribution and correlation structure) that is in agreement with an interestingly large

variety of turbulent (such as thermal jet of positively buoyancy processes using laser-induced-fluorescence techniques as well as grid-turbulence generated within a wind-tunnel) and hydroclimatic processes (such as temperature, atmospheric wind, dew-point/humidity, precipitation and atmospheric pressure in a global scale).

Keywords: generic stochastic methodology; second order dependence structure; marginal probability density function; intermittency; principle of maximized entropy; longterm persistence; climacogram; autocovariance; power spectrum; variogram; simulation and prediction stochastic algorithms; sum of independent Markov models; explicit moving-average generation scheme; explicit and implicit cyclostationary generation schemes; statistical uncertainty of deterministic models; process discretization; estimators adjusting statistical bias; fitting norms for both distribution tails; small to large scale analysis; experimental turbulent jets; grid-turbulence; global databases; temperature; dew-point/humidity; wind speed; precipitation; river-discharge; atmospheric pressure; Köppen-Geiger climatic classification.

Περίληψη

Η υψηλή πολυπλοκότητα και αβεβαιότητα της δυναμικής της ατμόσφαιρας έχει από καιρό αναγνωρισθεί μέσα από την εμπειρία και ανάλυση των υδρομετεωρολογικών διεργασιών, όπως θερμοκρασία, υγρασία, άνεμος, βροχόπτωση, ατμοσφαιρική πίεση, παροχές ποταμού κτλ. Συγκεκριμένα, όλες αυτές οι διεργασίες φαίνεται να εμπεριέχουν μεγάλη αβεβαιότητα στην πρόβλεψη που επιτείνεται λόγω της ομαδοποίησης ομοειδών φαινομένων. Αυτή η συμπεριφορά είναι πολύ διαφορετική από την εποχική περιοδικότητα που συμβαίνει σε υπο-ετήσια κλίμακα. Η ομαδοποίηση αυτή των φαινομένων ανιχνεύτηκε πρώτα από τον Η.Ε. Hurst το 1951 στο πλαίσιο μελέτης έργων στον ποταμό Νείλο. Η μαθηματική έκφραση αυτής της συμπεριφοράς αποδίδεται στον Α. Kolmogorov που την ανέπτυξε ενώ μελετούσε τυρβώδη φαινόμενα το 1940. Για να δοθεί εξίσου αναγνώριση και στους δύο επιστήμονες, το φαινόμενο και η δυναμική αυτή ονομάζεται Hurst-Kolmogorov (HK).

Για την σωστή μελέτη αυτής της ομαδοποίησης των φαινομένων και γενικά την στοχαστική συμπεριφορά των υδρομετεωρολογικών διεργασιών, θα χρειαζόμασταν άφθονες μετρήσεις σε ετήσια κλίμακα. Δυστυχώς, μεγάλα μήκη και υψηλής ποιότητας δεδομένα είναι δύσκολο να βρεθούν για υδρομετεωρολογικές διεργασίες. Ωστόσο, οι φυσικές διεργασίες μικρής κλίμακας που δημιουργούν και οδηγούν τις υδρομετεωρολογικές, διέπονται από τυρβώδη συμπεριφορά. Μελετώντας την μικροκλίμακα τυρβωδών φαινομένων σε εργαστήριο, μπορούμε να κατανοήσουμε ορισμένες εκφάνσεις των συγγενών μακροσκοπικών διεργασιών στο πεδίο. Υπάρχουν ορισμένες ομοιότητες μεταξύ της μικροκλίμακας της ταχύτητας του ανέμου και της θεωρίας τυρβώδους οριακού στρώματος. Επίσης, το μέγεθος των σταγόνων βροχής, που είναι συνυφασμένο με την μορφή και ένταση επεισοδίων βροχόπτωσης, επηρεάζεται από την τυρβώδη κατάσταση της μικροκλίμακας του ανέμου. Ορισμένα ισχυρά πλεονεκτήματα της μελέτης στη μικροκλίμακα τύρβης στο εργαστήριο είναι η καταγραφή χρονοσειρών μεγάλου μήκους, η υψηλή συχνότητα καταγραφής και το ελεγχόμενο περιβάλλον του εργαστηρίου. Η ανάλυση αυτών των χρονοσειρών μας δίνει τη δυνατότητα καλύτερης κατανόησης, ελέγχου και σύγκρισης των δύο επιστημονικών μεθόδων, της ντετερμινιστικής και της στοχαστικής ανάλυσης.

Σε αυτή την διατριβή, αναπτύσσουμε το πλαίσιο της στοχαστικής ανάλυσης για την εμπειρική αλλά και θεωρητική εκτίμηση περιθώριων χαρακτηριστικών και δομής συσχέτισης μιας διεργασίας. Επίσης, αναπτύσσουμε και εφαρμόζουμε αλγορίθμους στοχαστικής σύνθεσης μαθηματικών ανελίξεων αλλά και στοχαστικής πρόβλεψης φυσικών διεργασιών. Επίσης, συζητούμε και προτείνουμε έναν χαρακτηρισμό της τυρβώδους συμπεριφοράς μέσα από την παράμετρο Hurst και την μείωση της διασποράς με την αύξηση της χρονικής κλίμακας με βάση εργαστηριακά πειράματα θερμαινόμενης τυρβώδους φλέβας. Επιπρόσθετα, προτείνουμε ένα στοχαστικό μοντέλο συμπεριφοράς μιας διεργασίας από μικρές σε μεγάλες κλίμακες, που προκύπτει από την μεγιστοποίηση της εντροπίας. Τέλος, εφαρμόζουμε αυτό το μοντέλο και σε άλλες διεργασίες μικροκλίμακας τύρβης αλλά και σε χρονοσειρές θερμοκρασίας, βροχόπτωσης, υγρασίας, ατμοσφαιρικής πίεσης, παροχών ποταμού και ανέμου, από χιλιάδες σταθμούς ανά τον κόσμο.

Selected publications

1) Dimitriadis, P., and D. Koutsoyiannis, Climacogram versus autocovariance and power spectrum in stochastic modelling for Markovian and Hurst–Kolmogorov processes, *Stochastic Environmental Research and Risk Assessment*, 29, 1649–1669, 2015.
2) Dimitriadis, P., and D. Koutsoyiannis, Application of stochastic methods to double cyclostationary processes for hourly wind speed simulation, *Energy Procedia*, 76, 406–411, 2015.
3) Dimitriadis, P., D. Koutsoyiannis, and P. Papanicolaou, Stochastic similarities between the microscale of turbulence and hydro-meteorological processes, *Hydrological Sciences Journal*, 61, 1623–1640, 2016.
4) Dimitriadis, P., D. Koutsoyiannis, and K. Tzouka, Predictability in dice motion: how does it differ from hydro-meteorological processes? *Hydrological Sciences Journal*, 61, 1611–1622, 2016.
5) Dimitriadis, P., A. Tegos, A. Oikonomou, V. Pagana, A. Koukouvinos, N. Mamassis, D. Koutsoyiannis, and A. Efstratiadis, Comparative evaluation of 1D and quasi-2D hydraulic models based on benchmark and real-world applications for uncertainty assessment in flood mapping, *Journal of Hydrology*, 534, 478–492, 2016.
6) Deligiannis, E., P. Dimitriadis, O. Daskalou, Y. Dimakos, and D. Koutsoyiannis, Global Investigation of Double Periodicity of Hourly Wind Speed for Stochastic Simulation; Application in Greece, *Energy Procedia*, 97, 278–285, 2016.
7) Dimitriadis, P., K. Tzouka, D. Koutsoyiannis, H. Tyralis, A. Kalamioti, E. Lerias, and P. Voudouris, Stochastic investigation of long-term persistence in two-dimensional images of rocks, *Journal of Spatial Statistics*, 2017 (accepted with minor revisions).
8) Koutsoyiannis, D., P. Dimitriadis, F. Lombardo, and S. Stevens, From fractals to stochastics: Seeking theoretical consistency in analysis of geophysical data, Advances in Nonlinear Geosciences, edited by A.A. Tsonis, 237–278, Springer, 2018.
9) Dimitriadis, P., and D. Koutsoyiannis, Stochastic synthesis approximating any process dependence and distribution, *Stochastic Environmental Research and Risk Assessment*, 2017 (submitted in December 2016 and was accepted with major revisions in June 2017, and then again was accepted with major revisions in September 2017, and now we are still waiting).

Overture and acknowledgment

An accomplishment may be important (or not) to know, whereas the extreme conditions (if any) under which this was made are always important to know.

Theoretical model of the PhD

The initial plan involved (as is the often case for a young scientist) an effort to analyze everything and solve all problems of humanity. However, the title would be too short, so we had to come up with a more specialized one. A fair compromise was to combine my favorite courses from my Civil Engineer studies at NTUA (Stochastics and Applied Hydraulics) and from the MSc (by scholarship) in Hydrology at Imperial College of London (Stochastic and Hydrometeorology). But now it seemed too easy, so we added some Laboratory Experiments to link the areas. Fortunately, the described topic was not already taken, so along with Demetris, Panos and Christian, we formed the final title of my Ph.D thesis (which interestingly, remained the same until the end).

So, the original plan was simply to:

- Understand and improve the framework of Stochastics (from the statistical analysis of a timeseries to the introduction, application ad generation of a second-order stochastic model).
- Perform laboratory experiments at NTUA (velocity and concentrations of hydraulic jets)
- Find stochastic similarities between these two and among other hydrometeorological processes from analysis of thousands of stations around the globe.

Boundary conditions of the PhD

In the next Figure, we present my extended supervisory committee.

I am really honored and thankful to have collaborate with Demetris for so many years (I know him for 11 years now) not only as a scientist but also (and most importantly) as a friend. He is (without doubt) the greatest, most intuitive and well educated Scientist and Teacher in his fields of expertise I had ever met with such a universally recognized work. He has collaborated as equal to equal with all the members of his universal team (and outside his team), colleagues, scientists and students, and has created the (in situ) scientific community of ITIA (it is not by luck that all the members of ITIA have become great scientists in their fields of expertise).

I am also thankful and honored that I had the opportunity to collaborate with Panos, a great Teacher, an expert in fluid mechanics and the greatest experimentalist Civil Engineer (in situ and in field) I had ever met in our School at NTUA. I can't even remember how many hours he spent working with me at the Laboratory with classical and high-tech technologies (like the Laser-Induced-Fluorescence and PIV) and helping other colleagues, scientist and students to get familiar with the art of experimentalist. He also taught me the importance of experiments and measurements in every aspect of engineering work and particularly, in Hydraulics and Turbulence.

Also, I am thankful to Christian who is a great Mathematician in his field of expertise. I had the luxury of meeting him as my Teacher at Imperial College and his expertise in stochastics came at hand when higher mathematical knowledge were required as for example, when we were struggling to find some properties of the n-dimensional field of the second-order stochastic framework.

I am thankful to Nikos Mamasis, who is, without doubt, a great Teacher capable of explaining even difficult and sophisticated meanings to any person willing to listen. He had supported me in several situations during my PhD and his expertise in climate dynamics came at hand more times than I can remember. He has handled the largest number of undergraduate and graduate theses (almost 100), teaches at NTUA in 10 courses and has worked in over 30 projects.

I am thankful to Andreas Efstratiadis, who has scientifically (and philosophically) supported me several times during my PhD. He is the strongest non-academic I have ever met with publications and citations that are above the average of the academic community in our School. He has handled a very large number of undergraduate and graduate theses (almost 50), teaches at NTUA in 5 courses and has worked in over 17 projects.

Great Thanks are also due to the ITIA research group that besides the financial crisis has a large number of expertises in many fields, provides highly sophisticated open-software for hydrological management and modelling, and keeps inspiring young scientists. More particular, I would like to thank for their friendship, collaboration, exchange of ideas and support, Any Iliopoulou, Katerina Tzouka, Hristos Tyralis, Yiannis Markonis, Federico Lombardo and his beautiful family, Georgia Papacharalampous, Romanos Ioannidis, Simon Papalexiou, Antonis Koukouvinos, Antonis Christofides, Sandra Mpaki, George Karavokiros, Evangelos Rozos, Archontia Lykou, Yiannis Tsoukalas, Panagiotis Kossieris, Stefanos Kozanis, Vicky Tsoukala, Christos Markopoulos, and Dimitris Dermatas.

I am also thankful to Panos' strongest PhD students Dr. Elias Papakonstantis and Dr. Spyros Michas, and his PhD candidates Aris Mauromatis and Evgenios Retsinis, for helping me with the experiments at the laboratory (many Thanks also to Manolis and Giannis, the two tireless engineers of the laboratory). Also, I would like to thank Giannis Nikiforakis for his the exchange of ideas and collaboration in so many experiments held at the laboratory, Georgia Papadonikolaki for her collaboration, exchanging of ideas, and our interesting scientific talks drinking coffee and beers, and Anthi Gkesouli for our friendly talks and exchanging of ideas.

Also, smaller but Crucial contribution has been made by Marina Pantazidou (for the exchange of ideas and interesting questions), Andreas Langousis (for his strong but fair position towards science and always under a friendly, but sometimes rather high-tempered, concern), and Anastasios Stamou (he has introduced me to the Fluid Mechanics during my graduate thesis and our collaboration in several projects).

A Great contribution was offered by my Family (in general), my beautiful Kondylia and our new born Anne, and my Friends (old-time, colleagues, poets and students).

Figure: My (extended) supervisory committee. From left to right, Demetris Koutsoyiannis, Panos Papanicolaou, Christian Onof, Nikos Mamasis, Andreas Efstratiadis and the ITIA group.

Since "a man is known by the company he keeps" a successful PhD should be on the way.

Initial conditions of the PhD

The beginning of my PhD is placed at the beginning of the Financial Crisis in Greece. After my return to Greece we got at least 5 rejections concerning my PhD (2010 to 2017):

- Heraclitus II (2010, European Commission): "*The most excellent proposal in Hydrology*" (evaluated 9/10 by reviewer from Greece) vs. "*This is already done*" (evaluated 6/10 by anonymous Greek reviewer from USA, without providing any references justifying this statement) resulted in a final decision with rejection by a third reviewer (NTUA).
- NTUA (internal scholarship): rejection due to application at an early stage of my PhD.
- NTUA (internal scholarship): rejection due to application at a late stage of my PhD.
- NTUA (ΠΕΒΕ): successful (!) but NTUA unable to fund research due to the financial crisis.
- State Scholarships Foundation (ΙΚΥ, Greece): Rejection by mistake –my supervisor was accidently evaluated lower than me– (President of ΙΚΥ promised through email that he will never let the two anonymous reviewers participate in evaluations again).
- Laboratory of Applied Hydraulics of NTUA lacked of appropriate facilities for microscale turbulence experiments (e.g., a dark room was necessary for the calibration and application of the Laser-Induce-Fluorescence technique).

Numerical scheme of the PhD

The PhD typically started (part time) a little bit later (2012) due to the funding provided by several NTUA projects (supervised by Demetris, Panos and Nikos). In total, I gained great work experience by doing several tasks (such as land surveying and statistical analysis of medical data). I gained great scientific experience by meeting several challenges (such as working side-by-side with great scientists and colleagues, and co-supervising undergraduate and graduate theses). For example, turbulent experiments were held during the night (mostly 20:00 to 03:00 and sometimes even later) and also a few times at the University of Thessaly in Volos. Difficult numerical calculations

were performed mostly using open-software (or software provided by NTUA). For the above reasons, creativity was highly increased after giving up on the system and started giving trust to people that never failed me (I hope I didn't failed them). Note that the only problem was that I didn't have much time to exercise (so, I gained a little weight).

General output results of the PhD

The general results from the PhD thesis are:

- In total nine publications in scientific journals (some additional ones are still pending)
- Around 45 conference publications (in 17 conferences, mostly funded by NTUA)
- More than 25 co-supervised theses (undergraduate and graduate level)
- Participation in several projects, 5 Courses (3 at undergraduate level and 2 at graduate level) and challenging tasks (e.g., organizing tens of students for the EGU conference)
- Met great people! (see next Figure for a small sample)

Figure: A sample of the Great People I met during my PhD thesis.

Contents

1 Introduction 1

1.1 The complexity of nature 1

1.2 The stochastic approach 1

1.3 The Hurst-Kolmogorov dynamics 2

1.4 From the microscopic analysis to the macroscopic observation 2

1.5 Scientific innovations of the thesis 3

2 Definitions, methods and notation for stochastic analysis 4

2.1 The definition of Stochastics and related concepts 4

2.2 Observing a natural stochastic process 5

2.3 Stochastic metrics for identification of a stochastic process 6

2.3.1 Most common measures for the marginal characteristics of a process 6

2.3.2 Most common and uncommon metrics for the dependence structure of a process 7

2.3.3 Climacogram-based metrics 7

2.3.4 Autocovariance-based metrics 9

2.3.5 The power spectrum 11

2.4 Stochastic processes and estimators used in thesis 13

2.4.1 The Markov process 14

2.4.2 The HK-behaviour processes 15

2.4.3 A mixed dependence structure from entropy extremization 17

2.4.4 Distributions based on entropy extremization 19

2.4.5 Comparison between autocovariance-based and climacogram-based measures for common processes 20

2.5 Proposed methodology for stochastic modelling 24

3 Stochastic synthesis and prediction algorithms 27

3.1 Synthesis of a Markov process 27

3.2 Sum of Markov processes; the SAR process and algorithms 30

3.3 Synthesis of a stochastic process through the (S)MA scheme 33

3.3.1 The impracticality of using multi-parameter stochastic models in geophysical processes 34

3.3.2 The impracticality of estimating higher-order moments in geophysical processes 37

3.3.3 The SMA generation scheme 39
3.4 Synthesis of a multiple dimensional process through SMA scheme 42
3.5 Prediction algorithms 45
3.5.1 Analogue prediction algorithm 46
3.5.2 Stochastic prediction algorithm 46
4 Uncertainty and HK dynamics 48
4.1 Complex natural systems 48
4.1.1 Experimental setup of dice throw 50
4.1.2 Hydrometeorological processes of high resolution 55
4.1.3 Uncertainty evaluation and comparison 56
4.2 Deterministic systems 61
4.2.1 A classical deterministic system 61
4.2.2 Comparison between deterministic systems of high complexity 62
4.3 HK dynamic as a measure of uncertainty 75
5 Application to microscale turbulent processes 76
5.1 On the definition of turbulence 76
5.1.1 Stochastic properties of large-scale range 78
5.1.2 Stochastic properties of intermediate range 80
5.1.3 Stochastic properties of small-scale range 81
5.2 Proposed model 83
5.3 Applications to laboratory microscale turbulent processes 84
5.3.1 Laboratory measurements of grid-turbulence velocities 84
5.3.2 Laboratory measurements of turbulent thermal jet temperatures 88
5.4 Stochastic similarities between the microscale of turbulent processes and the mesoscale geophysical ones 95
6 Application to hydrometeorological processes 99
6.1 Stochastic analysis of a long daily precipitation timeseries 99
6.2 Stochastic analysis of the longest hourly wind timeseries in Greece 101
6.3 Global stochastic analysis of the hourly wind process 103
6.4 Global stochastic analysis of the hourly temperature process 109
6.5 Global stochastic analysis of hydrometeorological processes based on the Koppen-Geiger climatic-classification 111

7 Conclusions and summary of thesis major innovations 115

Table Captions

Table 1: Climacogram definition and expressions for a process in continuous and discrete time, along with the properties of its estimator. Source: Dimitriadis et al., (2016a). 8

Table 2: Climacogram-based variogram (CBV) definition and expressions for a process in continuous and discrete time, along with the properties of its estimator (source: Dimitriadis et al., 2016a). 8

Table 3: Climacogram-based spectrum (CBS) definition and expressions for a process in continuous and discrete time, along with the properties of its estimator (source: Dimitriadis et al., 2016a)..... 9

Table 4: Autocovariance definition and expressions for a process in continuous and discrete time, along with the properties of its estimator (source: Dimitriadis et al., 2016a)..... 10

Table 5: Variogram definition and expressions for a process in continuous and discrete time, along with the properties of its estimator (source: Dimitriadis et al., 2016a). 11

Table 6: Power spectrum definition and expressions for a process in continuous and discrete time, along with the properties of its estimator (source: Dimitriadis et al., 2016a)..... 13

Table 7: Climacogram, autocovariance and power spectrum expressions of a Markov process, in continuous and discrete time (source: Dimitriadis and Koutsoyiannis, 2015a)..... 15

Table 8: Climacogram, autocovariance and power spectrum expressions of a positively correlated gHK process, with $0 < b < 1$, in continuous and discrete time. 17

Table 9: Parameters p_1 and p_2 estimated to fit different types of HK and gHK processes (for $\lambda = 1$) with a sum of Markov processes for $n = 10^2$. 32

Table 10: Parameters p_1 and p_2 estimated to fit different types of HK and gHK processes (for $\lambda = 1$) with a sum of Markov processes for $n = 10^3$. 33

Table 11: Parameters p_1 and p_2 estimated to fit different types of HK and gHK processes (for $\lambda = 1$) with a sum of Markov processes for $n = 10^4$. 33

Table 12: Climacogram definition and expressions for an Ld continue process, a discretized one, a common estimator for the climacogram and the estimated value, based on this estimator. 43

Table 13: Autocovariance definition and expressions for an Ld continue process, a discretized one, a common estimator for the autocovariance and the estimated value, based on this estimator. 44

Table 14: Definition of variables x, y and z that represent proportions of each pair of opposite colours (source: Dimitriadis et al., 2016b). 53

Table 15: Variables used within sensitivity analysis and associated range of feasible values; all variables are uniformly distributed, except for the model resolution determined by the channel width, which takes three discrete values with equal probability (25, 50 or 100 m). 66

Table 16: Central moments' variation (denoted Cv), skewness (denoted Cs) and excess kurtosis (denoted Ck) coefficients (using the unbiased classical estimators) for each model applied as well as cross correlation coefficients between the input and output variables. Source: Dimitriadis et al. (2016b). 69

Table 17: First order correlation between various hydraulic models and schemes as estimated from the sensitivity analysis. Source: Dimitriadis et al. (2016c). 73

Table 18: 1d and 3d power spectrum for Markov, powered-exponential and gHK processes as well as their LLD, where λ is the parameter related to the true variance of the process, q the scale parameter and b is related to the power-type behaviour of the process (source: Dimitriadis et al., 2016a). 78

Table 19: Details of the experiments held at the Laboratory of Hydraulics at the NTUA on the period 1/5/09 to 1/10/10 (where C_o is the R6G initial concentration, D is the diameter of the nozzle, Q is the initial discharge of R6G, T_{amb} and T_{jet} are the ambient and jet temperature). Source: Dimitriadis and Papanicolaou (2010). 92

Table 20: General information of the meteorological stations and statistical characteristics of the hourly wind timeseries (downloaded from ftp.ncdc.noaa.gov). Source: Deligiannis et al. (2016). . 102

Table 21: Hurst parameter under Köppen-Geiger classification (source: Dimitriadis et al., 2016d). 112

Figure Captions

Figure 1: The steps for a stochastic analysis (source: Koutsoyiannis and Dimitriadis, 2016). 5

Figure 2: An example of realization (blue line) of a continuous time process $\underline{x}$ and a sample of $x_i^{(\Delta,D)}$ realizations (black dots) of the discretized process $\underline{x}_i^{(\Delta,D)}$ averaged at time scale Δ, with time intervals D and for a total period T (source: Dimitriadis et al., 2016a). 6

Figure 3: Ratio of the true Markov process at lag one for $D \neq \Delta$ over the one with $D = \Delta$ vs. D/q, for various values of the ratio Δ/q. 14

Figure 4: Ratio of the true HK process for $D/\Delta \geq 5$ vs. the one with $\Delta = D$ for various Hurst parameters. 16

Figure 5: The ELTP of an HK process with $H = 5/6$, a Markov process with $q = 1$ and two powered-exponential functions with $q = 1$, and $M = 2/3$ and $M = 1/3$ 18

Figure 6: Dimensionless error between the autocovariance of a Markov process and those of expressed through various AR(1) models. 28

Figure 7: Expected 5% and 95% quantiles of the climacogram for an HK process estimated from Monte-Carlo experiments using the SMA model (exact solution) and the sum of three AR(1) models (3AR1) through the SAR scheme. 35

Figure 8: Expected 5% and 95% quantiles of the climacogram for an HK process estimated from Monte-Carlo experiments using the SMA model (exact solution) and the sum of five ARMA(1,1) models (5ARMA11) through the SARMA scheme. 36

Figure 9: Expected 5% and 95% quantiles of the climacogram for two HHK processes, both with q = 10, b = 1/3, n = 2×10^3 and one with α = 2/3 < 1 (left) and the other with α = 3/2 > 1 (right), estimated from Monte-Carlo experiments using the SMA model (exact solution) and the sum of five ARMA(1,1) models (5ARMA11) through the SARMA scheme. 36

Figure 10: Dimensionless climacogram vs. scale for a synthetic HK process with n = 10^5, H = 0.8 and distribution N(0,1) as well as its transformation to U(0,1) and Pareto distribution with shape parameter equal to 4. 37

Figure 11: Standard deviation of the mean estimator of an HK process standardized by σ vs. the sample size (n) for various Hurst coefficients. 38

Figure 12: Standard deviation of the sample estimates of the mean (μ), standard deviation (σ), skewness coefficient (C_s) and kurtosis coefficient (C_k) of an HK process with H = 0.8 and N(0,1) distribution vs. the simulation length. 39

Figure 13: Various two-parameter distributions along with the fitted ME probability density function and the empirical probability density from one single simulation with n = 10^5 using the proposed generation scheme. 42

Figure 14: Images of sandstone as seen from the SEM (50 μm), from a polarizing microscope, (3.5 mm), from a hand specimen (with length approximately 5 cm) and a field outcrop (1 m). For more information on the source, description and processing of the images see in Dimitriadis et al., (2017). 45

Figure 15: Climacograms of sandstone images depicted at four different scales (source: Dimitriadis et al., 2017) 45

Figure 16: Mixed combination of frames taken from all die throw experiments for illustration (source: Dimitriadis et al., 2016b) 50

Figure 17: Selected frames showing the die trajectory from experiments (a) 48 and (b) 78; (c, d) their three Cartesian coordinates (denoted x_c, y_c and z_c for length, width and height, respectively); (e) standardized audio index representing the sound the die makes when colliding with the box; and (f) colour triplets (each of the 8 possible triplets corresponds to three neighbouring colours). Source: Dimitriadis et al. (2016b). 52

Figure 18: All experiments (a) standardized audios, showing the time the die collides with the box (picks) and (b) linear velocities. Source: Dimitriadis et al. (2016b).....53

Figure 19: Time series of variables x, y and z for experiments 48 (a, b, c) and 78 (d, e, f); in both experiments the outcome was green. Source: Dimitriadis et al. (2016b).....54

Figure 20: Plot of (a) all (x, y) and (u, v) points from all experiments and (b) the probability density function of (u, v). Source: Dimitriadis et al. (2016b). 55

Figure 21: (a) Rainfall events 1 and 7 from Georgakakos et al. (1994) and (b) wind events 3 and 5 provided by NCAR/EOL. 56

Figure 22: True, expected and averaged empirical climacograms for (a) u and v, (b) w, (c) ξ and (d) ψ. Source: Dimitriadis et al. (2016b).....58

Figure 23: Comparisons of B1, B2, stochastic and analogue models for the die experiment (a and b), the observed rainfall intensities (c and d) and the observed wind speed (e and f). The left column (a, c and e) represents the application of the models to all experiments and events and the right column (b, d and f) to individual ones. Source: Dimitriadis et al. (2016b).....59

Figure 24: Sensitivity analyses of the stochastic and analogue model parameters for the die experiment (a and b), the rainfall intensities (c and d) and the wind speed (e and f). Source: Dimitriadis et al. (2016b).....60

Figure 25: (a) Values of X_L, Y_L and Z_L, plotted at a time interval of 0.1, for the 5th timeseries produced by integrating the classical Lorenz's chaotic system of equations, (b) observed climacogram as well as its true and expected values for the fitted stochastic gHK model (average of X_L, Y_L and Z_L processes), (c) sensitivity analysis of the analogue and stochastic models and (d) comparison of the optimum stochastic and analogue models with B1 and B2. Source: Dimitriadis et al. (2016b).....62

Figure 26: Layout of benchmark tests and associated input variables: (a) perspective view, (b) plan view, and (c) cross sectional view, where solid lines represent the continuous geometry, implemented within HEC-RAS, while dashed lines represent the raster-based geometry, implemented within LISFLOOD-FP and FLO-2d (d_c represents the channel depth; for rest of symbols please refer to Table 15). Source: Dimitriadis et al. (2016b).....66

Figure 27: Moving average of (a) coefficient of variation, Cv, for all model configurations of the water depth of the channels' upstream and downstream cell/section, and (b) mean, μ, (c) standard deviation, σ, and (d) Cv for the flood volume. Source: Dimitriadis et al. (2016b).....68

Figure 28: qq-plots and box-plots of the water depth of the channels' (a-b) upstream and (c-d) downstream cell/section as well as of the (e-f) total volume of the flooded area. Note that the water depths and flood volume are first standardized (i.e., the residual from their average value is divided with their standard deviation). Source: Dimitriadis et al. (2016b).....71

Figure 29: Contour maps of (a) water depths, (b) lateral flows, and (c) longitudinal flows produced by LISFLOOD-FP (unsteady), for $Q = 2500$ m^3/s, $nf = 0.10$, $nc = 0.07$, $gl = 2.5\%$, $gf = 2.8\%$ and $c = 50$ m. Source: Dimitriadis et al. (2016c). 72

Figure 30: Variation coefficients of the flood volume vs. grouped input variables (coloured solid lines), averaged per model (coloured dashed lines) and averaged (overall) for all models (black line). Note that each variation coefficient is estimated from 500 (=1500/3), 1500 and 9000 (= 1500×6) values, respectively. Also note that the overall variation coefficients of HEC-RAS, for steady and unsteady conditions, coincide with each other. Source: Dimitriadis et al. (2016c). 74

Figure 31: (a) Example of loss of low frequency information caused by the application of the windowing technique, in a time-series provided by the Johns Hopkins University as well as the maximum cross correlations between the partitioned segments; (b) 1D autocorrelation function derived from the 3D power spectrum model (with parameters based on the fitting of the windowed 1D power spectrum with 1000 segments: $c\mathrm{E} = 2.5\ \mathrm{m} - 2,\ p = 4,\ c\mathrm{I} = 13.0\ \mathrm{m3/s2},\ c\mathrm{D} = 2 \times 10 - 4\ \mathrm{m}$); a Markov autocorrelation function, i.e., $e - \tau/q$, for reasons of comparison; and the corresponding (to the windowed 1D power spectrum with 1000 segments) empirical autocorrelation function. Source: Dimitriadis et al. (2016a). 80

Figure 32: Expected power spectrum resulted from a combination of a Markov and a gHK process (source: Dimitriadis et al., 2016a). 81

Figure 33: (a) Power spectra and (b) corresponding autocovariances, in continuous time as well as their expected values, with varying number of records n for a gHK process (source: Dimitriadis et al., 2016a). 82

Figure 34: Expected power spectra of a gHK process, with varying q/Δ (where Δ the sampling time interval). Source: Dimitriadis et al., 2016a. 83

Figure 35: [left] Standardization scheme for grid-turbulence data, where μ and σ are the mean and standard deviation, r is the distance from the grid, with the first 16 time series corresponding to transverse points abstaining $r = 20S$ from the source, the second 4 to $r = 30S$, the third 4 to $40S$ and the last 16 to $48S$, with S = 0.152 m the size of the grid; [right] empirical probability density function of the overall standardized time series (observed) along with that from a single synthetic time series produced by the SMA scheme to preserve the first four moments (simulation); for comparison the theoretical distributions N(0,1), skew normal and ME constrained on the four moments (corresponding weights for the ME distribution: 15%, 51%, 21% and 13%). Source: Dimitriadis and Koutsoyiannis (2017). 85

Figure 36: The empirical, true and expected values of the climacogram [upper left], CBF [upper right], CBS [lower left] and power spectrum [lower right] along with some important logarithmic slopes. Source: Dimitriadis and Koutsoyiannis (2017). 87

Figure 37: Empirical and simulated 3rd order structure function [left] and kurtosis coefficient [right] of the velocity increments vs. lag. Source: Dimitriadis and Koutsoyiannis (2017). 88

Figure 38: Empirical and simulated structure function for various orders of the velocity increments vs. lag. Source: Dimitriadis and Koutsoyiannis (2017). 88

Figure 39: Photograph of the experimental set-up on turbulent buoyant jets at the laboratory of Hydraulics at NTUA. 89

Figure 40: Initial concentration C_o vs. the initial intensity I_o for the red (top) and green (bottom) RGB intensity. Source: Dimitriadis et al. (2010).....91

Figure 41: From left to righr and top to bottom: (a) Raw picture taken from the video-camera for the TBJH01 experiment, (b) gray-scale and (c) RGB format of the raw picture, (d) average gray-scale and (e) RGB image of the experiment, and (f) average RMS image of the experiment. Source: Dimitriadis and Papanicolaou (2010).....93

Figure 42: Time series of the excess temperature over the maximum temperature at the jet centerline for the TBHJ01 experiment (Table 19). Source: Dimitriadis and Papanicolaou (2010).....93

Figure 43: Dimensionless average and standard deviation of the RGB intensity (1st and 3rd pictures) and of the red RGB intensity (2nd and 4th plots), for the experiment TBVJ01a. Source: Dimitriadis and Papanicolaou (2010).....94

Figure 44: True (unbiased, pink line) and empirical (biased, blue line) Hurst parameter along the jet axis. Source: Dimitriadis and Papanicolaou (2010).....94

Figure 45: Part of the wind speed time-series provided by NCAR/EOL (http://data.eol.ucar.edu/).....95

Figure 46: From top to bottom and from left to right: Averaged empirical (a) climacograms and autocovariances, (b) CBV and variograms, (c) CBS and power spectra (for the three sets) and (d) qq-plot of empirical pdf *vs* standard Gaussian pdf (for the original time-series), along with modelled distribution density function (all parameters in m/s).....96

Figure 47: True, expected and empirical (averaged) climacogram values for the wind process stochastic simulation.....97

Figure 48: Three precipitation episodes provided by the Hydrometeorology Laboratory at the Iowa University.....97

Figure 49: (a) Averaged empirical climacograms and autocovariances, (b) CBV and variograms, (c) CBS and power spectra for T1, T2 and T3, and (d) true, expected and empirical (averaged) climacogram values for the rainfall processes stochastic simulation. Source: Dimitriadis et al. (2016a).....98

Figure 50: Empirical, modelled and simulated marginal distributions [upper left] and climacograms [upper right] for the standardized precipitation process; the mode and several other essential statistical measures of the standardized climacograms estimated from 10^3 synthetic timeseries (in the figure we depict only 50 empirical climacograms) [lower left]; a 3000 days window of the observed precipitation record along with a simulated one [lower right]. Source: Dimitriadis and Koutsoyiannis (2017).....100

Figure 51:Empirical mean (v_m) vs. standard deviation of the nine timeseries along with the fitted model [upper left]; the empirical, model and simulated marginal distributions [upper right] and climacograms [lower left] for the standardized wind process; a 1000-day window of the observed

standardized wind process in Kos island along with a standardized simulated one [lower right]. Source: Dimitriadis and Koutsoyiannis (2017). 103

Figure 52: (upper) Distribution of the wind speed stations over the globe; (middle) sketch about the selection of the stations in the analysis; (lower) evolution of the frequency of measured extremes in the stations (where the 'start' year denotes the first operational year of the station and the 'first' and 'last' year denote the first and last year that an extreme value was recorded, respectively). Source: Koutsoyiannis et al. (2017). 104

Figure 53: Standard deviation vs mean (upper) and coefficient of kurtosis vs. coefficient of skewness of all time series (source: Koutsoyiannis et al., 2017). 106

Figure 54: Probability density function of the medium scale time series along with theoretical and Monte Carlo generated distributions (source: Koutsoyiannis et al., 2017). 107

Figure 55: Probability density function of the velocity of grid-turbulent data (small) and of the wind speed of the medium and global scale time series along with fitted theoretical distributions (source: Koutsoyiannis et al., 2017). 108

Figure 56: Climacogram of the wind speed process estimated from the medium and global series (source: Koutsoyiannis et al., 2017). 108

Figure 57: Locations of the selected hourly time series of air temperature from the global database along with the Koppen climatic zones. Source: Lerias et al. (2016). 109

Figure 58: Coefficient of skewness vs. coefficient of kurtosis for the 90% of the macro-scale temperature time series (source: Koutsoyiannis et al., 2017). 110

Figure 59: Climacogram of the normalized temperature for the micro-scale time series (small) and the set of hourly air temperature time series (global; upper: average climacogram; lower: climacograms of 100 different time series), compared to the fitted model (true and expected). Source: Koutsoyiannis et al. (2017). 110

Figure 60: Climacogram-based spectrum of the normalized temperature for the micro-scale time series (small) and the set of hourly air temperature time series (global; average from all time series), compared to the fitted model (true). Source: Koutsoyiannis et al. (2017). 111

Figure 61: (a) temperature and (b) dew point timeseries and HK model for a station located in Dallas, USA; (c) wind speed timeseries and HK model for a station located in Winter Trail, Alaska; and (d) precipitation timeseries and HK model for a station located in North-East Australia. Source: Dimitriadis et al. (2016e) and references therein. Source: Dimitriadis et al. (2016d). 113

Figure 62: Prediction intervals for the examined station described in the previous figure and the overall prediction error for (a) temperature, (b) dew point, (c) wind speed and (d) precipitation. Source: Dimitriadis et al. (2016e) and references therein. Source: Dimitriadis et al. (2016d). 114

1 Introduction

1.1 The complexity of nature

The word “complex” is attributed to “a whole comprised of parts” and comes from Latin but has been re-borrowed from ancient Greek (originated from the verb “συμπλέκω”). It constitutes of the Latin preposition “com” or “cum”, which is related to the Greek preposition “συν” and is used, usually at the beginning of a word, to declare union, ensemble etc.; and the Latin verb “plectere” which comes from the Greek verb “πλέκω” meaning “weave”, “twine” etc. In recent times, we characterize a process as complex if it is difficult to analyze or explain it in a simple way. Climate dynamics is characterized by high complexity since it is comprised by numerous geophysical processes interacting with each other in a non-linear way. However, most of the involved processes (will) remain unknown since it is impossible to fully analyze such complicated systems. Nevertheless, even if we could determine a set of physical laws that describe in full detail the complexity of climate dynamics it would be impossible to combine the equations for the purpose of predictability due to the existence of chaos, i.e., a non-predictive sensitivity to initial conditions. For example, consider the analysis of Poincaré (1890) for the three-body problem, where chaotic behaviour emerges from the equations of classical mechanics when studying the interacting gravitational forces between three bodies (e.g., planets). Similar results came into sight from Lorenz (1963) while applying a simplified set of equations for the analysis of atmospheric dynamics. E.N. Lorenz came across to the idea that non-linear dynamic systems may have a finite limit of predictability (which for weather prediction he estimated this limit to be around two weeks), even if the model is perfect and even if the initial conditions are known almost perfectly. Later on, numerous methodologies were initiated not for predicting the exact outcome of a non-linear system, which as we already explained may be trivial, but for rather estimating the limits of this prediction through an alternative approach of stochastic analysis.

1.2 The stochastic approach

The scientific interest on Stochastics has increased over the last decades as an alternative way of deterministic approaches, to model the so-called random, i.e., complicated, unexplained or unpredictable, fluctuations recorded in non-linear geophysical processes. Randomness can emerge even in a fully deterministic system with non-linear dynamics (Koutsoyiannis, 2010). Thus, Stochastics help develop a unified perception for all natural phenomena and expel dichotomies like random *vs*. deterministic. Particularly, there is no such thing as a ‘virus of randomness’ that infects some phenomena to make them random, leaving other phenomena uninfected. It seems that rather both randomness and predictability coexist and are intrinsic to natural systems which can be deterministic and random at the same time, depending on the prediction horizon and the time scale (Dimitriadis et al., 2016b). On this basis, the uncertainty in a geophysical process can be both aleatory (alea = dice) and epistemic (as in principle we could know perfectly the initial conditions and the equations of motion but in practice we do not). Therefore, dichotomies such as ‘deterministic vs. random’ and ‘aleatory vs. epistemic’ may be false ones and may lead to paradoxes.

The line distinguishing whether determinism (i.e. predictability) or randomness (i.e. unpredictability) dominates is related to the scale (or length) $l(\varepsilon)$ of the time-window within which the future state deviates from a deterministic prediction by an error threshold ε. For errors smaller than ε, we assume that the system is predictable within a time-window $l(\varepsilon)$ and for larger errors unpredictable (Dimitriadis and Koutsoyiannis, 2017). Therefore, by applying the concept of stochastic analysis we identify the observed unpredictable fluctuations of the system under investigation with the variability of a devised stochastic process. This stochastic process enables generation of an ensemble of realizations, while observation of the given natural system can only produce a single observed time series (or multiple ones in repeatable experiments).

1.3 The Hurst-Kolmogorov dynamics

The high complexity and uncertainty of climate dynamics has been long identified through plain observations as well as extended analyses of hydrometeorological processes such as temperature, humidity, surface wind, precipitation, atmospheric pressure, river discharges etc. Particularly, all these processes seem to exhibit high unpredictability due to the clustering of events, an example is large periods of high annual precipitation which are usually followed by large periods of annual droughts. Note that this behaviour should not be confused with seasonal effects that correspond to sub-annual scales. Interestingly, this clustering behaviour has been first identified in Nature by Hurst (1951) while analyzing water levels from the Nile for optimum dam design. However, the mathematical description and analysis of this behaviour through a power-law autocorrelation function (vs. lag) is attributed to Kolmogorov (1940) who developed it earlier while studying turbulence. To give credits to both scientists Koutsoyiannis (2010) named this behaviour as Hurst-Kolmogorov (HK) behaviour.

1.4 From the microscopic analysis to the macroscopic observation

In order to properly study the aforementioned clustering of events and, in general, the stochastic behaviour of hydrometeorological processes we would naturally require copious measurements in annual scale. Unfortunately, large lengths of high quality annual data are hardly available in observations of hydrometeorological processes (Koutsoyiannis, 2014). However, the microscopic processes driving and generating the hydrometeorological ones are governed by turbulent state, e.g., as identified in the field of Hydrology by Mandelbrot and Wallis (1968). For example, the size of drops which is highly linked to the form and intensity of precipitation events is strongly affected by the turbulent state of small scale atmospheric wind (Falkovich et al., 2002). Also in a physical-basis the rain rate is found to be a function of gradient level wind speed, the translational velocity of the tropical cyclone, the surface drag coefficient, and the average temperature and saturation ratio inside the tropical cyclone boundary layer (Langousis and Veneziano, 2009). Another example is the multifractal similarities between rainfall and turbulent atmospheric convection (Veneziano et al., 2006). Therefore, by studying turbulent phenomena (or other related small scale processes) in situ we may be able to understand certain aspects of the related macroscopic processes in field. Additional advantages of studying macroscopic processes in field through the microscopic

turbulent ones in situ could be the recording of very long time series, the high resolution of records and the controlled environment of a laboratory.

1.5 Scientific innovations of the thesis

In this thesis, the sections are organized as follows: (1) in the first section we introduce basic concepts of the thesis, such as the HK dynamics and we discuss on the motivation and the scientific interest of the thesis mostly from an engineer point of view; (2) in the second section we introduce and develop the statistical tools as well as the methods used in the thesis; (3) in the third section we introduce and develop the generation algorithms that are extensively used in the thesis; (4) in the fourth section we discuss on how and why the HK dynamics are related to uncertainty as well as on the dichotomy between randonmness and determinism, with plenty applications on deterministic and more complex processes; (5) in the fifth section we conduct a stochastic analysis on an isotropic and an anisotropic turbulent process and we discuss on some identified similarities to hydrometeorological processes; (6) in the sixth section we apply a stochastic analysis on several hydrometeorological processes from a local to a global scale and we show how simple stochastic models can simulate certain challenging aspects such as long-term persistence, and (7) in the seventh section we summarize our results by highlighting the most important ones, and we discuss on future investigations.

The major innovations of the thesis are the following: (a) further development and extensive application to numerous processes of the classical second-order stochastic framework (sections 2.1 to 2.3 and 2.5) and related monoschedastic processes; (b) the estimation of the dimensionless statistical error through Monte-Carlo analysis for a variety of Markov and HK models, regarding the power spectrum, autocovariance and climacogram (section 2.4.5); (c) the exact mathematical expression of the statistical bias of the autocovariance, variogram and power spectrum classical estimator as a function of the theoretical autocovariance and climacogram (sections 2.3.4 and 2.3.5); (d) the introduction of the Markov process for a different time interval and response time, and the expressions for its generation through an ARMA(1,1) model (section 2.4.1); (e) the further development of the Sum of Autoregressive (SAR) and Moving Average (SARMA) schemes that can generate a large variety of Gaussian processes approximated by a finite sum of AR(1) or ARMA(1,1) processes (section 3.2); (f) the further development of the Symmetric-Moving-Average (SMA) scheme that can explicitly (or implicitly) generate any process second-order dependence structure, approximate (or exactly) preserve any marginal distribution function as well as simulate certain aspects of the intermittent behaviour (sections 3.3); (g) the introduction and application of an extended HK model to various turbulent and hydroclimatic processes (sections 2.4.3, 5.3, 6.3 and 6.4); (h) estimation of the Hurst parameter based on the Köppen-Geiger climatic-classification for numerous hydroclimatic processes from global databases (section 6.5); and (i) the further development of the multi-dimensional classical second-order stochastic framework and HK process (section 3.4).

Incidental contributions and moderate innovations of this thesis are: (a) several illustrative comparisons between complex natural as well as purely deterministic processes and the emerging statistical uncertainty (section 4); (b) the further development and application of analogue and

stochastic prediction algorithms based on the climacogram (sections 3.5 and 4); (c) the estimation of the most uncertain parameters in flood inundation modelling based on commonly-used hydraulic models and on benchmark geometries (section 4.2); (d) the further development of how to deal with discretization and statistical bias in stochastic modelling by selecting appropriate climacogram-based estimators for the identification of the second-order dependence structure of a process in case of the analysis of a single time series and of several time series of the same process with different lengths and identical lengths (sections 2.5 and 6).

2 Definitions, methods and notation for stochastic analysis

In this section, we present the definitions and notations of the concepts used in the thesis as well as the statistical metrics, methods and models for the stochastic analysis.

2.1 The definition of Stochastics and related concepts

A.N. Kolmogorov (1931) is the first to mathematically define how a process can be stochastically determined based on the theory of continuous-time probability function (rather than discrete), a concept first visualized and applied by Bachelier (1900) while working on the evolution of price for his PhD thesis (Koutsoyiannis and Dimitriadis, 2016). Kolmogorov (1931) distinguishes a purely deterministic from a stochastic process by correspondingly letting a preceding state to uniquely define a subsequent state rather than by permitting only a certain probability of a possible event of a subsequent state to occur. Alternatively, the change of a physical system is deterministically (stochastically) defined if (the probability distribution for) every subsequent state is decisive by the knowledge of a preceding state. Therefore, a deterministic (stochastic) physical process can exactly predict (the probability of an event of) a future state given the present and/or past state. The purpose of stochastic analysis, or else the mathematical field of Stochastics, is to subject a natural process to a stochastic process, or in other words to predict real changes using a stochastic (i.e., not purely deterministic) mathematical scheme. Two additional concepts can arise from the above definition of Stochastics, these of stationarity and ergodicity (Koutsoyiannis and Montanari, 2015). The main scope of a stochastic analysis is the identification of the most parsimonious model in continuous time that adequately preserves the physical characteristics of the natural process in discrete time along with its statistical estimates from observed timeseries in order to investigate its future variability through the generation of synthetic timeseries (Figure 1).

The analysis presented in this thesis is based on both the assumption of (cyclo)stationarity (although it can be easily expanded to non-stationary processes following the methodology described in section 3.3) and ergodicity, so that we can estimate all the desired characteristics of the marginal distribution, dependence structure and combination thereof (e.g., intermittent behaviour) from a single time series and simulate all periodicities (e.g., seasonal, diurnal) of the process. Another important concept used in most of the applications is the homogenization, where all time series corresponding to a single physical process are treated as realizations of a single mathematical process, with a single marginal distribution and dependence structure. Therefore, by a simple homogenization scheme (which depends entirely on the expression of both the marginal

distribution and dependence structure) we can combine all related time series to a single one with a much larger length and thus, towards a better estimation of the statistical and stochastic characteristics (see sections 5 and 6 for such applications). Note that the homogenization should not be confused with the concept of standardization which corresponds to the dimensionalization of a process by simply dividing it with a parameter or to the concept of normalization which can be only applied to normal (or close to normal) processes in order to transform them properly to follow exactly (or approximately) the standard N(0,1) distribution.

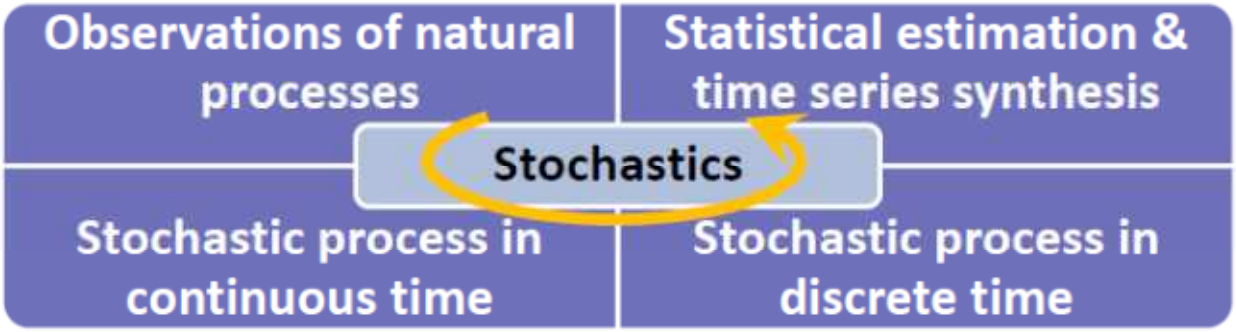

Figure 1: The steps for a stochastic analysis (source: Koutsoyiannis and Dimitriadis, 2016).

2.2 Observing a natural stochastic process

A stochastic analysis should imitate the physical procedure of data collection as much as possible rather than strictly the observations. *Nature is the most beautiful Being and although She might let you observe She will never reveal Her true secrets.* Observation of natural processes includes numerous technical and unsurpassed obstacles, mostly related to hydrometeorological and engineering processes, which are introduced by the complexity of numerous known and unknown interacting processes, such as (known) instrumental errors and the (unknown) hydroclimatic variability. This is of high importance in stochastic analysis and a stochastic analyst should be cautious with data as well as the technical properties of the instrument used for data collection in order not to end up simulating, without knowing it, the limitations of the instrument rather than the physical process.

Although natural processes evolve in continuous time all observed timeseries are subject to a response time $\Delta > 0$ of the instrument and a sampling time interval $D \geq \Delta$, often fixed by the observer. The corresponding discretized mathematical process can be estimated by averaging the continuous one over a time scale $\Delta \geq 0$ for every time interval $D \geq \Delta$. It should be noted that although the case $\Delta = 0$ is technically impossible, it is theoretically possible and can be used as an approximation for instruments of high resolution. Thus, the discrete time stochastic process $\underline{x}_i^{(\Delta,D)}$ can be calculated from the continuous one $\underline{x}(t)$ as:

$$\underline{x}_i^{(\Delta,D)} = \frac{\int_{(i-1)D}^{(i-1)D+\Delta} \underline{x}(\xi)\mathrm{d}\xi}{\Delta} \tag{1}$$

where $i \in [1,n]$ is an index representing discrete time, $n = \left\lfloor \frac{T-\Delta}{D} \right\rfloor + 1$ is the total number of realizations and $T \in [\Delta, \infty)$ is the time length of the realization sample (Figure 2). Note that underlined quantities denote random variables.

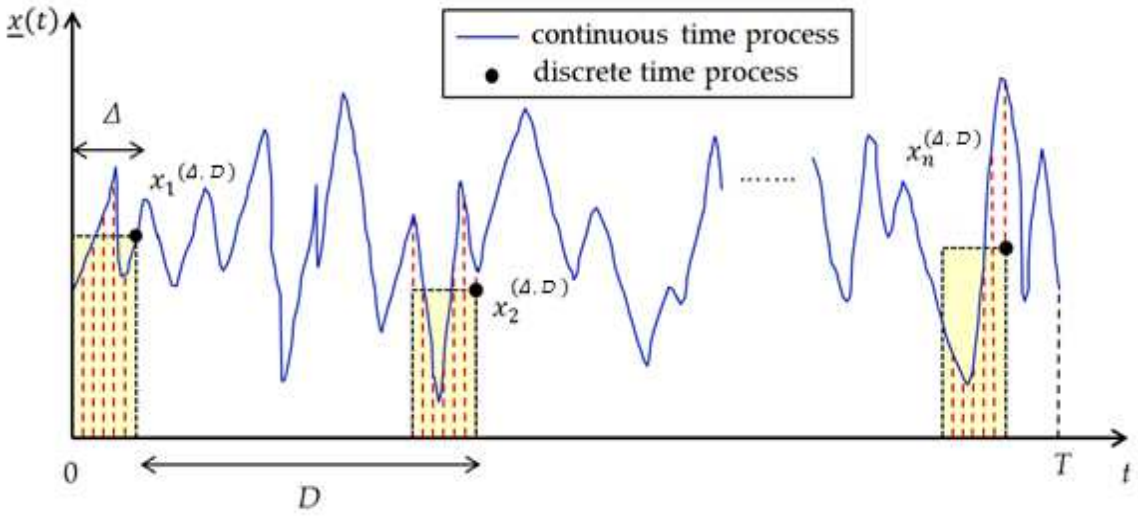

Figure 2: An example of realization (blue line) of a continuous time process $\underline{x}$ and a sample of $x_i^{(\Delta,D)}$ realizations (black dots) of the discretized process $\underline{x}_i^{(\Delta,D)}$ averaged at time scale Δ, with time intervals D and for a total period T (source: Dimitriadis et al., 2016a).

2.3 Stochastic metrics for identification of a stochastic process

During a stochastic analysis we first have to visualize certain behaviours of the natural process using the appropriate stochastic metrics, then to combine them for the identification of the mathematical process and finally, to estimate the parameters of the latter. For simplicity, we can investigate separately the probability distribution function and the dependence structure of the process.

2.3.1 Most common measures for the marginal characteristics of a process

The marginal characteristics of the process can be entirely described by the probability distribution function, i.e., $F(\underline{x}) := P(\underline{x} \le x)$, where $\underline{x}$ is the random process and x is a realization of the process. In this thesis, we also use the tail probability distribution function, i.e., $F^*(\underline{x}) := 1 - F(\underline{x})$, and the density distribution function, i.e., $f(\underline{x}) := \mathrm{d}F(\underline{x})/\mathrm{d}x$. The distribution function is estimated through $\hat{F}(\underline{x}) = n'/g(n)$, where n' is the empirical number of occurrence with values less or equal to x, n is the total number of observations, and typically $g(n) = n + 1$ is known as the Weibull estimator. For the density of the distribution function we use the forward difference quotient, i.e., $\hat{f}(\underline{x}) = \left(\hat{F}(\underline{x}+h) - \hat{F}(\underline{x})\right)/h$, where h is the length of the interval over which f is estimated. Note that the estimation of a marginal characteristic of a process through the distribution function has the drawback of preference of the function $g(n)$, whereas through the density distribution function that of the type of the derivative discretization. Other important marginal characteristics of the

process are the statistical moments (raw, central etc.) that can be estimated directly from the distribution density function, i.e. for the central ones $\mathrm{E}\left[\left(\underline{x}-\mu\right)^{i}\right] \coloneqq \int_{-\infty}^{\infty}\left(\underline{x}-\mu\right)^{i} f\left(\underline{x}\right) \mathrm{d}\underline{x}$., for $i > 1$, where $\mu = \mathrm{E}\left[\underline{x}\right]$ is the mean of the process. In case of large samples we can either use the above definition (i.e., provided that we know the theoretical distribution $f\left(\underline{x}\right)$ of the process) in discretized form or the classical estimators for the sample central moments, whereas for small samples lack of information on $f\left(\underline{x}\right)$ could lead to poor estimation of the sample moments.

2.3.2 Most common and uncommon metrics for the dependence structure of a process

For the second order dependence structure (we will refer to this as dependence structure) we present several metrics based on the correlation between variables as a function of lag as well as on the variance of averaged variables as a function of scale. The first presented metric is the climacogram $\gamma(k)$, i.e., the variance of the scaled process i.e., $\frac{1}{k}\int_{0}^{k}\underline{x}(t)\mathrm{d}t$ vs. scale k, where $k = \kappa\Delta$ is the continuous-time scale in time units and κ the dimensionless discrete one, assuming that $\Delta = D$ is a time unit that is used for discretization (Koutsoyiannis, 2000). The climacogram is directly linked to the autocovariance $c(h)$, i.e., $c(h) = \frac{1}{2}\partial^{2}\left(h^{2}\gamma(h)\right)/\partial h^{2}$, where h is the continuous-time lag in time units, and its power spectrum, i.e., $s(w) \coloneqq 2\int_{-\infty}^{\infty} c(h)\cos(2\pi w h)\,\mathrm{d}h$, where w is the continuous frequency in reverse time units (Koutsoyiannis, 2013). Thus, each of these three stochastic tools contains exactly the same information and either can be used for the estimation of the dependence structure. However, it has been shown that the former provides better estimates than the other two (Dimitriadis and Koutsoyiannis, 2015a) and therefore, all applications here are based on the climacogram. In Tables 1-3, we introduce the definitions of several climacogram-based measures and in Tables 4-6, the corresponding autocovariance-based ones. We show the definitions in case of a stochastic process in continuous time and in discrete time, widely used estimators and estimations based on the latter estimators, all expressed as a function only of the climacogram (Dimitriadis et al., 2016a).

2.3.3 Climacogram-based metrics for the dependence structure as a function of scale

First, we present the climacogram definition and expressions for a process in continuous and discrete time, along with the properties of its estimator (Table 1), for comparison with the autocovariance function.

Table 1: Climacogram definition and expressions for a process in continuous and discrete time, along with the properties of its estimator. Source: Dimitriadis et al., (2016a).

Type	Climacogram	
continuous	$\gamma(k) := \mathrm{Var}\left[\int_0^k \underline{x}(y)\mathrm{d}y\right]/k^2$	(T1-1)
	where $k \in \mathbb{R}^+$	
discrete	$\gamma^{(\Delta)}(k) := \frac{\mathrm{Var}\left[\sum_{l=1}^{\kappa} \underline{x}_l^{(\Delta)}\right]}{\kappa^2} = \gamma(\kappa\Delta)$	(T1-2)
	where $\kappa \in \mathbb{N}$ is the dimensionless scale for a discrete time process	
classical estimator	$\hat{\underline{\gamma}}^{(\Delta)}(k) = \frac{1}{n-1}\sum_{i=1}^{\lfloor n/\kappa \rfloor}\left(\frac{1}{\kappa}\left(\sum_{l=\kappa(i-1)+1}^{\kappa i} \underline{x}_l^{(\Delta)}\right) - \frac{\sum_{l=1}^{n} \underline{x}_l^{(\Delta)}}{n}\right)^2$	(T1-3)
expectation of classical estimator	$\mathrm{E}\left[\hat{\underline{\gamma}}^{(\Delta)}(k)\right] = \frac{1-\gamma(n\Delta)/\gamma(\kappa\Delta)}{1-\kappa/n}\gamma(\kappa\Delta)$	(T1-4)

Note that the climacogram can be estimated through other methods such as raw moments, L-moments etc. but for convenience in this thesis we choose the central classical moment estimator. Furthermore, we introduce a climacogram-based variogram (CBV) for comparison with the classical variogram defined in Table 5.

Table 2: Climacogram-based variogram (CBV) definition and expressions for a process in continuous and discrete time, along with the properties of its estimator (source: Dimitriadis et al., 2016a).

Type	Climacogram-based variogram	
continuous	$\xi(k) := \gamma(0) - \gamma(k)$	(T2-1)
discrete	$\xi_{\mathrm{d}}^{(\Delta)}(\kappa) := \gamma(0) - \gamma(\kappa\Delta)$	(T2-2)
classical estimator	$\hat{\underline{\xi}}_{\mathrm{d}}^{(\Delta)}(\kappa) = \gamma(0) - \hat{\underline{\xi}}_{\mathrm{d}}^{(\Delta)}(\kappa)$	(T2-3)
expectation of classical estimator	$\mathrm{E}\left[\hat{\underline{\xi}}_{\mathrm{d}}^{(\Delta)}(\kappa)\right] = \gamma(0) - \mathrm{E}\left[\hat{\underline{\xi}}_{\mathrm{d}}^{(\Delta)}(\kappa)\right]$	(T2-4)

Note that CBV includes the process variance at scale 0, i.e., $\gamma(0)$, and so, in cases where $\gamma(0)$ is infinite, we can use a slightly different estimator with $\gamma(\Delta)$ instead. Finally, we introduce a

climacogram-based spectrum (CBS) for comparison with the classical power spectrum (Koutsoyiannis, 2013) defined in Table 3.

Table 3: Climacogram-based spectrum (CBS) definition and expressions for a process in continuous and discrete time, along with the properties of its estimator (source: Dimitriadis et al., 2016a).

Type	Climacogram-based spectrum	
continuous	$\psi(w) := \frac{2\gamma(1/w)}{w}\left(1-\frac{\gamma(1/w)}{\gamma(0)}\right)$	(T3-1)
	where $w \in \mathbb{R}$ is the frequency for a continuous time process (in inverse time units) and is equal to $w=1/k$.	
discrete	$\psi_{\mathrm{d}}^{(\Delta)}(\omega) := \frac{2\gamma(1/\omega)}{\omega}\left(1-\frac{\gamma(1/\omega)}{\gamma(0)}\right)$	(T3-2)
	where $\omega \in \mathbb{R}$ is the frequency for a discrete time process (dimensionless; $\omega = w\Delta$)	
classical estimator	$\underline{\hat{\psi}}_{\mathrm{d}}^{(\Delta)}(\omega) = \frac{2\gamma(1/\omega)}{\omega}\left(1-\frac{\gamma(1/\omega)}{\gamma(0)}\right)$	(T3-3)
expectation of classical estimator	$\mathrm{E}\left[\underline{\hat{\psi}}_{\mathrm{d}}^{(\Delta)}(\omega)\right] = \frac{2\mathrm{E}[\gamma(1/\omega)]}{\omega}\left(1-\frac{\mathrm{E}[\gamma(1/\omega)]}{\gamma(0)}-\frac{\mathrm{Var}[\gamma(1/\omega)]}{\gamma(0)\mathrm{E}[\gamma(1/\omega)]}\right)$	(T3-4)

Note that in cases where $\gamma(0)$ is infinite, CBS simplifies to $\frac{2\gamma(1/w)}{w}$. Another useful metric is the dimensionless-climacogram which is defined as $\gamma(k)/\gamma(0)$ (to be used as an alternative tool to the autocorrelation function).

2.3.4 Autocovariance-based metrics for the dependence structure as a function of lag

The climacogram is useful to measure the variance of a process among scales (the kinetic energy, in case the variable under consideration is the velocity), and has many advantages in stochastic model building, namely small statistical as well as uncertainty errors (Dimitriadis and Koutsoyiannis, 2015a). It is also directly linked to the autocovariance function $c(h)$, h being the continuous-time lag, by the following equations (Koutsoyiannis, 2013):

$$\gamma(k) = 2\int_0^1 (1-x)c(xk)\mathrm{d}x \tag{2}$$

$$c(h) = \frac{\partial^2(h^2\gamma(h))}{2\partial h^2} \tag{3}$$

The autocovariance definition and expressions for a process in continuous and discrete time, along with the properties of its estimator can be seen in Table 4.

Table 4: Autocovariance definition and expressions for a process in continuous and discrete time, along with the properties of its estimator (source: Dimitriadis et al., 2016a).

Type	Autocovariance	
continuous	$c(h) := \mathrm{cov}\left[\underline{x}(t), \underline{x}(t+h)\right]$	(T4-1)
	where $h \in \mathbb{R}$ is the lag for a continuous time process (in time units)	
discrete	$c_{\mathrm{d}}^{(\Delta)}(v) \coloneqq \frac{\Delta^2[v^2\gamma(v\Delta)]}{2\Delta[v^2]} =$	(T4-2)
	$= \frac{1}{2}\left((v+1)^2\gamma\left((v+1)\Delta\right) + (v-1)^2\gamma\left((v-1)\Delta\right) - 2v^2\gamma(v\Delta)\right)$	
	where $v \in \mathbb{Z}$ is the lag for the process at discrete time (dimensionless)	
classical estimator	$\hat{\underline{c}}_{\mathrm{d}}^{(\Delta)}(v) = \frac{1}{\zeta(v)}\sum_{i=1}^{n-v}\left(\underline{x}_i^{(\Delta)} - \frac{1}{n}\left(\sum_{l=1}^{n}\underline{x}_l^{(\Delta)}\right)\right)\left(\underline{x}_{i+j}^{(\Delta)} - \frac{1}{n}\left(\sum_{l=1}^{n}\underline{x}_l^{(\Delta)}\right)\right)$	(T4-3)
	where $\zeta(v)$ is usually taken as: n or $n-1$ or $n-v$.	
expectation of classical estimator	$\mathrm{E}\left[\hat{\underline{c}}_{\mathrm{d}}^{(\Delta)}(v)\right] = \frac{1}{\zeta(v)}\left((n-v)c_{\mathrm{d}}^{(\Delta)}(v) + \frac{v^2}{n}\gamma(v\Delta) - v\gamma(n\Delta) - \frac{(n-v)^2}{n}\gamma\left((n-v)\Delta\right)\right)$ *	(T4-4)

* For proof see in (Dimitriadis and Koutsoyiannis, 2015a).

We then introduce the classical variogram or else the second-order structure function (Table 5).

Table 5: Variogram definition and expressions for a process in continuous and discrete time, along with the properties of its estimator (source: Dimitriadis et al., 2016a).

Type	Variogram	
continuous	$v(h) := c(0) - c(h)$	(T5-1)
discrete	$v_{\mathrm{d}}^{(\Delta)}(v) := \gamma(\Delta) - c_{\mathrm{d}}^{(\Delta)}(v)$	(T5-2)
classical estimator	$\underline{\hat{v}}_{\mathrm{d}}^{(\Delta)}(v) = \underline{\hat{\gamma}}(\Delta) - \underline{\hat{c}}_{\mathrm{d}}^{(\Delta)}(v)$	(T5-3)
expectation of classical estimator	$\mathrm{E}\left[\underline{\hat{v}}_{\mathrm{d}}^{(\Delta)}(v)\right] = \mathrm{E}\left[\underline{\hat{\gamma}}(\Delta)\right] - \mathrm{E}\left[\underline{\hat{c}}_{\mathrm{d}}^{(\Delta)}(v)\right]$	(T5-4)

2.3.5 The power spectrum

Finally, we define the power spectrum (or else spectral density) that was introduced as a tool to estimate the distribution of the power (i.e., energy over time) of a velocity sample over frequency, more than a century ago by Schuster (Stoica and Moses, 2005, p. xiii). Since then, various methods have been proposed and used to estimate the power spectrum, via the Fourier transform of the time series (periodogram) or its autocovariance or autocorrelation functions (for more information on these methods see in Stoica and Moses, 2005, ch. 2, and Gilgen, 2006, ch. 9). Most common (and also used in this thesis) is that of the autocovariance which corresponds to the definition of the power spectrum of a stochastic process. However, this accurate mathematical definition lacks immediate physical interpretation since the Fourier transform of a function is nothing more than a mathematical tool to represent the function in the frequency domain in order to identify any periodic patterns which are not easily tracked in the time domain. Historically the power spectrum is defined in terms of the Fourier transform of the process $\underline{x}(t)$ by taking the expected value of the squared norm of the transform for time tending to infinity, which for a stationary process converges to the Fourier transform of its autocovariance (this is known as the Wiener- Khintchine theorem after Wiener, 1930, and Khintchine, 1934). Both definitions can be used for the power spectrum; however the latter is simpler and more operational and has been preferred in modern texts (e.g. Papoulis and Pillai, 1991, ch. 12.4).

Several studies that evaluate the statistical estimator of the power spectrum conclude that its major disadvantage is that of its large variance (Stoica and Moses, 2005, p. xiv). Notably, this variance is not reduced with increased sample size (Papoulis and Pillai, 1991, p. 447). To remedy this, several mathematical smoothing techniques (e.g. windowing, regression analysis, see Stoica and Moses, 2005, ch. 2.6) have been developed. In cases of short datasets, trend-line approaches are most commonly used to obtain a very rough estimation of the model behaviour or rules of thumb to distinguish exponential and power-type behaviours (e.g., Fleming, 2008). In cases of long datasets, the most commonly used approach is the windowing (data partitioning), also known as the Welch

approach, where a certain window function (the simplest of which is the Bartlett window) is applied to nearly independent segments. In the latter method, one has first to divide the sample into several segments (but only after insuring these segments have very small correlations between them), to calculate the power spectrum for each segment and then to estimate the average. Assuming that the process is stationary, this average will be the power spectrum estimate. Unfortunately, the more segments we divide the sample into, the more the cross-correlations between segments are increasing as well as the more we lose in low frequency values (since the lowest frequency is determined by the length of the segments). Thus, this method could be indeed a robust one, but only for a very long sample (which is a rare case in geophysics), only when there is no interest in the low frequency values (which can reveal large-scale behaviours) and only for an unbiased power spectrum estimator or at least for an 'a priori' known bias, e.g. via an analytical equation (which, as can observe in Table 6, is rarely the case). Based on these limitations, Dimitriadis et al. (2012) and Koutsoyiannis (2013) provided some examples where this smoothing technique fails to detect the large scale behaviour (i.e., HK behaviour), gives small scale trends that are completely different from the ones characterizing the stochastic model and have several numerical calculation problems that could cause misinterpretation. These all are due to the fact that the power spectrum estimator has a large variance, is biased and it is difficult to estimate these analytically. Nevertheless, the power spectrum is a useful tool to analyze a sample in harmonic functions and so, to detect any dominant frequencies (this is the reason behind harmonic analysis introduced by Fourier, 1822, and not time series analysis). In Table 6, we summarize the basic equations for the power spectrum definition and estimation. Note that the identification and simulation of the dependence structure through frequency can be employed through the power spectrum (in this case frequency is defined as the inverse of lag) or equivalently through the CBS (Table 3) which is based on the climacogram (in this case frequency is defined as the inverse of scale).

Table 6: Power spectrum definition and expressions for a process in continuous and discrete time, along with the properties of its estimator (source: Dimitriadis et al., 2016a).

Type	power spectrum	
continuous	$s(w) := 4\int_0^{\infty} c(h)\cos(2\pi wh)\,\mathrm{d}h$	(T6-1)
discrete	$s_{\mathrm{d}}^{(\Delta)}(\omega) := 2\Delta\gamma(\Delta) + 4\Delta\sum_{v=1}^{\infty} c_{\mathrm{d}}^{(\Delta)}(v)\cos(2\pi\omega v)$ where $\omega \in \mathbb{R}$ is the frequency for a discrete time process (dimensionless; $\omega = w\Delta$)	(T6-2)
classical estimator	$\underline{\hat{s}}_{\mathrm{d}}^{(\Delta)}(\omega) = 2\Delta\underline{\hat{c}}_{\mathrm{d}}^{(\Delta)}(0) + 4\Delta\sum_{v=1}^{n}\underline{\hat{c}}_{\mathrm{d}}^{(\Delta)}(v)\cos(2\pi\omega v)$	(T6-3)
expectation of classical estimator**	$\mathrm{E}\left[\underline{\hat{s}}_{\mathrm{d}}^{(\Delta)}(\omega)\right] = 2n\Delta\left(\gamma(\Delta) - \gamma(n\Delta)\right)/\zeta(0) +$ $+4\Delta\sum_{v=1}^{n}\frac{\cos(2\pi\omega v)}{\zeta(v)}\left((n-v)c_{\mathrm{d}}^{(\Delta)}(v) + \frac{v^2}{n}\gamma(v\Delta) - v\gamma(n\Delta) - \frac{(n-v)^2}{n}\gamma((\mathrm{n}-v)\Delta)\right)$	(T6-4)

The continuous-time power spectrum can be solved in terms of c to yield (the inverse cosine Fourier transformation):

$$c(h) = \int_0^{\infty} s(w)\cos(2\pi wh)\,\mathrm{d}w \tag{4}$$

Also, it can be solved in terms of γ to yield (Koutsoyiannis, 2013):

$$\gamma(k) = \int_0^{\infty} s(w)\frac{\sin^2(\pi wk)}{(\pi wk)^2}\,\mathrm{d}w \tag{5}$$

$$s(w) = -2\int_0^{\infty} (2\pi wk)^2\gamma(k)\cos(2\pi wk)\,\mathrm{d}k \tag{6}$$

Note that the discrete-time power spectrum and the expectation of its classical estimator are more easily calculated with fast Fourier transform (fft) algorithms.

2.4 Stochastic processes and estimators used in thesis

Although numerous stochastic processes exist in literature, in this thesis we mostly focus on processes with mixed powered-exponential and power-type dependence structures as well as mixed forms of various distribution functions such as Gaussian-type and Pareto-type.

2.4.1 The Markov process

As shown above the time constants Δ and D affect the estimation of the statistical properties of the continuous time process. Two special cases, $\Delta = 0$ and $\Delta = D$, are analyzed by Koutsoyiannis (2013) who shows that in several tasks the differences are small. For samples with $\Delta << D$ (e.g., hourly timeseries with one minute resolution) we can assume $\Delta = 0$ and for samples with $\Delta/D \approx 1$ we can focus on the case $D = \Delta > 0$.

However, it is known that the discrete time representation of the Markov process corresponds to an ARMA(1,1) model (as mentioned in Dimitriadis and Koutsoyiannis, 2015a; Koutsoyiannis, 2002), denoted as $\underline{y}$. Its algorithm for the general case of $D \neq \Delta$, with discrete autocovariance:

$$c_d^{(\Delta,D)}(u) = \frac{1}{\Delta^2}\int_0^{\Delta}\int_{jD}^{uD+\Delta} c(x-y)\mathrm{d}x\mathrm{d}y = \frac{\lambda\left(1-e^{-\Delta/q}\right)^2}{(\Delta/q)^2}e^{-(Du-\Delta)/q} \tag{7}$$

where q is a scale parameter (with $\rho_1=e^{-\Delta/q}$) and λ is the true variance at zero lag.

In Figure 3, we show the discretization effect for the case $D \neq \Delta$ and for various Markov processes.

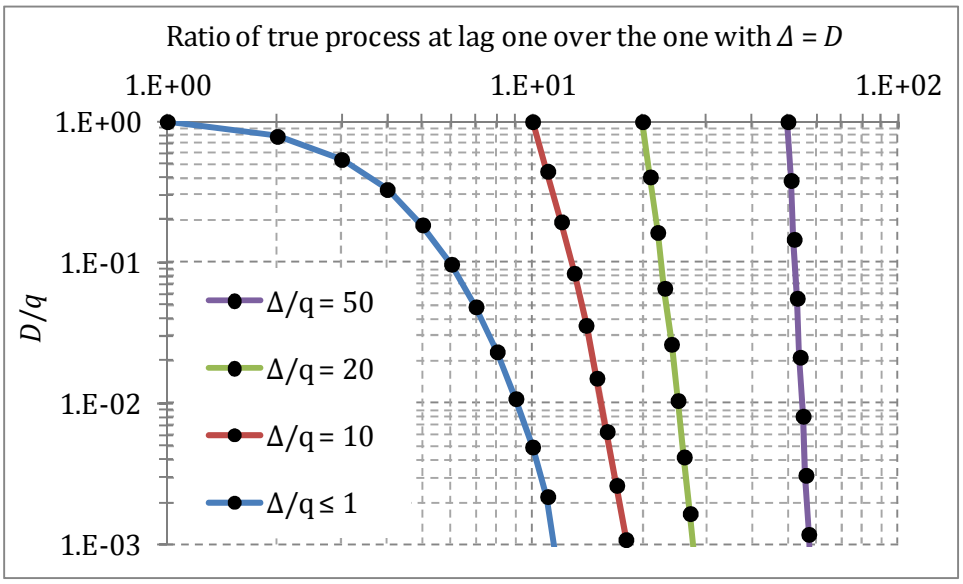

Figure 3: Ratio of the true Markov process at lag one for $D \neq \Delta$ over the one with $D = \Delta$ vs. D/q, for various values of the ratio Δ/q.

In Table 7, we provide the mathematical expressions of the climacogram, autocovariance and power spectrum for a Markov process, in continuous and discrete time for $D = \Delta > 0$.

Table 7: Climacogram, autocovariance and power spectrum expressions of a Markov process, in continuous and discrete time (source: Dimitriadis and Koutsoyiannis, 2015a).

Type	Markov process	
autocovariance (continuous)	$c(h) = \lambda e^{-\lvert h\rvert/q}$	(T7-1)
autocovariance (discrete)	$c_d^{(\Delta)}(v) = \frac{\lambda\left(1 - e^{-\Delta/q}\right)^2}{(\Delta/q)^2}\, \mathrm{e}^{-(\lvert v\rvert-1)\Delta/q}$ for $\lvert v\rvert \geq 1$ and $c_d^{(\Delta)}(0) = \gamma(\Delta)$	(T7-2)
climacogram (for continuous and discrete)	$\gamma(k) = \frac{2\lambda}{(k/q)^2}\left(k/q + e^{-k/q} - 1\right)$ with $\gamma(0) = \lambda$	(T7-3)
power spectrum (continuous)	$s(w) = \frac{4\lambda q}{1 + 4\pi q^2 w^2}$	(T7-4)
power spectrum (discrete)	$s_d^{(\Delta)}(\omega) = 4\lambda q\left(1 - \frac{1}{\Delta/q}\,\frac{(1 - \cos(2\pi\Delta\omega))\sinh(\Delta/q)}{\cosh(\Delta/q) - \cos(2\pi\Delta\omega)}\right)$	(T7-5)

2.4.2 The HK-behavioural processes

The term HK-behaviour corresponds to the behaviour of process at large scales while the process itself could not be necessarily an HK process or follow a Gaussian distribution. For example, both the fractional Gaussian noise (fGn; see section 3.2) and the generalized HK (GHK; see below) process are processes exhibiting an HK behaviour, but while the former's autocorrelation function is a power-law type at the whole range of lags, the latter's autocorrelation function is a power-law type only at large lags (at small lags behaves like a Markov process) and its distribution function is not necessarily Gaussian.

The HK process (for more details on the definition see in section 3.4) can be described via the climacogram in continuous time (with $\Delta = D$):

$$\gamma(\kappa\Delta) = \frac{\gamma(\Delta)}{\kappa^{2-2H}} \tag{8}$$

where $\kappa = k/\Delta$ denotes discrete time scale and $\gamma(\Delta)$ is the variance at the unit time scale Δ, and H is the Hurst parameter ($0 < H < 1$). Note that this process has infinite variance at zero scale and thus, should not be used to model the small scales of a physical process (e.g., the fGn process is widely but erroneously used to model several processes at small scales).

Another example that will be used in this thesis is the so-called Hybrid Hurst-Kolmogorov (HHK) process (Koutsoyiannis et al., 2017), whose climacogram is:

$$\gamma(k) = \frac{\lambda}{(1 + (k/q)^{2M})^{\frac{1-H}{M}}} \tag{9}$$

where λ is the variance of the continuous-time process $\underline{x}(t)$, M is a fractal parameter, H is the Hurst parameter and q is a characteristic time parameter. A particular case of the HHK, which is also used in this thesis and referred to as GHK process, is when M = ½, i.e.:

$$\gamma(k) = \frac{\lambda}{(1 + k/q)^{2-2H}} \tag{10}$$

Note that due to the discretization effect, an HK process for $D \neq \Delta > 0$ can be well represented by a GHK process. For example, an HK process with Δ = 0.1, D = 1, λ = 1 and H = 0.8, can be well represented by a GHK process with $\Delta = D$ = 1, λ = 2.2, q =0.14 and H = 0.8. In Figure 4, we show an example of comparison of an HK process with $D/\Delta \geq 5$ (which is approximately invariant and can be well represented by a process with Δ = 0) to the one with $D = \Delta$.

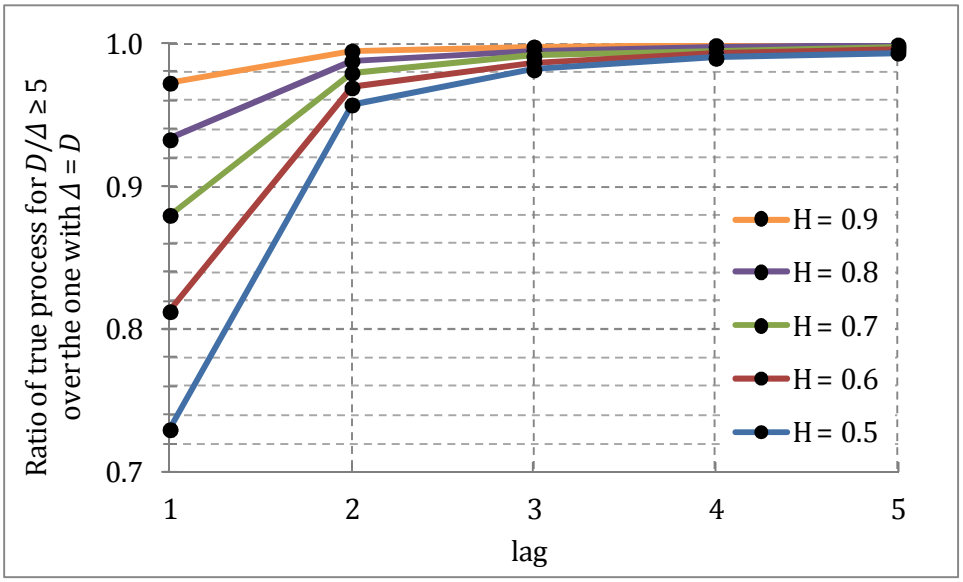

Figure 4: Ratio of the true HK process for $D/\Delta \geq 5$ vs. the one with $\Delta = D$ for various Hurst parameters.

We can also define another generalized HK process (gHK), similar to the HHK one, if we expand the HK process through the autocovariance rather than the climacogram. The expressions of climacogram, autocovariance and power spectrum for the gHK process are summarized in Table 8.

Table 8: Climacogram, autocovariance and power spectrum expressions of a positively correlated gHK process, with $0 < b < 1$, in continuous and discrete time.

Type	gHK process	
autocovariance (continuous)	$c(\tau) = \lambda((\lvert\tau\rvert/q)^{2M} + 1)^{-b/(2M)}$; Gneiting (2000)	(T8-1)
	with $b = 2 - 2H$	
autocovariance (discrete)	$c_d^{(\Delta)}(j) = \lambda \frac{\lvert j\Delta/q - \Delta/q + 1\rvert^{2-b} + \lvert j\Delta/q + \Delta/q + 1\rvert^{2-b} - 2\lvert j\Delta/q + 1\rvert^{2-b}}{(\Delta/q)^2(1-b)(2-b)}$	(T8-2)
for M=1/2	for $j \geq 1$, with $c_d^{(\Delta)}(0) = \gamma(\Delta)$	
climacogram (continuous and discrete)	$\gamma(m) = \frac{2\lambda((m/q+1)^{2-b} - (2-b)m/q - 1)}{(1-b)(2-b)(m/q)^2}$	(T8-3)
	with $\gamma(0) = \lambda$	
for M=1/2		
power spectrum (continuous)	$s(w) \approx \frac{4\lambda q^b\, \Gamma(1-b)\mathrm{Sin}\left(\frac{\pi b}{2} + 2q\pi\lvert w\rvert\right)}{(2\pi\lvert w\rvert)^{1-b}} - \frac{4\lambda q\ {}_1\mathrm{F}_2\left[1; 1-\frac{b}{2}, \frac{3}{2}-\frac{b}{2}; -\pi^2 q^2 w^2\right]}{1-b}$	(T8-4)
for M=1/2		
	(where ${}_1\mathrm{F}_2$ is the hyper-geometric function)	
power spectrum (discrete)	not a closed expression	
for q>0		

It should be noted that the gHK for M=1/2 (or the GHK) process can be considered as an HK process that gives a finite autocovariance value at zero lag, which is the common case in geophysical processes (an HK process with autocovariance $|h|^{-2+2H}$ gives infinity at zero lag). Thus, a parameter q is added to the HK process indicating the limit between HK processes ($q << |h|$) and those affected by the minimum scale limit of the process ($q >> |h|$). To switch to an HK process from the gHK (or GHK) we can replace λ with λq^{-2+2H} and then estimate the limit $q \to 0$ (see Dimitriadis and Koutsoyiannis, 2015a, section 2.1 of the supplementary material).

2.4.3 A mixed dependence structure from entropy extremization

In complex systems, entropy maximization (or extremization of entropy production) is a principle that can determine the thermodynamic equilibrium of a system (Koutsoyiannis, 2011). Therefore, it is a good practice when modelling a complex system, to first try-out processes that result from the

extremization of entropy, which is defined for a random process with a probability density function $f(\underline{x})$ as (Koutsoyiannis, 2011; Shannon, 1948):

$$\Phi(\underline{x}) = \mathrm{E}\left[-\ln\left(f(\underline{x})\right)\right] \tag{11}$$

Extremization of entropy is equivalent to extremization of entropy production (Koutsoyiannis, 2011). Such one-parameter processes that extremize the Entropy Production in Logarithmic Time (EPLT), i.e., $\varphi\big(\underline{x}(k)\big) = \mathrm{d}\Phi(\underline{x}(k))/\mathrm{dln}(k)$, are the Markov and HK processes. Particularly, the Markov process maximizes the EPLT in small scales while the HK process dominates in large scales (Koutsoyiannis, 2016). Interestingly, the EPLT for a Gaussian HK process is independent of scale and equals H (Koutsoyiannis, 2011), while for a Gaussian-Markov process it can be expressed as:

$$\varphi(k) = \frac{1}{2}\ln(\rho^{-k})(1-\rho^{k})/(\rho^{k} + \ln(\rho^{-k}) - 1) \tag{12}$$

Following the analysis in (Koutsoyiannis, 2011, 2016), we investigate the powered exponential dependence structure, i.e., with an autocovariance function (Gneiting, 2000):

$$c(h) = \lambda \mathrm{e}^{-(h/q)^{2M}} \tag{13}$$

In Figure 5, we observe that the HK process corresponds to a larger ELTP for large scales whereas for small scales the Markov process dominates. Also, the powered-exponential process for q tending to zero corresponds to a larger EPLT for $M < 0.5$ as compared to $M > 0.5$ (for $M = 0.5$ it coincides with the Markov process). Therefore, among processes with Markov, HK and mixed behaviour, we expect that an HHK process, with $M < 0.5$ and $H > 0.5$, should adequately describe a great variety of natural processes.

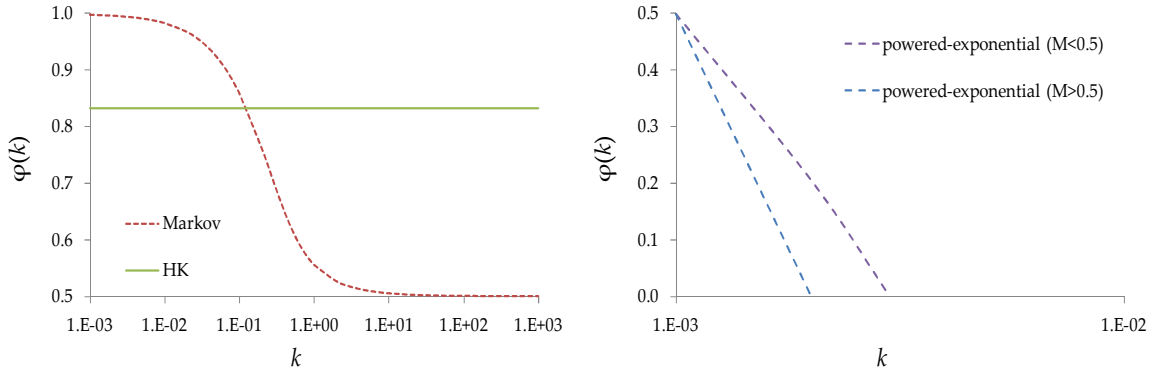

Figure 5: The ELTP of an HK process with $H = 5/6$ and a Markov process with $q = 1$ [left] and two powered-exponential functions with $q = 0.001$, and $M = 2/3$ (> 0.5) and $M = 1/3$ (< 0.5) [right].

2.4.4 Distributions based on entropy extremization

The extremization of entropy for a white noise process results in the so-called maximized entropy (ME) distribution, written as (Dimitriadis and Koutsoyiannis, 2017; Jaynes, 1957):

$$f(x;\boldsymbol{\lambda}) := \frac{1}{\lambda_0} \mathrm{e}^{-\left(\frac{x}{\lambda_1}+\mathrm{sign}(\lambda_2)\left(\frac{x}{\lambda_2}\right)^2+\left(\frac{x}{\lambda_3}\right)^3+\mathrm{sign}(\lambda_4)\left(\frac{x}{\lambda_4}\right)^4+\cdots+\left(\frac{x}{\lambda_l}\right)^l\right)} \tag{14}$$

where $\boldsymbol{\lambda} = [\lambda_0, \ldots, \lambda_l]$, with λ_l having same units as x, $\lambda_l \geq 0$ and with constraints: $\int_{-\infty}^{\infty} x^r f(x;\boldsymbol{\lambda})\mathrm{d}x = \mathrm{E}[\underline{x}^r]$, for $r = 0, \ldots, l$.

The ME for l = 2 results in the well-known Gaussian distribution (e.g., Koutsoyiannis, 2014). Another interesting distribution function for a real random variable is the Pareto-Burr-Feller (PBF) distribution (Koutsoyiannis et al., 2017) a generalization of the Cauchy distribution, i.e.:

$$f(x) = \left(1 + \left|\frac{x}{a} + d\right|^b\right)^{-c/b} \tag{15}$$

where α is a scale parameter in units of x, b and c are the dimensionless shape parameters of the marginal distribution, and d is a dimensionless scale parameter.

This distribution is similarly derived from the maximization of entropy as shown in the previous section, i.e., combination of exponential-type distributions for small values of x and heavy-tailed distributions for large values of x, maximizing the raw moment $\mathrm{E}[\underline{x}^b]$ and the entropic moment (cf., Costa, 2008) $\mathrm{E}[\ln(\underline{x}^c)]$, respectively.

It can be shown that the magnitude of independent and identically distributed variables (following the above distribution) follows the Pareto-Burr-Feller distribution (Dimitriadis and Koutsoyiannis, 2017):

$$F(x) = 1 - \left(1 + \left|\frac{x}{\alpha} + d\right|^b\right)^{-c/b} \tag{16}$$

where α is a scale parameter in units of $[x]$, b and c are the dimensionless shape parameters of the marginal distribution, and d is a dimensionless scale parameter.

The above distribution has been also derived with alternative methods, as for example from a generalization of the Rényi-Tsallis alternative definition of ME distribution (Bercher and Vignat, 2008; Yari and Borzadaran, 2010) or by adding a background measure to the original definition of entropy in order for the discretized entropy to diverge to a real value (Koutsoyiannis, 2014, and references therein). For this distribution we use the name Pareto- Burr-Feller (PBF) to give credit to (a) the engineer V. Pareto, who discovered the family of power-type distributions (while working on the size distribution of incomes in a society, Singh and Maddala, 1978), (b) to Burr (1942) who identified and analyzed (but without giving a justification) of its function first proposed as an algebraic form by Bierens de Haan, and (c) to Feller (1971) who linked it to the Beta function and distribution through a linear power transformation, which was further analyzed and summarized

by Arnold and Press (1983, sect. 3.2). Other names such as Pareto type IV or Burr type VII are also in use for the same distribution. Interestingly, the PBF distribution has two different asymptotic properties, i.e., the Weibull distribution for low wind speeds and the Pareto distribution for large ones. The PBF has been used in a variety of independent fields (Brouers, 2015). This distribution is in agreement with various geophysical processes such as magnitude of grid-turbulence and wind (see in sections 5 and 6 for applications).

Additionally, for non-Gaussian distributions or for cyclo-stationary processes (as in the atmospheric wind process, Dimitriadis and Koutsoyiannis, 2015b), we could apply a transformation scheme that approximately normalizes the process. If the distribution of the process is unknown, the transformation scheme should be based on maximum entropy (Dimitriadis and Koutsoyiannis, 2015b; Koutsoyiannis et al., 2008). If the distribution is known, then we can use the non-linear method (Lavergnat, 2016) to normalize (in case we wish to transform the process to Gaussian) or or homogenise (in case we wish to preserve the marginal distribution) the process.

2.4.5 On the uncertainty induced by the statistical bias

As we show above the true value of a statistical characteristic (e.g. variance) of a stochastic model may differ from the one estimated from a time series (with finite length). Therefore, the bias effect, i.e. the deviation of a statistical characteristic (e.g. variance) from its theoretical value in discretized time, should be taken into account not only for the marginal characteristics but also for the dependence structure. Therefore, to correctly adjust the stochastic model to the observed time series of the physical process we should always account for the bias effect since all time series are characterized by finite lengths. For example, in Tables 1-6, we present the expressions for the expected value of each stochastic metric as a function of their true values. Therefore, the bias of the expected value can be easily calculated by subtracting the expected value from its true value, e.g. the bias for the expected value of the classical estimator of the climacogram is equal to $\gamma(k) - \mathrm{E}[\underline{\hat{\gamma}}(k)] = (\gamma(n\Delta)/\gamma(\kappa\Delta) - k/n)/(1-\kappa/n)\,\gamma(\kappa\Delta)$. However, not all statistical characteristics have an analytical expression for the statistical bias and thus, Monte-Carlo techniques are usually applied. In sections 5 and 6, we present how the bias effect of the mode and the expected value can be simulated when we model a single time series, where the mode dependence structure should be analyzed and not the expected one as erroneously done in literature, and when we model several time series regarded as realizations of a single process and therefore, the expected value of the dependence structure should be analyzed.

Comparison between the bias introduced by the expected value of the classical estimator of the autocovariance, power spectrum and climacogram

Here, we investigate the bias in power spectrum estimator (evaluated via the autocovariance) that is caused by the bias of autocovariance and the finite sample size of the discretized-time process (often the discretization effect is also attributed to bias), complementing earlier studies (e.g., Stoica and Moses, 2005, ch. 2.4). We also examine the asymptotic behaviour when the sample size tends to infinity, investigating the question whether or not the discrete power spectrum estimator is asymptotically unbiased or not. For comparison, we perform similar investigations for the

autocovariance and climacogram (Dimitriadis and Koutsoyiannis, 2015a). The concepts of autocovariance, power spectrum and climacogram are examined using both exponential and power-type autocovariance, as well as combinations thereof, in order to obtain representative results for most types of geophysical processes.

The log-log derivative (LLD) is a measure of the scaling behaviour related to asymptotic coefficients such as the fractal and Hurst parameter. The LLD of a function $f(x)$ is defined as:

$$f^{\#}(x) \coloneqq \frac{\mathrm{d}\ln\big(f(x)\big)}{\mathrm{d}\ln x} = \frac{x}{f(x)}\frac{\mathrm{d}f(x)}{\mathrm{d}x} \tag{17}$$

and for the finite logarithmic derivative of $f(x)$, e.g. in case of discrete time process, we choose the backward log-log derivative, i.e.:

$$f^{\#}(x_i) \coloneqq \frac{\ln\big(f(x_i)/f(x_{i-1})\big)}{\ln(x_i/x_{i-1})} \tag{18}$$

Since the LLD is always negative for stationary mean processes, we also define for convenience the negative log-log derivative (NLD) as $-f^{\#}(x)$.

Based on Gneiting et al. (2012) analysis, the fractal parameter (F) can be defined as (cf., Beran et al., 2013, ch. 3.6):

$$F \coloneqq D + 1 - \frac{1}{2}\lim_{h\to 0}\xi^{\#}(h) \tag{19}$$

where D the dimension of the field (e.g. D = 1 for one-dimensional velocity field) and for a 1d HHK process is equal to M+2.

Based on Beran et al., (2013, ch. 1.3) analysis, the Hurst parameter (H) can be defined as (Dimitriadis et al., 2016a):

$$H \coloneqq 1 + \frac{1}{2}\lim_{k\to\infty}\gamma^{\#}(k) \tag{20}$$

Various physical interpretations of geophysical processes are based on the power spectrum and/or autocovariance behaviour. However, as mentioned above, the estimation of these tools from data may distort the true behaviour of the process and thus, may lead to wrong or unnecessarily complicated interpretation. To study the possible distortion we use the simplest processes often met in geophysics, which could be also used in synthesizing more complicated ones. Specifically, in Appendix A, we investigate and compare the climacogram, autocovariance and power spectrum of the Markov process and gHK one (for M = 0.5) in terms of their behaviour and of their estimator performance for different values of their parameters. The methodology we use to produce synthetic time series is through the SAR scheme (see in section 3.2). Some observations concluded from the graphical investigation of Appendix A as well as from the definitions of the stochastic metrics, are summarized as follows:

(a) In the definition of the climacogram, the continuous-time values are equal to the discrete-time ones (for $\Delta = D > 0$), while in case of the autocovariance and power spectrum they are different. More specifically, the discrete-time autocovariance is practically indistinguishable from the continuous-time one, but only after the first lags, while the power spectrum continuous and discrete time values vary in both small and large frequencies (where this variation is larger in the latter).

(b) The expectation of autocovariance departs from both the true one and the discrete-time one, for all the examined processes and its bias is always larger than that of the climacogram and the power spectrum. Also, the climacogram has smaller bias in comparison to the power spectrum.

(c) While in theory the NLD of the climacogram, autocovariance and power spectrum should be equal to each other, at least asymptotically, we observe from the graphical investigation (Appendix A ; Dimitriadis and Koutsoyiannis, 2015a) that in practice this correspondence may be lost.

(d) The expected value of the power can be estimated theoretically only up to frequency $w = 1/2$ (also known as the Nyquist frequency), due to the cosine periodicity. On the contrary, autocovariance and climacogram expected values can be estimated theoretically for scales and lags, respectively, up to $n - 1$.

(e) A high computational cost is involved in the calculation of the power spectrum as compared to the simple expressions of the climacogram and autocovariance. Although this is often dealt with fast-Fourier-transform algorithms, the involved large sums and large number of trigonometric products can often also cause numerical instabilities.

Some of the observations concerning the estimated power spectrum can be explained by considering the way the power spectrum is calculated through the autocovariance: when a sample value is above (below) the sample mean, the residual is positively (negatively) signed; thus, a high autocovariance value means that, in that lag, most of the residuals of the same sign are multiplied together (++ or --). In other words, the same signs are repeated (regardless of their difference in magnitude). The same 'battle of signs' process, is followed in the case of the power spectrum, but in this case, the sign is given by the cosine function. A large value of the power spectrum indicates that, in that frequency, the autocovariance values multiplied by a positive sign (through the cosine function) are more than those multiplied by a negative one. So, the power spectrum can often misinterpret an intermediate change in the true autocovariance or climacogram. A way to ttackle this could be through the autocovariance itself, i.e., not using the power spectrum at all, but this is also prone to high bias (especially in its high lag tail) which always results in at least one negative value (for proof see Hassani, 2010 and analysis in Hassani et al., 2012). These can be avoided with an approach based on the climacogram since the calculated variance is always positive. Also, the structure of the power spectrum is not only complicated to visualize and to calculate but also lacks direct physical meaning (opposite to autocovariance and climacogram), as it actually describes the Fourier transform of the autocovariance (Dimitriadis and Koutsoyiannis, 2015a)

Moreover, we investigate the performance of the estimators of climacogram, autocovariance and power spectrum for Gaussian distributed variables. For their evaluation we use mean square error

expressions as shown in the equations below. Assuming that θ is the true value of a statistical characteristic (i.e. climacogram, autocovariance, power spectral density and NLDs thereof) of the process, a dimensionless mean square error (MSE):

$$\varepsilon = \frac{\mathrm{E}\left[\left(\underline{\hat{\theta}} - \theta\right)^2\right]}{\theta^2} = \varepsilon_v + \varepsilon_b \tag{21}$$

where we have decomposed the dimensionless MSE into a variance and a bias term, i.e.:

$$\varepsilon_v = \mathrm{Var}\left[\underline{\hat{\theta}}\right]/\theta^2 \tag{22}$$

$$\varepsilon_b = \left(\theta - \mathrm{E}\left[\underline{\hat{\theta}}\right]\right)^2/\theta^2 \tag{23}$$

Note that θ is given by the true climacogram, the true autocovariance in discrete-time and the true power spectrum in discrete-time. ε_b can be found analytically through $\mathrm{E}\left[\underline{\hat{\theta}}\right]$, but ε_v cannot due to the lack of analytical solutions for $\mathrm{E}\left[\underline{\hat{\theta}}^2\right]$ and hence, $\mathrm{Var}\left[\underline{\hat{\theta}}\right]$, for the classical estimators of climacogram, autocovariance and power spectrum (hence, we use a Monte-Carlo analysis). This analysis (also presented in Appendix A) allows for some observations related to stochastic model building (Dimitriadis and Koutsoyiannis, 2015a):

(a) In general, the climacogram has lower variance than that of the autocovariance, which in turn is lower than that of the power spectrum (e.g., for the examined Markov and HK processes as well as in most scales for the gHK). Additionally, the climacogram has a smaller bias than that of the autocovariance but larger than that of the power spectrum (for all examined processes). Since, for the Markov and HK processes, the error component related to the variance, i.e., ε_v, is often larger than that from the bias, i.e., ε_b, or conversely for the gHK ones, the climacogram has a smaller total error ε. Thus, we can state that (for all the examined cases) the expression below holds:

$$\mathrm{E}\left[\left(\underline{\hat{\gamma}} - \gamma\right)^2\right]/\gamma^2 \leq \mathrm{E}\left[\left(\underline{\hat{c}}_d^{(\Delta)} - c_d^{(\Delta)}\right)^2\right]/{c_d^{(\Delta)}}^2 \leq \mathrm{E}\left[\left(\underline{\hat{s}}_d^{(\Delta)} - s_d^{(\Delta)}\right)^2\right]/{s_d^{(\Delta)}}^2 \tag{24}$$

(b) The total error for the NLD, i.e. $\varepsilon^{\#}$, increases with scale in the climacogram and with lag in the autocovariance for all examined processes. In case of a Markov process, the power spectrum NLD, i.e. $\varepsilon^{\#}$, first decreases and then increases in large inverse-frequency values, while the autocovariance and climacogram $\varepsilon^{\#}$ always increase. Also, climacogram and autocovariance $\varepsilon^{\#}$ are close to each other and in most cases smaller than the power spectrum $\varepsilon^{\#}$. For HK and gHK processes, where large scales/lags/inverse-frequencies exhibit HK behaviour, the power spectrum always decreases with inverse frequency under a power-law decay, in contrast to the autocovariance and climacogram $\varepsilon^{\#}$ which they always increase. Thus, in this type of processes, there exists a cross point between power spectrum $\varepsilon^{\#}$ and the other two, where behind this point, the power spectrum has a larger $\varepsilon^{\#}$ and beyond a smaller one.

(c) The density distribution function of the climacogram and autocovariance have small magnitude of skewness and can approximate a Gaussian density function for most of scales and lags, while the

power spectrum density has a larger skewness that results in non-symmetric prediction intervals (an important characteristic when it comes to stochastic modelling, e.g., see Lombardo et al., 2014). However, the NLD of the power spectrum has a negligible skewness in comparison to those of the autocovariance and climacogram, meaning that the expected NLD should be very close to the mode NLD.

2.5 Proposed methodology for stochastic modelling

As mentioned above, we should investigate the behaviour of a natural process by estimating separately its distribution function and marginal characteristics, and its dependence structure. A theoretically more valid approach for the estimation of the process parameters would be to apply estimators that take into account both marginal and dependence structures simultaneously. Such estimators can result to more accurate estimations. However, it is not advised to use them directly without having first visualized and identified candidates of mathematical processes, since this may result in an erroneous analysis due to the complex nature of geophysical processes, an often large number of included parameters and a high numerical burden. The best estimators of this kind certainly belong to the maximum-likelihood group of estimators.

In this thesis, we mostly focus to the dependence structure where the climacogram-based metrics are shown to be the most appropriate in terms of statistical uncertainty (section 2.4.5; Dimitriadis and Koutsoyiannis, 2015a; Dimitriadis et al., 2016a). An important issue in statistical estimation, which is sometimes misused or even neglected, is the discretization effect and statistical bias. The discretization effect can be easily tackled either by preferring the climacogram-based metrics or by following the methodology presented in section 2.4.

Furthermore, the accurate estimation of any characteristic of a timeseries corresponding to a stochastic process requires an infinite number of realizations, i.e., $T \to \infty$. However this is possible only in theory in the sense that all estimations from a timeseries are biased and therefore, cannot be accurately calculated. This can be illustrated through the estimation of raw moments from Gaussian-distributed processes with a power-law dependence structure, where statistical uncertainty is highly increased after the first two moments (Lombardo et al., 2014). Also, several researchers have commented on that higher order moments are underestimated from short finite samples (e.g., Ossiander and Waymire, 2000, 2002; Lashermes, 2004; Veneziano et al., 2006; Langousis and Veneziano, 2007; Veneziano and Furcolo, 2009; Langousis et al., 2009; Veneziano and Langousis, 2010, Langousis and Kaleris, 2014 and references therein).

Fortunately, although we cannot accurately estimate a statistical characteristic from a timeseries of a stochastic natural process, we can estimate the error induced by the bias effect of the stochastic mathematical process through theoretical calculations. In Tables 1 to 8, we show the equations for calculating the expected value for the most common dependence structures and metrics. In the cases where we cannot derive theoretically such relationships we can use as a fair approximation through the Monte Carlo method which is based on algorithms presented in section 3. Nevertheless, we can conclude that it is more likely for the sample climacogram to be closer to the theoretical one (considering also the bias) in comparison to the sample autocovariance or power spectrum to be

closer to their theoretical values. Thus, it is proposed to use the climacogram when building a stochastic model and estimate the autocovariance and power spectrum from that model, rather than directly from data. Particularly, we have to decide upon the large scale type of decay from the climacogram. If the large scale NLD is close to 1 then the process is more likely to exhibit either an exponential decay of autocovariance at large lags such as in Markov processes (scenario S1) or a white noise behaviour, i.e., $H = 0.5$ (scenario S2). In case where the large scale NLD deviates from 1 then the process is more likely to exhibit HK behaviour (scenario S3). The autocovariance can help us choose between scenarios S1 and S2, as in S1 we expect an immediate, exponential-like, drop of the autocovariance (which often has the smaller difference between its expected and mode value) whereas in S2 it is unbiased and therefore, the NLD should be close to 1. In case of the scenario S1, we can estimate the scale parameter of the Markov-type decay from the NLD of the climacogram while in case of the scenario S3 we should also look into the power spectrum decay behaviour in low frequencies. Thereafter, for the determination of the Hurst parameter, we can use various algorithms, e.g., the one of Tyralis and Koutsoyiannis (2011), which is based on the climacogram (usually taken up to 10%-20% of its maximum scale $n/2$), or that of Chen et al. (2007), which is based on the power spectrum. For the estimation of the rest of the properties, i.e., for intermediate and smaller scales, we should use the climacogram-based spectrum and climacogram-based variogram, respectively (Dimitriadis et al., 2016a).

A recipe for a robust second-order stochastic analysis includes the following steps:

1) Select a stochastic model based on parsimony (few parameters as possible it can be), theoretical justification (principle of maximized entropy) and physical interpretation (depending on the natural characteristics of the physical process) as done by Koutsoyiannis (2016) and Koutsoyiannis et al. (2017). From the analysis of this thesis, we find that the most appropriate models for the general case of both the second order dependence structure (in terms of the autocovariance or the climacogram) and the marginal distributions of several hydroclimatic processes (temperature, wind, precipitation, dew-point/humidity, river discharges, atmospheric pressure and turbulent processes) are in sections 2.4.3 and 2.4.4.

2) Handle the stochastic model for discretization and statistical bias in order to fit and emulate the sample statistical characteristics of the observed time series that can be estimated with metrics of low uncertainty. Note that the climacogram-based metrics (Koutsoyiannis, 2010; Dimitriadis and Koutsoyiannis, 2015b; Dimitriadis et al., 2016b) are the ones with the lowest statistical uncertainty and without a discretization effect. For the statistical bias, one should equate the mode of the climacogram-based metrics whereas for many time series one should use the expected value (see section 5 and 6 for many applications).

3) Using a Monte-Carlo analysis, generate as many time series as required (based on the uncertainty induced by the stochastic model) and perform a sensitivity analysis in order to certify the selection of model and parameters through the estimation of confidence intervals (Dimitriadis et al., 2016b). The generation scheme for the correlation structure can be the Sum of AR(1) or ARMA(1,1) models (known as the SAR model; Dimitriadis and Koutsoyiannis, 2015b) for correlation structures that are only based on autoregressive expressions or the Symmetric-Moving-

Average (SMA; Koutsoyiannis, 2000; Koutsoyiannis, 2016) model for any correlation structure. To approximate the marginal distribution an implicit (Koutsoyiannis, 2010; see also section 3) scheme can be used for simple applications whereas to adjust for intermittency an explicit scheme (Dimitriadis and Koutsoyiannis, 2017; see also section 3) is the most appropriate and parsimonious one.

3 Stochastic synthesis and prediction algorithms

The main purpose of stochastic analysis is the synthesis and prediction of a process. Here, we present several algorithms for generating and predicting the next values of a stochastic process by preserving both the marginal probability function and second order dependence structure. When applying the concept of stochastic analysis we model the observed unpredictable fluctuations of the system under investigation with the variability of devised stochastic processes. This stochastic process enables generation of an ensemble of its realizations, while observation of the given natural system can only produce a single or multiple (but always limited) observed timeseries. The most simple and yet powerful technique to reveal and analyze in total the system's variability, is the Monte-Carlo approach. However, this technique requires a generation algorithm capable of modelling any selected marginal probability distribution and dependence structure of the stochastic processes, appropriate for the investigated natural system.

3.1 Synthesis of a Markov process

In this section, we present a methodology to synthesize a discrete time representation of a continuous time Markov process, with parameters q and λ. We assume a sample size n and $D = \Delta \geq 0$. First, we try to approximate the continuous-time Markov process in discrete-time by an AR(1) model with variance λ_{AR}, shape parameter q_{AR} and autocovariance $\lambda_{\mathrm{AR}} e^{-j\Delta/q_{\mathrm{AR}}}$, for $\upsilon \geq 0$. We find that the AR(1) model either underestimates all autocovariances of the process for lags $\upsilon \geq 1$, when we set the variance correctly to:

$$\lambda_{\mathrm{AR}} = \gamma(\Delta) = \frac{2\lambda}{(\Delta/q)^2}\left(\Delta/q + e^{-\Delta/q} - 1\right) \leq \lambda \tag{25}$$

or overestimates this variance, when we set it equal to the continuous-time Markov variance, i.e., $\lambda'_{\mathrm{AR}} = \gamma(0) = \lambda$. Note that in both cases we apply the correct shape parameter $q_{\mathrm{AR}} = q$. Keeping the variance equal to λ_{AR} and setting the ratio of the lag-one autocovariance (or first-order autocorrelation coefficient) ρ_1 over the discrete variance to:

$$a' = \frac{c_d^{(\Delta)}(1)}{\gamma(\Delta)} = \frac{\left(1 - e^{-\Delta/q}\right)^2}{(\Delta/q + e^{-\Delta/q} - 1)} \tag{26}$$

instead of its proper value, i.e., $a = e^{-\Delta/q}$, the model correctly estimates the zero and one lags of the discrete-time autocovariances but leads to high overestimation for the rest autocovariances, i.e., for lags $\upsilon > 1$. Only in case of a very small Δ/q (or $\Delta \ll D$), i.e., when $a \approx a' \approx 1$, $c_d^{(\Delta)}(1) \approx a\gamma(\Delta)$ and $\lambda_{\mathrm{AR}} \approx \lambda$, a single AR(1) model can well approximate a discrete time representation of a continuous-time Markov process. In other words, only for the impossible case of $\Delta = 0$, the model AR(1) can exactly represent a Markov process. In practice, for $\Delta/q \lesssim 2.5\%$, we have $|a' - a|/a' \lesssim 1\%$ and thus, the AR(1) autocovariance deviates only a little from the Markov discretized one, while for large Δ/q, the error produced can be quite large. An example is shown in Figure 6 for $\Delta = D > 0$,

while for cases of $\Delta \neq D > 0$, the produced errors can be significant. Particularly, we plot the dimensionless error, between a Markov process in discrete time and various representations through the AR(1) model, defined as:

$$\varepsilon = \max_{j=0,\ldots,n-1} \left| \frac{c'^{(\Delta)}_d(v) - c^{(\Delta)}_d(v)}{c^{(\Delta)}_d(v)} \right| \tag{27}$$

where $c^{(\Delta)}_d(v > 0)$ is the Markov process and $c^{(\Delta)}_d(0)$ the zero-lag variance:

$$c^{(\Delta)}_d(v) = \frac{\lambda\left(1 - e^{-\Delta/q}\right)^2}{(\Delta/q)^2} \, \mathrm{e}^{-(|v|-1)\Delta/q} \tag{28}$$

$$c^{(\Delta)}_d(0) = \gamma(\Delta) = \frac{2\lambda}{(\Delta/q)^2}\left(\Delta/q + e^{-\Delta/q} - 1\right) \tag{29}$$

and $c'^{(\Delta)}_d(v)$ the AR(1) model, with $q_{\mathrm{AR}} = q$ and a scale parameter equal to the discrete-time variance λ_{AR} of the Markov process (blue line), the variance of the continuous time Markov process i.e., $\lambda'_{\mathrm{AR}} = \lambda$ (red line), the variance λ''_{AR} used to correctly estimate all autocovariances except the zero lag one (green line) and a variance $\lambda'''_{\mathrm{AR}} = (\lambda_{\mathrm{AR}} + \lambda)/2$ in between λ_{AR} and λ (black line). The λ_{AR} and λ''_{AR} can be expressed as:

$$\lambda_{\mathrm{AR}} = \frac{2\lambda}{(\Delta/q)^2}\left(\Delta/q + e^{-\Delta/q} - 1\right) \tag{30}$$

$$\lambda''_{\mathrm{AR}} = \frac{c^{(\Delta)}_d(1)}{e^{-\Delta/q}} = \frac{\lambda e^{\Delta/q}\left(1 - e^{-\Delta/q}\right)^2}{(\Delta/q)^2} \tag{31}$$

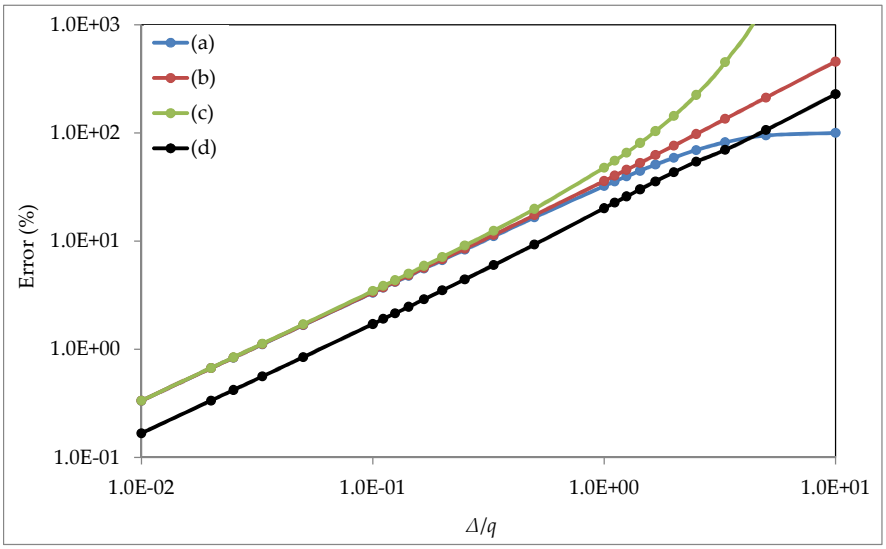

Figure 6: Dimensionless error between the autocovariance of a Markov process and those of expressed through various AR(1) models.

It is known that the discrete time representation of the Markov process corresponds to an ARMA(1,1) model (Koutsoyiannis, 2002). The ARMA(1,1) algorithm for generating a Markov

process $\underline{y}$, i.e., with continuous-time autocovariance $c(\tau) = e\exp(-|h| \,/\, q)$, in discrete-time, is the following:

$$\underline{y}_i^{(\Delta,D)} = a_1 \underline{y}_{i-1}^{(\Delta,D)} + \underline{v}_i + a_2 \underline{v}_{i-1} \tag{32}$$

where i=1, ..., n, $a_1 = e^{-D/q}$ is a parameter related to the shape of the process with $0 < a_1 \leq 1$, $\underline{v}_i = N\left(\mu_{\underline{v}}, \sigma_{\underline{v}}\right)$ is the discrete time Gaussian white noise process with mean value $\mu_{\underline{v}} = \frac{1-a_1}{1+a_2}\mu_{\underline{y}}$ with $\mu_{\underline{y}}$ the mean of $\underline{y}$.

The parameters a_2 and $\sigma_{\underline{v}}$ and can be found from the solution of two equations (Dimitriadis and Koutsoyiannis, 2015a):

$$c_d^{(\Delta)}(0) = a_1 c_d^{(\Delta)}(1) + (1 + a_1 a_2 + {a_2}^2){\sigma_{\underline{v}}}^2 \tag{33}$$

$$c_d^{(\Delta)}(1) = a_1 c_d^{(\Delta)}(0) + a_2 {\sigma_{\underline{v}}}^2 \tag{34}$$

where $c_d^{(\Delta,D)}(0)$ and $c_d^{(\Delta,D)}(1)$ are the discrete-time autocovariances of the Markov process for lag zero and one, respectively:

$$c_d^{(\Delta,D)}(0) = \gamma(\Delta) = \frac{2\lambda}{(\Delta/q)^2}\left(\Delta/q + e^{-\Delta/q} - 1\right) \tag{35}$$

$$c_d^{(\Delta,D)}(1) = \frac{\lambda\left(1 - e^{-\Delta/q}\right)^2}{(\Delta/q)^2} e^{-(D-\Delta)/q} \tag{36}$$

These equations result in a second-order polynomial, i.e.:

$${a_2}^2 + a_2 \frac{2a_1 c_d^{(\Delta,D)}(1) - (1 + {a_1}^2)\gamma(\Delta)}{c_d^{(\Delta,D)}(1) - a_1\gamma(\Delta)} + 1 = 0 \tag{37}$$

with $c_d^{(\Delta,D)}(1) \geq a_1\gamma(\Delta)$ (the equality holds only for $q \to \infty$). There are two real positive solutions:

$$a_2 = \frac{-B \pm \sqrt{B^2 - 4}}{2} \tag{38}$$

with $a_2 > 0$ and B and $\sigma_{\underline{v}}$ derived as:

$$B = \frac{2a_1 c_d^{(\Delta,D)}(1) - (1 + {a_1}^2)\gamma(\Delta)}{c_d^{(\Delta,D)}(1) - a_1\gamma(\Delta)} \leq -2 \tag{39}$$

$${\sigma_{\underline{v}}} = \sqrt{\frac{\gamma(\Delta) - a_1 c_d^{(\Delta,D)}(1)}{1 + a_1 a_2 + {a_2}^2}} \tag{40}$$

3.2 Sum of Markov processes; the SAR process and algorithms

In this section, we describe a methodology to produce synthetic Gaussian distributed timeseries of a target process $\underline{x}$ based on a sum of independent Markov processes. For a typical finite size n, the sum of a finite, usually small, number of Markov processes is capable of adequate representing a great variety of processes. For the HK-Gaussian process (else called fractional-Gaussian-noise and abbreviated as fGn) Mandelbrot (1971) introduced the idea of approximating the discrete fGn with a sum of finite AR(1)-Gaussian models. On the same principle Koutsoyiannis (2002) showed that the sum of three AR(1) models is adequate for representing an fGn process for $n < 10^4$. As accuracy requirements and n increase, a larger number of Markov processes maybe required that could be also applied for continuous processes as well as for processes different than fGn.

A general approach that can be applied to any autoregressive models (AR, ARMA etc.) has been introduced in Dimitriadis and Koutsoyiannis (2015a) based on the original ideas of Mandelbrot (1971) for the approximation of the fGn by a finite sum of Gaussian-AR(1) models, and that of Koutsoyiannis (2002) for a similar but simpler approach that can be also applied to other processes (i.e., with different dependence structures and probability distributions), with few parameters that can be analytically estimated rather than many parameters arbitrarily approximated and by also simulating the statistical bias. Note that although the methodology described below can be easily applied to the sum of higher order AR or ARMA models, it is highly not recommended, since the complexity increase could easily cause a model over-fit (e.g., Fig. 8), and present practical as well as psychological drawbacks (Mandelbrot, 1971). In other words, a three-parameter GHK model can be equivalently simulated by a sum of, as large as possible, finite number of AR(1) models, ARMA(1,1) models, as well as by a sum of high-order autoregressive models (e.g., AR(q_1), ARMA(q_1, q_2) etc., with arbitrarily large q_i >1), but only the former approach is recommended since it can provide the same (if not better) results with a simpler way and it can also deal with non-Gaussian distributions (see also section 3.3.1). In fact, *Everything should be as simple as it can be, but not simpler* (quote attributed to A. Einstein in 1933).

To explain how the SAR works, we seek the Markov climacograms whose sum fits the climacogram of our target true continuous time process, represented by a function $f(k\Delta)$, with κ the discrete-time scale and $D = \Delta > 0$ the time step. We could use the autocovariance or power spectrum but the climacogram for $D = \Delta > 0$ has the advantage of reduced computational cost due to the identical expressions for continuous and discrete-time. We denote $g(\kappa\Delta, q, \lambda)$ the true climacogram of a Markov process, i.e.:

$$g(\kappa\Delta, q, \lambda) := \frac{2\lambda}{(\kappa\Delta/q)^2}\left(\kappa\Delta/q + e^{-\kappa\Delta/q} - 1\right) \tag{41}$$

where λ and q are parameters corresponding to the variance and a characteristic time scale of the process, respectively.

The SAR has been applied to several processes such as the wind process (Deligiannis et al., 2016), for the process of solar radiation (Koudouris et al., 2017) or for the process of wave height and wave period (Moschos et al., 2017).

However, although the SAR algorithm can preserve any of the processes presented in 2.4 with M = ½ and additionally can preserve any distribution function (while the SARMA can be used only for Gaussian distributions), it is inappropriate for several processes with Δ>>0, such as precipitation.

Our target is to approximate $f(k\Delta)$ with the sum of a finite number N of functions $g(\kappa\Delta, q_l, \lambda_l)$ for l = 1 to N, i.e., for all integral scales from κ = 1 to n, where n is the number of data produced in the synthetic time series. We seek $q_l > 0$ and $\lambda_l > 0$ such as for all scales $\kappa \geq 1$ we have $f(\kappa\Delta) \approx \sum_{l=1}^{N} g(\kappa\Delta, q_l, \lambda_l)$. The basic assumption of this methodology is that the Markov parameters q_l are connected to each other in a predefined way, which can be even similar to the target process if we wish to preserve in an exact way the 2nd order dependence structure. Here, we choose a simple relationship based on two parameters p_1 and p_2 (Dimitriadis and Koutsoyiannis, 2015a):

$$q_l = p_1 {p_2}^{l-1} \tag{42}$$

If we know p_1 and p_2, we can calculate analytically parameters λ_l (expressed by the matrix $\boldsymbol{\Lambda} \geq \mathbf{0}$) from the equation below, since the ratio $g(k\Delta, q_l, \lambda_l)/\lambda_l$ is independent of λ_l for Markov processes:

$$\boldsymbol{A\Lambda} = \boldsymbol{I} \rightarrow \boldsymbol{\Lambda} = \boldsymbol{A}^{-1}\boldsymbol{I} \tag{43}$$

where $\boldsymbol{\Lambda}$=$[\lambda_1, \dots, \lambda_N]^T$, $\boldsymbol{I}$=$[1, \dots, 1]^T$ and $\boldsymbol{A}^{-1} = \left(\boldsymbol{A}^T\boldsymbol{A}\right)^{-1}\boldsymbol{A}^T$, the left inverse of $\boldsymbol{A}$ (for $n > N$), expressed as:

$$\boldsymbol{A} = \begin{bmatrix} \frac{g(\Delta, q_1, \lambda_1)/\lambda_1}{f(\Delta)} & \cdots & \frac{g(\Delta, q_N, \lambda_N)/\lambda_N}{f(\Delta)} \\ \vdots & \ddots & \vdots \\ \frac{g(n\Delta, q_1, \lambda_1)/\lambda_1}{f(n\Delta)} & \cdots & \frac{g(n\Delta, q_N, \lambda_N)/\lambda_N}{f(n\Delta)} \end{bmatrix} \tag{44}$$

As minimization objective for the above system of equations, in order to estimate the parameters p_1 and p_2, first we use the dimensionless error ε_s between the sum of Markov climacograms and $f(\kappa\Delta)$, to locate initial values and then, we use the error ε_m (maximum absolute dimensionless residual), for fine tuning and distributing the error equally to all scales:

$$\varepsilon_s = \sum_{\kappa=1}^{n} \left| \frac{\sum_{l=1}^{N} g(\kappa\Delta, q_l, \lambda_l) - f(\kappa\Delta)}{f(\kappa\Delta)} \right| \tag{45}$$

$$\varepsilon_m = \max_{\kappa=1,\dots,n} \left| \frac{\sum_{l=1}^{N} g(\kappa\Delta, q_l, \lambda_l) - f(\kappa\Delta)}{f(\kappa\Delta)} \right| \tag{46}$$

Thus, we can estimate parameters p_1 and p_2 by minimizing the above errors, then parameters q_l and λ_l can be easily found. Finally, the synthetic discrete time series for the $\underline{x}(t)$ process can be estimated as:

$$\underline{x}_i^{(\Delta)} = \sum_{l=1}^{N} \underline{y}_i^{(\Delta)}(l) \tag{47}$$

where $\underline{y}_i^{(\Delta)}(l)$ is the discrete time Markov process corresponding to the climacogram $g(\kappa\Delta, q_l, \lambda_l)$ with parameters q_l and λ_l.

The above methodology has been tested in simple processes such as HK, GHK, gHK and combination thereof as well as with Markov processes (Dimitriadis and Koutsoyiannis, 2015a) and therefore, for other types of processes (e.g. anti-correlated ones with $1 < b < 2$) one should be cautious when applying it. For the purpose of the analysis, we apply the above methodology for HK and gHK processes for λ = 1 and for a variety of b, q/Δ and n values. In Tables 9-11, we present the results from this analysis. Note that we choose N, for each n and each process, as the minimum value of the sum of Markov processes achieving $\varepsilon_{\mathrm{m}} \leq 1\%$.

Table 9: Parameters p_1 and p_2 estimated to fit different types of HK and gHK processes (for λ = 1) with a sum of Markov processes for $n = 10^2$.

process	b	q/Δ	p_1	p_2	N	ε_{m} (‰)
HK	0.2	-	0.069	47.358	3	6
HK	0.5	-	0.122	22.196	3	8
HK	0.8	-	0.101	17.045	3	9
gHK	0.2	1	2.888	10.656	3	5
gHK	0.2	10	11.424	27.168	2	1
gHK	0.2	100	611.13	-	1	2
gHK	0.5	1	1.789	7.695	3	9
gHK	0.5	10	9.232	12.514	2	2
gHK	0.5	100	243.46	-	1	4
gHK	0.8	1	1.373	6.559	3	9
gHK	0.8	10	7.676	8.807	2	2
gHK	0.8	100	151.54	-	1	6

Table 10: Parameters p_1 and p_2 estimated to fit different types of HK and gHK processes (for $\lambda = 1$) with a sum of Markov processes for $n = 10^3$.

process	b	q/Δ	p_1	p_2	N	ε_c (‰)
HK	0.2	-	0.379	10.356	5	2
HK	0.5	-	0.251	9.490	5	5
HK	0.8	-	0.103	8.958	5	4
gHK	0.2	1	2.656	11.873	4	3
gHK	0.2	10	0.852	43.042	3	6
gHK	0.2	100	111.54	27.331	2	1
gHK	0.5	1	1.964	10.505	4	7
gHK	0.5	10	8.744	5.801	4	2
gHK	0.5	100	89.976	12.591	2	2
gHK	0.8	1	1.362	8.240	4	7
gHK	0.8	10	6.900	5.112	4	2
gHK	0.8	100	74.712	8.861	2	3

Table 11: Parameters p_1 and p_2 estimated to fit different types of HK and gHK processes (for $\lambda = 1$) with a sum of Markov processes for $n = 10^4$.

process	b	q/Δ	p_1	p_2	N	ε_c (‰)
HK	0.2	-	0.665	18.217	5	7
HK	0.5	-	0.200	11.400	6	6
HK	0.8	-	0.053	17.044	5	8
gHK	0.2	1	2.695	12.006	5	4
gHK	0.2	10	20.809	12.793	4	5
gHK	0.2	100	7.743	44.342	3	7
gHK	0.5	1	2.226	12.176	5	10
gHK	0.5	10	14.831	10.788	4	10
gHK	0.5	100	84.308	5.835	4	2
gHK	0.8	1	1.115	6.220	6	3
gHK	0.8	10	10.132	8.149	4	9
gHK	0.8	100	66.249	5.123	4	2

3.3 Synthesis of a stochastic process through the (S)MA scheme

In this section, we present an extension of the symmetric-moving-average (SMA) generalized framework introduced by Koutsoyiannis (2000) and further advanced by Koutsoyiannis (2016) and

implemented within the Castalia computer package (Efstratiadis et al., 2014). Also, the SMA model for autocorrelation functions accounting for seasonal aspects is initially developed by Langousis (2003) and Langousis and Koutsoyiannis (2006). The generation scheme simultaneously preserves any type of (second order) dependence structure as well as an approximation of the marginal distribution function through the preservation of its statistical moments. Note that this scheme can be applied to any type of statistical moments such as raw, central, L-moments etc. as well as to any type of moving-average model such as backward (BMA), forward (FMA), symmetric (SMA) or mixed. More details about the computational scheme can be found in Dimitriadis and Koutsoyiannis (2017).

3.3.1 The impracticality of using multi-parameter stochastic models in geophysical processes

Several families of autoregressive models are used for stochastic generation with the most popular in literature to be the so-named AR, ARMA, ARIMA, FARIMA (cf., Koutsoyiannis, 2016). These models are easy to handle and fast in stochastic generation once their parameters are known and not too many. However, whenever the process exhibits long-range dependence these models require a large number of parameters to approximate the long-range dependence (except only in the FARIMA(0,d,0) case, where $d = H - 0.5$, with H the Hurst coefficient).

An additional difficulty may arise when estimating the prediction intervals (P.I.) of a long-range process (Papoulis, 1990, pp. 240-242; Tyralis et al., 2013). Even if the model parameters are calculated with adequate accuracy, this does not guarantee an adequate approximation of the prediction intervals. Here, we apply various Monte-Carlo experiments and we show that even a small deviation of the true process from the model one, may cause a larger deviation in the prediction intervals. In Figure 7, we compare the 5% and 95% P.I. of the climacogram for a Gaussian HK process with $n = 2\times10^3$, using a model consisted from the sum of three AR(1) models (through the SAR scheme) and the exact solution produced via the SMA model. We observe that although the expected value is very well approximated by the SAR model with approximately a 99% correlation coefficient, the 5% P.I. deviates from the true one by 1% and the 95% P.I. by 10%.

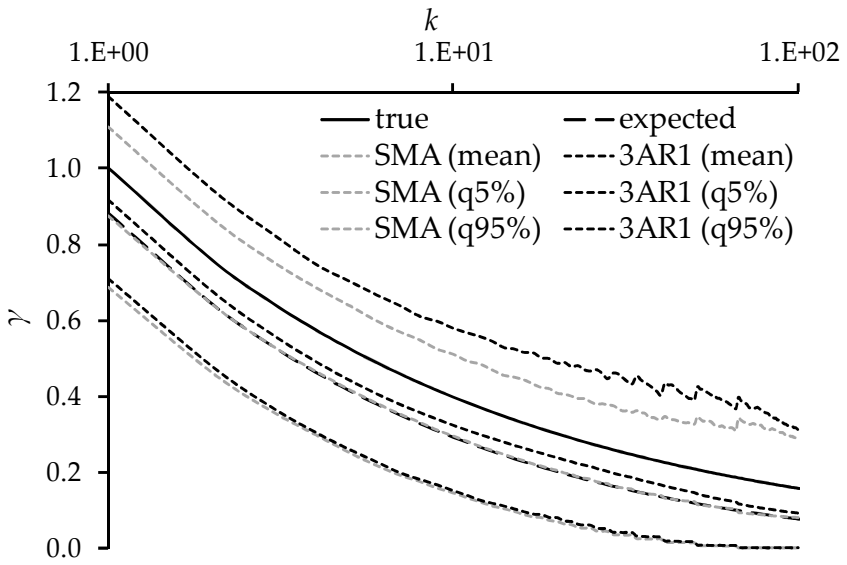

Figure 7: Expected 5% and 95% quantiles of the climacogram for an HK process estimated from Monte-Carlo experiments using the SMA model (exact solution) and the sum of three AR(1) models (3AR1) through the SAR scheme.

A practical solution could be to increase the number of AR(1) processes through the SAR scheme or to use higher order processes instead, such as ARMA models. However, in any case, it is often difficult to know a priori the true P.I. in order to decide whether the number of applied parameters is adequate. In Figure 8, we show that even when we extent the 3×AR(1) model to a 5×ARMA(1,1) model (through the SARMA scheme) for a simple HK process, the true 95% P.I. (defined through the SMA scheme) is still not reached (the fitting error is around 1%).

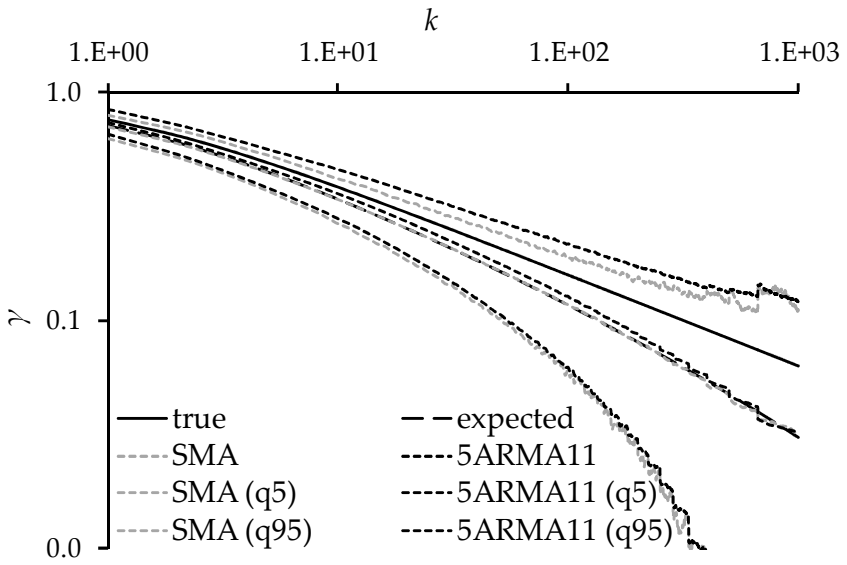

Figure 8: Expected 5% and 95% quantiles of the climacogram for an HK process estimated from Monte-Carlo experiments using the SMA model (exact solution) and the sum of five ARMA(1,1) models (5ARMA11) through the SARMA scheme.

Another limitation may arise for more complicated processes than that of the HK one. For example, the GHK process, which is an HHK process with $a = 1$, can be somehow simulated through the SAR algorithm. However, this simple algorithm is based on the sum of Markov processes and therefore, it can only preserve stochastic structures with an exponential short-term behaviour at large scales. In other words, the SAR scheme cannot accurately synthesize a process with a powered-exponential autocorrelation function, such as the HHK with $M \neq 1/2$ (Figure 9).

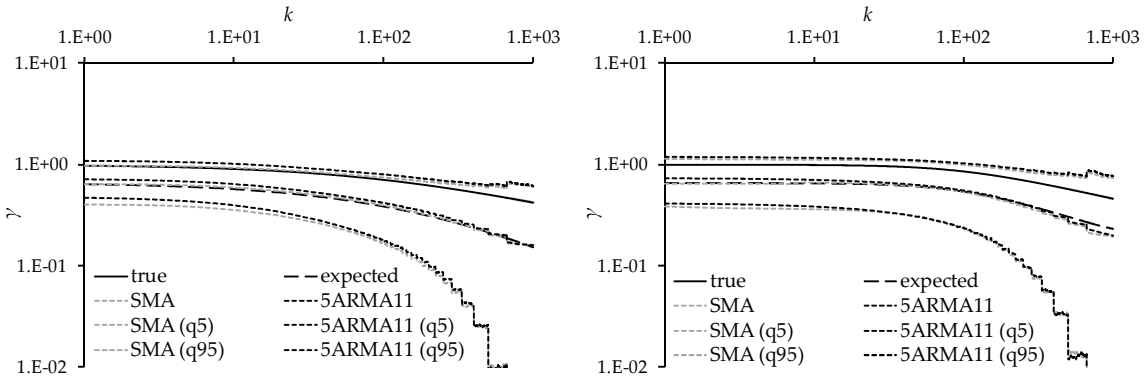

Figure 9: Expected 5% and 95% quantiles of the climacogram for two HHK processes, both with q = 10, b = 1/3 (H = 5/6), n = 2×10^3 and one with M = 1/3 < 0.5 (left) and the other with M = 3/4 > 0.5 (right), estimated from Monte-Carlo experiments using the SMA model (exact solution) and the sum of five ARMA(1,1) models (5ARMA11) through the SARMA scheme.

Additionally, another common practice is to use transformation schemes to indirectly simulate both the dependence structure and marginal distribution of a process. However, since the transformation of a Gaussian distributed process to a more complicated one is often non-linear,

there will be a non-linear distortion in the dependence structure especially in case of an HK process. In Figure 10, we show such a distortion in case of a Pareto distribution that leads to a non HK process resembling that of a cyclo-stationary HK process (i.e., causing a small increase of the climacogram at small scales) with the same Hurst parameter.

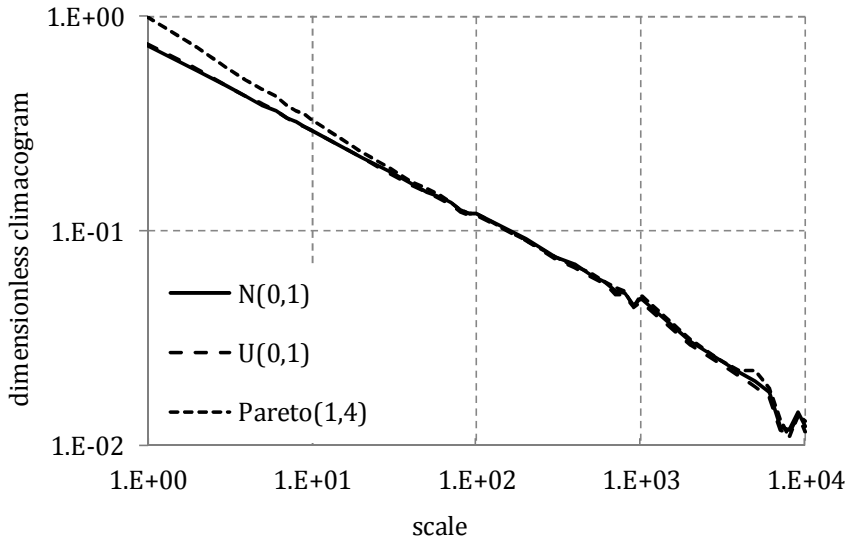

Figure 10: Dimensionless climacogram vs. scale for a synthetic HK process with $n = 10^5$, $H = 0.8$ and distribution N(0,1) as well as its transformation to U(0,1) and Pareto distribution with shape parameter equal to 4.

Finally, if the estimation of higher than the third moment is needed, for example the kurtosis, higher-order moments, i.e., $\mathrm{E}[\underline{x}^2\underline{V}^2]$, will emerge that are not possible to measure or handle for SARMA (or higher order) algorithms (Koutsoyiannis, 2016). In conclusion, the SMA algorithm overcomes all the above limitations and offers a strong tool for applying a Monte Carlo analysis.

3.3.2 The impracticality of estimating higher-order moments in geophysical processes

Non-Gaussianity of the marginal distribution is very common in geophysical processes. It has been shown (Lombardo et al., 2014) that the estimation of high raw moments corresponds to high uncertainty and thus, it is rather ambiguous to use the schemes described in the previous section to preserve higher moments for natural processes with only a few measurements, as for example in typical geophysical records. For example, in case of a continuous HK process the variance of the mean estimator is γ_Δ/n^{2-2H} (e.g., Koutsoyiannis, 2003), where n is the sample size. Consequently, for estimating the true mean μ of a process with a standard error $\pm\varepsilon$, we would require a timeseries of length of at least $(\sigma/\varepsilon)^{1/(1-H)}$, where $\sigma = \sqrt{\gamma_\Delta}$ is the standard deviation at scale Δ (Figure 11). For an HK process with $H = 0.8$, in order to estimate the mean of the process with an error $\varepsilon \approx \pm 10\%\sigma$, we would need a timeseries of length at least $n = 10^5$. Such lengths are hardly available in observations of geophysical processes, which are not only often characterized by HK behaviour but also include sub-daily and seasonal periodicities (e.g., Hasson et al., 1990; Dimitriadis and Koutsoyiannis, 2015b, for the atmospheric wind process) that complicate the estimation further.

Therefore, the preservation of solely the second order joint statistics is often adequate for capturing the most important attributes of a geophysical process but also it is often impractical to estimate higher-order statistics from observations of hydrometeorological processes since, the typically available observation records cannot support the estimation of a few parameters (Koutsoyiannis, 2016).

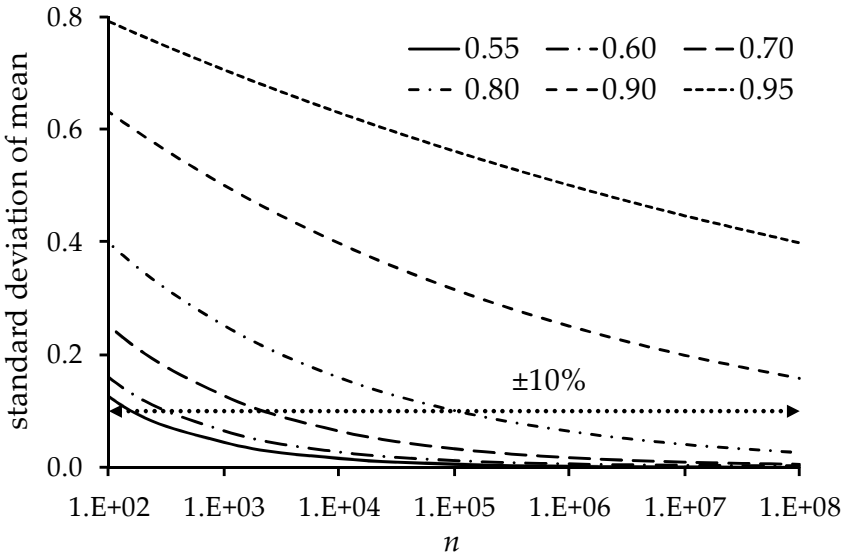

Figure 11: Standard deviation of the mean estimator of an HK process standardized by σ vs. the sample size (n) for various Hurst coefficients.

To give another example, we perform a Monte Carlo experiment for an HK process with $H = 0.8$ that follows a standard Gaussian distribution (i.e., $\mu = 0$ and $\sigma = 1$) and the results are shown in the Figure 12. For each synthetic timeseries we estimate the mean, standard deviation as well as skewness and kurtosis coefficients for six different lengths, i.e., $n = 10$, 10^2,..., and 10^6. This experiment shows that for $n = 10^6$ the uncertainty (measured in terms of the standard deviation of each measure) is below 10% for all measures. Therefore, to adequately estimate these measures from data we would need timeseries with similar lengths. The same experiment must be repeated for the estimated set of parameters to verify that the observed length was adequate for such estimation.

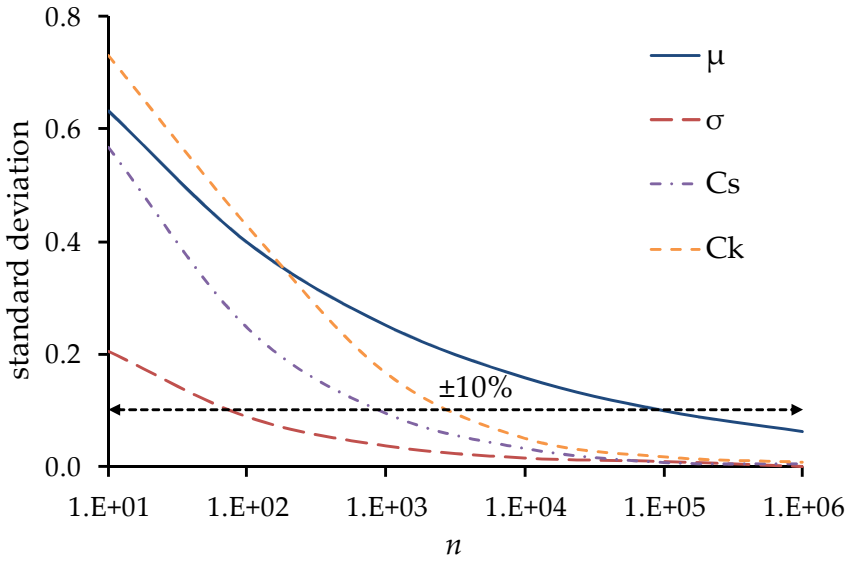

Figure 12: Standard deviation of the sample estimates of the mean (μ), standard deviation (σ), skewness coefficient (C_s) and kurtosis coefficient (C_k) of an HK process with H = 0.8 and N(0,1) distribution vs. the simulation length.

3.3.3 The SMA generation scheme

Although there are several methods for simulation of an arbitrary stochastic process each one has its own limitations and advantages (Lavergnat, 2016 and references therein). For example, the method of de-normalization (i.e., a Gaussian distributed process with the desired dependence structure is produced and then it is transformed to the desired distribution through a non-linear transformation) is often applied for synthesis of long-term processes (e.g., Koutsoyiannis et al., 2008) but it has a disadvantage of distorting the dependence structure (because of the transformation), while, in addition, the transformation cannot be invariant with respect to the time scale (Lombardo et al., 2013). A rigorous and general method is the SMA scheme that is able to fully preserve any (second order) dependence structure of a process and, simultaneously, the complete multivariate distribution function if it is Gaussian (because of the preservation of the Gaussian attribute within linear transformations). Koutsoyiannis (2000) also studied the application of the same scheme to non-Gaussian processes by preserving the skewness of the marginal distribution. In Dimitriadis and Koutsoyiannis (2017) the scheme is extended to precisely preserve the first four central moments of the distribution, while exactly and simultaneously preserving any type of (second-order) dependence structure, such as short-range (Markov) or long-range (Hurst-Kolmogorov, abbreviated as HK). In most problems preservation of four moments suffices for a very good approximation of the distribution function. In particular, the fourth moment has been regarded of great importance in some problems, e.g., in the characterization of intermittency in turbulence (Batchelor and Townsend, 1949).

In the SMA model, the simulated process is expressed through the sum of products of coefficients (not parameters) a_j and white noise terms $\underline{v}_i$, (Koutsoyiannis, 2000):

$$\underline{x}_i = \sum_{j=-l}^{l} a_{|j|}\underline{v}_{i+j} \tag{48}$$

in which for simplicity and without loss of generality we assume that $\mathrm{E}[\underline{x}] = \mathrm{E}[\underline{v}] = 0$ and $\mathrm{E}[\underline{v}^2] = \mathrm{Var}[\underline{v}] = 1$ and where j is an index ranging from 0 to infinity.

Derivation of the SMA dependence structure parameters

This scheme can be used for stochastic generation of any type of second order process structure represented by functions such as the climacogram, the autocovariance function, the power spectrum, and the variogram. It exhibits several advantages over widely used backward moving average (BMA) schemes (Koutsoyiannis, 2000). The most important is that for some processes (for example the HK) it allows closed expressions for the coefficients a_j, based on any of the above functions, which can yield a very fast generation algorithm, in case an explicit expression for the coefficients a_j is not possible (as for the GHK and HHK processes). Particularly, the coefficients can be numerically calculated through the Fourier transform of the discrete power spectrum of the coefficients which is directly linked to the discrete power spectrum of the process (Koutsoyiannis, 2000):

$$s_{\boldsymbol{a}_\mathrm{d}}(\omega) = \sqrt{2 s_\mathrm{d}(\omega)} \tag{49}$$

where $s_{\boldsymbol{a}_\mathrm{d}}$ and s_d are the SMA coefficients and process power spectra in discrete time, respectively.

As an example, for an HK process with $H > 0.5$, the SMA coefficients can be estimated from (Koutsoyiannis, 2016):

$$a_j = C\left(\frac{|j+1|^{H+\frac{1}{2}} + |j-1|^{H+\frac{1}{2}}}{2} - |j|^{H+\frac{1}{2}}\right) \tag{50}$$

where the coefficient C is:

$$C = \sqrt{\frac{2\Gamma(2H+1)\sin(\pi H)\gamma_\Delta}{\Gamma^2(2H+1)\big(1+\sin(\pi H)\big)}} \tag{51}$$

Derivation of the SMA distribution parameters

Koutsoyiannis (2000) estimated the first three moments of the marginal distribution of the white noise process $\underline{v}_i$ required to reproduce those of the actual process $\underline{x}_i$ using the SMA scheme. With the conventions used here the mean and variance of $\underline{v}_i$ are 0 and 1, respectively, while the third moment, which is equal to the coefficient of skewness is:

$$C_{\mathrm{s},v} = \frac{\left(\sum_{j=-l}^{l} a_{|j|}^2\right)^{3/2}}{\sum_{j=-l}^{l} a_{|j|}^3} C_{\mathrm{s},x} \tag{52}$$

where $C_{\mathrm{s},x}$ is the coefficient of skewness of $\underline{x}_i$.

Here, we expand the calculations to include the coefficient of kurtosis (Appendix B; Dimitriadis and Koutsoyiannis, 2017):

$$C_{\mathrm{k},v} = \frac{\left(\sum_{j=-l}^{l} a_{|j|}^2\right)^2}{\sum_{j=-l}^{l} a_{|j|}^4} C_{\mathrm{k},x} - \frac{\sum_{j=-l}^{l}\sum_{k=-l}^{l} a_{|j|}^2 a_{|k|}^2}{\sum_{j=-l}^{l} a_{|j|}^4} \tag{53}$$

where $C_{\mathrm{k},\underline{x}}$ is the coefficient of kurtosis of $\underline{x}_i$. Note that the constant term in the right-hand side depends only on the SMA coefficients and not on the marginal distribution of the process. Also, note that the kurtosis of the white noise is not proportional to the kurtosis of the process, which makes a difference from the case of the skewness.

For the generation scheme we need distributions that: (a) contain at least four parameters, creating in such way a large variety of combinations between the first four moments; (b) have closed analytical expressions for the first four central moments; and (c) can easily and quickly generate random numbers. Here, we propose one distribution mostly appropriate for generating thin-tailed distributions and another one for heavy-tailed ones (see Appendix C for the tail-classification of the applied distributions).

For illustration, we apply the described SMA model for white noise processes with various marginal distributions often met in geophysics, such as Weibull, gamma, lognormal and Pareto. Also, we estimate the ME distribution up to the fourth moment and we compare it to the theoretical and modelled distribution (through the SMA algorithm). The coefficients $1/\lambda_1, 1/\lambda_2, 1/\lambda_3, 1/\lambda_4$ of the ME distribution can be also regarded as weighting factors representing the dependence of the distribution on each raw moment. Interestingly, after standardizing these four parameters based on the sum of their absolute values, $1/\lambda_1$ contributes to the Weibull, gamma, lognormal and Pareto distributions in Figure 13, approximately 65%, 66%, 69% and 93%, respectively. Similarly, the contribution of $1/\lambda_2$ is approximately 20%, 20%, 18% and 4%, the contribution of $1/\lambda_3$, 11%, 10%, 9% and 2% and the contribution of $1/\lambda_4$, 4%, 4%, 4% and 1%, respectively. Therefore, we can use the ME probability density to approximately determine the weight for each statistical moment and justify whether the preservation up to the fourth moment is adequate. Additionally in Figure 13, we observe that the goodness of fit highly depends on the weighting factors of the ME distribution. Particularly, large weighting factors of λ_1 and small weighting factors of λ_2, λ_3 and λ_4 result in small fitting errors.

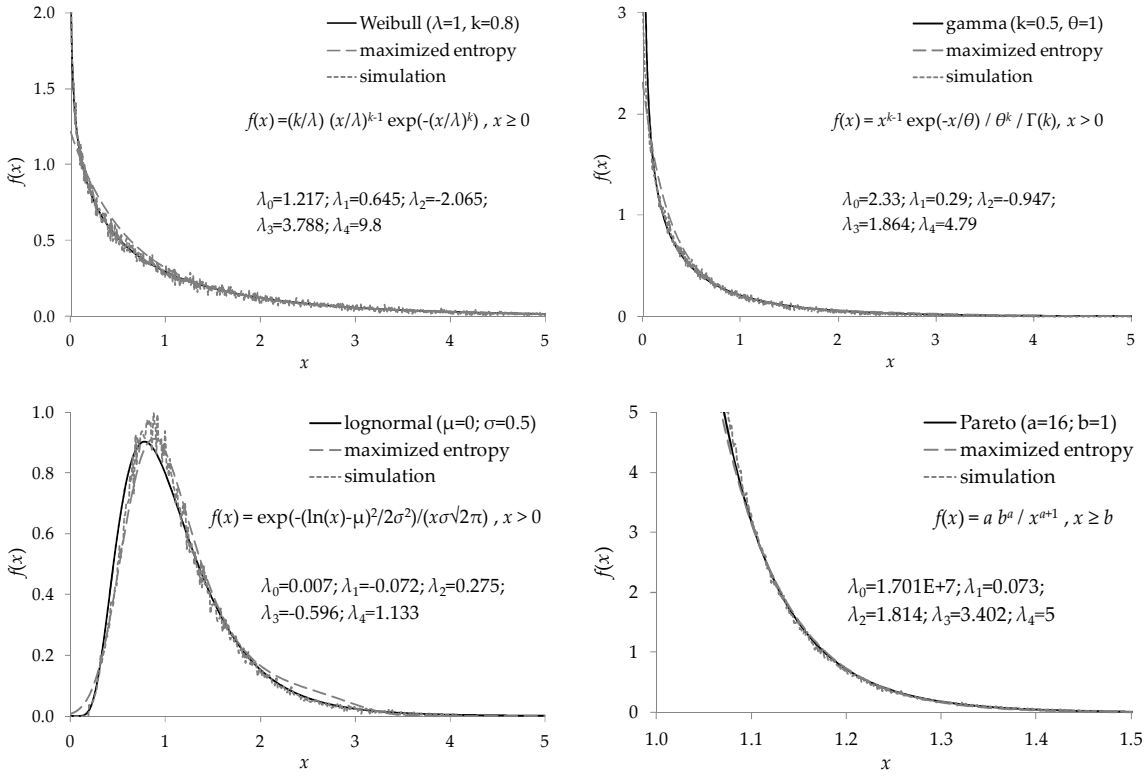

Figure 13: Various two-parameter distributions along with the fitted ME probability density function and the empirical probability density from one single simulation with $n = 10^5$ using the proposed generation scheme.

3.4 Synthesis of a multiple dimensional process through SMA scheme

Multi-dimensional stochastic processes are advantageous over multivariate ones in cases where the natural process is observed by images (e.g., produced by satellite or radar) rather than point measurements (e.g., temperature recorded at meteorological stations). In this section, we show the expansion of the 1d SMA algorithm to an L-dimensional (Ld) based on mathematical reasoning as well as numerical validation (Dimitriadis et al., 2013). We denote with $\underline{x}(\boldsymbol{t})$, the continuum random variable of a stochastic stationary and isotropic process of M dimensions with $\boldsymbol{t}$ a matrix of L variables and l varying from 1 to L, i.e., $\boldsymbol{t} \coloneqq \{t_1, \dots, t_L\}$ that is used to describe each dimension of the process (e.g., t_1 can be a temporal variable, t_2 a spatial one etc.). Note that in this analysis the M dimensions are considered independent to each other. Discretized processes are subject to a sampling frequency $\boldsymbol{D} \coloneqq \{D_1, \dots, D_M\}$ and a response time $\boldsymbol{\Delta} \coloneqq \{\Delta_1, \dots, \Delta_L\}$ as in the 1d case. Both $\boldsymbol{D}$ and $\boldsymbol{\Delta}$ have the same units with the corresponding variable $\boldsymbol{t}$ (e.g., if $t_1$is a temporal variable measured in seconds then D_i and Δ_i will be measured also in seconds). Here, we focus only in the case of $\boldsymbol{D}$=$\boldsymbol{\Delta}$>$\boldsymbol{0}$. Also, for simplicity, we assume that all elements in $\boldsymbol{D}$ have the same magnitude (e.g., D_1=1 sec, D_2=1 km etc.) and so, we can use a unique symbol for that magnitude, i.e., $|D_i| = D = \Delta$.

Finally, n denotes the total number of data in the Ld field. Thus, the discretized stochastic process $\underline{x}_{i_1,i_2,\dots,i_L}^{(\Delta_1,\Delta_2,\dots,\Delta_L)} := \underline{x}_{\boldsymbol{i}}^{(\Delta)}$, for $\Delta > 0$, can be estimated from $\underline{x}$ as (Dimitriadis et al., 2013):

$$\underline{x}_{\boldsymbol{i}}^{(\Delta)} = \frac{\int_{(i_1-1)\Delta}^{i_1\Delta}\int_{(i_2-1)\Delta}^{i_2\Delta}\dots\int_{(i_L-1)\Delta}^{i_L\Delta}\underline{x}(\xi_1,\xi_2,\dots,\xi_L)\mathrm{d}\xi_1\mathrm{d}\xi_2\dots\mathrm{d}\xi_L}{\Delta^L} \tag{54}$$

where $i_1 \in [1,n_1], i_2 \in [1,n_2], \dots, i_M \in [1,n_L]$, are indices representing the serial number of each observation associated with the corresponding variable t_m,

In Tables 12 and 13, we provide all necessary definitions and equations for the true continuous, true discrete and most common estimators and estimations for the expected climacogram and autocovariance for an Ld process (for the variogram and power spectrum see in Dimitriadis et al., 2013).

Table 12: Climacogram definition and expressions for an Ld continue process, a discretized one, a common estimator for the climacogram and the estimated value, based on this estimator.

Type	Md climacogram	
continuous space	$\gamma(\boldsymbol{k}) := \frac{\mathrm{Var}\left[\int_{t_1}^{t_1+k_1}\dots\int_{t_L}^{t_L+k_L}\underline{x}(\xi_1,\dots,\xi_M)\mathrm{d}\xi_1\dots\mathrm{d}\xi_L\right]}{(k_1k_2\dots k_L)^2}$	(T12-1)
	where $\boldsymbol{k} := (k_1,\dots,k_M)$, with $\boldsymbol{k} \in \mathbb{R}^+$, the vector of the scales.	
discretized space	$\gamma^{(\Delta)}(\boldsymbol{k}) := \gamma(\Delta k_1,\dots,\Delta k_L)$	(T12-2)
	where $\boldsymbol{\kappa} := (\kappa_1,\dots,\kappa_L)$, with $\boldsymbol{\kappa} \in \mathbb{N}^+$, the vector of all the dimensionless scales for a discretized process.	
classical estimator	$\hat{\underline{\gamma}}^{(\Delta)}(\boldsymbol{\kappa}) = \frac{1}{n'/\kappa'-1}\sum_{r_l=1}^{\lfloor n_l/\kappa_l \rfloor}\left(\frac{1}{\kappa'}\left(\sum_{i_l=\kappa_l(r_l-1)+1}^{\kappa_l r_l}\underline{x}_{\boldsymbol{i}}^{(\boldsymbol{\Delta})}\right) - \frac{\sum_{i_l=1}^{n_l}\underline{x}_{\boldsymbol{i}}^{(\Delta)}}{n'}\right)^2$	(T12-3)
	where $n' = n_1n_2\dots n_L$, $\kappa' := \kappa_1\kappa_2\dots\kappa_L$ and l ranges from 1 to L	
expected value of estimator	$\mathrm{E}\left[\hat{\underline{\gamma}}^{(\Delta)}(\boldsymbol{\kappa})\right] = \frac{1-\gamma^{(\Delta)}(\boldsymbol{n})/\gamma^{(\Delta)}(\boldsymbol{\kappa})}{1-\kappa'/n'}\gamma^{(\Delta)}(\boldsymbol{\kappa})$	(T12-4)

Table 13: Autocovariance definition and expressions for an Ld continue process, a discretized one, a common estimator for the autocovariance and the estimated value, based on this estimator.

Type	Md autocovariance	
continuous space	$c(\boldsymbol{h}) \coloneqq \mathrm{Cov}\left[\underline{x}(\boldsymbol{t}), \underline{x}(\boldsymbol{t}+\boldsymbol{h})\right] = \frac{\partial^{2L}((h_1 h_2 \dots h_L)^2 \gamma(\boldsymbol{h}))}{2^L \partial {h_1}^2 \partial {h_2}^2 \dots \partial {h_L}^2}$ where $\boldsymbol{h} := (h_1, \dots, h_L)$, with $\boldsymbol{h} \in \mathbb{R}$, the lag vector for the continue process.	(T13-1)
discretized space	$c_d^{(\Delta)}(\boldsymbol{u}) := \mathrm{Cov}\left[\underline{x}_{\boldsymbol{i}}^{(\Delta)}, \underline{x}_{\boldsymbol{i}+\boldsymbol{u}}^{(\Delta)}\right] = \frac{\Delta^{2L}\left[(u_1 u_2 \dots u_L)^2 \gamma_d^{(\Delta)}(\boldsymbol{u})\right]}{2^L \Delta[{u_1}^2]\Delta[{u_2}^2]\dots\Delta[{u_L}^2]}$ where $\boldsymbol{u} := (u_1, \dots, u_L)$, with $\boldsymbol{u} \in \mathbb{Z}$, the lag vector for the discretized process.	(T13-2)
classical estimator	$\underline{\hat{c}}_d^{(\Delta)}(\boldsymbol{u}) = \frac{1}{\zeta(\boldsymbol{u})} \sum_{i_1=1}^{n_1-j_1} \dots \sum_{i_L=1}^{n_L-j_L} \left(\underline{x}_{\boldsymbol{i}}^{(\Delta)} - \frac{\sum_{i_1=1}^{n_1} \dots \sum_{i_L=1}^{n_L} \underline{x}_{\boldsymbol{i}}^{(\Delta)}}{N}\right)\left(\underline{x}_{\boldsymbol{i}+\boldsymbol{u}}^{(\Delta)} - \frac{\sum_{i_1=1}^{n_1} \dots \sum_{i_L=1}^{n_L} \underline{x}_{\boldsymbol{i}}^{(\Delta)}}{N}\right)$ where $\zeta(\boldsymbol{u})$ is usually taken as: N or N-1 or $\prod_{r=1}^{L}(n_r - u_r)$.	(T13-3)
expected value of estimator	$\mathrm{E}\left[\underline{\hat{c}}_d^{(\Delta)}(\boldsymbol{u})\right] = \frac{1}{\zeta(\boldsymbol{u})}\left(c_d^{(\Delta)}(\boldsymbol{u}) \prod_{r=1}^{L}(n_r - j_r) + \frac{u^2}{n'}\gamma(\boldsymbol{u}\Delta) - u'\gamma(\boldsymbol{u}\Delta) - \frac{\prod_{r=1}^{L}(n_r - u_r)^2}{n'}\gamma((\boldsymbol{n}-\boldsymbol{u})\Delta)\right)$, where $u' = u_1 u_2 \dots u_L$.	(T13-4)

For example, the Ld HK process is subject to the equation below:

$$\left(\underline{x}_{\boldsymbol{i}}^{(\kappa\Delta)} - \mu\right) =_{\mathrm{d}} \kappa^{2L(1-H)}\left(\underline{x}_{\boldsymbol{j}}^{(\Delta)} - \mu\right) \tag{55}$$

where μ is the mean of the process $\underline{x}_{\boldsymbol{i}}^{(\Delta)}$ and $\underline{x}_{\boldsymbol{j}}^{(\kappa\Delta)}$ the same process at scale κ.

The Ld climacogram and autocovariance in the continuous domain can be expressed as:

$$\gamma(k_g) \coloneqq \frac{\lambda}{\left(k_g/a\right)^{2L(1-H)}} \tag{56}$$

$$c(h_m) \coloneqq \frac{\lambda\left(H(2H-1)\right)^L}{(h_m/a)^{2L(1-H)}} \tag{57}$$

where k_g is the geometric mean of scales k_1, k_2 ,..., k_M, i.e., $k_g = \sqrt{k_1 k_2 \dots k_L}$, a is a scale parameter in units of k_g, so that $\gamma\left(k_g\right) = \lambda$, and similarly, $h_m = \sqrt{h_1^2 + h_2^2 + \dots + h_L^2}$ is the lag magnitude.

For illustration we apply the HK process to the 2d climacograms of 2d images of sandstones depicted at different spatial scales (Figure 14), and we estimate a Hurst parameter equal to 0.83 (Dimitriadis et al., 2017). Note that the 2d SMA is initially suggested and implemented by Theodoratos (2004) and Koutsoyiannis et al. (2011).

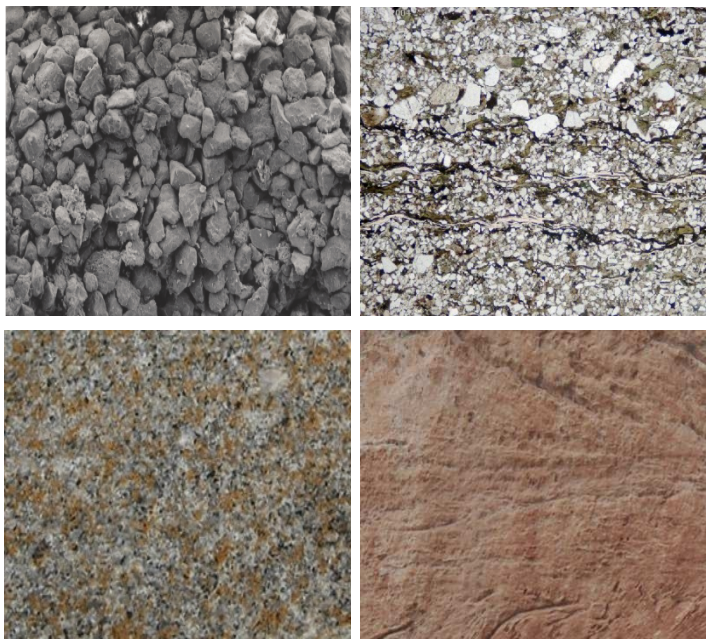

Figure 14: Images of sandstone as seen from the SEM (50 μm), from a polarizing microscope, (3.5 mm), from a hand specimen (with length approximately 5 cm) and a field outcrop (1 m). For more information on the source, description and processing of the images see in Dimitriadis et al., (2017).

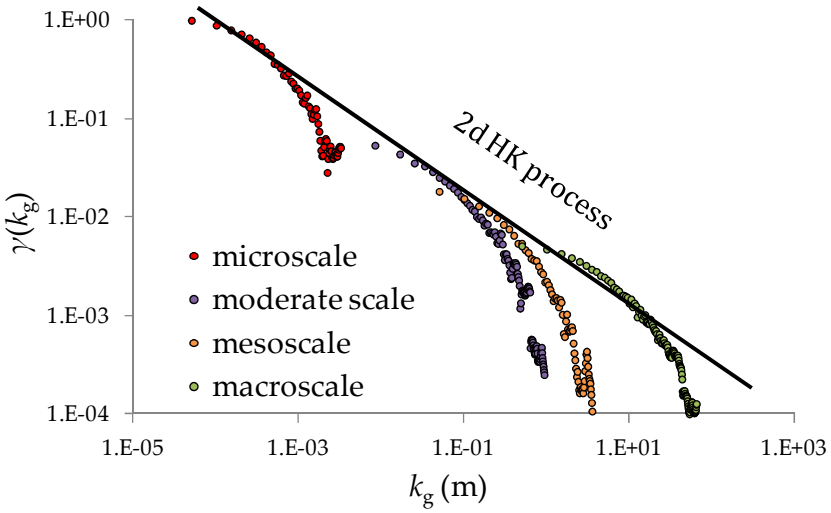

Figure 15: Climacograms of sandstone images depicted at four different scales (source: Dimitriadis et al., 2017).

3.5 Prediction algorithms

In this section, we apply two types of prediction algorithms, an analogue prediction algorithm based solely on observations without any use of models, and a stochastic prediction algorithm

based on the preservation of the marginal distribution and dependence structure of a process, permitting the prediction of unprecedented events such as extreme events.

3.5.1 Analogue prediction algorithm

Here, we apply a deterministic data-driven model known as the analogue model (e.g., Koutsoyiannis et al., 2008), which is often used in chaotic systems. This model is purely data-driven, as it does not use any mathematical expression between variables. Specifically, to predict a state $\boldsymbol{s}\big((t+l)\Delta\big)$ at future time $l\Delta$ and based on h past states $\boldsymbol{s}\big((t-r+1)\Delta\big)$, for r varying from 1 to h, we search the database of all experiments or events to find k similar states (called neighbours or analogues), $\boldsymbol{s}_j\left(\big(t_j-h+1\big)\Delta\right)$, so that for all j and r:

$$\left|\boldsymbol{s}_j\left(\big(t_j-r+1\big)\Delta\right)-\boldsymbol{s}\big((t-r+1)\Delta\big)\right|\le g \tag{58}$$

where g is an error threshold.

Then, we find for each neighbour the state at time $\big(t_j+l\big)\Delta$, i.e., $\boldsymbol{s}_j\left(\big(t_j+l\big)\Delta\right)$, and predicts the state at lead time $l\Delta$ as (e.g., Dimitriadis et al., 2016b):

$$\boldsymbol{s}\big((t+l)\Delta\big)=\frac{1}{k}\sum_{j=1}^{k}\boldsymbol{s}_j\left(\big(t_j-h+1\big)\Delta\right) \tag{59}$$

3.5.2 Stochastic prediction algorithm

Here, we describe the stochastic prediction model (Dimitriadis et al., 2016b), which is a linear stochastic model that predicts the state at lead time $l\Delta$, i.e., $\boldsymbol{s}\big((t+l)\Delta\big)$, based on the linear aggregation of weighted past states, $c_q\boldsymbol{s}\big((t-q+1)\Delta\big)$, c_q being the weighting factors. Before we calculate the weights, we need to assume a model for the stochastic structure of each process. For model fitting we choose the climacogram method since as already mentioned it results in smaller estimation errors in comparison to autocovariance (or autocorrelation) and power spectrum for this type of models. We then apply the best linear unbiased estimator (BLUE) method (Koutsoyiannis and Langousis, 2011, pp. 56-58), under the assumption of stationarity, to estimate the weighting factors c_q:

$$\boldsymbol{C}\coloneqq\begin{bmatrix}\boldsymbol{M_c} & \boldsymbol{1}\\ \boldsymbol{1}^{\mathrm{T}} & 0\end{bmatrix}^{-1}\begin{bmatrix}\boldsymbol{\eta_c}\\ 1\end{bmatrix} \tag{60}$$

where $\boldsymbol{C}=\big[c_1,\dots,c_p,\zeta\big]$; represents the vector of the weighting factors c_q (for q = 0,...,p) and ζ a coefficient related to the Lagrange multiplier of the method; $\boldsymbol{M_c}=\mathrm{Cov}\big[\underline{x}_{i-j}\big]$, for all i, j = q is the positive definite symmetric matrix whose elements are the true (included bias) autocovariances of $\underline{x}$, which represents the variable of interest (u, v, w, ξ or ψ) and now is assumed random (denoted by the underscore) for the application of this method; $\boldsymbol{\eta_c}=\mathrm{Cov}\big[\underline{x}_{l+q}\big]$ for all q; l is the index for the lead time (l > 0); the superscript T denotes the transpose of a matrix.

4 Uncertainty and HK dynamics

Although a white noise process is considered as the most unpredictable of all the processes, this is true only for very short-term predictions. For long-term predictions, which are of high interest from an engineering point of view, the maximization of entropy, and thus the uncertainty, shows that the most unpredictable process is the HK one (Koutsoyiannis, 2011). Therefore, it is only natural to assume that, eventually, a stationary process will exhibit HK behaviour. In this section, we show that the HK dynamics can arise not only in complicated deterministic systems, but in geophysical ones such as high-frequency precipitation and surface wind and even in a simple game such as die throw.

4.1 Complex natural systems

In principle, one should be able to predict the trajectory and outcome of a die throw solving the classical deterministic equations of motion; however, the die has been a popular symbol of randomness. This has been the case from ancient times, as revealed from the famous quotation by Heraclitus (ca. 540-480 BC; Fragment 52) 'Αἰών παῖς ἐστι παίζων πεσσεύων' ('Time is a child playing, throwing dice'). Die's first appearance in history is uncertain but, as evidenced by archeological findings, games with cube-shaped dice have been widespread in ancient Greece, Egypt and Persia (often in dice shaped bones). Often in history dice games were restricted or even prohibited by law perhaps for the fear of gamblers' growing passion to challenge uncertainty. Dice were also used in temples as a form of divination for oracles (Vasilopoulou, 2003). From ancient times, each side of the die represented one number from 1 to 6 so that the sum of two opposite sides was always seven. Despite dice games originating from ancient times, little has been carried out in terms of explicit trajectory determination through deterministic classical mechanics (cf., Kapitaniak et al., 2012; Nagler and Richter, 2008). Generally, a die throw is considered to be fair as long as it is constructed with six symmetric and homogenous faces (Diaconis and Keller, 1989) and for large initial rotational energy (Strzalko et al., 2010). However, statistical treatment of real experiments with dice has not been uncommon. In a letter to Francis Galton (1894), Raphael Weldon, a British statistician and evolutionary biologist, reported the results of 26,306 rolls from 12 different dice; the outcomes showed a statistically significant bias toward fives and sixes with an observed frequency approximately 0.3377 against the theoretical one of 1/3 (cf., Labby, 2009). Labby (2009) repeated Weldon's experiment (26,306 rolls from 12 dice) after automating the way the die is released and reported outcomes close to those expected from a fair die (i.e., 1/6 for each side). This result strengthened the assumption that Weldon's dice was not fair by construction. More recently, Strzalko et al. (2010) studied the Newtonian dynamics of a three dimensional die throw and noticed that a larger probability of the outcome face of the die is towards the face looking down at the beginning of the throw, which makes the die not fair by dynamics. The probability of the die landing on any face should approach the same value for any face for large values of the initial rotational and potential energy and large number of die bounces. Similar experiments of coin tossing have also been examined in the past (Diaconis et al., 2007; Jaynes, 1957,

ch. 10). According to Strzalko et al. (2010), a significant factor influencing the coin orientation and final outcome is the coin's bouncing. Specifically, they observed that successive impacts introduce a small dependence on the initial conditions leading to a transient chaotic behaviour. Similar observations are noticed in the analysis of Kapitaniak et al. (2012) in die trajectory, where lower dependency in the initial conditions is noticed when die bounces are increasing and energy status is decreasing. This observation allowed the speculation that as knowledge of the initial conditions becomes more accurate, the die orientation with time and the final outcome of a die throw can be more predictable and thus, the experiment tends to be repeatable. Nevertheless, in experiments with no control of the surrounding environment, it is impractical to fully determine and reproduce the initial conditions (e.g. initial orientation of the die, magnitude and direction of the initial or angular momentum). Although in theory one could replicate in an exact way the initial condition of a die throw, there could be numerous reasons the die path could change during its course and thus, so would the outcome. Since the classical Newtonian laws can lead to chaotic trajectories, this infinitesimal change could completely alter the rest of die's trajectory and thus, the outcome. For example, the smallest imperfections in die's shape or inhomogeneities in its density, external forces that may occur during the throw such as air viscosity or table's friction and elasticity etc., could vaguely alter dice orientation. Strzalko et al. (2010) and Nagler and Richter (2008) describe the die throw behaviour as pseudorandom since its trajectory is governed by deterministic laws while it is extremely sensitive to initial conditions. However, Koutsoyiannis (2010) argues that it is a false dichotomy to distinguish deterministic from random. Rather randomness is none other than unpredictability, which can emerge even if the dynamics is fully deterministic (see in section 4.1.2 for an example of a chaotic system resulting from the numerical solution of a set of linear differential equations). According to this view, natural process predictability (rooted to deterministic laws) and unpredictability (i.e., randomness) coexist and should not be considered as separate or additive components (see also section 1.2). A characteristic example of a natural system considered as fully predictable is the Earth's orbital motion, which greatly affects the Earth's climate (e.g., Markonis and Koutsoyiannis, 2013). More specifically, the Earth's location can become unpredictable, given a scale of precision, in a finite time-window (35 to 50 Ma, according to Laskar, 1999). Since die trajectory is governed by deterministic laws, the related uncertainty should emerge as in any other physical process and thus, there must also exist a time-window for which predictability dominates over unpredictability. In other words, die trajectory should be predictable for short horizons and unpredictable for large ones.

Here, we reconsider the uncertainty related to dice throwing (section 4.1.1). We conduct dice throw experiments to estimate a predictability window in a practical manner without implementing equations based on first principles. Furthermore, we apply the same models to high temporal resolution series of rainfall intensity and wind speed (sections 4.1.2), occurring during smooth and strong weather conditions, to acquire an insight on their similarities and differences in the process uncertainty. The predictability windows are estimated based on the aforementioned two types of models, the stochastic model fitted on experimental data using different time scales and the deterministic-chaotic model that utilizes observed patterns assuming some repeatability in the process (section 4.1.3). For validation reasons, the aforementioned models are also compared to

benchmark ones. Certainly, the estimated predictability windows are of practical importance only for the examined type of dice experiments and hydrometeorological process realizations; nevertheless, this analysis can improve our perception of what predictability and unpredictability (or randomness) are.

4.1.1 Experimental setup of dice throw

A simple mechanism is constructed with a box and a high speed camera in order to record the die motion for further analysis. For this experiment we use a wooden box with dimensions 30 cm x 30 cm x 30 cm and white colour painted to easily distinguish it from the die. The die is of acetate material with rounded corners, has dimensions 1.5 × 1.5 × 1.5 cm^3 and weighs 4 g. Each side of the die has been painted with a different colour: yellow, green, magenta, blue, red and black, for 1, 2, …, 6 pips, respectively. Instead of the primary colour cyan, we use black to be easier traceable contrasting to the white colour of the box. The height (30 cm) from which the die is released with zero initial momentum or thrown, remained constant for all experiments. However, the die is released or thrown with a random initial orientation and momentum, so that the results of this study are independent of the initial conditions. Specifically, 123 die throws are performed in total, 52 with initial angular momentum in addition to the initial potential energy as well as 71 with the initial potential energy only (Figure 16). Despite the similar initial energy status of the die throws, the duration of each throw varied from 1 to 9 s, mostly due to the die's cubic shape that allowed energy to dissipate at different rates.

Figure 16: Mixed combination of frames taken from all die throw experiments for illustration (source: Dimitriadis et al., 2016b).

Visualization of the die's trajectory is done via a digital camera with 0.045 cm/pixel density of and frame resolution rate of 120 Hz. The camera is placed in a fixed location and symmetrically at the top of the box. The video is analysed to frame by frame and numerical codes are assigned to

coloured pixels (based on the HSL system) and die's position inside the box. Specifically, three coordinates are recorded based on the Cartesian orthogonal system; the two horizontal ones are taken from the box's plan view while the die's height above the box bottom is estimated from die's image size (the higher the die, the larger the die's size in pixels). Moreover, the area of each colour traced by the camera is estimated and then is transformed to a dimensionless value divided by the total traced area of the die. In this manner, the orientation of the die in each frame can also be estimated (with some observational error) through the traced colour triplets. Note that pixels not assigned to any colour due to relatively low resolution and blurriness of the camera, are on average approximately 30% of the total traced die area in each frame.

Finally, the audio signal is transformed to a dimensionless index from 0 to 1 (with 1 indicating the highest sound produced in each experiment) and can be used to record the times the die bounces colliding with the bottom or the sides of the box, contributing in this way to sudden changes in die's orientation, to its orbit and thus, to the final outcome. We observe that die bounces decay faster than kinetic energy status (roughly estimated through linear velocity). Also, most of the die bounces and energy dissipation occur approximately during initial 1.5 s, regardless of the initial conditions of the die throw. Based on these observations, we expect predictability to improve after the first 1.5 s (Figure 18).

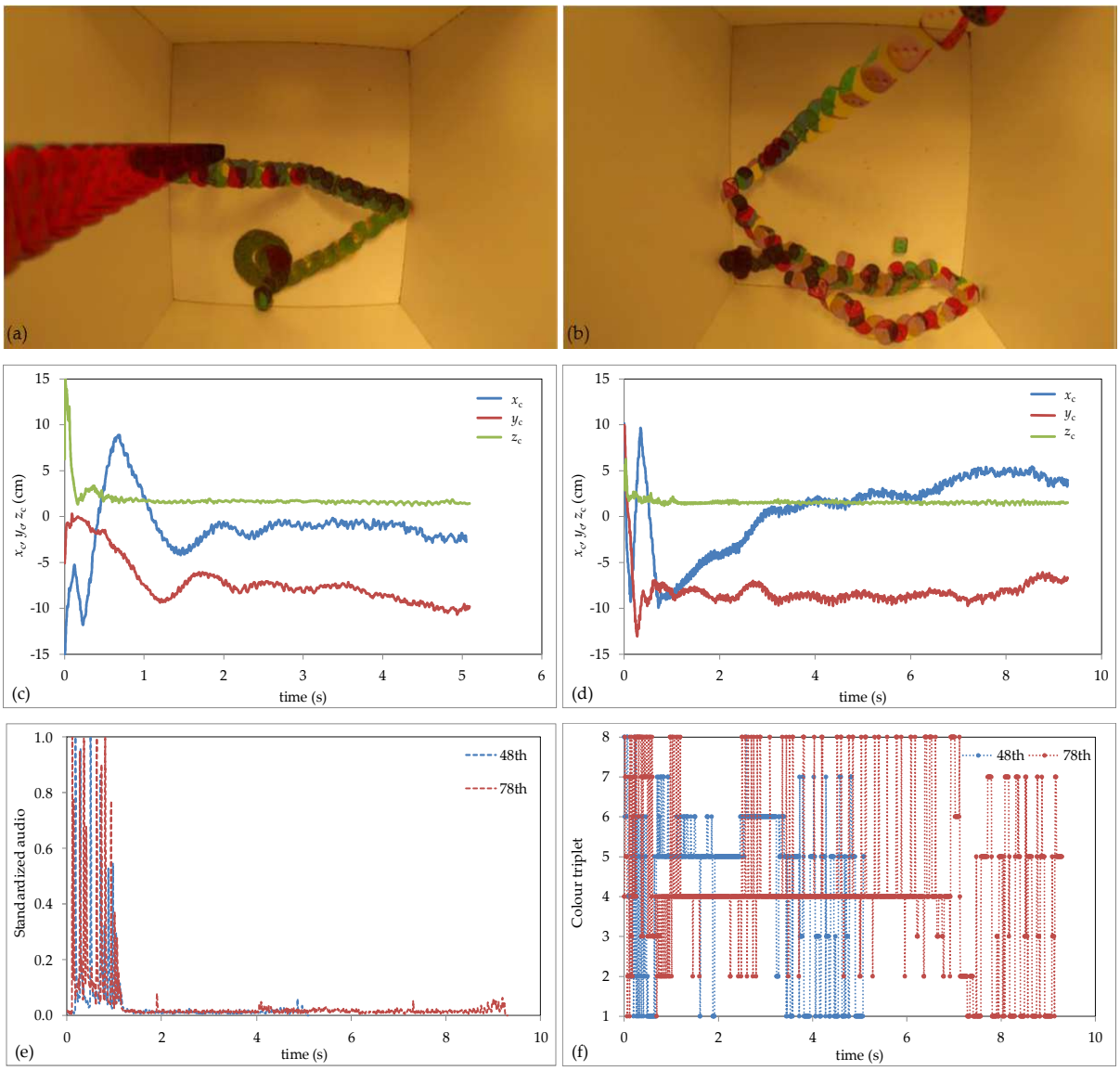

Figure 17: Selected frames showing the die trajectory from experiments (a) 48 and (b) 78; (c, d) their three Cartesian coordinates (denoted x_c, y_c and z_c for length, width and height, respectively); (e) standardized audio index representing the sound the die makes when colliding with the box; and (f) colour triplets (each of the 8 possible triplets corresponds to three neighbouring colours). Source: Dimitriadis et al. (2016b).

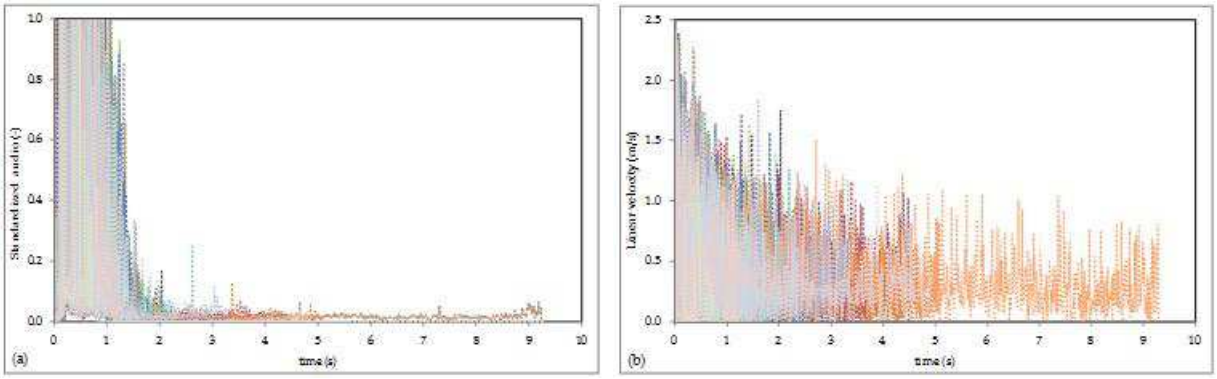

Figure 18: All experiments (a) standardized audios, showing the time the die collides with the box (picks) and (b) linear velocities. Source: Dimitriadis et al. (2016b).

To describe the die orientation we use three variables (x, y and z) representing proportions of each colour, as viewed from above, each of which varies in [-1,1], with x, y, z = 1 corresponding to black, blue or green, respectively, and with x, y, z = -1 to the colour of the opposite side, that is yellow, magenta or red, respectively (Table 14). In Figure 19 we show two examples of dice orientation recorded through colour identification.

Table 14: Definition of variables x, y and z that represent proportions of each pair of opposite colours (source: Dimitriadis et al., 2016b).

Value →	−1		+1	
Variable ↓	Colour	Pips	Colour	Pips
x	yellow	1	black	6
y	magenta	3	blue	4
z	red	5	green	2

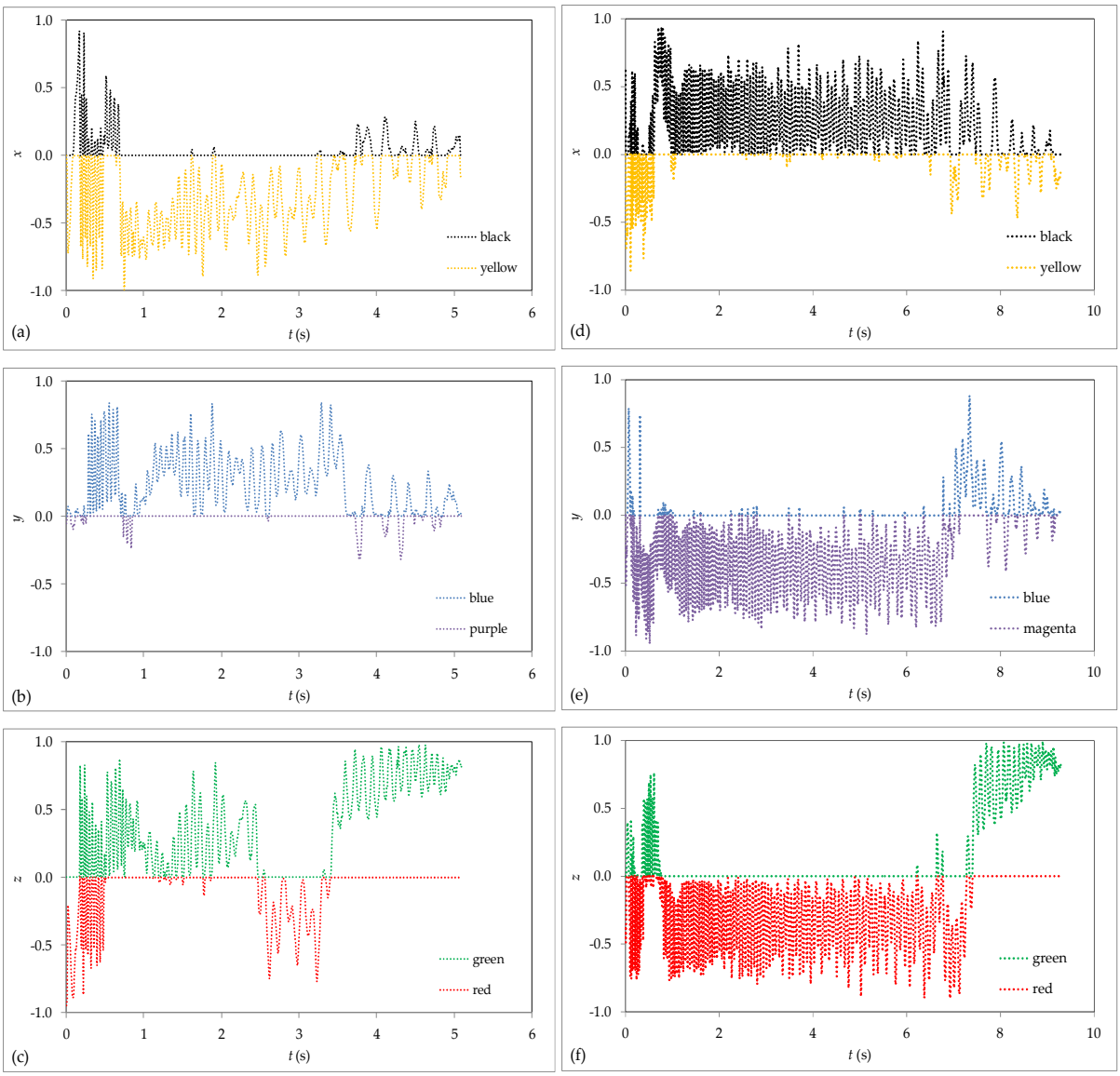

Figure 19: Time series of variables x, y and z for experiments 48 (a, b, c) and 78 (d, e, f); in both experiments the outcome was green. Source: Dimitriadis et al. (2016b).

However, these variables are not stochastically independent to each other because of the obvious relationship:

$$|x| + |y| + |z| = 1 \tag{61}$$

The following transformation produces a set of independent variables u, v and w, where u and v vary in [-1,1] and w is a two-valued variable taking either the value -1 or 1:

$$u = x + y, v = x - y, w = \text{sign}(z) \tag{62}$$

The inverse transformation is:

$$x = (u + v)/2, y = (u - v)/2, z = w\big(1 - \max(|u| + |v|)\big) \tag{63}$$

In Figure 20, the plot of all experimental points and the probability density function (pdf) show that u and v are independent and fairly uniformly distributed except the more probable states where $u \pm v = 0$ (corresponding to one of the final outcomes). Note that w outcomes are also nearly uniform with $P(w = -1) \approx 54\%$ and $P(w = 1) \approx 46\%$.

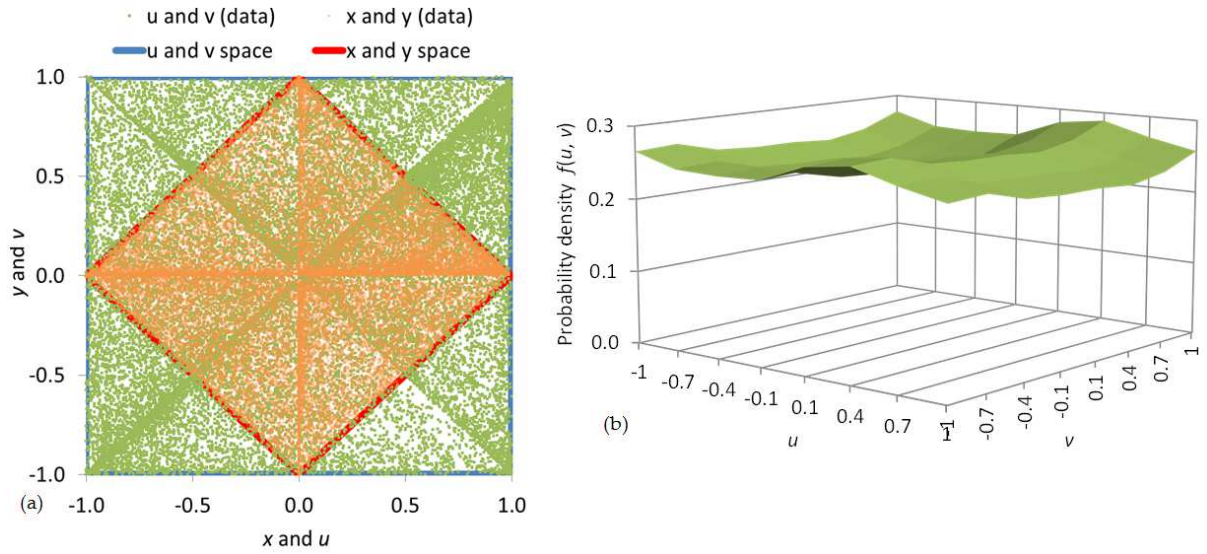

Figure 20: Plot of (a) all (x, y) and (u, v) points from all experiments and (b) the probability density function of (u, v). Source: Dimitriadis et al. (2016b).

4.1.2 Hydrometeorological processes of high resolution

Here, we choose a set of high resolution time series of rainfall intensities (denoted by ξ and measured in mm/h) and wind speed (denoted by ψ and measured in m/s). The rainfall intensities data set consists of seven time series with a 10 s time step recorded during various weather states (such as low precipitation and storm events) and are provided by the Hydrometeorology Laboratory at the Iowa University (for more information regarding the database see Georgakakos et al., 1994). The wind speed database consists of five time series with a 1 min time step recorded during various weather states (such as strong breeze and storm events) by a sonic anemometer on a meteorological tower located at Beaumont KS and provided by NCAR/EOL (http://data.eol.ucar.edu/). We have chosen these processes as they are of high interest in hydrometeorology and often are also regarded as random-driven processes. For illustration we show in Figure 21 a couple timeseries drawn from the above datasets.

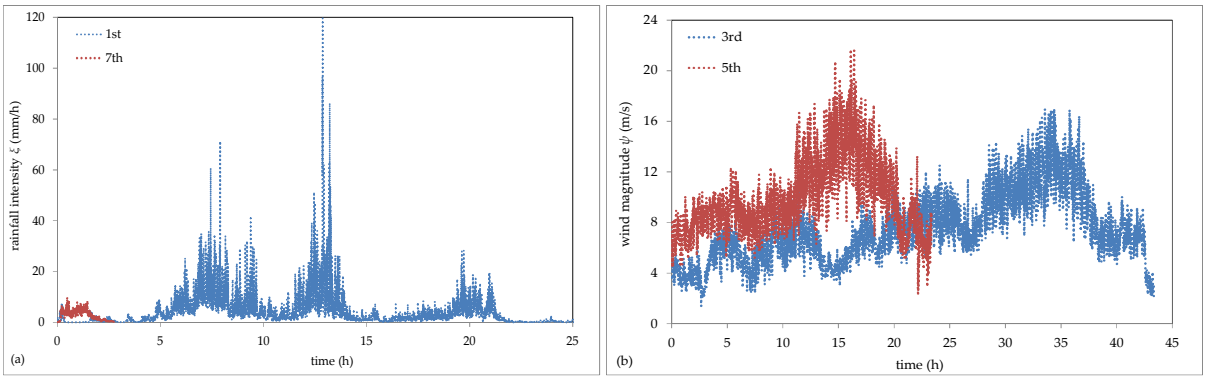

Figure 21: (a) Rainfall events 1 and 7 from Georgakakos et al. (1994) and (b) wind events 3 and 5 provided by NCAR/EOL.

4.1.3 Uncertainty evaluation and comparison

Here, we apply two types of prediction algorithms in each case and we compare them to each other for the same process and to the other processes, in terms of the Nash and Sutcliffe (1970) efficiency coefficient defined as:

$$F = 1 - \frac{\sum_{d=1}^{n}\sum_{i=0}^{b_d}\left(\boldsymbol{s}_d(i) - \hat{\boldsymbol{s}}_d(i)\right)^2}{\sum_{d=1}^{n}\sum_{i=0}^{b_d}\left(\hat{\boldsymbol{s}}_d(i) - \bar{\hat{\boldsymbol{s}}}_d(i)\right)^2} \tag{64}$$

where d is an index for the sequent number of the die experiments, rainfall or wind events; i denotes time; n is the total number of the experiments, or of recorded rainfall or wind events (n = 123 for the die throw experiment, n = 7 for the rainfall and n = 5 for the wind events); b_d is the total number of recorded frames in the dth experiment, rainfall or wind event; $\hat{\boldsymbol{s}}$ is the vector $\left(\hat{u}_d(i), \hat{v}_d(i), \hat{w}_d(i)\right)$, transformed from the originally observed $\left(\hat{x}_d(i), \hat{y}_d(i), \hat{z}_d(i)\right)$, for the die throw, the 1d rainfall intensity $\hat{\xi}_d(i)$ for the rainfall events and the 1d wind speed $\hat{\psi}_d(i)$ for the wind events, with $\bar{\hat{\boldsymbol{s}}}$ the corresponding mean empirical discrete-time vector; and $\boldsymbol{s}$ is the discrete-time vector estimated from the model.

Also, the prediction models described above are checked against two naïve benchmark models. At the first benchmark model (abbreviated B1), the prediction is the average state, i.e.:

$$\boldsymbol{s}\left((t+l)\Delta\right) = \boldsymbol{0} \tag{65}$$

where $t\Delta$ is present time in s, $l\Delta$ the lead time of prediction in s ($l > 0$) and Δ the sampling frequency (equal to 1/120 seconds per frame for the die throw game, 10 seconds per record for the rainfall events and 1 minute per record for the wind events). Although the zero state is not permissible per se, the B1 is useful, as any model worse than that is totally useless. At the second benchmark model (abbreviated B2), the prediction is the current state regardless of how long the lead time $l\Delta$ is, i.e.:

$$\boldsymbol{s}\big((t+l)\Delta\big) = \boldsymbol{s}(t\Delta) \tag{66}$$

The observed climacograms of the processes under investigation show the strong dependence of the die orientation, rainfall intensity and wind speed in time (long-term, rather than short-term persistence). This enables stochastic predictability up to a certain lead time. Here, we choose the gHK model for the mathematical process, i.e., with climacogram $\gamma(\kappa\Delta)$ as in Table 8, where Δ is the time resolution parameter, i.e., $1/120$ s for the die experiments, 10 s for the rainfall events and 1 min for the wind events. For consideration of the bias effect due to varying sample sizes n of the die experiments and rainfall and wind events, we estimate the average of all empirical climacograms for experiments and events of similar sample size. However, due to the strong climacogram structure of all three processes, the varying sample size has small effect on the shape of the climacogram for scales approximately up to 10% of the sample size (following the rule of thumb for this type of models, as analysed by Dimitriadis and Koutsoyiannis (2015a). Thus, we consider the averaged empirical climacogram to represent the expected one. The fitted models are shown in Figure 22 in terms of their climacograms. Their parameters are: for the u and v symmetric variables of the dice process $\lambda = 0.6, q = 0.013\ s$ and $b = 0.83$ ($H = 0.6$); for the w variable $\lambda = 1.635, q = 0.0082\ s$ and $b = 1.0$ ($H = 0.5$); for the rainfall process $\lambda = 12.874\ \text{mm}^2/\text{h}^2, q = 130\ \text{s}$ and $b = 0.22$ ($H = 0.9$); and for the wind process $\lambda = 65.84\ \text{m}^2/\text{s}^2$, $q = 86\ \text{min}$ and $b = 0.09$ ($H = 0.95$). We observe that the scale parameter q and Hurst coefficient H are largest in the wind process and smallest in the dice one.

Note that two additional criteria for the two above model parameters is that firstly, they should give an efficiency coefficient greater than that of the B2 model (at least for most of the lead times) and secondly, their efficiency values are estimated from a reasonable large set of tracked neighbours (>10% of the total number of realizations for each process). Due to high variances of the time averaged process (which correspond to high autocorrelations), it is expected that the B2 model will work well, for fairly small lead times. Next, we depict the results for the four models for the 48th die experiment, the 1st rainfall event and the 3rd wind event (Figure 23). The stochastic model provides relatively good predictions ($F \gtrsim 0.5$ and efficiency coefficients larger than the B2 and B1 models) for lead times $l\Delta \lesssim 0.1$ s for the die experiments (with a range of approximately 0.05 to 0.5 s), $\lesssim 5$ min for the rainfall events (with a range of approximately 1 min to 30 min) and $\lesssim 1$ h for the wind events (with a range of approximately 0.1 h to 2 h). The analogue model gives smaller F values than the B2 model for the die experiments and the wind process and larger in case of the rainfall process (but smaller than the stochastic model). Predictability is generally good for small lead times; however, the situation deteriorates for larger ones. Finally, we define and estimate the predictability-window (that is the window beyond which the process is considered as unpredictable), as the time-window beyond which the efficiency coefficient F becomes negative. Specifically, predictability is superior to the case of a pure random process (B1) for lead times $l\Delta \lesssim$ 1.5 s for the die throw process, $l\Delta \lesssim 1$ h for the rainfall process and $l\Delta \lesssim 4$ h for the wind process.

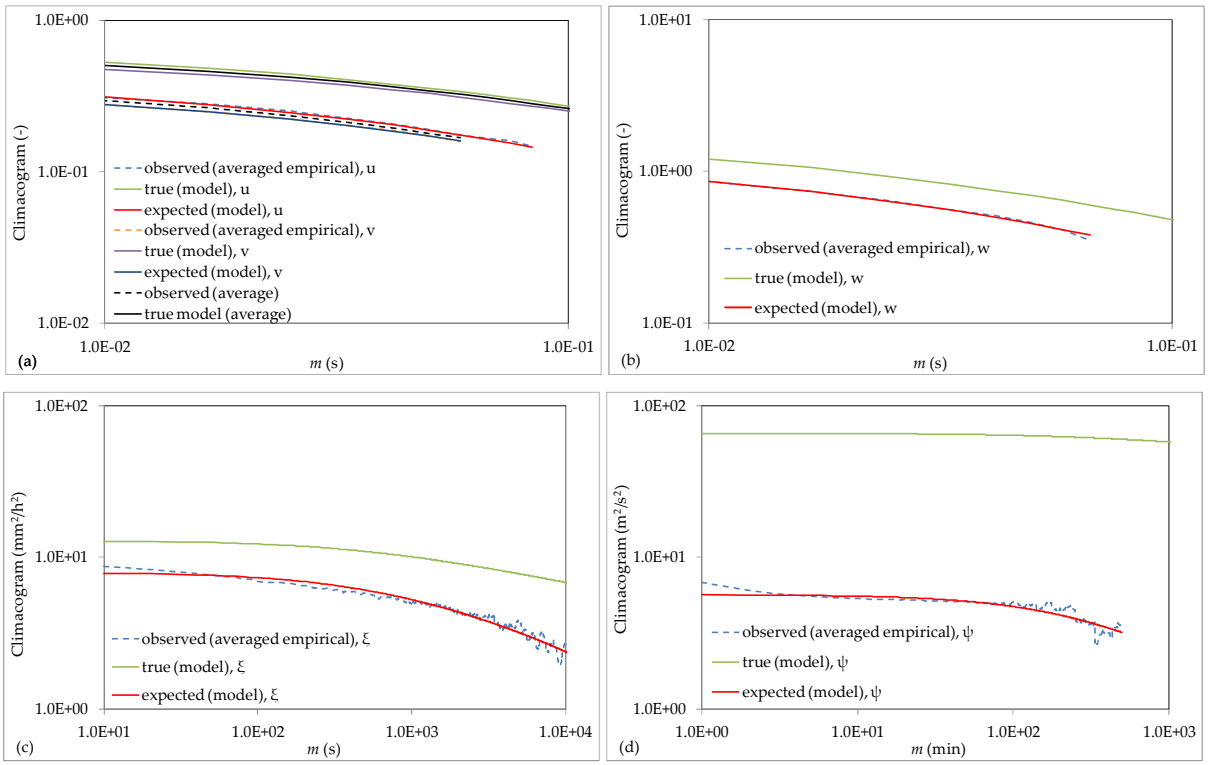

Figure 22: True, expected and averaged empirical climacograms for (a) u and v, (b) w, (c) ξ and (d) ψ. Source: Dimitriadis et al. (2016b).

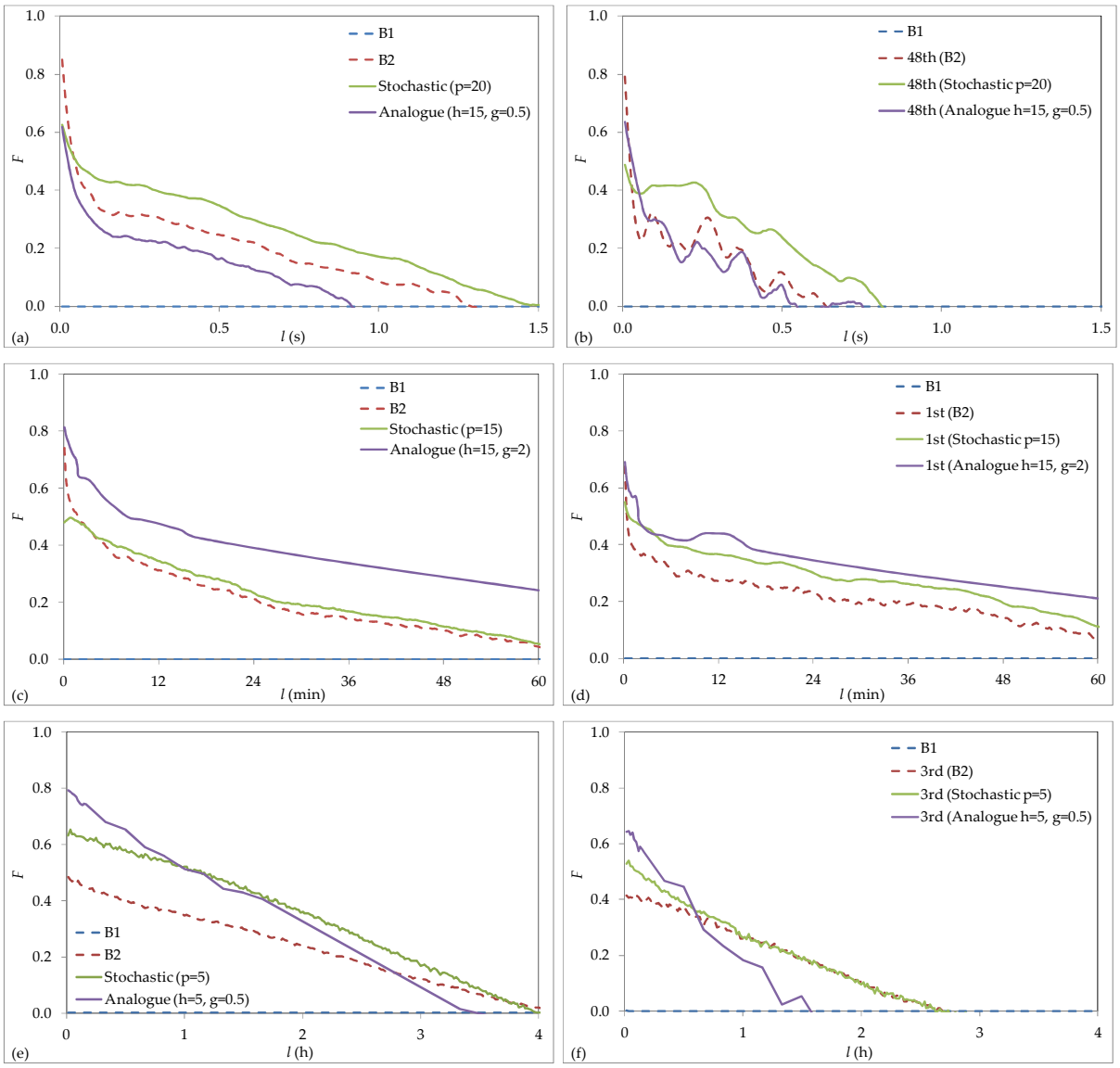

Figure 23: Comparisons of B1, B2, stochastic and analogue models for the die experiment (a and b), the observed rainfall intensities (c and d) and the observed wind speed (e and f). The left column (a, c and e) represents the application of the models to all experiments and events and the right column (b, d and f) to individual ones. Source: Dimitriadis et al. (2016b).

Next in Figure 24, we show the sensitivity analysis applied to each process and for both stochastic and analogue models. Specifically, we apply a variety of p values (i.e., number of past states that the model assumes the future state is depending on) for the stochastic model and combinations of h (same as p) and g (i.e., error threshold value for selecting neighbours) values for the analogue one. Employing a sensitivity analysis to the analogue model, we conclude that for the die process a value of p = 20 (which corresponds to time length ~0. 17 s) works relatively well (on the concept that it is a small value giving a large F), for l varying from 8 ms to 1.5 s (for larger values of p we have negligible improvement of the efficiency). Similarly, for the rainfall process, we concluded that p = 150 s is adequate, for l varying from 10 s to 1 h (the variation of l is set equal to half the minimum

duration between events). Finally, for the wind process, we concluded that p = 5 min works well, for l varying from 1 min to 6 h. Applying a sensitivity analysis for the stochastic model, we found that a number of past values h= 15 (which corresponds to time length ~0.125 s) and a threshold g = 0.5 work relatively well for the die process. Similarly, for the rainfall process, we conclude that h = 15 (which corresponds to time length 150 s) and a threshold g = 2 mm/h works well. Finally, we concluded that h = 5 (which corresponds to time length 5 min) and a threshold g = 0.5 m/s works well for the wind process.

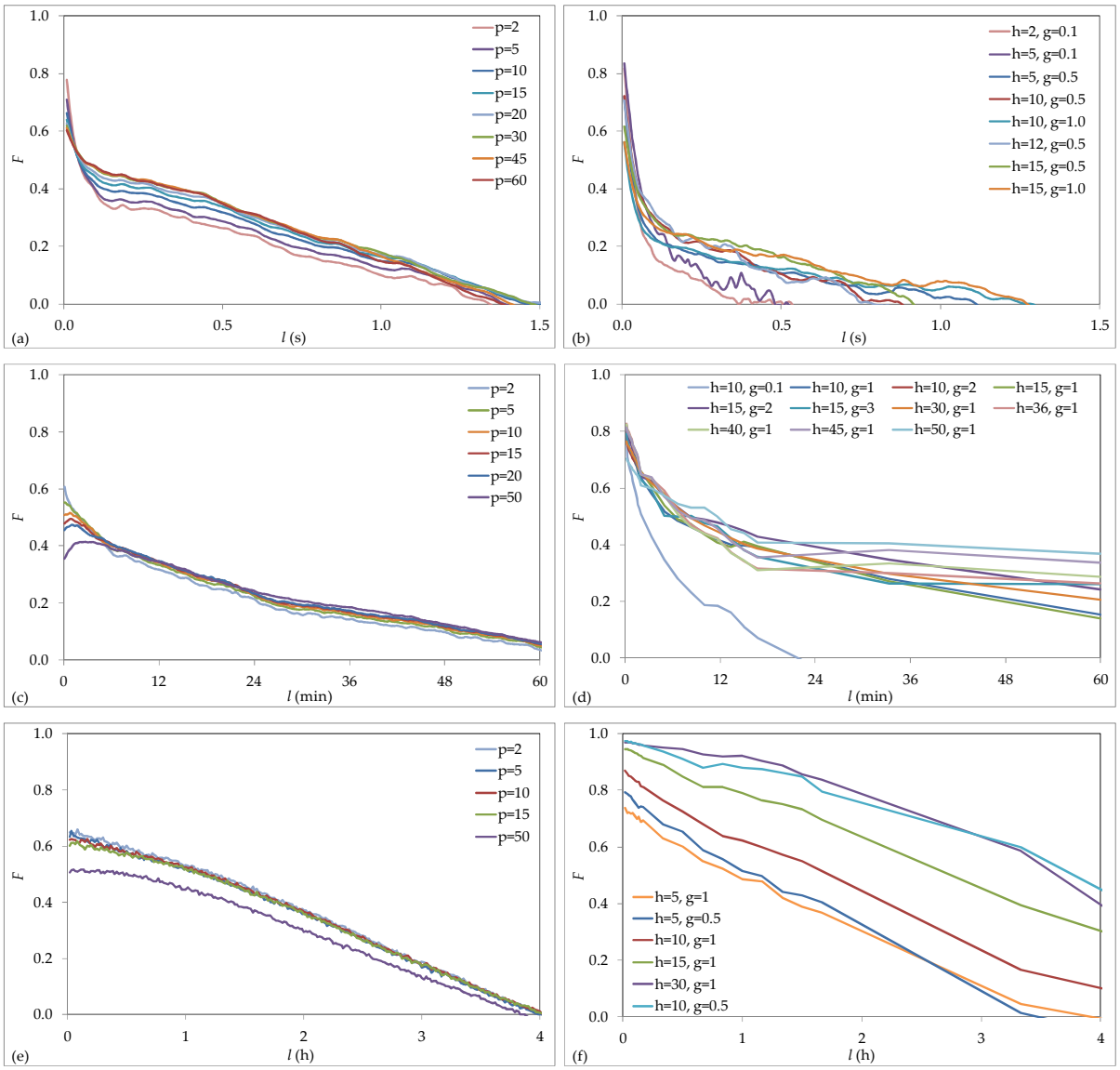

Figure 24: Sensitivity analyses of the stochastic and analogue model parameters for the die experiment (a and b), the rainfall intensities (c and d) and the wind speed (e and f). Source: Dimitriadis et al. (2016b).

4.2 Deterministic systems

Here, we show various examples of deterministic systems and application to benchmark and real-case scenarios. By definition, these systems will exhibit Markov behaviour rather than HK, and therefore, their window of predictability is expected to be short, a result which contradicts our experience and thus, reality.

4.2.1 A classical deterministic system

As a prelude example, we apply all models described above to a set of timeseries produced by numerically solving the classical Lorenz (1963) chaotic system of differential equations. Specifically, using the Runge-Kutta integration approach (Press et al., 2007), we produce n = 100 timeseries of the Lorenz-system dimensionless variables (denoted $\underline{X}_{\mathrm{L}}$, $\underline{Y}_{\mathrm{L}}$ and $\underline{Z}_{\mathrm{L}}$), with randomly varying initial values of variables between -1 and 1, a time step of d_t=Δ=0.01 (dimensionless), a total time length of T_{L}=10^3 (so, each timeseries contains N = 10^5 data) and with the classical Lorenz-system dimensionless parameters of σ_{L}=10, r_{L}=8 and b_{L}=8/3 (Lorenz, 1963):

$$\left\{\begin{array}{c} \frac{dX_{\mathrm{L}}}{dt} = \sigma_{\mathrm{L}}(Y_{\mathrm{L}} - X_{\mathrm{L}}) \\ \frac{dX_{\mathrm{L}}}{dt} = r_{\mathrm{L}}X_{\mathrm{L}} - Y_{\mathrm{L}} - X_{\mathrm{L}}Z_{\mathrm{L}} \\ \frac{dX_{\mathrm{L}}}{dt} = X_{\mathrm{L}}Y_{\mathrm{L}} - b_{\mathrm{L}}Z_{\mathrm{L}} \end{array}\right\} \tag{67}$$

The 5th timeseries is shown in Figure 25 along with the results from the stochastic and analogue models. The estimated parameters for the best fitted (Markov-type) stochastic model are $\lambda = 72.8, q = 0.13$ for X_{L} process, $\lambda = 93.1, q = 0.0836$ for Y_{L} and $\lambda = 272, q = 0.0007$ for Z_{L}, with $b = 1.0$ ($H = 0.5$) for all processes. From the analysis, we concluded that the analogue model, with h = 2 (which corresponds to time length 0.02 s) and a threshold of g = 0.1, works very well as opposed to the stochastic model whose efficiency factor is always lower than the one corresponding to B2 model. We believe this is because the system's dynamics is relatively simple and no other factors affect the trajectory. Such conditions are never the case in a natural process and thus, the performance of the analogue model is usually of the same order (given there are many data available, in contrast with the stochastic which can be set up with much fewer data). Finally, we can also see here, that predictability is generally superior to a pure random process (B1), for lead times $l\Delta \lesssim 1$.

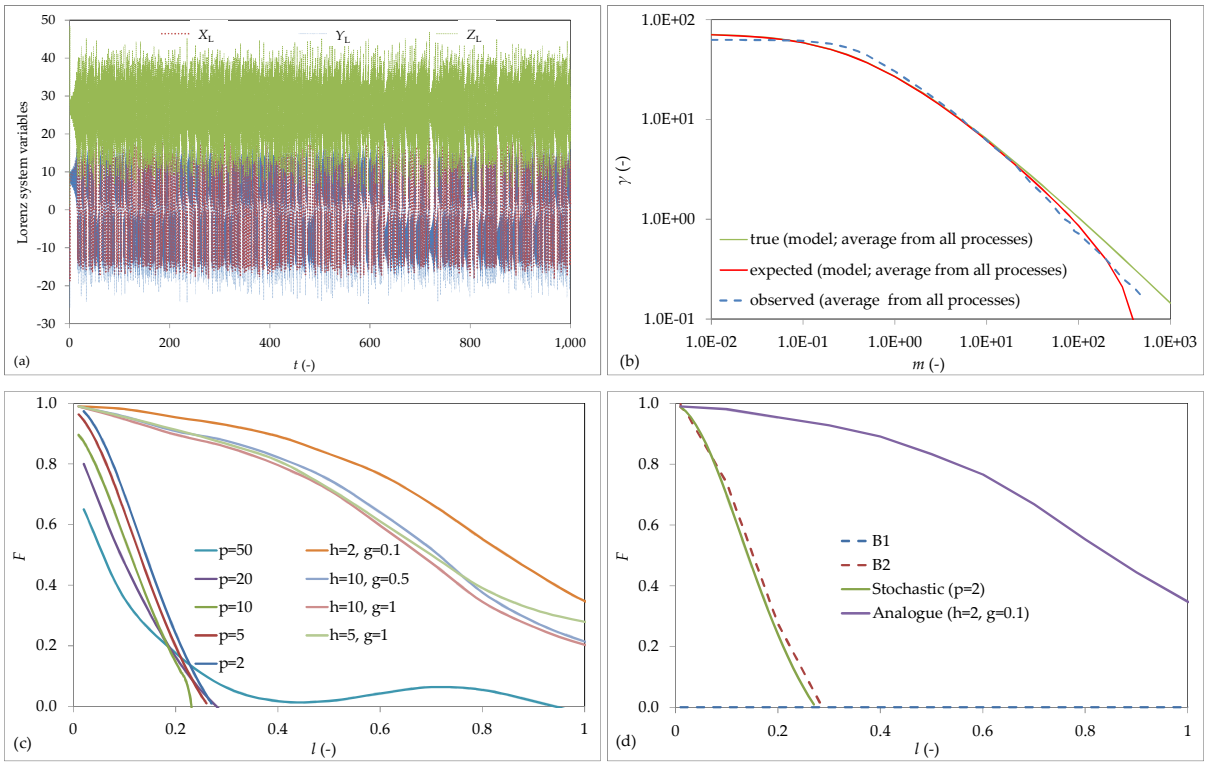

Figure 25: (a) Values of X_L, Y_L and Z_L, plotted at a time interval of 0.1, for the 5th timeseries produced by integrating the classical Lorenz's chaotic system of equations, (b) observed climacogram as well as its true and expected values for the fitted stochastic gHK model (average of X_L, Y_L and Z_L processes), (c) sensitivity analysis of the analogue and stochastic models and (d) comparison of the optimum stochastic and analogue models with B1 and B2. Source: Dimitriadis et al. (2016b).

4.2.2 Comparison between deterministic systems of high complexity

Here, we show some examples of deterministic systems of simplified hydraulic wave inundation models. Although all parameters and equations are a priori selected and known in an exact way, we show that they exhibit a large sensitivity to initial and boundary conditions as well as to discretization schemes (Dimitriadis et al., 2016c).

In general, flood routing models solve part or the full one-dimensional (1d) Saint-Venant continuity and momentum depth-averaged equations in the longitudinal direction (1d models) or, additionally, in the lateral direction (quasi-2d or 2d models). The 1d Saint-Venant continuity and momentum equations are (Chow et al., 1988, p. 279):

$$\frac{\partial A}{\partial t} + \frac{\partial Q}{\partial x} = 0 \tag{68}$$

$$\frac{1}{A}\frac{\partial Q}{\partial t} + \frac{1}{A}\frac{\partial (Q^2/A)}{\partial x} + \mathrm{g}\frac{\partial w}{\partial x} = \mathrm{g}(S_\mathrm{o} - S_\mathrm{e}) \tag{69}$$

where Q is the discharge, A is the wetted area, g is the gravity acceleration, w is the water depth, S_o is the longitudinal bed slope (expressing the gravitational force), S_e is the energy (or else friction) slope, and $\partial w/\partial x$, $\frac{1}{A}\partial Q/\partial t$ and $\frac{1}{A}\partial(Q^2/A)/\partial x$ represent the pressure gradient and the local and convective acceleration terms of the momentum equation.

The three hydraulic tools that are used in next analyses are described briefly below.

HEC-RAS

HEC-RAS is a widely used hydraulic software tool developed by the U.S Army Corps of Engineers, which is usually combined with the HEC-HMS platform for hydrological simulations (hec.usace.army.mil). HEC-RAS employs 1d flood routing in both steady and unsteady flow conditions by applying an implicit-forward finite difference scheme between successive sections of flexible geometry. Due to the 1d nature of the model, the discharge is distributed within the whole cross section in the longitudinal direction. This can create difficulties when multiple flow directions are required or when the flow exchange between the channel and the floodplain cannot be neglected. However, it can sufficiently represent the topography since it is not raster-based, it has quite low computational cost and it is very powerful in 1d steady flow simulations. The steady flow scheme is based on the solution of the 1d energy equation (for gradually-varied conditions) or the momentum equation (for rapidly-varied conditions) between two successive cross sections:

$$\Delta Y + a_2\frac{{V_2}^2}{2g} - a_1\frac{{V_1}^2}{2g} = L\bar{S}_e + C\left|a_2\frac{{V_2}^2}{2g} - a_1\frac{{V_1}^2}{2g}\right| \tag{70}$$

$$b_2\frac{Q_2}{A_2} - b_1\frac{Q_1}{A_1} + g\frac{(A_2\bar{Y}_2 - A_1\bar{Y}_1)}{\bar{A}L} = g(S_o - \bar{S}_e) \tag{71}$$

where Y is the water surface elevation and ΔY is the residual between the upstream and downstream cross sections, Q_1, A_1 and Q_2, A_2 are the discharge and wetted area of the upstream and downstream cross sections, a_1, b_1 and a_2, b_2 are velocity and momentum correction coefficients (for a non-uniform distribution), L is the flow-weighted reach length, $\bar{S}_e$ is the representative energy slope between two cross sections and C is the expansion or contraction loss coefficient (representing the magnitude of the loss of energy between two expanding or contracting cross sections).

For unsteady conditions, the model uses the 1d Saint-Venant set of equations:

$$\frac{\partial A}{\partial t} + \frac{\partial(\varphi Q)}{\partial x_c} + \frac{\partial\big((1-\varphi)Q\big)}{\partial x_f} = 0 \tag{72}$$

$$\frac{\partial Q}{\partial t} + \frac{\partial(\varphi^2 Q^2/A_c)}{\partial x_c} + \frac{\partial((1-\varphi)^2 Q^2/A_f)}{\partial x_f} + g\left(A_c\left(\frac{\partial Y_c}{\partial x_c} + S_{ec}\right) + A_f\left(\frac{\partial Y_f}{\partial x_f} + S_{ef}\right)\right) = 0 \tag{73}$$

where the subscripts c and f refer to the channel and floodplain, a variable specifying how flow is partitioned between the channel and floodplain:

$$\varphi = 1/\left(1 + \frac{n_c}{n_f}(A_f/A_c)^{5/3}(P_f/P_c)^{-2/3}\right) \quad (74)$$

with A_c, P_c and A_f, P_f are the wetted area and perimeter, related to the channel and floodplain, respectively. Note that the energy slope is approximated by the Manning's equation.

Further details on the mathematical background of HEC-RAS are provided in the associated documentation manual (Brunner, 2010).

LISFLOOD-FP

LISFLOOD-FP (bristol.ac.uk) is a quasi-2d, raster-based model that is appropriate for both steady and unsteady flow conditions. It allows using a high resolution grid-based topographic terrain and is more suitable for large basins with wide and shallow channels, since it assumes a rectangular channel section and so, it approximates the wetted perimeter by the channel width. It can process up to 10^6 grid cells, thus being suitable for implementing probabilistic investigations based on Monte Carlo approaches. The channel's flood routing is handled using the 1d kinematic wave (in case of positively varying channel gradient) or the diffusive wave (in case of negative channel gradient), which are solved with a backward-implicit numerical scheme. The diffusive wave scheme is also used for lateral flow propagation (floodplain inundation), where the 1d channel and floodplain routings are linked via a quasi, two-dimensional continuity equation (Bates et al., 2013):

$$\frac{\partial A}{\partial t} + \frac{\partial Q}{\partial x} = q \quad (75)$$

$$\frac{\partial w}{\partial x} + \frac{Q^2 n^2 P^{4/3}}{A^{10/3}} - S_o = 0 \quad (76)$$

where q is the flow exerting from the channel to the floodplain. In this approach, it is assumed that the flow between two adjacent cells is linearly interpolated between the known water depths of the cells.

FLO-2d

FLO-2d basic (flo-2d.com) is also raster-based and allows for flexible geometry of the channel and the floodplain terrain. It solves the 1d Saint-Venant set of equations using an explicit-central finite difference scheme and, thus, it can describe in a more detail the flow wave propagation along the channel and floodplain. It is more suitable for large grid cell size since it may be time consuming when processing a high number of cells. For the floodplain, the equations of motion are applied by computing independently the average flow velocity across each one of eight potential flow directions (O'brien, 2007):

$$\frac{\partial w}{\partial t} + \frac{\partial (wV)}{\partial x} = 0 \quad (77)$$

$$\frac{\partial V}{\partial t} + V\frac{\partial V}{\partial x} + g\frac{\partial w}{\partial x} = g(S_o - S_e) \quad (78)$$

where V is the depth-averaged velocity in one of the eight flow longitudinal directions, while the energy slope component S_e is based on the Manning's equation.

Flow and boundary conditions

In all the above models, two boundary conditions are required, which are usually set at the upstream end of the channel through an imposed inflow as well as the assumption of uniform water depths at the upstream and downstream end (kinematic wave condition). Although an imposed depth would result in more stable solutions than the condition of uniform flow, we choose the latter since, in practice, it is rare to know the temporal evolution of the water depth at a specific location. The models compute the appropriate time step based on the Courant number stability criteria (Courant et al., 1959).

It can be illustrated that the uncertainty of the flood volume (which can be regarded as the wetted area over length) corresponding to a triangular cross section is often larger than that of a rectangular one. This is due to the fact that the area of a triangular cross section is a function of the square of the water depth $\underline{w}$, i.e., $\underline{A}_t = \underline{z}\,\underline{w}^2$, where $\underline{z}$ is the tangent of the interior angle of the section, in contrast to the rectangular one which is linear function of $\underline{w}$, i.e., $\underline{A}_r=\underline{b}\,\underline{w}$, where $\underline{b}$ is the section width. Considering the uncertainty associated to a random variable as being proportional to its variation coefficient C_v and the water depth as being stochastically independent of the geometrical characteristics of the channel, we get for the rectangular cross section that $C_v{}^2[\underline{A}_r] = \mathrm{Var}[\underline{A}_r]/\mathrm{E}^2[\underline{A}_r]$, which after algebraic manipulations $C_v{}^2[\underline{A}_r] = (C_v{}^2[\underline{b}] + 1)\mathrm{E}[\underline{w}^2]/\mathrm{E}^2[\underline{w}] - 1$ and equivalently for the triangular cross section we get $C_v{}^2[\underline{A}_t] = (C_v{}^2[\underline{z}] + 1)\mathrm{E}[\underline{w}^4]/\mathrm{E}^2[\underline{w}^2] - 1$. Furthermore, considering the water depth as being uniformly distributed, i.e., $\underline{w}\sim U(0,2\mu)$, with μ its mean value, we have that $C_v{}^2[\underline{A}_r] = \frac{4}{3}(C_v{}^2[\underline{b}] + 1) - 1$ and equivalently, $C_v{}^2[\underline{A}_t] = \frac{9}{5}(C_v{}^2[\underline{z}] + 1) - 1$. Thus, if we assume that $C_v[\underline{b}] \approx C_v[\underline{z}]$, then $C_v[\underline{A}_r] < C_v[\underline{A}_t]$. For this reason, we apply a triangular-like cross section (Figure 26), which appears quite often in field (compared to the rectangular one). Moreover, this type of section permits the development of lateral flow wave propagation (as opposed to the rectangular one) and thus, is convenient for observing the differences between 1d and 2d models.

Benchmark experiments

Initially, we test the above models in theoretical applications to identify the impacts of the different mathematical schemes and other assumptions in terms of uncertainty. In this respect, we employ sensitivity analysis against the most important hydraulic variables (inflow, channel and floodplain slope and roughness), as well as the model resolution (see Figure 26 and Table 15).

We consider six model configurations, by running HEC-RAS and LISFLOOD-FP in both steady and unsteady conditions, and FLO-2d with including or not the wave propagation along the channel. Note that when we omit the channel's flow propagation we still apply the channel's friction at the

same cells that would be overlaid by the channel. Also, when we refer to unsteady conditions (but with constant inflow), we mean that at the beginning we apply an increasing (i.e., starting from zero) inflow and then we stabilize it to the desired constant value, in order to achieve steady state conditions (i.e., no change in the water surface profile).

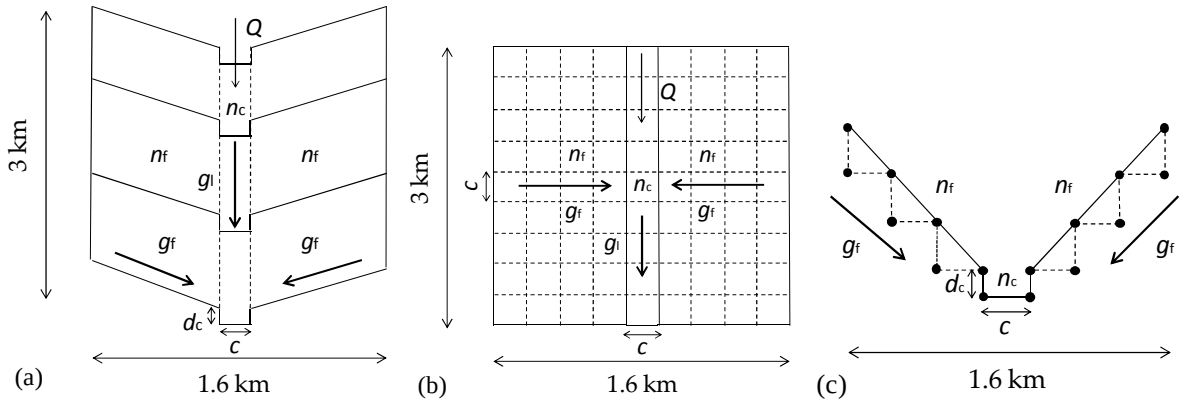

Figure 26: Layout of benchmark tests and associated input variables: (a) perspective view, (b) plan view, and (c) cross sectional view, where solid lines represent the continuous geometry, implemented within HEC-RAS, while dashed lines represent the raster-based geometry, implemented within LISFLOOD-FP and FLO-2d (d_c represents the channel depth; for rest of symbols please refer to Table 15). Source: Dimitriadis et al. (2016b).

Table 15: Variables used within sensitivity analysis and associated range of feasible values; all variables are uniformly distributed, except for the model resolution determined by the channel width, which takes three discrete values with equal probability (25, 50 or 100 m).

variable	symbol and units	min	max
upstream flow	Q (m^3/s)	100	5000
longitudinal gradient	g_l (%)	0.1	5
lateral gradient	g_f (%)	0.1	5
roughness coefficients (channel)	n_c	0.01	0.1
roughness coefficients (floodplain)	n_f	0.05	0.3
model resolution (= channel width)	c (m)	25, 50, 100	

Input data and model setup

The channel and floodplain geometry are chosen in such a way to be similar in all models. We consider the mixed section shown above, which is a typical approximation of a river and its floodplains. Its geometry is defined by the channel width c and the lateral gradient g_f. The channel width is equal to the size of the model resolution and is allowed to take three values, i.e., 25, 50 and

100 m. For simplicity, the depth of the channel d_c is determined by the intersection of the right and left floodplain section, thus it is set equal to $d_c = cg_f/2$. The longitudinal gradient g_l is constant along the channel and floodplain. The channel length is L_c = 3 km, in order to approximate uniform flow conditions downstream, while the floodplain width is L_f = 1.6 km, in order to ensure that it is never fully flooded, for the given geometry and the examined flow conditions. The representation of the actual layout differs according to the model structure. In HEC-RAS, we consider a discrete number of cross-sections at same distances, which are set equal to the channel width c, each one preserving the actual geometry. Therefore, the number of cross-sections is by definition $L_c/c + 1$. On the other hand, in LISFLOOD-FP and FLO-2d, the geometry is approximated by a grid of $(L_c/c) \times (L_f/c)$ cells, since the models are raster-based.

The inflow Q is applied to the upstream section, in HEC-RAS, or cell, in the other two models. In order to assess the performance of the three models against multiple flow conditions, we investigate a large range of inflow values, employing the steady flow scheme as well as the unsteady one. In the second case, we assign a synthetic hydrograph of 48 h duration, in which discharge slowly increases from zero to the desired value, within first 24 h, and then remains constant until reaching steady state conditions. We remark that FLO-2d is only examined for non-steady conditions, assuming both the full structure as well as the simplified structure in which the channel flow propagation is omitted. Next, these two configurations will be marked as "with channel" and "no channel", respectively. Finally, different Manning's roughness coefficients are set for the channel and floodplain, symbolized n_c and n_f, respectively.

Setup of Monte Carlo simulations

Sensitivity analysis in based on a Monte-Carlo approach, by generating 1500 random values for each of the six variables (resulting to 1500 parameter sets for each model configuration). For continuous variables, we generate independent random values from a uniform distribution in the range given in Table 15, while for the channel width, which also determines the model resolution, we generate three equally-distributed discrete values (25, 50 and 100 m). The number of simulations is chosen to ensure a satisfactory accuracy in statistical estimations.

For each simulation and each model configuration we record the water depths at the upstream and downstream section (or cell), symbolized w_u and w_d, respectively. We also record the flood volume, V_f, over the entire model domain. For each of the three output variables we employ typical statistical analysis, focused on the quantification of their uncertainty. In particular, we calculate the main statistical characteristics (mean, variance, skewness and kurtosis) and we extract their q-q and box-plots. Moreover, we calculate their cross-correlation coefficients with all inputs variables.

Monte Carlo simulation results

We chose to perform 1500 simulations for each one of the six model configurations, to balance the computational cost with an adequate quantitative analysis with an equivalent of more than three values per input variable (i.e., $1500^{1/6} \approx 3.4$). In Figure 27, we show the moving average of the coefficient of variation C_v (i.e., the ratio of standard deviation over mean) for the uniform depths upstream, w_u, and downstream, w_d, of the channel's cell/section, as well as for the flood volume V_f.

For comparison, we also show the depths estimated from the Manning's equation, assuming a triangular cross section, i.e., $w_o = \sqrt{2}\left(Q g_f n_f / \sqrt{g_l}\right)^{3/8}$. We observe that all variables have nearly reached a constant value, which strengthens the fact that the chosen number of simulations is adequate. We also underline that the HEC-RAS w_d and LISFLOOD-FP w_u lines for the steady-state scheme coincide to the HEC-RAS and LISFLOOD-FP unsteady ones, respectively. Additionally, we remark that the HEC-RAS steady and unsteady V_f lines coincide to each other.

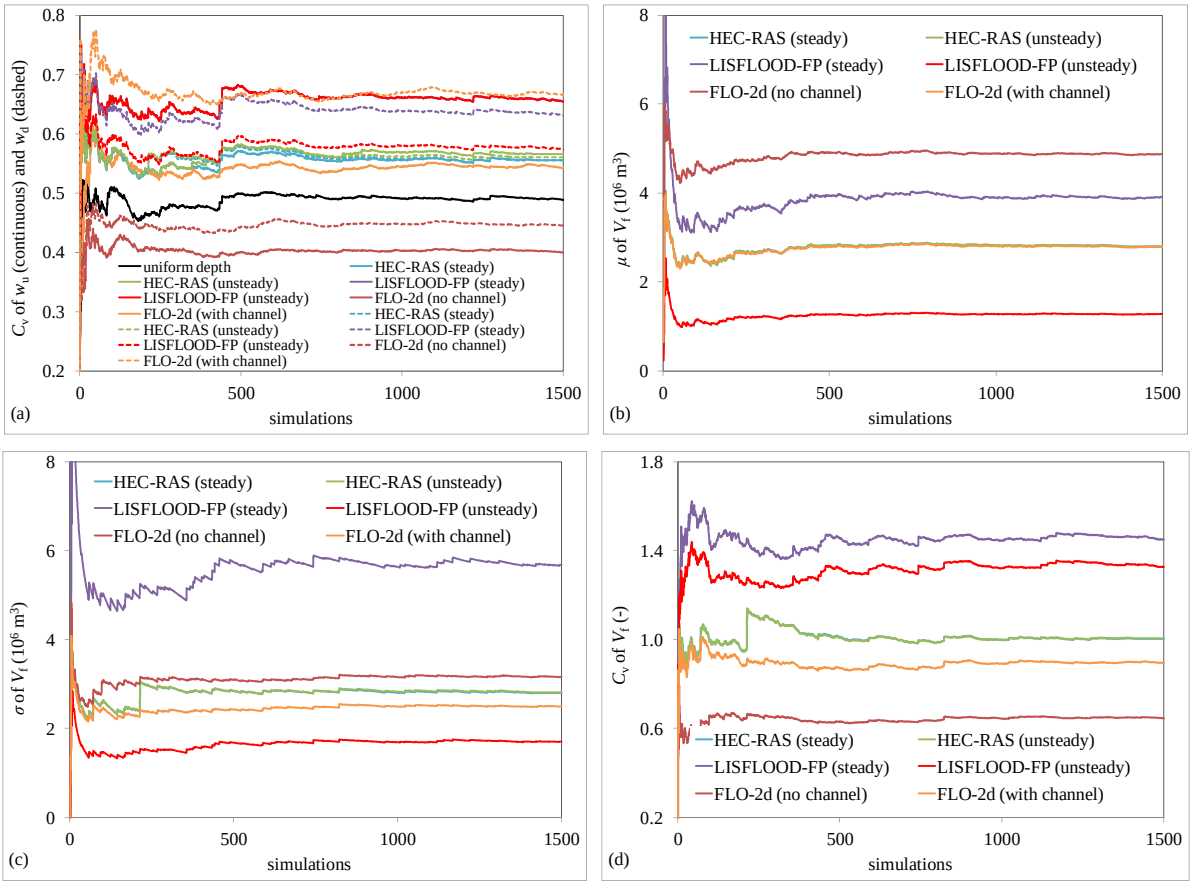

Figure 27: Moving average of (a) coefficient of variation, C_v, for all model configurations of the water depth of the channels' upstream and downstream cell/section, and (b) mean, μ, (c) standard deviation, σ, and (d) C_v for the flood volume. Source: Dimitriadis et al. (2016b).

In Table 16 we show the statistical characteristics (moment coefficients and cross correlations) of the examined output variables, estimated from the full samples (i.e., 1500 values per model configuration). The cross correlations between the input and output variables show that all output variables are an increase function of the inflow discharge and channel and floodplain roughness coefficients (same between the total flood volume and the lateral gradient) as well as a decrease function of the longitudinal gradient and model resolution (same between the upstream and downstream depths and the lateral gradient). Particularly, the largest correlations correspond to the inflow discharge, followed by the channel and floodplain slopes and roughness coefficients and with the model resolution having the smallest correlations.

Table 16: Central moments' variation (denoted C_v), skewness (denoted C_s) and excess kurtosis (denoted C_k) coefficients (using the unbiased classical estimators) for each model applied as well as cross correlation coefficients between the input and output variables. Source: Dimitriadis et al. (2016b).

variable	model	C_v	C_s	C_k	Q	g_l	g_f	n_c	n_f	c
w_o	uniform depth	0.5	0.9	1.4	0.5	-0.3	0.5	0.4	~ 0	~ 0
w_u	HEC-RAS (steady)	0.5	1.1	2.2	0.6	-0.4	0.3	0.4	0.1	-0.2
	HEC-RAS (unsteady)	0.6	1.4	4.0	0.5	-0.3	0.3	0.3	0.1	-0.2
	LISFLOOD-FP (steady)	0.7	1.8	5.1	0.6	-0.4	~ 0	0.4	~ 0	-0.3
	LISFLOOD-FP (unsteady)	0.7	1.8	5.1	0.6	-0.4	~ 0	0.4	~ 0	-0.3
	FLO-2d (no channel)	0.4	0.4	~ 0	0.7	-0.2	0.3	~ 0	0.5	-0.1
	FLO-2d (with channel)	0.5	0.3	~ 0	0.8	-0.3	0.1	0.1	0.2	-0.4
w_d	HEC-RAS (steady)	0.5	1.1	2.4	0.6	-0.4	0.3	0.3	0.1	-0.2
	HEC-RAS (unsteady)	0.5	1.2	2.4	0.6	-0.4	0.3	0.4	0.1	-0.2
	LISFLOOD-FP (steady)	0.7	1.8	5.0	0.6	-0.4	0.1	0.4	~ 0	-0.3
	LISFLOOD-FP (unsteady)	0.6	1.4	3.2	0.6	-0.3	0.1	0.4	~ 0	-0.3
	FLO-2d (no channel)	0.4	0.4	-0.5	0.7	-0.2	0.3	0.1	0.5	-0.2
	FLO-2d (with channel)	0.7	0.6	-0.1	0.6	-0.3	~ 0	0.4	0.2	-0.4
V_f	HEC-RAS (steady)	1.0	2.5	9.7	0.5	-0.4	-0.3	0.3	0.2	-0.1
	HEC-RAS (unsteady)	1.2	6.4	89.6	0.4	-0.3	-0.3	0.2	0.1	-0.1
	LISFLOOD-FP (steady)	1.7	4.4	30.7	0.4	-0.4	-0.3	0.3	~ 0	-0.3
	LISFLOOD-FP (unsteady)	1.5	4.3	29.0	0.4	-0.4	-0.2	0.3	~ 0	-0.2
	FLO-2d (no channel)	0.7	1.5	3.6	0.7	-0.4	-0.2	~ 0	0.4	0.1
	FLO-2d (with channel)	0.9	1.9	5.8	0.6	-0.4	-0.4	0.3	0.2	-0.1

In Figure 28 we show the q-q and box-plots for each output variable and each model configuration. All variables are characterized by positive skewness, with the larger one corresponding to the total flood volume. Additionally, the latter variable exhibits heavy positive tails, as also indicated by the kurtosis values shown in Table 16. In particular, the more complicated the model structure is the less heavy is the positive tail of the empirical distribution. These outcomes are of major importance in hydrological design and therefore, the application of average values to crucial model inputs (e.g., discharge, roughness coefficients etc.) may lead to over-designing, while, in contrast, the application of the most probable values may result to severe underestimations. This can be even deteriorated when the above variables exhibit heavy tails, since the mean would further deviate from the mode value. Also, a heavy-tailed variable encloses higher uncertainty, since its prediction intervals are wider, thus there is a higher probability for extreme values to occur.

In Figure 28 we also observe that the prediction intervals of LISFLOOD-FP (unsteady conditions) are 1.5 times wider than the other models for the upstream depth, whereas for the downstream

depth, HEC-RAS and LISFLOOD-FP (steady-state) intervals are approximately double more wide (compared to the upstream depth). Also, FLO-2d exhibits similar intervals, compared to the uniform depth ones, for the upstream depth and two times narrower for the downstream one. Regarding flood volumes, HEC-RAS (unsteady) and LISFLOOD-FP (steady-state) exhibit wider intervals while the rest are close to each other. It is noted that wider intervals enclose larger variability and therefore, uncertainty. The aforementioned differences are due to the different schemes, initial and hydraulic conditions made by each model and highlight the large uncertainty that one should encounter in flood modelling.

Furthermore, in Figure 28 we observe that at the left and right tail of the flood volume distribution all models deviate from normality, with HEC-RAS exhibiting the largest deviation, followed by LISFLOOD-FP and FLO-2d. This can be explained by the fact that HEC-RAS is by construction 1d, while the other two models are quasi-2d. Therefore, they can better approximate the lateral flow attenuation along the floodplain, especially in mild topographic gradients, and thus, they would require less discharge to capture a target flooded area. Also, FLO-2d uses the dynamic wave and so, it can better approximate the floodplain attenuation in comparison to the diffusive wave of the LISFLOOD-FP, which omits the local and convective acceleration terms. However, the use of extra terms significantly increases the computational burden. In average, HEC-RAS (steady) requires approximately 1 s per simulation, HEC-RAS (unsteady) requires 5 s, LISFLOOD-FP (steady and unsteady) requires roughly 10 s for all cell sizes and FLO-2d requires up to 2 min, 15 min and 1.5 h for cell sizes 100, 50 and 25 m, respectively (all simulations are performed with an Intel Core i7-2600 @ 3.40GHz processor). Note that HEC-RAS includes 30, 60 and 120 cross sections and both LISFLOOD-FP and FLO-2d include approximately 500, 2000 and 8000 grid cells, for cell sizes 100, 50 and 25 m, respectively.

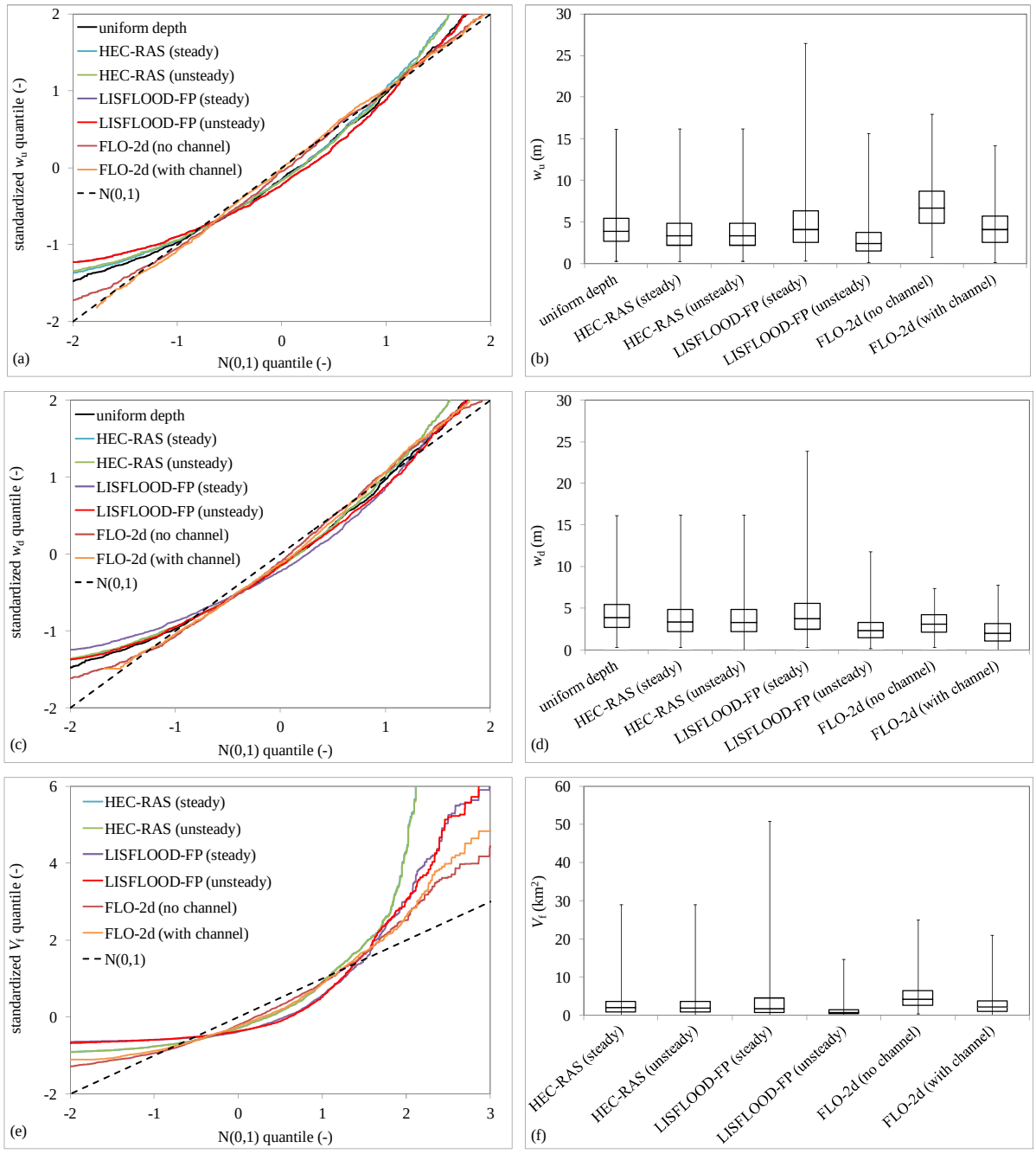

Figure 28: qq-plots and box-plots of the water depth of the channels' (a-b) upstream and (c-d) downstream cell/section as well as of the (e-f) total volume of the flooded area. Note that the water depths and flood volume are first standardized (i.e., the residual from their average value is divided with their standard deviation). Source: Dimitriadis et al. (2016b).

Model sensitivity against roughness coefficients

It is well-known that the roughness coefficient is one of the most difficult parameters to estimate in hydraulic modelling. A major issue is the different sensitivity of each model against the roughness

assigned to the channel and the floodplain. In general, we expect the flood inundation to exhibit a larger sensitivity to the channel friction rather than to the floodplain one, since the wave is carried primarily by the channel while the floodplain acts merely as additional storage (Cunge et al., 1980; Hunter et al., 2005). The above statement is in accordance with the computed correlation coefficients between the three output variables (upstream and downstream depths and flood volume) and all the input variables. Specifically, we observe that for the flood volume, HEC-RAS exhibits the largest correlation to the floodplain friction, followed by FLO-2d. On the other hand, LISFLOOD-FP exhibits minor only correlation. For the channel friction, HEC-RAS flood volume's correlation is larger than the floodplain one and similar for all models (except for the "no channel" configuration of FLO-2d, which is expected to be small). The differences in the sensitivity against the two roughness coefficients can be also illustrated through the estimation of the longitudinal and lateral momentums, where the former is expected to highly outrange the latter one. In Figure 29 we provide an example from LISFLOOD-FP, for the case of non-steady conditions.

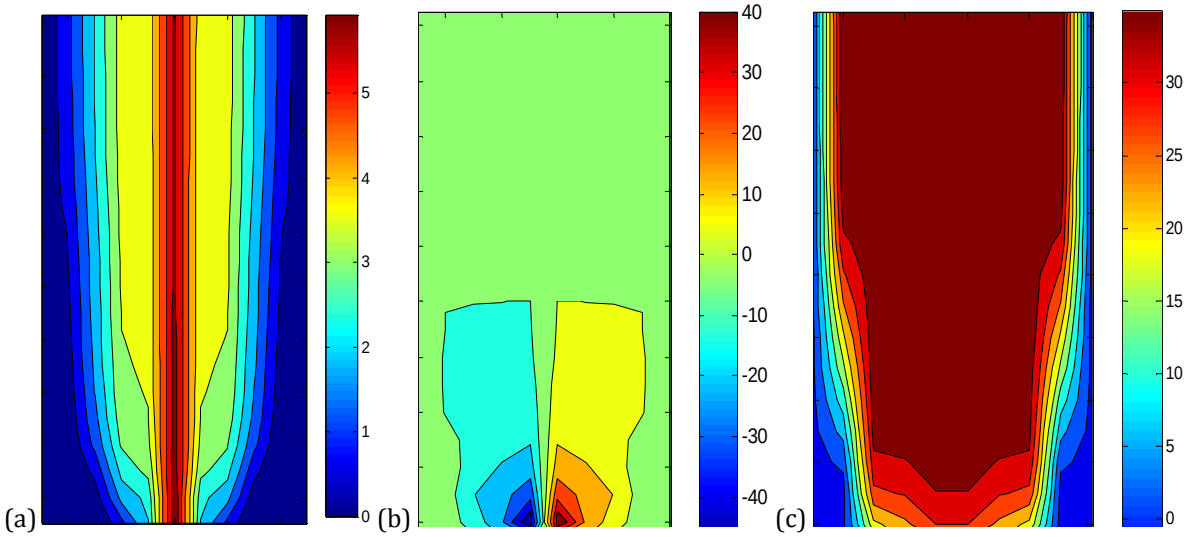

Figure 29: Contour maps of (a) water depths, (b) lateral flows, and (c) longitudinal flows produced by LISFLOOD-FP (unsteady), for $Q = 2500$ m^3/s, $n_\mathrm{f} = 0.10$, $n_\mathrm{c} = 0.07$, $g_\mathrm{l} = 2.5\%$, $g_\mathrm{f} = 2.8\%$ and $c = 50$ m. Source: Dimitriadis et al. (2016c).

Evaluation of uncertainty issues

In order to obtain a rough estimate on the uncertainty associated with the magnitude of each input variable, we calculate the variation coefficient for each model against clustered samples of each input variable. In particular, we formulate three equally sized clusters, with low, medium and high values. In Figure 30, we show the relationship between the flood volume uncertainty against each input variable and each model configuration. In general, we observe that for approximately all cases, uncertainty decreases with increasing Q, g_l and n_c, while it increases with increasing g_f, n_f and c. The most important source of uncertainty is the channel's roughness coefficient n_c, followed by the floodplain's one n_f and the inflow discharge Q. Regarding the rest of the examined inputs,

their range of variability is quite similar to each other and slightly smaller than the aforementioned three variables one. Finally, in Figure 30, we also show the variability coefficients for each model as well as the overall one, which is larger than their average value. A direct outcome of the above investigations is that since the uncertainty related to an output variable (e.g., water depth) varies significantly, so will do the hydraulic profile. This can completely alter the whole behaviour of the flow if for example the profile includes a hydraulic jump from a switch of super-critical flow to sub-critical one. It is interesting to remark that from the 1500 sets generated through the Monte Carlo method, we observe upstream sub-critical flow in 50% of simulations with HEC-RAS, 60% with LISFLOOD-FP (steady), 30% with LISFLOOD-FP (unsteady), 90% with FLO-2d (no channel) and 65% with FLO-2d (with channel). An important conclusion is that the uncertainty related to a specific input variable can sometimes outperform the uncertainty related to different models, schemes or conditions. The latter statement can be important in flood risk assessment, since it raises the question whether saving computational time can always outbalance the cost of in situ measurements (e.g., for accurate representation of geometry) in estimating a narrower variability range for an input variable or in choosing which modelling scheme or flow condition is the most appropriate for a particular case study.

As shown in the above analysis, uncertainty can be introduced in fully deterministic non-linear systems with however, a short-term persistent behaviour as shown in Table 17, where a strong lag-one cross-correlation between different models and schemes is apparent with all the larger lags corresponding to approximately zero values.

Table 17: First order correlation between various hydraulic models and schemes as estimated from the sensitivity analysis. Source: Dimitriadis et al. (2016c).

ρ_1 of downstream depth		steady (HecRac)	unsteady (HecRac)	steady (Lisflood)	unsteady (Lisflood)	no channel (Flo2d)	with channel (Flo2d)
steady	(HecRac)	1	0.998	0.510	0.519	0.455	0.484
unsteady	(HecRac)	0.998	1	0.512	0.521	0.452	0.485
steady	(Lisflood)	0.510	0.512	1	0.982	0.712	0.903
unsteady	(Lisflood)	0.519	0.521	0.982	1	0.733	0.922
no channel	(Flo2d)	0.455	0.452	0.712	0.733	1	0.799
with channel	(Flo2d)	0.484	0.485	0.903	0.922	0.799	1

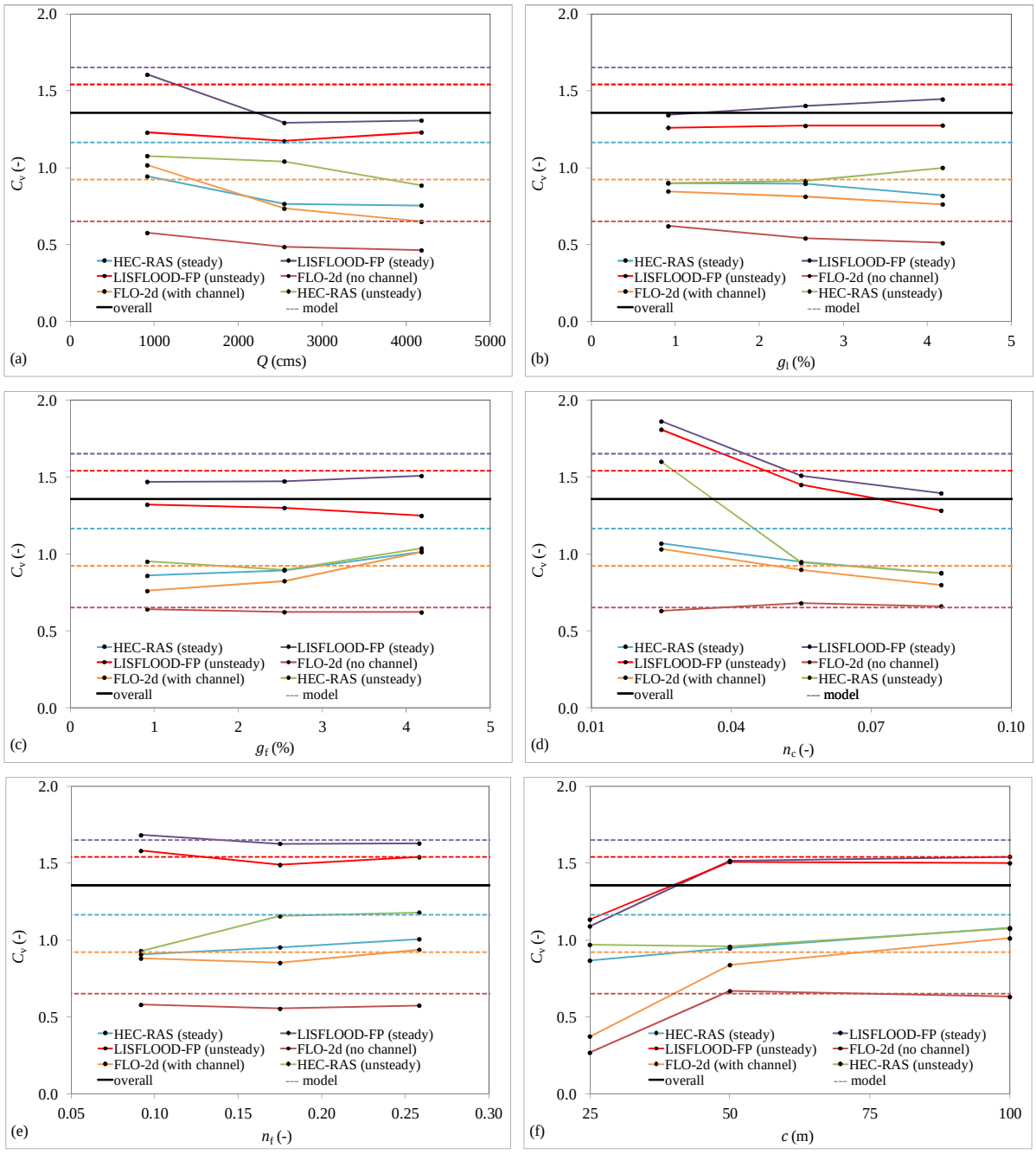

Figure 30: Variation coefficients of the flood volume vs. grouped input variables (coloured solid lines), averaged per model (coloured dashed lines) and averaged (overall) for all models (black line). Note that each variation coefficient is estimated from 500 (=1500/3), 1500 and 9000 (= 1500×6) values, respectively. Also note that the overall variation coefficients of HEC-RAS, for steady and unsteady conditions, coincide with each other. Source: Dimitriadis et al. (2016c).

4.3 HK dynamic as a measure of uncertainty

A key observation from the above analysis is that the more chaotic and complex a process is, the larger is the introduced uncertainty (unpredictability or equivalently the predictability time window) and the stronger is the HK behaviour (through the estimated Hurst parameter). Particularly, a die's trajectory is fairly predictable for time windows of approximately 0.1 s, and this time window becomes 5 min for rainfall intensity and 1 h for wind speeds. Thus, dice seems to behave like any other common physical system: predictable for short horizons, unpredictable for long horizons. The main difference of dice trajectories from other common physical systems is that they enable unpredictability very quickly. Also, the largest Hurst parameter corresponds to the process of local wind events (H = 0.95), the intermediate to the process of local rainfall events (H = 0.9) and the smallest one to the die process ($0.6 < H < 0.5$). Conversely, if averages at large time scales are considered, then the dice will become more predictable as it will soon develop a time average of 3.5; this is also strengthened by the fact that die is orientation-limited to a combination of six faces, while rainfall and wind processes have infinite possible patterns and thus, can be more unpredictable for long horizons and long time scales.

As far as the examined purely deterministic systems, it is well-known that solutions of stochastic differential equations cannot result in an HK behaviour and can be adequately approximated by Markov chain Monte-Carlo algorithms (e.g., Infante et al., 2016, and references therein). However, natural processes with HK behaviour abound in literature. For example, turbulent processes exhibit such long-term persistent behaviour (e.g., Dimitriadis et al., 2016a, and references therein), ecosystem variability (Pappas et al., 2017) as well as most geophysical processes as verified in several cases (Koutsoyiannis, 2003; O'Connell et al., 2016; Sakalauskienė, 2003), and specifically in key hydrometeorological processes such as: river discharges (Hurst, 1951; Koutsoyiannis et al., 2008); solar radiation and wind speed (Koutsoyiannis et al., 2017; Tsekouras and Koutsoyiannis, 2014); precipitation (Iliopoulou et al., 2016); paleoclimatic temperature reconstructions (Markonis and Koutsoyiannis, 2013); and temperature and dew point (Koutsoyiannis et al., 2017; Lerias et al., 2016). Interestingly, in most of the aforementioned processes the Hurst parameter is estimated at the range 0.8 to 0.85, as indicated by Hurst (1951) decades ago (Cohn and Lins, 2005).

5 Application to microscale turbulent processes

Stochastic modelling and probabilistic approaches in general, have been proven useful in the investigation of processes that resist a deterministic description, such as turbulence (e.g., Dimitriadis et al., 2016a; Frisch, ch. 3, 2006; Kraichnan, 1991, ch. 1; McDonough, ch. 1, 2004). For example, various physical interpretations of geophysical processes are based on the power spectrum and/or autocovariance behaviour (e.g., spectral density function of isotropic turbulence, see in Pope, 2000, p. 610), with both metrics belonging to the fields of Stochastics rather than classical mechanics. In this section, we apply the stochastic framework presented in the previous sections in microscale turbulent processes and we compare the results with the ones from applications in hydrometeorological processes in small scale and in larger scales.

5.1 On the definition of turbulence

Turbulence originates from the Greek word 'τύρβη' (cf. '…τὴν τύρβην ἐν ᾗ ζῶμεν':'…for the turbulence in which we live', Isokrates, 15.130) which means disorder, confusion, turmoil etc. Turbulence is considered to generate and drive most geophysical processes, e.g., wind turbulence giving birth and spatiotemporal variability in cloud rainfall (Falkovich et al., 2002), yet it is regarded as mystery within classical physics (McDonough, 2004, ch. 1). Studying turbulent phenomena is of high importance in hydrology (Mandelbrot and Wallis, 1969; Rinaldo, 2006) since the microscopic processes (related to turbulence) can help understand the macroscopic ones (related to hydrology), since they enable the recording of very long time series and with a high resolution, a rare case for hydrological processes (Koutsoyiannis, 2014). The simplest case of turbulent state (in terms of mathematical calculations) is the stationary, isotropic and homogeneous turbulence. While this is a physical phenomenon that has been recognized hundreds of years ago, still there is no universally agreed mathematical definition for the so-called 'turbulent state' (Tessarotto and Asci, 2010). Leonardo da Vinci tried to give a definition 500 years ago, based on his observations that water falling into a sink forms large eddies as well as rotational motion (Pedretti, 1977). Interestingly, Heisenberg (1985) commented on the definition of turbulent state of flow that it is just the result of infinite degrees of freedom developed in a liquid flowing without friction and thus, by contrast, laminar flow is a turbulent state of flow with reduced degrees of freedom caused by the viscous action. In 1880, Reynolds introduced one of the most important dimensionless parameters in fluid mechanics, the ratio of momentum over viscous forces which is called Reynolds number ever since. Based on this dimensionless parameter, it was observed that irrotationality in the streamlines occurred for values much greater than 1 and led to somehow confine the occurrence of turbulence to Reynolds number values greater than approximately 1000 to 2000. Richardson (1922) introduced the idea of turbulence 'energy cascade' by stating that turbulent motion, powered by the kinetic energy, is first produced at the largest scales (through eddies of size comparable to the characteristic length scale of the natural process) and then to smaller and smaller ones, until is dissipated by the viscous strain action. Taylor (1935) was the first to use stochastic tools to study this phenomenon modelling turbulence by means of random

variables rather than deterministic ones. Following this idea, Kolmogorov (1941a,b,c,d) managed to derive the famous '5/3' law (also known as K41 theory) through the Navier-Stokes equations. That law describes the energy cascade from larger to smaller turbulence scales within the inertial wavenumber sub-range, with the power spectrum no longer dependent on the eddy size and fluid viscosity. Since then, many scientists (including Von Karman, 1948; Heisenberg, 1985; Kraichnan, 1959; Batchelor, 1953 and Pope, 2000), have significantly contributed to the current power-spectrum-based models of turbulence. A general view of the stochastic approach of stationary and isotropic turbulence (in which the random variables describing turbulence have the same statistical properties in all directions) can be seen in many text books (e.g., Pope, 2000).

Following the stochastic framework in section 2, we derive in Table 18, the 1d and 3d isotropic power spectra as well as their LLD, for a Markov process, a special case of a powered-exponential process (e.g., Gneiting et al., 2012; Yaglom, 2004, ch. 10) and the gHK process. These positively-correlated mathematical processes enclose possible asymptotic behaviours in large and small scales. In particular, a positively-correlated natural process may approach zero or infinite scale, by a powered-exponential (e.g., Markov process) or a power-type (e.g., HK process) rise or decay, respectively. The 1d power spectrum and the 3d one, denoted as $s_{3\mathrm{D}}(\boldsymbol{w})$, are related by (Batchelor, 1953; Pope, 2000, pp. 226-227; Kang et al., 2003):

$$s(w) = \int_1^{\infty} \frac{x^2 - 1}{x^3} s_{3\mathrm{D}}(\|\boldsymbol{w}\| x)\, \mathrm{d}x \tag{79}$$

$$s_{3\mathrm{d}}(w) = \frac{w^3}{2} \frac{\partial\left(\frac{1}{w}\frac{\partial\left(s(w)\right)}{\partial w}\right)}{\partial w} \tag{80}$$

where $\boldsymbol{w}$ is the isotropic 3d frequency vector (wavenumber), with $\|\boldsymbol{w}\| = w \geq 0$.

As mentioned above, the most common used model for stationary and isotropic turbulence consists of the work of many scientists. Combining them into one equation, the power spectrum of isotropic and stationary turbulence can be expressed as (Pope, 2000; Kang et al., 2003):

$$s_{3\mathrm{D}}(w) = f_{\mathrm{E}}(w, c_{\mathrm{E}}, p) f_{\mathrm{I}}(w, c_{\mathrm{I}}) f_{\mathrm{D}}(w, c_{\mathrm{D}}) \tag{81}$$

where, from the work of Von Karman (1948), for the energy containing eddies (large scales):

$$f_{\mathrm{E}}(w, c_{\mathrm{E}}, p) = \left(\frac{w}{\sqrt{w^2 + c_{\mathrm{E}}}}\right)^{\frac{5}{3}+p} \tag{82}$$

combined with the work of Kolmogorov (1941a,b,c,d) for the inertial range (intermediate scales):

$$f_{\mathrm{I}}(w, c_{\mathrm{I}}) = c_{\mathrm{I}} w^{-\frac{5}{3}} \tag{83}$$

and from the work of Kraichnan (1959) for the dissipation range (small scales):

$$f_{\mathrm{D}}(w, c_{\mathrm{D}}) = \mathrm{e}^{-wc_{\mathrm{D}}} \tag{84}$$

where $c_{\mathrm{E}}, p, c_{\mathrm{I}}, c_{\mathrm{D}}$ are constants.

Table 18: 1d and 3d power spectrum for Markov, powered-exponential and gHK processes as well as their LLD, where λ is the parameter related to the true variance of the process, q the scale parameter and b is related to the power-type behaviour of the process (source: Dimitriadis et al., 2016a).

Markov		powered-exponential (special case)		gHK	
$c(\tau) = \lambda \mathrm{e}^{-\lvert\tau\rvert/q}$	(T18-1)	$c(\tau) = \lambda \mathrm{e}^{-(\tau/q)^2}$	(T18-2)	$c(\tau) = \lambda \frac{(1-b)(2-b)}{(1+\lvert\tau\rvert/q)^b}$ with $b \in (0,2)$	(T18-3)
$s(w) = \frac{4\lambda q}{1+4\pi^2 q^2 w^2}$ with $\lim_{w\to 0} s^{\#} = 0$ and $\lim_{w\to\infty} s^{\#} = -2$	(T18-4)	$s(w) = \frac{\lambda q\sqrt{\pi}}{2}\mathrm{e}^{-(qw\pi)^2}$ with $s^{\#}(w) = -2(qw\pi)^2$, $\lim_{w\to 0} s^{\#} = 0$ and $\lim_{w\to\infty} s^{\#} = -\infty$	(T18-5)	$\lim_{w\to 0} s \sim w^{b-1}$, with $\lim_{w\to 0} s^{\#} = b-1$	(T18-6)
				$\lim_{w\to\infty} s \sim w^{-2}$, with $\lim_{w\to\infty} s^{\#} = -2$	(T18-7)
$s_{3\mathrm{d}}(w) = \frac{4\lambda q(2\pi q w)^4}{(1+4\pi^2 q^2 w^2)^3}$ with $\lim_{w\to 0} {s_{3\mathrm{d}}}^{\#} = 4$ and $\lim_{w\to\infty} {s_{3\mathrm{d}}}^{\#} = -2$	(T18-8)	$s_{3\mathrm{D}}(w) \sim q^5 w^4 \mathrm{e}^{-(qw\pi)^2}$ with $s^{\#}(w) = 4 - 2(qw\pi)^2$ $\lim_{w\to 0} {s_{3\mathrm{d}}}^{\#} = 4$ and $\lim_{w\to\infty} {s_{3\mathrm{d}}}^{\#} = -\infty$	(T18-9)	$\lim_{w\to 0} s_{3\mathrm{d}} \sim w^{b-1}$, with $\lim_{w\to \mathbf{0}} {s_{3\mathrm{d}}}^{\#} = b-1$	(T18-10)
				$\lim_{w\to\infty} s_{3\mathrm{d}} \sim w^{-2}$, with $\lim_{w\to\infty} {s_{3\mathrm{d}}}^{\#} = -2$	(T18-11)

5.1.1 Stochastic properties of large-scale range

For the 3d and 1d (derived from the 3d one) power spectra at the energy containing range, we have that:

$$\lim_{w\to \mathbf{0}} s_{3\mathrm{d}} = \lim_{w\to 0} s \sim w^p \tag{85}$$

where Von Karman (1948) suggests $p = 4$ (or else known as 'Batchelor turbulence', cf., Davidson, 2000), while other works result in different values, e.g., Saffman (1967) suggests $p = 2$.

There are many arguments about the proper value of the p parameter and its relation to the Loitsyansky integral which controls the rate of decay of kinetic energy (Davidson, 2000). The main debate is whether points at a large distance in stationary, isotropic and homogeneous turbulent flow are statistically independent or show a correlation that decays either exponentially (e.g., Von Karman model for wind gust, cf., Wright and Cooper, 2008, ch. 16.7.1; Faisst and Eckhardt, 2004; Avila et al., 2010; Kuik et al., 2010; models for pipe flow) or with a power-type law.

Towards the stochastic properties of the aforementioned equation, we can see that the case $p = 2$ does not correspond neither to exponential (Markov or powered-exponential) nor to power-type (i.e., HK) decay of autocovariance. Hence, this model cannot be applied to asymptotic zero frequencies (or infinite scales). Interestingly, the case $p = 4$ can be interpreted by a Markov or a special case of the powered-exponential decay of autocovariance. However, this case also excludes the HK behaviour, i.e., long-range dependence, where p now equals $b - 1$ and is bounded to [-1, 1].

Although the aforementioned models do not include a possible power-law decay of autocovariance (HK behaviour), several works show strong indication that turbulence natural processes can exhibit such behaviour rather than Markov. Such works are reported by e.g., Nordin et al. (1972) for laboratory turbulent flume and turbulent river velocities, Helland and Van Atta (1978) for grid turbulence velocities, Goldstein et al. (1995) for magneto-hydrodynamic turbulent solar wind, Chamorro and Porté-Agel (2009) for wind turbulent wakes and grid-turbulence, Dimitriadis and Papanicolou (2012) and Charakopoulos et al. (2014a,b) for turbulent buoyant jets, Dimitriadis et al. (2016a) for grid turbulence.

We believe that the reason a possible HK behaviour is not detected in geophysical processes (which are often characterized by lack of measurements), is that mathematical smoothing techniques are applied, e.g., windowing or else Welch approaches, regression analysis, wavelet techniques (see other examples in (Stoica and Moses, 2005, ch. 2.6). Particularly, application of windowing techniques to any stochastic tool can be misleading since they eliminate a portion (depending on the type and length of the window applied) of the variance of the time series (which often is incorrectly attributed to 'noise', e.g., Koutsoyiannis, 2010). This elimination can lead to process misrepresentation in case of significant effects of discretization, small and/or finite record length and bias (examples of applications to the power spectrum can be seen in Lombardo et al., 2013; and Dimitriadis and Koutsoyiannis, 2015a). An example of smoothing out the HK behaviour by applying the Welch approach with a Bartlett window and no segment-overlapping to an observed time series is shown in Figure 31. Even though the smoothing technique decreases the variance of the power spectrum, it also causes low frequency loss of information. This loss of information may cause a process misinterpretation, as illustrated in Figure 31, where the autocorrelation function (derived from the 3D power spectrum model) exhibits a Markov-like decay, while the empirical one (derived from the windowed empirical power spectrum partitioned into 10^3 segments) exhibits HK behaviour. Also, this smoothing technique should be used in caution in strong-correlated processes, since an increase in the number of partitioned segments will cause an increase in their cross-correlation. Finally, processes with HK behaviour have usually large bias and in case this is not included in the model, the empirical rapid decay of autocovariance in large scales (or equivalently lags) may be erroneously interpreted as short-range dependence.

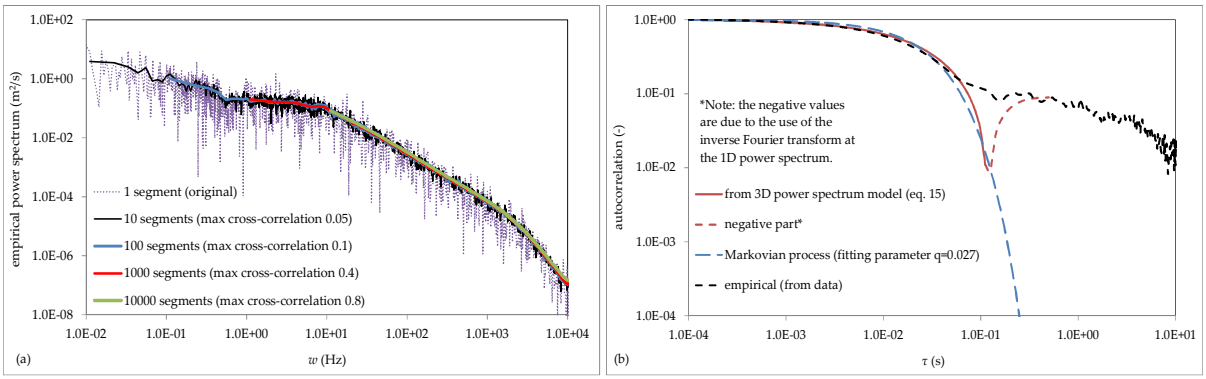

Figure 31: (a) Example of loss of low frequency information caused by the application of the windowing technique, in a time-series provided by the Johns Hopkins University as well as the maximum cross correlations between the partitioned segments; (b) 1D autocorrelation function derived from the 3D power spectrum model (with parameters based on the fitting of the windowed 1D power spectrum with 1000 segments: $c_{\mathrm{E}} = 2.5\ \mathrm{m}^{-2}$, $p = 4$, $c_{\mathrm{I}} = 13.0\ \mathrm{m}^3/\mathrm{s}^2$, $c_{\mathrm{D}} = 2 \times 10^{-4}$ m); a Markov autocorrelation function, i.e., $e^{-(\tau/q)}$, for reasons of comparison; and the corresponding (to the windowed 1D power spectrum with 1000 segments) empirical autocorrelation function. Source: Dimitriadis et al. (2016a).

To incorporate possible HK behaviour in the model, we may assume an autocovariance power-type decay at large scales, where the 3d and 1d power spectra at asymptotically zero frequency are of the form w^{b-1}, with b bounded to $(0, 2)$, for positively correlated processes (i.e., $0.5 < H < 1$), negatively-correlated processes (i.e., $0 < H < 0.5$) and for a process with a white-noise-like decay in large scales (i.e., $H = 0.5$).

5.1.2 Stochastic properties of intermediate range

One may observe that the power spectrum asymptotic LLD for various processes often coincident to each other. For example, for both a Markov and a gHK process with b =1, the power spectrum LLD is 0 for the low frequency tail and -2 for the high frequency one. This may be confusing and result in misinterpretation of the natural process. A solution to this may be to incorporate additional stochastic tools in the analysis. For the aforementioned example, if the autocovariance function asymptotic properties (local and global ones) are analyzed, one can decide upon powered-exponential lag decay (as in the Markov process) and a power-type one (as in the gHK process). Similarly, when a power-type behaviour appears in the intermediate frequencies of a power spectrum (as in the case of a -5/3 LLD), it may be misleading to interpret it as a power-law function (and thus, a power-type autocovariance decay), since this can be derived from different kind of processes with no power-type expressions for the intermediate scale-range. An illustrative example is shown in Figure 32, where the -5/3 LLD in the intermediate frequencies of the power spectrum results from a simple combination of a Markov and a gHK process, both of which have a purely stochastic interpretation and they do not include power-type expressions in the intermediate frequency-range.

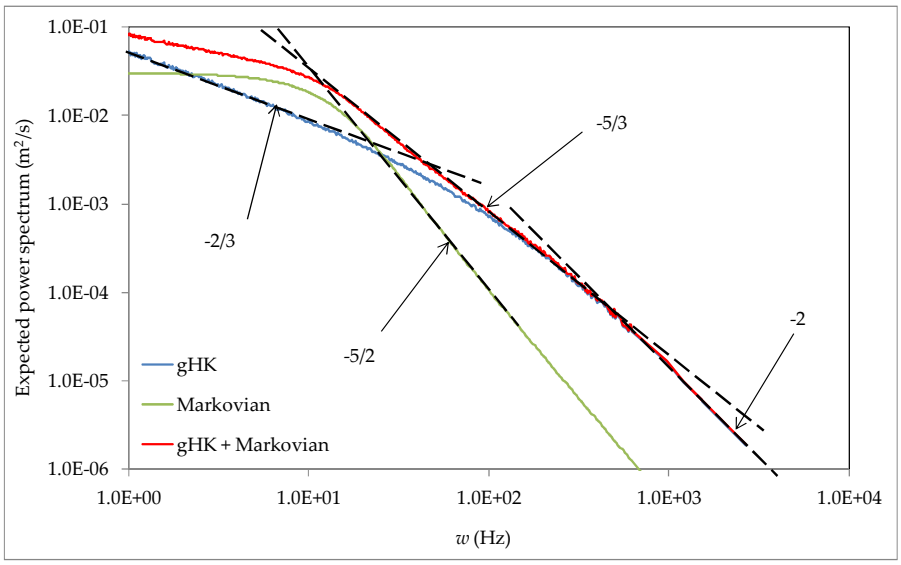

Figure 32: Expected power spectrum resulted from a combination of a Markov and a gHK process (source: Dimitriadis et al., 2016a).

Note also, that the Kolmogorov (1941a,b,c,d) power-type power spectrum refers only to intermediate frequencies and should not be also applied arbitrarily to low frequencies, since the corresponding asymptotic large-scale behaviour of the autocovariance, i.e., $c(\tau) \sim \tau^{5/3\text{-}1}$, is equivalent to an erroneous $H = 4/3 > 1$.

5.1.3 Stochastic properties of small-scale range

Similarly, for the 3d and 1d power spectra at the dissipation range, we have that (Figure 33):

$$\lim_{w\to\infty} s_{3\mathrm{d}}(w) = \lim_{w\to\infty} s(w) \sim \mathrm{e}^{-w} \tag{86}$$

This corresponds to an autocovariance function of the form:

$$c(\tau) \sim \frac{1}{\tau^2 + 1} \tag{87}$$

which corresponds to the Wackernagel (1995) process (also mentioned as an autocovariance-based Cauchy-class process resembling the Cauchy probability function). A generalized expression of this process can be found in Gneiting (2000), which we will refer to it as the Gneiting process (Table 8):

$$c(\tau) = \frac{\lambda}{(1 + (\tau/q)^{2M})^{\frac{1-H}{M}}} \tag{88}$$

Note that for M = ½ we have the gHK process and that if this process is expressed based on the climacogram rather than the autocovariance; it corresponds to the HHK process.

For small lags (and for $q = \lambda = M = 1$) this process behaves like (e.g., Gneiting and Schlather, 2004):

$$\lim_{\tau \to 0} c(\tau) \sim 1 - \tau^2 \sim e^{-\tau^2} \tag{89}$$

which corresponds to the special case of a powered-exponential process. Note that if this process is expanded directly to large scales it corresponds to an erroneous process with $H = 0$.

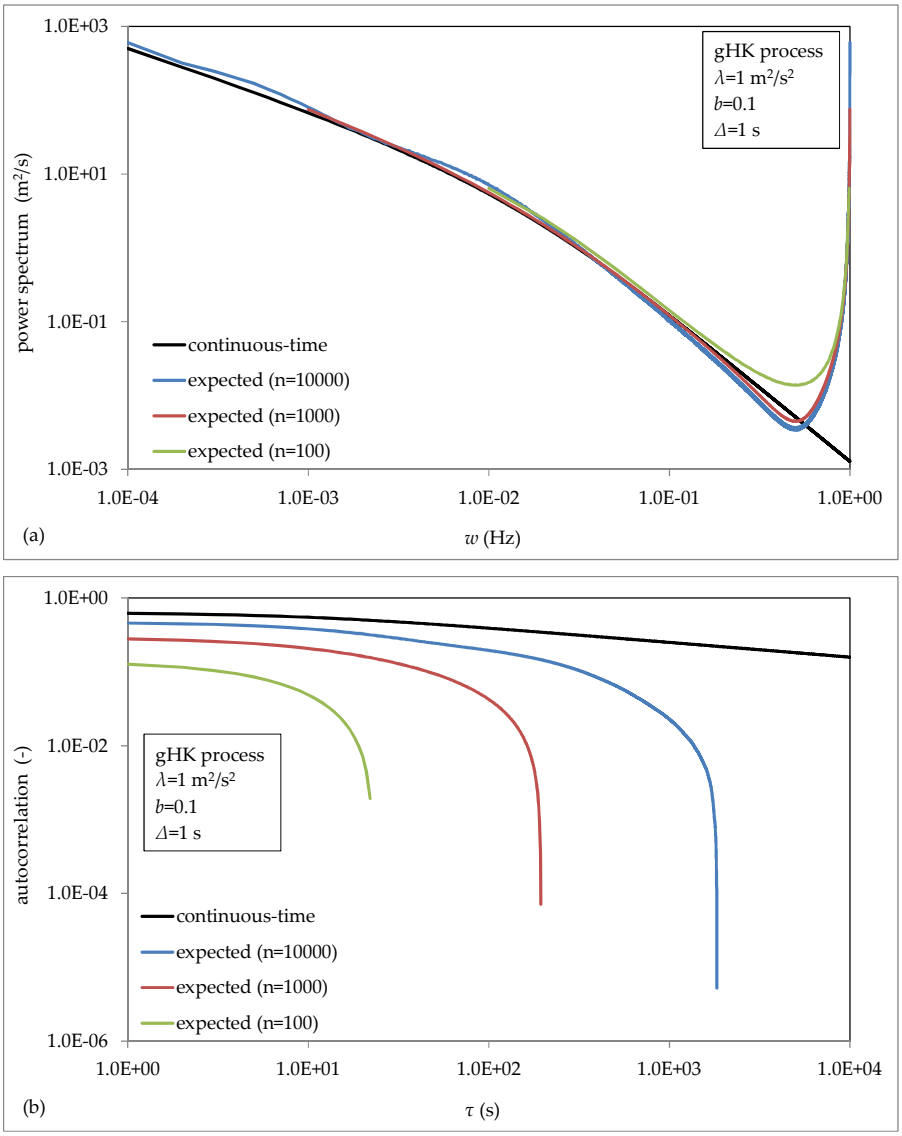

Figure 33: (a) Power spectra and (b) corresponding autocovariances, in continuous time as well as their expected values, with varying number of records n for a gHK process (source: Dimitriadis et al., 2016a).

Other models for the dissipation range are of the form of a powered-exponential power spectrum process that may result from a powered-exponential autocovariance function. However, there is evidence that these models cannot interpret the frequently observed spike in the high frequency power spectrum (e.g., Cerutti and Meneveau, 2000; Kang et al., 2003). This is usually ignored and

attributed to instrumental noise. In Figure 34, we show that this spike may appear in HK processes and is due to discretization and bias errors, in case the shape parameter q/Δ takes large values.

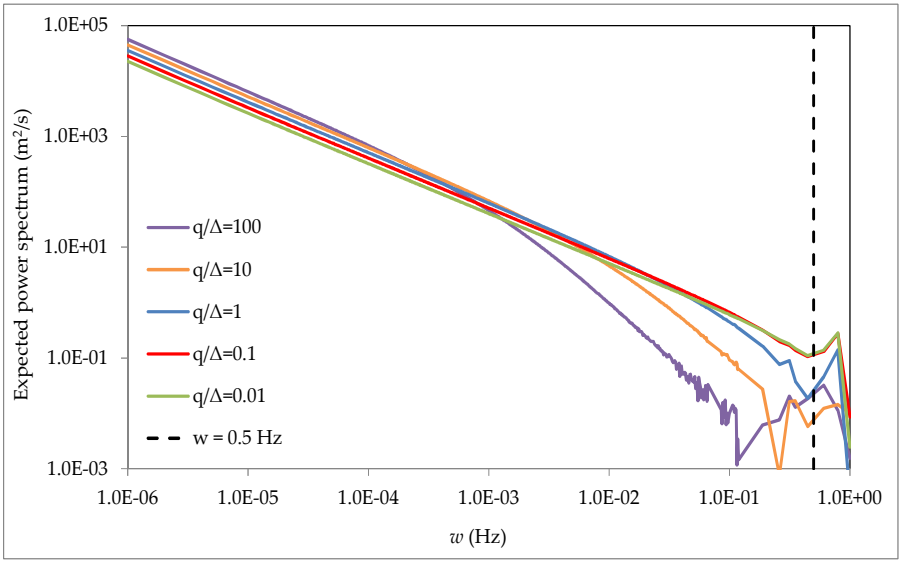

Figure 34: Expected power spectra of a gHK process, with varying q/Δ (where Δ the sampling time interval). Source: Dimitriadis et al., 2016a.

5.2 Proposed model

Here, we focus on the local and global stochastic properties of the most common three-dimensional power-spectrum-based models of stationary and isotropic turbulence in time domain and we detect some model weaknesses despite their widespread use. In the previous section, we present several limitations concerning the stochastic properties of proposed turbulent models from literature. Specifically, we see that they only include exponential decay in the energy containing area and thus, completely excluding possible HK behaviour. They also, describe the dissipation area decay with only a specific case of a powered-exponential process and thus, leaving out all other possible types of decay. Moreover, they interpret a possible power-type-like intermediate area (of the power spectrum) with power-type behaviour (and particularly, only that of the K41 theory) which can also result from intermediate non power-type processes. Furthermore, these models adequately represent only the power spectrum while failing to describe other tools like the climacogram and autocovariance. Moreover, these models are constructed based on multiplications between processes, an action with no mathematical or physical justification and which may cause numerical difficulties in stochastic generation. Since turbulence generates and drives most of geophysical processes, we expect geophysical processes to exhibit similar types of decay in small and large scales. Hence, a more robust, flexible and parsimonious model is required that can incorporate all the aforementioned microscale and macroscale behaviours linking turbulence to hydrology and beyond. Here, we choose the ergodic stochastic model that consists of two independent processes, these of a Markov and an HHK process (with $H > 0.5$ and $M < 0.5$), combined in such way to exhibit the desired behaviour in the intermediate scales. This model can describe a variety of combinations

between powered-exponential and HK processes, including the often observed intermediate quick drop of all the stochastic tools. This particular drop may be due to the interference of boundaries and/or the existence of multiple periodic functions, as for example in case of combinations of HK with cyclo-stationary processes (Dimitriadis and Koutsoyiannis, 2015b). Furthermore, although the proposed model results in a complicated expression for the power spectrum, it provides simpler expressions for the other tools. Additionally, the proposed model is also justified by the extremization of entropy production in logarithmic time, as shown in section 2. Finally, this model combines both fractal and HK dynamics using four parameters (Dimitriadis and Koutsoyiannis, 2017):

$$\gamma(k) = \frac{\lambda}{2(1+(k/q)^{2M})^{\frac{1-H}{M}}} + \frac{\lambda\left(k/q + \mathrm{e}^{-k/q} - 1\right)}{(k/q)^2} \tag{90}$$

For the estimation of the distribution parameters we minimize the error introduced in Dimitriadis and Koutsoyiannis (2017) and is based on the absolute value of the difference between the main body of the empirical and modelled distribution along with their left and right tails:

$$\varepsilon_f = \sum_i \left|1 - \frac{F_\mathrm{m}(x_i)}{F_\mathrm{e}(x_i)}\right| \sum_i |F_\mathrm{e}(x_i) - F_\mathrm{m}(x_i)| \sum_i \left|1 - \frac{F_\mathrm{e}(x_i)}{F_\mathrm{m}(x_i)}\right| \tag{91}$$

where f_m and f_e are the model and empirical distribution functions, respectively.

For the estimation of the parameters of the dependence structure we minimize a similarly defined error (Dimitriadis and Koutsoyiannis, 2017):

$$\varepsilon_\gamma = \sum_\kappa \left|1 - \frac{\gamma_\mathrm{m}(\kappa)}{\gamma_\mathrm{e}(\kappa)}\right| \sum_i |\gamma_\mathrm{e}(\kappa) - \gamma_\mathrm{m}(\kappa)| \sum_i \left|1 - \frac{\gamma_\mathrm{e}(\kappa)}{\gamma_\mathrm{m}(\kappa)}\right| \tag{92}$$

where γ_m is the model climacogram and γ_e is the empirical climacogram.

5.3 Applications to laboratory microscale turbulent processes

In this section, we use laboratory measurements of grid-turbulence velocities recorded within a wind-tunnel and of temperature differences recorded within a turbulent thermal jet.

5.3.1 Laboratory measurements of grid-turbulence velocities

As previously mentioned, high order moments cannot be reliably estimated from typically short time series of geophysical processes. However, in laboratory experiments with high sampling rates, very large time series of observations can be formed, which allow direct estimation of high order moments from data. Here, we use a grid-turbulence massive database provided by the Johns Hopkins University (www.me.jhu.edu/meneveau/datasets/datamap.html). This dataset consists of 40 time series, each with $n = 36\times10^6$ data points of longitudinal wind velocity along the flow direction, all measured by X-wire probes placed downstream of the grid and with a sampling time interval of 25 μs. (Kang et al., 2003). Due to the laboratory nature of the experiment we may apply

the Taylor's hypothesis of frozen turbulence (Taylor, 1938) and shift from the spatial to the temporal domain (Castro et al., 2011). We then use a standardization scheme illustrated in Figure 35 to homogenize all series (Dimitriadis et al., 2016a) and, by setting the empirical mean to zero, we calculate the standardized empirical variance as $\mathrm{E}\left[\underline{\hat{\gamma}}(D)\right] \approx 1$. By the standardization we are able to form a sample of 40 × 36 ×10^6 = 1.44 ×10^9 values for the estimation of the marginal characteristics of the process and an ensemble of 40 series, each with 36 ×10^6 values for the estimation of the dependence structure characteristics.

It can be observed that the time series are not precisely Gaussian but rather nearly-Gaussian as shown in Figure 35. This is also verified by the skewness and kurtosis estimates of 0.2 and 3.1, respectively. If those values were estimated from a small sample, for example n = 100, then the probability density function of the process would be regarded Gaussian and the divergence from normality would be attributed to statistical error, since for n = 100 the uncertainty measured through the standard deviation of the skewness and kurtosis, is as high as 30% and 50%, respectively (Figure 12). However, for $n \approx 1.5\times10^9$ the uncertainty of the mean will drop below 1% for H = 0.8 and therefore, it is expected that the uncertainty of skewness and kurtosis will be low too. Moreover, there are some theoretical arguments justifying the divergence of fully developed turbulent processes from normality (Wilczek et al., 2011).

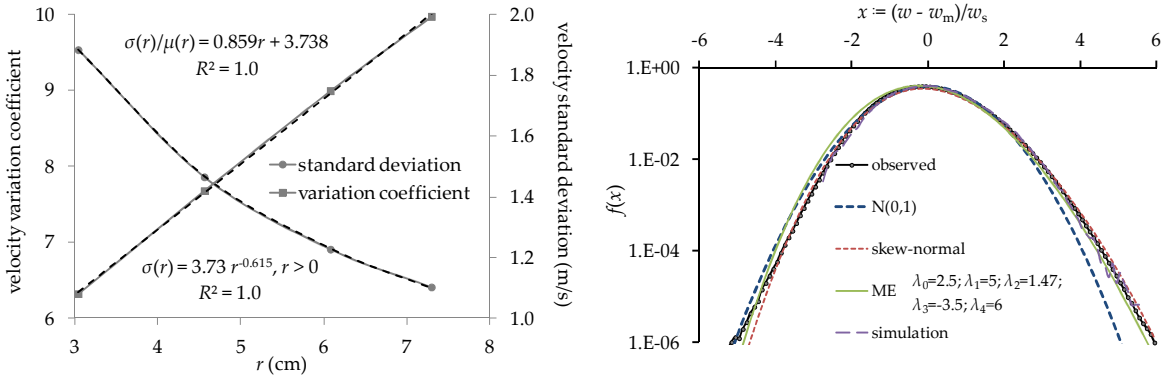

Figure 35: [left] Standardization scheme for grid-turbulence data, where μ and σ are the mean and standard deviation, r is the distance from the grid, with the first 16 time series corresponding to transverse points abstaining r = 20S from the source, the second 4 to r = 30S, the third 4 to 40S and the last 16 to 48S, with S = 0.152 m the size of the grid; [right] empirical probability density function of the overall standardized time series (observed) along with that from a single synthetic time series produced by the SMA scheme to preserve the first four moments (simulation); for comparison the theoretical distributions N(0,1), skew normal and ME constrained on the four moments (corresponding weights for the ME distribution: 15%, 51%, 21% and 13%). Source: Dimitriadis and Koutsoyiannis (2017).

For the estimation of the climacogram we apply the suggested methodology of fitting the expected model to the mean climacogram calculated from the 36 time series of identical length. However, to improve the fitting of the model, we include in the analysis the additional climacogram-based

metrics such as the CBF and CBS (see section 2.5). The climacogram is more representative of the large and intermediate scales, the CBF of the small and intermediate scales and the CBS of small and large scales and thus, by combining all three of them we can achieve a better fitting of the model (Dimitriadis et al., 2016a).

The model parameters are estimated as: $\lambda \approx 1$, $M \approx 1/3$, $H \approx 5/6$ and $q \approx 14$ ms. Here a large number of parameters could be justified due the large data size but the above model is quite parsimonious. Also, since the applied extended HHK model is theoretically justified through the maximization of entropy (as shown in section 2.4) each parameter has a physically-based interpretation. Moreover, we observe from Figure 36 that this model is also in agreement with the work on the turbulent power spectrum by Von Karman (1948) for the large scale range, by K41 model for the intermediate range and by Kraichnan (1959) for the dissipation range (cf., Pope, 2000, pp. 232-233), while here we also simulate the HK behaviour that clearly appears in the very small frequencies (very large scales) of the power spectrum and in the other stochastic tools. Additionally, certain aspects exhibited in the power spectrum such as the bottleneck effect (Kang et al., 2003) and the spike at large frequencies which is often ignored and attributed to instrumental noise (Cerutti and Meneveau, 2000) are also well represented. Finally, the preservation of kurtosis of the velocity increments (see below) enables to even simulate the effect that the intermittent behaviour of the process has on the marginal probability distribution, first discovered in turbulence by Batchelor and Townsend (1949).

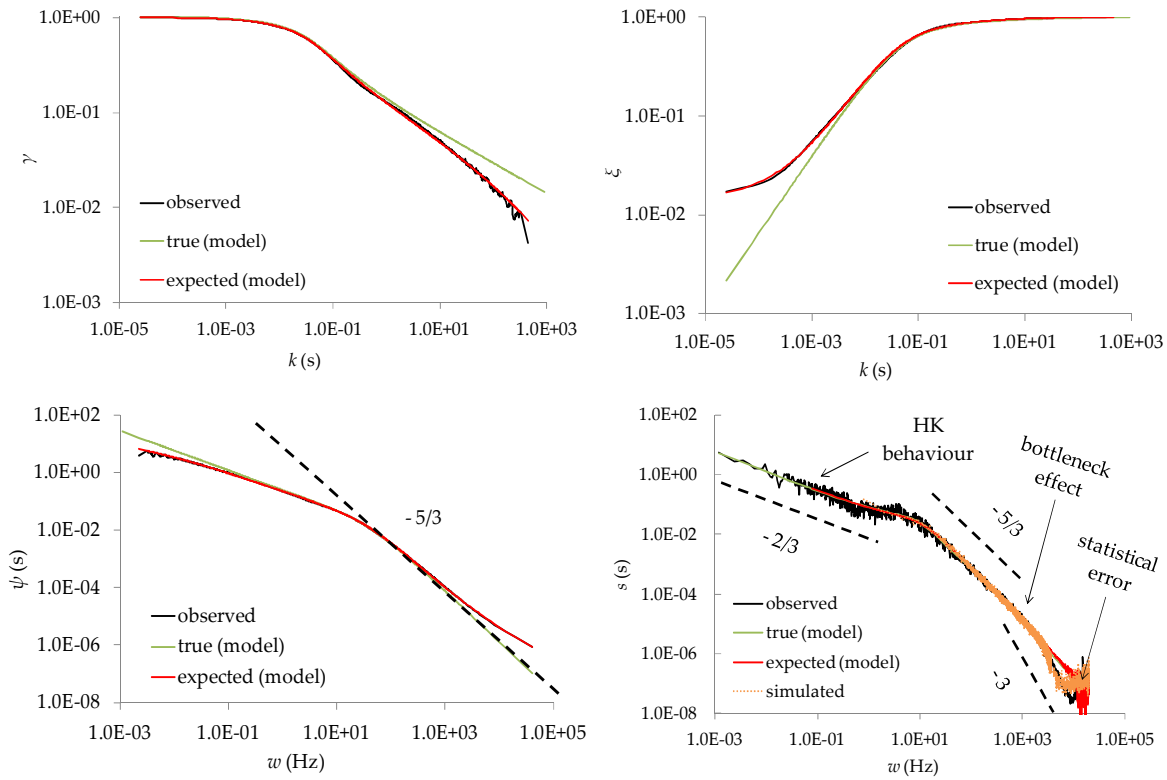

Figure 36: The empirical, true and expected values of the climacogram [upper left], CBF [upper right], CBS [lower left] and power spectrum [lower right] along with some important logarithmic slopes. Source: Dimitriadis and Koutsoyiannis (2017).

It is interesting to further investigate the latter issue through the behaviour of a generalized structure function $V_p(h) := \mathrm{E}\left[\left|\underline{x}_i - \underline{x}_{i+h}\right|^p\right]$ and in particular the power-law behaviour for the intermediate range of lags, i.e., $V_p(h) \approx h^{\zeta_p}$. Such behaviours have been attributed to intermittency (Frisch, 2006, sect. 8.3) which initiated the need for exploring models different from the K41 such as the multifractal ones (Frisch, 2006, sect. 8.5 to 8.9). As shown in Figure 37, the increase of $V_3(h)$ and the drop of kurtosis of the velocity increments for a wide range of lag (h), as well as the increase of the exponent ζ_p for a wide range of the p exponent, are impressively well preserved by the proposed model. This is achieved with no particular effort or provision (e.g., without using extra assumptions, parameters or models) but merely by simultaneously simulating the first four moments (with focus on the coefficient of kurtosis) and the stochastic structure of the process.

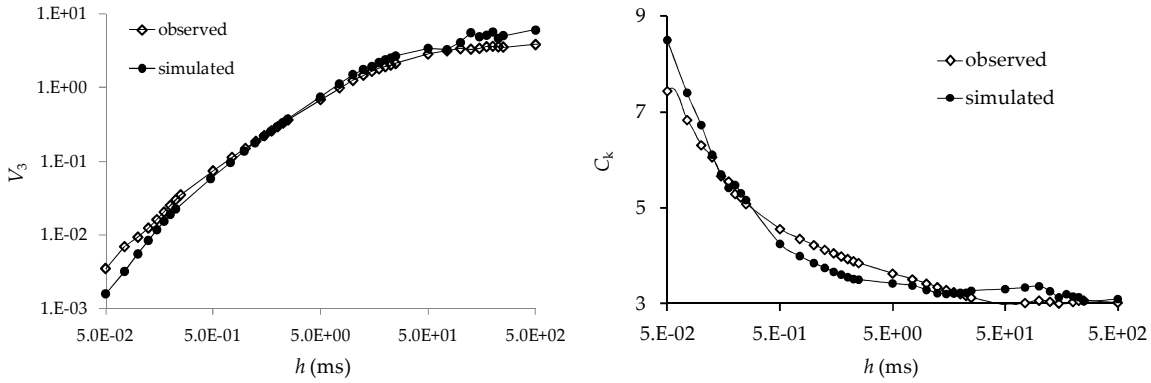

Figure 37: Empirical and simulated 3rd order structure function [left] and kurtosis coefficient [right] of the velocity increments vs. lag. Source: Dimitriadis and Koutsoyiannis (2017).

To further highlight this finding, we illustrate in Figure 38 that the HHK model alone cannot simulate the observed behaviour of the high order structure function but rather approaches the structure function as simulated by the K41 self-similarity model and reproduced by Frisch (2006, Fig. 8.8). Similar results are obtained in case a Markov dependence structure is adopted but by simultaneously preserving the empirical non-Gaussian marginal distribution. Interestingly, if both the proposed dependence structure and marginal distribution are combined, then the observed behaviour of the high order structure function is preserved and as a consequence the intermittent behaviour of turbulence. For comparison, we plot the She-Leveque model (She and Leveque, 1994) that behaves also exceptionally well and originates from the alternative assumption of independent identically distributed log-Poisson multiplicative factors (Frisch, 2006, sect. 8.6.4, 8.6.5).

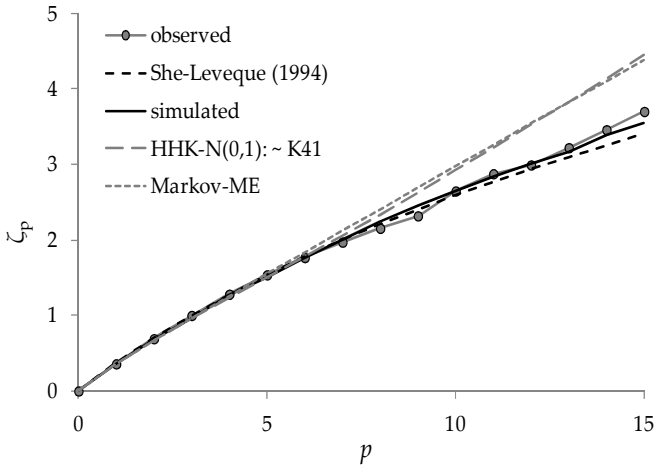

Figure 38: Empirical and simulated structure function for various orders of the velocity increments vs. lag. Source: Dimitriadis and Koutsoyiannis (2017).

5.3.2 Laboratory measurements of turbulent thermal jet temperatures

For the analysis of turbulence micro-scale through the measurement of concentration, a laser-induced fluorescence (LIF) technique is used, implemented at the laboratory of Hydromechanics

and Environmental Engineering at the University of Thessaly and at the laboratory of Applied Hydraulics at the NTUA. The measurements are based on the Laser-Induced Fluorescence (LIF) technique (Papanicolaou and List, 1987; 1988). Particularly, the buoyant jet is dyed with a rhodamine 6G (R6G) dye with low concentration that does not affect the buoyancy forces. The jet flow field is illuminated with a thin (order of 1 mm) plane sheet of laser light. A DPSS 1 W laser beam at 532 nm (green) is converted to a thin laser light sheet via a rotating prism mirror at 20 kHz. The rhodamine dye excited by the 532 nm wavelength emits (yellow) light at 556 nm, the intensity of which is proportional to the rhodamine concentration if it does not exceed 50 ppm, as indicated by Ferrier et al. (1993). Thus, laser based tomography of the buoyant jet flow-field can be obtained across any desired plane. Then, the experiment is videotaped using a high resolution video-camera pointing normal to the light sheet at 30 frames per second (fps). The experimental setup is illustrated in Figure 39.

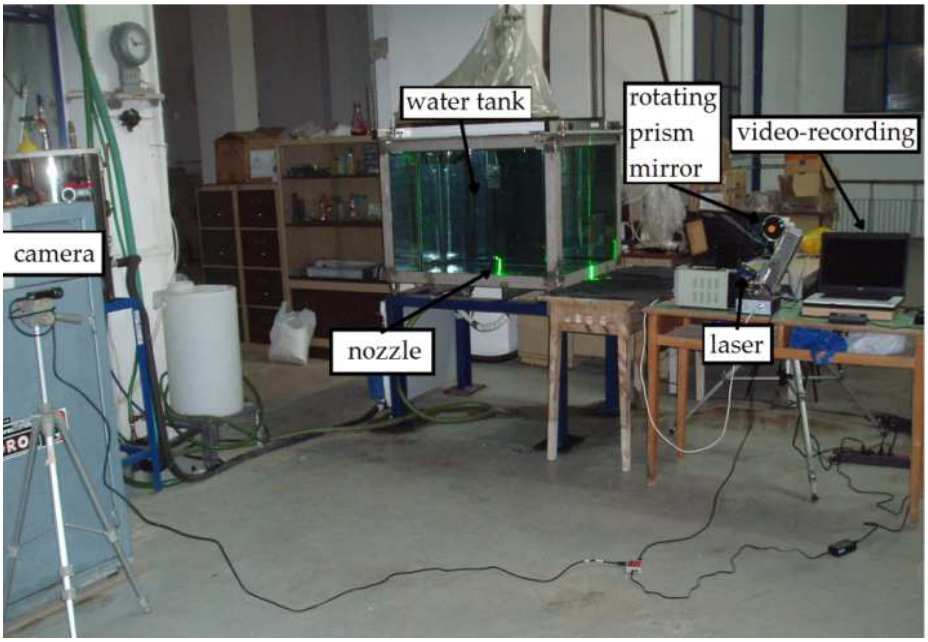

Figure 39: Photograph of the experimental set-up on turbulent buoyant jets at the laboratory of Hydraulics at NTUA.

For larger than 50ppm concentrations of R6G, the attenuation factor can no longer be assumed negligible and it should be taken into account (as shown in the equations below; Dimitriadis and Papanicolaou, 2010):

$$P(x) = P_{\mathrm{o}} \mathrm{e}^{-x\eta_P(x)} \tag{93}$$

$$I(x) = I_{\mathrm{o}} \mathrm{e}^{-x\eta_I(x)} \tag{94}$$

$$I(x) = \beta P(x) C(x) \tag{95}$$

$$\eta_P(x) = \eta_{P\mathrm{w}} + \varepsilon_P C(x) \tag{96}$$

$$\eta_I(x) = \eta_{I\mathrm{w}} + \varepsilon_I C(x) \tag{97}$$

where P_o and P is the, initial and at distance x(m) from the source laser power (W),

I_o and I is the initial, and at distance x(m) from the source, intensity of the radiation in units of wavelength (nm),

C is the concentration (μg/l) of the fluorescence element at distance x,

η_{Pw}, η_{Iw} and η_P, η_I are the attenuation parameters (m^{-1}) of laser power and radiation intensity resulting from clear water and from concentration C of the fluorescence upstream of the element at distance x, respectively,

ε_P and ε_I are coefficients (l/μg/m) that affect the attenuation of the laser power and radiation intension, respectively,

β (l nm/W/μg) is a coefficient indicating the measure of efficiency.

The coefficient η_{Iw} can be experimentally determined by estimating (via image processing methods) the distribution of the intensity along the laser beam in the water tank. The same method can be applied for the determination of ε_I and β by taking a threshold value of fluorescence (uniformly distributed in the tank). Afterwards, the coefficients η_{Pw} and ε_P can be also determined. The initial fluorescence light intensity Io is proportional to the R6G initial concentration Co if it does not exceed 50 ppm (or μg/L), as shown by Ferrier et al. (1993). Here, this is verified through the measurement of the intensity of several R6G concentrations samples fully mixed into the water-tank, for two camera shutter speeds (sp).of 50 and 100 Hz (see Figure 40). The curves in Ferrier et al. (1993) are adjusted to the measurements by multiplying with an arbitrary factor since the applied intensity is arbitrary. Finally, the emitted yellow light can be split to its components red and green light intensity (with the blue one being near zero), and therefore, to avoid the contribution of possible scattering from the green laser beam, one may compute the R6G concentration from the red light component intensity only.

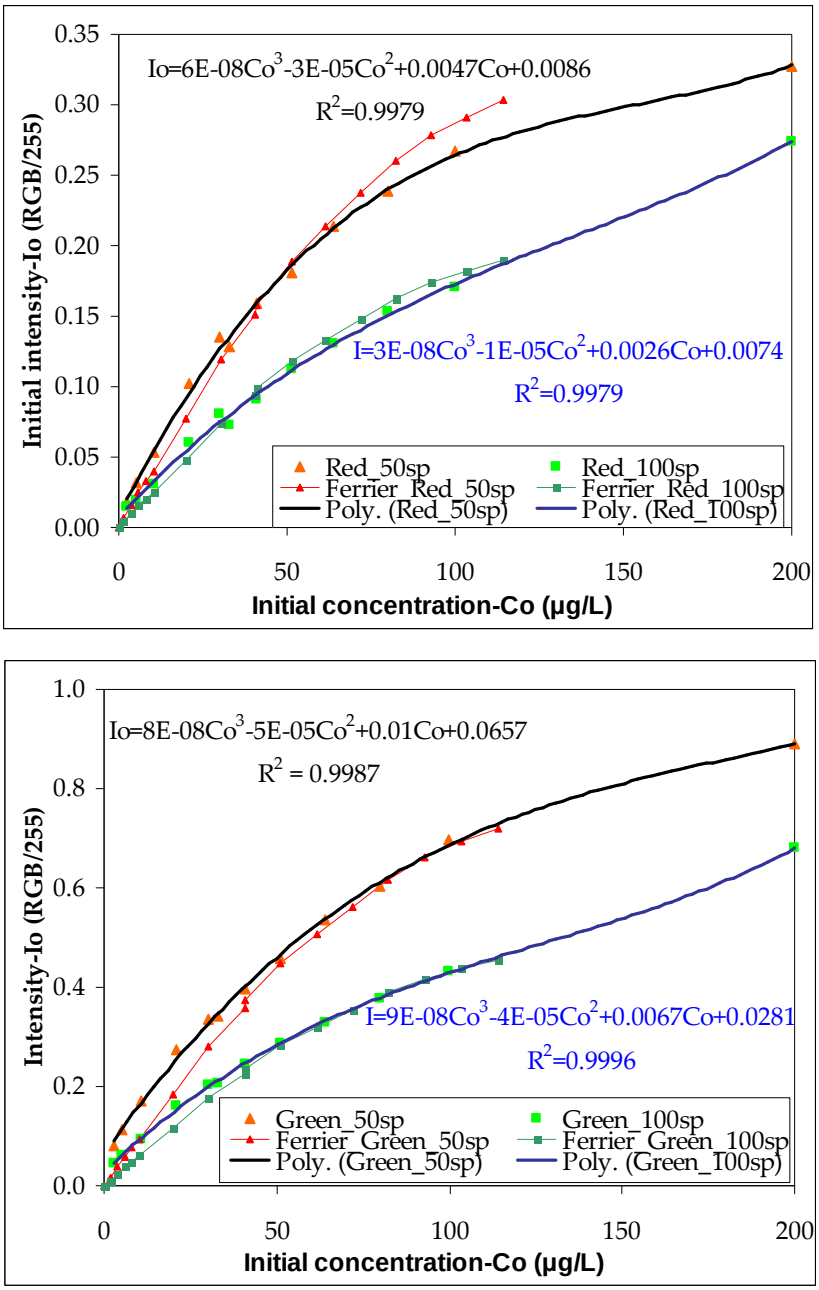

Figure 40: Initial concentration C_o vs. the initial intensity I_o for the red (top) and green (bottom) RGB intensity. Source: Dimitriadis et al. (2010).

A set of experiments is performed for buoyant jets discharging in the horizontal and vertical direction, for Richardson numbers in the range 0.01 to 0.20. Richardson number is determined from the initial jet volume, momentum and buoyancy fluxes Q, M and B, respectively, as $QB^{1/2}/M^{5/4}$ (Table 19) and is a measure of the relative strength of initial buoyancy and inertial forces applied at the jet. Note that the effect of laser attenuation due to light absorption from diluted rhodamine dye is not taken into account in the data analysis of this set of experiments. An image processing code is created in MATLAB for estimating certain turbulent characteristics based only on the ratio of

concentrations. Initially, the model zooms in the area of interest and removes the background noise (by setting a threshold intension value of R6G coloured radiation). Then, all the static objects that are not of interest (i.e., the nozzle) are removed from the video frames. Next, the model smoothes the gridline areas and rotates/enlarges the frames to adjust them to the real dimensions. Finally, the blue hue (from the RGB values) is removed as explained above. Following this initial frame elaboration, the concentration values are analyzed to examine if they are compatible with theoretical relationships resulting from dimensional arguments. The temperature difference between jet and ambient fluid ratio is assumed to be proportional to the rhodamine concentration for uniformly distributed R6G.

Table 19: Details of the experiments held at the Laboratory of Hydraulics at the NTUA on the period 1/5/09 to 1/10/10 (where C_o is the R6G initial concentration, D is the diameter of the nozzle, Q is the initial discharge of R6G, T_{amb} and T_{jet} are the ambient and jet temperature). Source: Dimitriadis and Papanicolaou (2010).

no	date	direction of flow	C_o (mg/l)	D (cm)	Q (ccs)	T_{jet} (oC)	T_{amb} (oC)	Reynolds number	Richardson number	l_M	type of flow
TBHJ01	8/2/2010	horizontal	6000	1.0	20.00	40.00	16.00	3851	0.094	9.45	Jet
TBVJ01a	9/7/2010	vertical	6000	1.0	15.26	38.50	25.10	2858	0.097	9.13	Jet
TBVJ01b	9/7/2010	vertical	6000	1.0	18.62	38.70	25.10	3499	0.080	11.04	Jet
TBVJ01c	9/7/2010	vertical	6000	1.0	21.97	38.80	25.10	4137	0.068	12.98	Jet
TBVJ02a	9/7/2010	vertical	6000	0.5	11.91	33.40	25.20	4040	0.017	26.69	Jet
TBVJ02b	9/7/2010	vertical	6000	0.5	18.62	33.40	25.20	6314	0.011	41.71	Jet
TBVJ02c	9/7/2010	vertical	6000	0.5	8.56	33.40	25.20	2903	0.023	19.18	Jet
TBVJ02d	9/7/2010	vertical	6000	0.5	5.21	33.40	25.20	1766	0.038	11.67	Jet

First, we analyze the horizontal turbulent buoyant jet (experiment TBH01) following the above analysis (Figures 41 and 42).

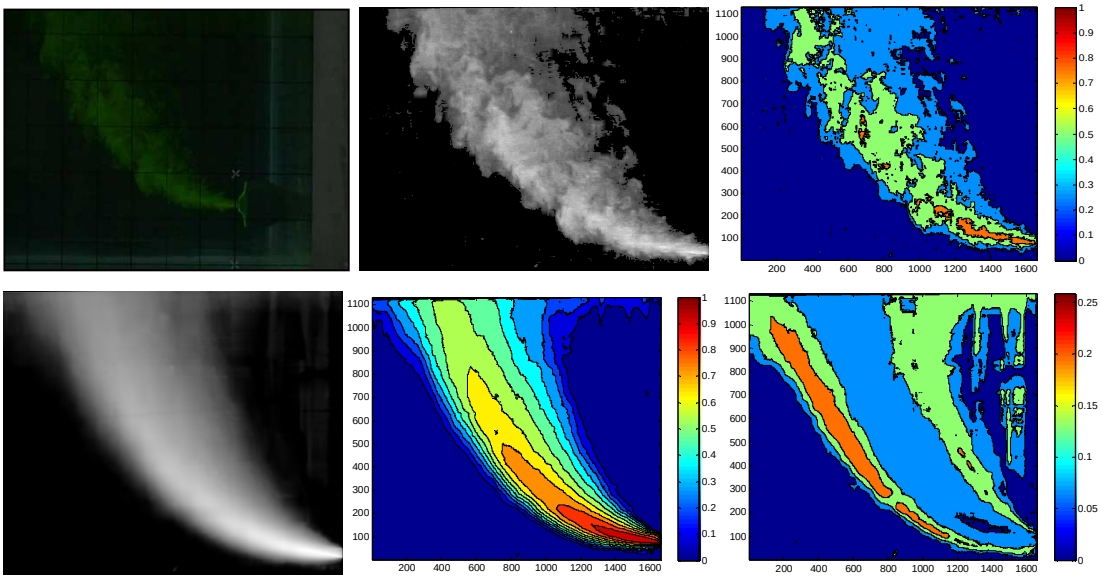

Figure 41: From left to righr and top to bottom: (a) Raw picture taken from the video-camera for the TBJH01 experiment, (b) gray-scale and (c) RGB format of the raw picture, (d) average gray-scale and (e) RGB image of the experiment, and (f) average RMS image of the experiment. Source: Dimitriadis and Papanicolaou (2010).

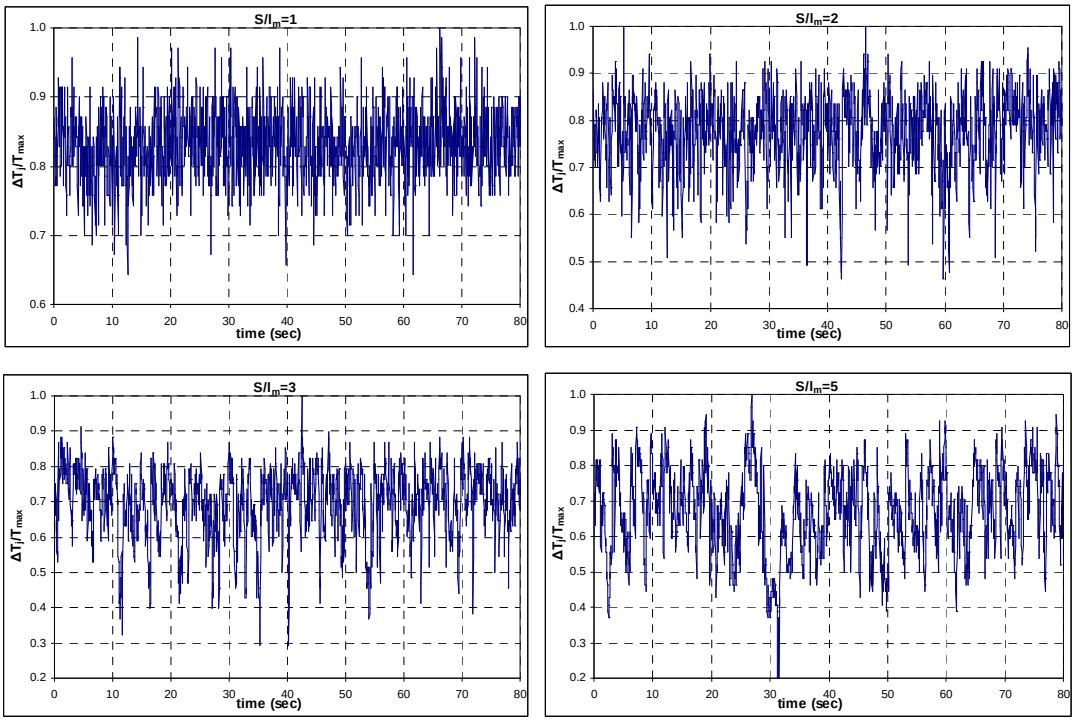

Figure 42: Time series of the excess temperature over the maximum temperature at the jet centerline for the TBHJ01 experiment (Table 19). Source: Dimitriadis and Papanicolaou (2010).

The same analysis is repeated for the vertical jets (Figure 43).

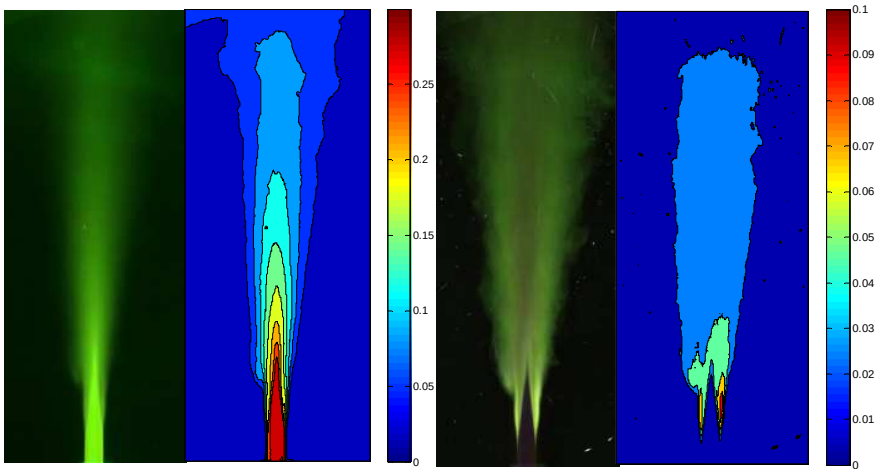

Figure 43: Dimensionless average and standard deviation of the RGB intensity (1st and 3rd pictures) and of the red RGB intensity (2nd and 4th plots), for the experiment TBVJ01a. Source: Dimitriadis and Papanicolaou (2010).

We then examine the HK behaviour of temperature as a function of the distance along the jet axis. Near the nozzle, the flow is dominated by the initial horizontal momentum and attains pure jet properties, while away from the nozzle the specific buoyancy flux dominates thus, the flow does not longer behave as a jet but as a plume. At the jet regime, the flow behaves irrationally and the fluctuations caused by turbulence are large. As a result of this, the temperature timeseries is expected to have a low Hurst coefficient close to 0.5. In contrast, in the plume regime the timeseries is expected to behave as a positively correlated process and thus, to have a larger Hurst coefficient. This state takes place for distances from the nozzle $S/l_{\mathrm{M}} > 1.5$ to 2 (Papanicolaou and List, 1987; Michas and Papanicolaou, 2009), where S is the distance from the nozzle, l_M is a characteristic length (indicating how far from the nozzle the buoyancy forces become significant).

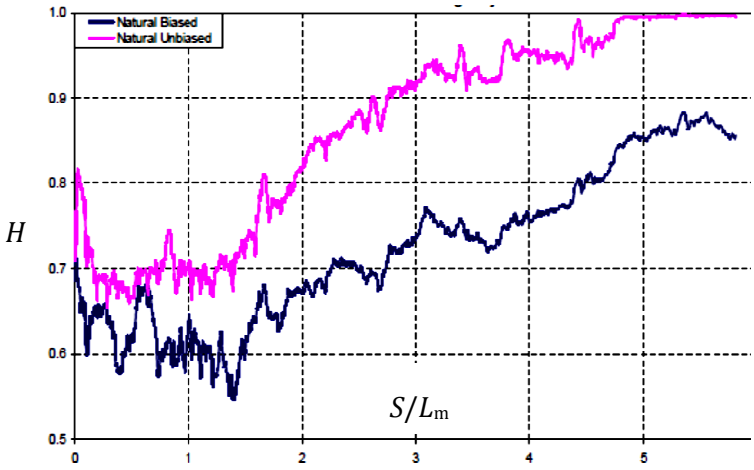

Figure 44: True (unbiased, pink line) and empirical (biased, blue line) Hurst parameter along the jet axis. Source: Dimitriadis and Papanicolaou (2010).

5.4 Stochastic similarities between the microscale of turbulent processes and the mesoscale geophysical ones

In this section, we show the stochastic analysis of a time-series of one month (Figure 45), consisted of high resolution ($\Delta \approx D$ = 0.1 s) atmospheric longitudinal wind speed (measured in m/s). This is recorded by a sonic anemometer on a meteorological tower, located at Beaumont KS and are provided by NCAR/EOL (http://data.eol.ucar.edu/).

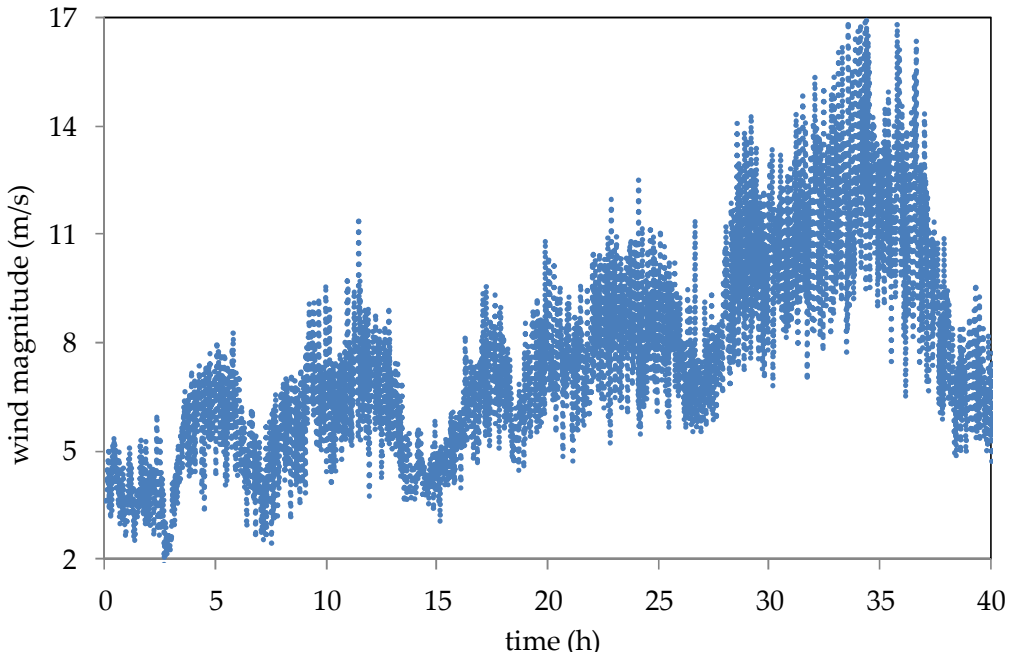

Figure 45: Part of the wind speed time-series provided by NCAR/EOL (http://data.eol.ucar.edu/).

First, we divide the time-series into three sets nearly Gaussian, each of which includes almost 1400 time-series of 10 min duration and of marginal empirical variances 0.15, 0.5 and 1.4 m^2/s^2, respectively, and we estimate the climacogram and autocovariance based metrics for each set (Figure 46).

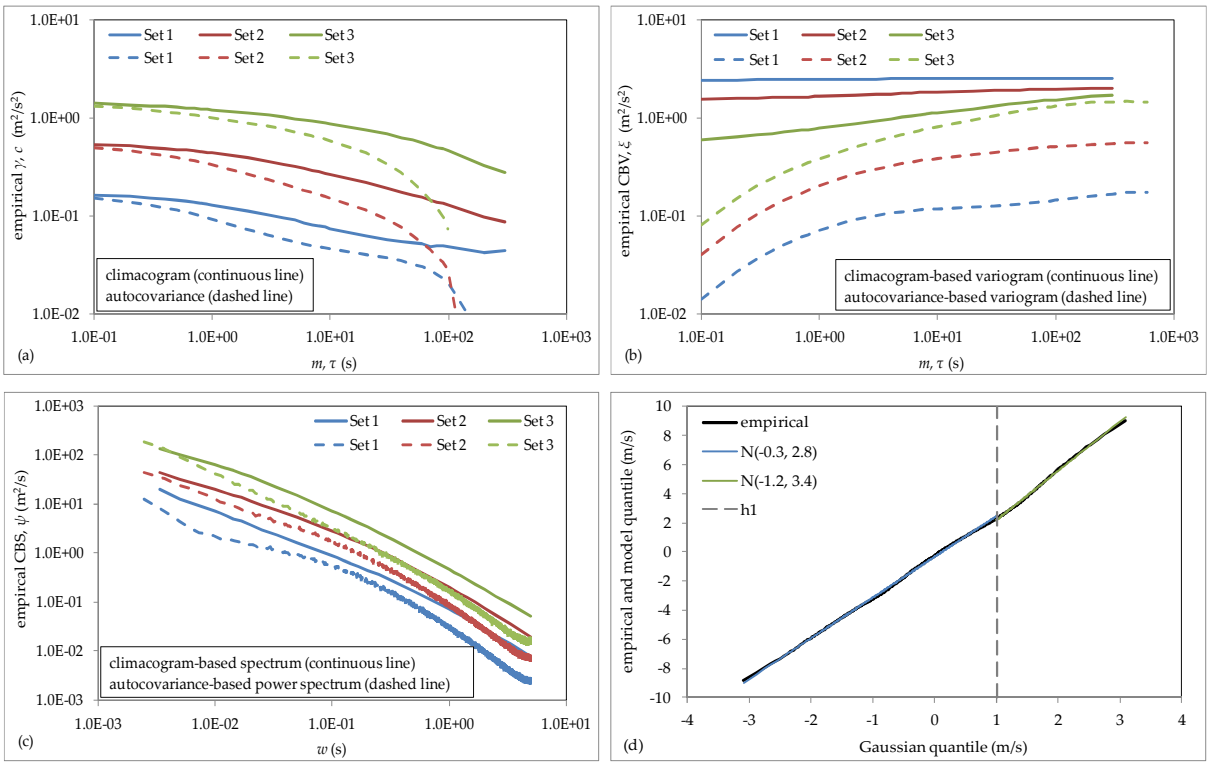

Figure 46: From top to bottom and from left to right: Averaged empirical (a) climacograms and autocovariances, (b) CBV and variograms, (c) CBS and power spectra (for the three sets) and (d) qq-plot of empirical pdf *vs* standard Gaussian pdf (for the original time-series), along with modelled distribution density function (all parameters in m/s).

Additionally, we apply a model with HK behaviour (for details see in Dimitriadis et al., 2016a) and we estimate the Hurst parameters as (Figure 47): $H = 0.99$ (first set), $H = 0.98$ (second set) and $H = 0.98$ (third set).

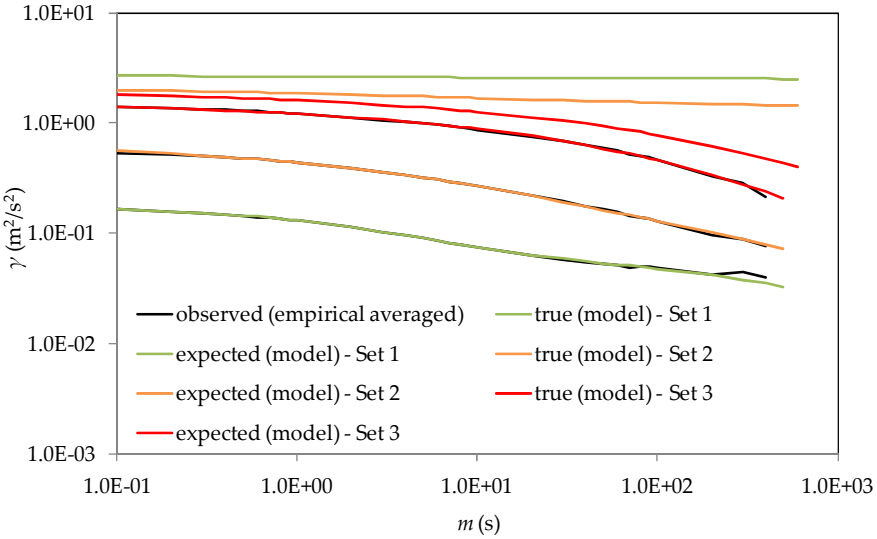

Figure 47: True, expected and empirical (averaged) climacogram values for the wind process stochastic simulation.

We also show the stochastic analysis of three time-series (Figure 48) with high resolution ($\Delta \approx D$ = 10 s) precipitation intensities (measured in mm/h). These episodes are recorded during various weather states (high and low rainfall rates) and provided by the Hydrometeorology Laboratory at the Iowa University (for more information concerning these episodes and various stochastic analyses, see Georgakakos et al. (1994) and Koutsoyiannis and Langousis (2011, ch. 1.5).

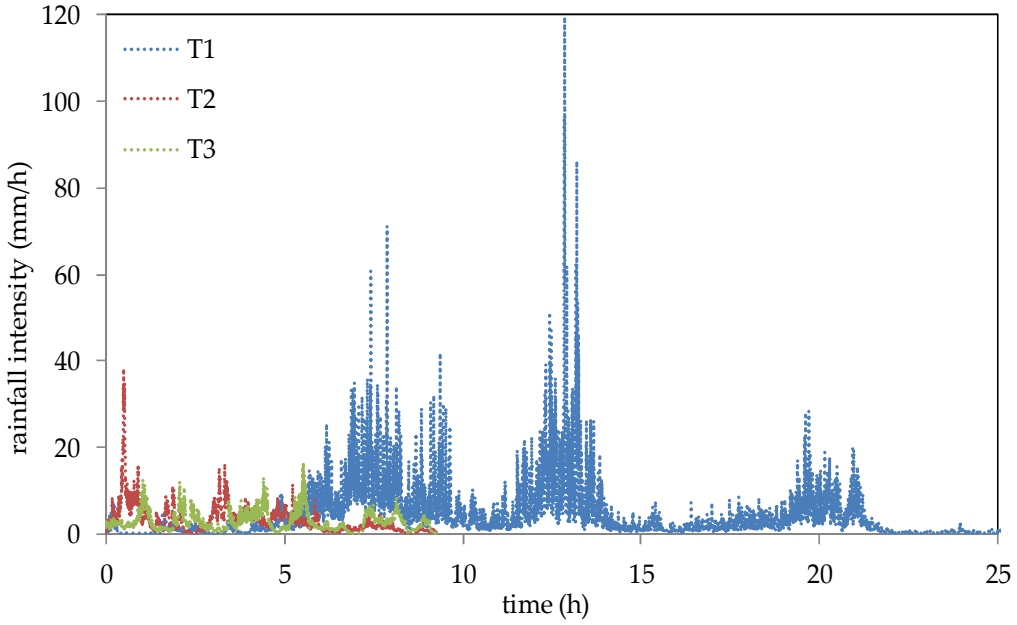

Figure 48: Three precipitation episodes provided by the Hydrometeorology Laboratory at the Iowa University.

Additionally, we estimate the climacogram and autocovariance based stochastic metrics for each time series (Figure 49). Finally, we apply a model with HK behaviour (for details see in Dimitriadis

et al., 2016a) and we estimate the Hurst parameters as (Figure 49): $H = 0.94$ (T1),$H = 0.95$ (T2) and $H = 0.93$ (T3).

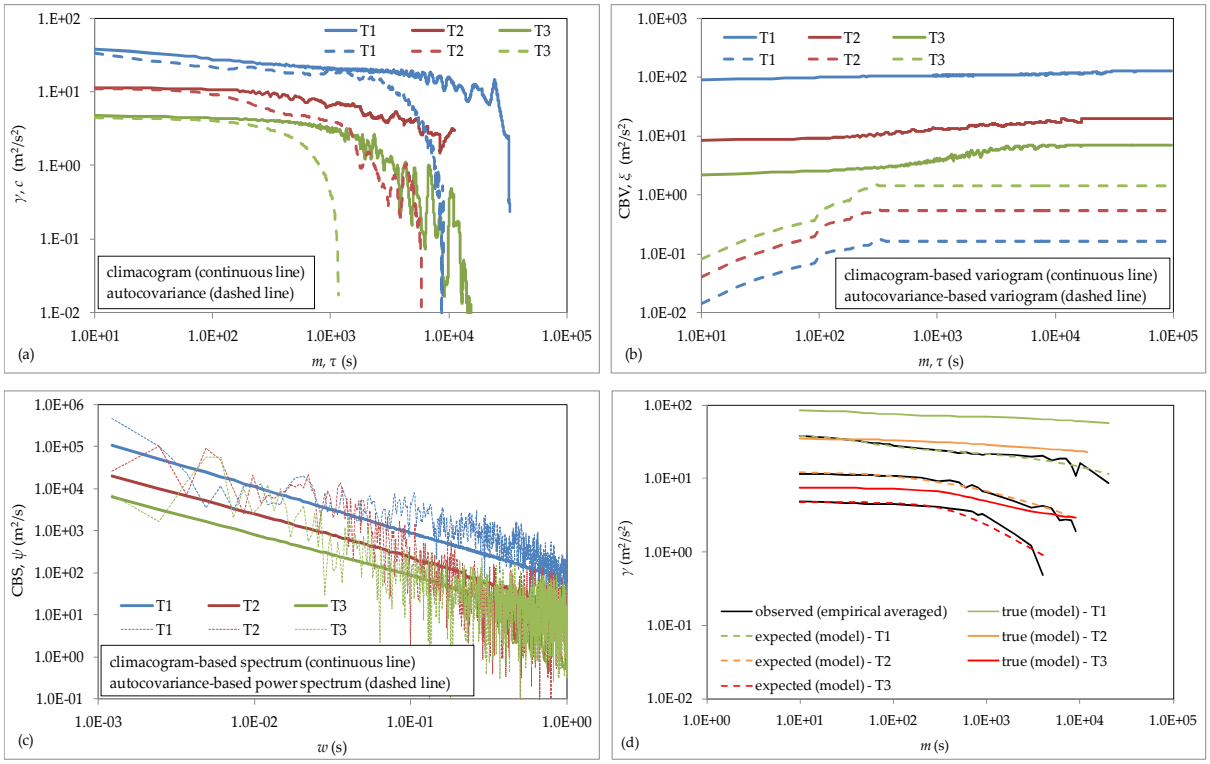

Figure 49: (a) Averaged empirical climacograms and autocovariances, (b) CBV and variograms, (c) CBS and power spectra for T1, T2 and T3, and (d) true, expected and empirical (averaged) climacogram values for the rainfall processes stochastic simulation. Source: Dimitriadis et al. (2016a).

We choose these two processes (wind and precipitation events) since they are of high importance in hydrometeorology. One may observe the transition from a process with low marginal variance having a power spectrum with a drop in the intermediate scales (like in the turbulent applications), to the one with larger marginal variance power spectrum (with no drop). Moreover, the similarities between the climacogram (and autocovariance) based metrics are again obvious. Although the above analysis can be considered quite simple, it highlights the deviation from Markov and white noise behaviours of the high resolution wind and precipitation events (as in the case of the examined turbulent processes). Particularly, the HK behaviour is apparent to all examined processes with an interestingly large fitting error (for more details see in Dimitriadis et al., 2016a). Therefore, although the physical mechanisms are considered to be substantially different between a laboratory small-scale turbulent process and an atmospheric meso-scale hydrometeorological process, the stochastic properties, such as the HK behaviour, seem to be quite similar.

6 Application to hydrometeorological processes

In this section, we show how the proposed model that adequately describes the examined small scale processes in the previous sections can be applied to macroscale hydrometeorological processes.

6.1 Stochastic analysis of a long daily precipitation timeseries

In this application, we analyze one of the longest daily precipitation timeseries recorded for over 130 years at the site of Hohenpeißenberg in Germany (latitude 47.801°N, longitude 11.011°E; data from www.gkd.bayern.de/). We apply a special case of the PBF marginal distribution (see section 2.4) introduced for its use in precipitation in Koutsoyiannis (2004a) and justified in Koutsoyiannis (2004b):

$$F(r) = 1 - \left(1 + \left(\frac{r}{a} - h\right)\right)^{-c} \tag{98}$$

where $r > ah$ is precipitation; $a > 0$ is a dimensionless scale parameter; $c > 0$ is a dimensionless parameter characterizing the right tail of the distribution and h is a dimensionless parameter representing a threshold value. Theoretically, $h = 0$ but values slightly different from zero highly improve fitting (Figure 50), while after the simulation we can set to zero any negative values of the synthetic timeseries. With this technique, the probability of zero rainfall can be also adequately preserved, i.e., $P(\underline{r} \leq 0) \approx P(r = 0)$. This technique can be justified through noticing that rainfall measurements are usually corrupted with significant uncertainties (Krajewski et al., 1998; Villarini et al., 2008) causing losses mainly due to wind effects (Sevruk and Nespor, 1998).

Note that here we ignore the seasonal periodicity of precipitation, which causes only a small increase in the dependence structure as depicted in the climacogram of Figure 50. Since we have a single timeseries we wish to estimate the dependence structure of the process through the mode climacogram rather than the mean one (for more details see in Dimitriadis and Koutsoyiannis, 2017). For this, we apply a Monte-Carlo analysis by generating one thousand daily timeseries of 130 years following the fitted marginal distribution and an HK process. We use the ESK distribution to simulate the white noise of the SMA scheme (section 3.3). From the Monte-Carlo ensemble, we calculate the mode for each scale with three-digit accuracy and thus, constructing the mode climacogram for the specified process. For the marginal distribution we use the same norm as in the previous section and for the climacogram we use its classical estimator (referred in this section as the E1 estimator):

$$\hat{\underline{\gamma}}(k\Delta) = \frac{1}{\lfloor n/k \rfloor - 1} \sum_{i=1}^{\lfloor n/k \rfloor} \left(\frac{1}{k} \left(\sum_{l=k(i-1)+1}^{ki} \underline{x}_l \right) - \frac{\sum_{l=1}^{n} \underline{x}_l}{n} \right)^2 \tag{99}$$

where $\lfloor n/k \rfloor$ is the integer part of n/k and $\underline{x}_l$ is the time-averaged process at scale Δ.

The parameters related to the dependence structure via the climacogram are estimated from data, based on the fitting norm, as: $\lambda = {r_s}^2$, where $r_s = 6.5$ mm is the standard deviation of r, and $H = 0.6$, whereas those of the marginal distribution are: $a = 42.25$ mm, $c = 7.7$ and $h = -0.1$, corresponding to $\mu = 2.1$ mm, $\sigma = 7.3$ mm, $C_s = 3.2$ and $C_k = 24$ (all estimations are based on the fitting norms in Equations 91 and 92). Also, we calculate their corresponding weights determined from the ME distribution (section 2.4.4) as 73%, 15%, 7% and 5%. Through a single synthetic timeseries of equivalent length and after setting any negative values to zero, the modelled marginal characteristics can be re-estimated as: $\mu = 3.3$ (3.1) mm, $\sigma = 6.5$ (6.5) mm, $C_s = 4.5$ (4.3), $C_k = 36.4$ (30.2) and dry probability 44% (48%), where inside parentheses are the empirical values that are adequately preserved. For illustration purposes, in Figure 50 we plot a 3000 days window of the observed vs. the simulated precipitation.

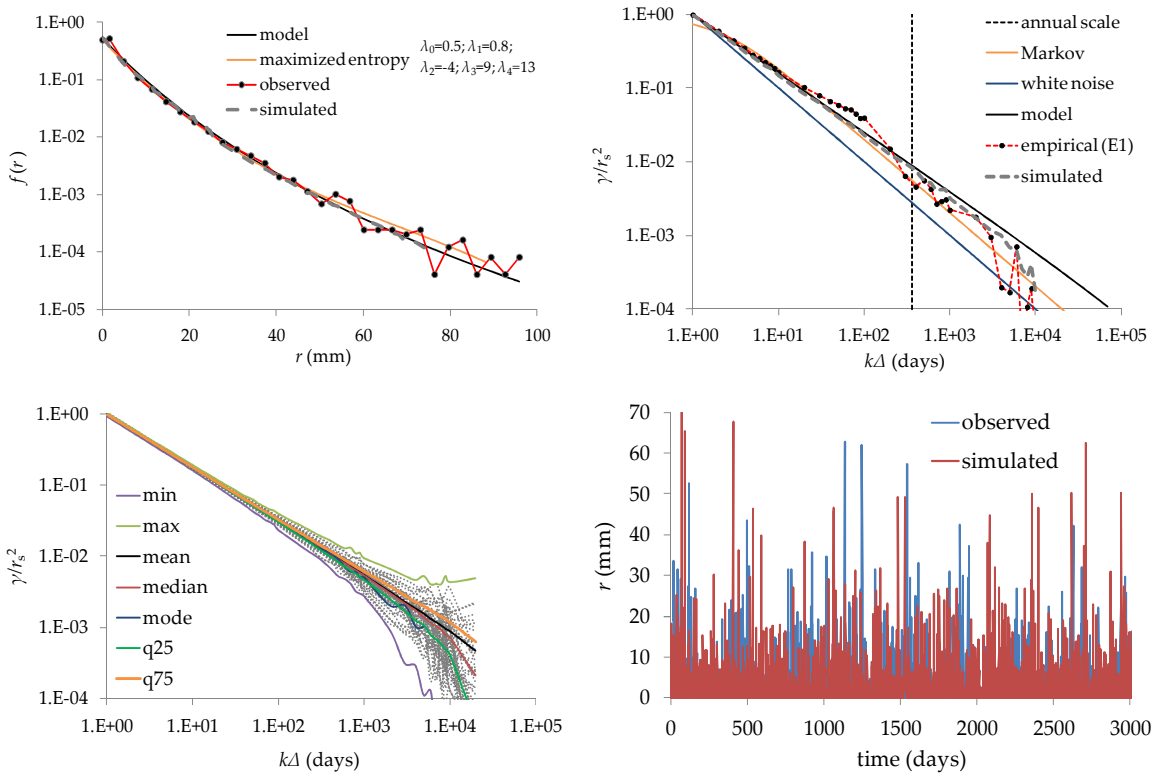

Figure 50: Empirical, modelled and simulated marginal distributions [upper left] and climacograms [upper right] for the standardized precipitation process; the mode and several other essential statistical measures of the standardized climacograms estimated from 10^3 synthetic timeseries (in the figure we depict only 50 empirical climacograms) [lower left]; a 3000 days window of the observed precipitation record along with a simulated one [lower right]. Source: Dimitriadis and Koutsoyiannis (2017).

6.2 Stochastic analysis of the longest hourly wind timeseries in Greece

For the hourly wind process we adopt the GHK process for the dependence structure. For the probability function we apply a special case of the PBF marginal distribution (section 2.4) which approximates the Weibull distribution for small hourly velocities and the Pareto distribution for larger ones (e.g., Aksoy et al., 2004; Brano et al., 2011). The dependence structure, marginal distribution and standardization scheme of wind are based on the preliminary analysis from thousands of stations around the globe, performed by Dimitriadis et al. (2015); Deligiannis et al. (2016); and Koutsoyiannis et al. (2017). A more thorough analysis justifying the above choices can be seen in Koutsoyiannis et al. (2017) and in section 6.3. The three-parameter GHK process and selected PBF marginal probability function can be written as:

$$\gamma(\kappa) = \frac{\lambda}{(1+\kappa/q)^{2-2H}} \tag{100}$$

$$F(v) = 1 - \left(1 + \left(\frac{v/v_s}{\alpha}\right)^b\right)^{-c/b} \tag{101}$$

where $v > 0$ is the wind process; $\kappa = k\Delta$ is the continuous time scale with Δ = 1 h the sampling time interval and k the discrete time scale; q is the scale parameter of the process; λ is the true variance of the continuous-time process; H is the Hurst coefficient; v_s is the standard deviation of the discretized process that should approximate the expected value of the square root of the climacogram for scale k = 1, i.e., $\sqrt{\gamma(\Delta)} = (1+\Delta/q)^{H-1}\sqrt{\lambda}$; and α is the scale parameter and b and c are the shape parameters of the marginal distribution, all dimensionless. Note that we standardize the wind process, in order to homogenize all timeseries recorded at different locations, altitude and climatic conditions.

We choose to apply the above stochastic model to nine hourly wind timeseries of different lengths located in Greece (Table 20). The expression for the bias of the classical estimator of the climacogram is derived in Tyralis and Koutsoyiannis (2011) for an HK process and generalized for all processes in Koutsoyiannis (2011). Here, we use the general expression and, since the timeseries have different lengths n, we apply an estimator of the climacogram adjusted for n (referred in this section as the E2 estimator):

$$\underline{\hat{\gamma}}(k\Delta) = \frac{1-k/n}{\lfloor n/k \rfloor - 1}\sum_{i=1}^{\lfloor n/k \rfloor}\left(\frac{1}{k}\left(\sum_{l=k(i-1)+1}^{ki}\underline{x}_l\right) - \frac{\sum_{l=1}^{n}\underline{x}_l}{n}\right)^2 + \gamma(n\Delta) \tag{102}$$

where $\underline{\hat{\gamma}}(k\Delta)$ is an unbiased estimator of the climacogram $\gamma(k\Delta)$, since $\mathrm{E}\left[\underline{\hat{\gamma}}(k\Delta)\right] = \gamma(k\Delta)$.

The parameters related to the dependence structure via the climacogram are estimated from data as: $\lambda = 1.3$, $q = 5$ h and $H = 0.75$, whereas for the marginal distribution as: $a = 6$, $b = 1.9$ and $c = 14.8$, corresponding to μ = 1.9, σ = 1.1 ($\approx \sqrt{\lambda}$), C_s = 1.2 and C_k = 4.8 (all estimations are based on the fitting norms in Equations 91 and 92). Also, we calculate their corresponding weights

determined from the ME density function as 43%, 32%, 16% and 9%. To emulate the observed wind timeseries one could set to zero any values of the synthetic timeseries that are below the corresponding recording threshold of an anemometer, which is in average around 0.5 m/s depending on the type of the anemometer (e.g., Conradsen et al., 1984)). For illustration purposes, in Figure 51 we plot a 1000-day window of the observed vs. the simulated wind speed at Kos Island. The empirical and modelled probability of wind speed less than or equal to 0.5 m/s are both around 20%.

Table 20: General information of the meteorological stations and statistical characteristics of the hourly wind timeseries (downloaded from ftp.ncdc.noaa.gov). Source: Deligiannis et al. (2016).

hourly wind station	longitude (deg)	latitude (deg)	above sea elevation (m)	no. years	mean (m/s)	stdev (m/s)	missing values (%)	zero values (%)
Heraκleio	25.183	35.333	39	39	4.583	2.918	8.8	6.3
N. Aghialos	22.8	39.217	15	17	3.258	2.331	28	19
Karpathos	35.417	27.15	20	17	7.506	4.074	30.4	3.9
Santorini	36.4	25.483	38	24	5.701	3.229	29.5	7.5
Kos	36.8	27.083	125	33	4.805	2.7	15	7
El. Venizelos	37.93	23.93	96	11	3.954	2.995	0.6	1.9
Limnos	39.917	25.233	5	38	4.458	3.546	23	17.5
Paros	37.02	25.13	36	11	5.567	3.265	46.8	6.5
Meganissi	38.95	20.767	4	40	3.571	2.746	36.3	19.4

Note that σ and λ should approximate unity but they are slightly larger due to the cyclo-stationary effect of the daily and seasonal periodicities of the wind process (Deligiannis et al., 2016; Dimitriadis and Koutsoyiannis, 2015b). These effects cause the small increase of climacogram around daily and annual scales. Here, for simplicity, we ignore these effects and we apply a stationary rather than cyclo-stationary model.

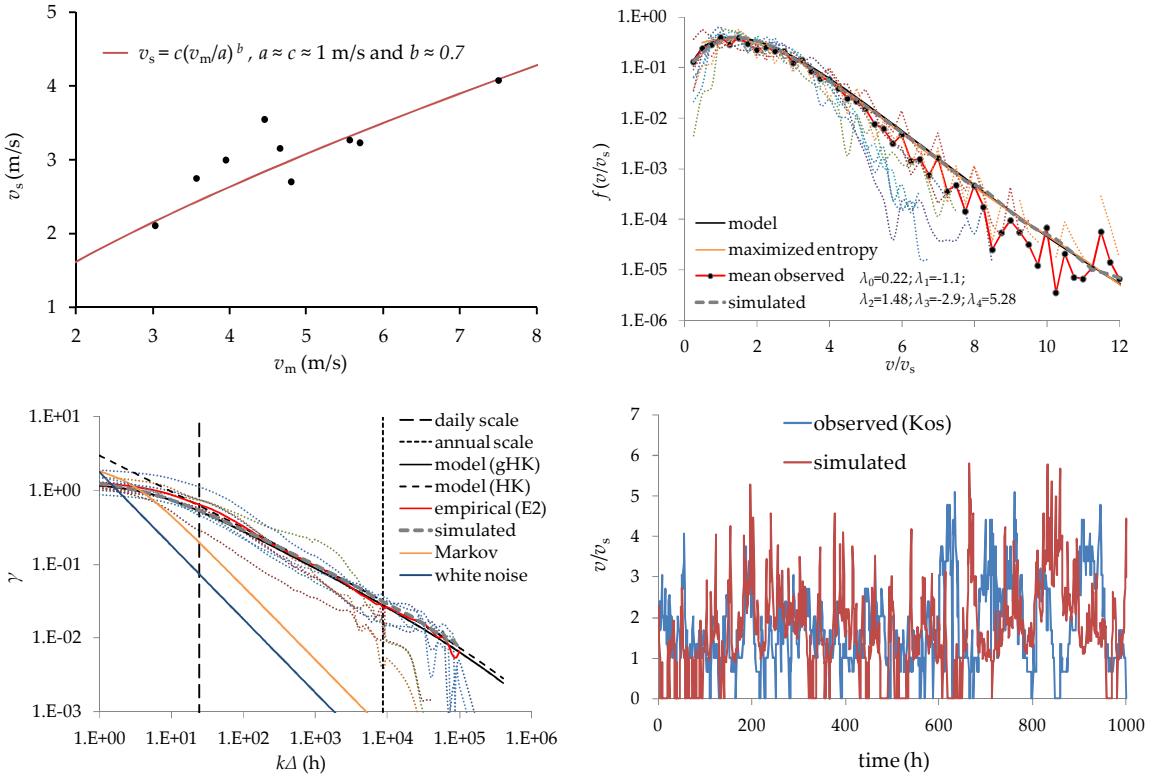

Figure 51:Empirical mean (v_m) vs. standard deviation of the nine timeseries along with the fitted model [upper left]; the empirical, model and simulated marginal distributions [upper right] and climacograms [lower left] for the standardized wind process; a 1000-day window of the observed standardized wind process in Kos island along with a standardized simulated one [lower right]. Source: Dimitriadis and Koutsoyiannis (2017).

6.3 Global stochastic analysis of the hourly wind process

Understanding atmospheric motion in the form of wind is essential to many fields in geophysics. Wind is considered one of the most important processes in hydrometeorology since, along with temperature, it drives climate dynamics. Currently, the interest for modelling and forecasting of wind has increased due to the importance of wind power production in the frame of renewable energy resources development. For the investigation of the large scale of atmospheric wind speed, we use over 15000 meteorological stations around the globe recorded mostly by anemometers and with hourly resolution (www.nooa.gov; GHCN database). In total, we analyze almost 4000 stations from different sites and climatic regimes by selecting time series that are still operational, with at least one year length of data, at least one non-zero measurement per three hours on average and at least 80% of non-zero values for the whole time series (Figure 52). This data set is referred below as “global”.

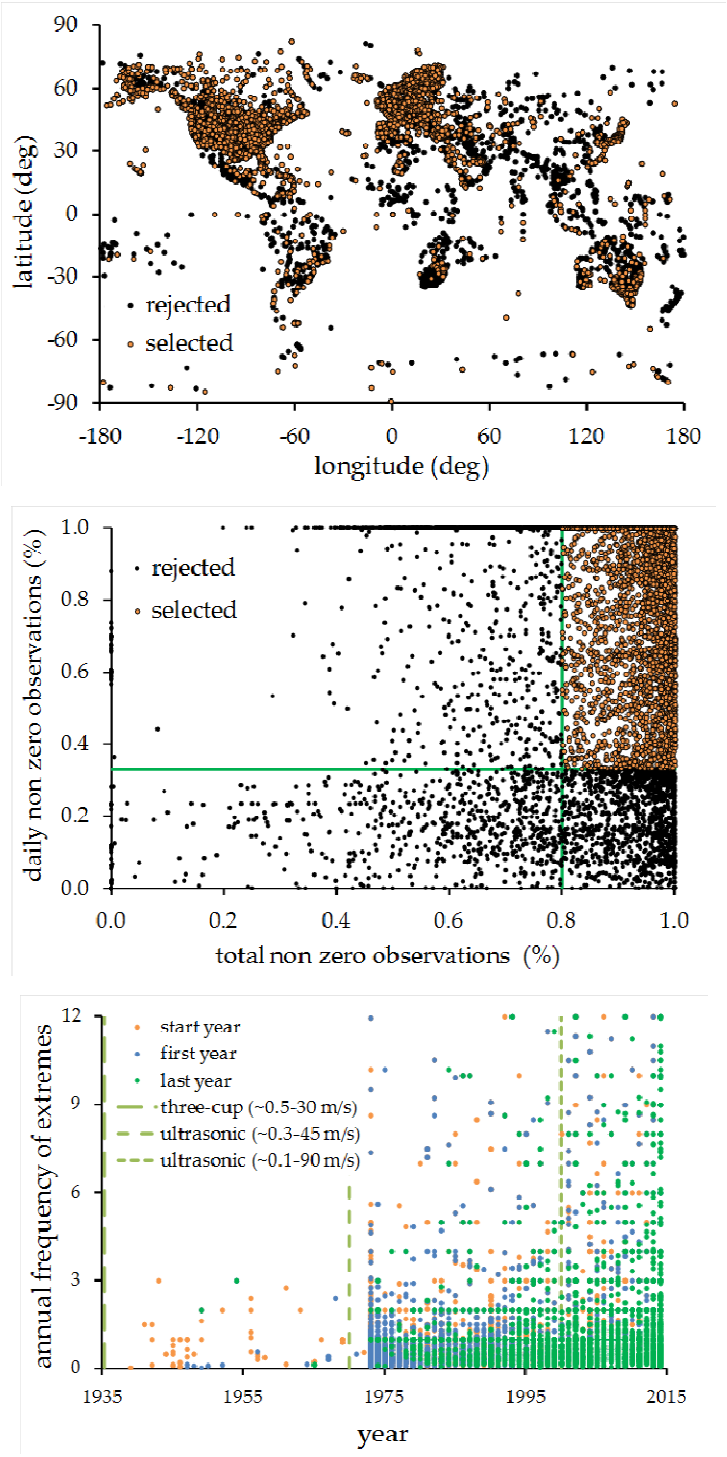

Figure 52: (upper) Distribution of the wind speed stations over the globe; (middle) sketch about the selection of the stations in the analysis; (lower) evolution of the frequency of measured extremes in the stations (where the 'start' year denotes the first operational year of the station and the 'first' and 'last' year denote the first and last year that an extreme value was recorded, respectively). Source: Koutsoyiannis et al. (2017).

By standardizing all series we formed a sample of $\sim 0.5\times10^9$ values to estimate the marginal distribution, and an ensemble of 3886 series, each with $\sim 10^5$ values on average, to estimate the dependence structure through the climacogram. A known problem of field measurements of wind (particularly those originating from over 70 years ago), is that the technology of measuring devices has been rapidly changed (Manwell et al., 2010, sect. 2.8.3). For example, in Figure 52 we illustrate a rather virtual increase of extreme wind events after the 1970s which is mainly due to the inability of older devices to properly measure wind speeds over 30 m/s (i.e., category I of Saffir–Simpson hurricane wind scale). Furthermore, in common anemometer instrumentation there is a lower threshold of speed that could be measured, usually within the range 0.1 – 0.5 m/s (e.g., www.pce-instruments.com). It should be noted that, as the recorded wind speed decreases, so does the instrumental accuracy and it may be a good practice to always set the minimum threshold to 0.5 m/s to avoid measuring the errors of the instrument (e.g., zero or extremely low values) in place of the actual wind speed that can never reach an exact zero value.

In an attempt to incorporate smaller scales, starting from the microscale of turbulence, we include again the dataset of the previous application of turbulence, using it as an indicator of the similar statistical properties of small scale wind (Castaing et al., 1990). In addition to the 40 time series of the longitudinal turbulent velocity in section 5.1, here we also use another 40 time series of transverse velocity, measured at the same points with the longitudinal one; again each time series has $n = 36\times10^6$ data points with a sampling interval of 25 μs (Kang et al., 2003). The coefficients of skewness and kurtosis are estimated as 0.1 and 3.1 for the transverse velocity, respectively. Stochastic similarities between small scale atmospheric wind and turbulent processes abound in the literature as for example in terms of the marginal distribution (Monahan, 2013, and references therein), of the distribution of fluctuations (Böttcher et al., 2007, and references therein), of the dependence structure (Dimitriadis et al., 2016a, and references therein) and of higher-order behaviour such as intermittency (e.g., Mahrt, 1989). This data set is referred below as “small”.

Finally, to link the large and small scale of atmospheric wind we analyse an additional time series, referred to as “medium”, provided by NCAR/EOL of one-month length and with a 10 Hz resolution. This time series has been recorded by a sonic anemometer on a meteorological tower located at Beaumont KS and it includes over 25×10^6 longitudinal and transverse wind speed measurements (http://data.eol.ucar.edu/; Doran, 2004).

The statistical characteristics based on moments up to fourth order are shown in Figure 53; interestingly, there appears to be a rather well defined relationship between mean and standard deviation. The plot of coefficient of kurtosis vs. coefficient of skewness indicates that Weibull distribution falls close to the lower bound of the scatter of empirical points.

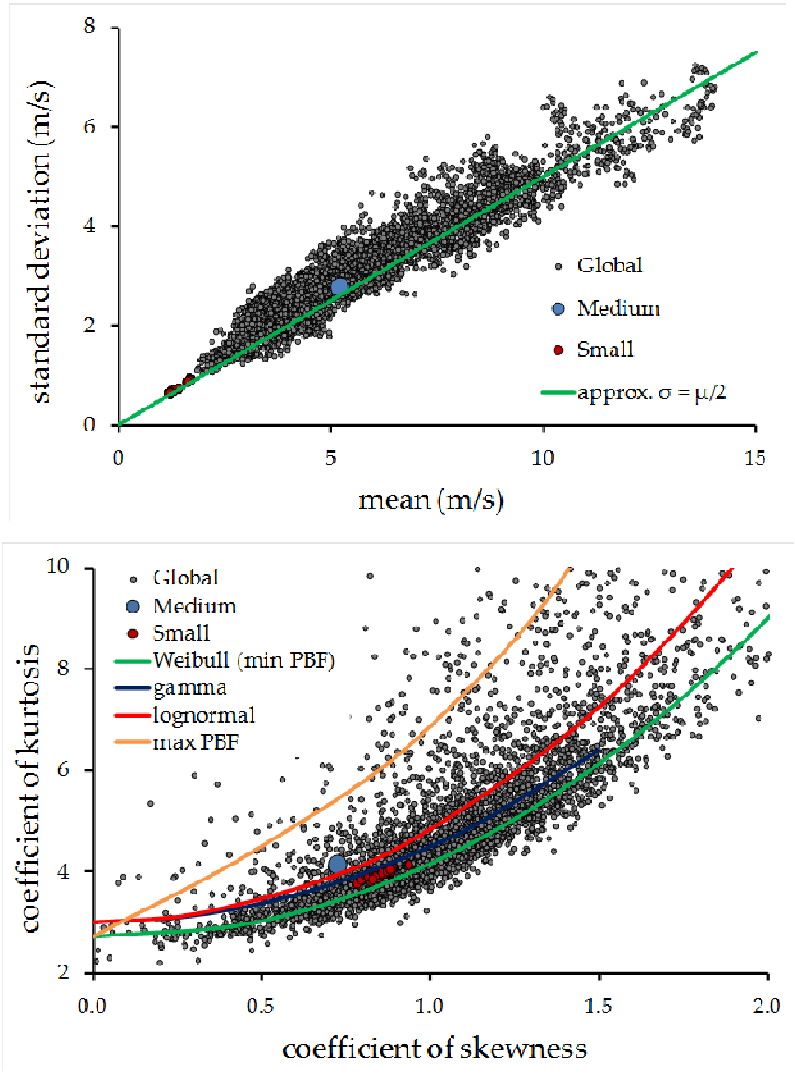

Figure 53: Standard deviation vs mean (upper) and coefficient of kurtosis vs. coefficient of skewness of all time series (source: Koutsoyiannis et al., 2017).

Numerous works have been conducted for the distribution of the surface wind speed (see in Koutsoyiannis et al., 2017, and references therein). The Weibull distribution is proven very useful in describing the wind magnitude distribution for over three decades (Monahan, 2013, and references therein). However, various studies illustrate empirical as well as physically-based deviations from the Weibull distribution (Drobinski and Coulais, 2012, and references therein). Due to the discussed limitations of properly measuring wind speed most studies have focused on a local or small scale. In such cases where there is limited empirical evidence, but we could search for a physical justification for the left and right tail of the probability function.

It can easily be proven that the magnitude of uncorrelated Gaussian distributions follows the Rayleigh distribution. However, there is empirical and theoretical evidence that the small-scale distribution of turbulence is not Gaussian and it is expected that this should also be the case for the

components of wind speed. Through Monte-Carlo experiments we illustrate in Figure 54 that correlated non-Gaussian components result in a distribution close to Weibull and is in agreement with small and medium scale observations.

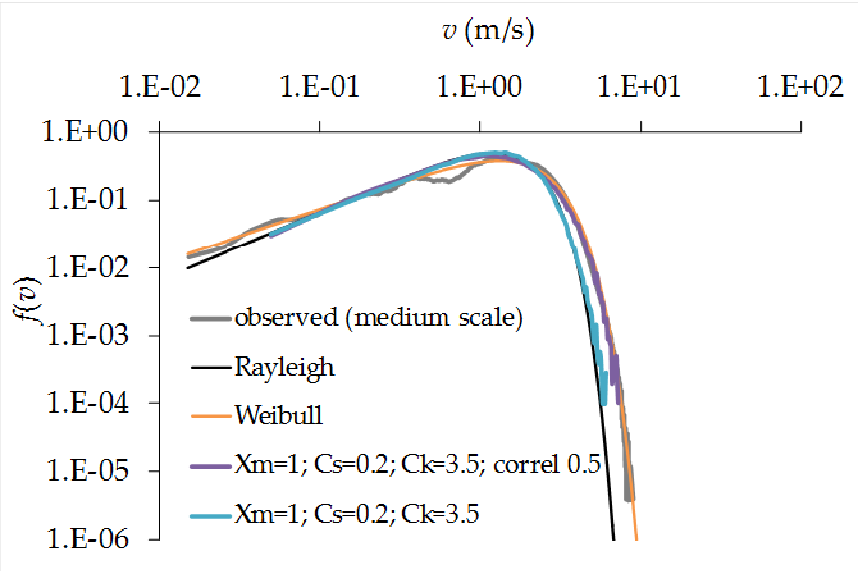

Figure 54: Probability density function of the medium scale time series along with theoretical and Monte Carlo generated distributions (source: Koutsoyiannis et al., 2017).

The distribution of the "global" time series appear to deviate from Weibull, gamma and log-normal distributions, and is closer to a distribution with a much heavier tail, such as the PBF:

$$F(v) = 1 - \left(1 + \left(\frac{v}{\alpha v_s}\right)^b\right)^{-c/b} \tag{103}$$

where $v > 0$ is the wind speed, v_s is the standard deviation of the wind speed process; α is a scale parameter and b and c are the shape parameters of the marginal distribution, all three dimensionless.

The fitted distribution to all data sets and the fitted parameters are α = 3.5, b = 1.9, c = 8.5 (see Figure 55).

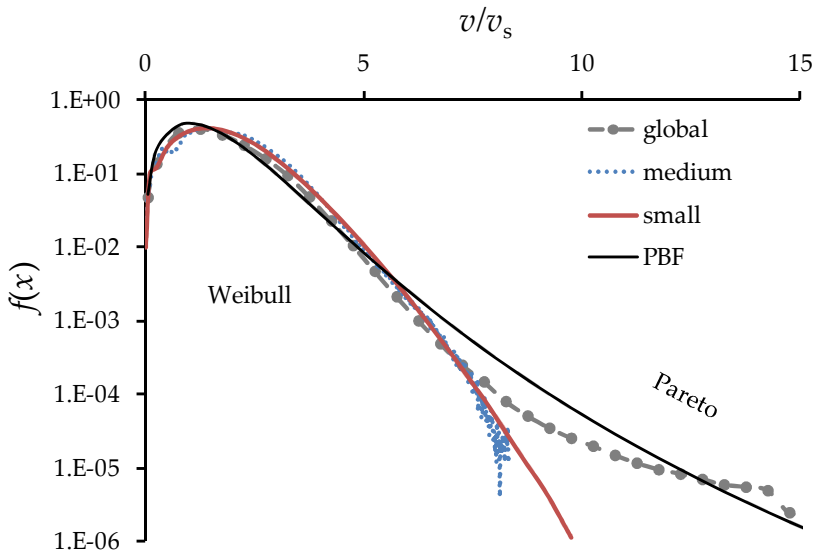

Figure 55: Probability density function of the velocity of grid-turbulent data (small) and of the wind speed of the medium and global scale time series along with fitted theoretical distributions (source: Koutsoyiannis et al., 2017).

The mean estimated climacograms from the data indicate that the model is also applicable for the wind speed at all scales with parameters estimated as $\lambda \approx 1$, $M = 1/3$, $H = 5/6$ and $\alpha = 6$ h (Figure 56).

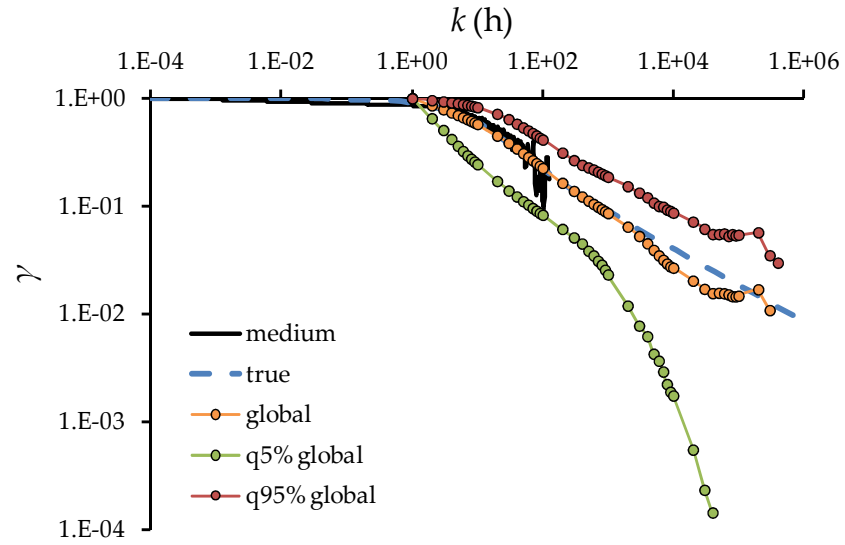

Figure 56: Climacogram of the wind speed process estimated from the medium and global series (source: Koutsoyiannis et al., 2017).

6.4 Global stochastic analysis of the hourly temperature process

In this last application we analyze the dependence structure of the air temperature process close to surface. For the microscale structure, we use a 10 Hz resolution timeseries recorded for a 2-month period by a sonic anemometer at Beaumont USA (https://data.eol.ucar.edu/dataset/45.910). For the macro-scale structure, we use a global database of hourly air temperature (www.nooa.gov; GHCN database). In total, we analyze over 5000 stations from different sites and climatic regimes by selecting time series with at least 1 year length and at least one measurement per three hours (Figure 57). It can be assumed that the air temperature process follows a Gaussian distribution (Koutsoyiannis, 2005). Indeed, the 90% of the time series have coefficient of skewness around 0 and of kurtosis around 3 with a standard deviation for both coefficients approximately equal to 1 (Figure 58). We normalize all time series and we estimate the dependence structure through the climacogram, autocovariance and power spectrum (Figure 59 and 60) following the methodology in Dimitriadis et al. (2016a).

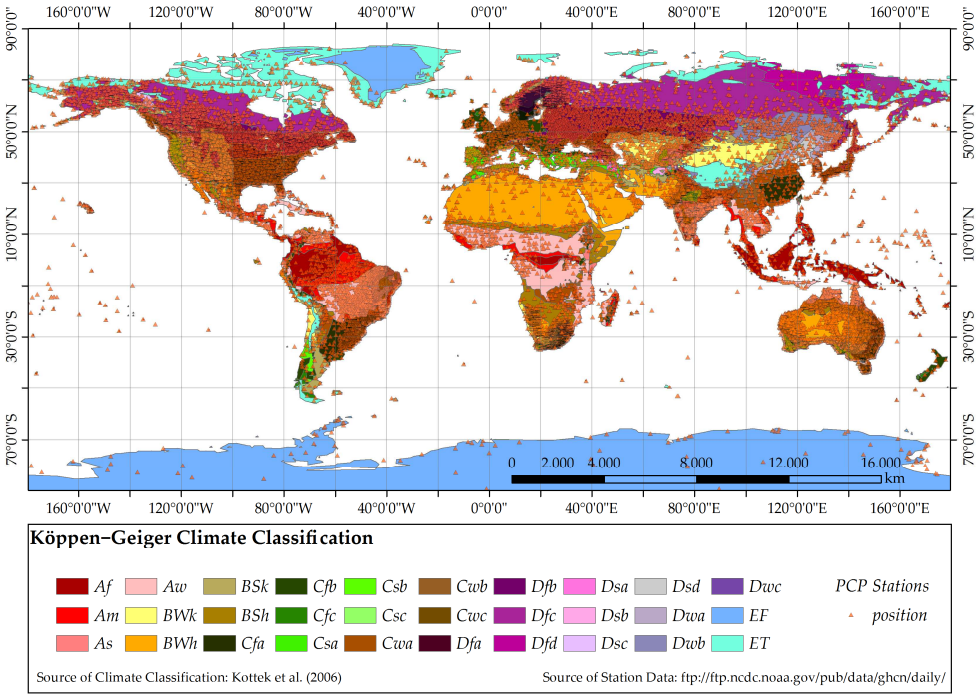

Figure 57: Locations of the selected hourly time series of air temperature from the global database along with the Koppen climatic zones. Source: Lerias et al. (2016).

The mean estimated climacograms and climacogram-based spectrum from the data indicate that, interestingly, the proposed mixed HHK model is also applicable here with parameters estimated as: $\lambda \approx 1$, $M = 1/3$, $H = 5/6$ and $\alpha = 3.3$ d (Figure 59).

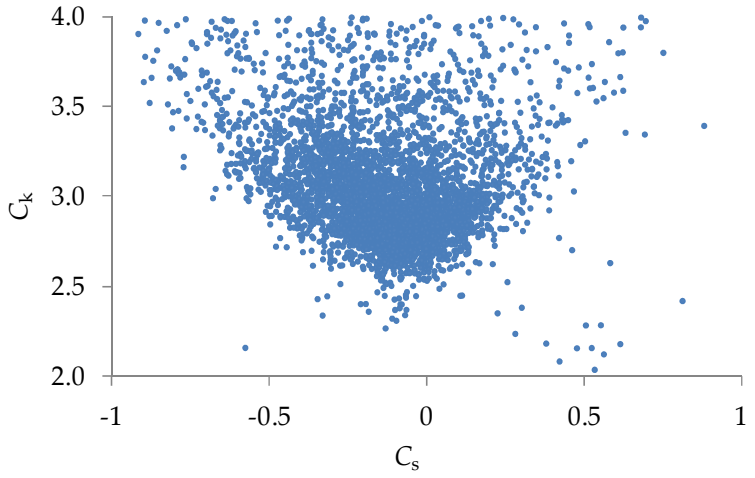

Figure 58: Coefficient of skewness vs. coefficient of kurtosis for the 90% of the macro-scale temperature time series (source: Koutsoyiannis et al., 2017).

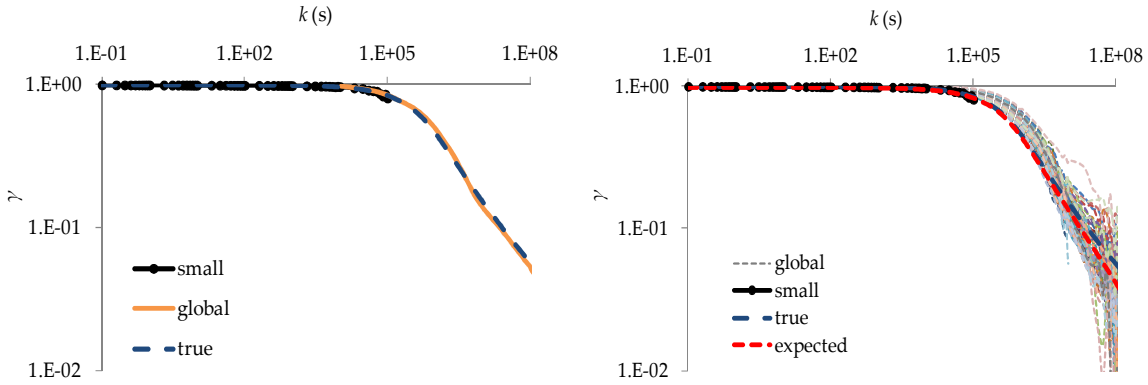

Figure 59: Climacogram of the normalized temperature for the micro-scale time series (small) and the set of hourly air temperature time series (global; upper: average climacogram; lower: climacograms of 100 different time series), compared to the fitted model (true and expected). Source: Koutsoyiannis et al. (2017).

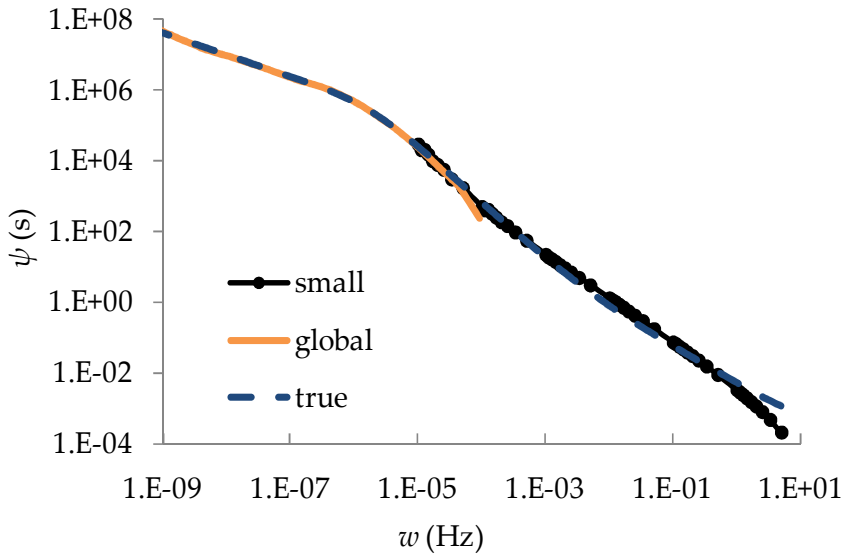

Figure 60: Climacogram-based spectrum of the normalized temperature for the micro-scale time series (small) and the set of hourly air temperature time series (global; average from all time series), compared to the fitted model (true). Source: Koutsoyiannis et al. (2017).

6.5 Global stochastic analysis of hydrometeorological processes based on the Koppen-Geiger climatic-classification

An annual change in hydroclimatic processes is commonly attributed to anthropogenic climatic change. However, most of the studies have not taken into consideration the possibility of the Hurst phenomenon. Usually, high (low) values of a hydroclimatic process are followed by high (low) ones, meaning that observations appear in groups. In other words, the autocorrelation coefficient remains quite high as the scale increases due to this clustering effect. Here, we analyze (additional to the analyses of the previous sections) several hydroclimatic processes classified by the Koppen-Geiger system of climatic zones and in terms of the climacogram in order to determine whether they exhibit such behaviours of Long-Term Persistence (LTP). Again, we use the hourly database GHCN with over 15,000 stations around the globe for the temperature, dew point, atmospheric wind, precipitation and atmospheric pressure. First, we estimate the Hurst parameter for various 30-year time periods to test that there are no suspicious changes in LTP behaviour. The results from this analysis are shown in Lerias et al. (2016) for the temperature and dew point processes, in Sotiriadou et al. (2016) and Tyralis et al. (2017) for the precipitation process, in Deligiannis et al. (2016) for the wind process and in Dimitriadis et al. (2016d) for the atmospheric pressure. In the Table below we show the average Hurst parameter for each climatic-zone.

Table 21: Hurst parameter under Köppen-Geiger classification (source: Dimitriadis et al., 2016d).

Hurst parameter / Koppen-Geiger classification	temperature	dew point	wind Speed	precipitation	atmospheric pressure
A	0.79	0.78	0.84	0.62	0.71
B	0.73	0.77	0.82	0.59	0.72
C	0.70	0.71	0.87	0.65	0.73
D	0.72	0.68	0.85	0.66	0.65
E	0.68	0.65	0.70	0.83	0.71

Finally, we estimate the prediction intervals for the 30 year period as well as the corresponding error (prediction error) as shown in the next Figures. If the prediction error is small for all examined 30-year periods and each station, then the model can describe adequately the climatic variability of the process and so, the changes observed during the last decades can be attributed to the Hurst phenomenon and not to anthropogenic factors. This should not be confused with the urbanization factor. For example, the major cause for the deterioration of the natural defence mechanism against floods and hurricanes is the destruction of forests. Indeed the damages from severe flood events and hurricanes have increased over the last decades but that does not mean that the human-kind has increased the severe storm events nor has changed the annual trend of global climatic processes such as temperature, humidity (through the dew point), wind and precipitation (similar to the atmospheric pressure).

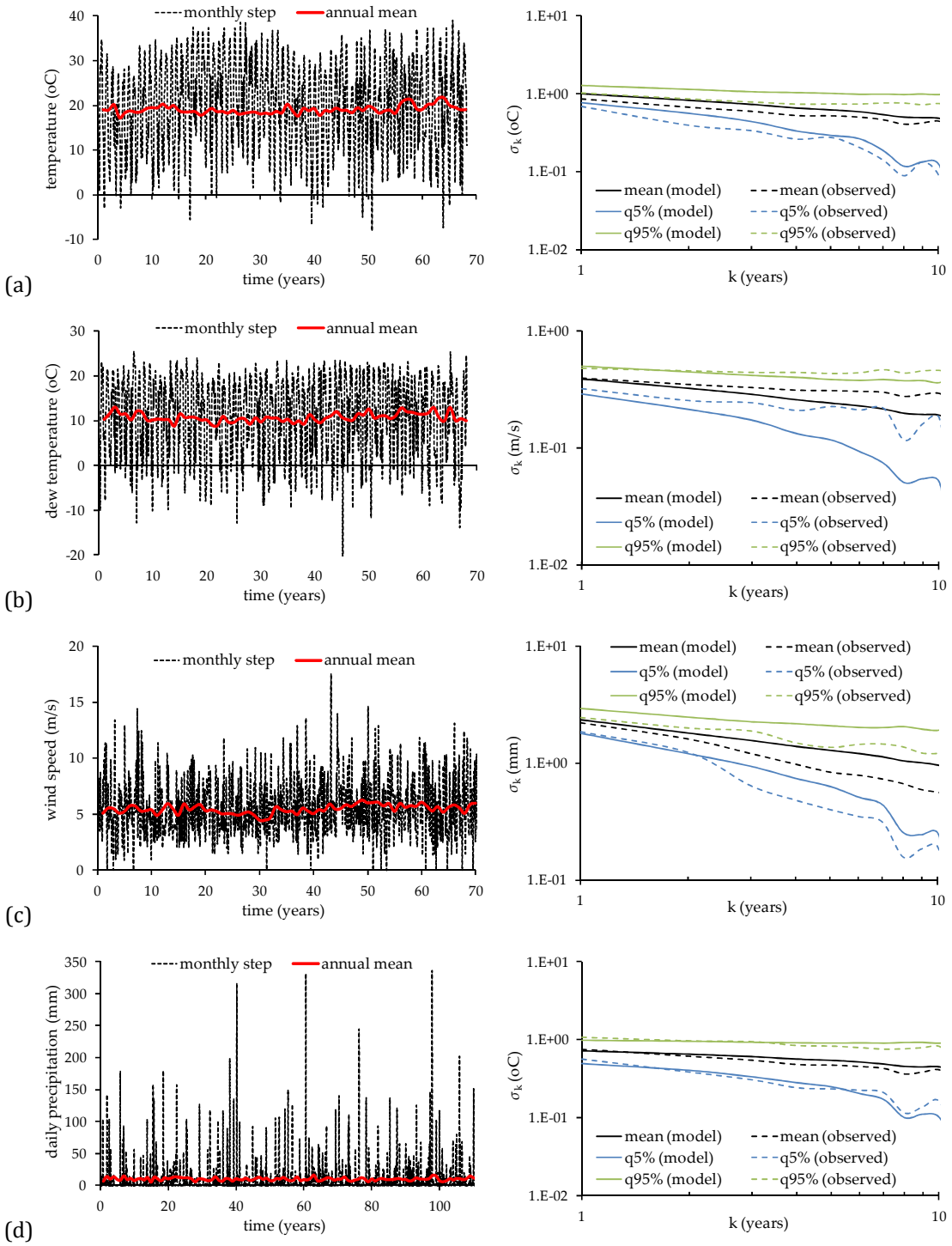

Figure 61: (a) temperature and (b) dew point timeseries and HK model for a station located in Dallas, USA; (c) wind speed timeseries and HK model for a station located in Winter Trail, Alaska; and (d) precipitation timeseries and HK model for a station located in North-East Australia. Source: Dimitriadis et al. (2016e) and references therein. Source: Dimitriadis et al. (2016d).

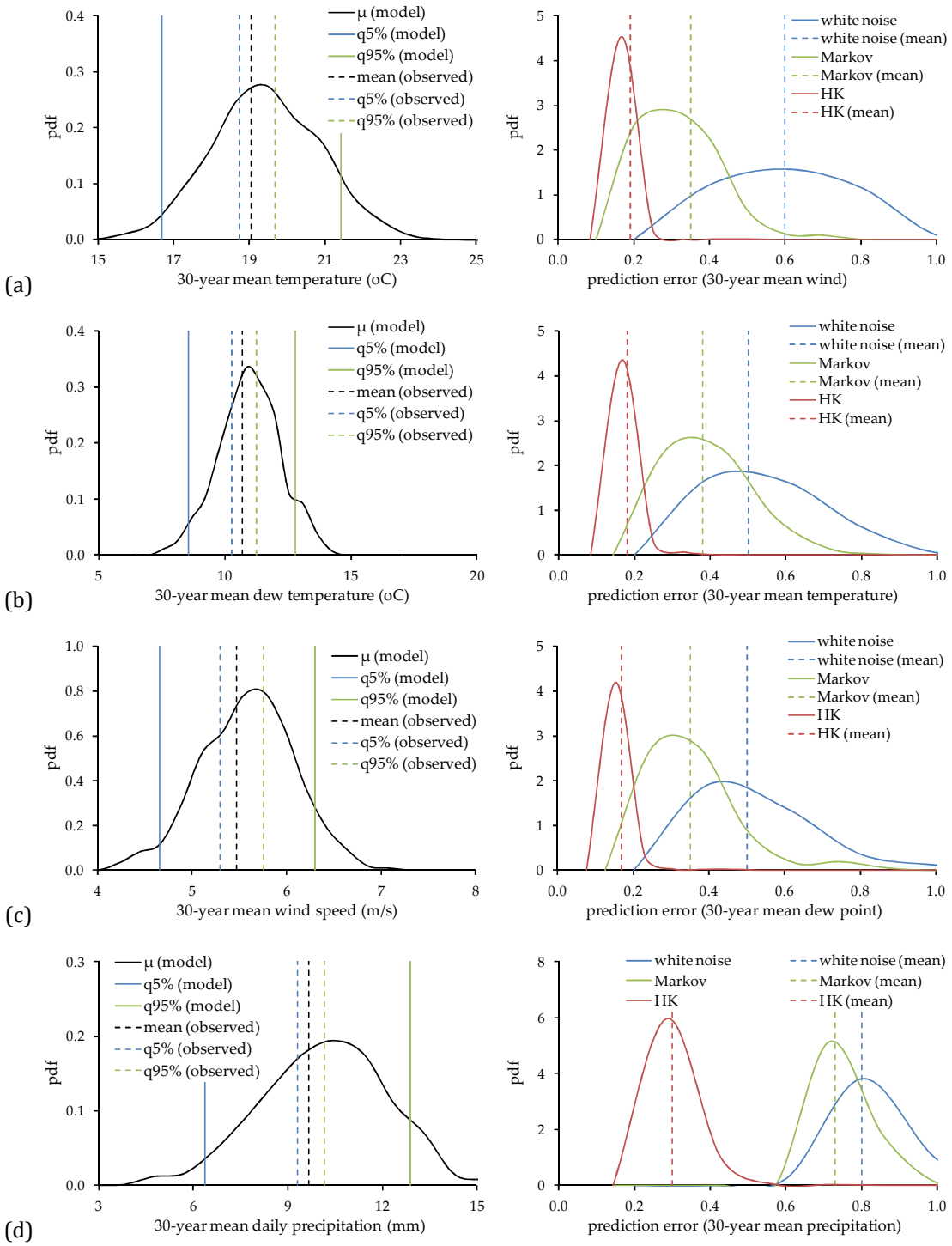

Figure 62: Prediction intervals for the examined station described in the previous figure and the overall prediction error for (a) temperature, (b) dew point, (c) wind speed and (d) precipitation. Source: Dimitriadis et al. (2016e) and references therein. Source: Dimitriadis et al. (2016d).

Overall, the Hurst parameter and the prediction errors are estimated from this analysis (following an atmospheric circulation pattern) as: (a) H = 0.85 for the temperature process, with a prediction error lower than 10% for the 73% of stations, (b) H = 0.83 for the wind process, with a prediction error lower than 10% for the 71% of stations, (c) H = 0.80 for the dew point process, with a prediction error lower than 10% for the 80% of stations, and (d) $H \approx 0.67$ for the precipitation and atmospheric pressure processes, with a prediction error lower than 20% for the 86% of stations.

7 Conclusions and summary of thesis major innovations

The deeper understanding of the high complexity of atmospheric dynamics has been the key factor towards the further enhancement of predictability of hydrometeorological processes. Although in the last decades there has been a substantial increase of measurements and of the number of meteorological stations, technological and theoretical advances on the recording devices, breakthroughs on the mathematical techniques etc., the predictability has not significantly improved. The latter conclusion is based on the simple observation that (extreme or mild) weather phenomena most of the times still remain unpredictable. Hurst-Kolmogorov dynamics, i.e., the dynamics causing random changes on the behaviour of a process that result in a clustering of events, maybe a simple but rather a vital explanation of this inability of accurate predictions. In this thesis, we analyze numerous of processes originating from the microscale of turbulence and extending to macroscale hydrometeorological processes and we identify stochastic similarities between them such as the HK behaviour with Hurst parameters considerably above 0.5. For this, we first develop the stochastic framework for the empirical as well as theoretical estimation of the marginal characteristic and second order dependence structure of a process, and by also developing algorithms for stochastic synthesis of mathematical processes as well as stochastic prediction of physical ones.

The major innovations of the thesis are (a) the further development and extensive application to numerous stationary and isotropic processes of the second-order stochastic framework including models in continuous and discrete time, expected values and classical estimators; (b) the estimation of the dimensionless statistical error (due to discretization and bias) through Monte-Carlo analysis of a variety of Markov and HK models, for the power spectrum, autocovariance and climacogram, with the latter exhibiting the smaller error and the former the larger one for all examined processes; (c) the exact mathematical expression of the statistical bias of the autocovariance and power spectrum classical estimator as a function of the theoretical autocovariance; (d) the introduction of the Markov process for a different time interval and response time, and the expressions for its generation through an ARMA(1,1) model; (e) the further development of the Sum of Autoregressive (SAR) and Moving Average (SARMA) schemes that can generate a large variety of Gaussian processes approximated by a finite sum of AR(1) or ARMA(1,1) processes; (f) the further development of the Symmetric-Moving-Average (SMA) scheme that can generate any process second-order dependence structure as well as certain aspects of the intermittency behaviour, and any marginal distribution by approximating a finite number of statistical moments; (g) the introduction and application of an extended Hybrid HK model that is in agreement with an

interestingly large variety of turbulent flows, such as grid-turbulence (analyzing ~1.5 billion of data) and turbulent thermal jets of positive buoyancy (by performing numerous laboratory experiments following the laser-induced-fluorescence technique, and by analyzing ~15,000 data), as well as hydrometeorological processes, such as atmospheric wind and temperature (analyzing ~0.5 billion of data for each process and at various micro and macro scales); (h) estimation of the Hurst parameter based on the Köppen-Geiger climatic-classification for numerous hydrometeorological processes, such as temperature, atmospheric wind, precipitation, atmospheric pressure and dew point (analyzing almost 5000 stations for each process with at least 30 years of records); and (i) the further development of the multi-dimensional classical second-order stochastic framework and HK process, and application to turbulence and geostatistics. Incidental contributions and moderate innovations of this thesis are: (a) several illustrative comparisons between complex natural as well as purely deterministic processes; (b) the further development of analogue and stochastic prediction algorithms based on the climacogram; (c) the estimation of the most uncertain parameters in flood inundation modelling based on commonly-used hydraulic models and on benchmark geometries; (d) the introduction of an optimization target function and the further development of the climacogram-based estimators, for the identification of the dependence structure of a process, in case of the analysis of a single time series and of several time series of the same process with different lengths and identical lengths.

An overall conclusion is that a simple model (from the view of Stochastics) can adequately explain (and thus, predict) several aspects of turbulence in microscale and hydrometeorological processes. Future investigations will mainly include the further investigation of the generating schemes for simulating cyclostationary processes and of the HK behaviour for additional atmospheric and hydroclimatic processes as well as of their marginal characteristics and additional aspects such as their intermittent behaviour, and the deeper understanding towards a physical justification of the origins of the HK dynamics in Nature.

Some scientific and philosophical questions to the Readers are:

- Will Determinism ever be able to fully describe (and predict) Natural phenomena?
- Will Stochastics ever be acceptable by scientists as well as non-scientists?
- Assuming that the world continues at the same course; will Stochastics be useful in many years from now where observations will be abundant?
- Is Stochasticity an intrinsic property of Nature?

References

Aksoy, H., Toprak, Z.F., Aytek, A., Ünal, N.E., 2004. Stochastic generation of hourly mean wind speed data. Renew. Energy 29, 2111–2131.

Arnold, B.C., Press, S.J., 1983. Bayesian inference for Pareto populations. J. Econom. 21, 287–306.

Avila, M., Willis, A.P., Hof, B., 2010. On the transient nature of localized pipe flow turbulence. J. Fluid Mech. 646, 127–136.

Bachelier, L., 1900. Theory of speculation. Dimson E M Mussavian 1998 Brief Hist. Mark. Effic. Eur. Financ. Manag. 4, 91–193.

Barndorff-Nielsen, O., 1978. Hyperbolic distributions and distributions on hyperbolae. Scand. J. Stat. 151–157.

Batchelor, G.K., 1953. The theory of homogeneous turbulence. Cambridge university press.

Batchelor, G.K., Townsend, A.A., 1949. The nature of turbulent motion at large wave-numbers, in: Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences. The Royal Society, pp. 238–255.

Bates, P., Trigg, M., Neal, J., Dabrowa, A., 2013. LISFLOOD-FP User Manual. Univ. Bristol Bristol.

Beran, J., Feng, Y., Ghosh, S., Kulik, R., 2013. Long-memory processes. Monogr. Stat. Appl. Probab.

Bercher, J.-F., Vignat, C., 2008. An entropic view of Pickands' theorem, in: Information Theory, 2008. ISIT 2008. IEEE International Symposium On. IEEE, pp. 2625–2628.

Böttcher, F., Peinke, J., others, 2007. Small and large scale fluctuations in atmospheric wind speeds. Stoch. Environ. Res. Risk Assess. 21, 299–308.

Brano, V.L., Orioli, A., Ciulla, G., Culotta, S., 2011. Quality of wind speed fitting distributions for the urban area of Palermo, Italy. Renew. Energy 36, 1026–1039.

Brouers, F., 2015. The Burr 12 Distribution Family and the Maximum Entropy Principle: Power-Law Phenomena are not necessarily Nonextensive. ArXiv Prepr. ArXiv151007489.

Brunner, G.W., 2010. HEC-RAS (River Analysis System), in: North American Water and Environment Congress & Destructive Water: ASCE, pp. 3782–3787.

Burr, I.W., 1942. Cumulative frequency functions. Ann. Math. Stat. 13, 215–232.

Castaing, B., Gagne, Y., Hopfinger, E.J., 1990. Velocity probability density functions of high Reynolds number turbulence. Phys. Nonlinear Phenom. 46, 177–200.

Castro, J.J., Carsteanu, A.A., Fuentes, J.D., 2011. On the phenomenology underlying Taylor's hypothesis in atmospheric turbulence. Rev. Mex. Física 57, 60–64.

Cerutti, S., Meneveau, C., 2000. Statistics of filtered velocity in grid and wake turbulence. Phys. Fluids 12, 1143–1165.

Chamorro, L.P., Porté-Agel, F., 2009. A wind-tunnel investigation of wind-turbine wakes: boundary-layer turbulence effects. Bound.-Layer Meteorol. 132, 129–149.

Charakopoulos, A.K., Karakasidis, T.E., Papanicolaou, P.N., Liakopoulos, A., 2014a. The application of complex network time series analysis in turbulent heated jets. Chaos Interdiscip. J. Nonlinear Sci. 24, 024408.

Charakopoulos, A.K., Karakasidis, T.E., Papanicolaou, P.N., Liakopoulos, A., 2014b. Nonlinear time series analysis and clustering for jet axis identification in vertical turbulent heated jets. Phys. Rev. E 89, 032913.

Chen, Y., Sun, R., Zhou, A., 2007. An improved Hurst parameter estimator based on fractional Fourier transform, in: ASME 2007 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. American Society of Mechanical Engineers, pp. 1223–1233.

Chhikara, R.S., Folks, J.L., 1989. The Inverse Gaussian Distribution: Theory. Methodol. Appl. CRC N. Y. 1988.

Chow, V.T., Maidment, D.R., Mays, L.W., 1988. Applied hydrology.

Cohn, T.A., Lins, H.F., 2005. Nature's style: Naturally trendy. Geophys. Res. Lett. 32.

Conradsen, K., Nielsen, L.B., Prahm, L.P., 1984. Review of Weibull statistics for estimation of wind speed distributions. J. Clim. Appl. Meteorol. 23, 1173–1183.

Cordeiro, G.M., de Castro, M., 2011. A new family of generalized distributions. J. Stat. Comput. Simul. 81, 883–898.

Costa, L. da F., 2008. Entropy Moments Characterization of Statistical Distributions. ArXiv Prepr. ArXiv08033348.

Courant, R., Friedrichs, K., Lewy, H., 1959. On the partial difference equations of mathematical physics. CALIFORNIA UNIV LOS ANGELES.

Cunge, J.A., Holly, F.M., Verwey, A., 1980. Practical aspects of computational river hydraulics.

Davidson, P.A., 2000. Was Loitsyansky correct? A review of the arguments. J. Turbul. 1, 006–006.

Deligiannis, I., Dimitriadis, P., Daskalou, O., Dimakos, Y., Koutsoyiannis, D., 2016. Global Investigation of Double Periodicity of Hourly Wind Speed for Stochastic Simulation; Application in Greece. Energy Procedia 97, 278–285.

Diaconis, P., Holmes, S., Montgomery, R., 2007. Dynamical bias in the coin toss. SIAM Rev. 49, 211–235.

Diaconis, P., Keller, J.B., 1989. Fair dice. Am. Math. Mon. 96, 337–339.

Dimitriadis, P., Koutsoyiannis, D., 2017. Stochastic synthesis approximating any process dependence and distribution. Stoch. Evironmental Res. Risk Assess. (accepted).

Dimitriadis, P., Koutsoyiannis, D., 2015a. Climacogram versus autocovariance and power spectrum in stochastic modelling for Markovian and Hurst–Kolmogorov processes. Stoch. Environ. Res. Risk Assess. 29, 1649–1669.

Dimitriadis, P., Koutsoyiannis, D., 2015b. Application of stochastic methods to double cyclostationary processes for hourly wind speed simulation. Energy Procedia 76, 406–411.

Dimitriadis, P., Koutsoyiannis, D., Markonis, Y., 2012. Spectrum vs Climacogram, in: EGU General Assembly Conference Abstracts. p. 993.

Dimitriadis, P., Koutsoyiannis, D., Onof, C., 2013. N-dimensional generalized Hurst-Kolmogorov process and its application to wind fields. Facets Uncertain. 5th EGU Leonardo Conf. – Hydrofractals 2013 – STAHY 2013 Kos Isl. Greece. doi:10.13140/RG.2.2.15642.64963

Dimitriadis, P., Koutsoyiannis, D., Papanicolaou, P., 2016a. Stochastic similarities between the microscale of turbulence and hydro-meteorological processes. Hydrol. Sci. J. 61, 1623–1640.

Dimitriadis, P., Koutsoyiannis, D., Tzouka, K., 2016b. Predictability in dice motion: how does it differ from hydro-meteorological processes? Hydrol. Sci. J. 61, 1611–1622.

Dimitriadis, P., Lazaros, L., Daskalou, O., Filippidou, A., Giannakou, M., Gkova, E., Ioannidis, R., Polydera, A., Polymerou, E., Psarrou, E., others, 2015. Application of stochastic methods for wind speed forecasting and wind turbines design at the area of Thessaly, Greece, in: EGU General Assembly Conference Abstracts. p. 13810.

Dimitriadis, P., Markonis, Y., Iliopoulou, T., Gournari, N., Deligiannis, I., Kastis, P., Nasika, X., Lerias, E., Moustakis, Y., Petsiou, A., others, 2016c. Stochastic similarities between hydroclimatic processes for variability characterization, in: EGU General Assembly Conference Abstracts. p. 15632.

Dimitriadis, P., Papanicolaou, P., 2010. Hurst-Kolmogorov dynamics applied to temperature field of horizontal turbulent buoyant jets, in: EGU General Assembly Conference Abstracts. p. 10644.

Dimitriadis, P., Papanicolaou, P., Koutsoyiannis, D., 2010. Hurst-Kolmogorov dynamics applied to temperature fields for small turbulence scales.

Dimitriadis, P., Papanicolou, P., 2012. Statistical analysis of turbulent positively buoyant jets, in: EGU General Assembly Conference Abstracts. p. 12672.

Dimitriadis, P., Tegos, A., Oikonomou, A., Pagana, V., Koukouvinos, A., Mamassis, N., Koutsoyiannis, D., Efstratiadis, A., 2016d. Comparative evaluation of 1D and quasi-2D hydraulic models based on benchmark and real-world applications for uncertainty assessment in flood mapping. J. Hydrol. 534, 478–492.

Dimitriadis, P., Tzouka, K., Koutsoyiannis, D., Tyralis, H., Kalamioti, A., Lerias, E., Voudouris, P., 2017. Stochastic investigation of long-term persistence in two-dimensional images of rocks. J. Spat. Stat.

Doran, J.C., 2004. Characteristics of intermittent turbulent temperature fluxes in stable conditions. Bound.-Layer Meteorol. 112, 241–255.

Drobinski, P., Coulais, C., 2012. Is the Weibull distribution really suited for wind statistics modeling and wind power evaluation? ArXiv Prepr. ArXiv12113853.

Efstratiadis, A., Dialynas, Y.G., Kozanis, S., Koutsoyiannis, D., 2014. A multivariate stochastic model for the generation of synthetic time series at multiple time scales reproducing long-term persistence. Environ. Model. Softw. 62, 139–152.

Faisst, H., Eckhardt, B., 2004. Sensitive dependence on initial conditions in transition to turbulence in pipe flow. J. Fluid Mech. 504, 343–352.

Falkovich, G., Fouxon, A., Stepanov, M.G., 2002. Acceleration of rain initiation by cloud turbulence. Nature 419, 151–154.

Feller, W., 1971. Law of large numbers for identically distributed variables. Introd. Probab. Theory Its Appl. 2, 231–234.

Ferrier, A.J., Funk, D.R., Roberts, P.J.W., 1993. Application of optical techniques to the study of plumes in stratified fluids. Dyn. Atmospheres Oceans 20, 155–183.

Fleming, S.W., 2008. Approximate record length constraints for experimental identification of dynamical fractals. Ann. Phys. 17, 955–969.

Fourier, J., 1822. Theorie analytique de la chaleur, par M. Fourier. Chez Firmin Didot, père et fils.

Frisch, U., 2006. Turbulence: the legacy of AN Kolmogorov. AIP.

Georgakakos, K.P., Carsteanu, A.A., Sturdevant, P.L., Cramer, J.A., 1994. Observation and analysis of midwestern rain rates. J. Appl. Meteorol. 33, 1433–1444.

Gilgen, H., 2006. Univariate time series in geosciences. Springer.

Gneiting, T., 2000. Power-law correlations, related models for long-range dependence and their simulation. J. Appl. Probab. 37, 1104–1109.

Gneiting, T., Schlather, M., 2004. Stochastic models that separate fractal dimension and the Hurst effect. SIAM Rev. 46, 269–282.

Gneiting, T., Ševčíková, H., Percival, D.B., 2012. Estimators of fractal dimension: Assessing the roughness of time series and spatial data. Stat. Sci. 247–277.

Goldstein, M.L., Roberts, D.A., Matthaeus, W.H., 1995. Magnetohydrodynamic turbulence in the solar wind. Annu. Rev. Astron. Astrophys. 33, 283–325.

Halliwell, L.J., 2013. Classifying the tails of loss distributions, in: Casualty Actuarial Society E-Forum, Spring 2013 Volume 2. p. 1.

Hassani, H., 2010. A note on the sum of the sample autocorrelation function. Phys. Stat. Mech. Its Appl. 389, 1601–1606.

Hassani, H., Leonenko, N., Patterson, K., 2012. The sample autocorrelation function and the detection of long-memory processes. Phys. Stat. Mech. Its Appl. 391, 6367–6379.

Hasson, A.M., Al-Hamadani, N.I., Al-Karaghouli, A.A., 1990. Comparison between measured and calculated diurnal variations of wind speeds in northeast Baghdad. Sol. Wind Technol. 7, 481–487.

Heisenberg, W., 1985. On the theory of statistical and isotropic turbulence, in: Original Scientific Papers Wissenschaftliche Originalarbeiten. Springer, pp. 115–119.

Helland, K.N., Van Atta, C.W., 1978. The 'Hurst phenomenon'in grid turbulence. J. Fluid Mech. 85, 573–589.

Hunter, N.M., Horritt, M.S., Bates, P.D., Wilson, M.D., Werner, M.G., 2005. An adaptive time step solution for raster-based storage cell modelling of floodplain inundation. Adv. Water Resour. 28, 975–991.

Hurst, H.E., 1951. Long-term storage capacity of reservoirs. Trans Amer Soc Civ. Eng 116, 770–808.

Iliopoulou, T., Papalexiou, S.M., Markonis, Y., Koutsoyiannis, D., 2016. Revisiting long-range dependence in annual precipitation. J. Hydrol.

Infante, S., Luna, C., Sánchez, L., Hernández, A., 2016. Approximations of the solutions of a stochastic differential equation using Dirichlet process mixtures and Gaussian mixtures. Stat. Optim. Inf. Comput. 4, 289–307.

Jaynes, E.T., 1957. Information theory and statistical mechanics. Phys. Rev. 106, 620.

Kang, H.S., Chester, S., Meneveau, C., 2003. Decaying turbulence in an active-grid-generated flow and comparisons with large-eddy simulation. J. Fluid Mech. 480, 129–160.

Kapitaniak, M., Strzalko, J., Grabski, J., Kapitaniak, T., 2012. The three-dimensional dynamics of the die throw. Chaos Interdiscip. J. Nonlinear Sci. 22, 047504.

Khintchine, A., 1934. Korrelationstheorie der stationären stochastischen Prozesse. Math. Ann. 109, 604–615.

Klugman, S.A., Panjer, H.H., Willmot, G.E., 2012. Loss models: from data to decisions. John Wiley & Sons.

Kolmogorov, A., 1931. Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung. Math. Ann. 104, 415–458 (English translation: On analytical methods in probability theory, In: Kolmogorov, A.N., 1992. Selected Works of A. N. Kolmogorov).

Kolmogorov, A.N., 1941a. On the degeneration of isotropic turbulence in an incompressible viscous fluid, in: Dokl. Akad. Nauk SSSR. pp. 319–323.

Kolmogorov, A.N., 1941b. Equations of turbulent motion in an incompressible fluid, in: Dokl. Akad. Nauk SSSR. pp. 299–303.

Kolmogorov, A.N., 1941a. Dissipation of energy in locally isotropic turbulence, in: Dokl. Akad. Nauk SSSR. JSTOR, pp. 16–18.

Kolmogorov, A.N., 1941b. The local Structure of turbulence in incompressible viscous fluid for very large Reynolds numbers [In Russian], in: Dokl. Akad. Nauk SSSR. pp. 299–303.

Kolmogorov, A.N., 1940. The Wiener spiral and some other interesting curves in Hilbert space, in: Dokl. Akad. Nauk SSSR. pp. 115–118.

Koudouris, G., Dimitriadis, P., Iliopoulou, T., Mamassis, N., Koutsoyiannis, D., 2017. Investigation of the stochastic nature of solar radiation for renewable resources management. Eur. Geosci. Union Gen. Assem. 2017 Geophys. Res. Abstr. 19.

Koutsoyiannis, D., 2016. Generic and parsimonious stochastic modelling for hydrology and beyond. Hydrol. Sci. J. 61, 225–244.

Koutsoyiannis, D., 2014. Reconciling hydrology with engineering. Hydrol. Res. 45, 2–22.

Koutsoyiannis, D., 2013. Encolpion of stochastics Fundamentals of stochastic processes. Lect. Notes Stoch. Version 5.

Koutsoyiannis, D., 2011. Hurst–Kolmogorov dynamics as a result of extremal entropy production. Phys. Stat. Mech. Its Appl. 390, 1424–1432.

Koutsoyiannis, D., 2010. HESS Opinions" A random walk on water". Hydrol. Earth Syst. Sci. 14, 585–601.

Koutsoyiannis, D., 2005. Uncertainty, entropy, scaling and hydrological stochastics. 1. Marginal distributional properties of hydrological processes and state scaling/Incertitude, entropie, effet d'échelle et propriétés stochastiques hydrologiques. 1. Propriétés distributionnelles marginales des processus hydrologiques et échelle d'état. Hydrol. Sci. J. 50.

Koutsoyiannis, D., 2004a. Statistics of extremes and estimation of extreme rainfall: I, Theoretical Investigation. Hydrol. Sci. J. 49.

Koutsoyiannis, D., 2004b. Statistics of extremes and estimation of extreme rainfall: II. Empirical investigation of long rainfall records. Hydrol. Sci. J. 49.

Koutsoyiannis, D., 2003. Climate change, the Hurst phenomenon, and hydrological statistics. Hydrol. Sci. J. 48, 3–24.

Koutsoyiannis, D., 2002. The Hurst phenomenon and fractional Gaussian noise made easy. Hydrol. Sci. J. 47, 573–595.

Koutsoyiannis, D., 2000. A generalized mathematical framework for stochastic simulation and forecast of hydrologic time series. Water Resour. Res. 36, 1519–1533.

Koutsoyiannis, D., Dimitriadis, P., 2016. From time series to stochastics: A theoretical framework with applications on time scales spanning from microseconds to megayears, in: Orlob Second International Symposium on Theoretical Hydrology, University of California Davis.

Koutsoyiannis, D., Dimitriadis, P., Lombardo, F., Stevens, S., 2017. From fractals to stochastics: Seeking theoretical consistency in analysis of geophysical data. Adv. Nonlinear Geosci. (accepted).

Koutsoyiannis, D., Langousis, A., 2011. Precipitation, Treatise on Water Science, edited by P. Wilderer and S. Uhlenbrook, 2, 27–78. Academic Press, Oxford.

Koutsoyiannis, D., Montanari, A., 2015. Negligent killing of scientific concepts: the stationarity case. Hydrol. Sci. J. 60, 1174–1183.

Koutsoyiannis, D., Yao, H., Georgakakos, A., 2008. Medium-range flow prediction for the Nile: a comparison of stochastic and deterministic methods/Prévision du débit du Nil à moyen terme: une comparaison de méthodes stochastiques et déterministes. Hydrol. Sci. J. 53, 142–164.

Kraichnan, R.H., 1991. Stochastic modeling of isotropic turbulence, in: New Perspectives in Turbulence. Springer, pp. 1–54.

Kraichnan, R.H., 1959. The structure of isotropic turbulence at very high Reynolds numbers. J. Fluid Mech. 5, 497–543.

Krajewski, W.F., Kruger, A., Nespor, V., 1998. Experimental and numerical studies of small-scale rainfall measurements and variability. Water Sci. Technol. 37, 131–138.

Kuik, D.J., Poelma, C., Westerweel, J., 2010. Quantitative measurement of the lifetime of localized turbulence in pipe flow. J. Fluid Mech. 645, 529–539.

Labby, Z., 2009. Weldon's dice, automated. Chance 22, 6–13.

Langousis, A., Veneziano, D., 2009. Long-term rainfall risk from tropical cyclones in coastal areas. Water Resour. Res. 45.

Laskar, J., 1999. The limits of Earth orbital calculations for geological time-scale use. Philos. Trans. R. Soc. Lond. Math. Phys. Eng. Sci. 357, 1735–1759.

Lavergnat, J., 2016. On the generation of colored non-Gaussian time sequences.

Lerias, E., Kalamioti, A., Dimitriadis, P., Markonis, Y., Iliopoulou, T., Koutsoyiannis, D., 2016. Stochastic investigation of temperature process for climatic variability identification, in: EGU General Assembly Conference Abstracts. p. 14828.

Lombardo, F., Volpi, E., Koutsoyiannis, D., 2013. Effect of time discretization and finite record length on continuous-time stochastic properties. IAHS-IAPSO-IASPEI Joint Assembly.

Lombardo, F., Volpi, E., Koutsoyiannis, D., Papalexiou, S.M., 2014. Just two moments! A cautionary note against use of high-order moments in multifractal models in hydrology. Hydrol. Earth Syst. Sci. 18, 243–255.

Lorenz, E.N., 1963. Deterministic nonperiodic flow. J. Atmospheric Sci. 20, 130–141.

Mahrt, L., 1989. Intermittency of atmospheric turbulence. J. Atmospheric Sci. 46, 79–95.

Mandelbrot, B. B. (1971), A Fast Fractional Gaussian Noise Generator, Water Resour. Res., 7(3), 543–553.

Mandelbrot, B.B., Wallis, J.R., 1969. Some long-run properties of geophysical records. Water Resour. Res. 5, 321–340.

Mandelbrot, B.B., Wallis, J.R., 1968. Noah, Joseph, and operational hydrology. Water Resour. Res. 909–918.

Manwell, J.F., McGowan, J.G., Rogers, A.L., 2010. Wind energy explained: theory, design and application. John Wiley & Sons.

Markonis, Y., Koutsoyiannis, D., 2013. Climatic variability over time scales spanning nine orders of magnitude: Connecting Milankovitch cycles with Hurst–Kolmogorov dynamics. Surv. Geophys. 34, 181–207.

McDonough, J.M., 2004. Introductory lectures on turbulence physics, mathematics and modeling.

Michas, S.N., Papanicolaou, P.N., 2009. Horizontal round heated jets into calm uniform ambient. Desalination 248, 803–815.

Monahan, A.H., 2013. The Gaussian statistical predictability of wind speeds. J. Clim. 26, 5563–5577.

Moschos, E., Manou, G., Georganta, C., Dimitriadis, P., Iliopoulou, T., Tyralis, H., Koutsoyiannis, D., Tsoukala, V., 2017. Investigation of the stochastic nature of wave processes for renewable resources management: a pilot application in a remote island in the Aegean sea. Eur. Geosci. Union Gen. Assem. 2017 Geophys. Res. Abstr. 19.

Nagler, J., Richter, P., 2008. How random is dice tossing? Phys. Rev. E 78, 036207.

Nordin, C.F., McQuivey, R.S., Mejia, J.M., 1972. Hurst phenomenon in turbulence. Water Resour. Res. 8, 1480–1486.

O'brien, J.S., 2007. FLO-2D users manual. Nutr. Ariz. June.

O'Connell, P.E., Koutsoyiannis, D., Lins, H.F., Markonis, Y., Montanari, A., Cohn, T., 2016. The scientific legacy of Harold Edwin Hurst (1880–1978). Hydrol. Sci. J. 61, 1571–1590.

Papanicolaou, P.N., List, E.J., 1987. Statistical and spectral properties of tracer concentration in round buoyant jets. Int. J. Heat Mass Transf. 30, 2059–2071.

Papanicolaou, P.N., List, E. J., 1988. Investigations of round vertical turbulent buoyant jets. J. Fluid Mech., 195:341–391.

Papoulis, A., 1990. Probability & statistics. Prentice-Hall Englewood Cliffs.

Papoulis, A., Pillai, S.U., 1991. Stochastic processes. McGraw-Hill New York.

Pearson, K., 1930. On a new theory of progressive evolution. Ann. Hum. Genet. 4, 1–40.

Pedretti, C., 1977. The literary works of Leonardo da Vinci. Univ of California Press.

Poincaré, H., 1890. Sur le probleme des trois corps et les équations de la dynamique. Acta Math. 13, A3–A270.

Pope, S.B., 2001. Turbulent flows. IOP Publishing.

Pope, S.B., 2000. Turbulent Flows, 771 pp. Cambridge Univ. Press, Cambridge, UK.

Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 2007. Numerical recipes: the art of scientific computing. Cambridge University Press, New York.

Richardson, L.F., 1922. Weather prediction by numerical methods. Cambridge University Press, London. experiment which prompted the community model idea, it is in no way sacrosanct if better ideas come forward. A model integration could be carried cut with.

Rinaldo, A., 2006. W06D01-Introduction to special issue on Rain, Rivers, and Turbulence: A view from hydrology (DOI 10.1029/2006WR004945). Water Resour. Res. 1.

Saffman, P.G., 1967. The large-scale structure of homogeneous turbulence. J. Fluid Mech. 27, 581–593.

Sakalauskienė, G., 2003. The Hurst phenomenon in hydrology. Environ. Res. Eng. Manag. 3, 16–20.

Sevruk, B., Nespor, V., 1998. Empirical and theoretical assessment of the wind induced error of rain measurement. Water Sci. Technol. 37, 171–178.

Shannon, C.E., 1948. A note on the concept of entropy. Bell Syst. Tech J 27, 379–423.

She, Z.-S., Leveque, E., 1994. Universal scaling laws in fully developed turbulence. Phys. Rev. Lett. 72, 336.

Shuaib, K.M., Robert, K., Lena, H.I., 2016. Transmuted Kumaraswamy Distribution. Stat. Transit. New Ser. 17, 183–210.

Singh, S.K., Maddala, G.S., 1978. A function for size distribution of incomes: reply. Econom. Pre-1986 46, 461.

Sotiriadou, A., Petsiou, A., Feloni, E., Kastis, P., Iliopoulou, T., Markonis, Y., Tyralis, H., Dimitriadis, P., Koutsoyiannis, D., 2016. Stochastic investigation of precipitation process for climatic variability identification, in: EGU General Assembly Conference Abstracts. p. 15137.

Stoica, P., Moses, R.L., 2005. Spectral analysis of signals. Pearson Prentice Hall Upper Saddle River, NJ.

Strzalko, J., Grabski, J., Stefanski, A., Kapitaniak, T., 2010. Can the dice be fair by dynamics? Int. J. Bifurc. Chaos 20, 1175–1184.

Taylor, G.I., 1938. The spectrum of turbulence, in: Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences. The Royal Society, pp. 476–490.

Taylor, G.I., 1935. Statistical theory of turbulence, in: Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences. The Royal Society, pp. 421–444.

Tessarotto, M., Asci, C., 2010. On the behavior of homogeneous, isotropic and stationary turbulence. ArXiv Prepr. ArXiv10031475.

Tsekouras, G., Koutsoyiannis, D., 2014. Stochastic analysis and simulation of hydrometeorological processes associated with wind and solar energy. Renew. Energy 63, 624–633.

Tyralis, H., Dimitriadis, P., Iliopoulou, T., Tzouka, K., Koutsoyiannis, D., 2017. Dependence of long-term persistence properties of precipitation on spatial and regional characteristics, in: EGU General Assembly Conference Abstracts. p. 3711.

Tyralis, H., Koutsoyiannis, D., 2011. Simultaneous estimation of the parameters of the Hurst–Kolmogorov stochastic process. Stoch. Environ. Res. Risk Assess. 25, 21–33.

Tyralis, H., Koutsoyiannis, D., Kozanis, S., 2013. An algorithm to construct Monte Carlo confidence intervals for an arbitrary function of probability distribution parameters. Comput. Stat. 28, 1501–1527.

Vasilopoulou, E., 2003. The child and games in ancient Greek art, Graduating thesis, Aristotle University of Thessaloniki.

Veneziano, D., Langousis, A., Furcolo, P., 2006. Multifractality and rainfall extremes: A review. Water Resour. Res. 42.

Villarini, G., Mandapaka, P.V., Krajewski, W.F., Moore, R.J., 2008. Rainfall and sampling uncertainties: A rain gauge perspective. J. Geophys. Res. Atmospheres 113.

Von Karman, T., 1948. Progress in the statistical theory of turbulence. Proc. Natl. Acad. Sci. 34, 530–539.

Wackernagel, H., 1995. Multivariate geostatistics. Introd. Appl. 235.

Wiener, N., 1930. Generalized harmonic analysis. Acta Math. 55, 117–258.

Wilczek, M., Daitche, A., Friedrich, R., 2011. On the velocity distribution in homogeneous isotropic turbulence: correlations and deviations from Gaussianity. J. Fluid Mech. 676, 191–217.

Wright, J.R., Cooper, J.E., 2008. Introduction to aircraft aeroelasticity and loads. John Wiley & Sons.

Yaglom, A.M., 2004. An introduction to the theory of stationary random functions. Courier Corporation.

Yari, G.-H., Borzadaran, G.M., 2010. Entropy for Pareto-types and its order statistics distributions. Commun. Inf. Syst. 10, 193–202.

Appendix A

In this Appendix, we investigate and compare the climacogram, autocovariance and power spectrum of the Markov process and gHK one for $M = 0.5$ in terms of their behaviour and of their estimator performance for different values of their parameters (Dimitriadis and Koutsoyiannis, 2015a). The methodology we use to produce synthetic time series is through the SAR scheme (see in section 3.2).

Graphical investigation

We start our comparison with graphical investigations, which are actually very common in model identification. We compare the true, continuous-time stochastic tools, along with their discrete-time versions as well as their expectation of classical estimators. For the estimator, a medium sample size $n = 10^3$ is used (apparently, as n increases the bias will decrease).

In particular, we investigate the climacogram, autocovariance and power spectrum for a Markov processes with $q = 1$, 10 and 100, and $\lambda = 1$ (Figure A-1).

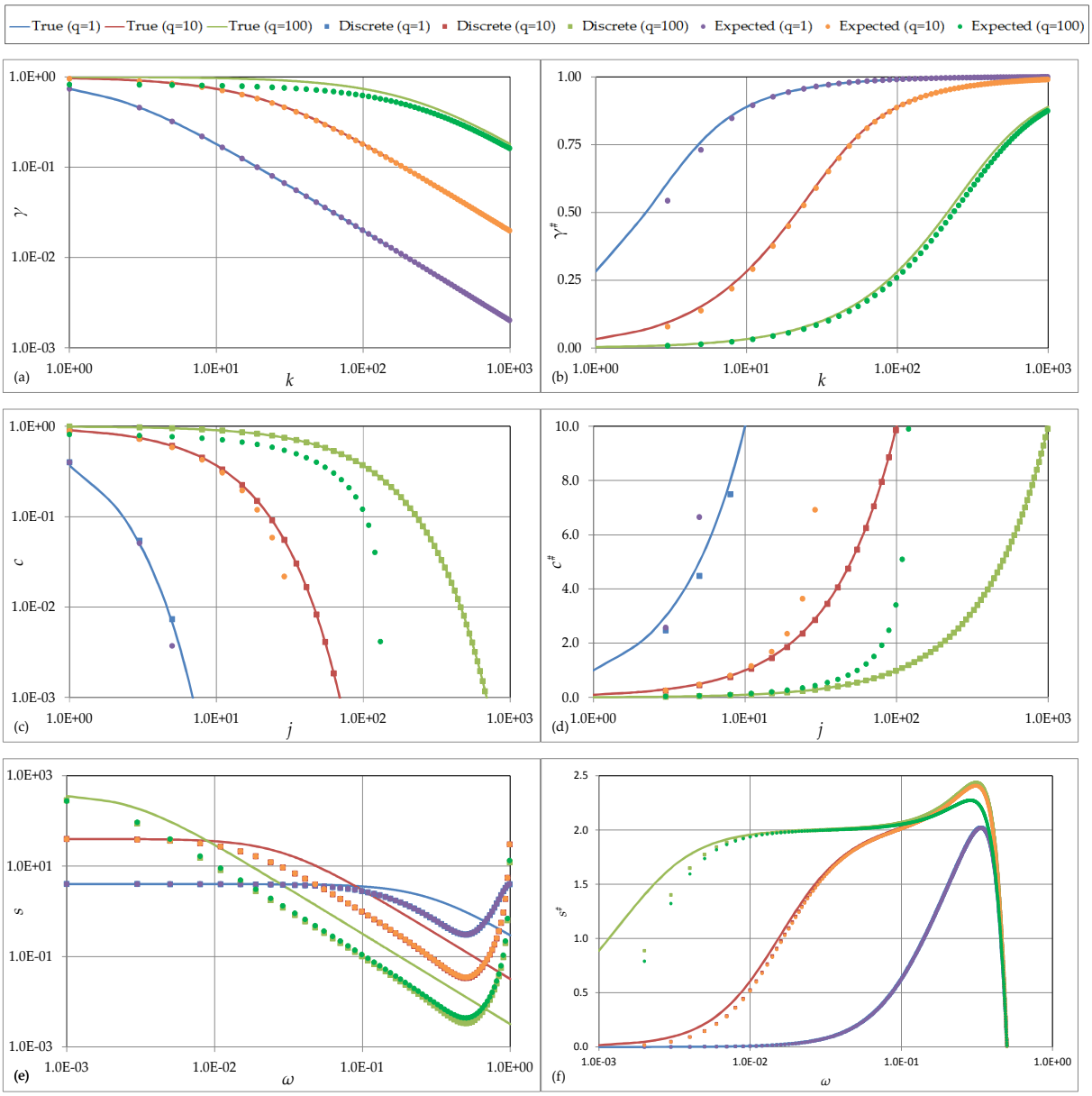

Figure A-1: True values in continuous and discrete time and expected values of the climacograms (a), autocovariances (c) and power spectra (e) as well as their corresponding NLDs (b, d and f, respectively) of Markov processes with q = 1, 10 and 100, λ = 1 and n = 10^3. Note that the continuous and discrete values of the climacogram are identical for $\Delta = D > 0$.

Additionally, we investigate the climacogram, autocovariance and power spectrum for a gHK processes with q = 1, 10 and 100, b = 0.2 and $\lambda = q^{-b}$, all with $D = \Delta = 1$ (Figure A-2).

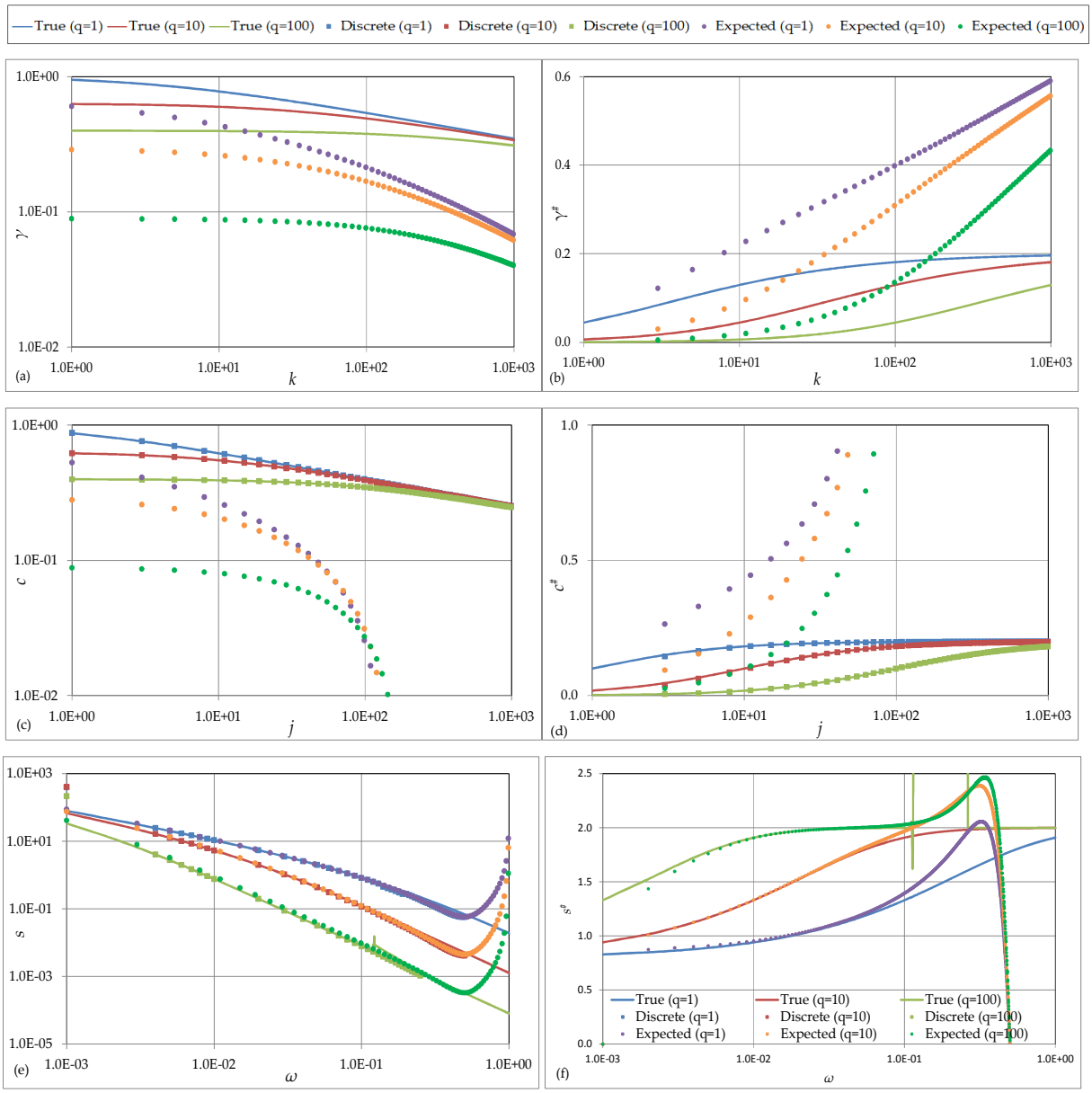

Figure A-2: True values in continuous and discrete time and expected values of the climacograms (a), autocovariances (c) and power spectra (e) as well as their corresponding NLDs (b, d and f, respectively) of gHK processes with b = 0.2 and q = 1, 10 and 100, $\lambda = q^{-b}$ (not λ = 1, for demonstration purposes) and n=10^3. Note that the continuous and discrete values of the climacogram are identical for $\Delta = D > 0$.

Comparison of statistical estimators

Thus, we produce synthetic time series for Markov processes with q = 1, 10 and 100 and gHK ones with q = 1, 10 and 100 and b = 0.2, all with $D = \Delta = 1$. Then, for each scale, lag and frequency and each synthetic timeseries, we calculate the mean, variance, mean of the NLD, and variance of the NLD, for the climacogram, autocovariance and power spectrum, as well as their corresponding errors (Figure A-3). Note that, on one hand, as n decreases, both bias and variance increase and

thus, for the point estimate and variance to be closer to the expected ones, we need more time series. On the other hand, as n increases, more Markov processes have to be added and with a larger bias and variance (due to larger q). So, for the examined processes, we conclude that in order to achieve a maximum error of about 1‰ between scales 1 and $n/2$, we have to produce approximate 10^4, 10^3 and 10^2 timeseries for n = 10^2, 10^3 and 10^4, respectively. The error is calculated as the absolute difference between the estimated and expected value, and divided by the expected value. Furthermore, the 1‰ error refers to the climacogram and corresponds to a gHK process with b = 0.2 and q = 100, which is considered the more adverse of the examined processes. Note that in the Figures below, we try to show all estimates within a single plot for comparison to each other. The inverse frequency in the horizontal axis is set to $1/(2\omega)$, in order to vary between 1 and $n/2$ and the lag to υ+1 and for the estimation of variance at υ = 0 to be also included in a log-log plot.

Moreover, we investigate the shape of the probability distribution density function for each stochastic tool, which, in many cases, differs from a Gaussian one, resulting in deviations between the mean (expected) and mode (Figure A-4). To measure this difference, we use the sample skewness (denoted g), where for $g \approx 0$, the difference is small and for any other case, larger. We show for each stochastic tool and for a gHK process with b = 0.2 and q/Δ = 10, an example of their 95% upper and lower prediction intervals (corresponding to exceedence probabilities of 2.5% and 97.5%), as well as their pdf for a specific scale, lag and frequency.

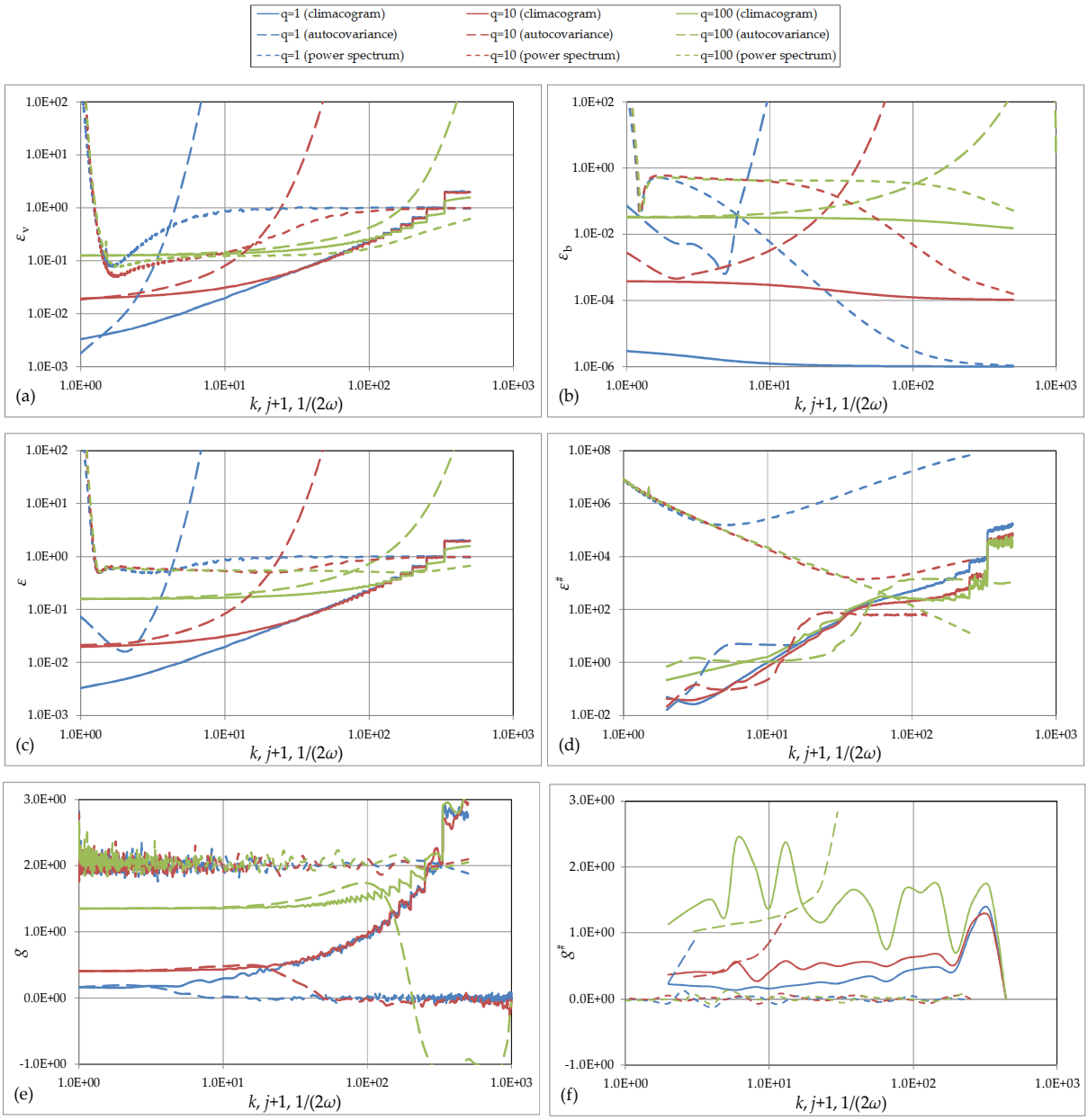

Figure A-3: Dimensionless errors of the climacogram estimator (continuous line), autocovariance (dashed line) and power spectrum (dotted line), calculated from 10^4 Markov synthetic series with $n = 10^3$ (for $b = 0.2$, $q = 1$, 10 and 100 and $\lambda = q^{-b}$): (a) ε_v (dimensionless MSE of variance); (b) ε_b (dimensionless MSE of bias); (c) ε (total dimensionless MSE); and (d) $\varepsilon^{\#}$ (total dimensionless MSE of NLD); as well as the sample skewness of each of the stochastic tools and their NLDs are also shown (e) and (f).

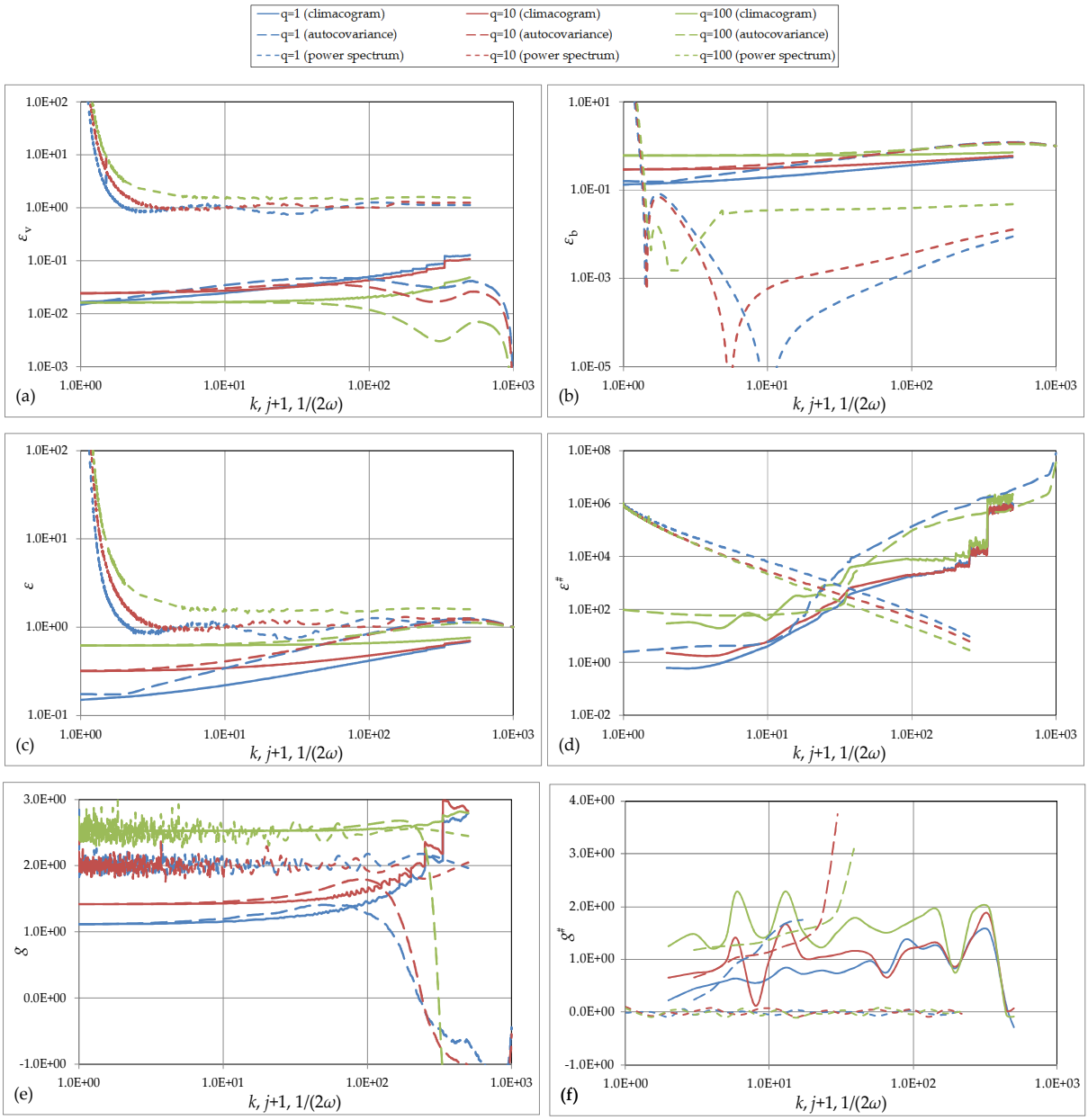

Figure A-4: Dimensionless errors of the climacogram estimator (continuous line), autocovariance (dashed line) and power spectrum (dotted line), calculated from 10^4 gHK synthetic series with n = 10^3 (for b = 0.2, q = 1, 10 and 100 and $\lambda = q^{-b}$): (a) ε_v (dimensionless MSE of variance); (b) ε_b (dimensionless MSE of bias); (c) ε (total dimensionless MSE); and (d) $\varepsilon^{\#}$ (total dimensionless MSE of NLD); as well as the sample skewness of each of the stochastic tools and their NLDs are also shown in (e) and (f).

Appendix B

Here, we estimate several statistical characteristics of the ESK and NIG distributions such as the mean, variance, and coefficients of skewness and kurtosis, as well as the minimum and maximum kurtosis as a function of skewness.

For random number generation from thin-tailed distributions we adopt an extended standardized version of the Kumaraswamy (1980) distribution (abbreviated as ESK) with probability distribution function:

$$F(x;\boldsymbol{p}) := 1 - \left(1 - \left(\frac{x-c}{d}\right)^{a}\right)^{b} \tag{B-1}$$

where $x \in [c, c+d]$, $\boldsymbol{p} = [a, b, c, d]$, the parameters of the distribution (see also Table C-1 and C-2), with $c, d \in \mathbb{R}$ (location and scale parameters, respectively, with units same as in x) and $a, b > 0$ (dimensionless shape parameters).

Below, we estimate several statistical characteristics of the ESK distribution such as the mean, variance, and coefficients of skewness and kurtosis, as well as the minimum and maximum kurtosis as a function of skewness. A detailed analysis on the general expansion of the Kumaraswamy distribution can be found in Cordeiro and de Castro (2011), and Shuaib et al. (2016). The ESK distribution has simple, analytical and closed expressions for its statistical central moments. Notably, we find through numerical investigation that ESK has a low kurtosis boundary based on its skewness and approximately expressed by $C_{\mathrm{k}} \geq {C_{\mathrm{s}}}^2 + 1$, which is also the mathematical boundary for the sample skewness and kurtosis (Pearson, 1930).

The central moments of the ESK distribution can be expressed as (Dimitriadis and Koutsoyiannis, 2017):

$$\mathrm{E}\left[(\underline{x} - \mu)^p\right] = d^p \sum_{\xi=1}^{p+1} \left((-1)^{p+1-\xi} \binom{p}{\xi - 1} B_1^{p+1-\xi} B_{\xi-1}\right) \tag{B-2}$$

for $p > 1$ and where $\mu = c + dB_1$, $\binom{p}{\xi-1}$ the binomial coefficient and $B_\xi = b\mathrm{B}(1 + \xi/a, b)$, with B the beta function.

Thus, the variation, skewness and kurtosis coefficients can be expressed as (Dimitriadis and Koutsoyiannis, 2017):

$$C_{\mathrm{v}} = \frac{B_2 - {B_1}^2}{(B_1 + c/d)^2}, C_{\mathrm{s}} = \frac{2{B_1}^3 - 3B_1B_2 + B_3}{\left(B_2 - {B_1}^2\right)^{3/2}}, C_{\mathrm{k}} = \frac{-3{B_1}^4 + 6{B_1}^2B_2 - 4B_1B_3 + B_4}{\left(B_2 - {B_1}^2\right)^2} \tag{B-3}$$

respectively. After the numerical estimation of a and b, the parameters c and d can be analytically calculated as (Dimitriadis and Koutsoyiannis, 2017):

$$d = \sigma / \sqrt{b\mathrm{B}\left(1+\frac{2}{a}, b\right) - b^2\mathrm{B}^2\left(1+\frac{1}{a}, b\right)},\ c = \mu - bd\mathrm{B}\left(1+\frac{1}{a}, b\right) \tag{B-4}$$

Therefore, we can use the ESK distribution to approximate a variety of thin-tailed distributions based on the estimation of a, b, c and d parameters from data.

For heavy-tailed distributions we use the Normal-Inverse-Gaussian (abbreviated as NIG) distribution with probability density function (cf., Barndorff-Nielsen, 1978):

$$f(x;\boldsymbol{p}) \coloneqq \frac{\sqrt{a^2+b^2}\mathrm{e}^{b+\frac{a(x-c)}{d}}}{\pi d\sqrt{1+\left(\frac{(x-c)}{d}\right)^2}} K_1\left(\sqrt{a^2+b^2}\sqrt{1+\left((x-c)/d\right)^2}\right) \tag{B-5}$$

where $x \in \mathbb{R}$, $\boldsymbol{p} = [a, b, c, d]$, the parameters of the distribution with $c \in \mathbb{R}$, $a \neq 0$ and $b, d > 0$ (see also Table C-1 and C-2); again c, d are location and scale parameters, respectively, with units same as in x, and $a, b > 0$ are dimensionless shape parameters.

The NIG distribution has similar advantages to the ESK, such as closed expressions for the first four central moments. Also, it enables a large variety of skewness-kurtosis combinations and its random numbers can be generated almost as fast as the ESK ones through the normal variance-mean mixture:

$$x = c + \frac{a}{d}z + \sqrt{z}g \tag{B-6}$$

where

$$g \sim N(0,1),\ z \sim f(y; b/d\,, d) = d/\sqrt{2\pi y^3}\,\mathrm{e}^{-\frac{b^2(y/d-d/b)^2}{2y}} \tag{B-7}$$

The latter is the Inverse Gaussian distribution which can be easily generated (e.g., Chhikara and Folks, 1989, ch. 4.5).

Below, we estimate the statistical characteristics of the NIG and we justify the use of the NIG distribution as a heavy-tailed distribution. Note that the central moments of the NIG function cannot be expressed as closed and analytical forms and thus, we can estimate them through the NIG characteristic function (cf., Barndorff-Nielsen, 1978):

$$\varphi_X(t) = \mathrm{E}[\mathrm{e}^{itX}] = \mathrm{e}^{ict+b-\sqrt{\left(\frac{b}{d}\right)^2 - i\frac{2a}{d}t - it^2}} \tag{B-8}$$

where the p^{th} raw moment corresponds to

$$\mathrm{E}[X^p] = (-i)^p \lim_{t\to 0}\left(\frac{\mathrm{d}^p\varphi_X(t)}{\mathrm{d}t^p}\right) \tag{B-9}$$

Particularly, the first moment and the sequent three central moments are given by:

$$\mu = c + ad/b \tag{B-10}$$

$$\mathrm{E}\left[\left(\underline{x} - \mu\right)^2\right] = (a^2 + b^2)d^2/b^3 \tag{B-11}$$

$$\mathrm{E}\left[\left(\underline{x} - \mu\right)^3\right] = \frac{3a\left((a^2 + b^2)d^2/b^3\right)^{3/2}}{\sqrt{b(a^2 + b^2)}} \tag{B-12}$$

$$\mathrm{E}\left[\left(\underline{x} - \mu\right)^4\right] = \frac{3\left((a^2 + b^2)d^2/b^3\right)^2}{b}\left(1 + \frac{4}{1 + (b/a)^2}\right) + 3\left((a^2 + b^2)d^2/b^3\right)^2 \tag{B-13}$$

After algebraic manipulations the coefficients of variation, skewness and kurtosis can be expressed as (Dimitriadis and Koutsoyiannis, 2017):

$$C_\mathrm{v} = \frac{a^2+b^2}{b(a+bc/d)^2}, C_\mathrm{s} = \frac{3a}{\sqrt{b(a^2+b^2)}}, C_\mathrm{k} = \frac{3}{b}\left(1 + \frac{4}{1+(b/a)^2}\right) + 3 \tag{B-14}$$

respectively. The NIG parameters can then be calculated from these equations as:

$$d = \frac{3\sigma\sqrt{3C_\mathrm{k} - 5{C_\mathrm{s}}^2 - 9}}{3C_\mathrm{k} - 4{C_\mathrm{s}}^2 - 9}, b = \frac{d}{\sigma}\sqrt{\frac{3}{C_\mathrm{k} - \frac{5}{3}{C_\mathrm{s}}^2 - 3}}, a = \frac{b^2 C_\mathrm{s}\sigma}{3d},\ c = \mu - ad/b \tag{B-15}$$

Also, we can derive theoretically the maximum kurtosis of NIG for a given skewness:

$$C_\mathrm{k} \geq \frac{5}{3}{C_\mathrm{s}}^2 + 3 \tag{B-16}$$

For the classification of tails we use the test based on the functions proposed by (Klugman et al. 2012, sect. 3.4.3; see also Halliwell, 2013) and here defined as:

$$\tau_\mathrm{r} \coloneqq -\lim_{x\to\infty}\left(\frac{\mathrm{d}f(x;\boldsymbol{p})}{f(x;\boldsymbol{p})\mathrm{d}x}\right), \tau_\mathrm{l} \coloneqq \lim_{x\to-\infty}\left(\frac{\mathrm{d}f(x;\boldsymbol{p})}{f(x;\boldsymbol{p})\mathrm{d}x}\right) \tag{B-17}$$

After calculations we get:

$$\tau_\mathrm{r} = \sqrt{a^2 + b^2}/d - a/d \geq 0,\ \tau_\mathrm{l} = \sqrt{a^2 + b^2}/d + a/d \geq 0 \tag{B-18}$$

and hence the NIG is expected to represent a large variety of heavy-tailed distributions.

In Fig. B-1 and B-2, we observe that the smaller possible kurtosis of the ESK distribution for a given skewness coincides with the theoretical limit defined by Pearson (1930). Also, the larger kurtosis of the ESK includes a variety of sub-Gaussian and thin-tailed distributions. On the contrary, the

smaller kurtosis of the NIG distribution is very close to the larger one of the ESK and thus, it can include a variety of heavy-tailed distributions.

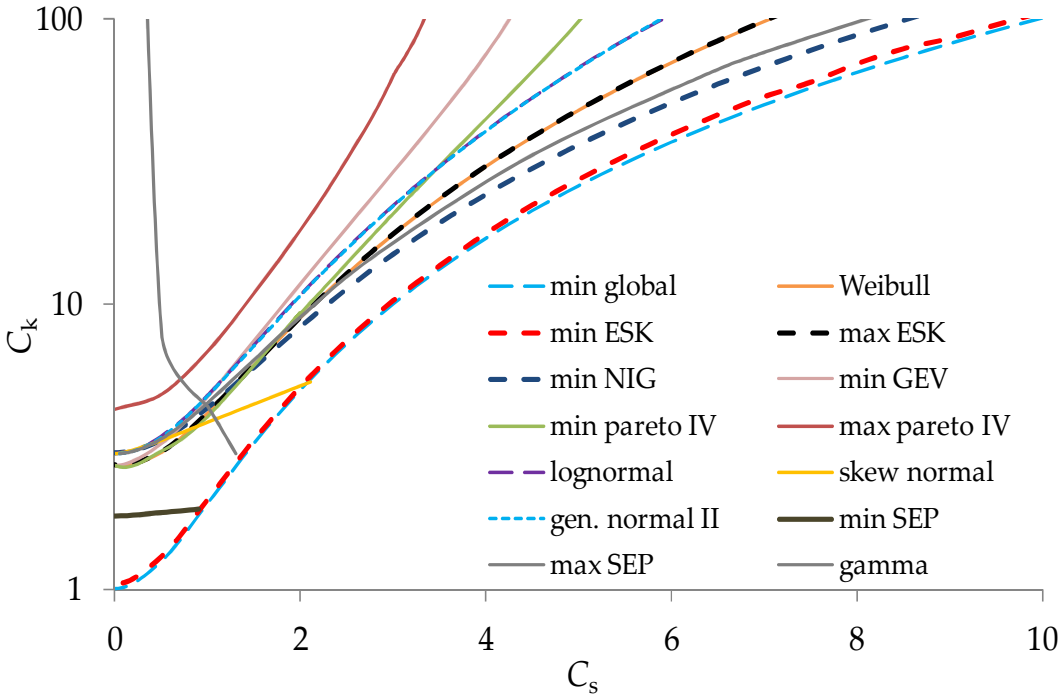

Figure B-1: Combinations of skewness and kurtosis coefficients for various two-parameter (Weibull, GEV, lognormal, generalized normal I, skew-exponential-power —SEP— and gamma), three-parameter (generalized normal II and skew normal) and the four-parameter Pareto-Burr-Fuller (PBF, further described in section 4) distribution functions along with the thin-heavy tailed separation based on the ESK and NIG functions, respectively. Source: Dimitriadis and Koutsoyiannis (2017).

Table B-1: Mean, variance, and coefficients of skewness and kurtosis for the ESK and NIG distributions. Note that $B_i = b\mathrm{B}(1 + i/a, b)$, where $\mathrm{B}(x, y)$ is the beta function and i an integer. Source: Dimitriadis and Koutsoyiannis (2017).

	ESK	NIG
μ	$c + dB_1$	$c + ad/b$
σ^2	$d^2\left(B_2 - {B_1}^2\right)$	$\frac{(a^2 + b^2)d^2}{b^3}$
C_s	$\frac{2{B_1}^3 - 3B_1B_2 + B_3}{\left(B_2 - {B_1}^2\right)^{3/2}}$	$\frac{3a}{\sqrt{b(a^2 + b^2)}}$
C_k	$\frac{-3{B_1}^4 + 6{B_1}^2B_2 - 4B_1B_3 + B_4}{\left(B_2 - {B_1}^2\right)^2}$	$\frac{3}{b}\left(1 + \frac{4}{1 + (b/a)^2}\right) + 3$
min C_k	$\approx {C_\mathrm{s}}^2 + 1$	$= \frac{5}{3}{C_\mathrm{s}}^2 + 3$
max C_k	$\approx \frac{5}{3}{C_\mathrm{s}}^2 + 3$ *	$+\infty$

* This is a fair approximation only for $C_\mathrm{s} \leq -2$. A more exact but empirical approximation for $-10 \leq C_\mathrm{s} \leq 10$, can be given by: $0.039{C_\mathrm{s}}^3 + 1.724{C_\mathrm{s}}^3 + 0.032{C_\mathrm{s}}^3 + 2.7$. Note that the max kurtosis for the ESK for a given skewness coincides with the kurtosis of the Weibull distribution (Fig. B-1).

Table B-2: Parameters of the ESK and NIG distributions in terms of the mean, standard deviation, and coefficients of skewness and kurtosis (see also Fig. B-2). Source: Dimitriadis and Koutsoyiannis (2017).

distribution	ESK4	NIG
a	non-analytical *	$\frac{b^2C_\mathrm{s}\sigma}{3d}$
b	non-analytical *	$\frac{d\sqrt{3}}{\sigma\sqrt{C_\mathrm{k} - \frac{5}{3}{C_\mathrm{s}}^2 - 3}}$
c	$\mu - dB_1$	$\mu - ad/b$
d	$\frac{\sigma}{\sqrt{\left(B_2 - {B_1}^2\right)}}$	$\frac{3\sigma\sqrt{3C_\mathrm{k} - 5{C_\mathrm{s}}^2 - 9}}{3C_\mathrm{k} - 4{C_\mathrm{s}}^2 - 9}$

* The two parameters of the ESK distribution a and b can be found by solving numerically the equations: $C_\mathrm{s} = \left(2{B_1}^3 - 3B_1B_2 + B_3\right)/\left(B_2 - {B_1}^2\right)^{3/2}$, $C_\mathrm{k} = \left(-3{B_1}^4 + 6{B_1}^2B_2 - 4B_1B_3 + B_4\right)/\left(B_2 - {B_1}^2\right)^2$.

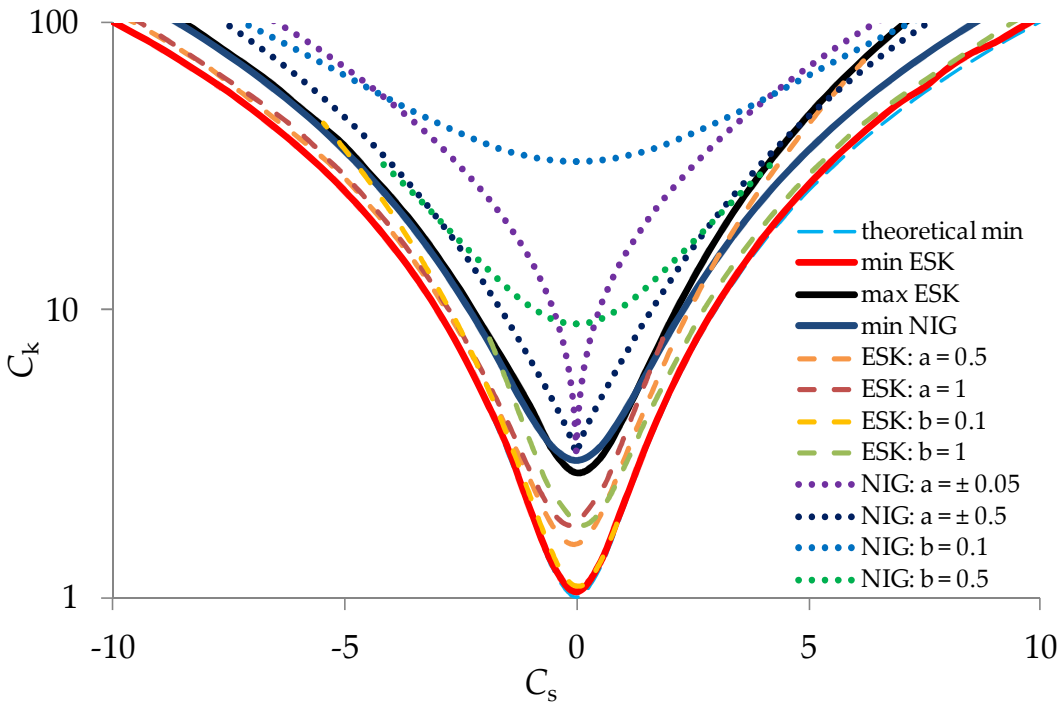

Figure B-2: Isopleths for estimated coefficients of skewness and kurtosis for the specified values of parameters *a* and *b* of the ESK and NIG distributions. Source: Dimitriadis and Koutsoyiannis (2017).

Appendix C

Here, we describe how the SMA scheme can preserve an approximation of the marginal distribution of a process through the preservation of its first four moments. Although this scheme can be extended to preserve any number of moments, here we present the solution for preservation up to the fourth moment corresponding to kurtosis. The p^{th} raw moment that coincides with the corresponding central moment for $\mathrm{E}[\underline{v}] = 0$, can be expressed through the SMA scheme as (Dimitriadis and Koutsoyiannis, 2017):

$$\mathrm{E}[\underline{x}_i{}^p] = \mathrm{E}\left[\left(\sum_{j=-l}^{l} a_{|j|}\underline{v}_{i+j}\right)^p\right] \tag{C-1}$$

Therefore, assuming that $\mathrm{E}[\underline{v}^2] = 1$, the second and third raw moments can be expressed as (Koutsoyiannis, 2000):

$$\mathrm{E}[\underline{x}^2] = \left(a_0{}^2 + 2\sum_{j=1}^{l} a_j{}^2\right) \tag{C-2}$$

$$\mathrm{E}\left[\underline{x}^3\right] = \left({a_0}^3 + 2\sum_{j=1}^{l} {a_j}^3\right)\mathrm{E}\left[\underline{v}^3\right] \tag{C-3}$$

For the fourth raw moment ($p = 4$) we use the multinomial theorem:

$$\mathrm{E}\left[{\underline{x}_i}^4\right] = \mathrm{E}\left[\left(\sum_{j=-l}^{l} a_{|j|}\underline{v}_{i+j}\right)^4\right] = \sum_{k_{-l}+k_{1-l}+\cdots+k_l=4} \binom{4}{k_{-l}, k_{1-l}, \ldots, k_l} \mathrm{E}\left[\prod_{-l\le j\le l} \left(a_{|j|}\underline{v}_{i+j}\right)^{k_j}\right] \tag{C-4}$$

where the multinomial coefficient can be expressed as:

$$\binom{4}{k_{-l}, k_{1-l}, \ldots, k_l} = \frac{4!}{k_{-l}!\, k_{1-l}! \ldots k_l!} \tag{C-5}$$

We notice that all combinations with $k_j = 1$ are zero and thus, after algebraic manipulations we obtain:

$$\mathrm{E}\left[\underline{x}^4\right] = \mathrm{E}\left[\underline{v}^4\right]\left({a_0}^4 + 2\sum_{j=1}^{l} {a_j}^4\right) + \sum_{j=-l}^{l}\sum_{k=-l}^{l} a_{|j|}^2\, a_{|k|}^2 \tag{C-6}$$

Thus, the skewness and kurtosis coefficients can be estimated as (Dimitriadis and Koutsoyiannis, 2017):

$$C_{\mathrm{s},\underline{x}} = C_{\mathrm{s},\underline{v}} \frac{\left({a_0}^3 + 2\sum_{j=1}^{l} {a_j}^3\right)}{\left({a_0}^2 + 2\sum_{j=1}^{l} {a_j}^2\right)^{3/2}} \tag{C-7}$$

$$C_{\mathrm{k},\underline{x}} = \frac{C_{\mathrm{k},\underline{v}}\left({a_0}^4 + 2\sum_{j=1}^{l} {a_j}^4\right) + 6\sum_{j=1}^{l} {a_j}^4 + 12{a_0}^2\sum_{j=1}^{l} {a_j}^2 + 24\sum_{j=i+1}^{l}\sum_{i=1}^{j} a_j^2\, a_k^2}{\left({a_0}^2 + 2\sum_{j=1}^{l} {a_j}^2\right)^2} \tag{C-8}$$