NATIONAL TECHNICAL UNIVERSITY OF ATHENS
School of Civil Engineering
Department of Water Resources and Environmental Engineering

Stochastic investigation of hydrological extremes: influence of temporal variability and dependence

For the degree of Doctor of Philosophy
by Theano Iliopoulou

Athens, October 2020

ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
Σχολή Πολιτικών Μηχανικών
Τομέας Υδατικών Πόρων και Περιβάλλοντος

Στοχαστική διερεύνηση υδρολογικών ακροτάτων: επίδραση της χρονικής μεταβλητότητας και εξάρτησης

Διατριβή για την απόκτηση διδακτορικού διπλώματος από τη

Θεανώ Ηλιοπούλου

Πολιτικό Μηχανικό ΕΜΠ, MSc. Imperial College London.

Supervisor

Professor Demetris Koutsoyiannis (NTUA)

Advisory board

1. Professor Demetris Koutsoyiannis, National Technical University of Athens
2. Associate Professor Nikos Mamassis, National Technical University of Athens
3. Professor Alberto Montanari, University of Bologna

Examination Committee

1. Professor Demetris Koutsoyiannis, National Technical University of Athens
2. Associate Professor Nikos Mamassis, National Technical University of Athens
3. Professor Alberto Montanari, University of Bologna
4. Assistant Professor Andreas Efstratiadis, National Technical University of Athens
5. Reader Christian Onof, Imperial College London
6. Associate Professor Vissarion Papadopoulos, National Technical University of Athens
7. Associate Professor Elena Volpi, University of Roma


© Theano Iliopoulou, 2020.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

PREFACE

First and foremost, I wish to express my gratitude to my supervisor, Professor Demetris Koutsoyiannis for the invaluable support during these years, including the essential scientific guidance and the continuous moral encouragement. I truly thank him for the endless inspiration ever since my graduate studies.

My sincere thanks are due as well to Professor Nikos Mamassis for the support and encouragement and for always providing a very helpful, realistic and grounding point-of-view on all academic matters.

I am also indebted to Professor Alberto Montanari for hosting my two-month research period at the University of Bologna and the ever since support, encouragement and inspiring scientific collaboration.

I sincerely thank as well Professor Andreas Efstratiadis for the moral encouragement and the always pleasant and fruitful collaboration.

From the early days of conducting my Master thesis in Imperial College, I cherish the interactions with Professor Christian Onof, who ever since that period has been supportive in my path.

I am also very grateful for the positive evaluation and constructive comments on the PhD content that I received by Proressors Elena Volpi and Vissarion Papadopoulos. I am also thankful and acknowledge the contributions by my co-authors throughout these years whose names are listed in the publications in the end.

Everyday academic life was made immensely happier through the creative collaboration and friendship of Panayiotis Dimitriadis, Romanos Ioannidis and Fivos Sargentis. Special thanks are due as well to Olga Kitsou for all her help and positivity.

In non-academic life, I feel lucky and thankful for the love and support of my parents, Alexis and Hara, my sister Christina, Giannis, and my best friends.

As this Thesis is being completed in a globally challenging period from the pandemic, I am grateful to each and everyone for their ever present support despite the physical distancing.

Ευχαριστώ.

T.I.
Athens, 28-06-2020

ACKNOWLEDGEMENTS

Theano Iliopoulou acknowledges financial support by the European Regional Development Fund of the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call RESEARCH – CREATE – INNOVATE (project code: T1EDK-00956), project «ARCHYTAS: Archetypal telemetry and decision support system for the protection of monumental structures» and by Greece and the European Union (European Social Fund-ESF) through the Operational Programme «Human Resources Development, Education and Lifelong Learning 2014-2020» in the context of the project "Development of Stochastic Methods for Extremes (ASMA): identification and simulation of dependence structures of extreme hydrological events" (MIS 5049175).

ABSTRACT

The understanding and modelling of hydrological extremes is a classic endeavor in hydrology and engineering, one which has received renewed interest during the past decades under climate change theory. Long before concerns regarding intensification of extremes became prominent, their inherent variability and uncertainty sufficed to make their understanding and modelling challenging. Stochastics, integrating probability, statistics, and the theory of stochastic processes, offer a uniquely appropriate and consistent framework to deal with the uncertain nature of extremes. While the marginal properties of extremes have been extensively studied in the literature, the same does not hold for their temporal properties, since extremes are traditionally treated as temporally independent. As a consequence, their temporal behaviours have been either largely overlooked, or approached via deterministic reasoning. Yet, there are both empirical and theoretical grounds that question the independence assumption, namely the fact that hydrological extremes originate from natural processes characterized by marked dependence at various scales.

This Thesis aims to stochastically investigate and model the temporal variability and dependence of hydrological extremes from seasonal to climatic scales. The key innovation of the analysis is the identification of the temporal behaviours of the extremes and their stochastic linkage to the inherent properties of the parent hydrological process. Such an approach creates new perspectives on understanding the temporal dynamics of hydrological extremes that can significantly improve the perception of related risk over time and inform advanced mitigation practices. Two complementary objectives are pursued in this respect: (a) the characterization of their temporal properties, including the multi-scale dependence dynamics, from long-term hydrological records, and (b) the development of hydrologically relevant modelling frameworks that reproduce the observed extremal patterns. These objectives unfold at the following three scales: (i) the seasonal scale, pertaining to extreme rainfall seasonality and dependence dynamics of seasonal streamflow extremes, (ii) the annual scale, with respect to the propagation of long-term persistence, i.e. Hurst-Kolmogorov (HK) dynamics, from the parent process to the extremes and properties thereof, and last, (iii) the climatic-scale, regarding the theoretical and empirical basis of climatic projections of future rainfall.

ΕΚΤΕΝΗΣ ΠΕΡΙΛΗΨΗ

ΑΝΤΙΚΕΙΜΕΝΟ ΕΡΕΥΝΑΣ

Ο όρος «ακρότατα» υποδηλώνει σπάνια και σημαντικά γεγονότα που είναι δύσκολο ή ακόμη και αδύνατο να προβλεφθούν με βάση την ιστορική εμπειρία. Κατ' αναλογία, τα ακραία υδρολογικά φαινόμενα σχετίζονται με την εμφάνιση βροχοπτώσεων και απορροών ασυνήθιστα μεγάλης έντασης ή/και συχνότητας που αποτελούν εν δυνάμει κίνδυνο για την ανθρώπινη κοινωνία. Συγκεκριμένα, οι ακραίες βροχοπτώσεις και πλημμύρες μπορούν να βλάψουν το δομημένο περιβάλλον, συμπεριλαμβανομένων σημαντικών έργων υποδομής, να διαταράξουν την οικονομική και αγροκτηνοτροφική δραστηριότητα προκαλώντας οικονομικές απώλειες, ενώ αποτελούν άμεση απειλή για τη δημόσια υγεία. Έτσι, η κατανόησή τους που έχει ως στόχο το μετριασμό του σχετικού κινδύνου αποτελούσε ανέκαθεν πρόκληση, γενικά για την κοινωνία, και ειδικά για την επιστήμη. Τις τελευταίες δεκαετίες όμως, αυτή η πρόκληση έχει γίνει ακόμα μεγαλύτερη καθώς αμφισβητείται πλέον ριζικά η ίδια η επάρκεια της συμβατικής αντίληψης για τη διακινδύνευση (Hall et al., 2014).

Από τη μία πλευρά, η παγκόσμια κάλυψη των καταστροφών από τα μέσα ενημέρωσης αύξησε τη διαθεσιμότητα παραδειγμάτων καταστροφικών υδρολογικών γεγονότων (Barredo, 2007), προκαλώντας ολοένα αυξανόμενες ανησυχίες σχετικά με την εντατικοποίηση των ακροτάτων. Οι ανησυχίες αυτές εντείνονται περαιτέρω από τις προβλέψεις της θεωρίας της ανθρωπογενούς κλιματικής αλλαγής. Σύμφωνα με την τελευταία, οι αυξημένες ανθρωπογενείς εκπομπές αερίων θερμοκηπίου τις τελευταίες δεκαετίες έχουν προκαλέσει συστηματικές αλλαγές στη δυναμική του κλίματος που οδηγούν στη θέρμανση του πλανήτη και εντατικοποίηση του κύκλου του νερού (IPPC, 2014; Πέμπτη έκθεση αξιολόγησης AR5). Σε αυτή τη βάση, έχει υποστηριχθεί από μέρος της επιστημονικής κοινότητας ότι είναι αναγκαία μια ριζική αναδιατύπωση των υποθέσεων και των μεθόδων μοντελοποίησης προκειμένου να συμπεριληφθεί αιτιοκρατικά η επίδραση των νέων ανθρωπογενών παραγόντων (Milly et al., 2008). Αν και αυτή η θέση έχει επικριθεί σε μεγάλο βαθμό στην υδρολογία (Cohn and Lins, 2005; Montanari and Koutsoyiannis, 2014; Koutsoyiannis and Montanari, 2015; Serinaldi et al., 2018; Koutsoyiannis, 2020a), οι ανησυχίες σχετικά με την εντατικοποίηση των ακροτάτων στο μέλλον είναι κυρίαρχες στη σχετική βιβλιογραφία.

Από την άλλη πλευρά, είναι ευρέως παραδεκτό ότι η δυναμική της διακινδύνευσης έχει αλλάξει κατά τον περασμένο αιώνα ως αποτέλεσμα συστηματικών αλλαγών στην έκθεση του ανθρώπου στον υδρολογικό κίνδυνο. Για παράδειγμα, η έκθεση του ανθρώπου στον υδρολογικών κίνδυνο έχει αυξηθεί διαχρονικά ως αποτέλεσμα της ανθρώπινης τάσης για οργάνωση της ζωής στην εγγύτητα υδρολογικών δικτύων (Ceola et al., 2014). Ταυτόχρονα, η ευπάθεια έναντι ακραίων υδρολογικών γεγονότων έχει επίσης αυξηθεί λόγω της υψηλής πυκνότητας πληθυσμού στα αστικά κέντρα και της άναρχης αστικοποίησης. Μόνο στην Ελλάδα, πάνω από 200 θάνατοι λόγω ακραίων πλημμυρών έχουν αναφερθεί

από το 1960, με την πλειοψηφία τους να συγκεντρώνεται στην έντονα αστικοποιημένη περιοχή της Αττικής.

Η υπεκτίμηση τόσο του μεγέθους όσο και της συχνότητας εμφάνισης των ακροτάτων είναι, δυστυχώς, συχνός παράγοντας σε περιπτώσεις υδρολογικών καταστροφών άνευ προηγουμένου (Mimikou and Koutsoyiannis, 1995; Coles et al., 2003; Koutsoyiannis et al., 2012; Ntigkakis et al., 2018). Ανησυχίες εκφράζονται επίσης ως προς το ενδεχόμενο υπεκτίμησης της πιθανότητας αστοχίας υποδομών και έργων μεγάλης κλίμακας από ακραίες βροχοπτώσεις και πλημμύρες, καθώς πλήθος σχετικών αστοχιών έχει αναφερθεί τα τελευταία χρόνια. Συγκεκριμένα, ακραία υδρολογικά φαινόμενα έχουν προκαλέσει καταστροφές φραγμάτων, καταρρεύσεις γεφυρών και σοβαρές φθορές στις υποδομές μεταφορών, συμπεριλαμβανομένων δρόμων και σιδηροδρόμων, προξενώντας τεράστιες οικονομικές ζημιές και ανθρώπινες απώλειες (Wardhana και Hadipriono, 2003; Serra-Llobet et al., 2013; Koskinas et al. , 2019; Kellermann et al., 2019; Pizarro et al., 2020).

Είναι φανερό ότι η αναβάθμιση του υδρολογικού σχεδιασμού έναντι ακραίων γεγονότων αποτελεί πλέον επιτακτική ανάγκη υψηλής κοινωνικής προτεραιότητας. Από αυτήν την άποψη, υποστηρίζεται ότι, προτού εξετάσουμε την επίδραση παγκόσμιων τάσεων ή εξωτερικών παραγόντων, είναι κρίσιμο να βελτιωθεί πρωτίστως το υπάρχον επίπεδο κατανόησης των υδρολογικών ακροτάτων, ξεκινώντας από τη διεργασία της βροχόπτωσης. Για το σκοπό αυτό, είναι σημαντικό να επανεξετάσουμε πρώτα και κύρια, και υπό το φως νέων δεδομένων, την εγκυρότητα των κλασικών υποθέσεων που διέπουν τη μελέτη των ακροτάτων.

Η χρονική ανεξαρτησία είναι η κυρίαρχη υπόθεση σε εφαρμογές θεωρίας ακραίων τιμών, όπως συναντάται στα περισσότερα επιστημονικά εγχειρίδια. Έτσι, ενώ η περιθώρια κατανομή των ακροτάτων έχει μελετηθεί εκτενώς στη σχετική βιβλιογραφία, η χρονική μεταβλητότητα τους είτε παραγνωρίζεται εξ ολοκλήρου είτε μελετάται μέσω ντετερμινιστικών θεωρήσεων, εφόσον τα ίδια αντιμετωπίζονται εκ των προτέρων ως ανεξάρτητες τυχαίες μεταβλητές. Η υπόθεση ωστόσο της χρονικής ανεξαρτησίας επιδέχεται σημαντικής αμφισβήτησης επί τη βάσει τόσο εμπειρικών όσο και θεωρητικών λόγων. Ο κυριότερος από αυτούς είναι το γεγονός ότι τα υδρολογικά ακρότατα είναι προϊόν φυσικών διεργασιών που χαρακτηρίζονται οι ίδιες από ισχυρές δομές εξάρτησης σε διάφορες κλίμακες. Η θεωρία των στοχαστικών μεθόδων η οποία ενσωματώνει τις έννοιες της πιθανότητας, της στατιστικής και των στοχαστικών ανελίξεων, αποτελεί το πλέον πρόσφορο θεωρητικό πλαίσιο για τη κατανόηση και διερεύνηση της χρονικής μεταβλητότητας ακροτάτων που αποκλίνουν από την ιδεατή συνθήκη της ανεξαρτησίας (Koutsoyiannis, 2020b).

Ο κεντρικός στόχος της παρούσας διατριβής είναι η στοχαστική διερεύνηση των υδρολογικών ακροτάτων ως προς την χρονική τους μεταβλητότητα και εξάρτηση, και η συνεπαγόμενη επανεξέταση των κλασικών υποθέσεων της μελέτης τους. Η βασική καινοτομία εστιάζεται στην αναγνώριση των χρονικών συμπεριφορών των ακροτάτων και στη στοχαστική σύνδεσή τους με τις εγγενείς ιδιότητες της μητρικής υδρολογικής διεργασίας. Μια τέτοια προσέγγιση δημιουργεί ένα νέο πρίσμα κατανόησης της δυναμικής των υδρολογικών ακροτάτων που μπορεί

να βελτιώσει σημαντικά την αντίληψη του σχετικού κινδύνου στο χρόνο και να συνεισφέρει στην αναβάθμιση πρακτικών μετριασμού της υδρολογικής διακινδύνευσης. Για το σκοπό αυτό, καταρτίζεται ένα δυσεύρετο σύνολο χρονοσειρών μεγάλου μήκους και τίθενται δύο επιμέρους στόχοι: (α) ο αποτελεσματικός χαρακτηρισμός των χρονικών ιδιοτήτων και της εξάρτησης των ακροτάτων από εποχικές έως κλιματικές κλίμακες, όπως προκύπτουν από τις ιστορικές χρονοσειρές, και (β) η διαμόρφωση μεθοδολογικών πλαισίων μοντελοποίησης για την αναπαραγωγή των παρατηρημένων χρονικών προτύπων. Η ανάλυση αφορά υδρολογικά ακρότατα σε τρεις χρονικές κλίμακες: την εποχική, την ετήσια, και την κλιματική. Το κύριο σώμα της ανάλυσης αφορά στη διερεύνηση των ακραίων βροχοπτώσεων, οι χρονικές ιδιότητες των οποίων έχουν μελετηθεί ελάχιστα σε σχέση με αυτές των απορροών. Η χρονική εξάρτηση των ακραίων απορροών διερευνάται επίσης με δύο σκοπούς: (α) τον εντοπισμό εποχικών δομών εξάρτησης που βελτιώνουν την προβλεψιμότητα των ακραίων απορροών, και (β) την σύγκριση των ιδιοτήτων τους με αυτές της βροχόπτωσης ώστε να εξαχθούν ευρύτερα υδρολογικά συμπεράσματα. Παρακάτω γίνεται συνοπτική αναφορά στους στόχους και στα ευρήματα των επιμέρους κεφαλαίων.

ΧΡΟΝΙΚΗ ΕΞΑΡΤΗΣΗ ΑΠΟ ΤΗ ΜΗΤΡΙΚΗ ΣΤΟΧΑΣΤΙΚΗ ΑΝΕΛΙΞΗ ΣΤΑ ΑΚΡΟΤΑΤΑ ΤΗΣ: ΕΠΙΣΚΟΠΗΣΗ ΤΗΣ ΒΙΒΛΙΟΓΡΑΦΙΑΣ

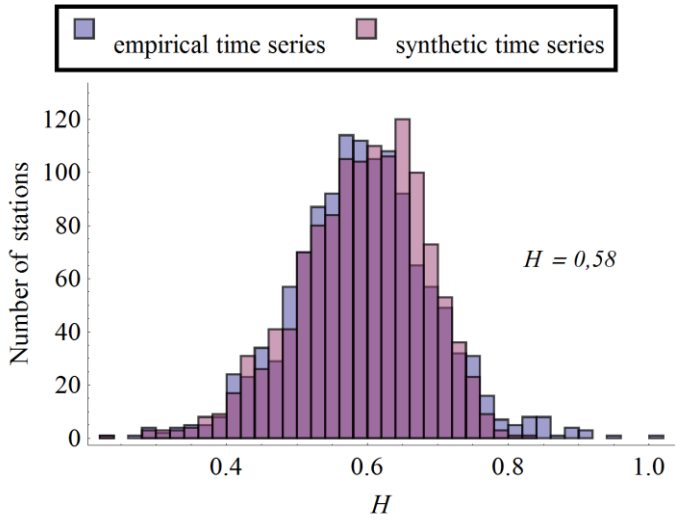
Στο κεφάλαιο 2 αναπτύσσεται το θεωρητικό πλαίσιο της διατριβής. Αρχικά, παρουσιάζονται οι θεμελιώδεις έννοιες των στοχαστικών ανελίξεων και περιγράφονται τα βασικά εργαλεία ανάλυσης της χρονικής εξάρτησης που χρησιμοποιούνται. Πραγματοποιείται εκτενής επισκόπηση της βιβλιογραφίας σχετικά με τη θεωρία ακραίων τιμών και αναδεικνύονται μερικά από τα λιγότερο γνωστά θεωρητικά αποτελέσματα που σχετίζονται με την εξάρτηση των ακροτάτων. Στη συνέχεια, παρουσιάζονται κριτικά οι κυρίαρχες μεθοδολογικές προσεγγίσεις στη μοντελοποίηση των ακροτάτων. Τέλος, προσδιορίζονται ανοιχτές θεωρητικές ερωτήσεις σχετικά με τις ιδιότητες ακροτάτων από στοχαστικές ανελίξεις με υψηλή χρονική συσχέτιση και «βαριές» ουρές κατανομής.

ΕΠΑΝΕΞΕΤΑΣΗ ΤΗΣ ΥΠΑΡΞΗΣ ΕΜΜΟΝΗΣ ΣΤΗ ΔΙΕΡΓΑΣΙΑ ΤΗΣ ΕΤΗΣΙΑΣ ΒΡΟΧΟΠΤΩΣΗΣ

Στο κεφάλαιο 3 διερευνάται η ύπαρξη εμμονής, αλλιώς δυναμικής Hurst-Kolmogorov, στη διεργασία της ετήσιας βροχόπτωσης από μία παγκόσμια βάση δεδομένων. Η αναγνώριση της ύπαρξης εμμονής σε μια φυσική διεργασία αποτελεί ισχυρό κίνητρο για τη διερεύνηση των μακροπρόθεσμων χρονικών ιδιοτήτων των ακροτάτων της. Ενώ η εμμονή έχει αναγνωριστεί ως δυναμική σε πλήθος φυσικών διεργασιών, η ύπαρξή της στη διεργασία της βροχόπτωσης παραμένει αδιευκρίνιστη. Έτσι, επανεξετάζεται εδώ χρησιμοποιώντας ένα παγκόσμιο σύνολο βροχομετρικών σταθμών.

Από την ανάλυση του παγκόσμιου συνόλου δεδομένων προκύπτουν αξιοσημείωτες ενδείξεις ύπαρξης εμμονής στην ετήσια βροχόπτωση. Η θεωρητική τιμή της παραμέτρου Hurst $H \approx 0.58$ που εκτιμήθηκε μέσω ανάλυσης Monte Carlo

μπορεί να θεωρηθεί αντιπροσωπευτική για την πλειονότητα (97.5%) των 1265 σταθμών (Εικόνα 1). Η δομή της αυτοσυσχέτισης βρέθηκε επίσης γενικά ισχυρότερη από τη δομή ενός μοντέλου Markov και συνεπής με την αυτήν ενός μοντέλου με εμμονή. Ορισμένες μελέτες που είχαν χρησιμοποιήσει μικρότερα μήκη δεδομένων (Potter, 1979; Fraedrich and Blender, 2003; Kantelhardt et al., 2006) υποστήριξαν την καταλληλότητα της δομής Markov, χωρίς όμως να διερευνούν τις διαφορές μεταξύ πραγματικής και θεωρητικής αυτοσυσχέτισης για διάφορες χρονικές υστερήσεις. Αυτές οι διαφορές μπορεί να είναι μικρές, ωστόσο, έχει αποδειχθεί ότι ενδέχεται να έχουν σοβαρές επιπτώσεις όσον αφορά την αβεβαιότητα εκτίμησης (Koutsoyiannis and Montanari, 2007). Για παράδειγμα, όσον αφορά τη στατιστική σημαντικότητα των τάσεων, οι παρατηρούμενες αλλαγές στις βροχοπτώσεις μπορεί να θεωρηθούν πολύ πιο σπάνιες από ό, τι είναι στην πραγματικότητα (Cohn and Lins, 2005). Τέλος, καταδείχθηκε ότι η δομή της αυτοσυσχέτισης αποκλίνει σημαντικά από την περίπτωση της ανεξαρτησίας.

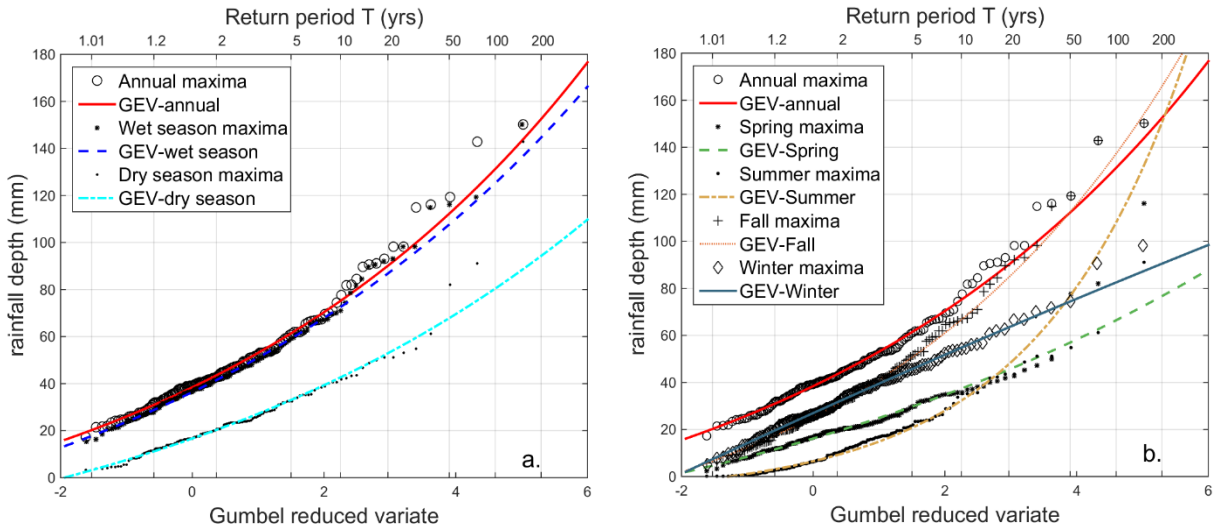


Εικόνα 1 Κατανομή της παραμέτρους Hurst από παρατηρημένες χρονοσειρές (μπλε χρώμα) και από συνθετικές χρονοσειρές (μωβ χρώμα) από στοχαστική ανέλιξη με παράμετρο $\boldsymbol{H} = 0.58$.

Αν και τα παραπάνω ευρήματα τάσσονται υπέρ της ύπαρξης μιας ισχυρότερης δομής εξάρτησης από αυτήν που συνήθως υποτίθεται στη βιβλιογραφία (Potter, 1979; Fraedrich and Blender, 2003; Kantelhardt et al., 2006), φαίνεται ότι υπάρχει μια ασυμφωνία μεταξύ των μικρότερων και μεγαλύτερες χρονικές κλίμακες (Fraedrich and Larnder, 1993; Pelletier and Turcotte, 1997; Poveda, 2011; Ault et al., 2013). Για το σκοπό αυτό, δεν πρέπει να παραβλέπεται η πιο σημαντική πηγή αβεβαιότητας στον προσδιορισμό της εμμονής, που είναι το μήκος της χρονοσειράς (Koutsoyiannis, 2002; Koutsoyiannis and Montanari, 2007). Ένα σημαντικό συμπέρασμα που προκύπτει από την ανάλυση είναι ότι οι πραγματικές χρονοσειρές μπορεί να αποκλίνουν σημαντικά από απλοποιητικές υποθέσεις που χρησιμοποιούνται στην πράξη, όπως η χρονική ανεξαρτησία, και ως εκ τούτου, η εμπεριστατωμένη μελέτη των ιδιοτήτων εξάρτησης είναι απαραίτητη, ειδικά όταν ενδιαφέρουν μακροπρόθεσμοι χρονικοί ορίζοντες.

ΧΑΡΑΚΤΗΡΙΣΜΟΣ ΚΑΙ ΜΟΝΤΕΛΟΠΟΙΗΣΗ ΤΗΣ ΑΚΡΑΙΑΣ ΕΠΟΧΙΚΗΣ ΒΡΟΧΟΠΤΩΣΗΣ

Το κεφάλαιο 4 μελετά τη χρονική μεταβλητότητα των ακραίων βροχοπτώσεων που προκαλείται από την εποχικότητα. Για το σκοπό αυτό αναλύεται ένα σύνολο 27 σταθμών βροχής με δεδομένα ημερήσιων καταγραφών που υπερβαίνουν τα 150 έτη. Προτείνεται μια καινοτομική μεθοδολογία που επιτρέπει τον αντικειμενικό προσδιορισμό της εποχικότητας στις ακραίες ημερήσιες βροχοπτώσεις και τη μοντελοποίηση της περιθώριας κατανομής των ακροτάτων κάθε εποχής. Η μεθοδολογία αναγνώρισης εποχής είναι σε θέση να προσδιορίσει τη βέλτιστη επιλογή μοντελοποίησης για την εποχική ακραία βροχόπτωση, προσδιορίζοντας τόσο το βέλτιστο αριθμό εποχών όσο και τη χρονική διάρκεια τους, βάσει του κριτηρίου πληροφορίας Akaike (AIC). Αντί για την αυθαίρετη και μαζική εφαρμογή των 4 εποχών του έτους, η μέθοδος προκρίνει φειδωλή και κατά περίπτωση μοντελοποίηση της εποχικότητας κάθε σταθμού που αποδεικνύεται συνεπέστερη και αποτελεσματικότερη για τη μελέτη των ακροτάτων στην πράξη (Εικόνα 2).



Εικόνα 2 Πιθανοτικά διαγράμματα Gumbel της προσαρμογής της Γενικευμένης Κατανομής Ακροτάτων (GEV distribution) στα ετήσια και εποχικά μέγιστα βάσει (a) της προτεινόμενης μεθοδολογίας και (b) του συμβατικού επιμερισμού σε 4 εποχές για τη χρονοσειρά βροχοπτώσεων της Αθήνας.

Καταδεικνύεται επίσης ότι η ύπαρξη ισχυρής εποχικότητας επηρεάζει την περιθώρια κατανομή των ακροτάτων ως προς τις παραμέτρους θέσης και κλίμακας της Γενικευμένης Κατανομής Ακροτάτων (GEV) οι οποίες είναι υψηλότερες την υγρή περίοδο. Αντίθετα, η παράμετρος σχήματος παρουσιάζει περιορισμένη εποχική μεταβλητότητα, και άρα η εποχικότητα δεν επηρεάζει ουσιαστικά το σχήμα της ουράς της κατανομής. Προκύπτει επίσης ότι η ανάλυση των ετήσιων μεγίστων, χωρίς επαρκή προσδιορισμό της εποχικότητας, είναι υπέρ συντηρητικού σχεδιασμού, δεδομένου ότι συμπεριλαμβάνει τις σπάνιες περιπτώσεις ακραίων γεγονότων σημαντικού μεγέθους που συμβαίνουν στην ξηρή περίοδο. Ωστόσο, για τον ενδοετήσιο υδρολογικό σχεδιασμό, είναι σημαντικό να ληφθεί υπόψη η εποχική μεταβλητότητα των ακροτάτων. Η σειρά μέγιστων της υγρής εποχής περιέχει πολύτιμες πληροφορίες σχετικά με το χρόνο εμφάνισης των πιο ακραίων

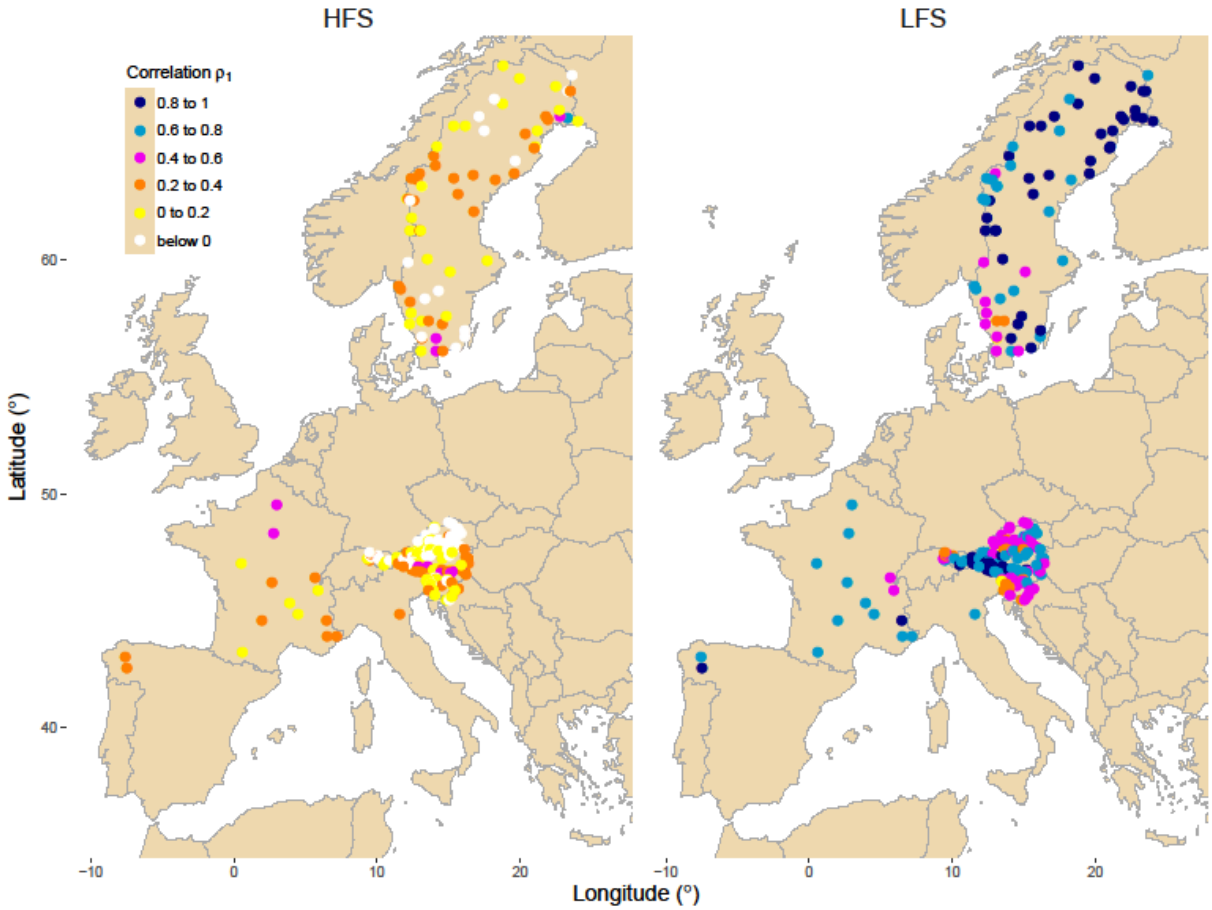
γεγονότων, παρότι το μέγεθος τους αναμένεται κοντά στο εκτιμώμενο από τη σειρά ετήσιων μεγίστων. Ωστόσο, η εκτίμηση της εποχικότητας στα μέγιστα είναι σημαντική για ξηρές υδρολογικές περιόδους ώστε να αποφευχθεί δαπανηρή υπερεκτίμηση των μεγεθών σχεδιασμού εποχικών έργων καθώς και άσκοπη διοχέτευση πλημμυρικών ροών σε περίπτωση ταμιευτήρων υδροδότησης.

Τα ευρήματα αυτά έχουν άμεσες εφαρμογές τόσο στη θεωρητική σύλληψη της εποχικότητας σε ακραίες βροχοπτώσεις όσο και στη μοντελοποίηση στην πράξη. Σε μεθοδολογικό επίπεδο, συμβάλλουν στην ευρύτερη καθιέρωση κριτηρίων επιλογής μοντέλων στην υδρολογία, όπως το AIC, ενώ αποτελούν σημαντικό βήμα προς τον «αντικειμενικό» προσδιορισμό της εποχικότητας τόσο σε τοπική όσο και σε παγκόσμια κλίμακα.

ΔΥΝΑΜΙΚΕΣ ΕΞΑΡΤΗΣΗΣ ΤΩΝ ΕΠΟΧΙΚΩΝ ΑΚΡΑΙΩΝ ΑΠΟΡΡΟΩΝ

Το Κεφάλαιο 5 μελετά την ύπαρξη εξάρτησης στις υψηλές και χαμηλές απορροές στην εποχική κλίμακα, τους πιθανούς γεωφυσικούς και υδρολογικούς παράγοντες που ευνοούν την εκδήλωσή της σε μια υδρολογική λεκάνη καθώς και τη δυνατότητα αξιοποίησης της σχετικής πληροφορίας για τη βελτίωση της πιθανοτικής πρόβλεψης πλημμυρικών αιχμών και χαμηλών ροών. Για το σκοπό αυτό, αναλύονται δεδομένα ημερήσιων παροχών άνω των 50 ετών από 224 ποταμούς στην Ευρώπη.

Τα αποτελέσματα δείχνουν ότι η πλειονότητα των ποταμών εμφανίζει χαρακτηριστικά εξάρτησης στην εποχική κλίμακα, που εκδηλώνονται ως γραμμική συσχέτιση μεταξύ των προγενέστερων μέσων μηνιαίων απορροών και των μεταγενέστερων (α) πλημμυρικών αιχμών στην υγρή εποχή (High Flow Season; HFS) και (β) μέσων απορροών στη ξηρή εποχή (Low Flow Season; LFS), αντίστοιχα. Η συσχέτιση για τις απορροές της ξηρής εποχής είναι υψηλότερη απ' ότι για της υγρής εποχής, ενώ και στις δύο περιπτώσεις εμφανίζεται αυξημένη χωρική μεταβλητότητα καθώς και ομαδοποίηση της συσχέτισης (Εικόνα 4).



Εικόνα 3 Χωρική κατανομή της γραμμικής συσχέτισης (α) των πλημμυρικών αιχμών την υγρή περίοδο (HFS) με τις μέσες απορροές του προηγούμενου μήνα (αριστερά) και (β) των μέσων απορροών την ξηρή περίοδο (LFS) με τις μέσες απορροές του προηγούμενου μήνα (δεξιά). Ο πίνακας δείχνει τα χρώματα που αντιστοιχούν στις κλάσεις των συσχετίσεων.

Από τη μελέτη πλήθους γεωφυσικών και υδρολογικών χαρακτηριστικών των λεκανών απορροής προκύπτει ότι βραδύτεροι χρόνοι υδρολογικής απόκρισης της λεκάνης, καθώς και η κυρίαρχη ύπαρξη βασικής ροής ευνοούν την εποχική συσχέτιση. Αντιθέτως, η συσχέτιση είναι χαμηλότερη στις καρστικές λεκάνες που ανταποκρίνονται γρήγορα καθώς σε λεκάνες με υψηλή ειδική απορροή και υψηλότερη μέση βροχόπτωση, πιθανώς λόγω της παρουσίας κορεσμένων συνθηκών σε υγρότερα κλίματα και της αυξημένης βραχυπρόθεσμης μεταβλητότητάς τους.

Ως προς την πρακτική αξία, παρουσιάζεται μια εφαρμογή αξιοποίησης της εποχικής συσχέτισης για τη μείωση της αβεβαιότητας των εποχικών προβλέψεων υψηλών και χαμηλών ροών, μέσω στοχαστικών μοντέλων που ενσωματώνουν νέες παρατηρήσεις σε μηνιαία κλίμακα για την ανανέωση των πιθανοτικών εκτιμήσεών τους.

Τα αποτελέσματα συνηγορούν στο ότι η ύπαρξη εποχικής «μνήμης» του ποταμού παρέχει σημαντικές πληροφορίες για την κατανόηση και πρόβλεψη των ακραίων απορροών, ενώ εξασφαλίζει σημαντικό χρόνο ορίζοντα σε κλίμακα μηνών, για τη λήψη αποφάσεων σχετικά με την πρόληψη υδρολογικών κινδύνων στη λεκάνη.

ΕΜΜΟΝΗ ΣΤΗΝ ΕΜΦΑΝΙΣΗ ΑΚΡΟΤΑΤΩΝ: ΟΜΑΔΟΠΟΙΗΣΗ ΣΕ ΠΟΛΛΑΠΛΕΣ ΚΛΙΜΑΚΕΣ

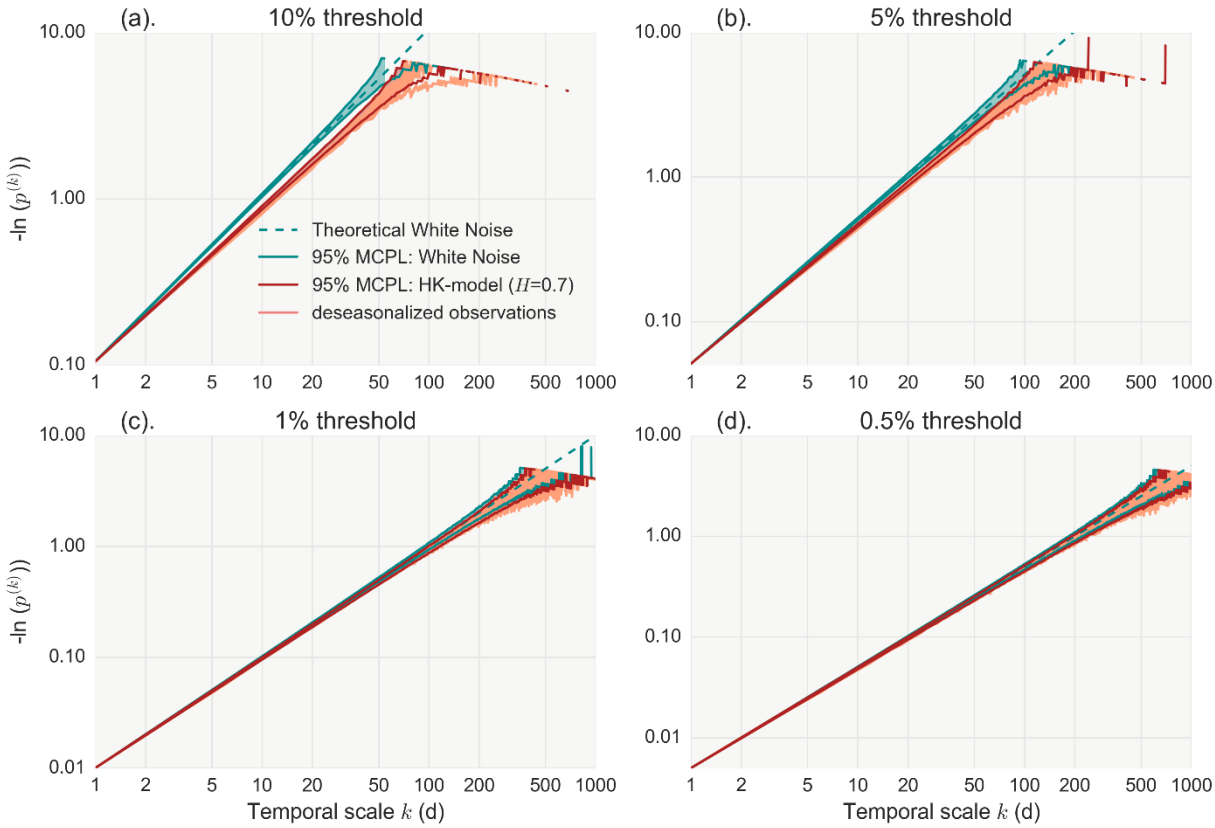
Το κεφάλαιο 6 πραγματεύεται τη δυναμική διάδοσης της εμμονής από τη μητρική στοχαστική ανέλιξη στα ακρότατά της, εστιάζοντας στο φαινόμενο της ομαδοποίησής τους. Η ύπαρξη δομής ομαδοποίησης των ακροτάτων αμφισβητεί την υπόθεση της χρονικής ανεξαρτησίας τους και έτσι, η κατανόησή της είναι κρίσιμη για τον υδρολογικό σχεδιασμό και την αντίληψη της χρονικής εξέλιξης του κινδύνου. Εδώ, επιδιώκεται (α) η θεωρητική και εμπειρική σύνδεση μεταξύ της δυναμικής ομαδοποίησης των ακροτάτων και της μακροπρόθεσμης εμμονής, δηλ. της δυναμικής HK, και (β) η μεθοδολογία ανάκτησης της δεύτερης, δηλαδή της εμμονής της στοχαστικής ανέλιξης, από την πρώτη, δηλαδή χρονική συμπεριφορά των ακροτάτων της. Για το σκοπό αυτό συλλέγεται μια δυσεύρετη βάση υδρολογικής πληροφορίας από 60 ημερήσιους βροχομετρικούς σταθμούς ανά τον κόσμο με πάνω από 150 έτη ημερήσιων καταγραφών.

Ένα γενικό συμπέρασμα που προκύπτει είναι ότι η δυνατότητα αναγνώρισης της εμμονής από χρονοσειρές μεγίστων είναι περιορισμένη και εξασθενεί σταδιακά καθώς αυξάνεται το κατώφλι πάνω από το οποίο συλλέγονται τα ακρότατα. Στη χρονική συμπεριφορά των ακροτάτων υπάρχει σημαντική επιρροή τόσο από τις ιδιότητες εξάρτησης δεύτερης τάξης (δυναμική HK) όσο και από τις ροπές υψηλής τάξης της μητρικής ανέλιξης, και επομένως είναι απαραίτητη η επίγνωση και των δυο για την κατανόησή τους. Έτσι, εκτιμήσεις της συσχέτισης των ακροτάτων μόνο με μεθόδους δευτέρας τάξης όπως το κλιμακόγραμμα και δείκτες ομαδοποίησης βασισμένοι στη διασπορά, υποεκτιμούν τη συσχέτιση ακροτάτων από ανελίξεις με μη-Γκαουσιανή συμπεριφορά, οι οποίες επηρεάζονται από ροπές υψηλότερης τάξης. Για το λόγο αυτό προτείνεται ένας νέος πιθανοτικός δείκτης που χαρακτηρίζει την ομαδοποίηση βάσει της πιθανότητας μη υπέρβασης ενός δεδομένου κατωφλίου σε κλίμακα και ονομάζεται δείκτης NEPvS (Non-Exceedance Probability vs Scale). Ο δείκτης μπορεί να χαρακτηρίσει την ομαδοποίηση των ακροτάτων από πλήθος ανελίξεων, συμπεριλαμβανομένων αυτών με ισχυρή εμμονή, τύπου HK, και βαριές ουρές κατανομής. Προτείνεται μάλιστα και προσαρμογή ενός μοντέλου που είχε αρχικά προταθεί για την περιγραφή της πιθανότητας μη βροχής (probability dry) από Koutsoyiannis (2006), το οποίο μπορεί να χαρακτηρίσει τον προτεινόμενο πιθανοτικό δείκτη ως εξής:

$$p^{(k)} = p^{\left(1+\left(\xi^{-1/\eta}-1\right)(k-1)\right)^{\eta}}, \qquad p = 1 - F(u) \tag{1}$$

όπου u είναι το κατώφλι για τα ακρότατα, F η περιθώρια συνάρτηση κατανομής, και η, ξ παράμετροι στο (0, 1). Για τιμές των παραμέτρων $\eta = 1$ και $\xi = 0.5$, προκύπτει η συμπεριφορά ανέλιξης Λευκού Θορύβου, δηλαδή χρονικής ανεξαρτησίας. Καθώς αυξάνεται η εμμονή της μητρικής ανέλιξης, αυξάνεται η πιθανότητα μη εμφάνισης ακροτάτων σε μια κλίμακα (αντίστοιχα, ο μείον λογάριθμος της πιθανότητας μειώνεται όπως φαίνεται στην Εικόνα 4) και οι παράμετροι του μοντέλου γίνονται $\eta < 1$ και $\xi > 0.5$. Αυτή η συμπεριφορά υποδηλώνει ότι η άλλη όψη της ύπαρξης ομαδοποίησης των ακροτάτων είναι η ύπαρξη παρατεταμένων χρονικών περιόδων χωρίς ακραίες παρατηρήσεις.

Εκτεταμένες προσομοιώσεις συνθετικών χρονοσειρών Monte Carlo για βροχομετρικούς σταθμούς έδειξαν αποκλίσεις από τη συμπεριφορά ανεξαρτησίας και συμφωνία της ομαδοποίησης των ακροτάτων με τις ιδιότητες εξάρτησης (δομή HK) και κατανομής (ακριβής διατήρηση 4 πρώτων ροπών) της μητρικής ανέλιξης, όπως φαίνεται στην Εικόνα 4.



Εικόνα 4 Μείον φυσικός λογάριθμος της πιθανότητας μη υπέρβασης του κατωφλίου προς την κλίμακα για 28 χρονοσειρές βροχής στην Ολλανδία με αφαίρεση εποχικότητας 2 πρώτων ροπών, καθώς και 95% όρια Monte Carlo για θεωρητικό μοντέλο με συσχέτιση HK με $\boldsymbol{H}$=0.7 και διατήρηση 4 πρώτων ροπών, για 4 διαφορετικά κατώφλια ακροτάτων: (a) 10%, (b) 5%, (c) 1% και (d) 0.5%.

Καταλήγουμε στο συμπέρασμα ότι τα ακρότατα τείνουν να «κρύβουν» την εμμονή της μητρικής στοχαστικής ανέλιξης, οδηγώντας συχνά σημαίνει λανθασμένα στο συμπέρασμα της χρονικής ανεξαρτησίας. Οι επιπτώσεις όμως της εμμονής στην εκτίμηση των ακροτάτων είναι υπαρκτές, παρόλο που η ισχύς των ενδείξεων από συνήθεις χρονοσειρές είναι συχνά αδύναμη.

ΕΠΙΔΡΑΣΗ ΕΜΜΟΝΗΣ ΣΤΗΝ ΠΕΡΙΘΩΡΙΑ ΚΑΤΑΝΟΜΗ ΤΩΝ ΑΚΡΟΤΑΤΩΝ ΚΑΙ ΙΔΙΟΤΗΤΩΝ ΤΟΥΣ

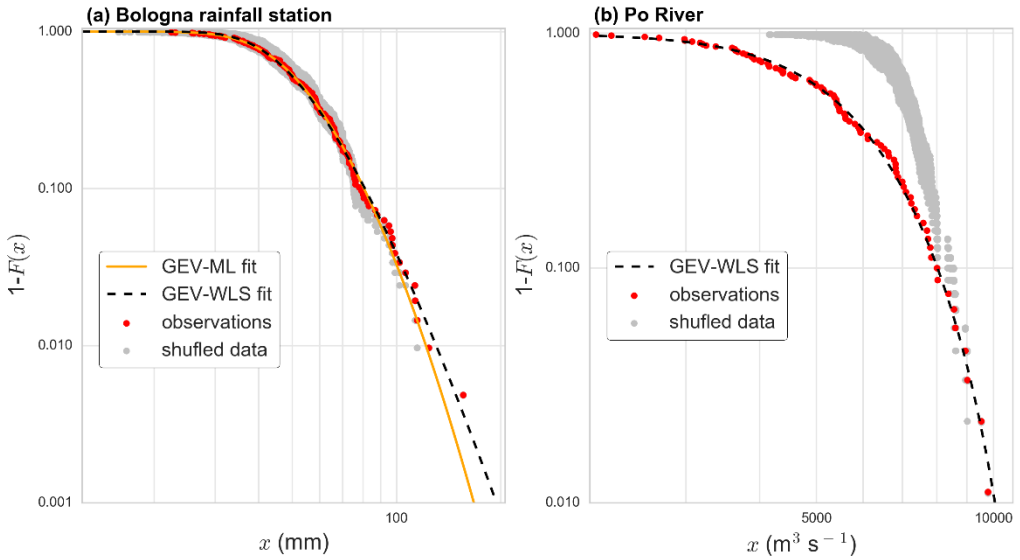
Το Κεφάλαιο 6 εξετάζει την επίδραση της εξάρτησης στην περιθώρια κατανομή των ακροτάτων καθώς και των χαρακτηριστικών ιδιοτήτων τους.

Στο πλαίσιο αυτό αξιολογείται αρχικά μέσω εκτεταμένων προσομοιώσεων Monte Carlo η εφαρμοσιμότητα των θεωρημάτων της θεωρίας ακραίων τιμών και ειδικά η ισχύς της Γενικευμένης Κατανομής Ακροτάτων για έμμονες στοχαστικές ανελίξεις με εξάρτηση τύπου HK. Φαίνεται ότι η προσαρμογή της κατανομής με τη

μέθοδο σταθμισμένων ελαχίστων τετραγώνων είναι πολύ ικανοποιητική ακόμη και για μη ασυμπτωτικές συνθήκες ανελίξεων με έντονη εξάρτηση. Από την άλλη πλευρά, είναι αμφισβητήσιμο εάν η θεωρία του δείκτη ακραίας συσχέτισης (extremal index) για ανελίξεις που παράγουν μόνο βραχυπρόθεσμη ομαδοποίηση ακροτάτων επαρκεί για να περιγράψει ακρότατα από έμμονες ανελίξεις. Οι τελευταίες εμφανίζουν τις εξής ιδιότητες κατανομής σε σχέση με πλήρως ανεξάρτητες ανελίξεις: (α) η παράμετρος κλίμακας, που σχετίζεται με τη μεταβλητότητα των ακροτάτων, αυξάνεται ως αποτέλεσμα της αυξημένης μεταβλητότητας της έμμονης ανέλιξης στην ίδια κλίμακα, (β) η παράμετρος θέσης της κατανομής μειώνεται ως αποτέλεσμα της ομαδοποίησης των ακροτάτων σε λιγότερες κλίμακες, αλλά (γ) η παράμετρος σχήματος δεν επηρεάζεται, καθώς θεωρητικά ταυτίζεται με την παράμετρο σχήματος της ουράς της περιθώριας κατανομής της μητρικής ανέλιξης.

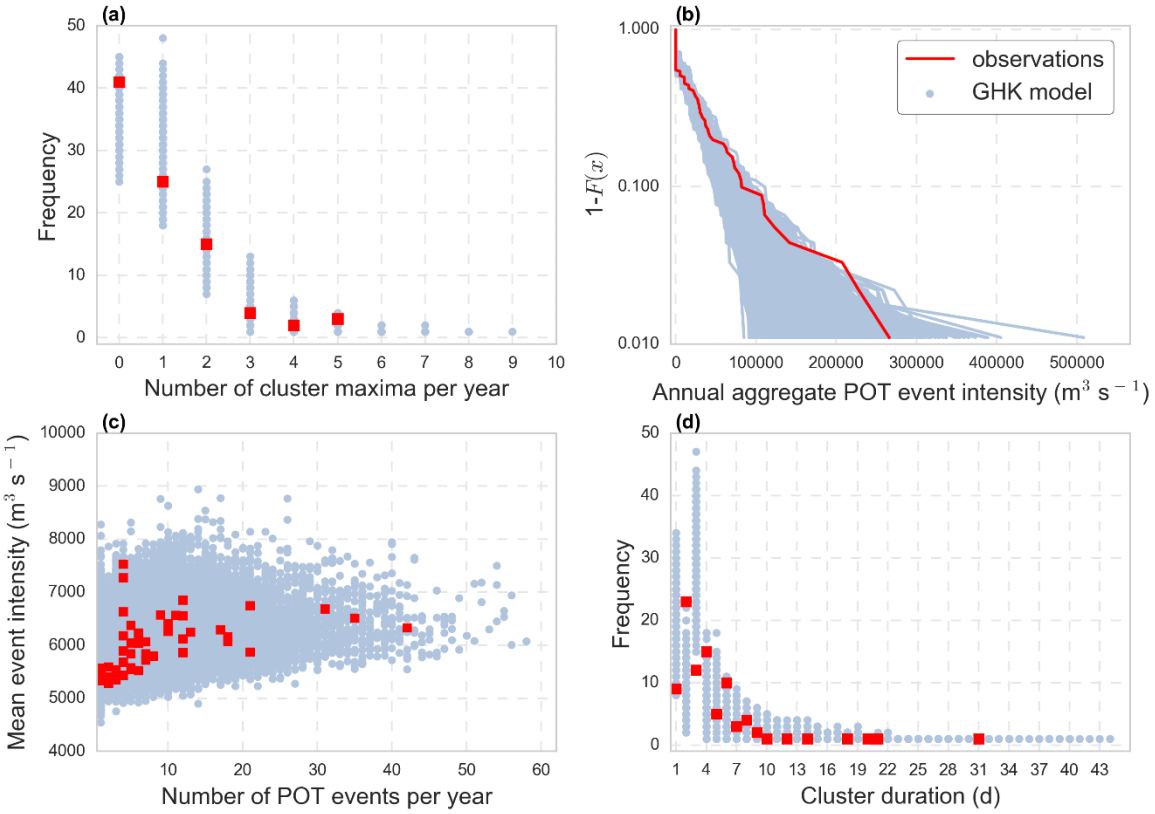
Αναδεικνύεται επίσης η επιρροή της εξάρτησης στις ιδιότητες των ακραίων παρατηρήσεων άνω κατωφλίου (Peaks Over Threshold; POT) σε ετήσια κλίμακα. Η απουσία γεγονότων POT σε ένα έτος είναι πιο πιθανή από την περίπτωση της πλήρους ανεξαρτησίας, αλλά στην περίπτωση που αυτά εμφανιστούν, αναμένεται υψηλότερη διάρκεια της «συστάδας» ακροτάτων και μεγαλύτερη ένταση. Οι επιπτώσεις αφορούν τόσο την εκτίμηση των κινδύνων πλημμύρας, ως προς την εκτίμηση της χρονικής περιόδου που μια περιοχή είναι υπό κατάκλιση (Dimitriadis and Koutsoyiannis, 2020), όσο και την εκτίμηση της συλλογικής διακινδύνευσης από αθροιστική έκθεση στον κίνδυνο που αφορά ιδιαίτερα τις ασφαλιστικές πρακτικές έναντι φυσικών καταστροφών (Serinaldi and Kilsby, 2016b; Goulianou et al., 2019; Manolis et al., 2020; Papoulakos et al., 2020).

Από τη μελέτη χρονοσειρών βροχοπτώσεων και απορροών στην ίδια υδρολογική περιοχή, προέκυψαν αποκλίσεις της χρονικής συμπεριφοράς των ακροτάτων τους από την τυπική συμπεριφορά ανεξαρτησίας, οι οποίες ήταν πιο έντονες στην ετήσια κλίμακα για τις απορροές, όπως φαίνεται στην Εικόνα 5.



Εικόνα 5 Πιθανότητες υπέρβασης των ετήσιων μεγίστων για τις παρατηρημένες χρονοσειρές και χρονοσειρές ίδια περιθώριας κατανομής αλλά απουσίας συσχέτισης μέσω τυχαίας αναδιάταξης (shuffled data) για το σταθμό βροχής στη Bologna και ημερήσιων απορροών του ποταμού Po.

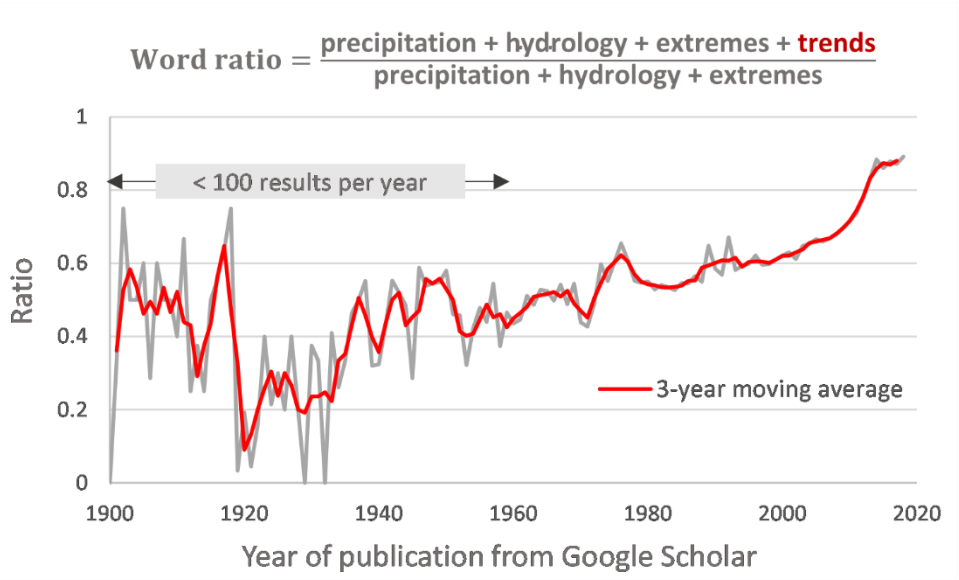
Μοντέλα τύπου HK προσαρμοσμένα στα δεδομένα αναπαρήγαγαν με επιτυχία τα παρατηρημένα χρονικά πρότυπα των ακροτάτων, όπως φαίνεται στην Εικόνα 6 για τον ποταμό Po, δείχνοντας τις δυνατότητες του στοχαστικού πλαισίου HK ως προς την εξήγηση και αναπαραγωγή της χρονικής μεταβλητότητας των ακροτάτων.



Εικόνα 6 Ιδιότητες των μεγίστων άνω κατωφλίου (POT) για την χρονοσειρά απορροών του ποταμού Po (90 έτη) και 1000 συνθετικές χρονοσειρές από στοχαστικό μοντέλο δομής HK: (a) κατανομή συχνότητας εμφάνισης των μεγίστων ανά χρόνο, (b) πιθανότητα υπέρβασης της αθροιστικής έντασης μεγίστων στο έτος, (c) σχέση του αριθμού μεγίστων ανά έτος με τη μέση έντασή τους, και (d) κατανομή της χρονικής διάρκειας ομάδας μεγίστων.

ΚΛΙΜΑΤΙΚΕΣ ΠΡΟΒΛΕΨΕΙΣ ΤΑΣΕΩΝ ΜΕΛΛΟΝΤΙΚΩΝ ΒΡΟΧΟΠΤΩΣΕΩΝ

Το κεφάλαιο 8 εξετάζει τη θεωρητική και εμπειρική βάση της μεθοδολογίας μοντελοποίησης και πρόβλεψης κλιματικών τάσεων στη βροχόπτωση. Καθώς η θεωρία της ανθρωπογενούς κλιματικής αλλαγής προβλέπει την εντατικοποίηση του κύκλου του νερού και των ακροτάτων στο μέλλον, ένα μεγάλο μέρος της σύγχρονης έρευνας για τα ακρότατα στην υδρολογία περιστρέφεται γύρω από τη ντετερμινιστική μελέτη των χρονικών αλλαγών των ακροτάτων σε κλιματική κλίμακα, όπως φαίνεται και στην Εικόνα 7.

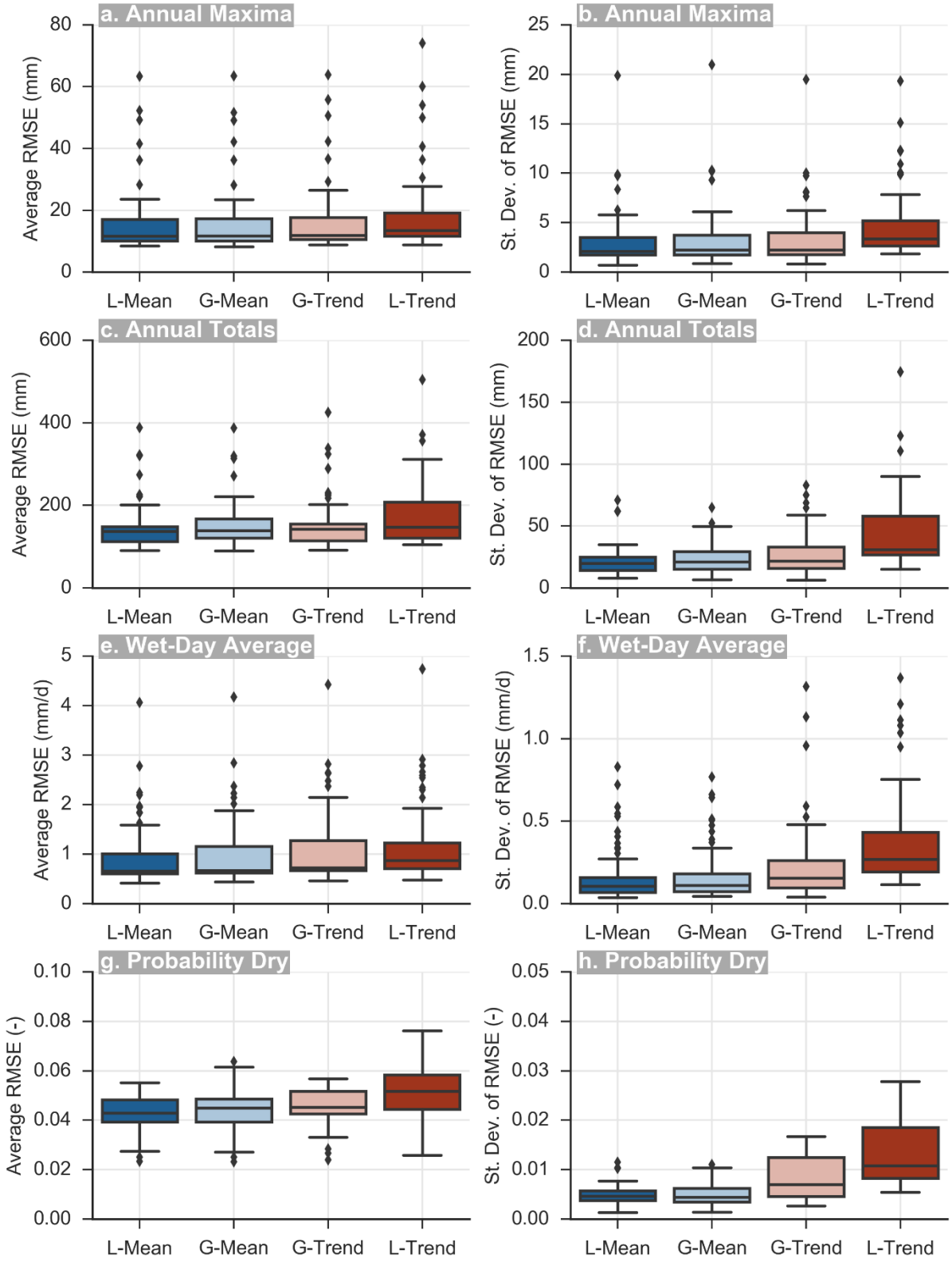


Εικόνα 7 Χρονική μεταβολή και κινούμενος μέσος όρος του λόγου εμφάνισης της λέξης 'trends' σε δημοσιεύσεις του Google Scholar που περιέχουν ήδη το συνδυασμό λέξεων 'precipitation', 'hydrology' και 'extremes'.

Ενώ η ανάλυση ιστορικών τάσεων ως προς την 'στατιστική σημαντικότητα' τους έχει κυριαρχήσει στη βιβλιογραφία, η αξιολόγηση της προγνωστικής ικανότητάς τους ως μοντέλων δεν έχει αξιολογηθεί, παρά την προφανή σημασία που ενέχει για τον μελλοντικό σχεδιασμό έναντι κινδύνων. Αυτή η έρευνα επανατοποθετεί το πρόβλημα της αξιολόγησης τάσεων, ως πρόβλημα επιλογής μοντέλου που προσανατολίζεται στην αναγνώριση του μοντέλου με τις καλύτερες προγνωστικές ιδιότητες, το οποίο δεν είναι ούτε ισοδύναμο με το «πραγματικό» μοντέλο ούτε με το μοντέλο που εξηγεί καλύτερα την περίοδο βαθμονόμησης.

Για το σκοπό αυτό, εισάγεται ένα συστηματικό πλαίσιο επικύρωσης των προγνώσεων των τάσεων μέσω της σύγκρισης του σφάλματος πρόβλεψης (RMSE) με αυτό που λαμβάνεται από απλούστερα μοντέλα μέσου όρου. Το μήκος της χρονοσειράς επιμερίζεται σειριακά σε περιόδους βαθμονόμησης και 30ετίες επικύρωσης και αξιολογούνται οι προγνώσεις με βάση τις τοπικές τάσεις (local trends), τις τάσεις από όλο το μήκος της χρονοσειράς (global trends), τον τοπικό μέσο όρο (local mean) και τον ολικό μέσο όρο (global mean). Τα μοντέλα αξιολογούνται ως προς τα σφάλματά τους στις προβλέψεις της ετήσιας μέγιστης, συνολικής και μέσης βροχόπτωσης καθώς και της ετήσιας πιθανότητας μη βροχής για 30 έτη μεταγενέστερα της περιόδου βαθμονόμησης των μοντέλων. Τα αποτελέσματα όπως προκύπτουν από το σύνολο των 60 σταθμών άνω των 150 ετών

δείχνουν ότι τα μοντέλα κατατάσσονται από το καλύτερο στο χειρότερο ως εξής: τοπικός μέσος όρος, ολικός μέσος όρος, ολική τάση και τοπική τάση (Εικόνα 8).



Εικόνα 8 Κατανομή του μέσου σφάλματος πρόβλεψης (RMSE) και της τυπικής του απόκλισης όπως εκτιμήθηκε από την επικύρωση των 4 προγνωστικών μοντέλων σε διαδοχικούς επιμερισμούς του συνολικού μήκους των χρονοσειρών και για τους 60 σταθμούς βροχόπτωσης.

Η ξεχωριστή εξέταση της πλέον πρόσφατης περιόδου 30 ετών για κάθε σταθμό επιβεβαίωσε επίσης την παραπάνω κατάταξη των μοντέλων, ενώ καμία συγκριτική αύξηση της προβλεψιμότητας δεν εντοπίστηκε για το μοντέλο τοπικής τάσης κατά την κοινή τελευταία περίοδο (1980-2009). Τα αποτελέσματα δείχνουν ότι η μελλοντική μεταβλητότητα της βροχόπτωσης προβλέπεται κατά μέσο όρο καλύτερα από τα πιο φειδωλά μοντέλα μέσου όρου, γεγονός το οποίο υποστηρίζεται και πειραματικά μέσω προσομοιώσεων έμμονων ανελίξεων. Για ίδιο μήκος χρονοσειράς στην περίπτωση ισχυρής εμμονής, ο τοπικός μέσος έχει καλύτερες ιδιότητες πρόβλεψης από τον ολικό. Σε σχέση με χρονικά ανεξάρτητες ανελίξεις, οι έμμονες ανελίξεις χαρακτηρίζονται από αυξημένη μεταβλητότητα σε μεγάλες κλίμακες. Χωρίς επίγνωση αυτής της ιδιότητάς τους, η μεταβλητότητα αυτή μπορεί εύκολα να παρερμηνευτεί ως «συστηματική τάση» ειδικά σε μικρά τμήματα των χρονοσειρών τους.

Σε κάθε περίπτωση προκύπτει ότι εμπειρικά δεν τεκμηριώνεται η πρακτική της προέκτασης κλιματικών τάσεων βροχόπτωσης στο μέλλον. Η εγγενής μεταβλητότητα της βροχόπτωσης καθιστά ούτως ή άλλως δυσχερή την πρόβλεψή της σε μεγάλες κλίμακες, πολύ περισσότερο δε όταν η πολυπλοκότητα των προγνωστικών μοντέλων αυξάνεται.

ΕΠΙΛΟΓΟΣ ΚΑΙ ΚΑΤΕΥΘΥΝΣΕΙΣ ΠΕΡΑΙΤΕΡΩ ΕΡΕΥΝΑΣ

Ο κεντρικός στόχος αυτής της διατριβής είναι η στοχαστική διερεύνηση και μοντελοποίηση της χρονική μεταβλητότητας και εξάρτησης των ακροτάτων από εποχικές έως κλιματικές κλίμακες. Βασική καινοτομία αποτελεί η αναγνώριση των χρονικών συμπεριφορών των ακροτάτων και η οργανική σύνδεσή τους με τις εγγενείς ιδιότητες της μητρικής υδρολογικής διεργασίας. Μια τέτοια προσέγγιση δημιουργεί ένα νέο πρίσμα κατανόησης της δυναμικής των υδρολογικών ακροτάτων που μπορεί να βελτιώσει σημαντικά την αντίληψη του σχετικού κινδύνου στο χρόνο και να χρησιμοποιηθεί για την αναβάθμιση πρακτικών μετριασμού της υδρολογικής διακινδύνευσης.

Περαιτέρω έρευνα είναι απαραίτητη για τον εντοπισμό κοινών δομών εξάρτησης ακροτάτων της βροχόπτωσης και απορροής σε πολλαπλές χωροχρονικές κλίμακες. Η συμπερίληψη των πρόσφατα αναπτυγμένων εκτιμητριών ροπών υψηλής τάξης στη μεθοδολογία εκτίμησης ακροτάτων (*k*-moments; Koutsoyiannis, 2019c) είναι εξίσου κρίσιμη και αναμένεται να βελτιώσει τη μοντελοποίησή τους σε ρεαλιστικές προ-ασυμπτωτικές συνθήκες, όπως αυτές που συναντώνται στην υδρολογία. Τέλος, είναι σημαντικό πλέον της κατανόησης της στοχαστικότητας του υδρολογικού κινδύνου, να μελετηθεί και η στοχαστικότητα της ανθρώπινης χρονικής έκθεσης και ευπάθειας (vulnerability) έναντι του κινδύνου. Νέοι τύποι δεδομένων που ενσωματώνουν πληροφορίες για τη διεπαφή ανθρώπου και υδατικού περιβάλλοντος καθίστανται ολοένα και περισσότεροι διαθέσιμοι και θα μπορούσαν να συνδράμουν προς την ολιστικότερη κατανόηση της υδρολογικής διακινδύνευσης.

1. INTRODUCTION

1.1 Motivation

The term 'extremes' signifies rare and consequential events that are difficult or even impossible to predict from past experience. By analogy, hydrological extremes pertain to the occurrence of rainfall and streamflow of *least* expected properties that are able to place human-water systems under severe stress. In particular, extreme rainfall and flooding can damage the built environment, including water-related and civil-engineering infrastructure, may disrupt financial activities triggering economic losses, while pose direct threats to public health. Naturally so, their understanding and mitigation of associated risk have always been a challenge to society and science. In the past decades though, this challenge has become ever prominent as the adequacy of the conventional perception of risk has been radically questioned itself (Hall et al., 2014).

On the one hand, worldwide media coverage of disasters has increased the availability of examples of catastrophic hydrological events (Barredo, 2007), causing growing concerns on intensification of extremes. These concerns have been scientifically corroborated by the anthropogenic climate change hypothesis. The latter suggests that the increased anthropogenic emissions of greenhouse gases over the past decades have induced systematic changes in the climate dynamics that lead to the intensification of the water cycle (IPPC, 2014; Fifth Assessment Report AR5). On this basis, it has been argued that a radical reformulation of assumptions and modelling practices would be required in order to explicit model the presence of new deterministic drivers (Milly et al., 2008). Although this position has been largely debated in hydrology (Cohn and Lins, 2005; Montanari and Koutsoyiannis, 2014; Koutsoyiannis and Montanari, 2015; Serinaldi et al., 2018; Koutsoyiannis, 2020a), the associated expectations of intensification of extremes in the future are omnipresent in the literature.

On the other hand, it is widely accepted that the risk dynamics may have altered during the past century as a result of systematic changes in human exposure and vulnerability to extreme events. In this respect, it has been found that human exposure to flooding, stemming from living in proximity to the river network, has consistently increased over the years (Ceola et al., 2014). At the same time, vulnerability to extreme events is further exacerbated by high population density and uncontrolled urbanization. In Greece alone, over than 200 fatalities due to extreme flooding have been reported since 1960, with the majority of them concentrated in the highly urbanized area of Attica. At the global scale, the series of the 1999 rainfall-induced flood events in the densely populated Vargas state of Venezuela, is considered one of the worst water-related disasters worldwide resulting in more than 15 000 fatalities. In a later study, it was found that the occurrence of such rainfall was extreme but not implausible based on the historical record and could have been anticipated if a probabilistic framework accounting for uncertainty was applied (Coles et al., 2003).

The latter case study highlights the essential role exerted by hydrology in terms of frequency estimation of extremes and hence, risk preparedness. The underestimation of both the design quantiles and the frequency of occurrence of the

extremes is unfortunately, a common factor in cases of unprecedented hydrologic disasters (Mimikou and Koutsoyiannis, 1995; Coles et al., 2003; Koutsoyiannis et al., 2012; Ntigkakis et al., 2018). Furthermore, increasing concerns are expressed regarding the possible underestimation of the probability of failure of aging infrastructure and large-scale engineering projects as numerous engineering disasters triggered by extreme rainfall and flooding have been reported in recent years. These include dam accidents and disasters, bridge collapses and severe damages to transport infrastructure, including roads and railways, that have caused massive economic losses and human fatalities (Wardhana and Hadipriono, 2003; Serra-Llobet et al., 2013; Koskinas et al., 2019; Kellermann et al., 2019; Pizarro et al., 2020).

Therefore, the improvement of hydrological design and risk mitigation emerges as a high societal priority. In this respect, it is argued that before considering the case for global trends or external drivers, the need to advance the understanding and modelling of extreme hydrological events, starting with rainfall, is imperative. To this aim, it is important to first and foremost revisit the classic assumptions governing the study of extremes in light of new empirical evidence.

1.2 Framing the research question

The realm of civil and environmental engineering is perhaps the scientific field most organically connected to the need for probabilistic estimation of extremal properties. Although the first statistical approach of hydrological extremes is found in Fuller (1914), it is widely acknowledged that it was the work of Gumbel (1941), who placed the probabilistic study of extremes at the core of hydrological science. Among a series of preceding theoretical works, Gumbel's work, popularized by his renowned book (1958), was the most influential to the engineering cycles, triggering a domino of studies in hydrology that provided probabilistic grounds to analysis of extremes in the decades to follow. Some landmark examples of probabilistic extreme value modelling include the estimation of the extreme sea level surge for the sea dike projects in the Netherlands (de Haan, 1994), the construction of consistent rainfall intensity-duration-frequency curves for engineering design (Koutsoyiannis et al., 1998) as well as the estimation of regional flood frequency in the US (Stedinger and Griffis, 2008). A major advance has been the shift from deterministic approaches towards fully probabilistic modelling of rainfall and streamflow extremes (Koutsoyiannis, 1999, 2004). The field of applications has since become so wide that the probabilistic methodologies of extremes are now integral to the National Flood Insurance Program in the US (FEMA, 2016).

Contrary to inference for regular quantities which involves estimation within the range of available data, inference for extremes is synonymous to extrapolation to the range of unobserved behaviours. The latter is heavily dependent on assumptions regarding their properties. Independence is the most central assumption in extreme value theory applications, encountered in most engineering textbooks. In Gumbel's words (1958):

> *The observations from which the extreme values are drawn ought to be independent. This condition may be met in an experimental setup. However, it is seldom met in natural observations. Still, the asymptotic theory gives a very*

good fit for such observations, because it is only the asymptotic behaviour of the initial distribution which counts.

In other words, Gumbel himself acknowledges the non-fulfilment of the independence assumption in observations, but suggests that the latter should be irrelevant to the estimation of the magnitude of extremes.

Often the nuance of such statements is missed by standard practice and the asymptotic arguments for the distribution of extremes are misinterpreted in favour of the perception that observed extremes are in fact independent. The adherence to the independence assumption is further supported by the widespread use of statistics in hydrology (Koutsoyiannis, 2019a), where the independence assumption is central. Yet while in the field of statistics dealing with controlled experiments and idealized theoretical conditions, independence is a tenable assumption, it is rather a misplaced one in the case of real-world unique observations. In hydrology, ever since the works of Hurst (1951), the presence of long-range temporal dependence in observations, else known as persistence of Hurst-Kolmogorov dynamics, has been widely acknowledged (e.g. Koutsoyiannis, 2003; Montanari, 2003; O'Connell et al., 2016; Dimitriadis, 2017). By now, it has been theoretically and empirically established that presence of dependence in the data renders the application of classic statistics erroneous, introducing bias in the estimation of moments and quantiles, and inflates confidence and prediction intervals. Thus, unawareness of dependence is bound to lead to severe underestimation of uncertainty, which in the case of long-range dependence could be of the order of magnitude (Cohn and Lins, 2005; Koutsoyiannis, 2005; Hamed, 2008; Koutsoyiannis and Montanari, 2007; Lombardo et al., 2014; Serinaldi et al., 2018).

Notably, the independence assumption in hydrology is contradicted by empirical evidence of extremes at small scales, as in the case of long-duration rainfall and flood events and seasonal clustering of events, and more rarely, even at greater scales, as in the case of multi-year droughts and flood-rich/flood-poor climatological periods (Hall et al., 2014; Merz et al., 2016). Yet for a combination of the above reasons, reinforced by the limited availability of long records and the peculiarities of extremes from dependent processes hiding their properties, the independence assumption continues to dominate the modelling of hydrological extremes.

Stochastics, integrating statistics, probability calculus and the theory of stochastic processes, offer a self-contained and powerful framework for the study and modelling of uncertain processes, including extremes deviating from idealized randomness (Koutsoyiannis, 2020b). This thesis aims to employ both stochastics and real-world empirical evidence to investigate the temporal properties and dependence dynamics of extremes, revisiting the relevance that the common assumptions thereof, i.e. of being independent and identically distributed (IID), bear for hydrological practice. In this respect, primary focus is placed upon integrating the understanding and modelling of extremes to that of the parent process by exposing the hidden links between the two, from seasonal to climatic scales. To this end, a rare dataset of long-term observational records is compiled and two specific objectives are developed: (a) to efficiently characterize the temporal properties and dependence dynamics of observed extremes from seasonal to climatic scales, and (b) to formulate hydrologically relevant and parsimonious modelling frameworks to reproduce them.

The main body of the analysis prioritizes the investigation of rainfall extremes, the temporal properties of which are generally understudied compared to the streamflow ones. Dependence dynamics of streamflow extremes are explored as well with two purposes: (a) to identify dependence dynamics of high operational relevance, as in the case of seasonal predictability of streamflow extremes, and (b) to gain additional hydrological insights by comparing their dependence properties to those of the rainfall process.

1.3 Structure of thesis

The remainder of this thesis is structured in eight chapters. More specifically, the thesis outline is presented below.

Chapter 2 presents the basic stochastic framework that is employed throughout the thesis and provides a thorough literature review on modelling extremes. It also identifies open questions in the literature and highlights present challenges.

Chapter 3 revisits the subject of long-term persistence in the annual rainfall process via a global database. Although persistence in the streamflow process is well studied, persistence in the rainfall process is much less acknowledged. The findings of this chapter form the empirical basis for tracing dependence dynamics in rainfall extremes through a consistent stochastic framework.

Chapter 4 deals with dropping the independently distributed (ID) assumption in modelling of rainfall extremes, by accounting for their seasonality. It resolves the open question of optimal identification of extreme rainfall seasonality by introducing a new model selection method for the characterization and modelling of seasonal rainfall extremes.

Chapter 5 switches the focus to the seasonal dynamics of streamflow extremes, exploiting short-term dependence for predictability of high and low flows. A large database of rivers is explored to identify potential for seasonal predictability and investigate the presence of physical drivers enhancing it.

Chapter 6 deals with the propagation of persistence dynamics from the parent process to its extremes. It exposes the shortcomings of existing indexes in revealing persistence in extremes and addresses this gap by introducing a multi-scale probabilistic dependence characterization for extremes. An empirical investigation of dependence properties of rainfall extremes is carried out using a long-term dataset.

Chapter 7 revisits theoretical results of Extreme Value Theory and approaches open questions pertaining to persistent processes through extensive simulations. The manifestations of dependence dynamics in the annual patterns of rainfall and streamflow extremes are investigated while the performance of the second-order stochastic framework is evaluated in terms of reproducing them.

Chapter 8 completes the body of analysis on the facets of extremal dependence. A novel methodological approach is introduced to evaluate the relevance of trend projections to the future, by examining the statistics of their predictive performance in the past. Classic modelling principles and current approaches are discussed and tested against both data and simulations.

Chapter 9 revisits the thesis motivation and objectives and outlines the most important contributions. A discussion on future research directions is provided.

Appendix A contains background information, sources and acknowledgments for the compilation of the long-term rainfall dataset, while **Appendix B** contains supplementary material for Chapters 5, 6 and 8. A list of related publications is provided at the end.

1.4 Innovation points

The thesis formulates innovative modelling frameworks, the specific contributions of which are summarized in Chapter 9. Methodologically, the most fundamental innovation points are the following:

(a) The research subject per se is an innovation point, as the mainstream practice is to focus solely on the marginal distribution of extremes, ignoring their temporal properties. It is shown that their temporal behaviours critically affect our perception of risk and may affect various design properties.

(b) The introduction of the notion of scale in the characterization of extremal dependence is novel on its own, as it is usually the lagged correlation that is explored.

(c) The attempt to link extremal properties to the second-order properties of a process, both in terms of characterization and modelling is also a novelty of this work. Usually, the extremal properties are studied under asymptotic arguments, which allow disregarding the properties of the parent process.

(d) The systematic evaluation of the predictive skill of trends is performed for the first time to the author's knowledge. Although the evaluation of models by their predictive performance is established in hydrological literature, it has not been employed in such context so far.

A few other points are considered innovative with respect to current practice, although they are in fact revisiting classic approaches. These include:

(e) The compilation of a long-term observational dataset (>150 years) for studying rainfall extremal properties. There is a well-justified increasing interest in studying global databases, as these are essential for identifying common properties and investigating their spatial distribution as well. Yet, such databases are mostly comprised of short and medium-length records (<100 years) that are insufficient for the study of long-term variability of extremes.

(f) The use of parsimony as the modelling principle of choice. The parsimony principle is embedded in various sophisticated statistical procedures, however its meaning is often obscured by the standardized and complicated character of the former. Herein, the virtue of parsimony is highlighted, in an original fashion, directly linked to prediction.

2. STOCHASTIC DEPENDENCE DYNAMICS FROM THE PARENT PROCESS TO THE EXTREMES: A REVIEW

This chapter lays the theoretical foundations of the thesis. It starts by presenting the fundamental theoretical concepts of the stochastic framework and the analytical tools for second-order dependence which are used throughout the thesis. It also presents a thorough review on the history of extreme value theory and highlights some of its less known results, related to extremal dependence. It provides a critical overview of common approaches in modelling extremes in hydrology and beyond, and identifies open theoretical questions and challenges.

2.1 Definitions in a stochastic framework

In this section, we define the pivotal probabilistic and stochastic concepts that are ubiquitous throughout the analysis. For a comprehensive presentation of stochastic theory reader is referred to Papoulis (1991) and Koutsoyiannis (2020b).

2.1.1 Random variables, stochastic processes and timeseries

A random variable is a function that maps outcomes of experiments from the non-empty set Ω, else called set of elementary events or states, to numbers. A formal definition of the concept, along with the axiomatic definition of probability, is owed to Kolmogorov (1933). To distinguish random variables from regular variables, we underline them following the Dutch convention. A stochastic process is then an arbitrarily large family of random variables $\underline{x}(t)$ (Papoulis, A., 1991). These variables are indexed by t, which in our case, represents time, either from the discrete set of integers $\mathbb{Z}$ (resulting to a discrete-time stochastic process), or from the continuous set of real numbers $\mathbb{R}$, (resulting to a continuous-time stochastic process). Following Koutsoyiannis (2020b) we denote a continuous time stochastic variable by $\underline{x}(t)$, and a discrete one by $\underline{x}_\tau$. The stochastic variables per se can be either discrete (discrete-state stochastic process), as in the wet or dry day, or continuous, e.g. rainfall amount (continuous-state stochastic process). The index can also be multidimensional by e.g. referring to space. A realization $x(t)$ of stochastic process $\underline{x}(t)$ is called a timeseries. Both its observation and simulation take place in discrete time, but for theoretical and physical consistency, it is desirable to deduce the theoretical properties thereof in continuous time. Then the discrete-time representation is derived from the integration of the continuous-time process as (Koutsoyiannis, 2020b):

$$\underline{x}_\tau := \frac{1}{D}\int_{(\tau-1)D}^{\tau D} \underline{x}(u)\,\mathrm{d}u \tag{1}$$

where $\tau \in \mathbb{Z}$ represents the continuous time interval $[(\tau-1)D, \tau D]$ and D is the time step. It is important to distinguish between the notion of a timeseries and the one of a stochastic process, as the former is a finite sequence of numbers (observations), whereas the latter is a family of infinite stochastic variables.

2.1.2 Distribution function and moments

The distribution function of a random variable $\underline{x}(t)$ is:

$$F(x;t) := P\{\underline{x}(t) \leq x\} \tag{2}$$

This is called the *first* order distribution function of the process, and generalizes for the *n*-th order as:

$$F(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) := P\{\underline{x}(t_1) \leq x_1, \underline{x}(t_2) \leq x_2, \dots, \underline{x}(t_n) \leq x_n\} \tag{3}$$

A stochastic process is fully determined if we know the *n*th order distribution, or else joint distribution, for any *n*. The most important moments of a process that we utilize herein are the following:

i. The process mean:

$$\mu(t) := \mathrm{E}[\underline{x}(t)] = \int_{-\infty}^{+\infty} x f(x;t)\,\mathrm{d}t \tag{4}$$

ii. The process variance:

$$\gamma_0(t) := \mathrm{var}[\underline{x}(t] = \int_{-\infty}^{+\infty} (x-\mu(t))^2 f(x;t)\,\mathrm{d}t \tag{5}$$

iii. The process autocovariance:

$$c(t;h) := \mathrm{cov}[\underline{x}\big(t), \underline{x}(t+h)\big] = \mathrm{E}[\big(\underline{x}(t) - \mu(t)\big)\,\big(\underline{x}(t+h) - \mu(t+h)\big)] \tag{6}$$

iv. The process autocorrelation coefficient:

$$r(t;h) := \mathrm{corr}[\underline{x}\big(t), \underline{x}(t+h)\big] = \frac{c(t;h)}{(\gamma_0(t)\gamma_0(t+h))^{1/2}} \tag{7}$$

v. The process coefficient of skewness:

$$C_s(t) := \frac{\mu_3(t)}{(\gamma_0(t))^{3/2}} \tag{8}$$

where $\mu_3(t) := \int_{-\infty}^{+\infty} (x-\mu(t))^3 f(x;t)\,\mathrm{d}t$, the third central moment of the process.

vi. The process coefficient of kurtosis:

$$C_k(t) := \frac{\mu_4(t)}{(\gamma_0(t))^2} \tag{9}$$

where $\mu_4(t) := \int_{-\infty}^{+\infty} (x-\mu(t))^4 f(x;t)\,\mathrm{d}t$, the fourth central moment of the process.

2.1.3 Stationarity, cyclostationarity and ergodicity

A process is defined as stationary if its statistical properties are invariant with respect to a shift of time origin, i.e. $\underline{x}(t)$ and $\underline{x}(t')$ have the same nth order distribution for any t and t' (Kolmogorov, 1931; Khintchine, 1933). Kolmogorov (1947) further defined the special case of wide-sense stationarity in which the mean is constant and the autocovariance depends only on the time lag. In these case, the time index t in equations (4)–(7) could be dropped. On the contrary, a nonstationary process is one whose statistical properties are deterministic functions of time. Recalling the distinction between a process and a timeseries, it is clear that (non)stationarity is a property of a process and it cannot be inferred from a timeseries alone.

A nonstationary process could have some of its properties depend on time in a periodic manner, in which case it is called cyclostationary, and by adequate modifications can be modelled by stationary models. Such is the case of a process exhibiting pronounced seasonality, examples of which are discussed in Chapter 4.

A central problem in the study of stochastic processes is the estimation of their parameters from data. The fundamental property of processes that allows estimation from data and is tacitly implied in all timeseries analyses, is ergodicity (Papoulis, A., 1991). Ergodicity is a wider property of dynamical systems which can also be defined in the context of stochastic processes based on the ergodic theorem (Birkhoff, 1931; Khintchine, 1933) as follows (Koutsoyiannis, 2010). A stochastic process $\underline{x}(t)$ is ergodic if the time-average of any integrable function $g(\underline{x}(t))$ equals the true expectation (ensemble average) as time tends to infinity, i.e. for a continuous process:

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T g(\underline{x}(t))\, \mathrm{d}t \;\; = \mathrm{E}[g\left(\underline{x}(t)\right)] \tag{10}$$

and for a discrete-time process:

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T} g(\underline{x}_\tau) \;\; = \mathrm{E}[g\left(\underline{x}_\tau\right)] \tag{11}$$

The equation of the true expectation (right-side), i.e. a number, to the ensemble average (left-side), i.e. a stochastic variable, implies zero variance of the latter as the sample grows infinite, which is precisely the condition for ergodicity (Koutsoyiannis, 2020b). If $g(\underline{x}(t)) = \underline{x}(t)$, then the fulfillment of equations (10)-(11) makes a process mean-ergodic, while other specifications exist depending on the type of the function g, e.g. covariance-ergodic (Papoulis, A., 1991). The relationship of stationarity and ergodicity is a delicate one which also depends on the systems dynamics being stochastic or deterministic. For a stochastic system, the two do not necessarily coincide, but it is possible and practical to formulate a stationary model that is ergodic too, considering that a nonstationary model is generally nonergodic (Koutsoyiannis, 2010; Koutsoyiannis and Montanari, 2015; Montanari and Koutsoyiannis, 2014). For a deterministic process, the two are theoretically connected (see Mackey, 2011; and discussion in Koutsoyiannis, 2020b).

2.1.4 Dependence in time

Unlike classical statistics dealing with samples of measurements and experimental outcomes which can be appropriately designed in order to be modelled as independent random variables, the study of timeseries introduces the notion of dependence in time, hence change in time. This is precisely the focus of stochastics, as the mere definition of a stochastic process involves its time evolution (Kolmogorov, 1931).

In hydrology, the study of dependence in time has a long history dating back to the works of Hurst (1951), who observed that the annual behaviour of the level of the Nile river deviated from that of a purely random process. It has since become a very active topic in hydrology even under deterministic interpretations, discussed in Chapter 8. The stochastic patterns of manifestation of dependence in the rainfall and the runoff process form the central subject of the thesis, and parts of the relevant theory are outlined throughout all Chapters. In the following section, we present and summarize the stochastic methodology for quantification of dependence in time for hydrological problems that forms the reference framework for the main body of analysis.

2.2 Second-order properties, scaling laws and HK dynamics

The complete determination of a stochastic process requires knowledge of its nth order properties. In terms of estimation from data, this is almost impossible considering the bias of higher-order classical moments (Lombardo et al., 2014). In this respect, the second-order moments of a stochastic process, i.e. the autocorrelation, autocovariance and functions thereof, provide robust information on their dependence, have lower estimation bias, and are useful in the simulation process. Not surprisingly, they are the most extensively used tool in stochastics (Papoulis, A., 1991).

A common characteristic of second-order properties is their association with asymptotic power laws as $t \to 0$ or $t \to \infty$, or else scaling behaviour (Koutsoyiannis, 2014). Power laws are functions of the form $f(t) \propto t^b$ and can be visualized in the form of a straight line with slope b from a doubly logarithmic plot of $f(t)$ on t. A power-law valid over the entire domain is called *simple scaling*, while power-laws valid in the domain of $t \to 0$ define *local* behaviour and the case $t \to \infty$, define *global* behaviour (Koutsoyiannis, 2020b). Power laws have been studied in many domains, being popularized by Mandelbrot (1983), although first mathematically described by Kolmogorov (1940).

In hydrology, scaling of the second-order properties in long-time horizons is a ubiquitous behaviour, which was first observed in the Nilometer data by Hurst (1951), and hence is also known by the term Hurst behaviour/phenomenon (O'Connell et al., 2016). In order to give credit to the mathematical representation by Kolmogorov, in hydrology it is also known as the Hurst-Kolmogorov dynamics (HK dynamics; Koutsoyiannis, 2010). In the wider literature, it is known as well as long-range dependence, long-term persistence, and long-memory (Beran, 1994). The latter term though has been disputed on the basis that the induced long-range dependence is a

product of long-term change instead of the result of a long-memory physical mechanism (Klemeš, 1974; Koutsoyiannis, 2011b).

2.2.1 Climacogram and climacogram-based modelling

A comprehensive characterization of a process's second-order scaling properties can be achieved by inspecting its variance behaviour when the process is averaged, or aggregated, over different scales. The function of the variance of the averaged process versus the scale is called the climacogram, while the function of the cumulative process versus the scale is called the cumulative climacogram (Koutsoyiannis, 2010). The climacogram of a process $\underline{x}(t)$ is defined as:

$$\gamma(k) := \operatorname{var}\left[\frac{\underline{X}(k)}{k}\right] = \frac{\Gamma(k)}{k^2} \tag{12}$$

where $\Gamma(k)$ is the cumulative climacogram, and $\underline{X}(k)$ is the process $\underline{x}(t)$ aggregated at timescale k:

$$\underline{X}(k) := \int_0^k \underline{x}(t)\,\mathrm{d}t \tag{13}$$

or for a discrete-time process, with climacogram γ_κ:

$$\underline{X}_\kappa := \underline{x}_1 + \underline{x}_2 + \cdots + \underline{x}_\kappa \tag{14}$$

The climacogram is theoretically equivalent to other second-order properties, namely the autocovariance, autocorrelation and the power-spectrum, but it is advantageous estimation-wise for having superior properties in terms of bias, discretization errors, and sampling uncertainty (Dimitriadis and Koutsoyiannis, 2015). For these reasons, it is the basic tool employed herein for second-order characterization.

The theoretical climacogram differs among processes with different second-order dependence structure. For three key types of stochastic processes the following hold (Koutsoyiannis, 2020b). In case of an independent White-noise process in discrete time, the climacogram is a function of the variance of the process σ^2 and the scale κ:

$$\gamma_\kappa := \frac{\sigma^2}{\kappa} \tag{15}$$

which generalizes for the continuous-time by changing the scale to a real number $k := \kappa D$

$$\gamma(k) := \frac{\sigma^2 D}{k} \tag{16}$$

For a continuous-time process $\underline{x}(t)$ with variance $\lambda = \gamma_0 = \gamma(0) = c(0)$ and short-range dependence, i.e. an AR(1), or Markov process:

$$\gamma(k) := \frac{2\lambda}{k/a}\left(1 - \frac{1 - e^{-\frac{k}{a}}}{k/a}\right) \tag{17}$$

where α is parameter with units of time. In the case of a continuous process with asymptotic scaling at ∞, the climacogram is:

$$\gamma(k) := \lambda \left(\frac{a}{k}\right)^{2-2H} \tag{18}$$

where α and λ are scale parameters, with dimensions $[t]$ and$[x^2]$ while H is the so-called Hurst parameter ranging in the interval (0,1). The latter equation is the definition for a Hurst-Kolmogorov process. The case $H = 0.5$ corresponds to an independent process, while for $0.5 < H < 1$ the process is persistent and for $0 < H < 0.5$ anti-persistent.

From equations (16) and (18) it is evident that when the scale tends to zero the process's variance reaches infinity, which is not plausible for natural process, as an infinite variance process would require infinite energy to materialize. In order to remedy this shortcoming, and improve flexibility of the model for dependence in shorter time scales, the filtered Hurst-Kolmogorov process is developed with several climacogram types (Koutsoyiannis, 2017). The generalized Cauchy-type climacogram is:

$$\gamma(k) := \lambda \left(1 + (\frac{k}{a})^{2M}\right)^{\frac{H-1}{M}} \tag{19}$$

where M is an added dimensionless parameter which controls the local scaling of the process (fractal behaviour), named M in honor of Mandelbrot (Koutsoyiannis et al., 2018). Values of $M < \frac{1}{2}$ indicate a rough process, while $M > \frac{1}{2}$ indicate a smooth process. For more, on the bounds of scaling the reader is referred to Koutsoyiannis (2017). The usefulness of this parameterization in simulation is discussed next in Section 2.4.

2.2.2 Scaling in time by the entropic view: from predictability to uncertainty

A counter-intuitive characteristic of dependence in time is its non-equivalence to predictability, even more its association with increased unpredictability at greater time scales. The biased perception of dependence in favour of predictability, reflected by the term 'memory', is fortified by the dominance of autocorrelation-based models in the literature. Autocorrelation is intuitive for prediction purposes but does not expose uncertainty, on the contrary to the variance-based characterization of the climacogram.

A rigorous way to investigate the case for (un)predictability is through the unifying notion of entropy, which represents degree of uncertainty or ignorance (Papoulis, A., 1991). The conditional entropy, i.e. the uncertainty about the future when the past is observed, is directly linked to predictive uncertainty, while the difference between entropy and conditional entropy equals the information gain. Koutsoyiannis (2005, 2011) showed that the HK dynamics is a product of conditional

entropy maximization at large time scales, whereas the AR(1) model maximizes conditional entropy at small time-scales (k=1,2). An increase in model autocorrelation, signifies a decrease of conditional entropy at lower scales (1,2), thus improves predictability, whereas at higher scales the exact opposite is true. As the timescale increase, the conditional entropy of all models also decreases, yet at a different rate compared to the unconditional. For a scaling process (HK dynamics) the information gain remains constant with the scale, as does the autocorrelation function, due to the same rate of decrease of the conditional and unconditional entropies. However, this does not imply greater predictability, as in fact, the conditional entropy still remains greater than the case of an AR(1) model (Koutsoyiannis, 2005). This view of predictability is particularly relevant when dealing with climate, which represents the average weather at scale k=30 years (Koutsoyiannis, 2010). For a given scale, it is the role of time window that becomes critical in determining the predictability horizon, as highlighted in Dimitriadis et al. (2016). The relation of dependence with predictability is examined both at short time-scales, in Chapter 5, and at climatic-scales in Chapter 8.

2.3 Dependence in extremes: theory and diagnostics

2.3.1 The development of classic extreme value theory

Before considering the case of dependence, it is worth recapitulating the fundamental results of extreme value theory which is now well established. If $\underline{y}_1, \underline{y}_2, \dots, \underline{y}_n$ is a sequence of identically and independently distributed (IID) random variables, then the maximum of them, i.e. the largest order statistic, $\underline{x}_n := \max(\underline{y}_1, \underline{y}_2, \dots, \underline{y}_n)$ has the following probability distribution function:

$$F_{x_n}(x) := (F_y(x))^n \tag{20}$$

Results concerning the asymptotic behaviour of this distribution as $n \to \infty$ were obtained in the early 20th century. Fréchet (1927) was the first to identify the homonymous limiting law, Fisher and Tippett (1928) showed that there are only three possible types of the limiting laws, von Mises (1936) identified sufficient conditions for convergence to the limiting laws and provided a common parameterization, while Gnedeko (1943) set the solid foundations for convergence to the limiting laws under weak conditions. Their results were lately popularized to the engineering community by the prominent book of Gumbel (1958). A detailed presentation of the early history of the contributions is provided in Kotz and Nadarajah (2000). Specifically, the asymptotic theory for extremes states that for extremes from IID random variables, if there exist rescaling constants $a_n > 0$ and b_n, so that for the linearly rescaled maximum $x'_n := \frac{x_n - b_n}{a_n}$ exists a non-degenerate limiting distribution, then this should be of the form:

$$G(x) := F_{x'_\infty}(x) = \exp\left(-\left(1 + \xi\left(\frac{x-b}{a}\right)\right)^{-\frac{1}{\xi}}\right), \text{with } 1 + \xi\left(\frac{x-b}{a}\right) \geq 0 \tag{21}$$

The latter is known as the generalized extreme value (GEV) distribution. In this parameterization, ξ is a shape parameter $\epsilon \mathbb{R}$, $\alpha > 0$ a scale parameter and b a location parameter $\epsilon \mathbb{R}$. The shape parameter is unique and is identical to the tail index of the parent process, but the scale and location parameters depend on n. The first limiting law known as extreme value type I (EV1) is retrieved for $\xi = 0$ and is the well-known Gumbel distribution. The second law, known as extreme value type II (EV2), emerges for $\xi > 0$ and is the Fréchet distribution. The third law (EV3) appears for $\xi < 0$ and is the reverse Weibull distribution, but it is less of interest in applications of extremes, as it yields an upper bound. A parent distribution F is said to belong to a domain of attraction if a linear transformation of its maxima follows one of the three limiting laws (Von Mises, 1936). For instance, light-tailed distribution and heavy-tailed distribution with tail index ξ=0 belong to the domain of EV1, heavy-tailed with tail index ξ>0 to the domain of attraction of EV2, and distributions bounded from above to EV3 (Koutsoyiannis, 2020b). For typical applications of finite n, as in the annual maxima case, equation (21) is only an approximation of the distribution of extremes, and its asymptotic validity could be questioned. Details on the strength of convergence in cases of interest to hydrology can be found in Koutsoyiannis (2004a, 2004b) and Papalexiou and Koutsoyiannis (2013), and are further discussed in Chapters 4 and 7.

The three limiting laws form exactly the class of max-stable distributions, meaning that if G is max-stable then it can be shown that it is of extreme value type, while the EV class has the property to be max-stable (Leadbetter et al., 2012). A max-stable distribution as originally defined by Fréchet (1927) is one which retains the same form under a linear transformation of its maxima, specifically, for any $n \in \mathbb{N}$ and $x \in \mathbb{R}$, there exists $a_n > 0$ and b_n such that:

$$(G(a_n x + b_n))^n = G(x) \tag{22}$$

2.3.2 Extreme value theory under dependence

It is straightforward to see that some sort of restriction of the dependence structure is required in order to obtain an asymptotical result for the type of extremal behaviour. Otherwise it could be assumed that all $\underline{y}_i$ are equal arising from an arbitrary distribution function, in which case the distribution of its extremes $\underline{x}_n$ would be this arbitrary distribution (Leadbetter and Rootzen, 1988). As discussed in the previous section, the IID assumption was fundamental in the early development of the classic extreme value theory. Juncosa (1949) was the first to generalize results in case of non-ideally distributed variables. Some years after a number of publications emerged in the direction of relaxing the independence assumption. Early considerations of the case for dependence originated in the literature in the works of Watson (1954) and Newell (1964) who studied asymptotic results for extreme value series from stationary sequences of m-dependent random variables, i.e. considering events that occur more than time m apart as independent. Berman (1964) identified conditions for asymptotic independence of Gaussian processes and Loynes (1965) showed that m-dependence could be replace by the uniform mixing (else referred to as strong mixing) assumption for the parent process, which was further generalized by O'Brien (1974). The uniform

or strong mixing assumption is first introduced in Rosenblatt (1956) and is defined in probabilistic terms requiring that *any* two events separated in time tend to being independent as the separating time grows larger.

A pivotal result concerning influence of dependence on extreme value theory was obtained by Leadbetter (1974) who studied weaker conditions under which the non-degenerate limits of dependent sequences are still extreme value distributions. In particular, Leadbetter introduced the distributional mixing condition known as $D(u_n)$ condition which states that under weak conditions that exclude long-range dependence only for high level exceedances G is still an extreme value distribution, or equivalently, it is max-stable. This is a condition much weaker than the strong mixing condition which applied to all exceedances (uniform mixing) as presented in Loynes (1965). Examples of moving-maxima processes for which this holds are provided in Berliant et al. (2006a), comprising cases of validity in even weaker conditions of the original $D(u_n)$, i.e. including periodic Markov Chains. These results further establish the use of the EV theory and justify its acceptable empirical performance through the years.

Although dependence under the $D(u_n)$ condition does not challenge the asymptotic validity of the extreme limit theorems, it may affect the choice of the limit distribution. Leadbetter (1974) also introduced a second stricter condition known as $D'(u_n)$ which limits the amount of short-range dependence in high-level exceedances, or else local clustering, by requiring the probability of more than one exceedance in a cluster to be negligible. This implies asymptotic independence and is satisfied in the early results obtained in the literature (e.g. Watson, 1954; Loynes, 1965). Together these conditions ensure that asymptotically the occurrence of exceedances form a Poisson process, while the possibility of clustering of events is limited. If the $D'(u_n)$ does not hold then the exceedances of the threshold can occur in clusters as a compound Poisson process. In particular, Leadbetter (1983) by extending a result of Chernick (1981), showed that the weak mixing $D(u_n)$ condition alone suffices for the asymptotic distribution of extremes from stationary processes to be precisely of the same type of that of an IID sequence with the same marginal distribution. In this case though, dependence affects the parameterization of the limit distribution in terms of the linearly rescaling constants. The quantification of extremal dependence is defined in terms of a constant $\theta \in [0, 1]$, which is called the *extremal index*, and for which the following statement holds. If a process $\underline{y}$ has an extremal index, then for each $\tau > 0$:

(i) there exists $u_n(\tau)$ such that $n(1 - F(u_n(\tau)) \to \tau$, which suggests that the mean number of exceedances is constant as $n \to \infty$, and

(ii) $$P\{\underline{x}_n \le u_n(\tau)\} \to e^{-\theta\tau} \tag{23}$$

which suggests that the extremal distribution converges to a generalization of the limiting form for the IID case, $\tilde{G}^{\theta}$, where $\tilde{G}$ the limiting distribution for the associated IID process $\underline{\tilde{x}}_n$.

If (i) holds and $D(u_n(\tau))$ holds for each τ and $P\{\underline{x}_n \le u_n(\tau)\}$ converges for some $\tau>0$ then (ii) holds for some $\theta \in [0, 1]$ and all $\tau>0$, and therefore the process has an extremal index (Leadbetter et al., 2012). The case of $\theta = 0$ is considered pathological as it leads to $P\{\underline{x}_n \le u_n(\tau)\} \to 1$ for all τ although it may have marginal meaning in specific cases. If $\theta = 1$ this corresponds to the form G takes for IID data, but in case $\theta > 0$, then the (ξ,

a, b) parameters of G are related to the ones of the limiting distribution of the associated independence sequence, $\tilde{G}$ (equation 21) by:

$$\xi = \widetilde{\xi}, \qquad \alpha = \tilde{\alpha}\theta^{\xi}, \qquad b = \tilde{b} + (\theta^{\xi} - 1)\frac{a}{\xi}, \tag{24}$$

Therefore, only the shape parameter remains the same. These results can relate to ones derived from the generalized Pareto distribution for over-threshold exceedances, whose connection to extreme value theory was established by Pickands (1975) while also previously found in an independent study by Balkema and de Haan (1974). For independent processes, $\theta = 1$, although the latter also holds for some cases of dependent processes under the stricter condition $D'(u_n)$ which limits the amount of local clustering, i.e. short-range dependence in high-level exceedances (Leadbetter, 1983a). This condition is equivalent to assuming asymptotic independence. On the contrary, values of θ <1 represent the tendency of exceedances to occur in clusters.

Therefore, the extremal index can be thought as way to link extremal clustering behaviour of the process to its parent dependence structure. It can be shown that the extremal index is related to various properties of clustering of events, e.g. θ^{-1} is the mean size of extremal clusters. Likewise, the extremal index is the reciprocal of the limiting mean number of exceedances in blocks with at least one exceedance, thus it can be estimated as the ratio of the total exceedances of the threshold vs the number of cluster with at least one exceedance, which is known as the blocks estimator (Beirlant et al., 2006). It can be also related to the conditional time between exceedances, as well as the distribution of the maxima of the process, as shown before. Based on these properties, various estimators have been proposed in the literature, as the runs and maxima methods. A comparison of various methods is provided in Ancona-Navarrete and Tawn (2000) who point out the strong dependence of the estimate on the selected threshold for extremes.

It is worth reiterating however that fulfilment of the $D(u_n)$ condition alone does not guarantee that a process has an extremal index. For the latter both $P\{\underline{x}_n \leq u_n(\tau)\}$ and the associated with the IID process $P\{\underline{\tilde{x}}_n \leq u_n(\tau)\}$, need converge to a nondegenerate distribution (and thus of extreme value type). O'Brien (1974) and Leadbetter et al. (2012) construct a few counter-examples for which $D(u_n)$ holds and $P\{\underline{x}_n \leq u_n(\tau)\}$ converges, but the process does not have an extremal index because the associated IID sequence does not converge; yet the latter are too artificial and unlikely to be of practical interest.

In practical terms the above results show that weak forms of extremal dependence alter the parameters of the limiting distribution but do not invalidate its appropriateness. A slight decay has been reported in the rate of convergence to the asymptotic laws (e.g. Eichner et al., 2011), yet for normal sequences Leadbetter et al. (2012) prove that it is the same as in the independence case. In practice since the parameters are typically estimated from the data, regardless of the user's awareness of it, the effect of such dependence is incorporated in the model for extremes. Yet awareness of dependence is still important in terms of the bias induced in the estimation of extreme value quantiles (Koutsoyiannis, 2020b).

2.3.3 Other measures of extremal dependence

An alternative way of expressing extremal dependence other than testing based on the $D'(u_n)$ condition, is in terms of the limiting behaviour of the joint distribution of extremes. This measure has found large applications in multivariate extreme value analysis of independent random variables, where it is known as tail dependence coefficient (Ledford and Tawn, 1996). Sibuya (1960) was among the first to examine tail dependence by providing the proof for the asymptotic independence of the bivariate normal distribution. Under the assumption that the multivariate variables have only weak long-range dependence, the concept can also be suited for the examining dependence in extremes, treated as lagged variables of a single process. If $[\underline{X}_1, \underline{X}_2]$ denotes a bivariate random vector with common marginal distribution function F representing distinct exceedances of a threshold, then the coefficient of tail dependence between $\underline{X}_1$ and $\underline{X}_2$ is defined as:

$$\chi = \lim_{x \to x^*} P[\underline{X}_2 > x | \underline{X}_1 > x] \tag{25}$$

where x^* denotes the upper end point of the common marginal distribution, given that the limit exists. The case of $\chi = 0$ signifies asymptotic independence, whereas cases of $0 < \chi \leq 1$ denote asymptotic dependence. Threshold-dependent variants of the original coefficient have also been formulated, able to characterize dependence at sub-asymptotic levels as well (Coles et al., 1999). A number of other summaries of multivariate dependence exist (see eg. Beirlant et al., 2006; p. 273) but a more relevant discussion for applications in univariate processes and particularly rainfall series is provided in Ledford and Tawn (2003). An obvious however limitation of this approach is the increase in dimensionality when one is interested in characterizing dependence beyond the bivariate case.

2.3.4 Cases of stronger dependence

It is recalled that the results and methods of the previous section are based on the assumption of some form of restriction of long-range dependence in high-level exceedances ($D(u_n)$ assumption). This is considered a weak assumption and it may be asymptotically valid for extremes even from classes of processes exhibiting long-range dependence, namely Gaussian linear processes (Embrechts et al., 1999). However, given the marked non-Gaussianity of natural processes and the fact that extremes from lower thresholds may also be of interest, inference based on related metrics is not as straightforward for persistent processes. We stress that results concerning the validity and uncertainty of these metrics in non-Gaussian long-term persistent processes are very scarce in the literature, in fact the topic is not covered at all in most textbooks (e.g. Galambos, 1994; Embrechts et al., 1999; Beirlant et al., 2006; Finkenstadt and Rootzén, 2003; Kottegoda and Rosso, 2008; Kotz and Nadarajah, 2000; Leadbetter et al., 2012; Resnick, 2007; Beran, 2004). A notable exception are the mathematical contributions on extremal properties of self-similar processes by O'Brien et al. (1990) showing that different limits may emerge for the extremes of self-similar processes, other than the ones suggested by EVT, and the former may not have a planar point process representation, i.e. as the Poisson point process for IID extremes.

A few other contributions on the subject mainly refer to Gaussian processes. Normality is a convenient condition for the study of extremes, as the joint moments of normal processes —which control the extremes' behaviour, are fully determined by their mean and covariance structure. For such processes, Mittal and Ylvisaker (1975) have shown that in cases of strong persistence defined in terms of the rate of decay of the covariance function, i.e. cases where the $D(u_n)$ condition is not satisfied, the limiting distribution of extremes from normal processes is a normal distribution too.

Therefore, the theoretical properties of extremes from heavy-tailed and persistent processes are understudied, despite their relevance to natural process. In practice though, simulation provides the means for circumventing this issue, as performed in Chapters 6-7. The modelling options to simulate extremal behaviour are discussed next.

2.4 Treatment of dependence and extremes in common modelling approaches

Having discussed the development of theory and inference tools for extremal dependence, in this section we examine how the latter is dealt by common modelling approaches of the wider statistical literature. Studies dealing directly with extremal dependence abound in the econometrics literature, relating to modelling insurance and finance data (Embrechts et al., 2013), but are much scarcer in environmental literature. Below we review the most relevant hydrological modelling approaches of two types; ones related to the joint modelling of extreme and nonextreme properties of the parent process, and ones focused on the tail of the distribution of the parent process. In the final section, we discuss potential for bridging the two.

2.4.1 Joint modelling of nonextreme and extreme properties of the parent process

A complementary approach to the use of asymptotic theory for studying extremes is the explicit modelling of the parent process generating the extremes. This approach is particularly useful in cases where modelling of the parent process is required but preservation of the extremal properties is also essential, as in the case of streamflow simulation for reservoir management. In such cases, achieving an efficient modelling of the whole process including the tails, is not only practical, but improves theoretical consistency of the model estimates' as well. In general however, in parent process modelling extremal dependence is rarely explicitly dealt by, rather the behaviour of extremes is assessed in terms of model validation. As preservation of the multi-scale properties of a process is rather challenging, often this approach differentiates between placing the modelling focus either on finer or larger time-scales. Attempts to achieve consistency among these scales typically make use of disaggregation techniques (e.g. Koutsoyiannis and Manetas; 1996). Below we revisit the most classic and some relevant emerging approaches with regard to their accounting for extremes and dependence. Covering asymptotic results on general classes of stochastic models is not within the scope of this review; namely the interested reader is referred to Rootzén (1986; 1988) for extremal dependence in Markov chains and moving-average processes with non-Gaussian tails. Rather the aim here is to outline the basic characteristics of hydrological models with respect to flexibilities in modelling long-term persistence

and extremes. In this respect, non-parametric models are not discussed as they do not allow extrapolation to events beyond the range of the observed.

Point process models

Fine scale (sub-hourly, hourly or daily) rainfall requires a different modelling approach to that of large scale rainfall (monthly, annual or inter-annual) due to of the distinctive characteristics of rainfall at fine scales, i.e. prevalence of clustering mechanisms, including intermittence (Koutsoyiannis, 2006). Waymire and Gupta (1981) demonstrate mathematically that fine-scale rainfall properties are consistent with point process theory. According to Cox and Isham (1980), a point process defined in the set of positive real numbers, $\mathbb{R}^+$ is "a stochastic process for which each realization consists of a collection of points, each point having a well-defined position, usually in one-dimensional space, but possibly in some higher dimensional space". The main categories of point processes are Poisson-cluster processes, Cox processes and renewal processes. The first category is the simplest and most widely used in literature as established by Rodriguez-Iturbe et al. (1987a; 1987b). In the general case of this approach, storms arrive according to a Poisson process of rate λ triggering the generation of clusters of cells associated with each storm according to another process. Cells are characterized by duration usually following an exponential distribution, and a random depth described in terms of its first three moments. Depending on the type of process that is employed for the cell clustering mechanism, two main models are identified in literature, the Neyman-Scott and the Bartlett-Lewis processes (Onof et al., 2000). In the Neyman-Scott processes, the number of cells in a storm follows a random distribution, usually Poisson or geometrical, and the cell arrival times are exponentially distributed. In the Bartlett-Lewis processes, the cell arrival process is another Poisson process of rate β, associated with the origin of each storm and terminated at an exponential rate γ. Thus, in the first case the arrival times of cells are modelled with respect to the storm origins, while in the second case the inter-arrival times between successive cells are of interest. The model is fitted upon minimizing the difference between theoretical properties of the model and observed rainfall statistics. Typically, the latter include first and second-order statistics (mean, variance, autocovariance) as well as the probability dry, from timescales ranging from 1 h to 24 h, while inclusion of the third moment has been proposed as well in order to improve the fit to the extremes (Cowpertwait, 1998).

Reproduction of the extremes is the most challenging task for this type of models, as they tend to underestimate hourly and sub-hourly extremes (Verhoest et al., 2010), often followed by an overestimation the daily (Onof and Wang, 2019). Inclusion of the skewness in the calibration set along with various re-parameterizations of the original model and coupling with disaggregation schemes have been found to contribute to better fitting to the extremes (Cowpertwait, 1998; Kaczmarska et al., 2014; Kossieris et al., 2018; Onof and Wang, 2019). Onof and Wang (2019) argue that it is of high importance to capture the fat-tailedness of the storm intensity distribution as well, by employing a heavier tail distribution in the model.

The other critical issue with respect to the extremes relates to capturing their long-term variability. By construction Poisson-cluster processes are characterized by short-term dependence induced by the clustering of cells within a storm, but lack long-term dependence due to the use of a Poisson process for the generation of the storm arrivals and the independence of the cells among different storms (e.g. Rodriguez-Iturbe et al., 1987b). Marani (2003) highlighted the fact that accordingly Poisson-cluster models are expected to underestimate the variance for scales larger than those of calibration, which typically extends from one to few days; which is also confirmed by Onof and Wang (2019) in spite of various amendments to the original model. A remedy proposed by Park et al. (2019) is the coupling of the Bartlett-Lewis model with a seasonal autoregressive integrated moving average (SARIMA) model in order to capture the observed long-term rainfall variability. The model showed an improved fit to multi-scale extremes, which however came at the cost of a substantial increase in the model parameters compared to the original version. Recently, Kim and Onof (2020) also attributed the underestimation of the variability and extremes at large-scales to the fundamental structure of Poisson cluster models and proposed the use of adequate reshuffling procedures to induce long-term dependence in the model output. These works along with other prior studies (Kim et al., 2013; Paschalis et al., 2014) converge to the fact that accounting for rainfall variability across scales in this type of models is crucial for a better reproduction of the extremes.

Two-part models

Two-part models are based on the decomposition of the modelling of the rainfall process to the explicit modelling of the occurrence process, i.e. dry or wet state, and that of the rainfall intensity process, i.e. the nonzero rainfall of wet days (Srikanthan and McMahon, 2001). This class of models is one of the oldest in rainfall modelling (Gringorten, 1966; Todorovic and Woolhiser, 1975) and also became popular under the term 'stochastic rainfall generators' (Wilks, 1999). The occurrence process is typically modelled by a 'chain-dependent' process, comprising two states, a wet and a dry, the alternation between which is determined by a matrix of transition probabilities. The latter is usually assumed to be a Markov chain of order p, where p in case of daily rainfall indicates the number of days which are taken into account for the estimation of the transition probabilities. The common choice is that of a first-order Markov chain, although the resulting dependence pattern of dry spells is often underestimated (e.g. Wilks, 1999). This may be in part amended by considering higher-order Markov chains, or different parameterizations for the transition probability scheme, incorporating stronger dependence (Koutsoyiannis, 2006). An important drawback however is that the modelling of the rainfall intensity process typically entails the assumption of independence (Wilks, 1999). In fact, it is well-known that this class of models underrepresents the inter-annual rainfall variability (e.g. Buishand, 1978; Katz and Parlange, 1998). Possible improvements may come from the incorporation of hidden state Markov models which have shown potential in capturing inter-annual variability (Sansom, 1998; Thyer and Kuczera, 2000).

Multifractal models

Multi-fractal models also known as multiplicative random cascade models, arise from the concept of self-similarity, else scale-invariance, who shaped in the mid 20th century and became popular by the works of Mandelbrot (1974, 1983). A detailed review on the history of the concept, its definitions and its applications in hydrology is given by Veneziano et al. (2006b) and Veneziano and Langousis (2010). These models are mainly phenomenological based upon the empirical observation of scaling in nature (e.g. Newman, 2005). They are built upon the concept of perfect scaling of moments, and have been popular simulation algorithms of rainfall during the past decades (Schertzer and Lovejoy, 1987; Gupta and Waymire, 1993; Marshak et al., 1994; Over and Gupta, 1994; Menabde et al., 1997; Langousis and Veneziano, 2007). A desirable feature of these models is that they reproduce variability in a parsimonious fashion that also captures other statistical properties including the extremes (Veneziano and Langousis, 2010). At the same time though they encompass some fundamental limitations. First, typical models assume a single scaling exponent for all the moments of the process, which has been questioned in the literature for the rainfall process (e.g. Veneziano et al., 2006a; Molini et al., 2009; Serinaldi, 2010). In particular regarding the autocovariance structure more complicated behaviours have been observed (Marani, 2003; Markonis and Koutsoyiannis, 2016; Iliopoulou and Koutsoyiannis, 2019). Furthermore, the basis in multi-fractal analysis is the determination of the moment scaling function from the data, which is impacted by enormous estimation uncertainty considering classical moment estimators. In particular, estimation beyond the order of three is shown to be highly unreliable (Lombardo et al., 2014). Apart from the latter, Koutsoyiannis et al. (2018) highlighted a number of theoretical inconsistencies in the fractal approach and its applications, most notably the fact that scale invariance is a mathematical abstraction that violates certain natural laws as finiteness of energy and space. Rather the existence of scales beyond which a certain power law exists or ceases to hold appears to be a more applicable assumption for natural processes (Gneiting and Schlather, 2004; Gneiting et al., 2012; Koutsoyiannis, 2016). Considering the latter points, Koutsoyiannis et al. (2018) suggest that the useful concepts of fractal theory can be incorporated into existing stochastic models, while related estimation issues may too be more rigorously treated within the framework of stochastics.

Linear stochastic models

The class of stochastic models has a long history dating back to early 20th century; a classification of dominant approaches is provided by Koutsoyiannis (2019a). The most widely known modelling approach is autoregressive models which originated in the works of Yule (1927) and Walker (1931) and gained stochastic foundations by the works of Wold (1938, 1948) and Whittle (1952, 1953). They became however popular by the acronyms —AR(p) (autoregressions of order p), MA(q) (moving-averages of order q), ARMA(p,q) (linear combination of the latter models) and ARIMA (p,d,q) (autoregressive integrated moving average), given in the famous book of Box and Jenkins (1970). By construction they are short-range dependence models, with the exception of the ARFIMA(p,d,q) model able of modelling long-range dependence through the use of a real valued d parameter, instead of the integer one (Granger and Joyeux, 1980; Hosking, 1981)

Despite their large popularity, this class of models suffers from a number of issues, namely definition in discrete time in contrast to the continuous-time evolution of natural systems, definition in terms of the autocorrelation structure whose estimation is negatively biased, and overparameterization, with the exception of simple AR(1), ARMA(1,1), and ARFIMA(0, d, 0) versions (Koutsoyiannis, 2016). Koutsoyiannis (2000, 2002, 2016) developed an alternative parsimonious approach for model identification and fitting based on a generalized form of the autocovariance structure, and proposed a simulation algorithm, the symmetric moving-average scheme (SMA), suitable both for short- and long-range dependent processes. Another approach for long-range dependence is the approximation of the second-order structure by an infinite sum of Markov processes (Mandelbrot, 1971), an approach parameterized for HK processes by Koutsoyiannis (2002). In contrast though to the flexibility of the SMA scheme, the former is a simple scaling approach and cannot preserve the dependence structure at timescales tending to zero (fractal behaviour; cf Dimitriadis and Koutsoyiannis, 2018). A further advance to the moving-average scheme has been the development of its asymmetric variant (AMA) which enables preservation of temporal irreversibility of the process (Koutsoyiannis, 2019a). The latter may be profound in atmospheric process at fine time-scales while it is particularly relevant for the simulation of the streamflow processor time scales up to several days (Ribatet et al., 2009; Mathai and Mujumdar, 2019; Serinaldi and Kilsby, 2016a; Koutsoyiannis, 2019b).

A marked challenge for the above models is the preservation of the marginal distribution of the process in cases of non-Gaussianity (see Tsoukalas, 2018; for a review of different approaches in hydrology). The latter is particularly relevant for the reproduction of the extremes. The original version of the SMA model (Koutsoyiannis; 2000, 2002, 2016) explicitly models the second-order properties of the process and approximates the marginal distribution by preserving the first three moments (thus, up to skewness). An extension of the model enabling preservation of four moments (up to kurtosis) has been provided by Dimitriadis and Koutsoyiannis (2018a). An alternative approach was followed by Papalexiou (2018) performing the simulation of the dependence structure in the Gaussian domain using autoregressions and back-transforming to the non-Gaussian domain through the inverse transformation. The known effect of the non-linear marginal transformation on the autocorrelation of the process (Embrechts et al., 2002) is dealt by prior to the model application, by inflating the correlation structure of the parent Gaussian process in order to preserve the target correlation of the arbitrary process. A similar modelling approach based on the Gaussian auxiliary process but using the SMA model for the generation scheme instead, is developed by Tsoukalas et al. (2018). A general form of this approach is reviewed in Lavergnat (2016). Finally, an alternative approach of performing the simulation of the dependence structure in the frequency domain, instead of the time-domain, using phase randomization and coupling with a parametric distribution is suggested by Brunner et al.(2019).

It is worth recapitulating that short-range dependent Gaussian processes asymptotically do not exhibit extremal dependence (θ = 1), while the same might be true even for Gaussian processes with weak long-range dependence (Leadbetter et al., 1983; Embrechts et al., 1999; Ancona-Navarrete and Tawn, 2000). Therefore, in

presence of clustering of extremes and long-term rainfall variability, from the above class of linear models only the ones able to simulate long-range dependence and heavy tails should be of interest.

Autoregressive conditional heteroscedasticity models

These models are mainly developed in the econometrics literature around the concept of stochastic volatility. Stochastic volatility refers to random changes of the variance as a function of time, defined in the context of stochastic recurrence (difference) equations (de Haan et al., 1989). These models were introduced in econometrics because usual linear models of the ARMA-type exhibited light tails and conditional constant variance and could not capture the peculiarities of financial timeseries. The latter refer to presence of heavy tails, changes in volatility, high correlations in the squares and absolute values of the data, clustering of high-threshold exceedances, while showing almost no correlation in the actual values of the data (Embrechts et al., 2013). Hydrological processes share some similarities with the above properties, therefore although applications of this class of models are very limited in hydrology, they represent a potentially interesting class for consideration. A relevant application of such a model for daily rainfall was performed by Laux et al. (2011).

In discrete-time, models reproducing stochastic volatility are referred to as 'conditionally heteroscedastic' models and are of two general types. Engle (1982) introduced the AutoRegressive Conditionally Heteroscedastic process of order p (ARCH), extended by Bollerslev (1986) who developed the generalized version (GARCH(p, q)). In contrast to the linear case where the noise is additive, the noise in these models appears multiplicatively. The variance however is changing linearly, conditionally on the values of past observations for the ARCH type, as well as on their conditional variance for the GARCH type, in a way that high volatility may arise either as a result of large absolute values of past data or from previous periods of large volatility (e.g. Embrechts et al., 2013). Therefore, a squared ARCH process can be represented as an ARMA process. Different definitions of the way in which the variance changes conditionally abound giving rise to many variants of the type, while extensions to continuous time have also been proposed (Klüppelberg et al., 2004). An attractive property of these models have been the possibility to generate heavy tails using light-tailed noise terms, i.e. Gaussian innovations (Kesten, 1973). Conditions for the existence of stationary versions and for the existence of moments are discussed in Embrecths et al. (1996). A review on different types of ARCH models and fitting methods is provided by Shephard (1996).

The distinctive feature of ARCH processes is that they exhibit extremal clustering, which is the reason for their wide popularity in finance and econometrics. The degree of clustering may be difficult to obtain analytically, but can be approximated through Monte Carlo simulations as in de Haan (1989). Therefore, GARCH models implicitly capture some properties of persistent timeseries but in principle they do not preserve the correlation structure of the original process and they are not designed to reproduce long-range dependence, either. The reason is that in standard ARCH and GARCH modelling the behaviour of the conditional variance is modelled and this may appear persistent, irrespectively of the behaviour of the unconditional one, which is the modelling focus in case of long-range dependence

(Mikosch and Starica, 2002). Modifications however have been proposed in order to capture LRD-type behaviour defined in the absolute values and their squares of log-returns (Breidt et al., 1998; Giraitis et al., 2000; Ibragimov and Lentzas, 2008). Overall, the merits of the ARCH framework pertain to modelling heavy tails and clustering of extremes. Yet ARCH theoretical properties are not as developed as in the case of linear stochastic processes, and the modelling of the unconditional second-order dependence structure, which is important in hydrology, is not straight-forward.

Copula models

The copula representation is a way of modelling continuous multivariate distributions by separating the modelling of the univariate marginal distributions and that of their dependence structure (Joe, 2014). The latter is modelled through the copula which is a multivariate distribution comprising univariate dependent random variables uniformly distributed U(0, 1). The theoretical foundations are based upon Sklar's theorem (Sklar, 1959), showing that every multivariate cumulative distribution function of a random vector can be expressed in terms of its marginals and the copula, and the works of Fréchet (1951) and Hoeffding (1940) who derived the bounds of the copula. Applications of copulas abound in the statistical literature particularly due to their flexibility in modelling dependence structures other than the linear case; for instance dependence measures of monotonic association such as the Kendalls' tau (1938) and Spearmans' rank correlation (1904) as well as more general dependence structures (cf. Joe, 2014). The linear dependence is argued to be counter-intuitive and too restrictive for non-elliptical multivariate distributions (Embrechts et al., 2002). A desirable property of copulas is that under strictly increasing transformations of the random variables, the copula properties, including dependence between extremes, remain invariant. The bivariate Gaussian copula is the most commonly applied due to the desirable properties of the multivariate joint normal distribution. Yet with regard to capturing the behaviours of extremes, the Gaussian copula is not suited; e.g. see Sibuya (1960) for the proof of lack of tail dependence in the bivariate Gaussian case. In such cases, the Gaussian copula will underestimate the joint tail probability, and therefore non-Gaussian copulas, such as the Archimedean (Genest and MacKay, 1986) are often employed.

The great flexibility of the copula framework renders it a possible modelling option for a wide range of hydrological issues. Although in practice it has not been particularly popular for full process modelling, it is discussed in this section due to its potential of coupling with a wide range of modelling approaches. Notable examples are studies using the copula approach coupled with the Markov-chains in order to simulate intermittent rainfall (Laux et al., 2009; Serinaldi, 2009a, 2009a), coupled with linear stochastic models for simulation of hydrological processes (Lee and Salas, 2011; Papalexiou, 2018; Tsoukalas et al., 2018), use of multidimensional copulas to characterize various dependence structures in hourly rainfall (Salvadori and De Michele, 2006), as well as copula applications in the spatio-temporal modelling of rainfall (Villarini et al., 2008; Serinaldi, 2009; Bárdossy and Pegram, 2009; Laux et al., 2011). A greater deal of copula applications are found in the field of multivariate hydrological frequency analysis, as in intensity-duration rainfall models (De Michele and Salvadori, 2003; Zhang and Singh, 2007; Vandenberghe et al., 2010), and

multivariate flood modelling (Favre et al., 2004; Grimaldi and Serinaldi, 2006; Serinaldi and Grimaldi, 2007). Currently, the field of applications is still rapidly growing (Salvadori and De Michele, 2010), following the similar trend in finance and insurance (Embrechts, 2009).

Overall, the copula literature consists of a number of ad hoc approaches for modelling dependence structures and extremes. In terms of long-range dependence, Ibragimov and Lentzas (2008) employ a copula-based definition and show that there exists a range of copula-based Markov-processes that exhibit such dependence on the copula-level. However, these approaches do not form a well-understood and stand-alone framework for modelling long-range dependence, and tracking its effect on extremes. This shortcoming is prominent when compared to the self-contained theory of stochastic processes, where both theoretical properties and sample estimation procedures for dependent data are established. In this respect, a critical point of view on the applications of copulas is provided by Mikosch (2006).

2.4.2 Extreme-oriented modelling

Bortot and Tawn (1998) identify four critical components of the behaviours of extremes for stationary sequences: (i) the probability of exceeding the threshold, (ii) the distribution of the exceedances of the threshold, (iii) the long-range dependence between exceedances and (iv) the local clustering of exceedances within any set of dependent exceedances of the threshold. The basic theory behind impact of dependence in modelling of extreme, presented in Section 2.3., provides an asymptotic characterization of (ii), i.e. the distributional behaviour of extremes, by setting conditions on the rest of these properties. For high thresholds of extremes, it assumes presences of local clustering, quantified through the extremal index, but only weak long-range dependence between clusters. In general, the approaches to characterize between-cluster dependence fall under the following categories: (a) some formulation of probability mixing conditions, as Leadbetter's (1983) $D(u_n)$, which require that for extremes over an adequately high threshold a separation time exists above which they can be viewed as forming independent clusters, and (b) second-order dependence properties, suitable to characterize extremal dependence for Gaussian processes, as Berman's (1964) condition on the rate of autocorrelation decay. For within-cluster dependence the natural characterization is the extremal index approach as detailed above, but other conditional probability approaches, as the ones discussed in Section 2.3.3, have been formulated as well.

The asymptotic properties of extremes provide the theoretical basis for extrapolation irrespective of knowledge of their parent distribution. In practice though, issues in terms of statistical estimation from data arise. The main issues relate to determining the rate of convergence to the asymptotic behaviour, quantifying estimation bias and uncertainty, and identifying the impact of the threshold for which dependence or independence is manifested. These are particularly relevant for the frequency analysis of hydrological data. For instance, it has been argued that the use of GEV distribution via the blocks method, entails wasteful usage of data, as only the maximum per block is retained for modelling (Volpi et al., 2019), which is a central arguments in favour of the peaks over threshold method instead (Pickands III, 1975).

The peaks over threshold method may increase the sample of observations but in theory, it too attempts to restrict the presence of persistence by selecting only the maximum of a certain cluster, in order for the generalized pareto distribution to hold as an approximation of the distribution of exceedances (Coles et al., 2001; Ferro and Segers, 2003). In doing so, information on local clustering of data is discarded, without even considering the ambiguity in defining independent clusters. On the other hand, even for cases of asymptotic independence where the extremal index equals θ=1, it is possible that at finite levels of exceedances clustering is observed, with cluster size decreasing to 1 as the threshold increases; in essence, at sub-asymptotic levels the threshold is important in determining the behaviour of the process. Also the bias and the variability of the estimates are highly dependent on the model assumption, thus without formal modelling it is possible to misinterpret empirical results. Apart from the functionals of extremes identified at the beginning, aggregate exceedances are also very important and heavily rely on the assumed model (Smith et al., 1997). In terms of convergence to the asymptotic distributions, it has been shown to be very slow for hydrological data (Koutsoyiannis, 2004a).

The latter are important arguments in favour of non-asymptotic methods for modelling extremes, even though in this case as well, asymptotic results are useful to infer the properties that should be retained in close-form modelling (Koutsoyiannis, 2020b). Below, we review literature contributions regarding sub-asymptotic methods for extremes, which share three broad aims: i) modelling of exact (instead of asymptotic) extremal properties based on the properties of the parent process, ii) modelling dependent exceedances of a given threshold by multivariate analysis of their upper joint tail, iii) modelling both dependence and marginal distribution of extremes through autoregressive maxima models.

In terms of exact results, a number of studies on extremes have been published in the hydrological literature. De Michele (2019) provides a review of approaches to derive the exact distribution of maxima without assuming 'identically distributed' extremes. With a similar rationale but focused on relaxing the independence assumption, Lombardo et al. (2019) derive the exact distribution of maxima taken from low threshold POT with magnitudes characterized by an arbitrary marginal distribution and first-order Markovian dependence, and negative binomial occurrences. Volpi et al. (2015) derive the distribution function of the waiting time for processes with Markovian dependence, while Serinaldi and Lombardo (2020) derive the probability distribution of the waiting time till the kth extreme also under long-range dependence. A sub-asymptotic treatment of dependence in rainfall extremes in the framework of multifractal models is also described in Veneziano et al. (2006b). An explicit derivation of ombrian models, i.e. generalized intensity-duration-frequency curves, incorporating persistence has been provided by Koutsoyiannis (2020b), using the extreme-oriented modelling framework based on k-moments (Koutsoyiannis, 2019c).

The second type of studies aim at explicitly modelling local clustering of extremes. To this aim, a lot of studies model the joint distribution of threshold exceedances using bivariate distributions and typically assuming a Markov-chain dependence structure (Ledford and Tawn, 1997; Smith et al., 1997; Bortot and Tawn, 1998; Ribatet et al., 2009). Common choices for the joint distribution are the bivariate

t-tail model, as well as bivariate extreme value models (Ledford and Tawn, 1997; Salvadori and De Michele, 2010). A marked issue however is the fact that Markovian dependence may underestimate properties of extremes, e.g. Ribatet et. al (2009) find that flood durations are under-estimated in the case of first-order Markov chain. On the other hand, use of higher-order Markov chains increases complexity. A second approach is based on hierarchical modelling by using a latent stochastic process to infer parameters of the distribution of threshold exceedances in order to simulate extremal clustering. Such models for rainfall extremes are proposed by Bortot and Gaetan (2014, 2016), who also provide a detailed discussion on the degree of extremal clustering that the models can produce.

As far as autoregressive maxima approaches are concerned, to our knowledge applications in the hydrological literature are very scarce. The moving-maxima process is a representation of the max-stable processes introduced by de Haan (1984). This approach refers to replacing sums by maxima in the linear time series approach and using a Fréchet distribution for the innovation terms, which lead to a max-stable process, since all its finite dimensional distributions are max-stable. Moving-maxima processes bare connections to multivariate extreme value distributions as well as to stable and moving-average processes; e.g. these are discussed in Hall et al. (2002). A special class of moving maxima process are max-ARMA processes, whose theoretical properties are studied by Davis and Resnick (1989). For a first-order moving maxima process (ARMAX), the extremal index equals $\theta = 1-\alpha$, while Berliant et al. (2006) show that for a general type of moving maxima processes, asymptotically it holds that $\theta = \alpha(1) = \max_{j\geq 0} a_j$. Although the theoretical properties are well developed (de Haan, 1984; Davis and Resnick, 1989), the statistical applications are not as established in the hydrological domain. An exception is the work of Tyralis and Langousis (2019) modelling intensity-duration-frequency curves through max-stable processes. Notably more applications of max-stable process can be found for the spatial modelling of rainfall extremes (e.g. Davison et al., 2012).

2.4.3 Overview of approaches

The development of extreme value theory has enabled the decoupling of the modelling of the extremes from that of the parent process. The limiting laws of maxima provide the basis for theoretically consistent extrapolation to the range of unobserved events requiring estimation only of the three first moments of the sample maxima. The latter is convenient because most available records are of short length and cannot support the any-order estimation of the process's distribution, while even for longer records, estimation based on higher-order ordinary moments is very uncertain. This issue has been resolved only recently, by the development of approximately unbiased higher-order moment estimators that employ all data and are known as knowable (*k*-) moments (Koutsoyiannis, 2019c). The inclusion of higher-order moments in the model calibration may bridge the modelling of nonextreme properties of the parent process with a faithful representation of its extremes.

Although the decoupling of the distributional modelling of extremes from the parent process under EVT has been very practical, it has also led to a general disregard for the links of extremes to the parent process. Consequently, in the modelling of the

parent process, behaviours of extremes other than their marginal distribution are usually overlooked, by implicitly assuming independence. On the other hand, in extreme-oriented modelling approaches, absence of long-range dependence is commonly taken for granted, whereas short-term clustering, i.e. local dependence, is treated as a separate behaviour, rather than a byproduct of temporal dependence in the parent process. In a more subtle way, the attachment to the assumption of independence is also manifested by the increasing number of trend studies invoking deterministic causality in case of non-IID extremes (Koutsoyiannis and Montanari, 2015). The deviations of hydrological extremes from the IID assumption, their implications and modelling are the focus of Chapters 4-8, while Chapter 3 revisits the second-order structure of the rainfall process.

3. REVISITING LONG-TERM PERSISTENCE IN THE RAINFALL PROCESS

In this chapter, the second-order dependence theory is applied to the rainfall process with the aim to investigate its persistence dynamics. The identification of persistence in a natural process provides empirical grounds for investigating the long-term temporal properties of its extremes. While persistence is identified in various natural processes, it is usually less acknowledged in the rainfall process. Thus, it is revisited here using a global rainfall dataset.

3.1 Introduction

Since Hurst (1951) brought long-term persistence, also known as long-range dependence (LRD), into scientific discourse, the interest in this behaviour has been rising. This is mainly due to its serious implications into the modelling and design processes in various scientific fields and particularly in water resources (O'Connell et al.). Another fact contributing to its growing popularity is that LRD has been identified in many climatic variables, such as temperature (Pelletier, 1998; Koutsoyiannis, 2003), rainfall (Fraedrich and Larnder, 1993; Pelletier and Turcotte, 1997), wind power (Haslett and Raftery, 1989) and the North-Atlantic oscillation index (Stephenson et al., 2000). The Hurst behaviour has also a strong physical basis, as it is derived from the principle of entropy maximization (Koutsoyiannis, 2011a), a principle which can be used to determine the theoretical probability distribution model for rainfall (Papalexiou and Koutsoyiannis, 2012). More detailed discussion on the history and relevance of the Hurst behaviour can be found in the review paper by O'Connell et al. (2016).

In this analysis, we aim to investigate the dependence properties of annual rainfall. Studies regarding LRD in annual rainfall are usually limited in a specific area and/or utilize datasets of relatively short lengths (Kantelhardt et al., 2006; Bunde et al., 2013; Zhai et al., 2014). Short record lengths can introduce bias into the estimation of long-term persistence properties, which in general, need more than 100 years in order to avoid underestimation (and, in cases of very strong dependence, even more than 1000) (Koutsoyiannis and Montanari, 2007). A majority of other studies investigate the dependence structure of rainfall at sub-annual or even smaller scales (Papalexiou et al., 2011), but in that case, the phenomenon gets complicated due to the combined effects of seasonal variation and intermittency. On the other hand, paleoclimatic reconstructions suggest strong LRD behaviour in multi-decadal to centennial time scales (Pelletier and Turcotte, 1997; Markonis and Koutsoyiannis, 2015). Evidently, there are still ample grounds for research on the existence of LRD in annual precipitation.

Herein, we analyze more than one thousand annual precipitation records of length of a hundred years or more from different areas of the world. To quantify LRD, we estimate the Hurst coefficient, through the variance-based method (climacogram) and employ Monte Carlo method to identify a common Hurst coefficient for all the records. Additionally, we perform a simple test on the autocorrelation structure of the

first few lags to examine whether the hypothesis of a Markovian autocorrelation structure is justified or not. Finally, we discuss the effect of time-scale and record length on LRD estimation.

3.2 Dataset

The instrumental data are obtained from the Global Historical Climatology Network (GHCN-Daily, https://www.ncdc.noaa.gov/ghcnd-data-access), which contains daily data from more than 50 000 land surface stations around the globe. A significant percentage of these records exhibit the typical issues of most datasets available, i.e. missing values, short record length and rainfall values of questionable quality, such as unrealistic outliers. In order to restrict data quality to a significantly high level, we filter the dataset using certain criteria.

We study only the stations satisfying the following conditions: a) record length over 100 years, b) missing values less than 20% and, c) suspect values with quality flags less than 0.1%. Initially, in order to construct the annual series we delete all daily values assigned quality flags, indicating unrealistically large values, and then estimate the average daily value per year. Notably, because of the existence of missing values within most records, summing up all daily values of a year would result to smaller annual estimates than the real ones; to a degree dependent on the number of missing values. It would be clearly more robust to estimate the daily mean values per year. This is only performed for the years having less than 20 missing daily values while the rest are considered missing. Then, all stations having more than 20% missing yearly values are removed. This screening results in 3477 stations with lengths varying from 100 years to 173 years. Among the 3447 stations there are different combinations of record lengths and missing values, e.g., 558 stations having 100 years in a sequence with no missing values, 1474 stations with more than 100 values and only eight stations without any missing values. We choose to analyze 1265 stations having more than 100 values and a missing values percentage less than 15%. Obviously, this choice ensures a higher quality dataset for our analysis.

3.3 Methodology and results

3.3.1 Variance scaling method

The method employed herein is based on the study of the variability of the data averaged at different timescales. The method is sometimes referred to as aggregated variance method, but what is actually aggregate is the timescale and not the variance. Specifically, let $\underline{X}_j$ be a stationary process on discrete time j (referring to years in our case) with standard deviation σ and let:

$$\underline{X}_j^{(k)} = \frac{1}{k} \sum_{l=(j-1)k+1}^{jk} \underline{X}_l, \qquad k = 1,2,3 \ldots \tag{26}$$

denote the averaged process at timescale k, with standard deviation $\sigma^{(k)}$. In the case of an uncorrelated process, the standard deviation of $\underline{X}_j^{(k)}$ is obtained by $\sigma^{(k)} = \frac{\sigma}{k^{1/2}}$. In the

opposite case, i.e. if the process is a Hurst-Kolmogorov process, as introduced in Section 2.2, the abovementioned law is invalid. Instead one obtains the elementary scaling property:

$$\sigma^{(k)} = k^{H-1}\sigma \tag{27}$$

where H is the Hurst coefficient, which for stationary and positively correlated processes varies in the range (0.5, 1). The value of $H = 0.5$ denotes time independence, while smaller values are indicative of anti-persistence. The autocorrelation of the aggregated process is independent of the scale of aggregation k and is given as follows:

$$\rho_j^{(k)} = \rho_j = \frac{1}{2}\left[(j+1)^{2H} + (j-1)^{2H}\right] - j^{2H}, \qquad j > 0 \tag{28}$$

To apply the method to the data we use the climacogram tool (Koutsoyiannis, 2011b), which is the double-logarithmic plot of the standard deviation $\sigma^{(k)}$ of the aggregated time series at scale k versus the time scale k. The H value is estimated as the slope of the fitted line (least squares regression). In a variant of that method, the estimation bias of the standard deviation, which depends on the time-scale of aggregation, is also considered.

Each averaged time series is constructed as follows. For every scale k, the data are divided into n groups, the number of which is obtained as the fraction of the data length L versus the scale value k. For example in time scale $k = 4$, 120 years would be divided in 30 non-overlapping groups of 4 years. Subsequently, the values within each group are averaged according to equation (26). However, when missing values are encountered, the process of averaging may become problematic depending on the number of missing values; if more than a half of the values is missing, then the estimate would be quite uncertain (Markonis, 2015). To overcome the issue, we use a simple criterion on the number of missing values before estimating the averaged series within each group: a) for scale $k = 2$ the average value is estimated only when both values exist b) for scales $k \geq 3$ the average value is estimated only when there are at least three values within the group. According to the latter rule, we estimated the averaged series for all the scales between $k_{\min}$ and $k_{\max}$, where $k_{\min} = 1$ and $k_{\max} \leq L/10$ so that the variance in the maximum scale is estimated from at least 10 values (Koutsoyiannis, 2003). For a 100-year record length this would be the variance of the decadal means.

The results of the algorithm implementation for the instrumental data are shown in Table 3.1 and Figure 3.1, suggesting evidence of weak long range dependence. More specifically, it was found that 85% of the data exhibit $H \geq 0.5$, yet with notable variation. For example, only half of the data show $H \geq 0.59$, i.e. a more pronounced dependence structure. A very strong dependence structure, $H \geq 0.80$ is reported for the 2.5% of the records, while for 15% of them we observe lack of dependence. For the 95% interval, H values fluctuate between 0.4–0.8.

Table 3.1 Summary statistics of the Hurst parameter as estimated from the climacogram method applied to the 1265 records. Q indicates the empirical quantile

Min	$Q_{2.5}$	Q_{25}	Median	Q_{75}	$Q_{97.5}$	Max	Mean	SD
0.23	0.40	0.53	0.59	0.65	0.80	0.99	0.59	0.1

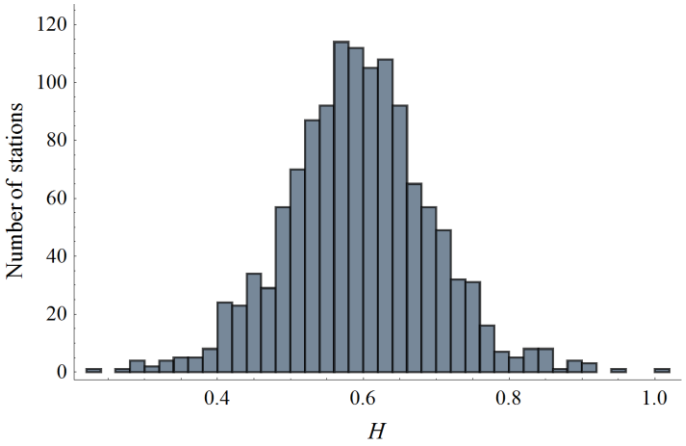


Figure 3.1 Empirical distribution of the Hurst coefficient $\boldsymbol{H}$ as resulted by applying the aggregated variance method to the 1265 annual rainfall records.

In order to test the effects of our parametric choices for the value of the minimum and maximum scale, we examined how the median and the variance of H estimates vary for different $k_{\min}$ and $k_{\max}$. As can be seen in Figure 3.2, the variance of the Hurst parameter estimate becomes larger as the value of the minimum scale $k_{\min}$ increases; yet the value of the median in the estimate remains the same. Therefore, our choice of $k_{\min} = 1$ is well-justified, since greater values of $k_{\min}$ only amplify the uncertainty in H estimation. In addition, the observation of the same median strengthens our hypothesis of the LRD structure, because in the alternative hypothesis of short term dependence, we would notice some change in the climacogram curvature and correspondingly to the logarithmic slope. The results for the $k_{\max}$ were similar. It can be seen in Figure 3.3 that the decrease in the number of values in the last scale increases the variance of the Hurst parameter estimate in this case too. Therefore, the choice of $n \geq 10$ leads to more reliable results compared to using smaller values of n.

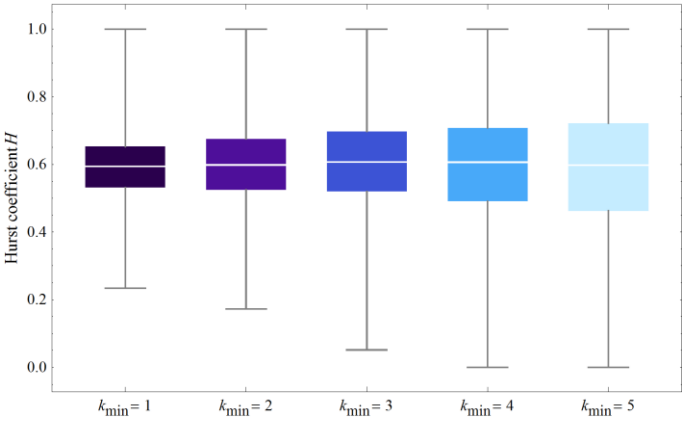


Figure 3.2 Box-plots depicting the sample differences resulting from variations in the value of minimum scale k_{min} when applying the climacogram method.

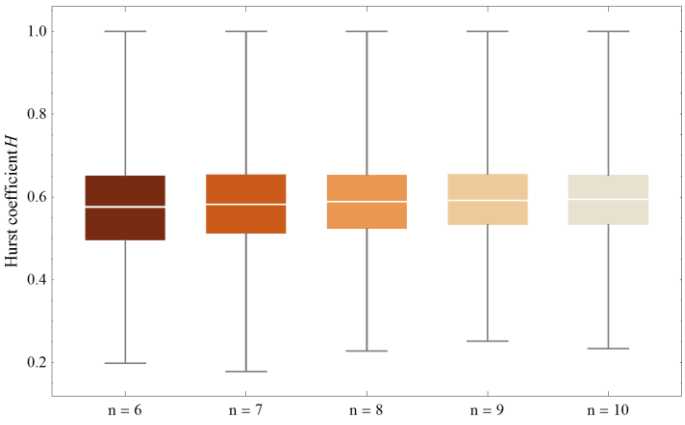


Figure 3.3 Box-plots depicting the sample differences resulting from variations in the number of minimum values n in k_{max} when applying the climacogram method.

3.3.2 Least Square Based on Standard Deviation Method (LSSD)

Koutsoyiannis (2003) demonstrated how the use of the classical estimator for the standard deviation can introduce significant negative bias in the estimation of the Hurst parameter by the variance scaling method. This is because the hypothesis of independence, which is a necessary condition for the use of the estimator, is violated in the case of processes with strong LRD behaviour. This shortcoming may be overcome by the use of the Least Square Based on Standard Deviation Method (LSSD) (Koutsoyiannis, 2003; Tyralis and Koutsoyiannis, 2011b), which performs a simultaneous estimation of the Hurst parameter H and the standard deviation σ using an approximately unbiased estimator for the latter.

Here, for simplicity reasons we applied the LSSD method (Tyralis and Koutsoyiannis, 2011) only to the sample of the 558 (44% of the total) stations with no missing values and then, compared our estimate with the one obtained by the simple climacogram method for the same sample. As shown in Table 3.2 and Figure 3.4 the two methods show small deviations from each other. Overall, the value of the bias

fluctuates between 1-2% with the bias in the estimate of the average being approximately 1%. The bias is negligible in this case because the estimated Hurst parameter is not very high.

Table 3.2 Summary statistics of the Hurst parameter as estimated from the climacogram method and the LSSD method both applied to the 558 records without missing values. Q indicates the empirical quantile.

	Climacogram method	LSSD method
Mean	0.56	0.58
SD	0.10	0.09
Min	0.28	0.33
$Q_{2.5}$	0.37	0.40
Q_{25}	0.50	0.52
Median	0.56	0.57
Q_{75}	0.63	0.64
$Q_{97.5}$	0.78	0.79
Max	0.90	0.92

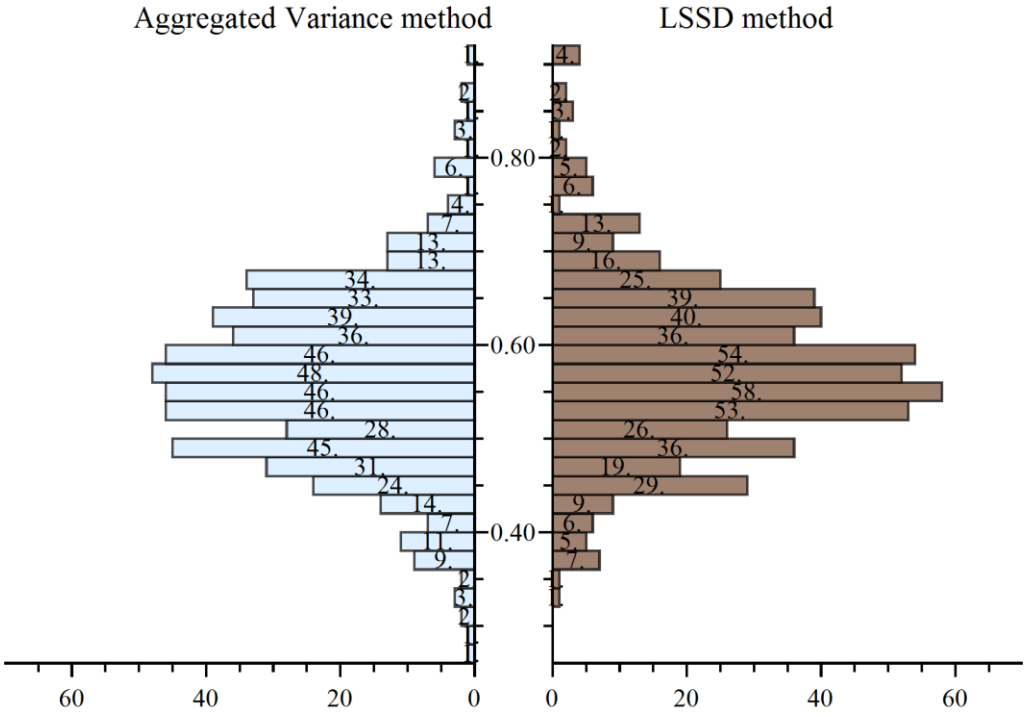


Figure 3.4 Double histogram depicting the empirical distribution of the Hurst coefficient ***H*** resulting from the climacogram method (left) and from the LSSD method (right), both applied to the 558 annual rainfall records without missing values.

3.3.3 Monte Carlo testing

We also investigate the existence of a theoretical distribution of the Hurst coefficient that can satisfactorily match the empirical one; i.e. whether there is a unique Hurst coefficient which could be considered representative for all the records. In order to produce a theoretical sample of time series exhibiting HK dynamics, we use a simple algorithm that generates Fractional Gaussian Noise based on a multiple timescale fluctuation approach (Koutsoyiannis, 2002). We generated 1265 time series that reproduce the record length, the mean and the standard deviation of the empirical sample, repeat the same procedure for several theoretical H values and then estimated the empirical ones. The distribution of the empirical estimates for the synthetic time series was compared to the distribution of the empirical estimates for the historic time series used in the analysis. It appears that the value of $H = 0.58$ (Figure 6) yields the most satisfactory match. However, it is worth noticing that that 2.5% of the stations, exhibiting $H > 0.8$, are outside the range of the theoretical distribution.

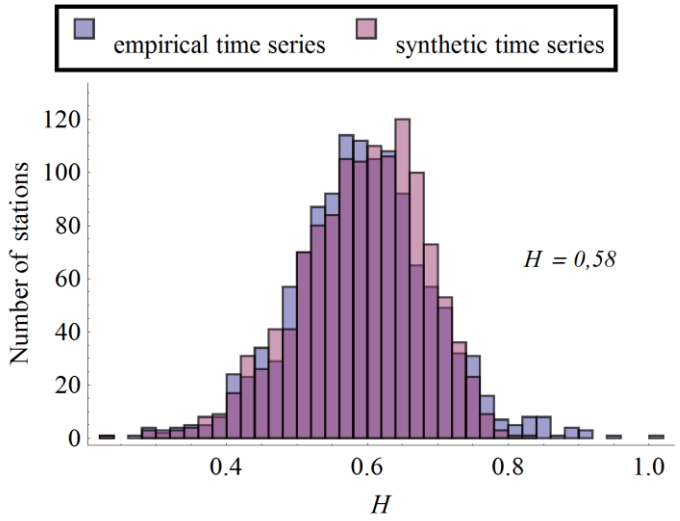


Figure 3.5 Paired histogram depicting the match of the empirical (blue) and theoretical (purple) distribution of the Hurst coefficient $\boldsymbol{H}$ resulting from applying the aggregated variance method to the 1265 historical records and 1265 synthetic records respectively. The synthetic records are realizations of a stochastic process characterized by a theoretical Hurst coefficient $\boldsymbol{H} = 0.58$.

3.3.4 Autocorrelation analysis

The estimated Hurst coefficient is not high enough to allow for any certain conclusion on the type of the dependence structure, since relatively low Hurst coefficients (0.5-0.6) can be estimated when there is short range dependence or no dependence at all due to algorithmic inadequacies, sample bias and estimation uncertainty. To this end, we also employ the autocorrelation function, to further examine the dependence properties of rainfall. Still one should keep in mind that the classical autocorrelation estimator, as in the case of standard deviation, is biased downwards (Koutsoyiannis, 2003; Dimitriadis and Koutsoyiannis, 2015). However, since the estimator is biased downwards, any result in favour of LTP, would mean that in reality, the LTP is even stronger.

The autocorrelation coefficients of the first three lags for the instrumental data are low (Table 3.3).

Table 3.3 Summary statistics of the estimated autocorrelation coefficients for lags 1, 2, 3. Q indicates the empirical quantile.

	ρ_1	ρ_2	ρ_3
Mean	0.12	0.03	0.05
SD	0.11	0.12	0.11
Min	−0.19	−0.35	-0.32
$Q_{2.5}$	−0.10	−0.16	-0.15
Q_{25}	0.05	−0.05	-0.02
Median	0.11	0.02	0.05
Q_{75}	0.18	0.10	0.12
$Q_{97.5}$	0.37	0.29	0.27
Max	0.62	0.59	0.47

On further investigation, we test whether independence is a plausible scenario for the dependence structure of our data. We produced 1265 independent, i.e. uncorrelated, time series of the same sample size and estimated the sample autocorrelation coefficients (Figure 3.6). It can be seen that for all three lags the value of the median of the historic data is greater than the one estimated from uncorrelated synthetic data. This is more obvious in the case of autocorrelation of lag-1 where for the 95% interval the values of the independent data fluctuate in the range –0.175 to 0.173, while the historic ones are in the range –0.09 to 0.37. In addition, in all three cases, the historic samples exhibit significantly fewer negative values than the uncorrelated ones.

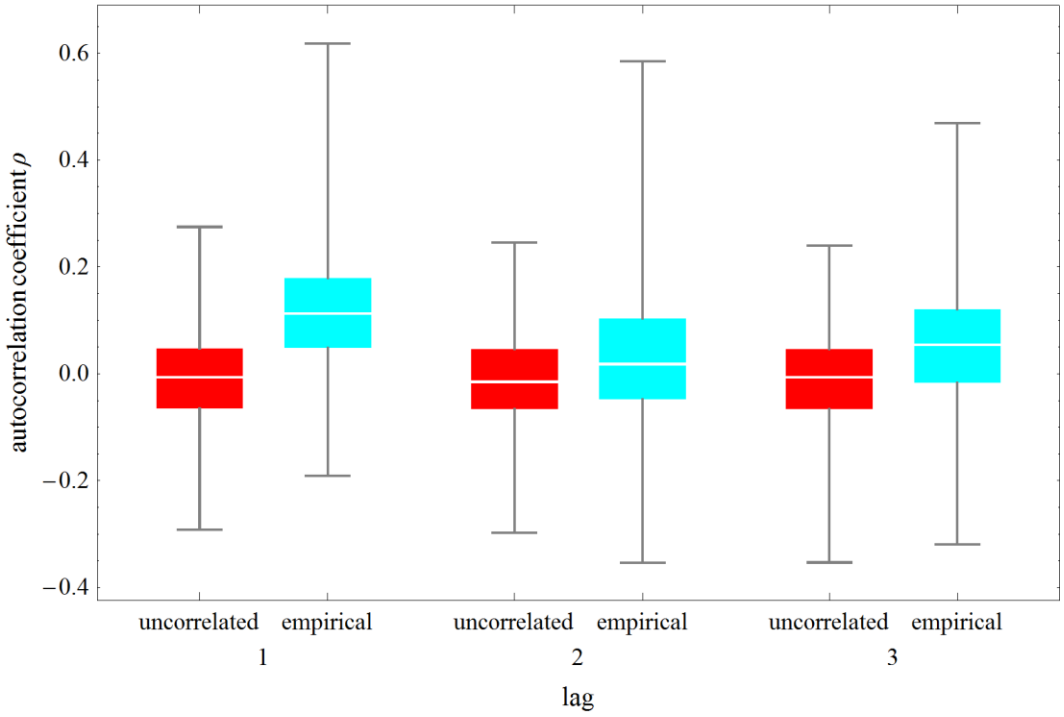


Figure 3.6 Box-plots depicting the resulting sample differences of the autocorrelation coefficient ρ between the empirical series and uncorrelated series for lags 1, 2, 3.

The above results could be typical for a Markov process too, also known as AR(1) process. To address this issue, a simple ad hoc test, which exploits the distinctive properties of Markov processes, was designed. Under the Markov hypothesis, the theoretical autocorrelation coefficient for lag 2 would be estimated as $\rho_2 = \rho_1^2$, where ρ_1 is the known empirical autocorrelation. Likewise, the Markovian autocorrelation coefficient for lag 3 would be given as $\rho_3 = \rho_1^3$. The resulting theoretical estimate is compared to the empirical one for the same lag; if the empirical value is higher than the theoretical AR(1) one, then the Markov hypothesis weakens.

We applied this comparison to the 52% of the stations for which all the autocorrelation coefficients for lags 1-3 are positive (Figure 3.7). It is evident that the empirical estimates are considerably higher than the theoretical ones resulting from an AR(1) structure and therefore, the Markov assumption becomes less likely. In addition, the empirical estimates do not follow the exponential convergence to zero of the Markovian ones, but instead, remain approximately stable for lags 2 and 3; this is in agreement with the theoretical behaviour of LRD whose distinctive feature is the existence of slowly decaying autocorrelation function (Beran, 1994).

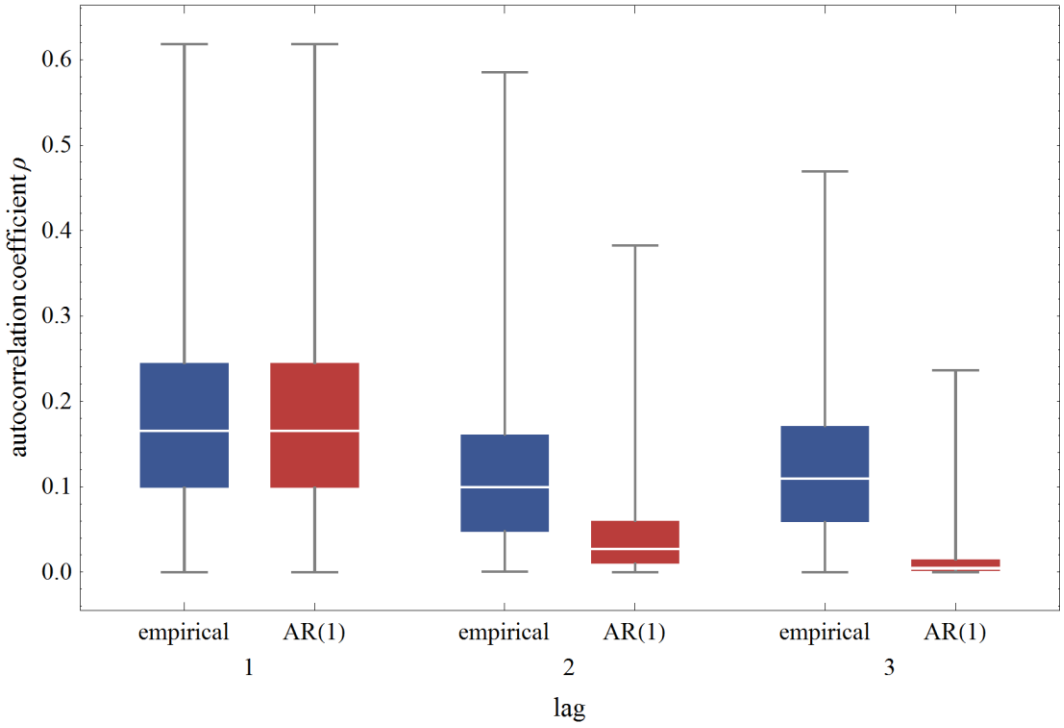


Figure 3.7 Box-plots showing the sample differences of the autocorrelation coefficient ρ between the empirical series and synthetic series generated from an AR(1) model for lags 1, 2, 3.

Having tested the cases of independence and short-range dependence, we finally examine whether the autocorrelation structure is consistent with that of a FGN model. In Figure 3.8 the empirical autocorrelation coefficient ρ_1 is plotted against the corresponding estimated Hurst coefficient H as obtained from equation (27). The diagram shows that the autocorrelation structure is consistent with that of a FGN model. The deviation between the theoretical and the empirical estimates becomes greater in the region of high values of H; still this is justified due to the increased negative bias in the autocorrelation estimation in that case.

Figure 3.8 Observed Hurst coefficient H vs. autocorrelation coefficient ρ_1 points of the 1265 annual rainfall records and the theoretical line typical of a HK model.

3.4 Discussion and conclusions

The analysis of the global instrumental dataset shows that there are notable indications of weak LRD in the annual rainfall. As the Hurst parameter is not very high, the simple application of the climacogram method induces only 1-2% negative bias in the Hurst coefficient estimation and therefore, the estimated via Monte Carlo, theoretical common value of $H = 0.58$, may be considered accurately representative for instrumental data.

The study of the autocorrelation function shows that it is consistent with the autocorrelation of a FGN model, even though for a certain percentage of the stations the Markov hypothesis could not be falsified. Specifically, the existence of negative correlations in all three lags examined did not permit the performance of the abovementioned method in the case of the 48% of the stations. Some studies using smaller data sets (Potter, 1979; Fraedrich and Blender, 2003; Kantelhardt et al., 2006) supported the appropriateness of the Markov structure, but they did not investigate the differences between actual and theoretical auto-correlation in larger lags (Figure 3.7). These differences might be quite small, it has been shown though, that they might have serious implications when it comes to the estimation uncertainty (Koutsoyiannis and Montanari, 2007). For instance, in terms of trend significance, the observed changes in rainfall might be considered quite rarer than they actually are (Cohn and Lins, 2005). Lastly, it was shown as well, that the autocorrelation function significantly departs from the case of independence.

Although the above findings are in favour of the existence of a stronger dependence structure than the one typically assumed in literature (Potter, 1979; Fraedrich and Blender, 2003; Kantelhardt et al., 2006), it seems that there is a

discrepancy between smaller and larger time scales (Fraedrich and Larnder, 1993; Pelletier and Turcotte, 1997; Poveda, 2011; Ault et al., 2013). To this end, the most important source of uncertainty in the determination of LRD, which is the record length, should not be overlooked (Koutsoyiannis, 2002; Koutsoyiannis and Montanari, 2007). Although using stations with relatively high —compared to the majority of the existing rainfall data records— record length, the accurate detection of long range dependence cannot be guaranteed because this behaviour may require even longer record length to be revealed. Subsequently, the low estimates of Hurst parameter in instrumental time series could be attributed to the limited record length available in some cases and therefore, should be considered characteristic only for this time horizon of approximately 100 years. This is also suggested by the work of Markonis and Koutsoyiannis (2015), which emphasizes the influence of time-scale when it comes to the analysis and reveal of the dependence of a time-series. An additional analysis of longer-term records is presented in Chapter 6.

It is also important to consider the uncertainty induced due to measurement errors or false homogenization techniques which may introduce bias to the estimation of LRD (Steirou, 2011). GHCN-Daily highlights the potential bias provoked by changes in instrumentation over the years and it is possible that this kind of bias could also affect the estimation of H.

Ultimately, the high variability of the results is in accordance with the inherent uncertainty of the phenomenon, apart from algorithmic or data choices. An important conclusion drawn from the analysis is that simplifying assumptions commonly used in practice, such as inter-annual independence, may, in cases, significantly, depart from reality and hence, a thorough and careful study of the dependence properties of the dataset, as performed here, is recommended, especially when longer time horizons are of interest.

4. CHARACTERIZATION AND MODELLING OF EXTREME RAINFALL SEASONALITY

This Chapter is the first of the two dealing with the temporal dynamics of hydrological extremes induced by seasonality. A novel framework is formulated to address the question of objectively characterizing and modelling seasonality of rainfall extremes. The effects of seasonality in the distributional modelling of rainfall extremes are discerned using extreme value theory. A robust parameterization approach is proposed to resolve consistency issues reported in the literature. The effectiveness of the proposed scheme for seasonal characterization and modelling is highlighted when contrasted to results obtained from the conventional approach of using fixed climatological seasons. To these aims, a dataset comprising long-term daily rainfall records (>150 years) is employed.

4.1 Introduction

Seasonality is a dominant feature of most hydrological processes including extreme rainfall (Hirschboeck, 1988). It implies intra-annual periodic variability which pertains to both timing and magnitude of extreme rainfall. An accurate and effective characterization of seasonality is critical to a wide variety of hydrological applications. For instance, it is useful in the scheduling of various flood preparedness measures, including management of stormwater infrastructures (Dhakal et al., 2015) and reservoir operation (Chiew et al., 2003; Fang et al., 2007; Chen et al., 2010a). Similarly, seasonality characterization is exploited in advanced schemes of flood-frequency analysis incorporating causative mechanisms (e.g. Sivapalan et al., 2005; Li et al., 2016) and may be useful for medium-range flood prediction (e.g. Koutsoyiannis et al., 2008; Wang et al., 2009; Aguilar et al., 2017), for which inclusion of seasonal extreme rainfall may increase prediction skill. Modelling of seasonal rainfall extremes – which typically implies some sort of frequency analysis – may also inform the selection of design values for related infrastructure. Additionally, the latter provides support to within-year operation of water resources systems, design rainfall estimation (Golian et al., 2010; Efstratiadis et al., 2014) and probabilistic assessment of extreme events occurring in a given season. Nowadays, extreme rainfall seasonality also prompts renewed scientific interest as a field of trend analyses (Ntegeka and Willems, 2008; Dhakal et al., 2015; Tye et al., 2016; Wu and Qian, 2017).

Characterization of extreme rainfall seasonality is scarcely dealt with by the relevant literature. Most of the established methods are devised to identify the temporal span of a wet season and assess its significance, typically by a priori identifying a single wet season. For example, directional statistics are typically applied to identify the high flow season (Cunderlik et al., 2004a; Baratti et al., 2012; Chen et al., 2013) and have also been applied to characterize the timing of seasonal rainfall (Parajka et al., 2009, 2010a; Lee et al., 2012). However, directional statistics are inefficient when extremes occur over multiple seasons, which is very likely in the case of rainfall (Cunderlik and Burn, 2002). Recently, Dhakal et al. (2015) provided an improvement to the traditional method of directional statistics by adopting a non-

parametric approach to capture multiple modes in the timing of annual rainfall maxima. Yet they noted that the proposed method is sensitive to the subjective selection of threshold values to assess significance of circular density estimates. Multimodality of the seasonal regime is also dealt with by analyzing the monthly relative frequencies of extreme occurrences (Cunderlik and Burn, 2002; Cunderlik et al., 2004a). This approach, however, relies on the subjective identification of the monthly time step to characterize seasonality. The latter along with the four climatological seasons are often used when large-scale or global analyses are performed (e.g. Rust et al., 2009; Villarini, 2012; Serinaldi and Kilsby, 2014; Papalexiou and Koutsoyiannis, 2016) but lead to disregarding the large spatial variability of atmospheric patterns and may not align well with local behaviours (Pryor and Schoof, 2008; Dhakal et al., 2015). Moreover, the fixed partitions do not resolve the crucial question of the identification of the optimal number of seasons, therefore resulting in over-parameterization of the seasonal model of extremes due to the large number of seasons that is adopted, particularly in the 12 month model.

A sub-optimal characterization of seasonality could be a reasonable compromise when one is interested in characterizing the timing of the most extreme events only. However, technical applications often require the modelling of the frequency of extremes during the whole course of the year. In this regard, several previous studies have either considered climatological information or employed statistical methods along with some degree of subjective judgement to estimate the optimal number of seasons and their displacement in time (e.g. Durrans et al., 2003; Chen et al., 2010a; Baratti et al., 2012; Bowers et al., 2012). Coles et al. (2003) adopted a different approach by treating seasonal temporal limits as unknown parameters to be identified within a Bayesian framework. Yet, they also identified the number of seasons a priori through subjective inference.

The above literature review highlights a methodological gap in the objective identification of the optimal number of extreme rainfall seasons and their duration. To the best of the authors' knowledge, existing methods are not suitable for directly inferring multimodality from the seasonal regime and concurrently identifying segmentation points between seasons in an objective manner.

The research herein presented proposes a two-purpose framework for (a) objective seasonality identification and (b) modelling of rainfall extremes in order to effectively estimate the seasonal probability of extreme events. To this end, we introduce two alternative methods for season identification, which are characterized by different levels of parsimony in terms of data requirements, therefore providing two options for practical applications. Our approach employs an information-theoretic framework (Akaike Information Criterion, AIC) to estimate the optimal number of seasons. In order to describe the frequency of extremes in each identified season we use the GEV probability distribution. We discuss the consistency of the model at different time scales. Finally, in order to demonstrate the efficiency of our framework we present a comparison with the traditional 4-season approach.

An extended dataset of long daily rainfall records is herein investigated, as detailed in the next section. The length of the records, the shortest one covering an observation period of 150 years, allows us to inspect the impact of uncertainty, which may be relevant for seasonal extreme value analyses (Cunderlik et al., 2004b). To

reduce uncertainty we propose a robust parameterization approach of seasonal-annual distributions which is supported by empirical evidence.

4.2 Dataset

Our dataset includes 27 daily rainfall records each one spanning over 150 years. Eighteen of them are collected from global databases, namely, the Global Historical Climatology Network Daily database (Menne et al., 2012) and the European Climate Assessment & Dataset (Klein Tank et al., 2002). Figure 4.1 shows the geographical location of the stations, while Table A.1 (Appendix A) reports the coordinates of each station, the observation period, as well as the number of years that are fully covered by observations after quality control and screening of missing values. For the extraction of the annual maxima we employ a methodology proposed by Papalexiou and Koutsoyiannis (2013); accordingly, an annual maximum is not accepted if (a) it belongs to the lowest 40% of the annual maxima values and (b) 30% or more of the observations for that year are missing. For seasonal and monthly maxima we compute statistics only if number of missing values is less than 10% of the total sample (season or month). The longest series is that of Padua, spanning a period of 275 years, that is the longest rainfall record existing worldwide (Marani and Zanetti, 2015).

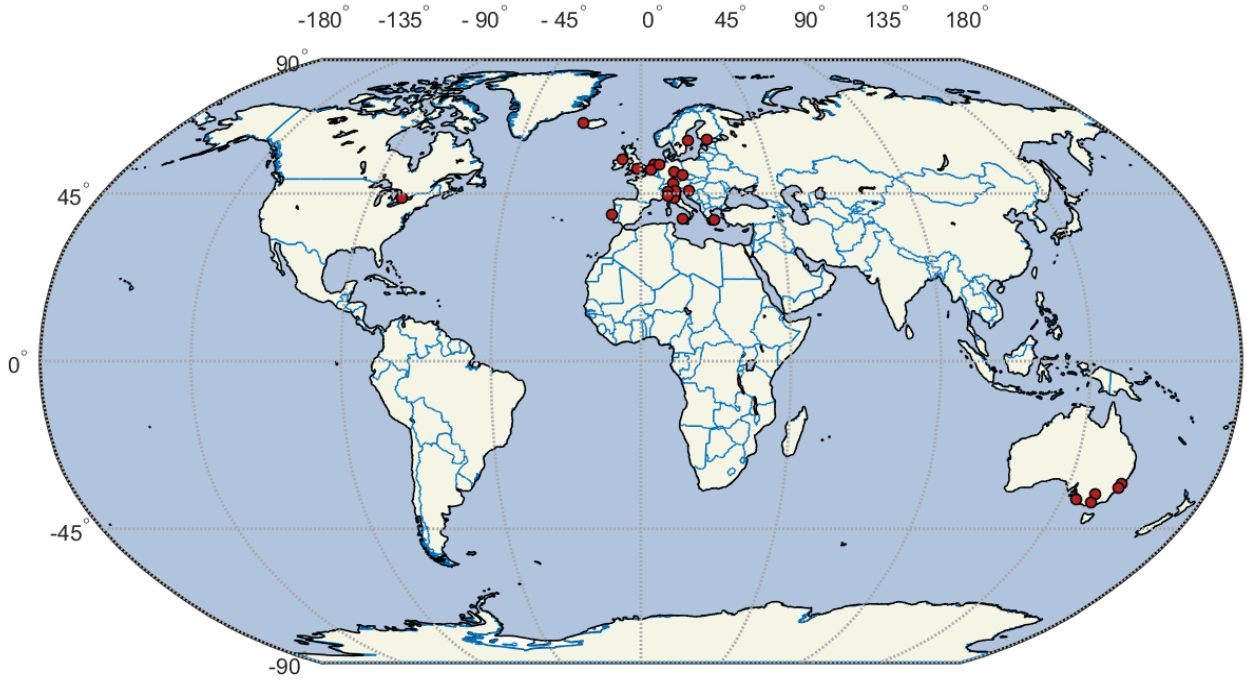


Figure 4.1 Map of the 27 analyzed stations with daily rainfall records spanning over 150 years.

4.3 A new method for identifying seasonality of extreme rainfall

The methodology that we propose to identify seasons is inspired by cluster analysis and model selection techniques. Seasons are regarded as groups (clusters) of consecutive months with similar behaviour of extremes. The question of selecting the

number of seasons that best describe the dataset is addressed here via a model selection process under the assumption that different numbers of clusters (seasons) represent alternative plausible models for the dataset. Two alternative methods for season identification characterized by different level of parsimony are considered here and described below.

In what follows, we denote random variables by underlined symbols and their realizations by plain form, respectively. We also use bold characters for vectors. We denote season, month and year with the indexes $i = 1,..n$, $j = 1,...,12$, and $k = 1,..k_{\max}$, respectively, where n is the number of seasons and $k_{\max}$ is the record length in years. We assume that n is fixed a priori and denote with $\boldsymbol{C}_i$ the vector containing the j values of contiguous months belonging to the same season i, and with s_i its size. Accordingly, we define the following random variables:

- $\underline{R}_{i,j,k}$ is the maximum daily rainfall amount of season i, month j and year k;
- $\underline{R}_{i,j}$ is the temporal average of maximum daily rainfall of month j of season i along the record, namely, $\underline{R}_{i,j} := \frac{1}{k_{\max}} \sum_{k=1}^{k_{\max}} \underline{R}_{i,j,k}$;
- $\underline{R}_i$ is the temporal average of the $\underline{R}_{i,j}$ values along the season i, $\underline{R}_i := \frac{1}{s_i} \sum_{j \in \boldsymbol{c}_i} \underline{R}_{i,j}$.

For instance, $R_{2,5,12}$ for season i=2 defined by $\boldsymbol{c}_2$ =(5,6,7) denotes the maximum daily rainfall observed in May of the 12th year of a given record and belonging to the 2nd identified season of the year, which also includes months June and July; likewise, $R_{2,5}$ is the sample average of maximum rainfall observed in all May days of the record, while, R_2 is the sample average of all monthly averages belonging to season 2, in this case of May, June and July.

We call the first method for season identification the SSD algorithm. It is based on the computation of Sum of Squared Deviations (SSD) of the $R_{i,j}$values from their seasonal average, R_i for all seasons according to the equation:

$$\mathrm{SSD} = \sum_{i=1}^{n} \sum_{j \in \boldsymbol{c}_i} (R_{i,j} - R_i)^2 \tag{29}$$

This metric is evaluated for each possible clustering combination $\boldsymbol{C}_i$ of consecutive months for the given number of seasons, thus enabling the identification of the lower value of SSD, which identifies the optimal partition of the year into n seasons. We require a season to span at least two months and allow the algorithm to group months across different calendar years. The requirement for a season to span at least two months implies that the maximum number of seasons is 6, but preliminary investigations showed that more than three seasons are rarely present in extreme rainfall. Therefore, we limit our attention to n values ranging in the interval (1-4).

Essentially, the SSD algorithm minimizes the within-cluster variance of the average value over the years of the monthly rainfall maxima and can be considered as a simplification of the well-known k-means algorithm (MacQueen, 1967). Since seasons may include contiguous months only, and the algorithm deals with only 12 data points to cluster —the average over the years of daily maximum rainfall values for each month— the number of possible combinations is relatively low and the method is parsimonious.

In order to identify the optimal number of seasons we define alternative probabilistic models, with different level of parsimony, to describe the frequency of occurrence of extreme events in each season and assess their ability to optimally fit the observed record. Accordingly, we first select a trial value for the number n of seasons in the range (1-4) and partition them by applying the above SSD algorithm. To describe the probability distribution of rainfall in each season and the whole year we form a mixture model with n seasonal components, each described by its own probability distribution. Hence, according to the law of total probability, the probability distribution of the seasonal model for a generic seasonal random variable $\underline{U}$ takes the form:

$$f_{\underline{U}}(u; \boldsymbol{a}_1, \dots, \boldsymbol{a}_n) = \sum_{i=1}^{n} w_i f_{\underline{U}_i}(u_i; \boldsymbol{a}_i) \tag{30}$$

where w_i are weights adding up to 1. They are obtained as the ratio of the season's length in months, s_i, versus the whole twelve-month period, i.e. $w_i = s_i/12$; and $\boldsymbol{a_i}$ is a seasonal parameter vector. Here $f_{\underline{U}_i}$ is a seasonal probability distribution for $\underline{U}$ describing realizations u_i in season i. Note that by applying the law of total probability instead of deriving the annual probability distribution as the product of the seasonal ones, we avoid relying on the assumption of independence of the random variables $\underline{U}_i$, which was adopted in other studies (Durrans et al., 2003). Therefore, this is a more general approach also appropriate for the cases of rainfall maxima being correlated among seasons.

The above step requires identifying and fitting a candidate model for the $f_{\underline{U}_i}$ probability distribution. We propose two alternative models for the seasonal probability distribution $f_{\underline{U}_i}(u_i, \boldsymbol{a}_i)$ which are characterized by different level of complexity.

The first option, which we call Average Based (AB) method, identifies the random variable $\underline{U}_i$, as the monthly temporal average $\underline{R}_{i,j}$. Then, we assume that $f_{\underline{R}_{i,j}}(R_{i,j}, \boldsymbol{a}_i)$ is a uniform distribution given by:

$$f_{\underline{R}_{i,j}}\left(R_{i,j}, \boldsymbol{a}_i\right) = \frac{1}{b_i} \tag{31}$$

where in this case $\boldsymbol{a}_i$contains only one parameter, namely, $b_i = \max_{j \epsilon \boldsymbol{c}_i} R_{i,j}$. Preliminary analyses showed that the uniform distribution provides an efficient representation of the frequency of the $R_{i,j}$ realizations, by minimizing the number of involved parameters. The above approach imposes an upper limit to the average value of the monthly maximum rainfall depth and sets the lower limit to zero.

The second option, which we call Complete Data (CD) method, identifies the random variable $\underline{U}_i$, as the maximum daily rainfall in each month j of the season i for the year k, which has been previously introduced as $R_{i,j,k}$. Then, we assume that $f_{\underline{R}_{i,j,k}}(R_{i,j,k}, \boldsymbol{a}_i)$ is described by two alternative probability distributions with a different tail behaviour, i.e. one characterized by a lighter and one by a heavier right tail, in order to allow flexibility in fitting the observed rainfall maxima. The first is the two-parameter Gamma distribution, given by:

$$f_{\underline{R}_{i,j,k}}\left(R_{i,j,k}, \boldsymbol{a}_i\right) = \frac{R_{i,j,k}^{\xi_i - 1} e^{-\frac{R_{i,j,k}}{\theta_i}}}{\theta_i^{\xi_i}\, \Gamma(\xi_i)} \tag{32}$$

where $\boldsymbol{a}_i = (\xi_i, \theta_i)$ is the parameter vector with ξ and θ being shape and scale parameters, respectively. The second is the two-parameter Weibull distribution:

$$f_{\underline{R}_{i,j,k}}\left(R_{i,j,k}, \boldsymbol{a}_i\right) = \frac{\mu_i}{\lambda_i}\left(\frac{R_{i,j,k}}{\lambda_i}\right)^{\mu_i - 1} e^{-(R_{i,j,k}/\lambda_i)^{\mu_i}} \tag{33}$$

where $\boldsymbol{a}_i = (\mu_i, \lambda_i)$ i is the parameter vector with μ and λ being shape and scale parameters, respectively. By working on the monthly maximum rainfall instead of their averages along the season, the CD method allows one to base the estimation of the probability model on a more extended dataset.

The above methodology allows several modelling options, which differ for the number of seasons, the application of either AB or CD method and the selection of either the Gamma or the Weibull distribution in the CD method. The best modelling option and the related optimal number of seasons is identified by applying the Akaike Information Criterion (AIC, Akaike, 1973, 1974). The criterion statistic for the pth candidate model, AICp, is given by:

$$\mathrm{AIC}_p = 2m_p - 2\ln L_p \tag{34}$$

where m_p is the number of parameters and L_p is the likelihood of the pth candidate model. The application of the criterion is straightforward as it only requires estimation of the likelihood function for the candidate probability models defined by equation (30). The minimum AIC value identifies the best candidate model by evaluating the bias versus variance trade off; i.e., the condition in which as the model parameters increase the bias of the model estimates decreases, yet their variance increases (Burnham and Anderson, 2002). Hence, AIC provides an implicit interpretation of the principle of parsimony which is pivotal in model selection (Box and Jenkins, 1970). Although AIC has a solid foundation in information theory both in mathematical terms and also from a philosophical point of view, its use is not still widely established in hydrological applications (Laio et al., 2009). For an insightful review of AIC's properties, the reader is referred to Burnham and Anderson (2002).

Therefore, the workflow for season identification is as follows:

1. A trial value is adopted for the number n of seasons in the range (1-4);
2. The n seasons are partitioned by applying the SSD algorithm therefore identifying the vectors $\boldsymbol{c}_i$, $i = 1,\ldots, n$, of the indices of the months that are included in each season;
3. AB and CD methods are applied to estimate the probability distribution of $\underline{R}_{i,j}$ and $\underline{R}_{i,j,k}$, respectively, in each season;
4. AIC is computed for candidate models;
5. The procedure is repeated for the other values of n in the range (1-4);
6. The resulting AIC values are compared therefore identifying the optimal number of seasons, and their partition, for AB and CD methods.

7. If n values resulting from AB and CD methods are the same, then the procedure is terminated and the optimal partition of seasons is uniquely identified;
8. If the estimated n values differ, then the user is allowed to select the preferred partition of seasons based on the suitability of $\underline{R}_{i,j}$ instead of $\underline{R}_{i,j,k}$, for the considered design problem.

4.4 Extreme value analysis

4.4.1 Fitting the GEV distribution

Once the optimal number of seasons and their partition have been identified, to estimate seasonal extremes one needs to fit a suitable probabilistic model for the seasonal block maxima series. The latter is formed by extracting from each identified season the maximum daily rainfall observed in each year. It is worth noting that distributions that were previously considered for seasonal partitioning (the Gamma and the Weibull) are not suited for fitting extreme values and therefore are not an option for the current target.

Extreme Value Theory (EVT) suggests that the distribution of the maximum of independent and identically distributed (IID) random variables asymptotically converges to three limiting laws (Fisher and Tippett, 1928), which are the Gumbel distribution (Type I), the Fréchet distribution (Type II) and the reversed Weibull (Type III), that can be unified under the single analytical form provided independently by von Mises (1936) and Jenkinson (1955) and known as Generalized Extreme Value (GEV) distribution, given by equation (21). As discussed in Section 2.3, in the case a limiting distribution exists for extremes from any parent distribution of the underlying stochastic process, then this is the GEV. Therefore, it could be the limiting distribution also in the case of monthly rainfall maxima described by the Gamma and Weibull distributions as in the CD method above. Leadbetter (1974) showed that convergence to GEV is guaranteed even in the presence of short-range correlation in the underlying stochastic process. In our case, the implication is that GEV emerges as limiting distribution even if rainfall maxima are weakly correlated. Koutsoyiannis (2004a) has shown mathematically that GEV still emerges as asymptotical distribution in the presence of different parent distributions from season to season. In practical applications, though, in which a maximum value is extracted from a small number of events, the asymptotic condition is unlikely to hold. In this respect, Koutsoyiannis (2004a) demonstrated that the convergence of the distribution of maxima to the GEV with a positive shape parameter (Type II) is good even for a small number of events and also for parent distributions belonging to the domain of attraction of the Gumbel (Type I), due to the increased flexibility of the three-parameter distribution. On the contrary, convergence rates to the Gumbel distribution are very slow even for distributions belonging to the domain of attraction of the Gumbel family (see also Papalexiou and Koutsoyiannis, 2013).

Here, we assume that the underlying stochastic process is given by the series of the monthly maxima of daily rainfall in each season. We aim to fit with the GEV distribution the seasonal samples that are obtained by extracting from each season i and each year k the maximum daily value $R^*_{i,k}$ therefore obtaining a block maxima series, which is assumed to be a realization of the random variable $\underline{R}^*_{i,k}$. We also fit the

series of the annual maxima R_k^*which is assumed to be a realization of the random variable $\underline{R}_k^*$. This approach shall allow one to estimate the extremes for the seasonal periods and the total annual period and ensures that both the seasonal and annual approaches refer to the same sample size when fitting the GEV, as the block maxima sampling method is used, i.e., one extreme event is sampled on a yearly basis for both the seasonal and annual periods.

4.4.2 Investigating consistency of seasonal and annual distributions

A considerable part of related literature (e.g. Buishand and Demaré, 1990; Durrans et al., 2003; Chen et al., 2010b; Baratti et al., 2012) has focused on the estimation of seasonal and annual flood frequency distributions and their inter-relationship. Usually, it is suggested that an independent fitting of seasonal and annual distributions may lead to inconsistency among them, manifested as a ''crossing over'' effect. The latter means that for extremely rare events seasonal quantiles may be higher than their annual counterparts. To resolve this inconsistency, a variety of methods for the joint estimation of the seasonal and annual distributions has been proposed.

Durrans et al. (2003) attributed distributional inconsistencies in seasonal-annual frequency analysis to three possible reasons: (a) the arbitrary parameterization of seasonal and annual distributions, (b) stochastic dependence among them and (c) estimation uncertainty. In this respect, we believe that the arbitrary specification of seasonal samples is also a major reason causing distributional inconsistencies (such a case is discussed and illustrated later in section 4.5). In our case though, we argue that the above inconsistency should rather be viewed as an empirical evidence of estimation uncertainty, which is particularly relevant in extreme value studies (Coles et al., 2003; Koutsoyiannis, 2004c). This is further supported by observing that the crossing over effect is manifested in the domain of extremely rare events, where uncertainty is prominent.

To inspect the impact of estimation uncertainty, we fit the GEV probability distribution by applying three different methods, namely, maximum likelihood (ML), method of moments (MM) and a least squares estimation method (LS) for an improved fitting of the extremes (Koutsoyiannis, 2004d). We further investigate estimation uncertainty in each of the three methods by computing 95% Monte Carlo Prediction Limits (MCPL) for the resulting GEV quantiles. MCPL are estimated by applying a Monte Carlo simulation which is structured according to the following steps: (1) we estimate the GEV parameters by each method, (2) produce 1000 synthetic GEV series for each derived parameter set, (3) re-estimate the parameters by the same method, (4) compute the resulting GEV quantiles for each of them and then (5) identify the 95% confidence region for each quantile value. The scope is to assess whether the crossing over falls within the limits of the estimation uncertainty as evaluated from applying a set of different parameter estimation methods. To further reduce fitting uncertainty, we propose a simpler alternative to joint parameterization, i.e. the joint estimation of a common shape parameter among seasonal-annual distributions – since the shape parameter is the most difficult to estimate accurately − and we discuss how this choice is supported by empirical evidence.

4.5 Results

4.5.1 Season identification for the observed records

Table 4.1 shows the AIC values resulting from season identification for the available stations. Following Burnham and Anderson (2004), we denote with ΔAIC the difference in the AIC value of each model with respect to the best one. Therefore, the zero ΔAIC model is the best model, while models with ΔAIC<2 and ΔAIC>10 are assumed to have good and little support, respectively. An example of seasonal partition for the case of Florence is shown in Figure 4.2a and Figure 4.2b for 2 and 3 seasons. We refer to this type of figures as *climatograms*, though the term is typically used for plots depicting both rainfall and temperature climatological regimes.

The results point out that both methods identified the one-season (annual) model as the best solution for 11 stations (with 6 stations being the same for both methods). In four stations, the one-season model was preferred by the CD method, while the two-season solution was indicated by the AB method. On further investigation, it was found that neither the Gamma, nor the Weibull provided satisfactory likelihood values for these stations. As a result, the more parsimonious one-season model was preferred by the AIC. The three-season model is identified as the best solution for five stations with the CD method, while the AB method did not select $n = 3$ for any station. This result was expected as the AB method exploits information from a limited dataset and therefore parsimonious models are likely to provide better AIC values. The Gamma distribution is selected as the best model in 21 cases and the Weibull for the remaining 6.

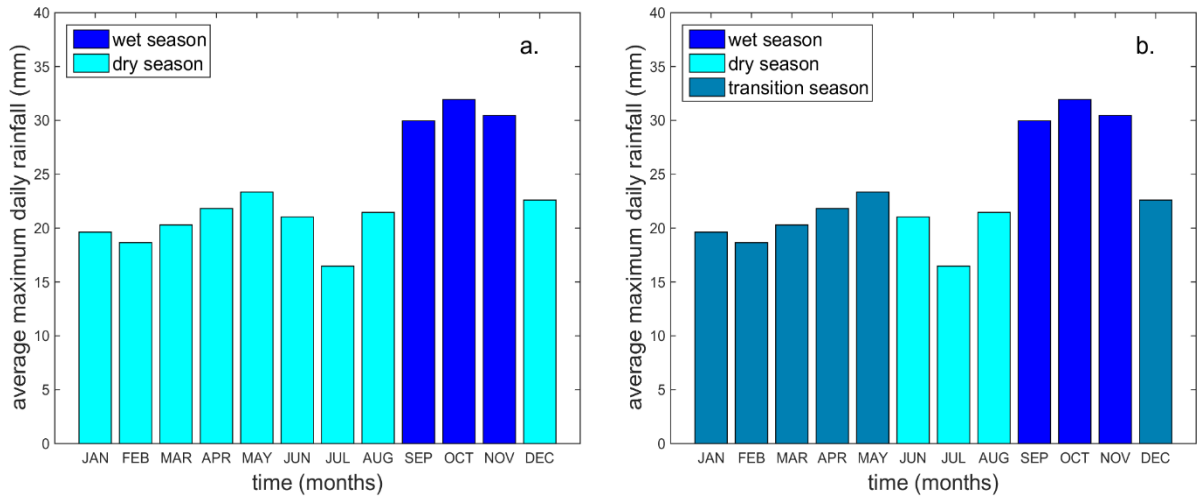


Figure 4.2 Climatograms showing the partition in two seasons (a) and three seasons (b) after application of the SSD clustering algorithm for the station of Florence.

Table 4.1 ΔAIC differences among the seasonal models (one, two or three seasons) under Average Based (AB) and Complete Data (CD) methods. A zero ΔAIC value indicates the model with the smallest AIC value which stands for the best model.

Stations	AB method			CD method					
	Uniform distribution			Weibull distribution			Gamma distribution		
	Number of seasons			*Number of seasons*					
	1	2	3	1	2	3	1	2	3
Bologna	0	0.047	2.9	27.55	27.51	35.12	0	4.91	9.04
Palermo	5.29	0	1.72	0	50.32	41.37	0.372	9.6	5.15
Mantova	0	0.414	3.92	28.22	25.76	13.43	0	8.58	7.12
Milan	0	0.639	3.86	8.633	10.75	0	30.72	33.5	34.7
Genoa	3.38	0	0.67	9.215	0	6.641	5.797	10.5	18
Florence	1.85	0	3.62	15.52	8.421	0	48.55	54.8	9.9
Padua	0	1.058	4.914	0	9.72	11.12	100.19	56.43	65.84
Newcastle	0.75	0	3.88	60.23	34.25	40.68	8.414	0	15.7
Deniliquin	0	2.894	6.41	18.11	16.3	20.18	0	3.68	9.11
Melbourne	0	0.903	4.58	139.1	76.04	78.35	37.31	0	5.27
Robe	1.45	0	5.09	3.475	36.71	40.45	0	1.7	7.29
Sydney	0	0.556	2.88	37.92	40.93	41.35	0	0.02	2.24
Jena Sternwarte	4.6	0	3.69	208.1	123.2	131.6	46.77	0	1.45
Hohenpeissenberg	5.5	0	3.25	85.83	63.77	67.7	0	8.93	6.95
Armagh	0	0.78	3.71	161.3	103.2	106.9	3.412	0	3.04
Radcliffe	0	0.682	3.68	129.6	70.5	77.79	0	0.74	0.75
Zagreb	0.5	0	3.68	42.95	17.99	23.94	0	19.9	25.87
Vlissingen	0.66	0	3.34	78.03	36.79	36.18	0	4.02	9.84
Eelde	1.79	0	3.55	135.8	61.08	67.77	3.338	0	6.24
Den Helder	0.36	0	3.7	201.4	137.5	118.7	27.93	3.58	0
Helsinki	1.9	0	3.69	108.9	48.3	35.58	1.161	0	6.95
Lisbon	8.18	0	7.07	58.26	0	3.646	72.71	7.33	3.04
Prague	2.7	0	1.44	133.3	64.84	58.67	36.75	0.06	0
Uppsala	4.73	0	3.17	187	72.35	58.84	27.02	0.38	0
Stykkisholmur	0	0.657	4.39	103.1	75.99	82.57	2.13	0	6.42
Athens	8.11	0	2.81	104	19.73	23.16	65.54	0	1.45
Toronto	0	1.704	5.14	183.5	121.3	113.1	18.24	0	2.2

To inspect the spatial coherence of the results, we present maps of the two regions of the dataset having neighboring stations, i.e. Europe and Australia (Figure 4.3). We group the stations in six clusters of similarity in their seasonal patterns and we also mark single stations for which similarity falls below the accepted threshold. As similarity index we define the ratio of the number of the wet season months that the stations in the cluster have in common versus the span of each wet season and we require it to be at least 60% for each station in the cluster. More specifically, Clusters 1 and 2 have 67% and 80% similarity, respectively, for both methods, while, Clusters 3, 4 and 5 exhibit 100%-75%, 60%-75% and 80%-67% for the AB and CD methods,

respectively. On top of the maps, we also plot Köppen maps of climate classification by Chen and Chen (2013) covering the period 1901-2010, in order to allow a direct comparison of the observed spatial patterns to the climatological ones. Some interesting insights can be derived. First, spatial coherence does not fully coincide with climatological coherence and vice versa, and this is especially true in regions with complex topography/climatology. For example, in the wider Alpine region, where climate shows great diversity, the stations are less spatially consistent than in Central Europe. On the contrary, stations belonging to a Mediterranean climate (Cluster 1) show consistent patterns. In general, we notice that patterns are coherent on both levels: neighboring stations show very high similarity (e.g. Cluster 3) and far apart stations belonging to a climatically homogenous region show medium to high similarity (see, e.g., Cluster 2).

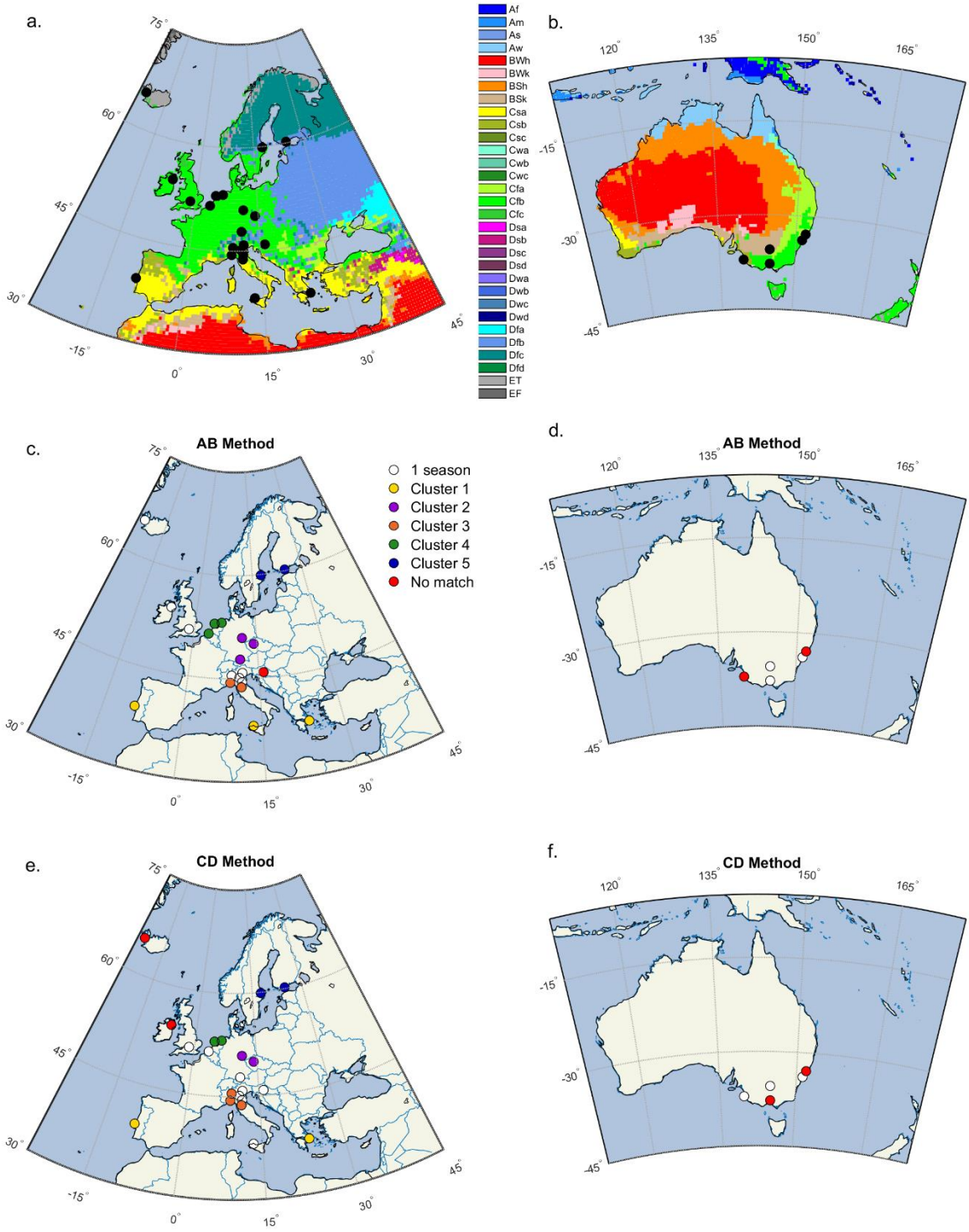


Figure 4.3 Spatial and climatological coherence of the identified seasons for the regions of Europe (a,c,e) and Australia (b,d,f). Figures a,b show the location of the stations on a Köppen climatological map, while the rest show the stations clustered by similarity. White dots represent stations having one season; the remaining dots denote stations having at least 60% overlap of months belonging to the wet season. Red dots denote stations with a lower percentage of similarity to their neighboring stations.

4.5.2 Assessing temporal change in observed seasonality

To demonstrate the applicability of the proposed season identification method in the inspection of temporal changes in seasonality, we analyze the four longest records of

the dataset, i.e. the stations of Padua (275 years), Prague (211 years), Bologna (195 years) and Radcliffe (188 years). We split the observation period into equally sized sub-periods and apply the methodology independently to each period. We employ four sub-periods for the significantly longer station of Padua and three for the other records.

Results are shown in Table 4.2. It can be seen that changes, both in the number and duration of seasons, are likely to emerge within each sub-period. For example, seasonality in Prague during the 2nd period changed in terms of the span of the wet season, but a two-season regime was selected for all sub-periods. Results for 3rd and 1st window coincide. These characteristics of the methodology make it useful for analysis of climatic changes.

Table 4.2 Temporal changes in seasonality identified by application of Average Based (AB) and Complete Data (CD) methods for non-overlapping sub-periods for the four longest stations of the dataset. For the longer station of Padua, an additional sub-period is investigated (4th window).

Station	Record length	Number of Seasons							
		1st window		2nd window		3rd window		4th window	
		Method		Method		Method		Method	
		AB	CD	AB	CD	AB	CD	AB	CD
Padua	1725-2013	1	1	1	1	1	1	1	1
Bologna	1813-2007	1	1	2	1	1	1	–	–
Radcliffe	1827-2014	1	1	1	1	1	1	–	–
Prague	1804-2014	2	1	2*	2*	2	2	–	–
		Span of wet season in months for Prague							
		5-8	–	5-9	5-9	5-8	5-8	–	–

4.5.3 Fitting the GEV distribution

Subsequently to the identification of seasons, we fitted the GEV distribution via maximum likelihood (ML) estimation to each of the seasonal sets (or the annual set if one season was identified). Table 4.3 contains summary statistics of the GEV fitting for wet and dry seasons, as well as for the whole year, for the cases where the two- and three-season model were found prevalent under AB and/or CD methods. Summary statistics for the transition season (placed between the wet and dry season) in the three season model are omitted since the sample is small (5 stations). The main differences in the seasonal distributions lie in the values of the scale and location parameters, which are in their vast majority (93.8% and 100%, respectively, under AB method and 100% and 100%, respectively, under CD method) higher for the wet season compared to the dry. What might be less anticipated is that there is limited seasonal variation in the value of the shape parameter κ, which is related to the shape of the tail of the seasonal maxima distribution. Hence, it is justifiable to represent the two seasons and the whole year by a common value for the shape parameter, therefore increasing robustness of the method, which is a desirable feature. Additionally, for the majority of the stations, the shape parameter takes positive values indicating the

appropriateness of heavier-tailed distributions for modelling of extremes. It is also clear that the wet extreme properties are quite close to the annual maxima ones, which indicates that the annual maxima distribution is dominated by the wet season.

Table 4.3 Comparative statistics of the GEV annual and seasonal parameters, i.e., shape parameter κ, scale parameter σ and location parameter ψ, as estimated via Maximum Likelihood method for the stations in which two or three seasons are identified by Average Based (AB) and Complete Data (CD) methods. The last column of each table shows the percentage (%) of stations in which the parameter value for the wet season is higher than the corresponding value for the dry season.

AB method (16 stations)						CD method (16 stations)			
Parameter		Annual	Wet Season	Dry Season	(wet>dry) %	Annual	Wet Season	Dry Season	(wet>dry) %
	Mean	0.112	0.091	0.097	62.5	0.115	0.106	0.104	45
κ	Percent Positive	93.8	93.8	87.5	-	93.8	93.8	75	-
σ	Mean	12.207	12.706	8.747	93.8	12.187	13.238	9.287	100
ψ	Mean	39.265	35.602	23.772	100	39.652	40.998	34.33	100

The singular cases of the stations of Prague from Czech Republic and Florence from Italy are plotted in Figures 4.4a and 4.4b. In the second case, there is small deviation between the wet season and the annual period, while in the first case the two lines are almost identical. In both cases, the dry-season probability line lies considerably lower. In the second case, in which the three-season model is preferred by the CD method (while two seasons were preferred by the AB method), the probability line of the transition season lies in the space between the wet and dry seasons' probability lines, as expected.

.

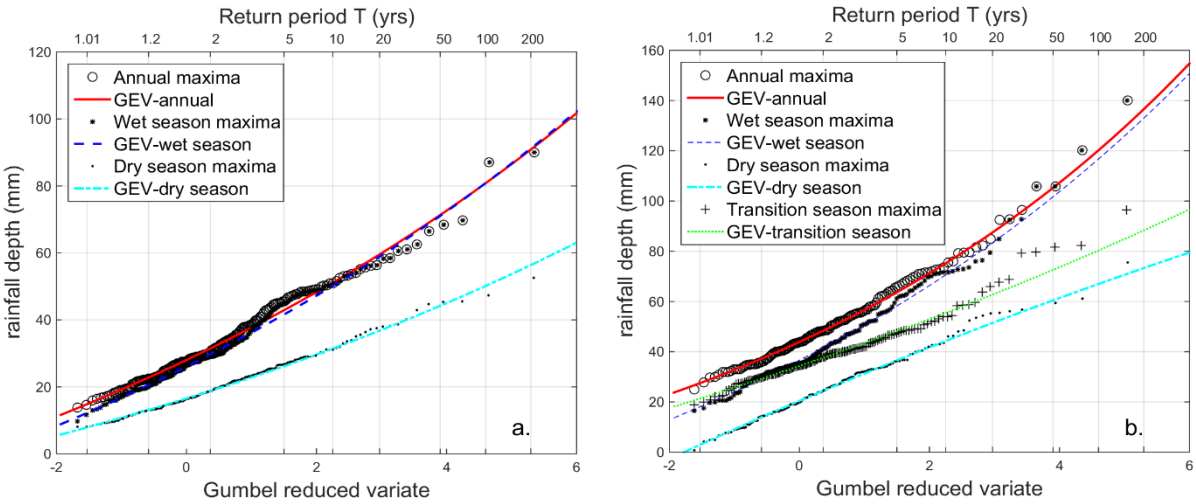


Figure 4.4 Gumbel probability plots of the fitting of the GEV distribution to the annual maxima (red solid line), to the wet season maxima (blue dashed line) and to the dry season maxima (cyan dash-dotted line) for the stations of Prague (a) and Florence (b). For the station of Florence (b), the fitting of the GEV distribution to the transition season maxima (green dotted line) is also shown.

4.5.4 Assessing estimation uncertainty in seasonal-annual GEV parameterization

The crossing over effect mentioned in Section 4.4 is observed in five cases (Eelde, Genoa, Hohenpeissenberg, Milan and Zagreb), where we found that the wet-season probability distribution lies higher than the annual one in the area of extremely rare events. We focus on the station in Genoa where the effect is more pronounced. We perform additional parameter estimation by applying the method of moments (MM) and the least squares algorithm (LS). Figure 4.5a shows results from the application of the three estimation methods for the annual maximum series along with uncertainty bounds computed within each method by means of Monte Carlo analysis. Uncertainty bounds in the area of extremely rare events, where the crossing over effect is also observed, are large. The larger annual maxima fall within the 95% limits of the annual maxima GEV distribution only for the LS method. This is due to the better fitting capability of the LS algorithm for extremely rare events (Koutsoyiannis, 2004d). To further improve the fitting we also estimate via LS a common shape parameter for the three distributions (two seasonal GEV and the annual one). In these cases as well, the choice of a common shape parameter is supported by empirical evidence from the previous independent fitting. The crossing over effect is significantly mitigated (Figure 4.5b), with a remaining positive difference between the quantiles of the wet season and annual distribution of 10 mm for the 0.5% annual exceedance probability, which is considered not significant in view of the large uncertainty in the high-quantile domain. The results for the other cases also showed that the crossing over effect was resolved.

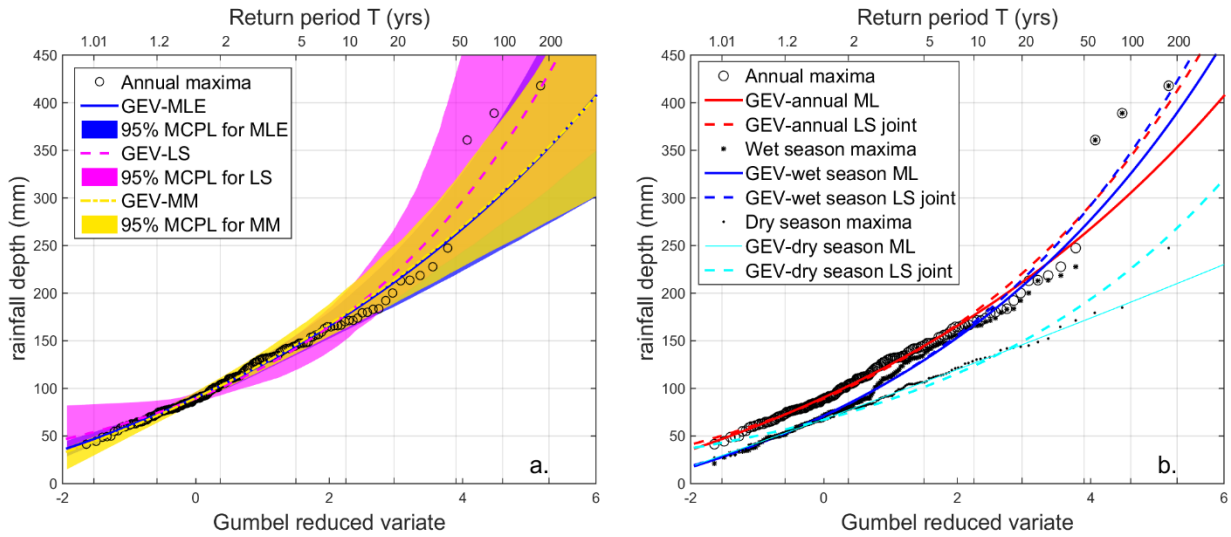


Figure 4.5 Gumbel probability plot of the fitting of the GEV distribution to the annual maxima by the maximum likelihood method (blue color), least-squares method (magenta color) and method of moments (yellow color) along with 95% Monte Carlo Prediction Limits (MCPL) for each method for the station of Genoa (a). The crossing over distance observed in the area of high return periods, where the wet-season probability line (blue solid line) crosses the annual probability line (red solid line), is greatly eliminated when a common shape parameter is employed via the least-squares method (b).

The importance of taking estimation uncertainty into consideration is additionally showcased by applying ML, MM and LS estimation methods to the entire set of stations, as shown in Table 4.4. One notices that uncertainty is higher in the estimation of the shape parameter, as already discussed in literature (Koutsoyiannis, 2004d; Papalexiou and Koutsoyiannis, 2013). The fact that this result is empirically confirmed for the long rainfall records considered here is a further confirmation that for practical applications uncertainty in the estimation of extremes is unavoidable even when dealing with long records.

Table 4.4 Statistics of the GEV parameters, i.e., shape parameter κ, scale parameter σ and location parameter ψ, as estimated for the Annual Maxima series for all stations (27) via Maximum Likelihood (ML), method of moments (MM) and Least Squares method (LS).

Parameter of the annual model		ML	MM	LS
κ	Mean	0.099	0.062	0.120
	Percent Positive	92.6	88.9	96.3
σ	Mean	12.638	10.500	12.732
ψ	Mean	40.510	42.246	40.295

4.5.5 A comparison to traditional methods of seasonal clustering

We compare our method to the climatological 4-season approach, which divides the annual period in Winter, Spring, Summer and Fall seasons. First, to highlight that site-specific season identification is important, we compare the monthly maxima plots for two stations in Europe for our method and the fixed seasonal partition (Figure 4.6). It is clear that climatological seasons are an inefficient partition for analyzing the extreme rainfall properties, and may also be a rather crude method for delineating the extreme's properties in multi-site analyses where seasonal differences in climate may be very pronounced among stations. As an example, seasonality of maximum rainfall in Jena (Germany) is completely out of phase with respect to Athens (Greece). The same could be argued for trend studies employing fixed characterizations of seasonality. For instance, the question of whether winter rainfall has increased is potentially ill-conceived, as it mostly pertains to a subjective interpretation of seasonality. A more relevant question is whether rainfall in the major rainy season has significantly changed, but such a change is unlikely to be identified by considering an arbitrary partition in seasons.

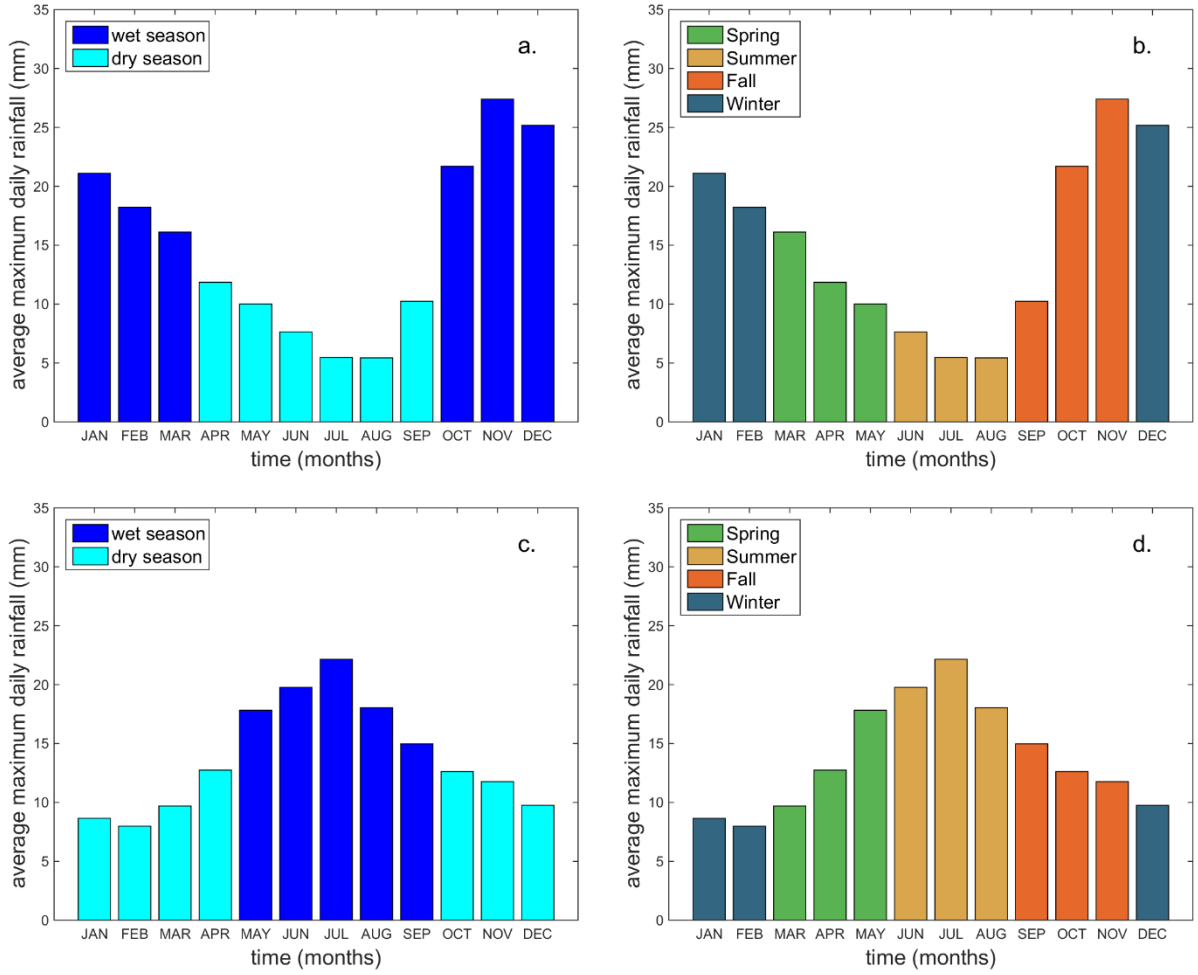


Figure 4.6 Partition in seasons resulting from application of the proposed season identification method versus the fixed 4-season partition for the stations of Athens (a, b respectively) and Jena (c, d respectively).

To demonstrate the effect that a fixed 4-season partition could have on the estimation of extreme value properties, we focus on the rainfall record of Athens (Figure 4.7). By applying the 4-season partition one obtains an apparent overfitting, as the seasonal lines are not clearly separated and even cross each other at several points (Figure 4.7b). It is evident that an inappropriate characterization of seasonality provides no valuable and practical information for seasonal planning and decision-making while, in fact, it obscures the presence of the existing seasonal regime (Figure 4.7a). Additionally, in the presence of parameter uncertainty and given the short record lengths that are usually available, adopting subjective characterizations of seasonality for the study of extreme values entails the risk of disproportionately increasing estimation uncertainty. The consequences of overfitting are even more obvious in stations with very low or no seasonality.

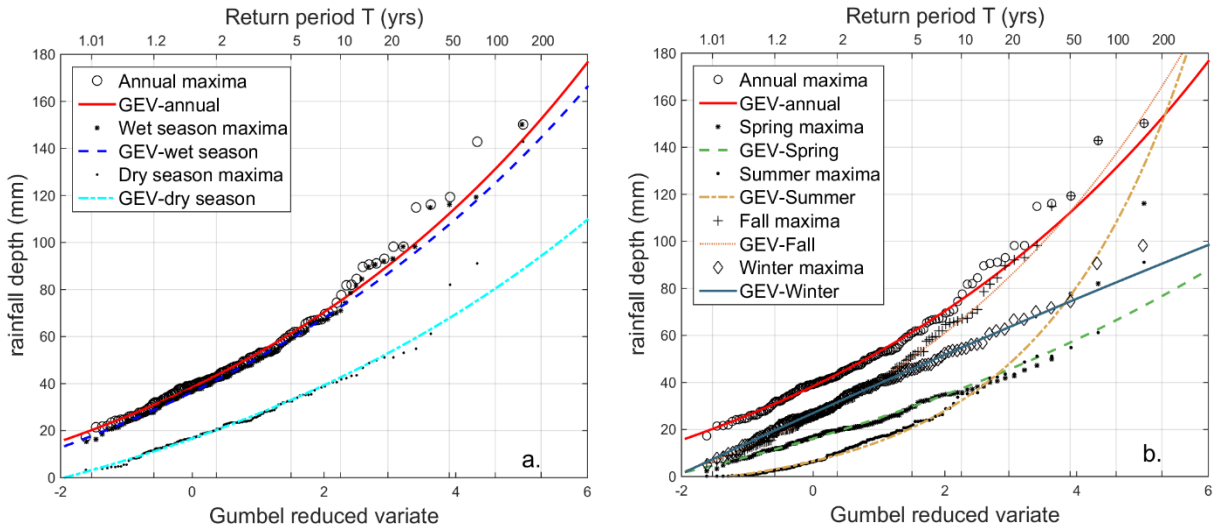


Figure 4.7 Gumbel probability plots of the fitting of the GEV distribution to the annual and seasonal maxima for the station of Athens resulting from (a) the proposed season identification method and (b) from the fixed 4-season partition.

4.6 Discussion and Conclusions

An objective methodology is proposed to allow season identification in extreme daily rainfall and the study of the resulting extreme properties in each season. The methodology is evaluated on an extended dataset comprising 27 rainfall stations covering a period of more than 150 years of daily observations. In the following, we discuss methodological and modelling issues, the results of the extreme value analysis and their comparison to the no-seasonality approach, as well as relative strengths and potential limitations of our method.

The season identification methodology herein proposed is based on the SSD algorithm, a simplified version of the *k*-means clustering algorithm, whose results are evaluated by exploiting the model selection properties of the Akaike Information Criterion (AIC). The method is able to identify the optimal modelling option for the seasonal extreme rainfall for a given dataset, discerning among the existence of 1 (no dominant season) to 4 seasons in the extreme rainfall properties and identifying their temporal span. Since AIC is a measure of relative performance of models, this task should be performed after thorough consideration of the appropriateness of the candidate seasonal distributions to be assessed. In that respect, our methodology

provides additional flexibility as multiple probabilistic models may be simultaneously assessed. Overall, the methodology shows good spatial coherence, which makes it potentially appropriate for regionalization studies, and its flexibility allows one to inspect temporal changes in a range of ways, which is also a desirable feature concerning climatic variability and trend studies.

In terms of generated results, the adopted scheme proved to be successful for the long rainfall records considered here, by both visual evaluation of the plots of the monthly maximum rainfall values (climatograms) and assessment of the resulting extreme seasonal distributional properties. For the cases where two or three seasons are identified, the differences in the distributional properties are reflected mainly in the value of the scale and location parameters of the GEV which are significantly higher for the wet season. The shape parameter shows limited seasonal variability, which implies that the seasonal distributional properties do not differ substantially in the shape of the distribution tail. Our results also confirm other studies regarding the prevalence of heavy-tailed distributions for daily rainfall extremes (Koutsoyiannis, 2004d; Villarini, 2012; Papalexiou and Koutsoyiannis, 2013; Serinaldi and Kilsby, 2014; Mascaro, 2018). Some of these studies have also argued that a positive shape parameter emerges for extremes caused by multiple types of synoptic patterns, whereas a zero exponent (i.e. an exponential tail) may occur for a single-type of events. Apart from pronounced intra-annual variability, a positive shape parameter may be also portraying increased inter-annual variability in the extremes which has been linked to the presence of large-scale circulation patterns, i.e. the NAO, for certain stations of our dataset (Kutiel and Trigo, 2014; Marani and Zanetti, 2015a). In principle, we believe that our findings are in agreement with previous research and strengthen the assumption that a heavier-tail behaviour better captures conditions of enhanced natural variability and complex atmospheric forcing, as revealed by the inspection of our long and spatially sparse dataset.

In comparison to the no-seasonality approach, in some cases the annual maxima series are found to be dominated by extreme events occurring in the wet season. This result is pointed out by the closeness in the estimated GEV parameter values between the annual and the wet season's probability distribution of extreme events. It also indicates that annual frequency analyses may suffice for studying the annual maxima (AM). Actually, studying the AM series is more in favour of a conservative design approach, since the former takes into account the rare cases of extreme events of significant magnitude happening in the dry season. Furthermore, since the majority of AM in records with pronounced seasonality still stems from the wet season, strong seasonality is not significantly violating the IID assumption in the GEV approach. A similar remark was also made by Allamano et al. (2011). However, for intra-annual hydrological design and management, it is crucial to take seasonal variability into account. The wet season maxima series contain valuable information on the timing of occurrence of the most extreme events, although it is likely that in some cases, their magnitude will be close to the AM estimated one. Yet when dry periods are of interest, using the AM series instead, i.e. adopting a no-seasonality approach, is likely to lead to costly overestimation of design values and floodwater waste.

A few key strengths of our methodology should be underlined. In general, estimation uncertainty in extreme studies is a known issue especially for the estimation of the shape parameter of the GEV distribution. Here, we show how an alternative choice of estimation methods, improving the model performance in the domain of extreme events, may resolve inconsistencies deriving from an independent seasonal and annual fitting. Given the latter, we consider the need for the laborious joint estimation of seasonal-annual distributions to be questionable and we propose a simpler procedure based on the estimation of a common shape parameter for the seasonal-annual parameterization, which is shown to increase robustness of the statistical model. On the whole, the entire methodology is compared to a conventional partition in fixed seasons and its advantageous features are highlighted both in that it enables consistent identification of seasonal regimes at single-site and multi-site levels, as opposed to arbitrary partitions, and that it consequently allows a more informed and parsimonious fitting of the GEV distribution to seasonal extremes.

A few limitations should be taken into account. We note that in case where the Average Based (AB) and the Complete Data (CD) methods diverge, there is some remaining degree of subjectivity in the choice for the most appropriate scheme. This constitutes a potential limitation of our method as results may not be fully conclusive. Yet this may be resolved if an equifinality framework is adopted and both options are considered. Additionally, it should be noted that the performance of AIC largely depends on the quality of the considered candidate models. Although the chosen distributions are representative of a variety of statistical behaviours, it is possible that there may be exceptions for which they do not perform well. Increasing the set of candidate distributions is another option to achieve a greater degree of confidence within a multi-model approach.

Despite these limitations, we believe that our findings have direct applications both in the theoretical conceptualization of seasonality in extreme rainfall and in engineering applications. On a methodological level, they contribute to a wider establishment of model selection techniques, in this case AIC, in hydrological studies and pave the way for the objective identification of seasonality via automated schemes which are required for global-scale hydrology.

5. SEASONAL DEPENDENCE DYNAMICS OF STREAMFLOW EXTREMES

It has been shown that the geophysical and hydrological processes governing river flow formation exhibit persistence at several timescales, which may manifest itself with the presence of positive seasonal correlation of streamflow at several different time lags (Aguilar et al. 2017). This Chapter builds upon this idea dealing with the presence of dependence dynamics in river flow at the seasonal scale, the associated physical drivers, and the related potential for employing this information to improve probabilistic prediction of high and low flows. A dataset of 224 rivers from six European countries spanning more than 50 years of daily flow data is exploited. The practical benefit of the methodology is demonstrated by updating the frequency distribution of high and low flows one season in advance in a real-world case. Results suggest that there is a traceable physical basis for river memory which, in turn, can be statistically assimilated into high- and low-flow frequency estimation to reduce uncertainty and improve predictions for technical purposes.

5.1 Introduction

Recent analyses for the Po River and the Danube River highlighted that catchments may exhibit significant correlation between peak river flows and average flows in the previous months (Aguilar et al., 2017). Such correlation is the result of the behaviours of the physical processes involved in the rainfall–runoff transformation that may induce memory in river flows at several different timescales. The presence of long-term persistence in streamflow has been known for a long time since the pioneering works of Hurst (1951a) and has been actively studied ever since (e.g. Koutsoyiannis, 2011b; Montanari, 2012; O'Connell et al., 2016 and references therein). While a number of seasonal flow forecasting methods have been explored in the literature (e.g. Bierkens and van Beek, 2009; Dijk et al., 2013), attempts to explicitly exploit streamflow persistence in seasonal forecasting through information from past flows have been, in general, limited. Koutsoyiannis et al. (2008) proposed a stochastic approach to incorporate persistence of past flows into a prediction methodology for monthly average streamflow and found the method to outperform the historical analogue method (see also Dimitriadis et al., 2016, for theory and applications of the latter) and artificial neural network methods in the case of the Nile River. Similarly, Svensson (2016) assumed that the standardized anomaly of the most recent month will not change during future months to derive monthly flow forecasts for 1–3 months lead time and found the predictive skill to be superior to the analogue approach for 93 UK catchments. The above-mentioned persistence approach has also been used operationally in the production of seasonal streamflow forecasts in the UK since 2013, within the framework of the Hydrological Outlook UK (Prudhomme et al. 2017). A few other studies have included past flow information in prediction schemes along with teleconnections or other climatic indices (Piechota et al., 2001; Chiew et al., 2003; Wang et al., 2009). Recently, it was shown that streamflow persistence, revealed as seasonal correlation, may also be relevant for prediction of extreme events by allowing

one to update the flood frequency distribution based on river flow observations in the pre-flood season and reduce its bias and variability (Aguilar et al., 2017). The above previous studies postulated that seasonal streamflow correlation may be due to the persistence of the catchments storage and/or the weather, but no attempt was made to identify the physical drivers.

The present study aims to further inspect seasonal persistence in river flows and its determinants, by referring to a large sample of catchments in six European countries (Austria, Sweden, Slovenia, France, Spain, and Italy). We focus on persistence properties of both high and low flows by investigating the following research questions: (i) what are the physical conditions, in terms of catchment properties, i.e. geology and climate, which may induce seasonal persistence in river flow, and, (ii) can floods and droughts be predicted, in probabilistic terms, by exploiting the information provided by average flows in the previous months? These questions are relevant for gaining a better comprehension of catchment dynamics and planning mitigation strategies for natural hazards. To reach the above goals, we identify a set of descriptors for catchment behaviours and climate, and inspect their impact on correlation magnitude and predictability of river flows.

A few studies have analysed physical drivers of streamflow persistence on annual and deseasonalized monthly and daily time series (Mudelsee, 2007; Hirpa et al., 2010; Gudmundsson et al., 2011; Zhang et al., 2012; Szolgayova et al., 2014; Markonis et al., 2018) but the topic has been less studied on intra-annual scales relevant to seasonal forecasting of floods and droughts.

To demonstrate the high practical relevance of the identified seasonal correlations we present a technical experiment for one of the studied rivers (Section 5.7) in which the frequency distribution of both high and low flows is updated one season in advance by exploiting real-time information on the state of the catchment.

5.2 Data and catchment description

The dataset includes 224 records spanning more than 50 years of daily river flow observations from gauging stations, mostly from non-regulated streams. A few catchments are impacted by regulation. Among the 224 rivers, 108 are located in Austria, 69 in Sweden, 31 in Slovenia, 13 in France, two in Spain and one in Italy. Catchment areas vary significantly, the largest being the Po River basin in Italy (70 091 km^2) and the smallest being the Hålabäcken River basin in Sweden (4.7 km^2). The geographical location of the river gauge stations as well as their climatic classification are shown in Fig. 5.1. Most of the examined rivers belong to either a warm temperate (C) or a boreal/snow climate (D) with a subset impacted by polar climatic conditions (E), according to the updated World Map of the Köppen–Geiger climate classification (Fig. 5.1) based on gridded temperature and precipitation data for the period 1951-2000 (Kottek et al., 2006). More specifically, the majority of French and Slovenian and approximately one third of the Swedish basins belong to the warm temperate Cfb category characterized by precipitation distributed throughout the year (fully humid) and warm summers. The rest of the Swedish catchments are impacted by a Dfc climatic type, i.e. a snow climate, fully humid with cool summers. The Austrian catchments belonging to the region impacted by the European Alps have the most complicated

regime due to their topographic variability. At the lowest altitudes, Cfb is the prevailing regime, but as proximity to the Alps increases, a Dfc regime dominates and progressively, in the highest altitude basins, the climate becomes a polar tundra type (Et), characterized primarily by the very low temperatures present. The characteristics of all the climatic regimes of the studied rivers are given in the legend of Fig. 5.1. A summary of the river basins under study in terms of the selected descriptors is also provided in Table 5.1, showing that the investigated rivers cover a wide range of catchment area sizes, flow regimes and climatic conditions.

It is relevant to note that 16 of the Austrian rivers are subject to regulation, which may alter the persistence properties of river flows. This relates to generally 'mild' forms of regulation, i.e. upstream regulation with a very low degree of flow attenuation, hydropower operations and flow diversions to and from the basin. A preliminary examination of these rivers did not reveal any significant change during time of the flow regime. The presence of regulation does not preclude the exploitation of correlation for predicting river flows in probabilistic terms, but it may affect the analysis of physical drivers, as it may enhance or reduce persistence in the natural river flow regime. Given that detailed information is generally lacking on the impact of regulation (Kuentz et al. 2017), we assume stationarity of the river flows for all the catchments herein considered and additionally, assume that river management does not significantly affect the identification of the physical drivers.

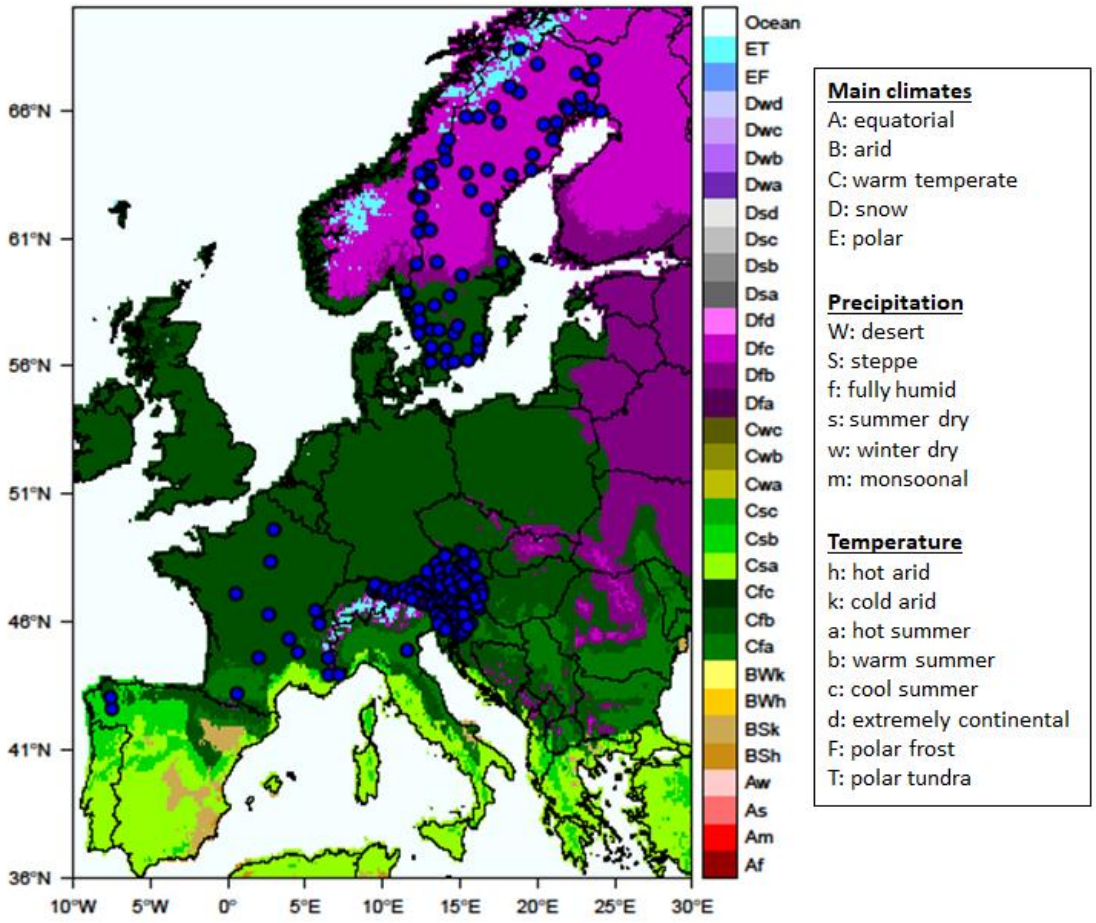


Figure 5.1 Updated Köppen–Geiger climatic map for period 1951–2000 (Kottek et al., 2006) showing the location of the 224 river gauge stations.

Table 5.1 Summary statistics of the river descriptors. Summary statistics for PL, PG and PF variables are computed only for the subset of catchments with positive values (the total number of catchments is also reported in brackets). PK is used as a categorical variable (PK is either higher or lower than 50% of catchment area), therefore sample statistics are not computed in this case, but the number of stations with PK ≥ 50% is reported as 'positive' presence of karst.

Descriptor (Units)	A (km^2)	BI (–)	SR (m^3 s^{-1} km^{-2})	PL (%)	PG (%)	PF (%)	PK (–)	P (mm year^{-1})	T (°C)	IDM (–)
Min value	4.7	0.29	0.004	0.5	0.1	0.3	–	444	–1.8	29.41
Max value	70091	0.99	0.088	19.5	56.5	100	–	1500	13.7	153.40
Standard deviation	5904.3	0.14	0.018	4.04	15.54	32.56	–	288.22	3.59	24.53
Sample size	224	224	224	69 (69)	39 (108)	18 (108)	21 (31)	224	224	224

5.3 Methodology

The investigation of the persistence properties of river flows focuses separately on both high and low discharges and is articulated in the following steps: (a) identification of the high- and low-flow seasons, (b) correlation assessment between the peak flow in the high-flow season (average flow in the low-flow season) and average flows in the previous months; (c) analysis of the physical drivers for streamflow persistence and its predictability through principal component analysis (PCA), (d) real-time updating of the frequency distribution of high and low flows for a selected case study with significant seasonal correlation by employing a meta-Gaussian approach. The above steps are described in detail in the following sections.

5.3.1 Season identification

Season identification is performed algorithmically to identify the high-flow season (HFS) and low-flow season (LFS) for each river time series. For the estimation of HFS, we employ an automated method recently proposed by Lee et al. (2015), which identifies the high-flow season as the 3-month period centred around the month with the maximum number of occurrences of peaks over threshold (POT), with the threshold set to the highest 5% of the daily flows. To evaluate the selection of HFS, a metric constructed as the percentage of annual maximum flows (PAMF) captured in the HFS is used. The PAMFs are classified in the subjective categories of "poor" (<40%), "low" (40–60%), "medium" (60–80%) and "high" (>80%) values, denoting the probability that the identified HFS is the dominant high-flow season in the record. If the identified peak month alone contains more than or equal to 80% of annual maxima flows, a unimodal regime is assumed and the identification procedure is terminated. In all other cases, the method allows for the search of a second peak month and the

identification of a minor HFS, but we do not further elaborate on this analysis here, because we are only interested in the most extreme seasons for the purpose of predicting high and low flows.

The method proposed by Lee et al. (2015) has several advantages that make it suitable for the purpose of this research. Most importantly, it is capable of handling conditions of bimodality, which is usually a major issue for traditional methods like, e.g. directional statistics (Cunderlik et al., 2004a). A potential limitation is the assumption of symmetrical extension of HFS around the peak month, along with the uniform selection of its length (3-month period). The degree of subjectivity in the evaluation of the second HFS is another limitation, which is not relevant here as we focus on the main HFS.

The LFS is herein identified as the 1-month period with the lowest amount of mean monthly flow. An alternative approach of estimating the relative frequencies of annual minima of monthly flow and selecting the month with the highest frequency as the LFS is also considered.

5.3.2 Analysis of streamflow correlation and its physical drivers

Correlation analysis

In the case of HFS, a correlation is sought between the maximum daily flow occurring in the HFS period and the mean flow in the previous months, before the onset of HFS. For LFS, correlation is computed between the mean flow in the LFS itself and the mean flow in the previous months. We use the mean flow in the previous month as a robust proxy of 'storage' in the catchment that is expected to reflect the state of the catchment, i.e., wetter or drier than usual. Since we are interested in seasonal persistence, we compute the Pearson's correlation coefficient for HFS lag up to 9 months and for LFS lag up to 11 months.

Analysis of physical drivers

a. Catchment, geological, and climatic descriptors.

An extensive investigation is carried out to identify physical drivers of seasonal streamflow correlation, in terms of catchment, geological, and climatic descriptors.

As catchment descriptors, we consider the basin area (A), the baseflow index (BI), the mean specific runoff (SR), the percentage of basin area covered by lakes (percentage of lakes PL) and glaciers (percentage of glaciers PG) and altitude as candidates for explanatory variables for streamflow correlation.

The area A (km^2) is primarily investigated, as it is representative of the scale of the catchment, under the assumption that in larger basins the impact of the climatological and geophysical processes affecting river flow becomes more significant and may lead to a magnified seasonal correlation.

The BI is considered based on the assumption that high groundwater storage may be a potential driver of correlation. BI is calculated from the daily flow series of the rivers following the hydrograph separation procedure detailed in Gustard et al. (2009). Flow minima are sampled from non-overlapping 5-day blocks of the daily flow series and turning points in the sequence of minima are sought and identified when the 90% value of a certain minimum is smaller or equal to its adjacent values. Subsequently, linear interpolation is used in between the turning points to obtain the

baseflow hydrograph. The BI is obtained as the ratio of the volume of water beneath the baseflow separation curve versus the total volume of water from the observed hydrograph, and an average value is computed over all the observed hydrographs for a given catchment. A low index is indicative of an impermeable catchment with rapid response, whereas a high value suggests high storage capacity and a stable flow regime.

SR ($m^3\ s^{-1}\ km^{-2}$) is computed as the mean daily flow of the river standardized by the size of its basin area. It may be an important physical driver as it is an indicator of the catchment's wetness. PL (%) and PG (%) are investigated for the Swedish and Austrian catchments, respectively, as lakes and glaciers are expected to increase catchment storage thus affecting persistence. Lake coverage data are based on cartography and are available from the Swedish Water Archive (https://www.smhi.se/), while glacier coverage data are estimated from the CORINE land cover database (https://www.eea.europa.eu/publications/COR0-landcover).

The effect of catchment altitude is also inspected using relief maps from the Shuttle Radar Topography Mission (SRTM) data (http://srtm.csi.cgiar.org/). The data are available for the whole globe and are sampled at 3 arch-seconds resolution (approximately 90 meters). Topographic information is available for all catchments located at latitudes lower than 60 degrees north, while a 1 km resolution digital elevation model is available for Austria.

As geological descriptors we consider the percentage of catchment area with the presence of flysch (percentage of flysch PF) and karstic formations (percentage of karst PK) for Austrian and Slovenian catchments, respectively, where this type of information is available. A subset of Austrian catchments is characterized by the dominant presence of flysch, a sequence of sedimentary rocks characterized by low permeability, which is known to generate a very fast flow response. Karstic catchments, characterized by the irregular presence of sinkholes and caves, are also known for having rapid response times and complex behaviour; e.g. initiating fast preferential groundwater flow and intermittent discharge via karstic springs (Ravbar, 2013; Cervi et al., 2017). Geological features are also presumed to be linked to persistence properties because geology is the main control for the baseflow index across the European continent (Kuentz et al. 2017). PK (%) and PF (%) are estimated from geological maps of Slovenia and Austria, respectively.

As climatic descriptors, the mean annual precipitation P (mm $year^{-1}$) and the mean annual temperature T (°C) are selected. Corresponding gridded data are retrieved from the WorldClim database (http://www.worldclim.org/) at a spatial resolution of 10 arcminutes (approximately 18.55 km at the equator). We note that low mean temperature regimes are also associated with snow, the presence of which is also considered in the interpretation of the results. We also adopt the De Martonne index (IDM; De Martonne, 1926), as a climatic descriptor, which is given by $\mathrm{IDM} = P/(T + 10)$, and enables classification of a region into one of the following 6 climate classes, i.e., arid (IDM ≤ 5), semi-arid (5 < IDM ≤ 10), dry subhumid (10 < IDM ≤ 20), wet subhumid (20 < IDM ≤ 30), humid (30 < IDM ≤ 60) and very humid (IDM ≥ 60). Additionally, the Köppen–Geiger climatic classification (Kottek et al., 2006) of the rivers is assessed.

Principal component analysis

To identify which catchment, physiographic and climatic characteristics may explain river memory we attempt to regress the seasonal streamflow correlation on the physical descriptors introduced above. We expect the presence of multicollinearity among the predictor variables and therefore PCA (Pearson, 1901; Hotelling, 1933) was applied to construct uncorrelated explanatory variables. In essence, PCA is an orthonormal linear transformation of p data variables into a new coordinate system of $q \leq p$ uncorrelated variables (principal components, PCs) ordered by decreasing degree of variance retained when the original p variables are projected into them (Jolliffe, 2002). Therefore, the first principal axis contains the greatest degree of variance in the data, while the second principal axis is the direction which maximizes the variance among all directions orthogonal to the first principal axis and each succeeding component in turn has the highest variance possible while satisfying the condition of orthogonality to the preceding components. Specifically, let $\boldsymbol{x}$ be a random vector with mean μ and correlation matrix $\boldsymbol{\Sigma}$, and the principal component transformation of $\boldsymbol{x}$ is then obtained as follows:

$$\boldsymbol{y} = \boldsymbol{C}^T \boldsymbol{x}' \tag{35}$$

where $\boldsymbol{y}$ is the transformed vector whose kth column is the kth principal component (k =1, 2..p), $\boldsymbol{C}$ is the $p \times p$ matrix of the coefficients or loadings for each principal component and $\boldsymbol{x}'$ is the standardized $\boldsymbol{x}$ vector. Standardization is applied in order to avoid the impact of the different variable units on selecting the direction of maximum variance, when forming the PCs. The $\boldsymbol{y}$ values are the scores of each observation, i.e. the transformed values of each observation of the original p variables in the kth principal component direction.

PCA has useful descriptive properties of the underlying structure of the data. These properties can be efficiently visualized in the biplot (Gabriel, 1971), which is the combined plot of the scores of the data for the first two principal components along with the relative position of the p variables as vectors in the two-dimensional space. Herein, the distance biplot type (Gower and Hand, 1995), which approximates the Euclidean distances between the observations, is used. Variable vector coordinates are obtained by the coefficients of each variable for the first two principal components. After construction of the PCs, a linear regression model is explored for the case of HFS and LFS lag-1 correlation.

5.3.3 Technical experiment: real-time updating of the frequency distribution of high and low flows

In order to evaluate the usefulness of the information provided by the 1-month-lag seasonal correlation for flow signatures in HFS and LFS, we perform a real-time updating of the frequency distribution of high and low flows based on the average river flow in the previous month. A similar analysis for the high flows was carried out by Aguilar et al. (2017) for the Po and Danube Rivers. In principle, this is a data assimilation approach, since real-time information, i.e. observations of the average river flow, is used in order to update a probabilistic model and inform the forecast of the flow signature for the upcoming season.

In detail, a bi-variate meta-Gaussian probability distribution (Kelly and Krzysztofowicz, 1997; Montanari and Brath, 2004) is fitted between the observed flow signatures, i.e., peak flow in the HFS, Q_P and average flow in the LFS, Q_L, and the average flow in the pre-HFS and LFS months, Q_m, respectively. The peak HFS flow and the average LFS flow are the dependent variables and are extracted as the peak river discharge observed in the previously identified HFS and the average river discharge observed in the previously identified LFS, respectively. The average flow in the month preceding the HFS and the LFS is the explanatory variable in both cases. In the following, random variables are denoted by underscore and their outcomes are written in plain form.

The normal quantile transform (NQT; Kelly and Krzysztofowicz, 1997) is used in order to make the marginal probability distribution of dependent and explanatory variables Gaussian. This is achieved as follows: a) the sample quantiles Q are sorted in increasing order e.g. Q_{m_1}, Q_{m_2}...Q_{m_n} , b) the cumulative frequency, e.g. FQ_{m_i} is computed via a Weibull plotting position, and c) the standard normal quantile, e.g., NQ_{m_i} is obtained as the inverse of the standard normal distribution for each cumulative frequency, e.g., $G^{-1}(FQ_{m_i})$. Therefore, all sample quantiles are discretely mapped into the Gaussian domain. To get the inverse transformation for any normal quantile, we connect the points in the above mapping with linear segments. The extreme segments are extended to allow extrapolation outside the range covered by the observed sample.

In the Gaussian domain, a bi-variate Gaussian distribution is fitted between the random explanatory variable $\underline{NQ}_m$ and the dependent variables $\underline{NQ}_P$ and $\underline{NQ}_L$ by assuming the stationarity and ergodicity of the variables. We define the generic random variable $\underline{NQ}_{fs}$ to represent any dependent flow signature, i.e.; $\underline{NQ}_P$ and $\underline{NQ}_L$ in our case. Then, the predicted signature at time t can be written as:

$$\underline{NQ}_{fs}(t) = \rho\left(\underline{NQ}_m, \underline{NQ}_{fs}\right)\underline{NQ}_m(t-h) + N\varepsilon(t) \tag{36}$$

where $\rho\left(\underline{NQ}_m, \underline{NQ}_{fs}\right)$ is the Pearson's cross-correlation coefficient between $\underline{NQ}_m$ and $\underline{NQ}_{fs}$, h is the selected correlation lag with $h = 1$ in the present application, and $N\varepsilon(t)$ is an outcome of the stochastic process $\underline{N\varepsilon}$, which is independent, homoscedastic, stochastically independent of $\underline{NQ}_m$ and normally distributed with zero mean and variance $1 - \rho^2\left(\underline{NQ}_m, \underline{NQ}_{fs}\right)$. Then, the joint bi-variate Gaussian probability distribution function is defined by the mean ($\mu(\underline{NQ}_m) = 0$ and $\mu(\underline{NQ}_{fs}) = 0$), the standard deviation ($\sigma(\underline{NQ}_m) = 1$ and $\sigma(\underline{NQ}_{fs}) = 1$) of the standardized normalized series, and the Pearson's cross correlation coefficient between the normalized series, $\rho\left(\underline{NQ}_m, \underline{NQ}_{fs}\right)$. From the Gaussian bi-variate probability properties, it follows that for any observed $\underline{NQ}_m(t-h)$, the probability distribution function of $\underline{NQ}_{fs}(t)$ conditioned on NQ_m is Gaussian, with parameters given by:

$$\mu(\underline{NQ}_{fs}(t)) = \rho\left(\underline{NQ}_m, \underline{NQ}_{fs}\right)\underline{NQ}_m(t-h) \tag{37}$$

$$\sigma(\underline{NQ}_{fs}(t)) = \left(1 - \rho^2\left(\underline{NQ}_m, \underline{NQ}_{fs}\right)\right)^{0.5} \quad (38)$$

To derive the probability distribution of $Q_{fs}(t)$ conditioned to the observed $Q_m(t-h)$, we first apply the inverse NQT, i.e., we use linear segments to connect the points of the previous discrete quantile mapping of the original quantiles into the Gaussian domain, and accordingly, obtain $Q_{fs}(t)$ for any $NQ_{fs}(t)$. Subsequently, we estimate the parameters of an assigned probability distribution for the obtained quantiles in the untransformed domain. This is referred to as the updated probability distribution of the considered flow signature ($\underline{NQ}_P$ and $\underline{NQ}_L$ in our case). We use the Extreme Value Type I distribution for the peak flows and calculate the differences in the magnitude of estimated maxima for a given return period between the unconditioned and the updated distribution. The latter is conditioned by the 95% sample quantile of the observed mean flow in the previous month. To model the low flows we use the lognormal distribution, which was found to exhibit the best fit for the river in question among other typical candidates for average flows, i.e. the Weibull and Gamma distribution. The low flows are conditioned by the lower 5% sample quantile of the observed mean flow in the previous month.

5.4 Seasonal correlation of high and low flows

5.4.1 Season identification

Approximately half of the 224 rivers are characterized by at least one high-flow season with medium or higher significance (PAMF of HFS ≥ 60%). Among them, very strong unimodal regimes (PAMF of HFS ≥ 80%) are observed in 63 rivers, the majority of which are located in Sweden. For 25% of the rivers, a high-flow season of low significance is found (PAMF of HFS between 40–60%), while for the remaining 25% the high-flow distribution looks uniform throughout the year. Bimodality regimes are found with low and moderate significance in rivers located mostly in Austria and Sweden, but we focus here on the major high-flow season, as we are interested in the most extreme events. A minor HFS analysis would be perhaps relevant in other regions of the world where bimodal flood regimes are more prominent, as suggested by the analysis of Lee et al. (2015).

Regarding the LFS identification, the two considered approaches (see Section 5.3.1) agree for 139 out of 224 stations but the first method, i.e. the 1-month period with the lowest amount of mean monthly flow is selected for being more relevant to the purpose of computing mean flow correlations.

5.4.2 Seasonal correlation

LFS correlation is markedly higher than the corresponding HFS correlation for lags 1–5 and its median remains higher than 0 for more lags (Fig. 5.2). For the case of HFS correlation, we focus only on the most significant first lag, for which 73 rivers are found to have correlation significantly higher than 0 at 5% significance level. In Fig. 5.3, the autocorrelation of the whole monthly series is compared to the LFS correlation for lag of 1 and 2 months, in order to prove that the seasonal correlation for LFS is

significantly higher than its counterpart computed by considering the whole year. The latter is also confirmed by the Kolmogorov–Smirnov test for both LFS lags (corresponding p values, $p_{lag1} < 2.2 \times 10^{-6}$ and $p_{lag2} < 2.2 \times 10^{-6}$ for the null hypothesis that the LFS correlation coefficients are not higher than the corresponding values for the monthly series autocorrelation; Conover, 1971).

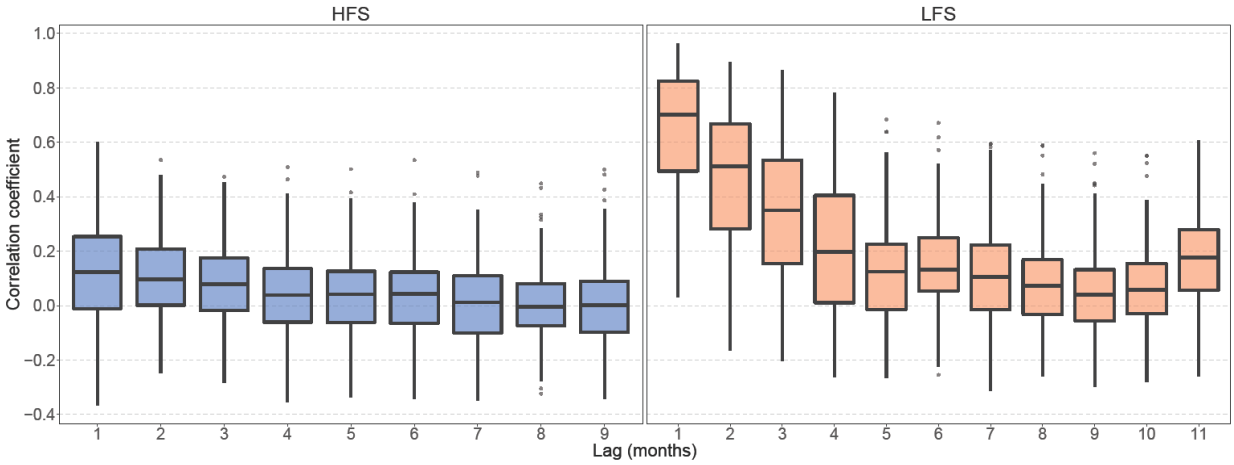


Figure 5.2 Boxplots of seasonal correlation coefficient against lag time for HFS (left panel) and LFS (right panel) analysis for the 224 rivers. The lower and upper ends of the box represent the 1st and 3rd quartiles, respectively, and the whiskers extend to the most extreme value within 1.5 IQR (interquartile range) from the box ends; outliers are plotted as filled circles.

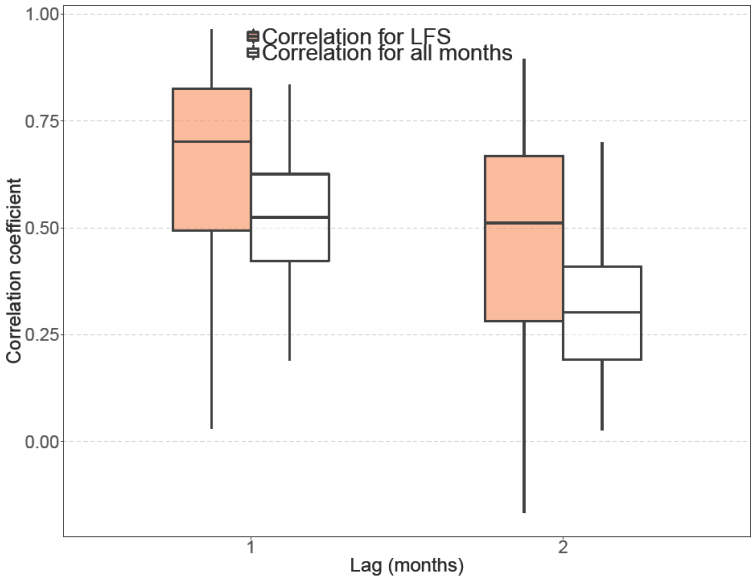


Figure 5.3 Boxplots of lag-1 and lag-2 correlation coefficients for LFS analysis (orange) and the whole monthly series (white) for the 224 rivers. The lower and upper ends of the box represent the 1st and 3rd quartiles, respectively, and the whiskers extend to the most extreme value within 1.5 IQR (interquartile range) from the box ends.

Figure 5.4 shows the spatial pattern of HFS and LFS streamflow correlations. It is interesting to notice the emergence of spatial clustering in the correlation magnitude, which implies its dependence on different spatially varying physical mechanisms. For

example, for HFS, a geographical pattern emerges within France, since the highest correlation coefficients are located in the northern part of the country, which is characterized by an oceanic climate and higher baseflow indices.

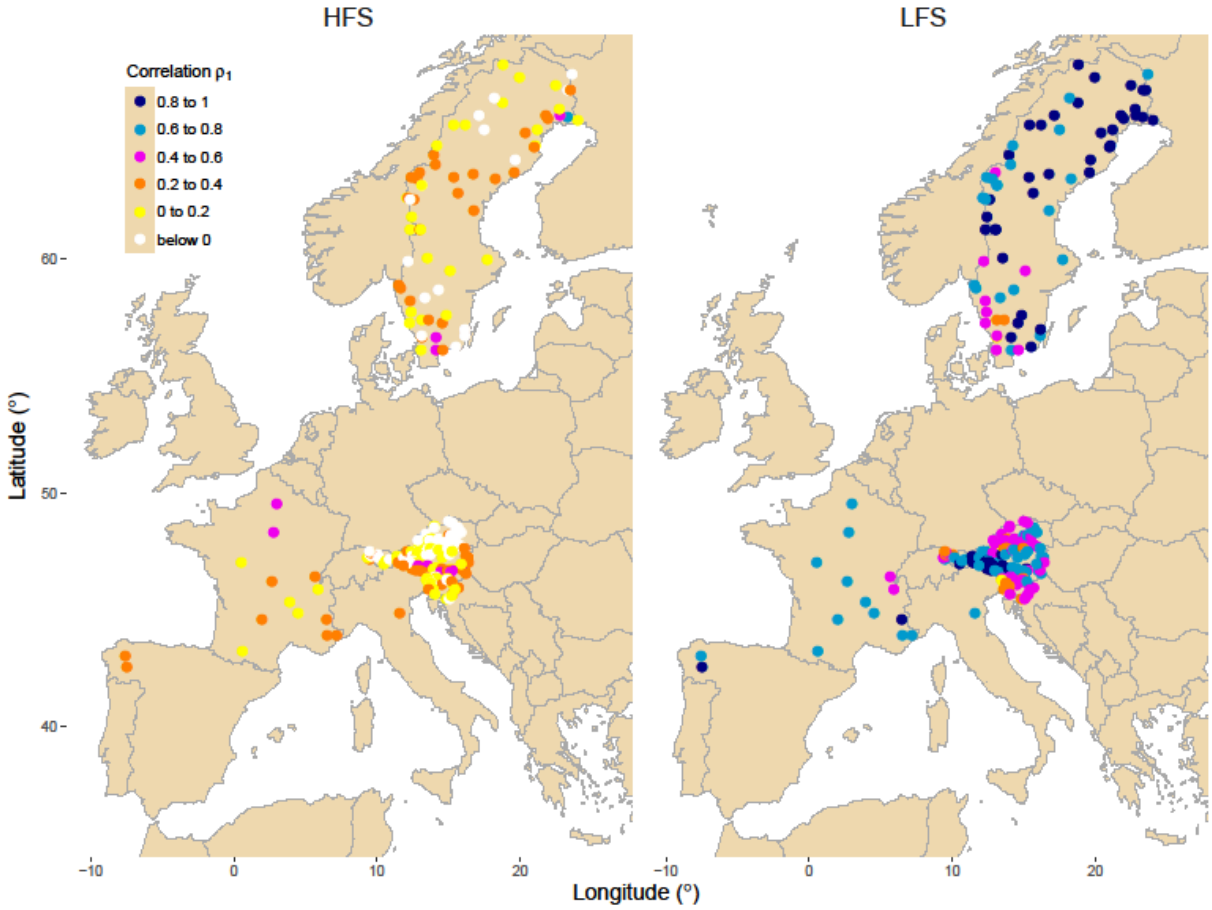


Figure 5.4 Spatial distribution of the lag-1 correlation coefficients for HFS (left) and LFS (right) analysis. Legend shows the colour assigned to each class of correlation for the data.

5.5 Physical interpretation of correlation

To attribute the detected correlations to physical drivers, we define six groups of potential drivers of seasonal correlation magnitude, which are: basin size, flow indices, the presence of lakes and glaciers, catchment elevation, catchment geology, and hydro-climatic forcing. For some of the descriptors the information is only available for a few countries.

In what follows, we will use the term "positive (negative) impact on correlation" to imply that an increasing value of the considered descriptor is associated with increasing (decreasing) correlation. For each descriptor, we also report, between parentheses, the Spearman's rank correlation coefficient r_s (Spearman, 1904) between its value and the considered (LFS or HFS) correlation, and the p value of the null hypothesis r_s=0. Spearman's coefficient is adopted in view of its robustness to the presence of outliers and its capability of capturing monotonic relationships of non-linear type.

Catchment area – descriptor A

Figure 5.5 shows that there is only a weak positive impact of the catchment area (log-transformed) on correlation for HFS (r_s= 0.17, p = 0.01) but a more significant positive one for LFS (r_s = 0.27, p = 5.5 × 10^{-5}). The presence of relevant scatter in the plots also indicates that it is not a key determinant of correlation.

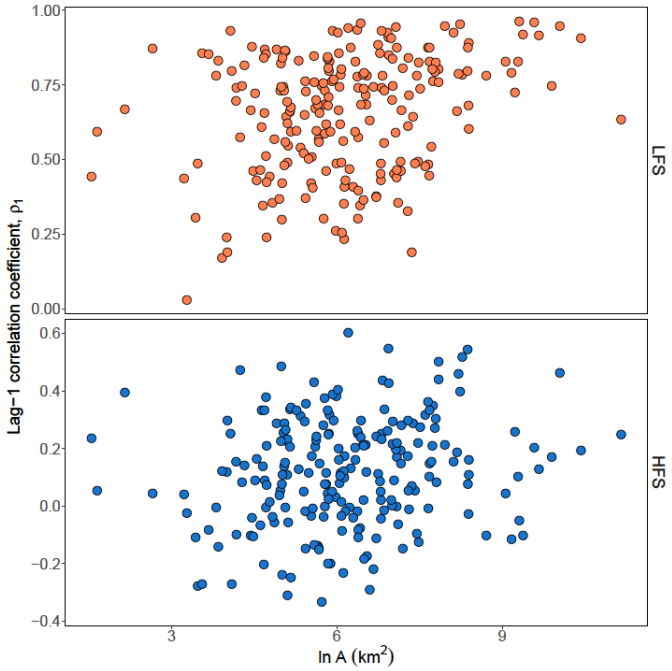


Figure 5.5 Scatterplots of lag-1 HFS (bottom panel) and LFS (top) streamflow correlation versus the natural logarithm of basin area **ln *A***.

Flow indices – descriptors BI and SR

The effect of the BI and SR is shown in Fig. 5.6. The BI (Fig. 5.6a) appears to be a marked positive driver for LFS (r_s = 0.6, p = 1.8 × 10^{-23}) while its effect for HFS is less clear, being weakly positive (r_s = 0.21, p = 0.001). For SR (Fig. 5.6b), it appears that both LFS and HFS streamflow correlations drop for increasing wetness (r_s = –0.4, p = 4 × 10^{-10} and r_s = –0.28, p = 2.8 × 10^{-5} respectively).

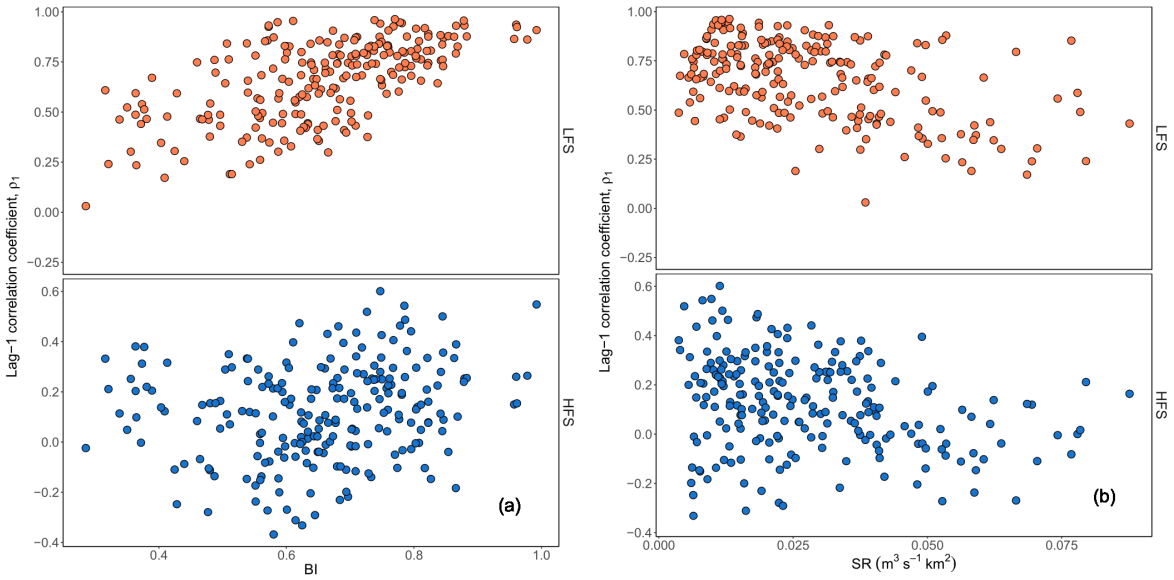


Figure 5.6 Scatterplots of lag-1 HFS (bottom panels) and LFS streamflow correlation (top panels) versus baseflow index BI (a) and specific runoff SR (b).

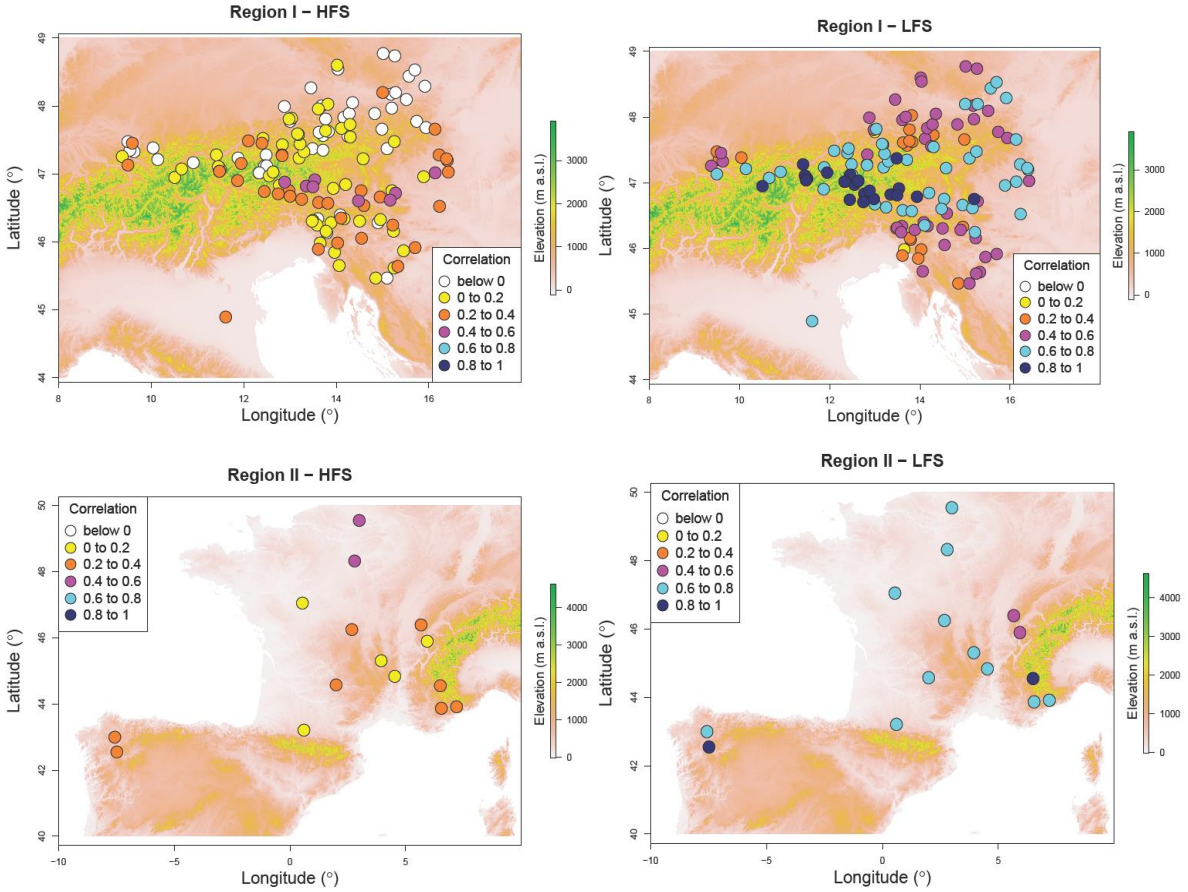
Presence of lakes and glaciers – descriptors PL and PG

Detailed information on the presence of lakes is available for the 69 Swedish catchments while areal extension of glaciers is known for the 108 Austrian catchments. Figure B.1 in Appendix B.1 shows that the impact of lake area (Fig. B.1a) on correlation for LFS and HFS is not significant but positive (r_s = 0.10, p = 0.399 and r_s = 0.12, p = 0.347). The results for glaciers show a positive impact for LFS (r_s = 0.28, p = 0.081) but a negative impact for HFS (r_s = –0.34, p = 0.032). For a meaningful interpretation, these results should be considered in conjunction with the seasonality of flows for the Austrian catchments. Low flows for the glacier-dominated catchments typically occur in winter months, when glaciers are not contributing to the flow (Parajka et al., 2009). Thus the observed result for LFS more likely portrays the impact of low temperature (low evapotranspiration) and snow accumulation, the latter generally being a slowly varying process. For HFS, which typically occurs in the summer months for the considered catchments, flows are mainly determined by snowmelt, which is associated to reduced persistence (Appendix B.1; Fig. B.1b).

Catchment elevation

The areal coverage of the SRTM data is limited to 60° N and 54° S and therefore, data for the northern part of the Swedish catchments are not available. The rest of the rivers are divided in three regions based on proximity: Region I including the central and eastern part of the Alps and encompassing Austrian, Slovenian and Italian catchments; Region II including the western part of the Alps and encompassing French and Spanish territory; and Region III including the southern part of Sweden. Figure 5.7 shows elevation maps along with the location of gauge stations and magnitude of correlations. Elevation seems to enhance LFS correlation, which is more evident in the mountainous Region I (Fig. 5.7). For HFS correlation there is not a prevailing pattern.

In the case of Austrian catchments, a 1 km resolution digital model is also used to extract information on elevation. Figure 5.8 confirms that there is a positive correlation pattern emerging with elevation for LFS. Based on local climatological information, it can be concluded that the spatial pattern for LFS correlation is reflective of the timing and strength of seasonality of the low flows in Austria, where dry months occur in lowlands during the summer due to increased evapotranspiration and in the mountains during winter (mostly February) due to snow accumulation which is characterized by stronger seasonality compared to the lowlands flow regime (Parajka et al., 2016; see Fig. 1). Concerning HFS in the same region, high flows are significantly impacted by the seasonality of extreme precipitation (Parajka et al., 2010b), which is highly variable, with the exception of the rivers where high flows are generated by snowmelt. Therefore, a spatially consistent pattern does not clearly emerge.



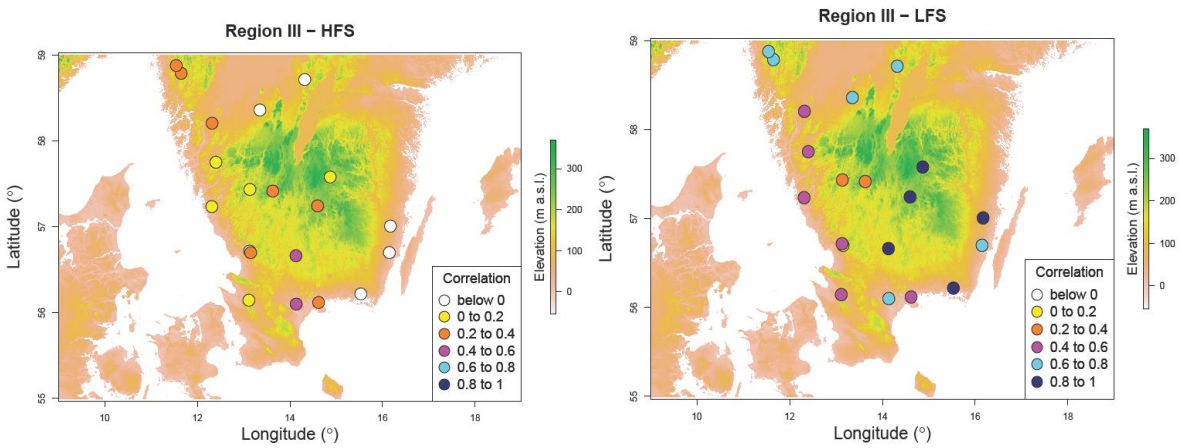


Figure 5.7 Relief maps from SRTM elevation data for the HFS and LFS lag-1 correlations of the rivers. Note that elevation scale is different for each region. Legend shows the colour assigned to each class of correlation for the data.

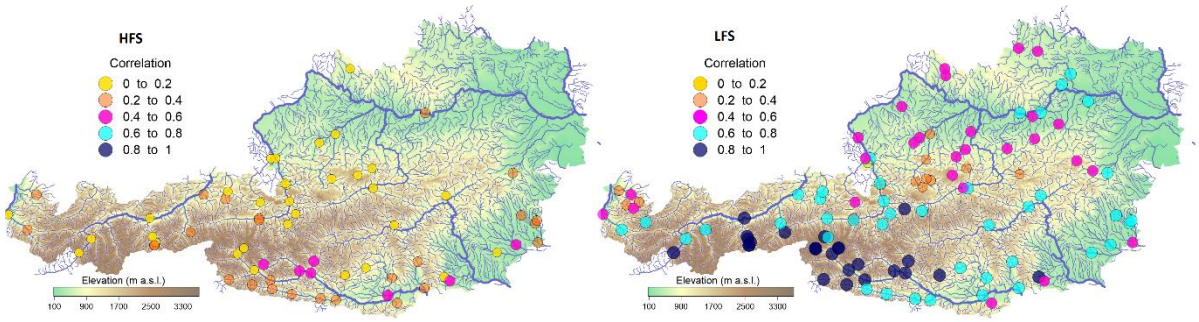


Figure 5.8 Digital elevation model of the Austrian river network depicting the spatial distribution of lag-1 positive correlation for HFS (left) and lag-1 positive correlation for LFS (right). Legend shows the colour assigned to each class of correlation for the data.

Catchment geology – descriptors PK and PF

Two different geological behaviours are identified which may impact river correlation. We first focus on 21 Slovenian catchments (out of 31) where more than 50% of the basin area is characterised by the presence of karstic aquifers (percentage of karstic areas PK ≥ 50%). Figure 5.9 shows boxplots of the estimated lag-1 correlation coefficient for both HFS and LFS against rivers where PK < 50%. It is clear that there is a significant decrease in correlation where karstic areas dominate for both for HFS and LFS.

In a second analysis, we focus on Austrian catchments and investigate the relationship between correlation and percentage of flysch coverage, PF. Figure B.2 in the Appendix B.1 shows that there is not a prevailing pattern in either case (r_s = 0.13, p = 0.6 for LFS and r_s = –0.19, p = 0.446 for HFS).

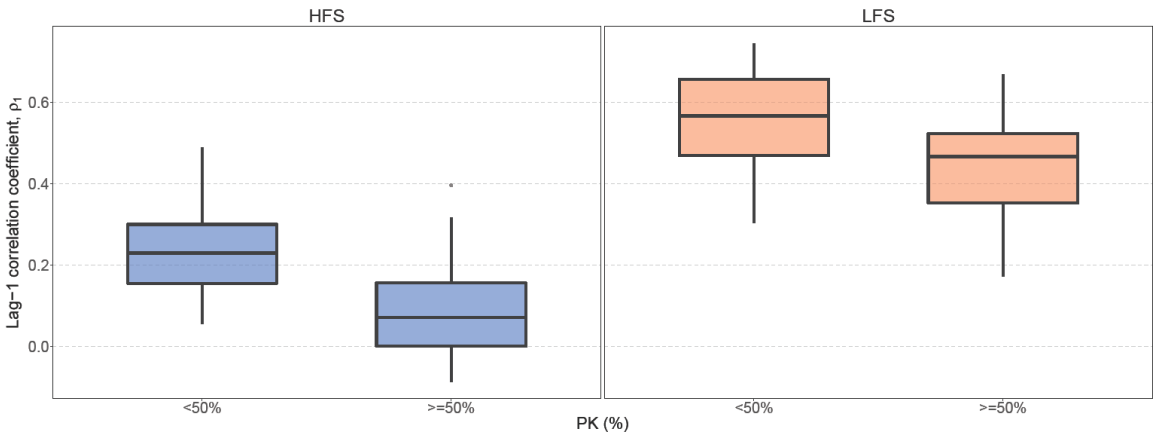


Figure 5.9 Boxplots of lag-1 correlation for Slovenian rivers with more than 50% presence of karstic formations PK and rivers with no or less presence for HFS analysis (left) and LFS analysis (right). The lower and upper ends of the box represent the 1st and 3rd quartiles, respectively, and the whiskers extend to the most extreme value within 1.5 IQR (interquartile range) from the box ends.

Atmospheric forcing – descriptors *P* and *T*

Figure 5.10 shows the lag-1 HFS and LFS correlations against estimates of the annual precipitation P and annual mean temperature T as well as the IDM. LFS correlation appears to be more sensitive than HFS to the above climatic indices, showing a decrease with increasing temperature and also a decrease with increasing precipitation ($r_s = -0.44$, $p = 3.1 \times 10^{-12}$ for P and $r_s = -0.57$, $p = 1.8 \times 10^{-20}$ for T). HFS correlation is scarcely sensitive to these variables ($r_s = -0.17$, $p = 0.011$ for P and $r_s = 0.08$, $p = 0.208$ for T). The IDM (Fig. 5.10 c) shows a mild decrease of both LFS ($r_s = -0.06$, $p = 0.368$) and HFS correlation with increasing IDM ($r_s = -0.17$, $p = 0.01$), while for the latter there seems to be a clearer trend (lower correlation with higher IDM) in very humid areas (dark blue points in Fig. 5.10c).

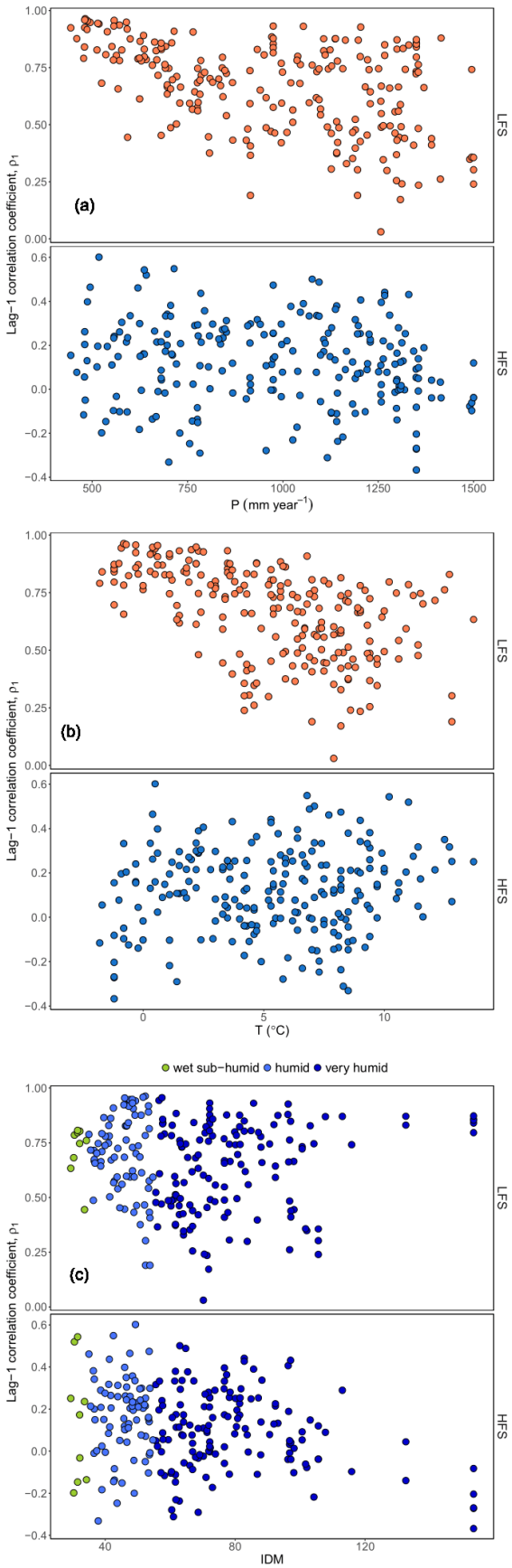


Figure 5.10 Scatterplots of lag-1 HFS and LFS correlation versus annual precipitation P (a), mean annual temperature T (b), and Index De Martonne IDM (c).

Physical drivers of high correlation

To gain further insight into the results we select the 20 catchments with the highest streamflow seasonal correlation coefficients for both HFS and LFS periods in order to investigate their physical characteristics in relation to the remaining set of rivers. Table 5.2 summarizes statistics for selected descriptors in order to identify dominant behaviours. We also compare the number of rivers with distinctive features, i.e. lakes N_L (number of rivers with lakes), glaciers N_G (number of river with glaciers), flysch N_F (number of rivers with flysch formations) and karst N_K (number of rivers with karstic areas) for the highest correlation group with those obtained from 1000 randomly sampled 20-catchment groups from the whole set of considered catchments to assess whether higher correlation implies distinctive features.

By focusing on HFS, one can notice that the catchments with higher seasonal correlation are characterized by larger catchment area, higher baseflow index and temperature with respect to the remaining catchments, and lower specific runoff, precipitation and wetness. The presence of lake, glacier, karstic and flysch areas do not appear significantly effective at a 5% significance level. More robust considerations can be drawn for the LFS: higher seasonal correlation is found for larger catchments with a higher baseflow index and lower specific runoff, precipitation and wetness. Decreasing temperature is strongly associated with higher correlation for the LFS. The presence of lakes plays a significant role both for lag-1 and lag-2 correlations with the latter also being significantly influenced by the presence of glaciers.

Table 5.2 Differences in the mean values between the descriptors of the 20-highest-correlation-river group for HFS and LFS versus the remaining rivers (204). $\boldsymbol{N_L}$, $\boldsymbol{N_G}$, $\boldsymbol{N_F}$ and $\boldsymbol{N_K}$ columns contain the absolute number of rivers in the higher correlation group with the specific descriptor (presence of lake, glacier, flysch and karst) with $*$ denoting significance at 5% significance level (two-sided test) and brackets containing the mean value from the 1000 resampled 20-catchment subsets.

Descriptor (Units)	A (km^2)	BI (–)	SR (m^3 s^{-1} km^{-2})	N_L (–)	N_G (–)	N_F (–)	N_K (–)	P (mm year^{-1})	T (°C)	IDM (–)
HFS lag1	+38.7%	+9.6%	–36.5%	5 (6)	5 (3)	1 (2)	1 (2)	–6.7%	+11.7%	–11.3%
LFS lag1	+358%	+20.2%	–47.3%	17$*$ (6)	3 (3)	0 (2)	0 (2)	–37.9%	–80%	–17.3%
LFS lag2	+139.7%	+18.9%	–40.8%	12$*$ (6)	7$*$ (3)	0 (2)	0 (2)	–26.5%	–64.2%	–8.8%

5.6 Principal component analysis of the predictors and linear regression

We attempt to fit a linear regression model to relate correlation to physical drivers, in order to support correlation estimation for ungauged catchments. To avoid the impact of multicollinearity in the regression while additionally summarizing river information, we apply PCA (Section 5.3.2). Although correlation effects are efficiently dealt with via the PCA, we avoid including highly correlated variables in the analysis.

For example, the De Martonne index, precipitation and SR are mutually highly correlated (all Pearson's cross-correlations are higher than 0.6) and therefore we only consider the SR in the PCA because it shows a more robust linear relationship with correlation magnitude. We select A, BI, SR and T as the variables to be considered in the PCA. A log transformation is applied to the basin area to reduce the impact of outliers. Table 5.3 shows the coefficients estimated for each component (the loadings) and the explained variance. The first principal component is primarily a measure of BI; the second principal component mostly accounts for T and the third principal component accounts for A. There is an evident geographical pattern emerging by the visualization of countries in the biplot (Fig. 5.11). Slovenian rivers cluster towards the direction of increasing SR and T, whereas Swedish rivers cluster towards the opposite direction of increasing BI and decreasing T. Austrian rivers, which are the majority, are the most diverse. The first two components together explain the 70% of the total variability in the data.

Naturally, the statistical behaviour of the indices reflects the known local controls for certain rivers. For example, the observed lowest BI in Slovenia is consistent with the presence of karstic formations for the majority of the Slovenian rivers, as is the higher BI in Sweden and Austria, which is related to the presence of lakes and glaciers in both countries.

In the case of HFS, all the examined linear models (combinations of ln A, SR, BI, P, T, IDM predictors) failed in explaining the streamflow correlation magnitude. On the contrary, the linear regression model performs fairly well in explaining the correlation for LFS, with an adjusted R^2 value of 0.58 and an F-test returning a p value $< 2.2 \times 10^{-16}$. The coefficients for the first three PCs are found significantly different from zero at a 0.1% significance level and are included in the regression (Table 5.4). The highest coefficient is obtained for the first PC, which mostly accounts for BI importance. Diagnostic plots from linear regression for LFS are shown in Fig. 5.12. There is no clear violation of homoscedasticity in linear regression, apart from the presence of a limited number of outliers. There is a certain departure from normality in the lower tail of the residuals, which relates to the fact that the model performs better in the area of higher seasonal streamflow correlations and overestimates the lower correlations.

Table 5.3 Loadings of the three Principal Components for ln A, SR, BI and T. The explained variance of each PC is denoted in parenthesis.

Predictor variables	PC1 (42.5%)	PC2 (28.2%)	PC3 (17%)	PC4 (12.2%)
ln A	–0.486	–0.427	0.748	0.145
SR	0.48	0.483	0.652	–0.332
BI	–0.619	0.262	–0.11	–0.731
T	0.385	–0.718	–0.04	–0.577

Table 5.4 Summary of Linear Regression results for the LFS model.

Predictor variables	Estimate	Standard Error	*t* value	Pr(>|*t*|)	Adjusted R^2	F-statistic
intercept	0.659	0.009	77.065	$< 2 \times 10^{-16}$	0.583	104.2
PC1	–0.111	0.007	–16.820	$< 2 \times 10^{-16}$		*p*-value:
PC2	0.0318	0.008	3.936	1.1×10^{-4}		$< 2.2 \times 10^{-16}$
PC3	–0.039	0.010	–3.754	2.2×10^{-4}		

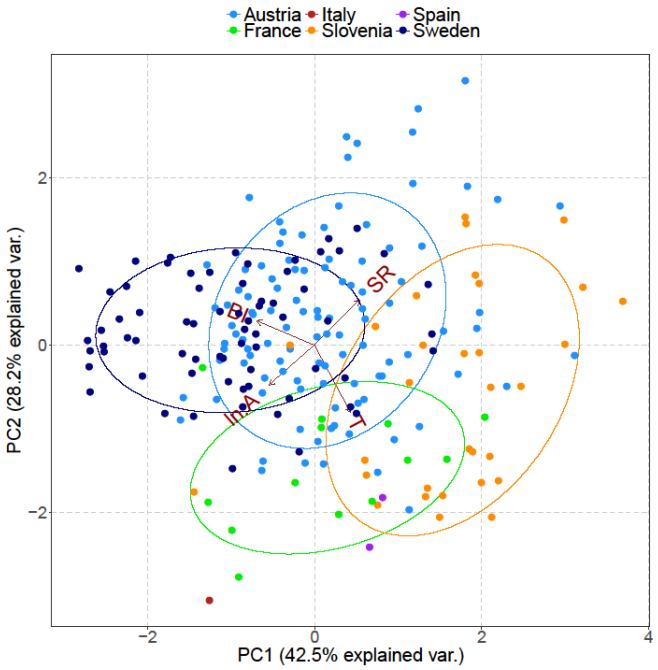


Figure 5.11 Principal component distance biplot showing the principal component scores on the first two principal axes along with the vectors (brown arrows) representing the coefficients of the baseflow index BI, specific runoff SR, natural logarithm of basin area ln A and mean annual temperature T variables when projected on the principal axes. Scores for the rivers are plotted in different colors corresponding to each country of origin and 68% normal probability contour plots are plotted for the countries.

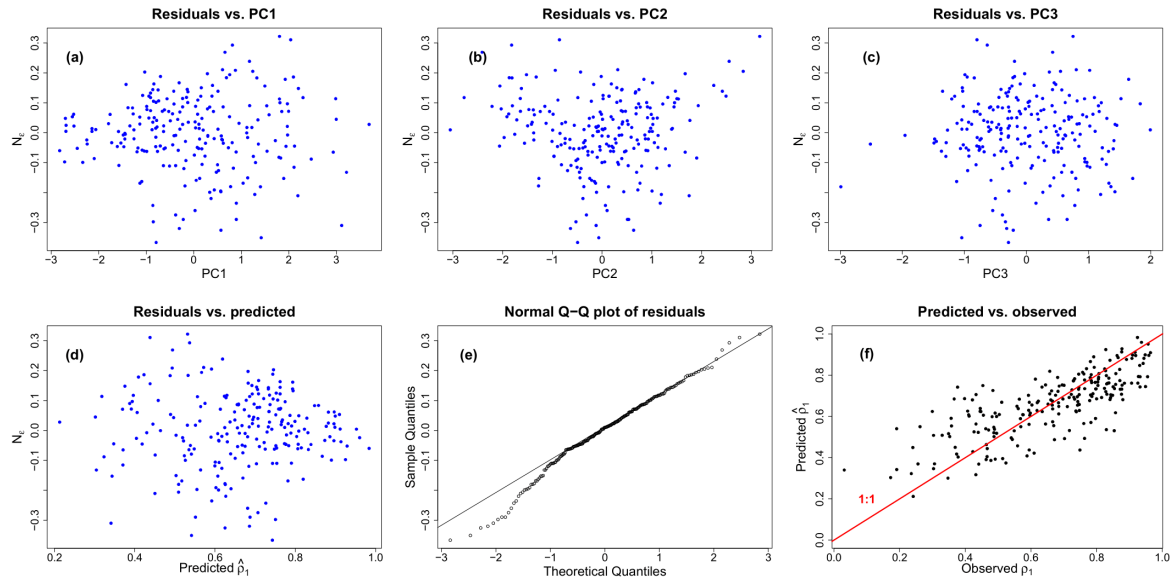


Figure 5.12 Diagnostic plots of linear regression for the LFS model. Residuals versus the first (a), the second (b) and the third principal component (c) and the predicted values (d). Normal Q-Q plot of the residuals (e). Plot of the predicted values from linear regression versus the observed ones; red line is the diagonal line 1:1 (f).

5.7 Real-time updating of the frequency distribution of high and low flows for the Oise River

We apply the technical experiment (Section 5.3.3) for high and low flows to the Oise River in France and assess the difference in the estimated flood and low-flow magnitudes. We update the probability distribution of high and low flows after the occurrence of the upper 95% and lower 5% sample quantile of the observed mean flow in the previous month, respectively.

The Oise River (55 years of daily flow values) at Sempigny in France has a basin area of 4320 km^2 and its gauging station at Sempigny is part of the French national real-time monitoring system (https://www.vigicrues.gouv.fr/), which is in place to monitor and forecast floods in the main French rivers. The selected river has a high technical relevance since it experiences both types of extremes with large impacts. For instance, a severe drought event in 2005 led to water restrictions impacting agriculture and water uses in the region (Willsher, 2005), while the river originated an inundation during the 1993 flood events in northern and central France, which was one of the most catastrophic flood-related disasters in Europe in the period 1950-2005 (Barredo, 2007). It is characterized by HFS correlation $\rho = 0.54$, which is the 3rd largest lag-1 correlation for the HFS in our dataset and LFS correlation $\rho = 0.80$, which stands for the 70% quantile of the sample lag-1 correlation for LFS.

A visual inspection of the residual plots is also performed (Fig. 5.13a, b) in order to evaluate the assumption of homoscedasticity of the residuals of the regression models given by equation (36). The residuals do not show any apparent trend and the Gaussian linear model is therefore accepted. Figure 5.13 (c, d) shows the conditioned

and unconditioned probability distributions of peak and low flows in the Gaussian domain. As follows from equation (38), the variance of the updated (conditioned) distributions decreases.

After application of the inverse NQT the conditioned peak flows are modelled through the EV1 distribution and compared to the unconditioned (observed) peak flows. The corresponding Gumbel probability plot for conditioned and unconditioned distributions is shown in Fig. 5.13e. For the return period of 200 years, the updated distribution shows a 6% increase in the flood magnitude for the Oise River (307.7 m^3 s^{-1} to 326.44 m^3 s^{-1}). Likewise, the conditioned low flows are modelled through the lognormal distribution. The two cumulative distribution functions are compared in Fig. 5.13f showing a major departure in the estimated quantiles for the updated distribution; the occurrence of the predefined 5% quantile flow in the pre-LFS month induces a decrease of the exceedance probability of an average LFS flow of 15 m^3 s^{-1} from a prior 43% (according to the unconditioned model) to 1%.

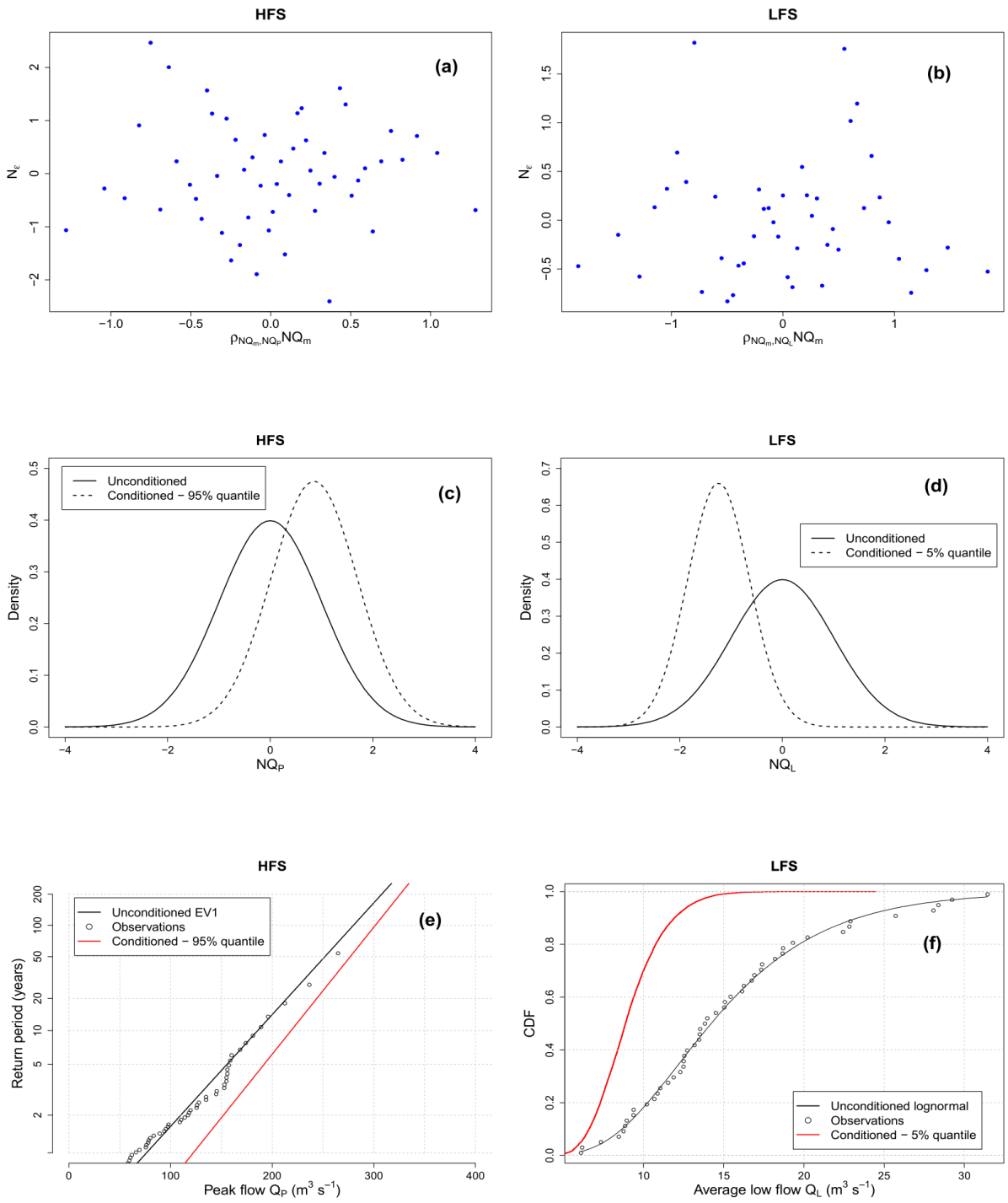


Figure 5.13 Conditioning the frequency distributions for high and low flows for the Oise River. Plots of the residuals of the linear regression given by Eq. (2) for the HFS (a) and LFS (b) models. Probability distribution of the unconditioned normalized peak flows NQ_P (solid line) and the normalized peak flows NQ_P conditioned to the occurrence of the 95% quantile (dotted line) for the HFS (c) and probability distribution of the unconditioned normalized low flows NQ_L (solid line) and the normalized low flows NQ_L conditioned to the occurrence of the 5% quantile (dotted line) for the LFS (d). Gumbel probability plots of the return period versus the unconditioned peak flows Q_P (black line) and the peak flows Q_P modelled by the EV1 distribution and conditioned to the occurrence of the 95% quantile (red line) for the HFS (e). Cumulative distribution function of the unconditioned low flows Q_L (black line) and the low flows Q_L modelled by the lognormal distribution and conditioned to the occurrence of the 5% quantile (red line) for the LFS (f).

5.8 Discussion

The methodology presented herein aims to progress our physical understanding of seasonal river flow persistence for the sake of exploiting the related information to improve probabilistic prediction of high and low flows. The correlation of average flow in the previous months with the LFS flow and HFS peak flow was found to be relevant, with the former prevailing over the latter. This result was foreseen since the LFS correlation refers to average flow while the HFS correlation is related to rapidly occurring events. We also aim to investigate physical drivers for correlation and quantify their relative impact on correlation magnitude. Therefore, a thorough investigation of the geophysical and climatological features of the considered catchments was carried out.

We found that the increasing basin area and baseflow index are associated with increasing seasonal streamflow correlation, yet the latter has a stronger impact. To this respect, Mudelsee (2007), Hirpa et al. (2010) and Szolgayova et al. (2014a) also found positive dependencies of long-term persistence on basin area, and Markonis et al. (2018) found a positive impact too but for larger spatial scales ($> 2 \times 10^4$ km^2), while Gudmunsson et al. (2011) found basin area to have negligible to no impact on the low-frequency components of runoff. Our results additionally point out that catchment storage induces mild positive correlation, not only for low discharges which are directly governed by base flow, but also for high flows, which is less anticipated.

Previous studies also pointed out that correlation increases for groundwater-dominated regimes (Yossef et al., 2013; Dijk et al., 2013; Svensson, 2016) and slower catchment response times (Bierkens and van Beek, 2009), which concurs with the impact of the baseflow index found herein as well as with the observed impact of fast responding karst areas. The latter findings are also in agreement with our conclusion that correlation decreases with increasing rapidity of river flow formation, which for instance occurs in the presence of karstic areas and wet soils, explaining why persistence decreases with high specific runoff, as also confirmed by other studies (Gudmundsson et al., 2011; Szolgayova et al., 2014).

Other contributions also reported higher streamflow persistence in drier conditions, either relating to lower specific runoff or mean areal precipitation estimates (Szolgayova et al., 2014; Markonis et al., 2018). It was postulated that this is due to wet catchments showing increased short-term variability compared to drier catchments (Szolgayova et al., 2014) and having a faster response to rainfall due to saturated soil. A similar conclusion has been reached by other previous studies reporting that low humidity catchments are more sensitive to inter-annual rainfall variability (Harman et al., 2011), therefore leading to enhanced persistence. Yet, these studies refer to generally humid regions and cannot be extrapolated to more arid climates. A related conclusion is proposed by Seneviratne et al. (2006) who found the highest soil moisture memory for intermediate soil wetness. These results do not contrast with our findings, which refer to a wide range of climatic conditions. In fact, our finding that increased wetness has a negative impact on seasonal memory of both high and low flows, extends the above results to the seasonal scale and interestingly, to both types of extremes.

We also confirm the role of lakes in determining higher catchment storage and therefore positive correlations for the LFS, which has only been reported for annual persistence in a few sites (Zhang et al., 2012).

The effect of snow cover for lag-1 LFS correlation is also revealed by the Austrian catchments. The mountainous rivers, directly affected by the process of snow accumulation, exhibit winter LFS and higher correlation than the rivers in the lowlands, which are more prone to drying out due to evapotranspiration in the hotter summer months. The inspection of elevation data confirmed the role of high altitudes in increasing LFS correlation, which is likely related to storage effects due to snow accumulation and gradual melting. In this respect, Kuentz et al. (2017) found that topography exerts dominant controls over the flow regime in the larger European region, controlling the flashiness of flow, and being a particularly important driver for other low flow signatures too. In fact, topography may affect the flow regime directly, through flow routing, but also indirectly, because of orographic effects in precipitation and hydroclimatic processes affected by elevation (e.g. snowmelt and evapotranspiration).

Regarding atmospheric forcing, we find LFS correlation to be negatively correlated to mean areal temperature and annual precipitation. The former result may be explained considering that increased evapotranspiration (higher temperature) is likely to dry out LFS flows while snow coverage (lower temperature) was found to be associated with higher LFS correlation. An apparently different conclusion was drawn by Szolgayova et al. (2014a) and Gudmundsson et al. (2011), who reported increasing persistence with increasing mean temperature postulating that snow-dominated flow regimes smooth out interannual fluctuations. Yet, it should be noted that they refer to interannual variability while we refer here to seasonal correlation and therefore to shorter time scales, which imply a different dynamic of snow accumulation and snowmelt; latitude may also play a relevant role in this, since in southern Europe the complete ablation of snow can occur more than once during the cold season, and sublimation may account for 20–30% of the annual snowfall (Herrero and Polo, 2016), decreasing the amount of snowmelt and impacting LFS flows in the summer season.

Snowmelt mechanisms are found to increase predictive skill during low-flow periods in some other studies (Bierkens and van Beek, 2009; Mahanama et al., 2011; Dijk et al., 2013). However, in the glacier-dominated regime of western Alpine and central Austrian catchments, it is unlikely that this is a relevant driver of higher correlation, since low flow occurs in the winter months. Yet the mountainous, glacier-dominated rivers still show increased LFS correlation compared to rivers in the lowlands, which agrees well with other studies that have found less uncertainty in the rainfall–runoff modelling in this regime owing to the greater seasonality of the runoff process and the decreased impact of rainfall compared to the rainfall-dominated regime of the lowlands (e.g Parajka et al., 2016).

Although the considerable uncertainty of areal precipitation estimates should be acknowledged, the contribution of annual precipitation interestingly complements the negative effect of increasing specific runoff –which is highly correlated to P estimates– on the correlation magnitude for both LFS and HFS. This outcome confirms that catchments receiving significant amount of rainfall do show less correlation than drier regimes as discussed before.

5.9 Conclusions and outlook

This research investigates the presence of persistence in river flow at the seasonal scale, the associated physical drivers and the prospect for employing the related information to improve probabilistic prediction of high and low flows by exploring a large sample of European rivers. The main findings are summarized below:

- Rivers in Europe show persistent features at the seasonal timescale, manifested as correlation between high- and low-flow signatures, i.e. peak flows in HFS and average flows in LFS, and average flows in the previous month. Correlation for LFS signatures is found to be consistently higher than HFS.
- Seasonal correlation shows increased spatial variability together with spatial clustering.
- Storage mechanisms, groundwater-dominated basins and slower catchment response time, as reflected by large basin areas, a high baseflow index and the presence of lakes, amplify seasonal correlation. On the contrary, correlation is lower in quickly responding karstic basins, and increased wetness conditions, as revealed by high specific runoff.
- Low mean areal temperature is associated with higher LFS correlation owing to the weaker drying-out evapotranspiration force and the mechanism of snow accumulation in higher altitudes. Higher mean areal precipitation is associated with lower LFS predictability, possibly due to the presence of saturated conditions and increased short-term variability in wetter climates.
- The drivers of LFS predictability are easier to identify and allow for the opportunity to construct regression models for possible application to ungauged basins (Section 5.6).
- HFS and LFS correlation may directly apply to the probabilistic prediction of 'extremes', i.e. high and low flows, as increased correlation can be exploited in various stochastic models. Such an application was performed in Section 5.7 in a data assimilation setting for a river of marked technical relevance.

Regarding the last point, once a significant correlation is identified, it may be exploited in other model variants as well, e.g. adding more dependent variables of lagged flow and/or coupling with other relevant explanatory variables, such as teleconnections or antecedent rainfall, in multivariate prediction schemes. Indeed, the presence of river memory at the seasonal scale represents a possible opportunity to improve the prediction of water-related natural hazards by reducing uncertainty of associated estimates and allowing significant lag time for decision-making and hazard prevention. Besides the high relevance for extremes, this type of seasonal predictability could also be of interest to water resources management by, for instance, exploring the memory properties of a minor HFS.

The inspection of the physical basis, apart from advancing our understanding of the catchment dynamics and enabling predictions in ungauged basins, is highly important as it may guide the search for other dependent variables and build confidence in the formation of process-based stochastic models (Montanari and Koutsoyiannis, 2012). A large sample of indices was herein inspected, yet more data are necessary in order to allow for more certain and generalized conclusions

worldwide. An important note is the effect of regulation, which, due to lack of objective data, is not completely understood. However, the opportunity of exploiting correlation is not affected by the presence of regulation, provided that the management of river flow does not change in time.

We conclude that our results point out that river memory provides interesting information that holds both theoretical and operational potential to improve the understanding and prediction of extremes, support decision-making and increase the level of preparedness for water-related natural hazards.

6. PERSISTENCE IN THE OCCURRENCE OF EXTREMES: MULTI-SCALE CLUSTERING OF RAINFALL EXTREMES

This Chapter deals with the propagation of persistence from the parent process to the occurrence process of its extremes. Clustering of extremes is critical for hydrological design and risk management and challenges the popular assumption of independence of extremes. Herein, we seek (a) the links between multi-scale clustering of extremes and long-term persistence, i.e. HK dynamics, and (b) the possibility to infer the former from the latter. To this aim, we highlight shortcomings of existing clustering indices and devise a new probabilistic index, which can reveal clustering, linking it to the persistence of the parent process. Its application to long-term rainfall records shows that the occurrence process of rainfall extremes may exhibit noteworthy departures from independence, which are consistent with the HK dependence structure of the parent process.

6.1 Introduction

The identification of clusters in series of extreme events is an ongoing research topic in geosciences, including hydrology, one that is particularly challenging due to the large estimation uncertainties involved when studying series of rare events. Regardless of the complications, this question has multiple important implications for earth sciences which range from understanding natural variability and process dynamics to correctly applying stochastic models for the purposes of inference and prediction. This is evident as most relevant hydrological and engineering applications require settling this issue at the early stage of the analysis, by either assuming independence (e.g. Coles et al., 2001; Kottegoda and Rosso, 2008) or 'ensuring' it through 'adequate' sampling techniques (Ferro and Segers, 2003). Thereby, the research focus can be uniquely placed on the more straightforward task of characterizing the probability distribution of extremes. For example, typical flood guidelines suggest that successive flood events have at least a certain separation lag time in order to be considered independent for the application of models (Lang et al., 1999).

In light of concerns for intensification of hydrological extremes due to anthropogenic forcing, the investigation of clustering receives additional interest (Ntegeka and Willems, 2008; Tye et al., 2018; Merz et al., 2016; Serinaldi and Kilsby, 2018a), as attribution of trends to an external deterministic forcing presupposes that at least the presence of natural inherent variability has been beforehand properly accounted for. In this respect, increasing evidence reporting the presence of persistence in various hydroclimatic variables (Hurst, 1951b; Koutsoyiannis, 2003; Montanari, 2003; Markonis and Koutsoyiannis, 2016; O'Connell et al., 2016; Iliopoulou et al., 2018b; Tegos et al., 2017; Dimitriadis, 2017) gives rise to the question of whether or not, and to what extent a regular behaviour of the extremes originating from persistent processes could be misinterpreted as a result of an anthropogenic cause.

This study deals with the investigation of clustering behaviour in records of maxima with a special focus on long-term daily rainfall observational records. As

recent studies reported evidence on the presence of persistence in annual rainfall (Iliopoulou et al., 2018b; Tyralis et al., 2018), the question of possible propagation of persistence to rainfall extremes naturally arises. Therefore, the research objectives can be articulated as follows: a) what are the links between persistence in the parent process and clustering of extreme events and can we infer the one from the other? and, b) what constitutes an informative characterization for clustering?

Typically, the assessment of clustering properties of extremes from a timeseries implies the selection of a threshold based on which the sampling of 'extreme' events is performed. Then, clustering is quantified based on the departure of the properties of extremes from the ones of a purely random process. This evaluation is performed either by considering the series of the inter-arrival times of extremes or equivalently, the series of counts of extreme events over counting windows. There is a direct correspondence between the two; it is well-known for example, that when the data come from a Poisson process, their inter-arrival times are exponentially distributed (Papoulis, A., 1991).

In the hydrological literature, various ad-hoc, sometimes visual and subjective approaches are used in order to quantify departures of extremes —typically floods— from independence and characterize clustering. The most systematic usually consist of some type of 'window' analysis, where the timeseries is split into subperiods which are examined for presence of perturbations in the statistics of extreme events, often corroborated by statistical testing (Marani and Zanetti, 2015b; Ntegeka and Willems, 2008; Willems, 2013). Avoiding the need for selection of time windows to study, Merz et al. (2016) applied a dispersion index, although mostly focused on a combination of kernel-based methods coupled with statistical significance tests to identify flood-rich and flood-poor periods in Germany. Yet, with a few exceptions only (Eichner et al., 2011; Serinaldi and Kilsby, 2016b, 2018a), the majority of clustering characterizations for hydrological extremes are not studied in relation to the dependence properties of the parent process, which is the focal point here.

To evaluate the clustering properties in a more comprehensive framework, two established indices are used in geophysical timeseries analysis, especially for the clustering analysis of earthquakes (Telesca et al., 2002) and storms (Vitolo et al., 2009) and are based on the 'counts' approach: the index of dispersion and the Allan factor. Both can be used to formally test the data against the Poissonian assumption (Serinaldi, 2013; Serinaldi and Kilsby, 2013) and it is reported that their scaling behaviour can also reveal the fractal properties of the underlying process for ideal rate fractal processes (Thurner et al., 1997). The latter is related to the asymptotic dependence property for large time horizons, long-term dependence, quantified by the Hurst parameter. For revealing the HK dynamics, a number of methods examining the original series also exist with the climacogram (Koutsoyiannis, 2010), i.e. the variance of the aggregated process over scales, shown to be the most robust (Dimitriadis and Koutsoyiannis, 2015).

We briefly review the above methods based on their performance on revealing the clustering of extremes sampled from synthetic timeseries generated in order to exhibit various degrees of persistence and different marginal distributions. We assess their degree of generality and showcase their shortcomings when extremes arrive from complex processes. We show how the interplay of persistence and moments of order

higher than 2 (skewness, kurtosis) can obscure the identification of the latter from extremes. Accordingly, we propose an alternative characterization of clustering based on a probabilistic index with distinctive features and test the proposed method on synthetic and real-world rainfall data. We find that the index exhibits some advantageous characteristics, namely it is capable of quantifying clustering by probabilistic means, linking it to the scaling behaviour of the parent process for a range of distributional and dependence properties. It also enables modelling the probabilities of threshold exceedances across multiple timescales, which can be used as a simulation tool, that being an important advance over existing methods that have mainly an inferential character.

6.2 Dataset

An extended dataset comprising the 60 longest available daily rainfall records is investigated in terms of its extreme properties. The data used in this study are collected from global datasets, i.e. Global Historical Climatology Network Daily database (Menne et al., 2012) and European Climate Assessment and Dataset (Klein Tank et al., 2002) and third parties acknowledged in the acknowledgments sections. They present an update of the previous dataset explored in Iliopoulou et al. (2018a) of long rainfall records surpassing 150 years of daily values. A detailed description of the dataset is provided in Appendix A.1. The geographic location of the rain gauges is shown in Figure 6.1. The length of the timeseries enables the investigation of clustering on extended time horizons from daily to yearly timescales.

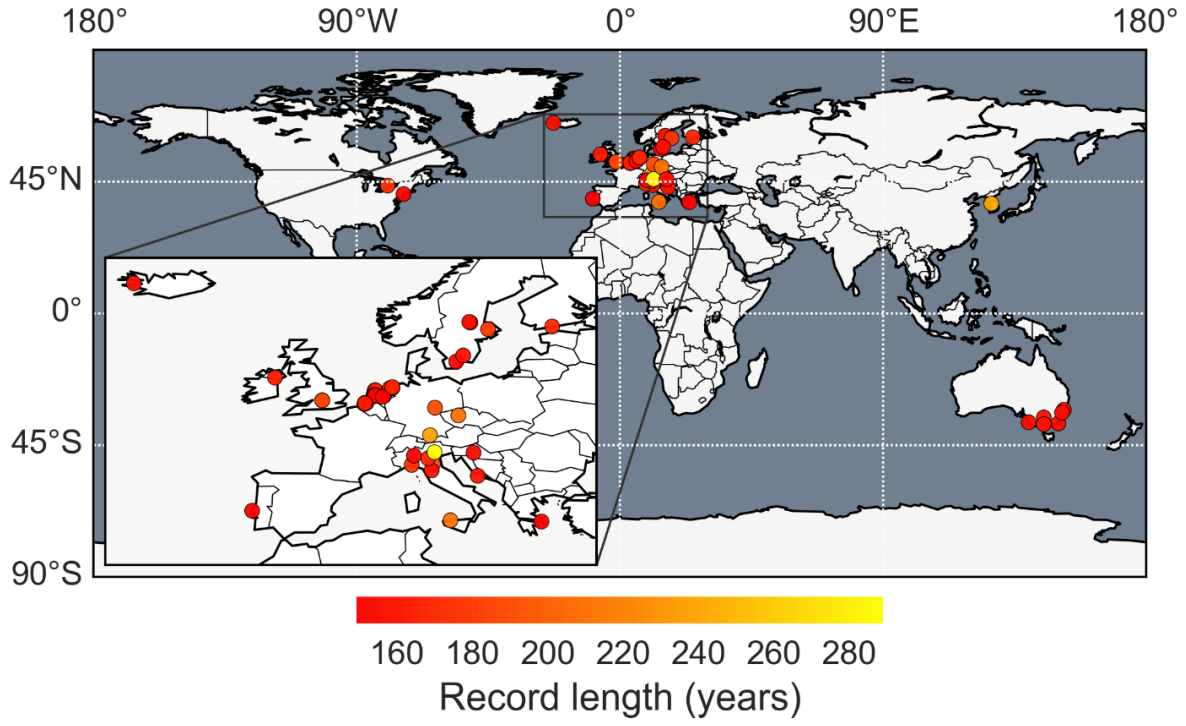


Figure 6.1 Map of the 60 stations with longest records used in the analysis.

6.3 Methodological framework

6.3.1 Definition of notation and mathematical formulation

Let $\underline{x_i}$ be a stationary stochastic process in discrete time i, i.e. a collection of random variables $\underline{x_i}$, and $x := \{x_1, \dots x_n\}$ a single realization (observation) of the latter, i.e. a timeseries. Now for u being a threshold, $u \in \mathbb{R}$, we define the process of peaks over the threshold (POT) consisting of events surpassing the threshold u, i.e,

$$\underline{y_i} := \begin{cases} \underline{x_i}, & \underline{x_i} > u \\ 0, & \underline{x_i} \leq u \end{cases} \tag{39}$$

Let also $\underline{N}(t)$ be a counting process of POT occurrences in time which is an increasing function of time t. We then define the process $z_q^{(k)} := \underline{N}(qk) - \underline{N}((q-1)k$ as the number of occurrences of POT at timescale k and at discrete time $q = 1, \dots, n/k$.

We also define by $\underline{m}_q^{(k)} := \max_{(q-1)k \leq j \leq qk}\{\underline{x_j}\}$ the block maxima series, which is formed by extracting the maximum order statistic of the observations divided in non-overlapping equally sized periods of length (timescale) k. In the following, we call the timescale k as timescale of filtering of the maxima. Figure 6.2 visualizes all the above at two temporal scales for a realization of a random process with Hurst parameter H=0.8 and the first four moments following a generalized Pareto distribution.

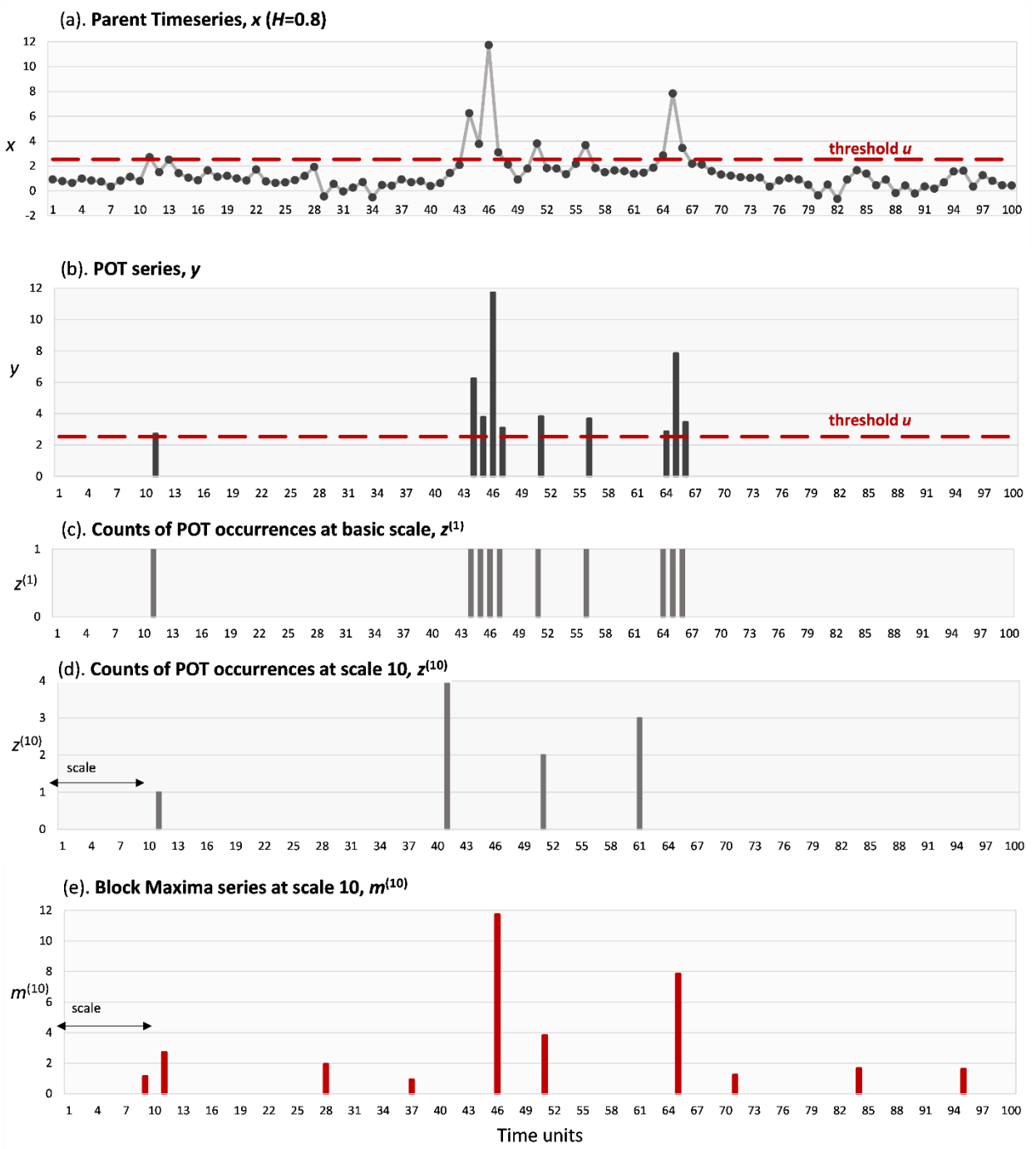


Figure 6.2 Explanatory graph of mathematical formulation. (a) Parent timeseries, (b) POT series, (c) temporal distribution of counts of POT at basic scale k=1, (d) temporal distribution of counts of POT occurrences at scale k=10 and (e) block maxima series at scale k=10.

6.3.2 Generation of benchmark synthetic timeseries

To evaluate the ability of clustering indices to discern the dependence characteristics of the parent (extreme generating) process, we first produce a set of synthetic timeseries with different dependence properties and marginal distributions. For the generation scheme, we employ a simulation procedure proposed by Dimitriadis and Koutsoyiannis (2018) which is capable of generating timeseries explicitly reproducing chosen theoretical moments up to any order together with any (long-term) persistence structure, i.e. the HK dynamics. We focus here only on processes exhibiting persistence

as these are the ones assumed consistent with the natural phenomena studied and also known to produce long-term clustering. For the marginal distribution, we generate timeseries preserving up to the 4^{th} order moments following the normal, generalized Pareto and gamma distributions. The higher-order moments of the generated timeseries follow the entropic distribution. Because the generation scheme preserves up to a specific number of moments from a distribution, the final shape may be slightly distorted with respect to the theoretical one, and therefore, we denote the generated series as type-gamma and type-Pareto, instead of gamma and generalized Pareto, respectively. For a detailed explanation of the generation scheme, the reader is referred to the Dimitriadis and Koutsoyiannis (2018). We focus only on the first four moments as higher-order classical moments cannot be reliably estimated from ordinary sample sizes (Lombardo et al., 2014).

The properties of these timeseries are chosen in order to cover a range of statistical and stochastic characteristics in terms of skewness, kurtosis and H parameter, and therefore, provide a good benchmark sample for testing the indices in typical but also more 'extreme' cases. Their properties are summarized in Table 6.1.We note that these timeseries are meant as theoretical case studies to test the appropriateness of the indices and are not to be considered as synthetic series of daily rainfall, which are the real-world data in question. However, since only the sequence of counts of extremes is of interest, and not their actual values, it is not necessary to strictly preserve other properties of daily rainfall, i.e. intermittency, and therefore in this sense comparison to the synthetic series is allowed. A sample of the timeseries is plotted in Figure 6.3.

Additionally to the above benchmark timeseries, we generate ensembles of shorter timeseries having lengths of 150×365 values, i.e. equal to the minimum record length of the rainfall data, and preserving the same moments as the benchmark timeseries. These series are produced using fewer weights for the SMA scheme, up to 2000, but applying proper weight adjustment scheme (Koutsoyiannis, 2016). They reproduce two dependence structures, white noise, and HK with H parameter 0.7, considered a representative value for hydrological processes. The purpose of the second benchmark sample is to test the methods in 'realistic' record lengths and to evaluate estimation uncertainty by Monte Carlo simulations that require significantly less computational effort compared to the first benchmark sample, which is generated using 10^6 weights, i.e. equal to the series length.

Table 6.1 Distributional properties of the benchmark samples used in the experiments with length 10^6 and $\boldsymbol{H}$ values in the range 0.5–0.99.

Distribution type	Parameters			Mean	Variance	Skewness	Kurtosis
	Shape	Scale	Location				
Normal	-	2.6	1.25	1.25	2.6	0	3
Gamma	0.1	5.1	-	0.51	2.6	6.325	63
Gamma	0.01	16.125	-	0.16	2.6	20	603
Pareto	0.1	1	0	1.11	1.54	2.81	17.83
Pareto	0.2	1	0	1.25	2.6	4.65	73.8

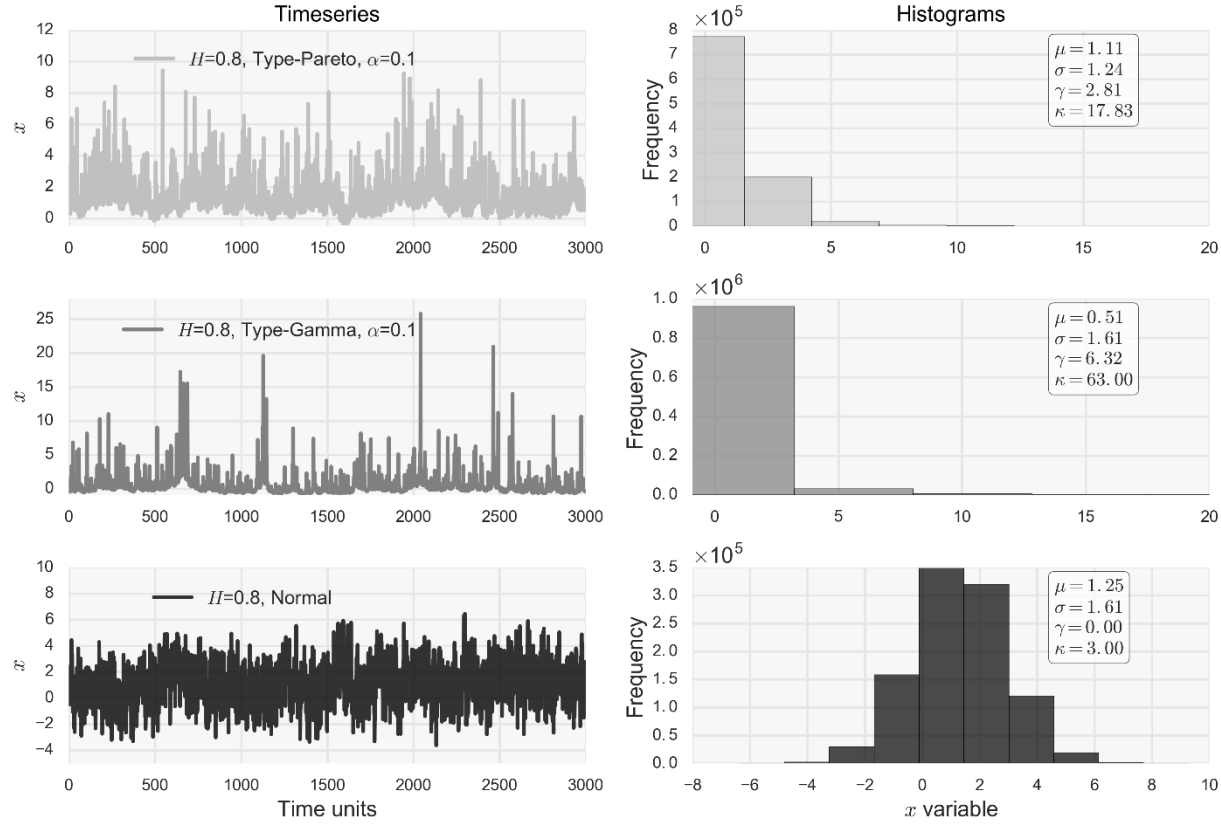


Figure 6.3 Visualization of three timeseries with $\boldsymbol{H}$=0.8 and different marginal distributions generated from the 4-moment SMA scheme (Dimitriadis and Koutsoyiannis, 2018). The legends report the mean, standard deviation, coefficient of skewness and coefficient of kurtosis of each distribution.

6.3.3 Second-order characterization of extremes

The Hurst parameter is a well-established measure of persistence. It can be estimated from the slope of the double logarithmic plot of the standard deviation of the averaged process versus the averaging timescale, i.e. the climacogram (Koutsoyiannis, 2010; see Section 2.2.1). To test how the estimator is impacted when extremes are used instead

of the original values, we compute the H parameter for extremes extracted from windows (scales) of length 1 to $N/10$ where N is the timeseries length. An example is provided in Figure 6.4. The first H value (scale = 1) is the value for the original data (the parent timeseries) and as the scale increases progressively the time series is filtered to show only the most 'extreme' data. For instance, if the basic timescale is daily, the estimated H parameter at timescale k=365 corresponds to the H parameter of the annual maxima. To reduce computational time, we perform estimation every 50 scales. The results are shown in Figure 6.4 are for the normal and the other benchmark timeseries. The impact of skewness and kurtosis on the estimator is striking as in the case of non-Gaussian timeseries, the H parameter quickly decays to 0.5, as if there was independence. On the contrary, for the normal timeseries it yields almost a stable value. To verify that this is not due to the impact of standard deviation bias induced by dependence, we performed estimation for selected timescales with the unbiased with respect to standard deviation, LSSV estimator (Koutsoyiannis, 2003; Tyralis and Koutsoyiannis, 2011b) as well. We also repeat the estimation for the shorter timeseries and plot the average values at each scale obtained from the Monte Carlo experiments. The same conclusion can be drawn. The climacogram estimator is severely biased downward for extremes originating from non-Gaussian processes and falsely indicates independence after a few scales of filtering. Therefore, we do not consider the climacogram estimator for the rest of the analysis on empirical data. Since it has been shown that the climacogram is closely related to other second-order characterizations, i.e. spectrum and autocovariance (Dimitriadis and Koutsoyiannis, 2015), we also expect similar results from the latter. Furthermore, Barunik and Kristoufek (2010) have shown that even for the underlying process (the parent), the sampling properties of the Hurst parameter estimation by some other approaches, i.e. the multifractal detrended fluctuation analysis and the detrending moving average, are also greatly impacted by heavy tails.

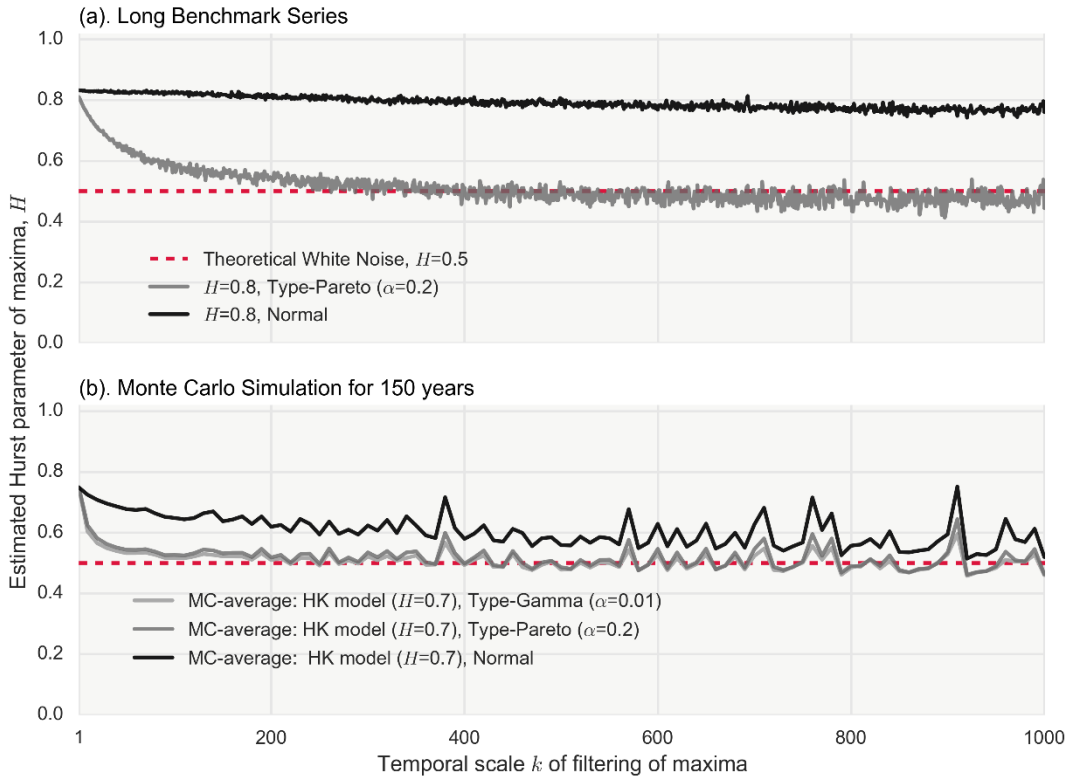


Figure 6.4 H parameters estimated from block maxima series at increasing scale of filtering for (a) benchmark series of length 10^6 from HK models with H =0.8 following normal and type-Pareto distributions and (b) average H values from 10^3 Monte Carlo simulations for HK models with H=0.7 and three different marginal distributions, type-gamma, type-Pareto and normal.

6.3.4 Clustering indices: the dispersion index

A well-known measure of clustering of events is the index of dispersion of counts, also known as the Fano factor (e.g. Thurner et al., 1997), which is defined as the ratio of the variance of the counts of events versus their mean number at a specific timescale k, i.e.:

$$\underline{d}^{(k)} = \frac{E\left[\underline{z}_k^2\right] - E[\underline{z}_k]^2}{E[\underline{z}_k]} \tag{40}$$

For a Poisson point process, the dispersion index is unity for all timescales. According to the literature (Thurner et al., 1997) the dispersion index exhibits power-law scaling behaviour which is linked to the underlying persistence structure. Although the exact form of the equation provided could not be theoretically validated *per se* at small scales, we have confirmed the power-law scaling at large scales, which by revising the original equation (Thurner et al., 1997), can be expressed as:

$$\underline{d}^{(k)} \approx ck^{2H-1}, k \geq k_0 \tag{41}$$

where c a real parameter and k_0 denotes the scaling onset timescale (a minimum time scale, for which the above scaling law applies). It follows that the exponent $2H - 1$ can

be obtained as the slope of the double logarithmic plot of the dispersion index versus the timescale for $k \geq k_0$ and therefore the Hurst parameter H, ranging in the (0,1) interval can be estimated accordingly. An example is provided in Figure 6.5.

We test the dispersion index against samples of Gaussian and non-Gaussian timeseries exhibiting HK dynamics. Namely, we use (a) the two long benchmark series (N=10^6), the normal and the type-Pareto, both exhibiting H = 0.8, and (b) the ensemble of simulations of shorter length (equal to daily values for 150 years) for three different distributions, normal, type-gamma with shape parameter α=0.01 and type-Pareto with α=0.2, all exhibiting H = 0.7. For the second sample, we provide the average value estimated from the 10^3 Monte Carlo simulations of the dispersion index at each scale. Results are shown in Figure 6.5.

At first, it is worth noting that the onset scale, from which scaling arises, appears to be smaller for the long compared to the shorter timeseries. The related H parameters are estimated from equation (41) for onset scale k_0=500, for both cases, in order to ensure a more robust estimate (yet fitted lines are extrapolated backwards to scale 365). It can be seen that the index yields satisfactory approximations of H only for the normal distribution and the long benchmark series (estimated H = 0.77, theoretical H = 0.8), whereas results are biased downward for the non-Gaussian one (estimated H = 0.67, theoretical H = 0.8). In the case of the shorter record length, the bias severely increases as the index yields H parameters falsely denoting independence (average H = 0.54). There is also a considerable degree of ambiguity regarding the selection of the onset time, a task that requires visual examination and subjective judgement. Due to the above reasons, and namely, to the observed underestimation of persistence for common record lengths, we do not consider the index for the rest of the analysis. A more sophisticated use of the dispersion index as well as bias correction methods may be possible but remain out of the scope of the paper. For more information on a related index, the Allan factor, and its properties for testing independence the reader is referred to Serinaldi and Kilsby (2013).

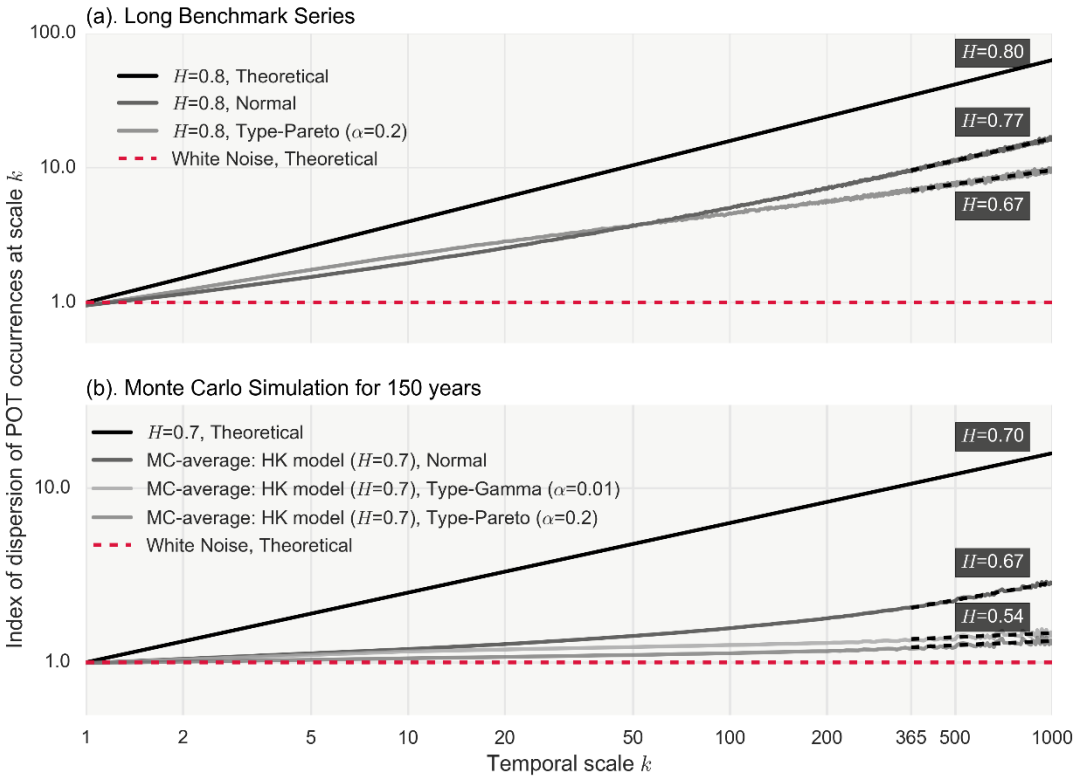


Figure 6.5 Index of dispersion of POT occurrences versus scale (double logarithmic axes) and estimated $\boldsymbol{H}$ parameters for scales>500 for (a) benchmark series of length 10^6 from HK models with theoretical $\boldsymbol{H}$=0.8 following normal and type-Pareto distributions and (b) average values from 10^3 Monte Carlo simulations for HK models with theoretical $\boldsymbol{H}$=0.7 and three different marginal distributions, type-gamma, type-Pareto and normal

6.3.5 A new probabilistic index to characterize multi-scale clustering behaviour

The above review highlights the complexity involved in identifying clustering of extremes and the need to devise an informative and objective characterization able to reveal persistence even for non-Gaussian series, which are usually the ones of interest in geophysical studies. To address this, we formulate a straightforward and assumption-free representation of clustering by estimating the probability of occurrence of extreme events across multiple scales. The proposed probabilistic index is defined as follows.

We set a threshold to the original timeseries and select the data surpassing the threshold as extreme events, hence, forming the Peaks Over Threshold series, y_i. Accordingly, we form the series of counts of the POT events for each scale, $z^{(k)}$, as explained in Section 6.1 (see also Fig.6.2). We additionally, define the binary process $\underline{r}_q^{(k)}$ to denote the event of exceedance of the threshold at each time interval q of size k, $q = 1, \ldots, \lfloor n/k \rfloor$:

$$\underline{r}_q^{(k)} := \begin{cases} 1, & \underline{z}_q^{(k)} > 0 \\ 0, & \underline{z}_q^{(k)} = 0 \end{cases} \tag{42}$$

Then, the probability of exceedance of the threshold for timescale k is obtained as the frequency of exceedances estimated from all $\lfloor n/k \rfloor$ intervals:

$$\bar{p}^{(k)} = \frac{\sum_1^{\lfloor n/k \rfloor} \underline{r}_q^{(k)}}{\lfloor n/k \rfloor} \tag{43}$$

The latter is the exceedance probability (of the threshold) versus the scale (EPvS) and its complement, $p^{(k)} = 1 - \bar{p}^{(k)}$ is the non-exceedance probability versus scale (NEPvS). Evidently, at scale $k = 1$ the EPvS is an estimate of the probability of the threshold value, $\bar{p}^{(1)} = F(u)$, and the NEPvs is $p^{(1)} = 1 - F(u)$. For example, in the previous applications, the threshold value was selected so that $F(u) = 0.05$. For a purely random process, the NEPvS is obtained as:

$$p^{(k)} = p^k \tag{44}$$

where p is the probability of non-exceedance at the basic scale $k = 1$ and equals $1 - F(u)$. Therefore, for white noise processes, the probability of occurrence of extremes across scales is fully determined by the choice of the threshold (controlling its probability at the basic scale) and the scale. For HK processes though, a different behaviour is revealed, with the probabilities of non-exceedance of the threshold being larger than those obtained under independence. This property of HK is discussed and investigated extensively in the following Section 6.5.

To model the NEPvS, we revisit a probabilistic model proposed by Koutsoyiannis (2006) to describe the clustering behaviour of dry spells in rainfall timeseries. The model derives from an entropy-maximization framework and was originally proposed to describe the probability dry across different timescales. The latter, according to our definition, corresponds to a threshold taking the value of 0. Therefore, in a similar manner to the probability dry, we obtain the probability of non-exceedance of the threshold at scale k as:

$$p^{(k)} = p^{\left(1+\left(\xi^{-1/\eta}-1\right)(k-1)\right)^\eta}, \qquad p = 1 - F(u) \tag{45}$$

where u is the threshold parameter and $\eta, \xi \in (0, 1)$. For $\eta = 1$ and $\xi = 0.5$, equation (45) describes the white noise process. To allow backward extendibility to scale $k = 0$, the positivity of the base should be ensured and therefore the following inequality should hold: $\xi \geq 1/2^\eta$. We apply both the index and the proposed model to the synthetic series as well as to the rainfall data and assess their performance in characterizing clustering. We evaluate the index's ability to reveal dependence by examining its performance for all the benchmark timeseries and we test its robustness by varying all the involved factors, i.e. sample size, marginal distribution's properties and threshold value.

6.4 Linking multi-scale clustering to persistence

We estimate the NEPvS index for the synthetic benchmark timeseries setting the threshold of extremes to 5%. The benchmark series have length 10^6 and therefore for a 5% threshold we obtain 50 000 extreme values (POT events). We investigate the temporal scales 1 to 1000, since the index's applicability to larger scales is to some

extent also conditioned by the available sample size (this feature is discussed in Section 6.5.1).

Results from the NEPvS application are demonstrated on a double logarithmic plot of minus natural logarithm of the non-exceedance probability of the threshold versus the scale, which for most cases yields a straight line (Fig. 6.6). Some interesting insights can be derived. As persistence increases, the probability of no occurrences of extremes in a scale progressively increases (equivalently, its minus logarithm —shown in the plots— decreases), which is true for all the examined distribution types. As already mentioned, there is a maximum temporal scale until which the index is informative. The latter, which we will call the 'max-discernible' scale, is the scale for which the estimated (from the simulated series) non-exceedance probability equals zero as at least one extreme event is encountered in every one of the $\lfloor n/k_{\max} \rfloor$ intervals. In this case, the minus logarithm of the NEPvS tends to infinity and is not shown on the plots. For a given number of extremes and thus, sample size, the max-discernible scale depends on the H parameter; the larger the persistence, the more timescales are required in order to 'encounter' the extremes. This can be explained by considering that another manifestation of clustering of extremes is the existence of prolonged periods of time with no extreme occurrences.

It is worth noticing that the marginal properties are irrelevant for the NEPvS of the white noise process. The latter is also proved in Fig. 6.6 as the lines of all the white noise timeseries with different marginals are completely identical, for which there is a theoretical justification. Likewise, for H parameters no far from 0.5 the different non-Gaussian distributions (Fig. 6.6a) yield negligible differences on the NEPvS plots. However, notable differences appear for $H > 0.7$. Specifically, the non-Gaussian NEPvS plots evidently differ from the NEPvS of the normal distribution, especially for large H values, with the latter showing more apparent clustering behaviour.

The NEPvS model, i.e. equation (45), fits perfectly all the range of non-Gaussian distributions, with a slight exception for the normal timeseries at small scales ($k < 50$) and very large H parameter ($H = 0.9$).

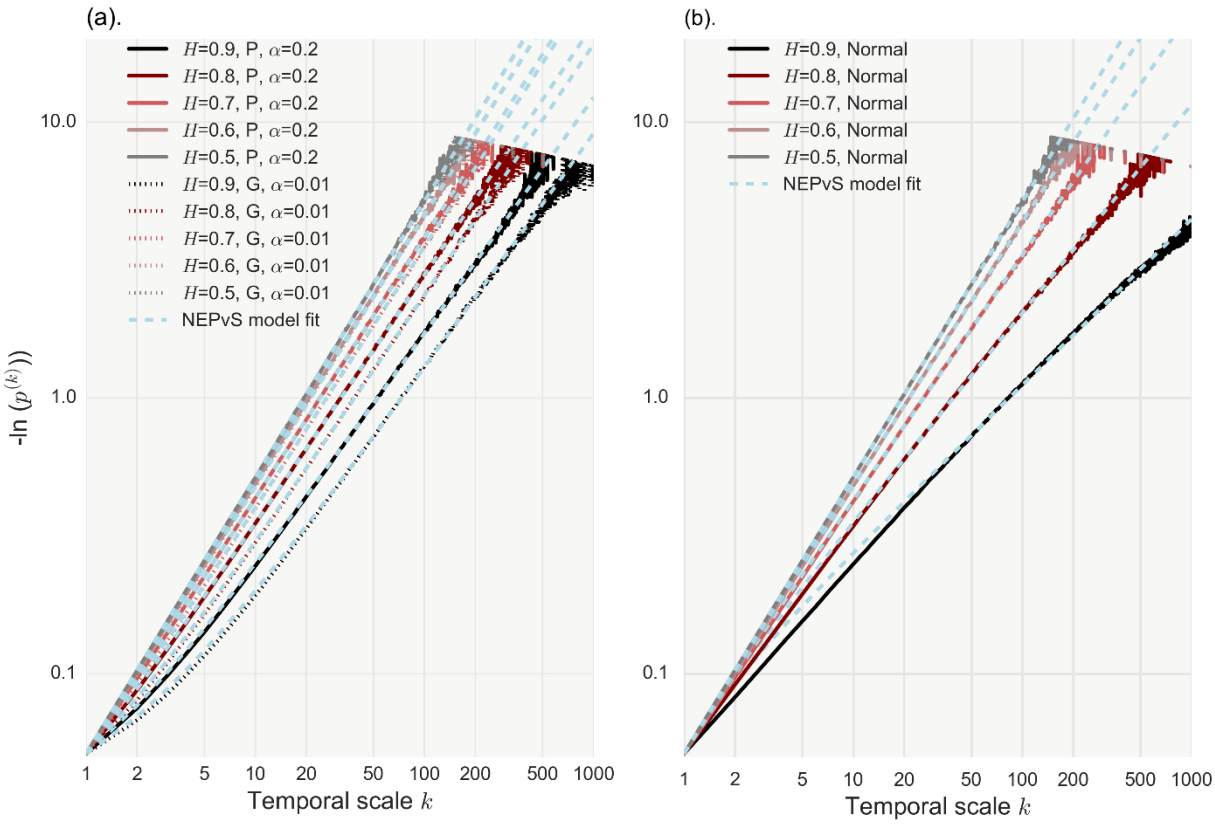


Figure 6.6 Minus natural logarithm of non-exceedance probability versus scale (NEPvS) index on double logarithmic axes along with the fit of the proposed model (Eq. 2) for (a) benchmark non-Gaussian timeseries (type-gamma and type-Pareto) and (b) benchmark normal timeseries, for a range of $\boldsymbol{H}$ parameters.

Having established that a representation in terms of the minus logarithm of probability vs. timescale, like that of Fig. 6.6, reflects the presence of persistence for a range of distribution types, we aim to frame its statistical behaviour for different configurations of extreme value analysis. For this purpose, the statistical behaviour of this graph is investigated by means of Monte Carlo simulation starting from the white noise case, which will serve as a benchmark model for identifying dependence from the rainfall data.

6.4.1 Sample size impact

We generate two ensembles of 10^3 white noise timeseries with sample sizes 150 years (150×365 daily values) and 300 years respectively, thus covering all the range of observed record lengths of our data set, and we produce the NEPvS plots for both lengths, shown in Figure 6.7. As expected, the larger sample size produces narrower Monte Carlo Prediction Limits (MCPL), yet the difference is almost negligible. The fact that sample sizes of this order of magnitude yield only minimal differences in the MCPL gives confidence in attributing the differences between the models that are examined next to other factors instead. The essential change however, between the two sample sizes is the propagation of the max-discernible scale to a larger timescale for the longer timeseries (Fig. 6.7). The latter is due to the fact that 'extremes' are

distributed in longer time periods for the longer series, and therefore, the longer the series the more timescales may be inspected for clustering.

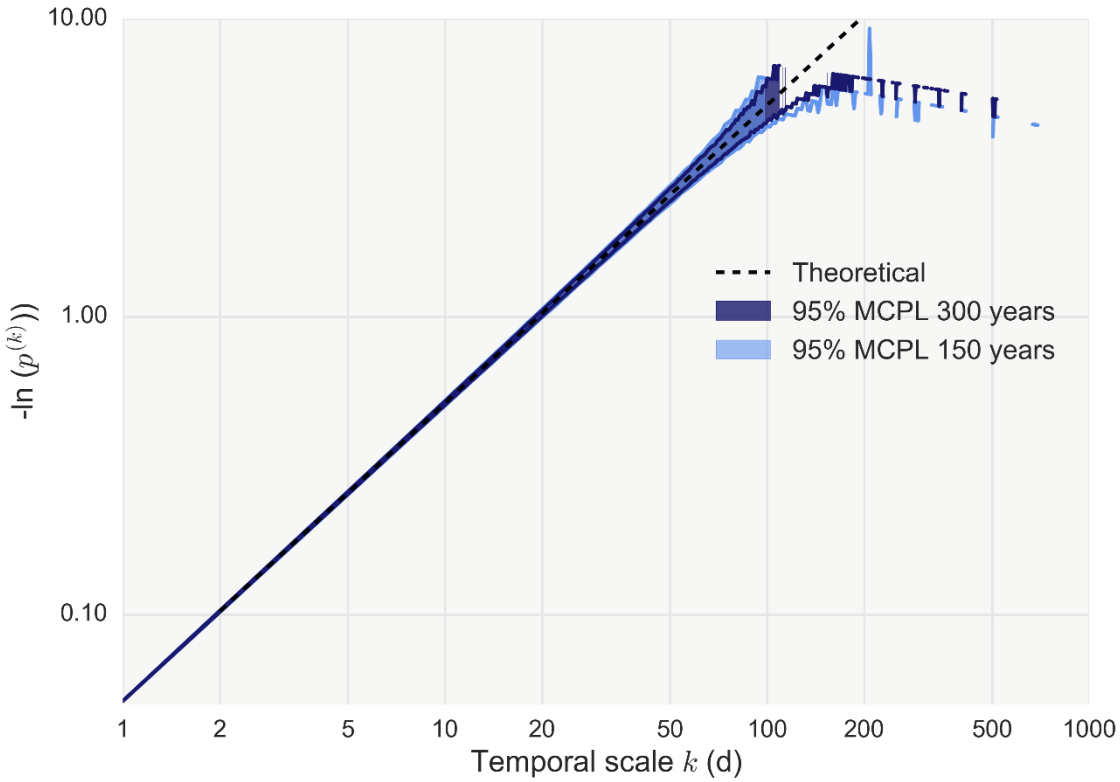


Figure 6.7 Minus natural logarithm of non-exceedance probability versus scale (NEPvS) index on double logarithmic axes for white noise timeseries and two sample lengths, 150×365 and 300×365.

6.4.2 Threshold impact

The selection of the threshold is the most important choice when analysing records of maxima. It is generally acknowledged that choosing 'high' thresholds for the extremes results to observations that are located far in the right tail of the distribution, and therefore they are of interest, but simultaneously, increases uncertainty as the sampled observations are fewer. The exact opposite is true for lower thresholds. Therefore, one has to seek an optimal threshold compromising this trade-off.

We first evaluate the choice of the threshold by examining four different thresholds associated with exceedance probabilities 0.5%, 1%, 5% and 10% respectively, applied for the benchmark case of independence, as seen in Figure 6.8. It is interesting to note that the main effect of the threshold for the IID case is the opportunity to apply the index to larger scales if the threshold is increased (smaller probability of exceedance). This is due to the fact that for the same record length, fewer extreme events are likely to be more separated in time and therefore, require longer timescales to be grouped.

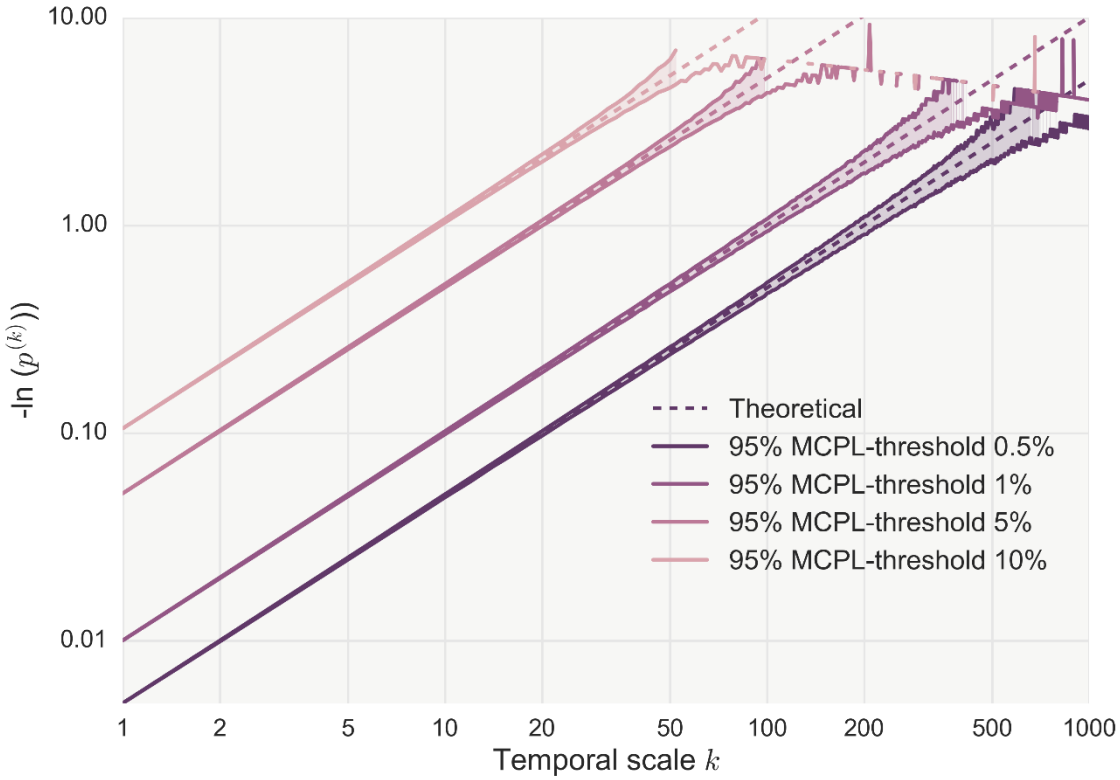


Figure 6.8 Minus natural logarithm of non-exceedance probability versus scale (NEPvS) index on double logarithmic axes for white noise timeseries (length 150×365) and variations of the sampling threshold of extremes

We also inspect the impact of the threshold in relation to the H parameter of the parent process for three distribution types from the benchmark series, type-Pareto with $a = 0.2$, type-gamma with $a = 0.01$ and the normal. In this case, we evaluate three different thresholds, 5%, 10% and 20%. Although the latter threshold would be considered 'low' for most extreme value analyses, here it is of interest, as by varying the threshold we aim to investigate the limits of identifiability of the HK behaviour, and not to focus on the exact shape of the distribution tail. To this aim, we fit the probabilistic model introduced in equation (45) to each timeseries and evaluate the ability to reveal persistence through the identifiability of the fitted parameters, η and ξ. In Fig. 6.9, the impact of the threshold is striking within the same distribution with lower threshold values (e.g. 20%) increasing identifiability of the parameters more than 10%. Additionally, it can be seen that the η parameter is more sensitive to the normal distribution, while on the contrary the ξ parameter is sensitive to increasing skewness and kurtosis.

By performing the above experiments, we have demonstrated the twofold effect of the threshold: 'lower' thresholds (higher probability of exceedance) enable better identifiability of persistence, yet they limit application of the index to less scales, and *vice versa*.

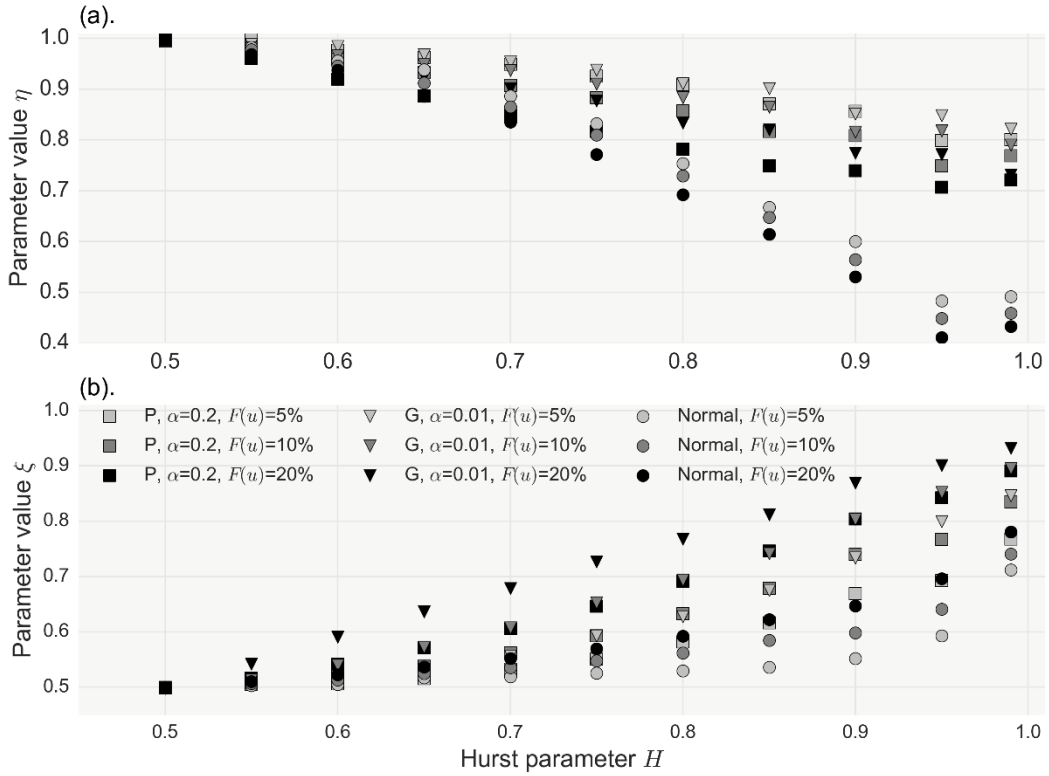


Figure 6.9 (a) Parameter $\boldsymbol{\eta}$ variation for increasing $\boldsymbol{H}$ parameter and different combinations of the sampling threshold and distribution type. (b) Parameter $\boldsymbol{\xi}$ variation for increasing $\boldsymbol{H}$ parameter and different combinations of the sampling threshold and distribution type.

6.4.3 Tail impact

At this stage, for the same threshold (5%), sample size (150 years) and H (0.7) parameter, we estimate the NEPvS index for the shorter benchmark series characterized by different marginal properties, and thus distribution tails, so as to focus solely on the impact of skewness and kurtosis on the index. Results are plotted in Figure 6.10. Two important conclusions can be drawn: (a) clustering of extremes and its identifiability is, in this case too, greater for the normal distribution (Fig.6.10a) and (b) for a specified non-Gaussian distribution, clustering is greater and also more visible for increasing skewness and kurtosis (Fig. 6.10b). The latter is a significant advance as the reviewed tools in section 6.3 showed very high downward bias for increasing higher order moments of the non-Gaussian distributions and practically no difference among them for the record lengths available (150 years). We also provide the plots of the fitting of η and ξ parameters computed for the long benchmark series with H parameters ranging in (0.5, 0.99) as well as their comparison in Fig.6.11 and Appendix B.2. All three plots confirm the above observations.

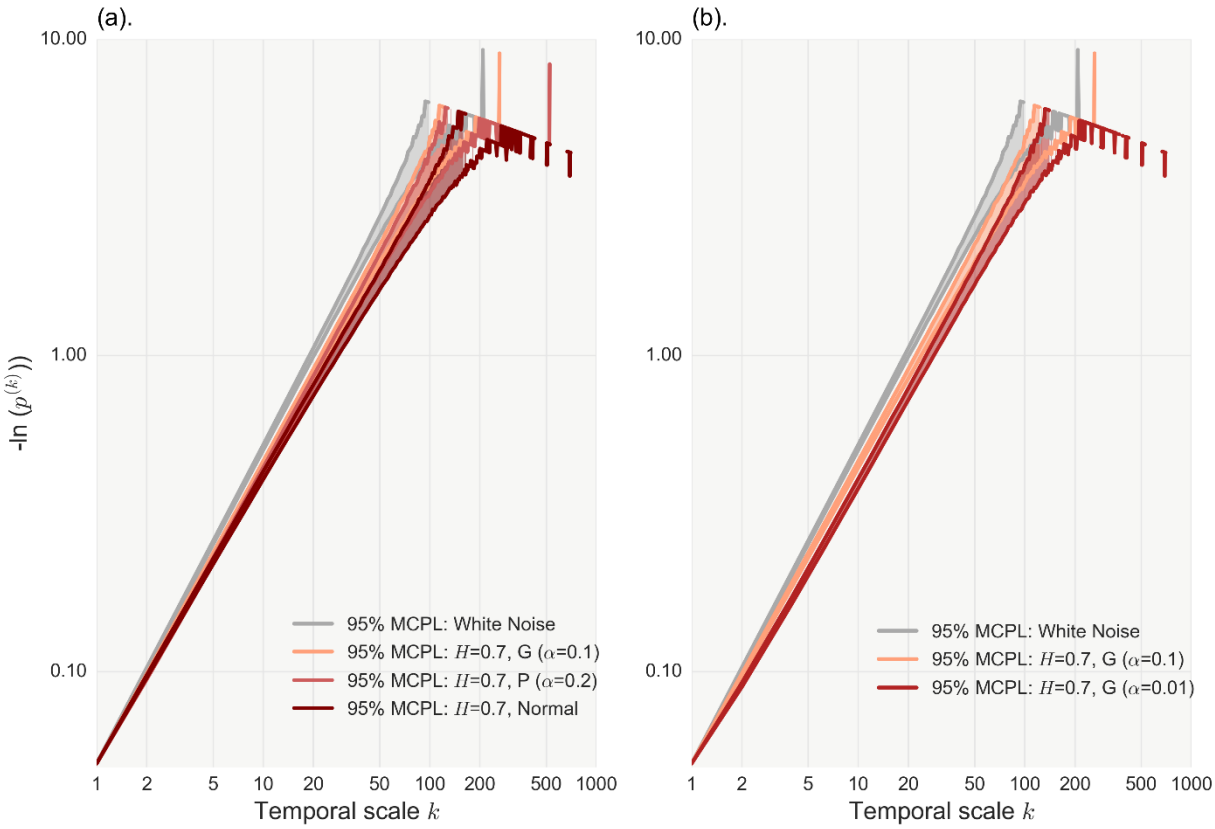


Figure 6.10 Minus natural logarithm of non-exceedance probability versus scale (NEPvS) index on double logarithmic axes along with 95% MCPL for (a) H=0.7 with type-gamma (α=0.1) and type-Pareto (α=0.2), and white noise and (b) H = 0.7 for two type-gamma distributions with α=0.1 and α=0.01.

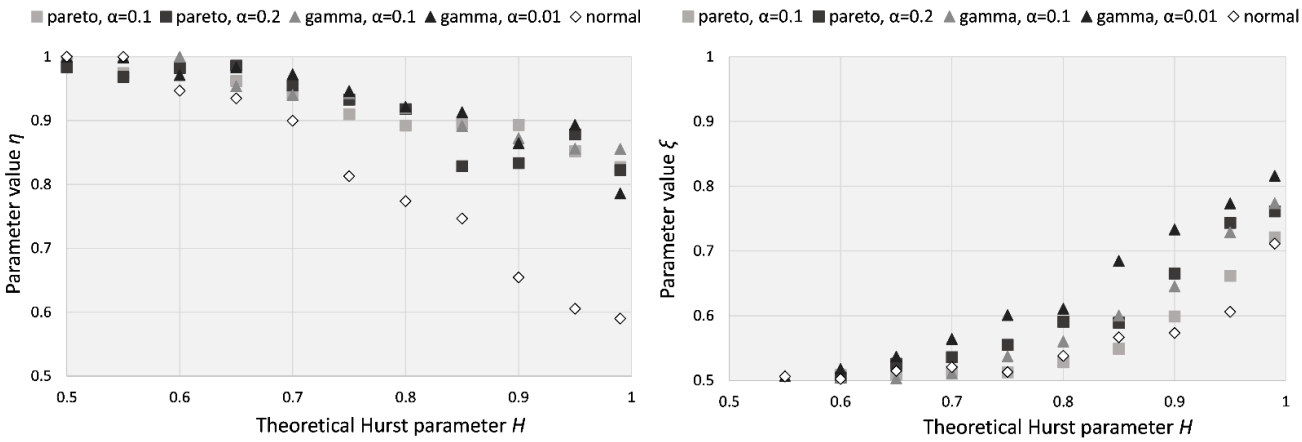


Figure 6.11 Plots of η and ξ parameters versus the H parameter for the type-Pareto with α=0.1 and α=0.2, type-gamma with α=0.1 and α=0.01 and the normal.

6.5 Clustering in real world rainfall extremes I: identifying clustering mechanisms in the parent process

Rainfall is a complex geophysical process for the stochastic modelling of which it is necessary to take into account its mixed-type marginal distribution (due to intermittency), the presence of cyclo-stationarity (seasonality and also diurnal cycle for sub-daily scales) as well as its scale dependence structure (Markonis and Koutsoyiannis, 2016). It is expected that all these mechanisms affect the clustering process of extremes.

In the following, we investigate their impact separately, although we note that the interplay among them may not necessarily allow the robust disentanglement of their effects at the different scales.

6.5.1 Influence of probability dry

The most distinctive feature of the rainfall process is its highly intermittent nature at fine temporal scales (Koutsoyiannis, 2006). To statistically account for intermittency, the marginal distribution is formed as a mixed (discrete-continuous) type one, having a probability mass function concentrated at 0 and a probability distribution function to describe the nonzero values. Therefore, if p_{d} is the probability of no-rain, termed probability dry, then the cumulative distribution function for the whole rainfall record $F_{\underline{x}}(x)$ can be defined in terms of the conditional distribution of wet days $F_{\underline{x}|\underline{x}>0}(x)$ as:

$$F_{\underline{x}}(x) = (1 - p_d)F_{\underline{x}|\underline{x}>0}(x) + p_d \,, \quad x \geq 0 \tag{46}$$

Since the threshold of extremes u is obtained as the quantile with a chosen probability of exceedance, it is evident that in the case of mixed-type processes, as in daily rainfall, the same threshold value will have a different probability of exceedance for the whole process and for the wet process (the nonzero rainfall). By simple probabilistic statements, it follows that the two exceedance probabilities of the threshold u for the compound and the wet process, $p_{\mathrm{c}}(u)$ and $p_{\mathrm{w}}(u)$, respectively, are related as:

$$p_w(u) = \frac{p_c(u)}{1 - p_d} \tag{47}$$

where $p_{\mathrm{d}} = 1 - p_{\mathrm{c}}(0)$ is the probability dry. Therefore, the exceedance probability for the same threshold is higher for the wet series, which means that depending on the probability dry, the values surpassing the same threshold may not necessarily belong to the right tail of the wet series as 'extremes'. For instance, a threshold u with associated exceedance probability 5% for the whole rainfall record with probability dry equal to 80% yields exceedance probability 25% for the wet series, and therefore the resulting series of POT events would also include lower rainfall values. While this is not a limitation of the methodology, it should be properly accounted for in order to (a) ensure that the resulting extremes are indeed towards the right end of the wet series tail and (b) to make meaningful comparisons among stations with different values of the probability dry. For this reason, we compute p_{d} for all stations in order to make sure that the resulting extremes are surpassing relevant thresholds. As previously shown, the latter is important since the threshold is the key control on the results.

6.5.2 Influence of seasonality

Seasonality may be in cases an important attribute of extreme rainfall impacting the central tendency of rainfall maxima belonging to different seasons and inducing temporal clustering in the series of extremes (Iliopoulou et al. 2018a). Since our aim is to focus on the impact of HK dynamics on clustering of extremes, we apply deseasonalization schemes to the original series in order to smooth out the seasonal components and reduce associated clustering. By doing so, we may perform Monte Carlo simulations with one marginal distribution per station for the validation of the chosen models. We note that a perfect separation of the impact of seasonality from HK dynamics may not always be possible, as in stations exhibiting strong seasonality we anticipate interplay between the two.

We consider two different methods for removing seasonality. The first one, termed M1, is a simple standardization scheme performed on a monthly basis. The daily values x_i belonging to each month m = 1,..,12 are transformed by subtracting the mean and dividing by the standard deviation of all daily values belonging to the same month, as follows: $y_i = (x_i - \mu_m)/\sigma_m$, $i \in m$. This method effectively removes seasonality from the first two moments of the data. In order to deal with higher order moments, we apply a second deseasonalization scheme denoted M2, which is based on the Normal Quantile Transformation (NQT) also applied on a monthly basis. The daily series for each month m are transformed to standard Gaussian quantiles through the inverse function of the standard Gaussian cumulative distribution, $y_i = {G_m}^{-1}(F(x))$ with their cumulative probability $F(x)$ estimated via their Weibull plotting position. Consequently, after the transformation, all daily values of each month follow the standard normal distribution. We found that the two schemes show minimal differences in the index's behaviour, with the most apparent ones belonging to the stations of Athens, Palermo and Lisbon.

In Figure 6.12, we plot three characteristic cases of the NEPvS behaviours found in the data: a) in a typical station with minimal to no seasonality (Oxford), extremes are not affected by deseasonalization schemes (Fig.6.12a), b) in a station with prominent seasonality (Athens, Fig.6.12b), a stronger deseasonalization scheme (M2) maybe required, and c) in an intermediate case (Helsinki, Fig.6.12c), the seasonal component in extremes is effectively dealt by with the simpler scheme (M1). The majority of the stations (40) belong to the third category, while for 17 stations accounting for seasonality yields minimal to no difference. These findings are consistent in general with the analysis of Iliopoulou et al. (2018a) on the presence of seasonality in extreme rainfall.

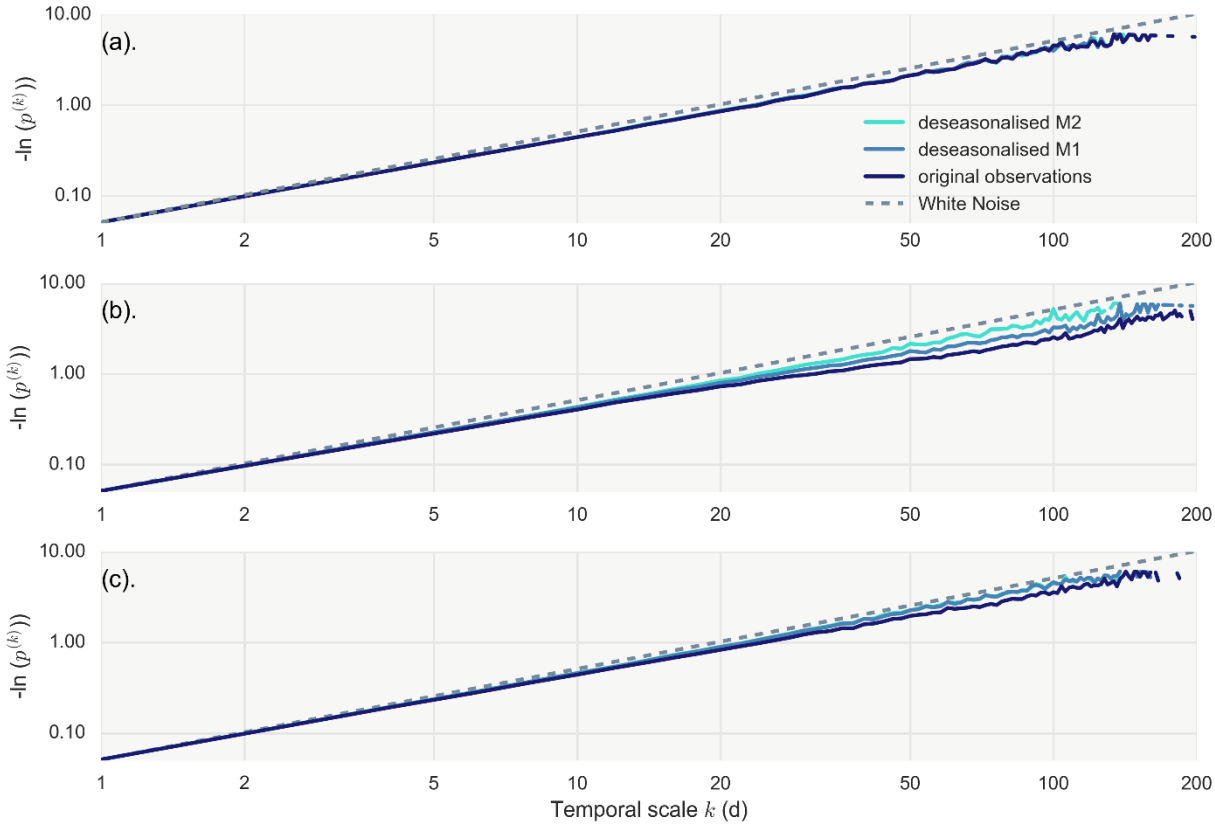


Figure 6.12 Minus natural logarithm of non-exceedance probability versus scale (NEPvS) index on double logarithmic axes for white noise timeseries and seasonal and deseasonalized series by methods 1 (M1) and 2 (M2) for the stations of Oxford (a), Athens (b) and Helsinki (c).

6.5.3 Rainfall scaling regimes

In order to highlight the motivation behind selecting the daily rainfall as a case study for the method and establish the 'target' persistence structure that we aim to reveal, we estimate the persistent properties of the previously deseasonalized daily rainfall series. To this aim, we compute the H parameter through the climacogram as introduced in Section 2.2.1. All the empirical climacograms are plotted in Figure 6.13. The estimated average persistence (Table 6.2) is close but even larger than the global estimate (H≈0.6) of Iliopoulou et al. (2018b) concerning annual rainfall. Remarkably, in many stations we observe a change of the scaling regime, namely an intensification of persistence, at scales above yearly. A similar result was observed in the work of Markonis and Koutsoyiannis (2016) for rainfall records at the over-decadal scale. This behaviour is also evident in the Table 6.2 reporting the estimated H parameters for the daily and above-yearly scales.

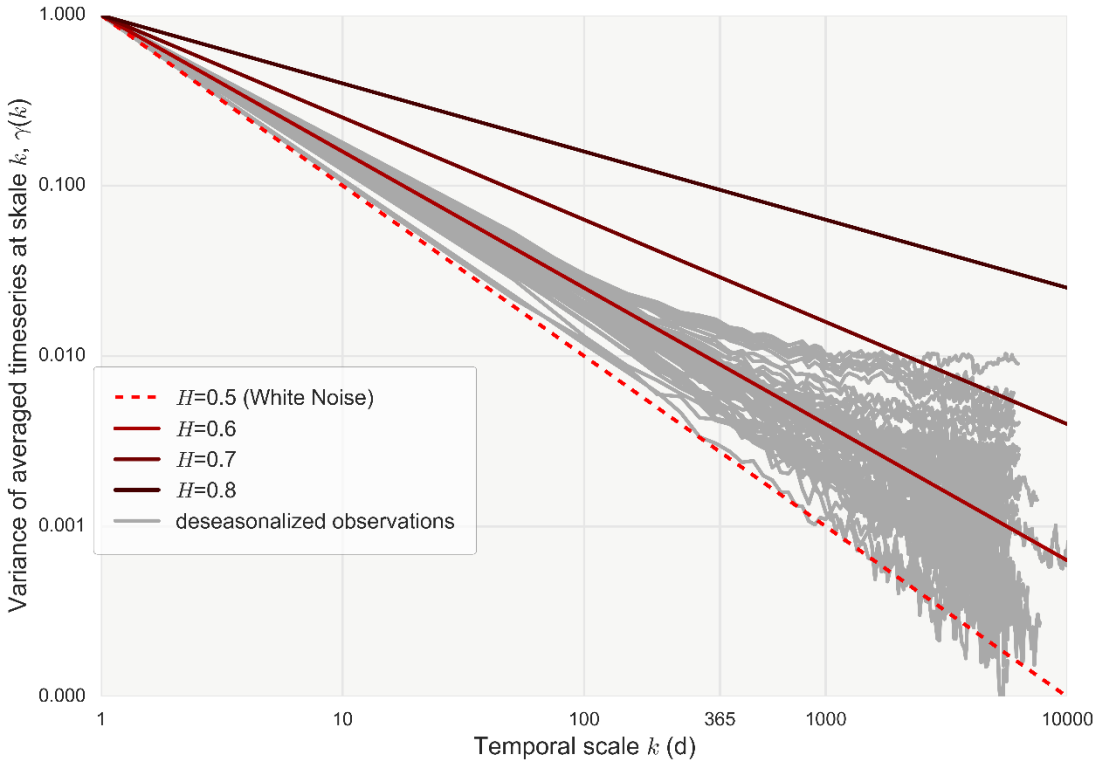


Figure 6.13 Empirical climacograms of the 60 daily rainfall series used in the analysis along with theoretical lines for $\boldsymbol{H}$=0.5, 0.6, 0.7, 0.8.

Table 6.2 Summary statistics (first and third quantiles, Q1 and Q3, mean and standard deviation, St.Dev.) of the properties of the rainfall dataset. Mean, Variance, Skewness and Kurtosis are estimated for the wet record.

Statistic	Mean	Variance	Skewness	Kurtosis	Prob. Dry	H_{daily}	H_{annual}	Years	Missing %
Q1	3.68	24.85	2.9	17.28	0.47	0.56	0.55	153	0.75
Mean	4.98	64.85	3.39	24.03	0.55	0.63	0.67	169.25	2.62
Q3	5.91	64.64	3.54	25.85	0.61	0.7	0.77	173	1.31
St.Dev.	2.27	94.15	0.72	10.94	0.11	0.09	0.13	24.66	5.11

6.6 Clustering in real world rainfall extremes II: HK dynamics?

6.6.1 Analysis of daily rainfall extremes in the Netherlands

It should be evident by now that the clustering dynamics of extremes depend not only on the persistence properties of the parent process but on its higher-order moments as well. The identifiability of clustering also varies depending on the choice of the threshold, which may need to be modified for mixed type processes, as discussed before. In our case, this means that depending on the probability dry of each station

the chosen threshold will correspond to a different one for the 'wet' record of each station. Therefore, a blind comparison of different stations with the obtained MCPL for a given threshold could be uninformative depending on the variability of probability dry in the sample of the stations. In order to apply the methodology effectively in as many stations as possible we assume a climatically homogenous regions in which the rainfall timeseries can be regarded as realizations of a single process. For this purpose, we select the region of the Netherlands in which 28 out of the 60 stations are located and preliminary analysis showed small variability of the summary statistics. We estimate the average values of the first four moments of the deseasonalized records for all 28 stations and we also estimate the H parameter resulting from the analysis of the daily values. We form an ensemble of 10^3 Monte Carlo simulations for the average number of years of the sample (160 years) with an HK-model preserving the first four moments and subsequently, compare its clustering behaviour with the one observed from the sample of the stations. We also repeat the Monte Carlo simulation for a white-noise process. We present both analyses in Fig. 6.14. It is evident that the assumed model is consistent with the majority of the observed records, with only a few stations located at the south-west of the Netherlands exhibiting even stronger clustering outside of the 95% region of the assumed HK model. As expected, as the threshold increases evidence of persistence is progressively 'lost' and the probabilistic behaviour of POT occurrences resembles a purely random one.

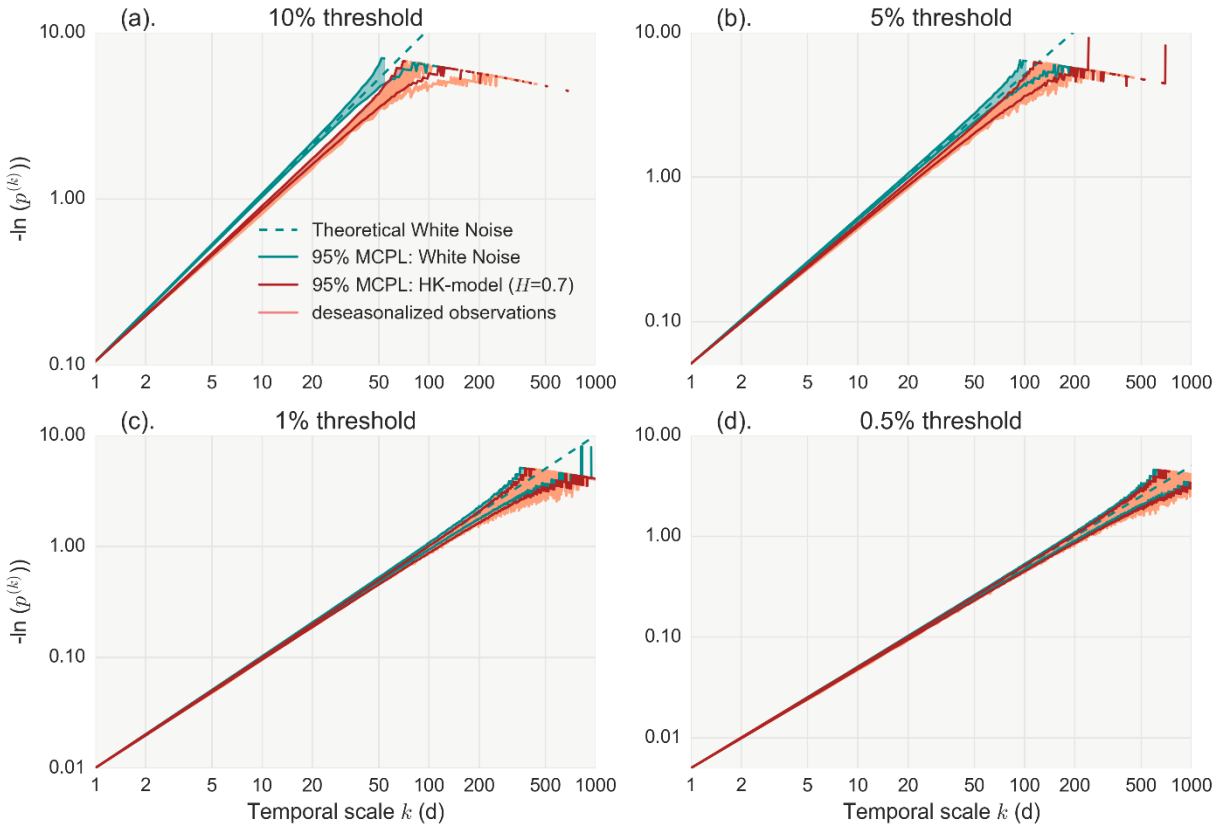


Figure 6.14 Minus natural logarithm of non-exceedance probability versus scale (NEPvS) index on double logarithmic axes for deseasonalized series for the 28 rainfall records in the Netherlands along with 95% MCPL of the fitted model with $\boldsymbol{H}$=0.7, for four different thresholds: (a) 10%, (b) 5%, (c) 1% and (d) 0.5%.

6.6.2 Case study of daily rainfall in Stykkisholmur

As a second case study we select a single station, located in Stykkisholmur (Iceland), which is the station with the most peculiar behaviour among all those we analysed. We repeat the Monte Carlo analysis for both a white noise process and a HK process preserving the first four moments and the H (= 0.65) parameter of the record. Results are shown in Figure 6.15. It is interesting to note that clustering in this case appears stronger than predicted by the HK model. The Monte Carlo experiment is repeated for H = 0.7 to explore the possible impact of estimation uncertainty due to the standard deviation bias in finite sample sizes (Koutsoyiannis and Montanari, 2007). In this case, the MCPL approach the observed data for the lower threshold, yet the impact is lower for the higher threshold. A similar behaviour was found in the station of Uppsala. We hypothesize that this 'discrepancy' between the persistence found in the parent process and the stronger one implied by the extremes might be explained by the impact of large-scale atmospheric circulation patterns (as the NAO) on rainfall extremes, which might need even longer record lengths in order to be effectively summarized by the second-order characterization provided by the H parameter.

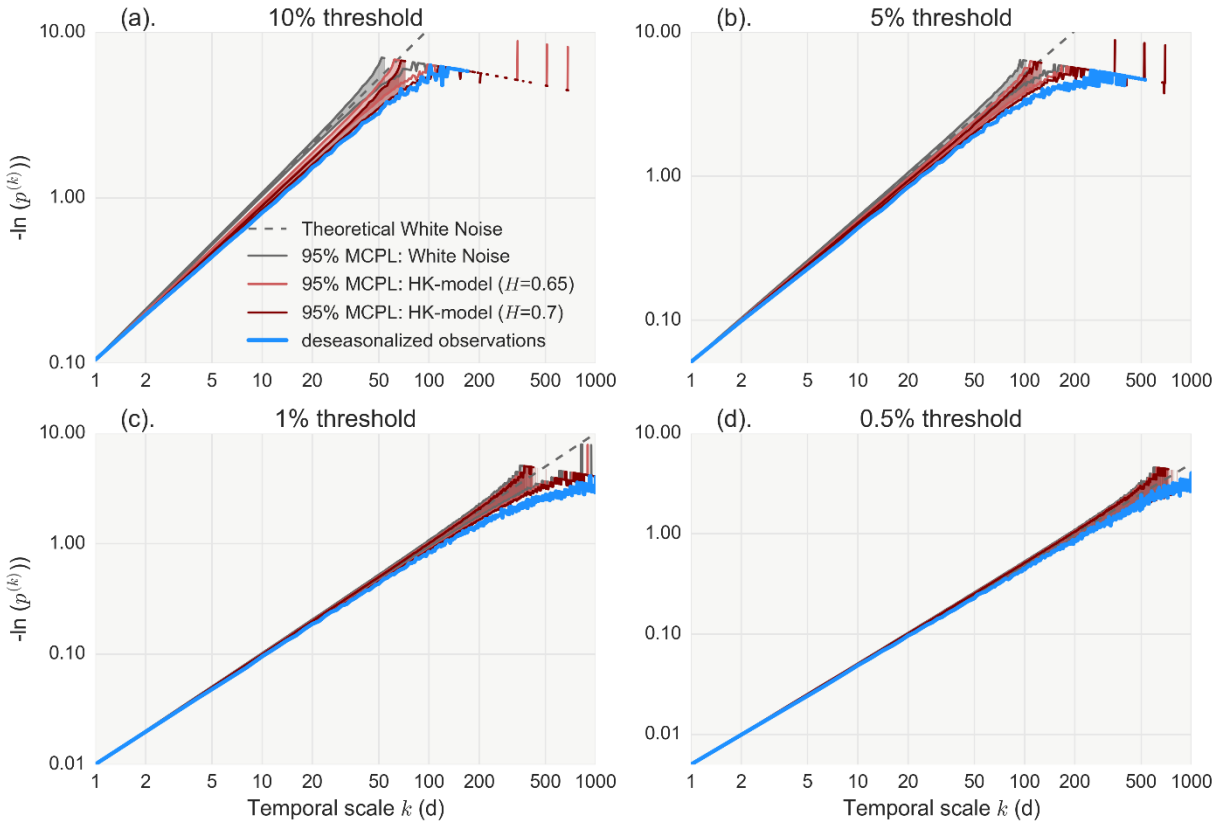


Figure 6.15 Minus natural logarithm of non-exceedance probability versus scale (NEPvS) index on double logarithmic axes for the deseasonalized series of Stykkisholmur in Iceland along with 95% MCPL of the fitted models with $\boldsymbol{H}$=0.65 and $\boldsymbol{H}$=0.7, for four different thresholds: (a) 10%, (b) 5%, (c) 1% and (d) 0.5%.

6.6.3 Modelling the clustering behaviour of all records

We apply the NEPvS model to both seasonal and deseasonalized timeseries of the rainfall data of all 60 stations in order to assess its applicability in all cases. We employ the deseasonalized scheme M1. In Figure 6.16 we plot the boxplots of the estimated parameters η and ξ as well as the RMSE for the seasonal and the deseasonalized series for three different threshold, 1%, 5% and 10%. From the fitted parameters, it is reaffirmed by this analysis as well that as the threshold decreases the estimates of the parameters deviate from the ones obtained for the IID case ($\xi = 0.5$ and $\eta = 1$). From the RMSE (Fig. 6.16c), it can be seen that the proposed model describes very well the deseasonalized data and fairly well the original observations, and in both cases the modelling efficiency improves for lower thresholds. Seasonality is associated with increased temporal clustering in the intermediate scales (approx. 20-150 days), which manifests with a curvature in the NEPvS plots that the model captures less efficiently compared to the deseasonalized case, typically producing a straight line plot. Also, it is evident that results concerning the threshold are not as robust for this case, since the impact of the threshold on seasonal clustering may vary depending on the specific seasonal regime. For instance, it is expected that for stations with prominent seasonality, high thresholds will show increased clustering only in the wettest season, whereas lower threshold will enable inspection of clustering in more seasons. However, depending on the characteristics of the seasonal regime and the intensity of the specific seasons, the temporal mixture of extremes from the different seasons differs from case to case, and thus, it is not straightforward to discern the impact of seasonality from a bulk fitting to all cases. On the other hand, for the deseasonalized cases it is clear that 'dependence' emerges as the threshold lowers.

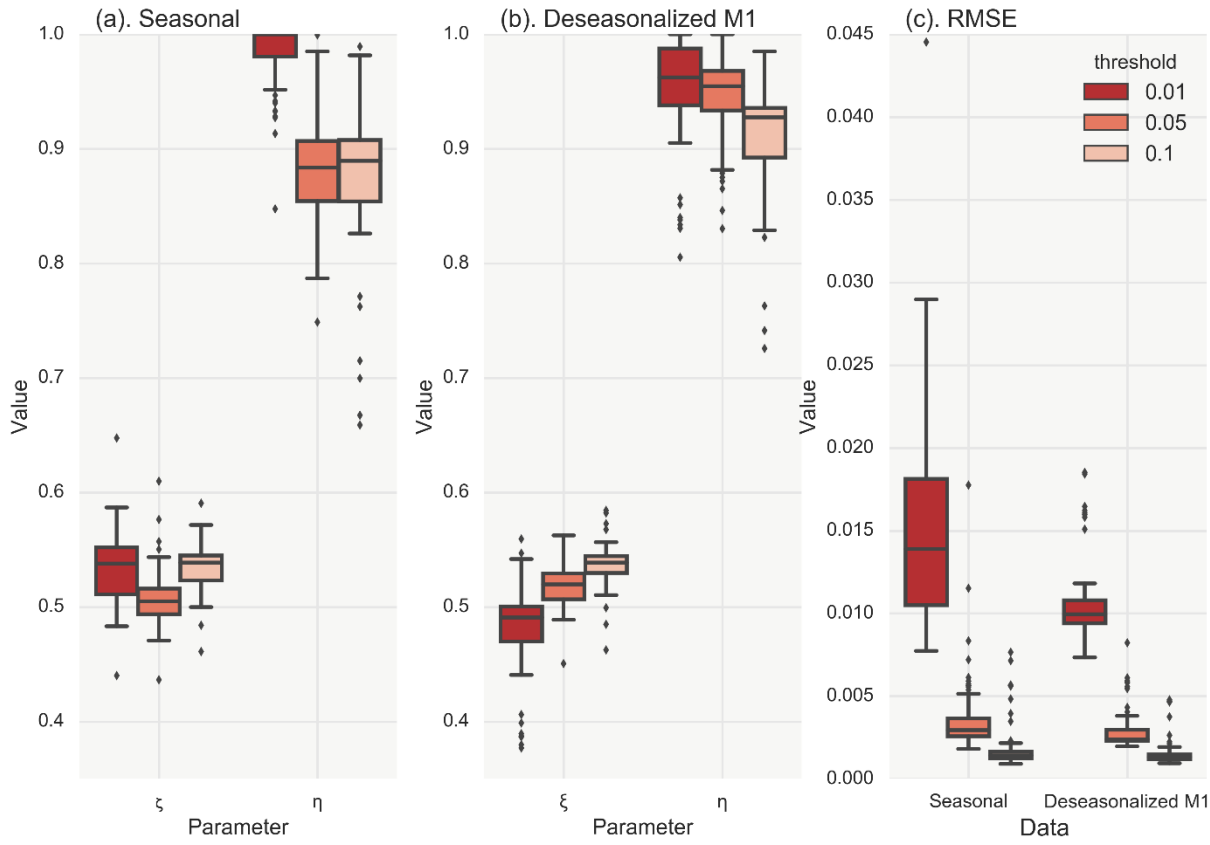


Figure 6.16 Boxplots of (a) parameter η, (b) parameter ξ and (c) RMSE from the fitting of the model to the seasonal and deseasonalized series by M1 for three different thresholds (1%, 5% and 10%).

6.7 Discussion

Clustering of extreme events is related to the presence of persistence, or HK dynamics, in natural processes. Here we approached this relationship with a twofold intention; first to 'retrieve' persistence from records of maxima, and second, to characterize it by probabilistic means. To this aim, we have introduced the NEPvS index, for which we also propose a model. The index examines the probabilistic behaviour of POT occurrences across multiple scales and proved successful in revealing persistence from extremes from various non-Gaussian timeseries, for which well-known tools performed poorly.

It seems, though, to be difficult to establish general analytical relationships linking the NEPvS behaviour to the H parameter of the parent process, which is true without even considering the uncertainty involved in estimating H from small record lengths in the first place. As the H parameter is a second-order characterization of a process, generation schemes reproducing H behaviour but coupled with different marginal distributions (having different high-order moments), will yield different behaviours of extremes. For instance, clustering of extremes and its identifiability appears to be much more prominent in Gaussian processes. The task therefore, of linking clustering of extremes to the H parameter, without also accounting for the specific high-order moments of the timeseries seems infeasible. We showed though, that the threshold is a key determinant in this respect, as lowering the threshold, i.e. moving towards the central tendency of the data, enables better identification of

persistence. On the contrary, as the threshold increases, evidence of persistence is progressively lost and the behaviour of extremes may falsely suggest independence of the parent process.

Application of the NEPvS index to daily rainfall data showed that there may exist significant departures from the case of independence, particularly for lower thresholds, which are dependent on the location and specific climatic region. In general, the behaviour of rainfall extremes in multiple case studies (28 stations in the Netherlands and 1 in Iceland) was found by means of extensive Monte Carlo simulations, to be consistent with HK dynamics characterized by moderate *H* parameters (in the range 0.6-0.7). The NEPvS model showed a very good fit to the probabilistic behaviour of exceedances for the seasonal and deseasonalized observations across multiple scales for all 60 stations. As a similar version of the model has been previously proposed to describe the probability dry across multiple scales (Koutsoyiannis, 2006), this result suggests that there exists a probabilistic law which effectively describes the multi-level exceedances of rainfall thresholds across scales, from zero-crossings (wet days) to high-level crossings, as the ones examined here.

From a theoretical point of view, these findings suggest that it is important to study change and clustering in a consistent stochastic framework examining the whole process behaviour, in order to better understand the process dynamics and avoid retaining 'preconceived' assumptions, such as IID, which may be inconsistent with the physical reality. For instance, various trend tests assume IID for the examined process, while modified tests accounting for persistence (Hamed, 2008), also do not consider its interplay with the higher order moments. Therefore, it is likely that they fail to account for extremes from complex processes, leaving aside issues regarding problematic applications due to misinterpretation of stationarity (Koutsoyiannis and Montanari, 2015; Montanari and Koutsoyiannis, 2014). Overdispersion in POT rainfall events has been also studied lately and attributed to a mixture of Poisson models, representing different climate regimes (Tye et al., 2018) as well as seasonality mechanisms (Serinaldi and Kilsby, 2013). Although, we have found as well that in some cases seasonality accounts for most of the observed clustering in the rainfall extremes, by performing multiple MC experiments focusing on the deseanonalized extremes, we have revealed consistency with HK dynamics. We note though that as the *H* parameter for rainfall revolves around the value of 0.6 and rainfall is a heavily skewed process, it is expected that identifiability of persistence from extremes will be limited, except if ones lowers the threshold. Nevertheless, this highlights an alternative scientific hypothesis to be considered in 'attribution' studies, which is the emergence of clustering and overdispersion of extremes from persistence in the parent process.

From a practical point of view, the presence of persistence in the parent process affects estimation of extreme values, and therefore various design outcomes, in multiple ways. Although the theoretical definition of return period is still valid under presence of persistence (Koutsoyiannis 2008; Volpi et al., 2015), the statistical estimates of distribution quantiles for a specified return period are severely impacted. Other important implications concern flood risk underestimation under persistence (Serinaldi and Kilsby, 2016b), as well as underestimation of IDF curves when the temporal dependence is disregarded (Roy et al. 2018). Therefore, although persistence of the parent process is less evident in the series of its extremes, and it is highly unlikely

that it can be fully retrieved except for very low thresholds, its impact cannot be disregarded when studying extremes, even if the latter appear independent. It is worth recalling the existence of theoretical arguments concerning the validity of the fundamental EVT results (limiting distributions etc.) under weak presence of persistence (Leadbetter, 1983b; see also Section 2.3). However, for scientific applications, which involve estimation from data of finite, and typically small record lengths, the presence of persistence in the process induces uncertainty in the estimation, as the actual information content of the data is lower than that for IID conditions (Koutsoyiannis and Montanari, 2007). This uncertainty inevitably propagates into the extreme value estimates.

The existence of clustering also increases the arguments towards the use of the POT method for sampling of extremes, instead of block maxima approaches which tend to hide dependence, as also evident in Fig.6.2. As the threshold plays a vital role, using POT approaches with more than one event per year on average, which is the common practice, is also equally important. Empirical declustering approaches (Lang et al., 1999) may as well be non-effective if they do not take into account each process characteristics. In this regard, we argue that instead of seeking to resort to independence, often at the cost of reducing the available information (e.g. by discounting 'dependent' data), accounting for dependence is a more viable and consistent way forward. In fact, the use of all the set of observations has been recently advocated (Volpi et al., 2019), while the emergence of new types of high-order moments (Koutsoyiannis, 2019c) that exploit the whole set of observations, provide an improved stochastic framework for applying this principle.

6.8 Conclusions

This research deals with the question of identifying the links between persistence in the parent process and clustering of extremes, with the specific aim to 'rediscover' the usually 'lost' persistence when one examines records of maxima. This is achieved by devising a probabilistic characterization of clustering of extremes. The main findings are summarized below:

- There is significant influence from both the second-order properties and the high-order moments of the parent process on the generated extremes, and therefore characterizations of clustering of extremes need to account for both.
- Identifiability of persistence from records of maxima is in general limited and weakens as the threshold for extremes increases.
- The estimates of the Hurst parameter by the climacogram and dispersion index analyses are found to be severely biased downward when derived from extremes originating from non-Gaussian processes.
- A new probabilistic index is proposed to represent clustering based on the probability of non-exceedance of a given threshold across scales, called the NEPvS (non-exceedance probability vs scale) index.
- The NEPvS exhibits scaling behaviour which is described by a proposed model accurately simulating the probability of exceedance of a threshold at multiple temporal scales.

- The index is transparent and can be directly used for statistical testing of departures from independence. Case-specific Monte Carlo simulations are needed to validate more complicated models coupling persistence with different marginal properties.
- The POT approach applied with 'low' thresholds is a robust and informative way to reveal the clustering dynamics of extremes, in contrast to the block maxima method which hinders identifiability of persistence.
- Deseasonalized daily rainfall POT events may show prominent departures from independence especially at lower thresholds, which may become important depending on the climatic region. Extensive station-specific Monte Carlo experiments showed consistency of clustering of extremes for various examined thresholds with assumed HK models fitted based on the properties of the parent process.

Further research is required in order to obtain analytical mathematic results for extremes arising from persistent processes, with the aim of constructing estimators for any distribution type and dependence structure without the need for Monte Carlo validations. However, the latter is doubtful as a task, since extremes over scales are controlled by higher order moments, which are also difficult to estimate correctly from data (Lombardo et al., 2014). Recently proposed moment types with unbiased estimators across all orders that can also model joint properties of processes could provide a way to circumvent this (Koutsoyiannis, 2019c).

We conclude that extremes tend to 'hide' the persistence of the parent process, often falsely signalling independence. Regardless however of the strength of the evidence, the impact of persistence in the parent process on the estimation of extreme values is nonetheless present. In this respect, more research should focus on the stochastic properties of extremes from natural processes, where dependence mechanisms manifest themselves across various temporal scales and challenge common assumptions and practices.

7. MANIFESTATION OF PERSISTENCE IN THE DISTRIBUTION OF EXTREMES AND THEIR PROPERTIES

This Chapter examines the applicability of the EVT results, presented in Chapter 2, for the case of extremes arising from persistent processes. To this aim, the appropriateness of the GEV distribution for modelling maxima from persistent processes is assessed through extensive Monte Carlo simulations and the effect of persistence in the parameterization of the GEV distribution is identified. Manifestations of dependence in the distributional properties of POT events at the annual scale are also highlighted. The theoretical results are corroborated by real-world evidence from records of rainfall and streamflow extremes. The second-order HK stochastic framework is applied for the modelling of the parent process and its ability to reproduce the empirical extremal patterns is evaluated.

7.1 Introduction

In Chapter 6, persistence (HK dynamics) is shown to manifest itself in the occurrence process of extremes across scales. Compared to extremes from a purely random process, persistent extremes have a lower probability of being encountered at a scale as a result of clustering. The change in the multi-scale occurrence behaviour of extremes under persistence implies that the properties of the extremes at a given scale might also be affected. This is the focus of this chapter.

In a first examination, we investigate the distribution of block maxima under various degrees of dependence. The relevant theory was outlined in the review section 2.3 of Chapter 2. Here, we specifically aim to assess the validity of the GEV distribution as an approximation in cases of long-range dependence, and to identify the impact of the latter in the GEV distribution parameterization, as well. In order to complement the analysis of temporal properties of extremes in a given scale, we also study the behaviours of POT events as these are indicative of the short-term clustering properties of extremes. These local dependence patterns cannot be revealed by the block maxima approach due to the disruption of short-term clustering by the indiscriminate sampling of one event per time-window.

Empirical investigation of rainfall and streamflow series is performed to showcase the hydrological relevance of extremal dependence. Finally, we carry out a preliminary investigation of the performance of the HK stochastic framework in modelling stochastic patterns of extremal dependence.

7.2 Assessing impacts of dependence on block maxima and their modelling by the GEV distribution

The asymptotic results of the EVT and the set of underlying assumptions were reviewed in Chapter 2. It was shown that by virtue of equation (23) for a process with limited long-range dependence in extremes (fulfilment of $D(u_n)$ condition), but with presence of local dependence in exceedances (non-fulfilment of $D'(u_n)$ condition) quantified by a positive extremal index, the non-exceedance probability of maxima of a given scale increases compared to maxima derived from an IID process with the

same marginal distribution. We test this theoretical result by comparing the probabilistic behaviour of block maxima derived from processes with HK dependence structure and AR(1) dependence. The HK and AR(1) series have exactly the same marginal distribution for the normally-distributed processes, while share the first four moments for the case of the Gamma-distributed processes.

It can be seen (Fig. 7.1) that indeed dependence lowers the exceedance probabilities of extremes of lower magnitude yet the impact becomes negligible as the magnitude increases. This behaviour however breaks down for extremely dependent process (H=0.95) and heavy-tailed processes (Fig. 7.1j, k, l), in which case the exceedance probability of larger extremes appears greater. Yet these cases are severely impacted by estimation uncertainty and therefore this effect is likely the result of insufficient record length. Gaussian autoregressions (Fig. 7.1a, b, c) appear to only marginally impact the behaviour of extremes, which verifies the relevant theory suggesting extremal index θ=1, whereas in the case of autoregressions with non-Gaussian innovations a considerably higher impact of dependence is observed. This implies that heavy-tails and dependence have synergistic effects on the behaviour of extremes.

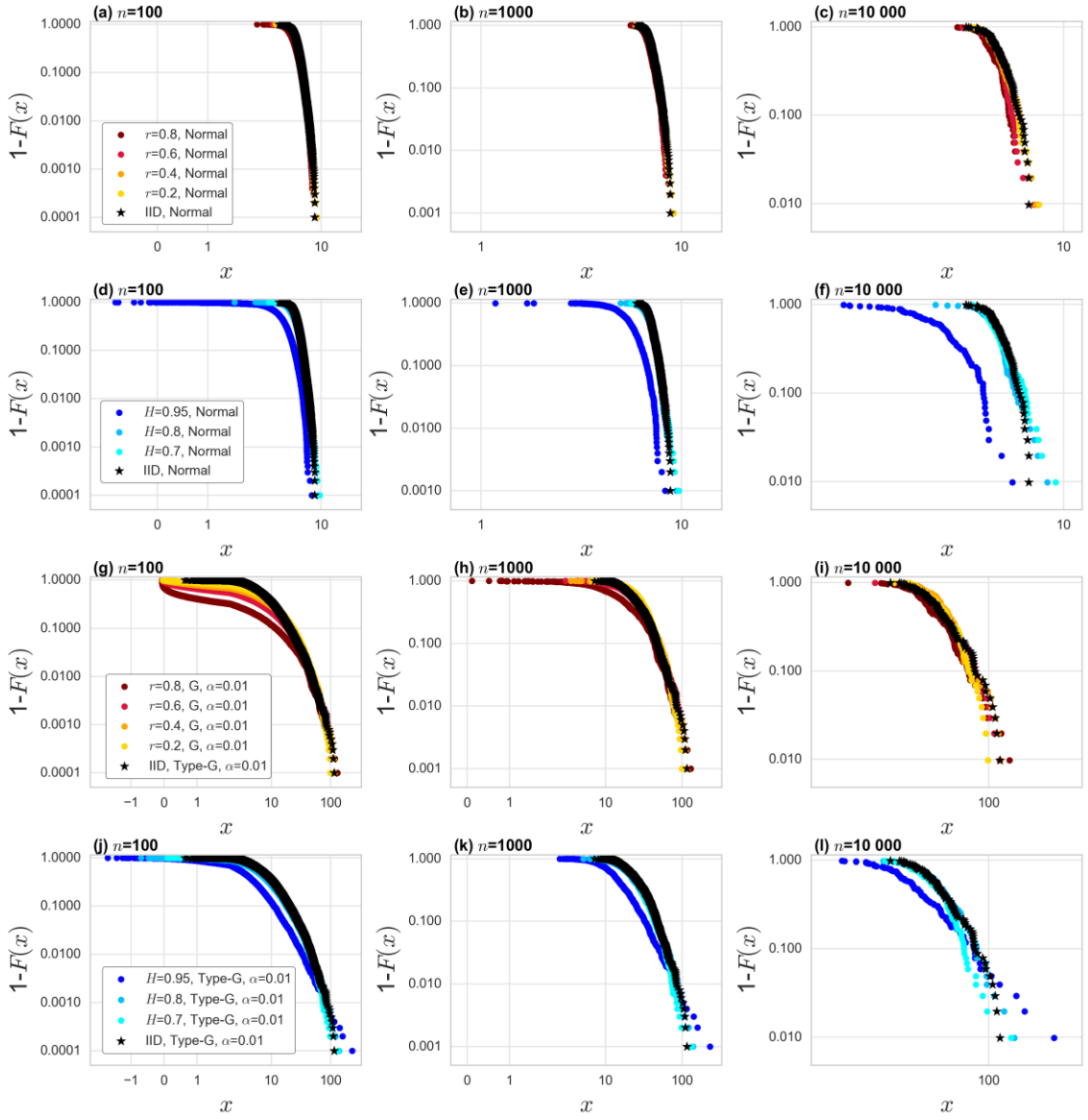


Figure 7.1 Exceedance probabilities of maxima from blocks of length n=100, 1000, 10 000 for timeseries generated from an AR(1) model with standard normal distribution (a, b, c) and gamma distribution (g, h, i) and an HK model standard normal distribution (d, e, f) and gamma distribution (j, k, l) for various degrees of dependence.

To gain insights into possible impact of estimation uncertainty on the above results, we also examine shorter series of length 150 years and we specifically focus on the annual timescale. Results from 10^3 Monte Carlo simulations from an HK-process model with Type-G and Type-P distribution, shown in Fig. 7.2 and 7.3, reaffirm the previous remarks. In terms of the expected values, the IID process acts as the upper boundary of the distribution of the exceedance probabilities, yet dependence inflates the prediction limits (7.2 b and 7.3 b).

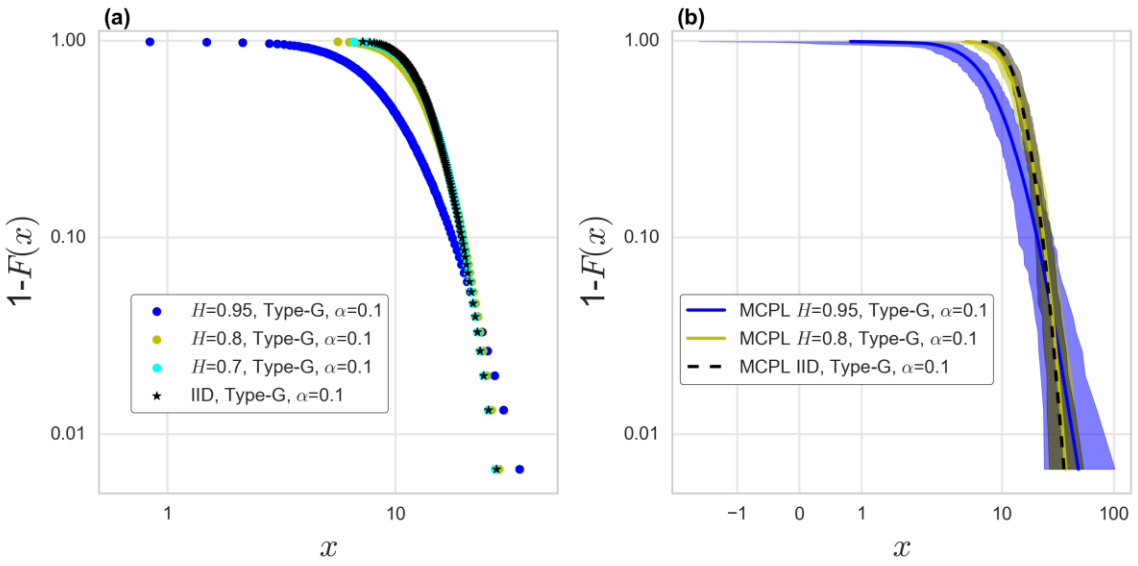


Figure 7.2 Expected values of the exceedance probabilities of maxima from blocks of length n=365 for timeseries generated from an HK-type model with Type-gamma distribution from 10^3 simulations (a) and MCPL for the cases of H=0.95, 0.8 and IID(b).

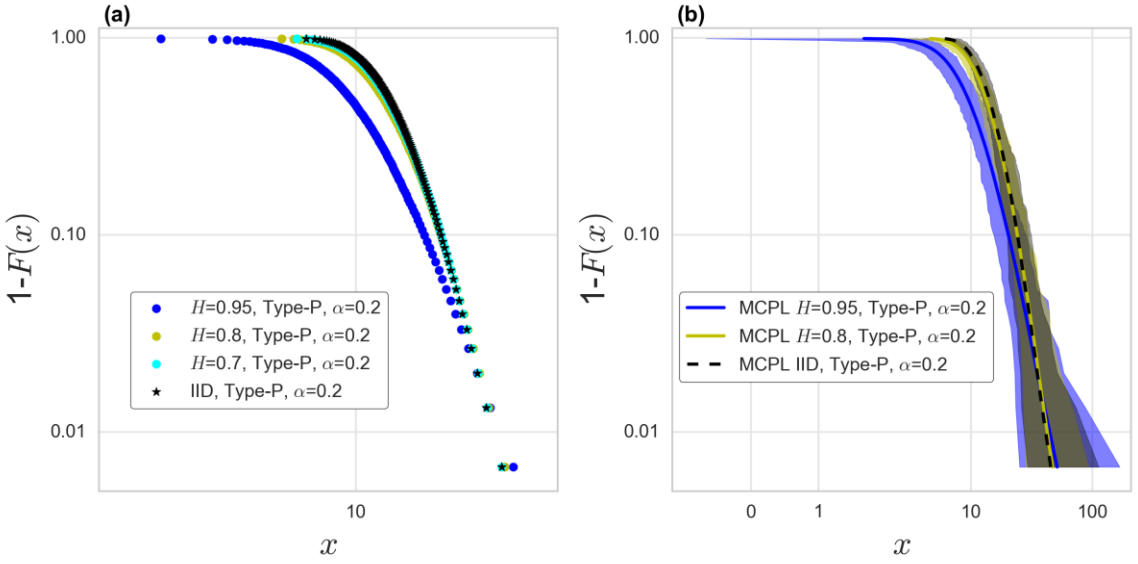


Figure 7.3 Expected values of the exceedance probabilities of maxima from blocks of length n=365 for timeseries generated from an HK-type model with Type-pareto distribution from 10^3 simulations (a) and MCPL for the cases of H=0.95, 0.8 and IID(b).

Before examining the effect of dependence on the parameterization of the GEV distribution, we first assess the quality of the latter as an approximation for extremes of persistent processes. Since it is difficult to analytically derive the exact distribution of maxima in such cases, we employ the empirical ones derived from the long benchmark series (10^6 length). For the fitting process, we assess both the fits of the maximum likelihood method and the weighted least-squares (WLS) method, with weights equal to the empirical quantiles, as the one used in Chapter 3 for fitting the GEV to seasonal extremes. Results from application to a series with H=0.7 and different marginal distributions are shown in Fig. 7.4, for block sizes n=10, 100, 1000.

It is apparent that the fit of the WLS outperforms the maximum likelihood fit for the smaller block sizes, due to the violation of the independence assumption of the method. On the contrary, the WLS algorithm yields a very good fit even for cases where the asymptotic arguments for the GEV certainly do not hold (n=10). A slightly less good performance is obtained for the case of the lower quantiles from n=10 of the series with Type-Gamma distribution and shape parameter α=0.01. Yet the latter is reasonable as the WLS algorithm favours a better fit to the larger quantiles over the fit to lower quantiles. Therefore, given an adequate fitting method, the GEV distribution appears flexible enough to model extremes from a wide range of persistent process. The latter was verified also for the extreme case of series with H=0.95, in which case the fit of both methods was even better.

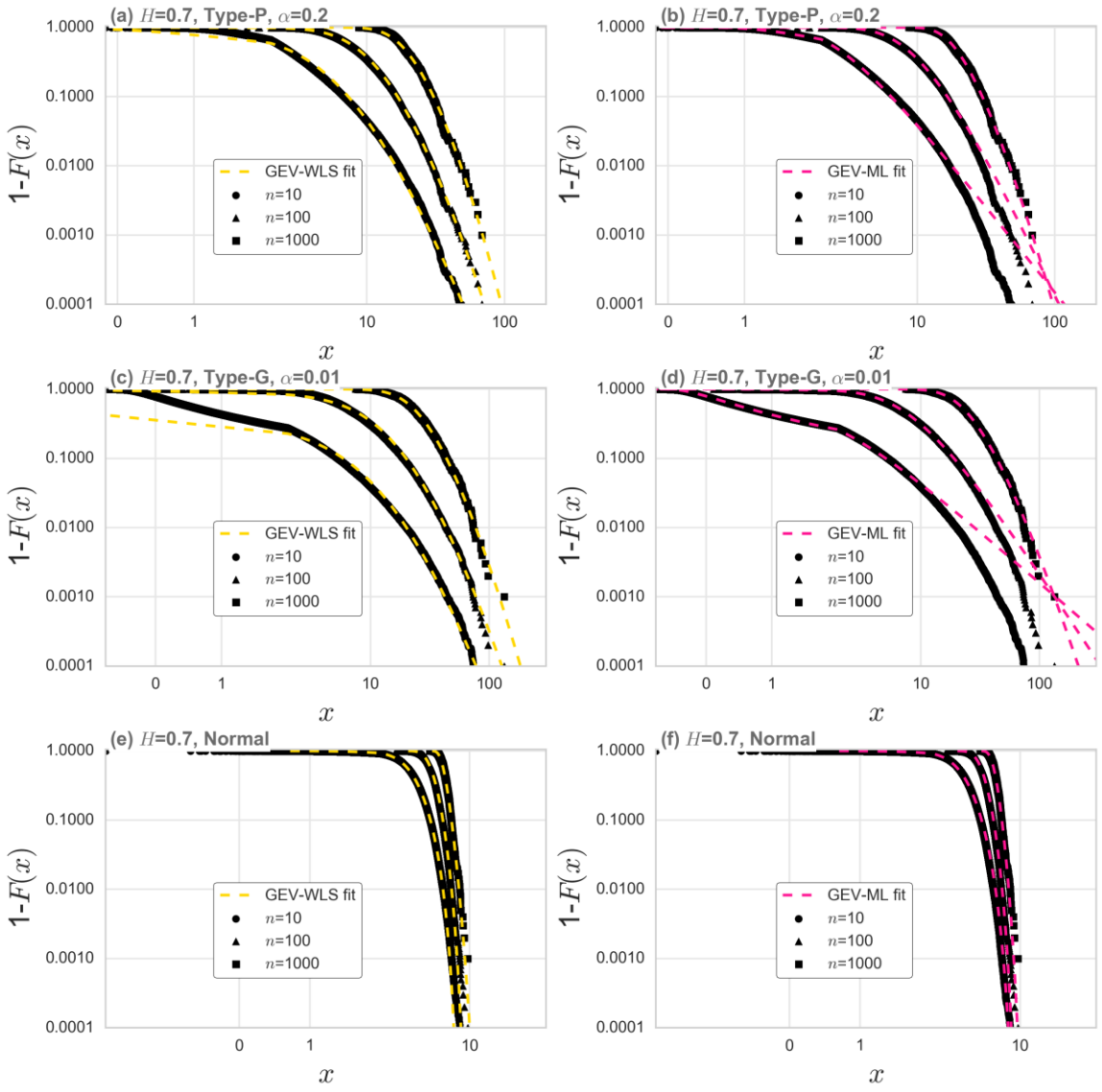


Figure 7.4 GEV fits to maxima from blocks of length n=100, 1000, 10 000 for timeseries generated from an HK model with H=0.7 and Type-Pareto distribution (α=0.2), Type-Gamma (α=0.01) and standard normal distribution, by the weighed-least squares method (a, c, e) and the maximum likelihood method (b, d, f).

To evaluate the effect of dependence on the parameterization of the GEV distribution, the timeseries of length 150 years (150×365) are employed and the GEV is fitted to the block maxima of block size n=365. Results from application to three different distributions with varying degree of HK dependence are shown in Fig. 7.5. As a first remark, it is interesting to observe that convergence of extremes from the normal distribution to their domain of attraction, i.e. the Gumbel distribution (shape parameter 0) is so slow, that it is not satisfied even for the IID series, at this block length. For the other two distributions, convergence is satisfactory, as in the case of the Gamma with shape parameter 0.1, the domain of attraction is the Gumbel and indeed the values of the shape parameter are correctly identified, whereas for the Pareto with shape parameter 0.2, again the shape parameter of the GEV is close, albeit slightly underestimated.

Regarding the parameterization of the GEV as dependence increases, the change is manifested in the values of the location and the scale parameter, as generally expected, yet not in complete agreement with the extremal index theory. In fact, while the location parameter indeed decreases with increasing dependence, the scale parameter, on the contrary increases. The shape parameter remains reasonably stable, as suggested by Chapter 2 theory, except for the cases of very strong persistence, where the increase of the shape parameter is balanced by the decrease of the scale parameter. The insensitivity of the shape parameter to dependence is theoretically justified as the former relates to the shape of the tail of the process, which is a property of the marginal distribution. On the contrary the rescaling parameters of the GEV depend on n, and thus on the scale. In this regard, the increase in the scale parameter arises from the increased variance of block maxima from persistent processes, which is intuitive if one considers the increased variability of the persistent process at different scales. The decrease in the location parameter is expected as well, as also shown in the Fig. 7.2-7.3, and directly stems from the clustering dynamics. Extremes are clustered in years, and only the larger of them is selected as the maximum, discarding the rest. On the contrary, in an IID process, extremes have a higher change to be encountered as separate maxima in any given year and hence, to be modelled.

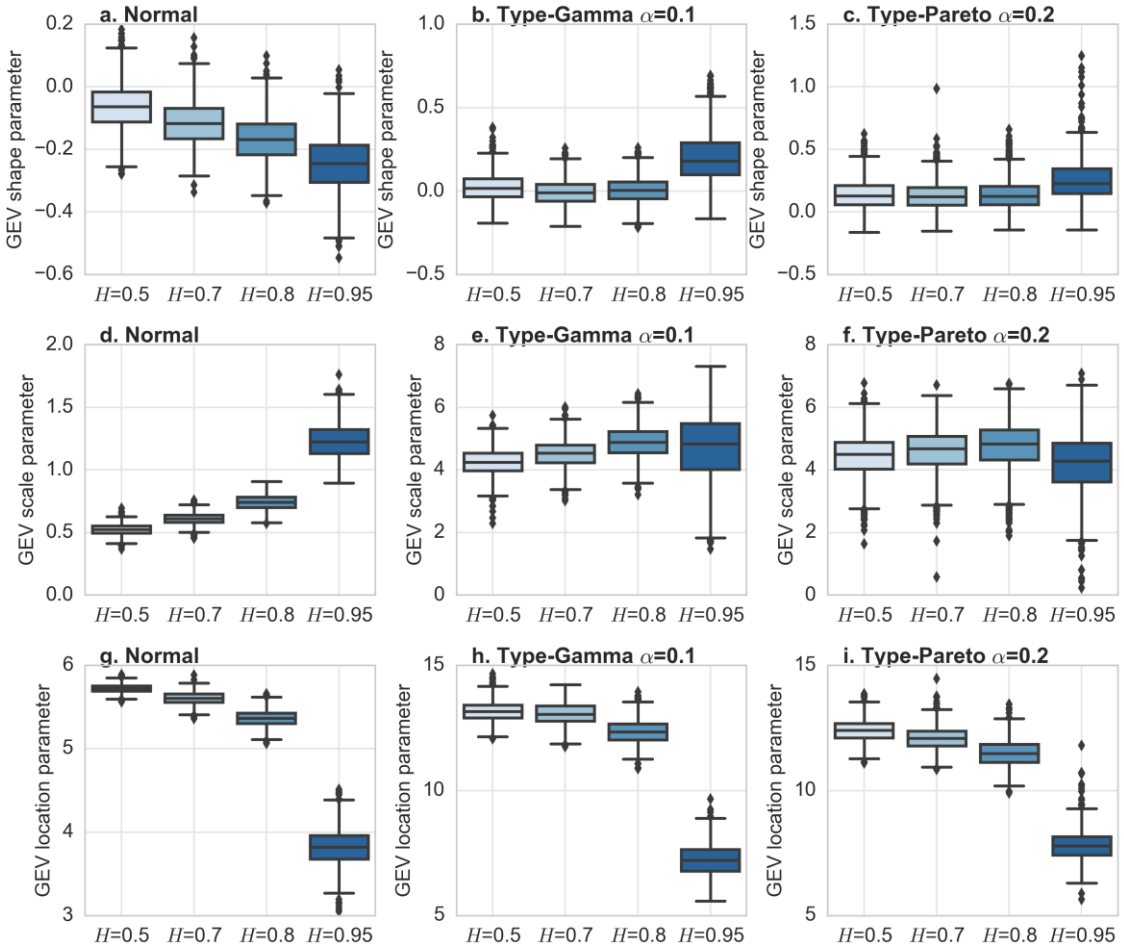


Figure 7.5 Shape, scale and location parameters from the fitting of the GEV distribution to the 1000 timeseries with normal distribution (a, d, g), Type Gamma distribution (b, e, h) and Type Pareto distribution (c, f, i) for varying $\boldsymbol{H}$ parameters.

7.3 Conditional properties of peaks over threshold under persistence

Switching the focus to the study of the behaviour of peaks over threshold in a given scale allows further insights into the effects of persistence. Namely, it enables direct inspection of clustering properties that are not revealed by the block maxima method, such as the cluster duration and the number of separate clusters in a year. For this analysis, we analyze a 500-year segment of the 10^6 Type-Pareto timeseries generated by the HK model with H=0.8, and subsequently, obtain prediction limits of the corresponding IID series produced by random shuffling of the original series. The shuffling technique has the advantage of ensuring that the generated random series has exactly the same distribution with the original. Therefore any changes in the extremal properties can be uniquely ascribed to the presence of dependence in the original series.

The following properties of POT events are examined: (a) the frequency distribution of the number of cluster maxima per year, (b) the exceedance probability distribution of the aggregate annual POT event intensity, (c) the number of annual POT events versus the mean POT event intensity, and (d) the frequency distribution of cluster duration. A cluster is defined here by one or more successive POT events separated from the next POT event by the occurrence of at least one event below the threshold. The threshold is chosen so at the number of cluster maxima (maxima of clusters of POT events, including single events) equals the number of years of the record, as in the annual maxima approach. Analysis of the aggregate intensity is inspired by the concept of collective risk in financial and insurance literature, which deals with the properties of sums of random variables (Iglehart, 1969). In hydrological literature, Serinaldi and Kilsby (2016b) suggested studying the behaviour of streamflow events under the collective risk viewpoint, treating streamflow POT as proxies for insurance claims.

Fig. 7.6 shows the stochastic patterns of POT events arising in the case of a long-term persistent process. In Fig.7.6a, is seen that both the frequencies of zero number of events and very large number of events tend to be greater under persistence. Fig 7.6b shows how exceedance probabilities of the aggregate annual intensity are generally higher compared to the IID case, as a result of clustering of events. Figure 7.6c shows that persistence also increases the positive association between the number of yearly events and their average intensity, compared to the IID case, where a lack of correlation is observed. In Fig. 7.6d, it is shown that the probability of observing clusters of large duration increases in the case of persistence. Recalling the findings of Chapter 6, it can be concluded that persistence increases the probability of experiencing large temporal periods with no extremes. On the other hand, it is shown that should such a period occur, it will probably last longer compared to the IID case, i.e. persistence induces a 'when it rains, it pours' dynamic.

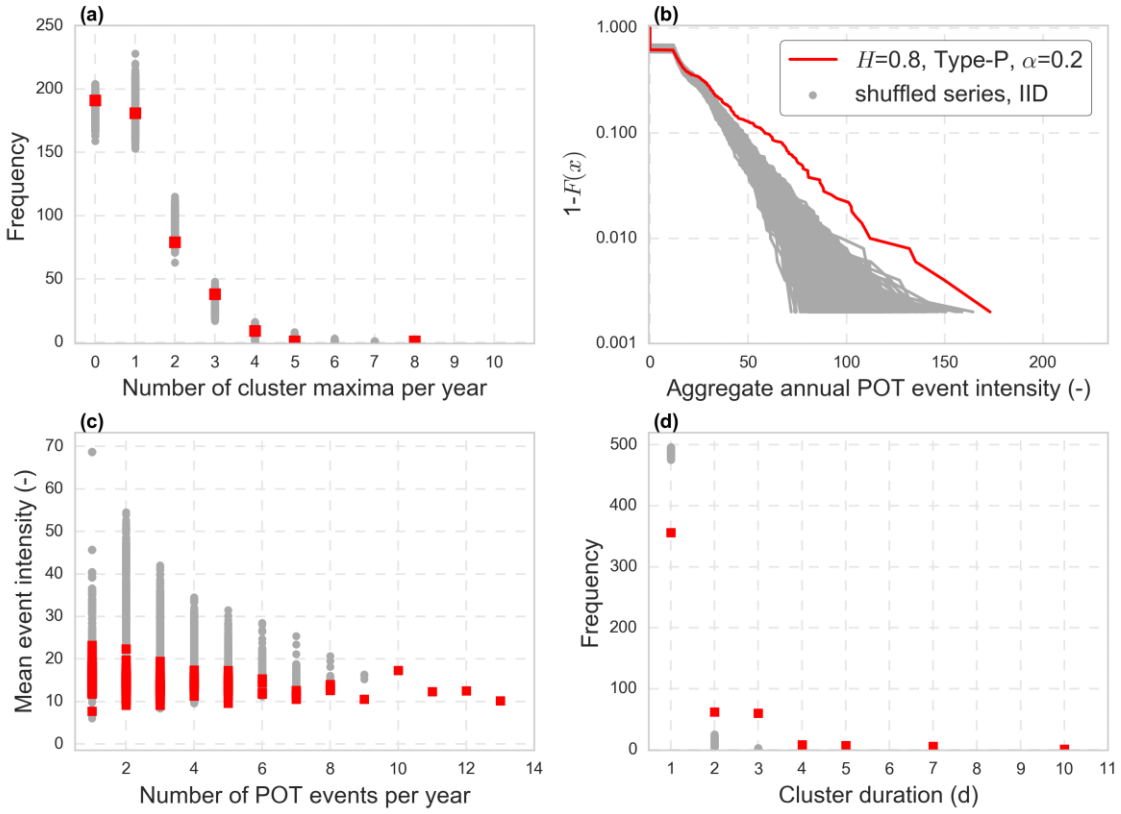


Figure 7.6 Plots of the following functionals of POT events from the original and 1000 shuffled timeseries (IID): (a) frequency distribution of the cluster maxima per year, (b) exceedance probability distribution of the aggregate annual POT event intensity, (c) number of POT events per year versus their mean annual intensity, and (d) frequency distribution of cluster duration.

7.4 Stochastic modelling of rainfall and streamflow extremal properties

In this section, we employ the previous shuffling methodology in order to trace the effect of dependence on real-world rainfall and streamflow extremes, derived from the Bologna daily rainfall series (206 years) and the Po river daily streamflow series (90 years). Further, we set up a preliminary modelling framework and test its effectiveness in capturing the observed patterns. We do not employ any deseasonalisation scheme for the rainfall and streamflow processes, since the 'distinct' effects of seasonality and persistence have been already investigated in Sections 7.1-7.3 as well as throughout Chapters 4-6. Rather the aim here is to assess the combined effect of dependence mechanisms on the extremal properties.

In Fig. 7.7, the annual maxima distributions of the daily rainfall in Bologna and the daily streamflow of the Po river record are shown. The behaviour rainfall maxima is within the MCPL range of the shuffled series although there is a tendency to deviate from the expected IID behaviour. The departure from the latter is however prominent in the case of the streamflow maxima, where the shift in the distribution is fully in line with the behaviour of a persistent process. In both cases the WLS fit of the GEV is very good and superior to the ML fit, which is not shown in the second case, due to being very poor.

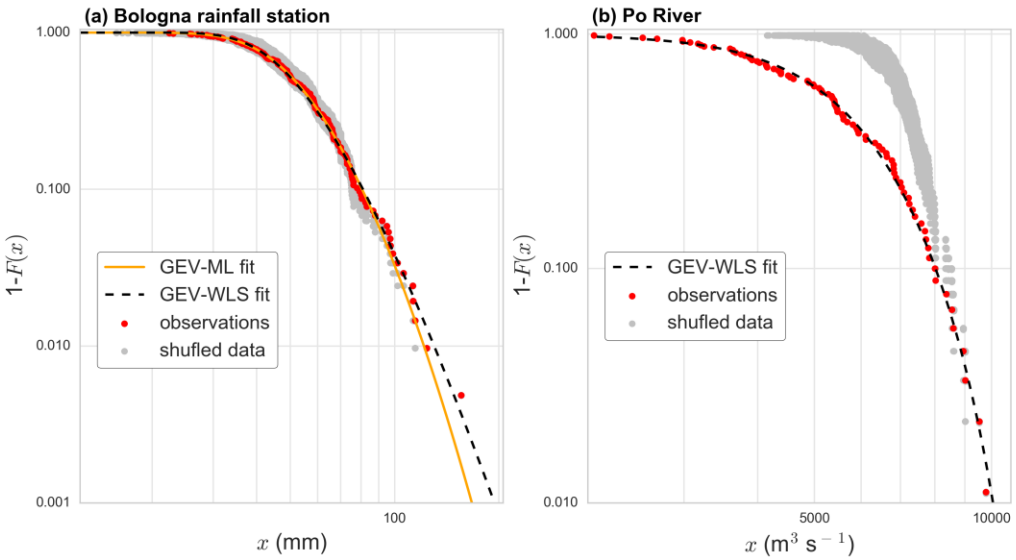


Figure 7.7 Exceedance probabilities of annual maxima for the original timeseries and the shuffled series for the Bologna rainfall series and the Po river daily streamflow series.

Proceeding to the second type of analysis, similar conclusions can be drawn. In this case, the number of POT declustered events (cluster maxima) equals the number of annual maxima for each record. In the rainfall series, dependence is manifested by an increase in the cluster duration distribution and thus, in the annual aggregate event intensity as well, as both deviate from the expected IID behaviour (Fig. 7.8). Dependence is overall much more prevalent in the streamflow series (Fig.7.9) where most studied functionals (Fig.7.9b,c,d) of extremes strongly deviate from their shuffled counterparts.

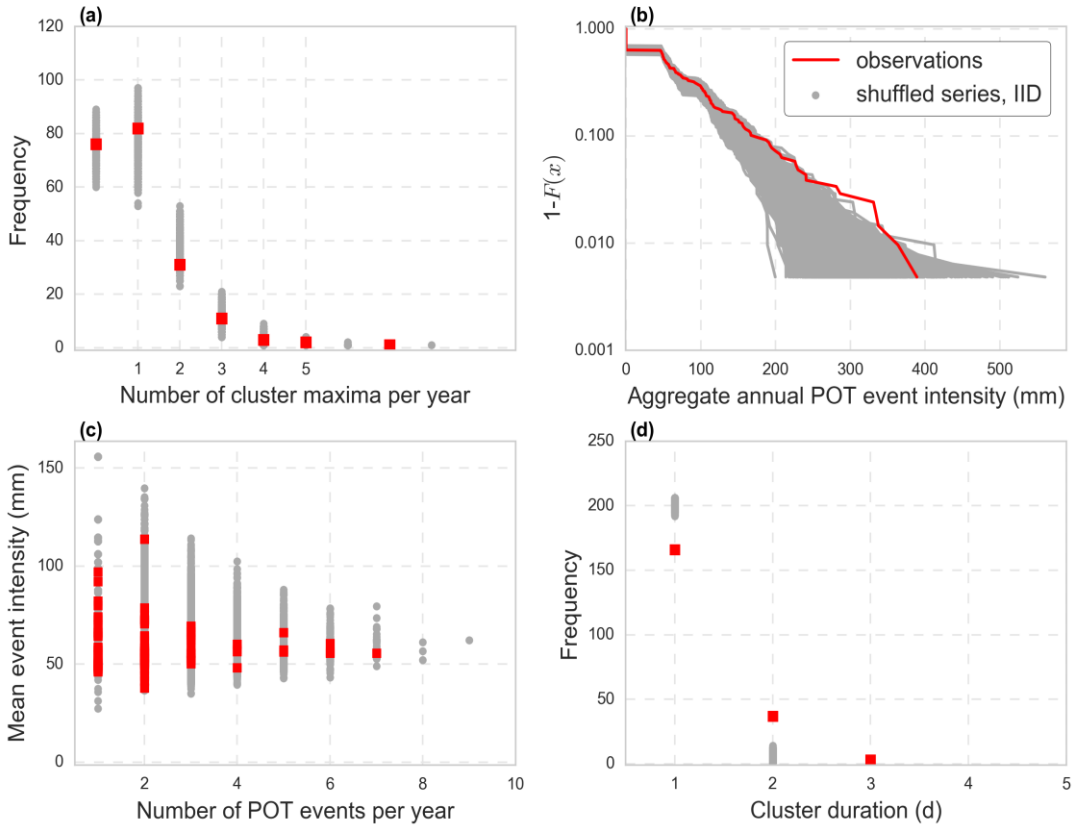


Figure 7.8 Plots of the following functionals of POT events from the original Bologna rainfall series (206 years) and 1000 shuffled timeseries (IID(a) frequency distribution of the number of cluster maxima per year, (b) exceedance probability of the aggregate annual POT event intensity, (c) number of POT events per year versus their mean annual intensity, and (d) frequency distribution of cluster duration.

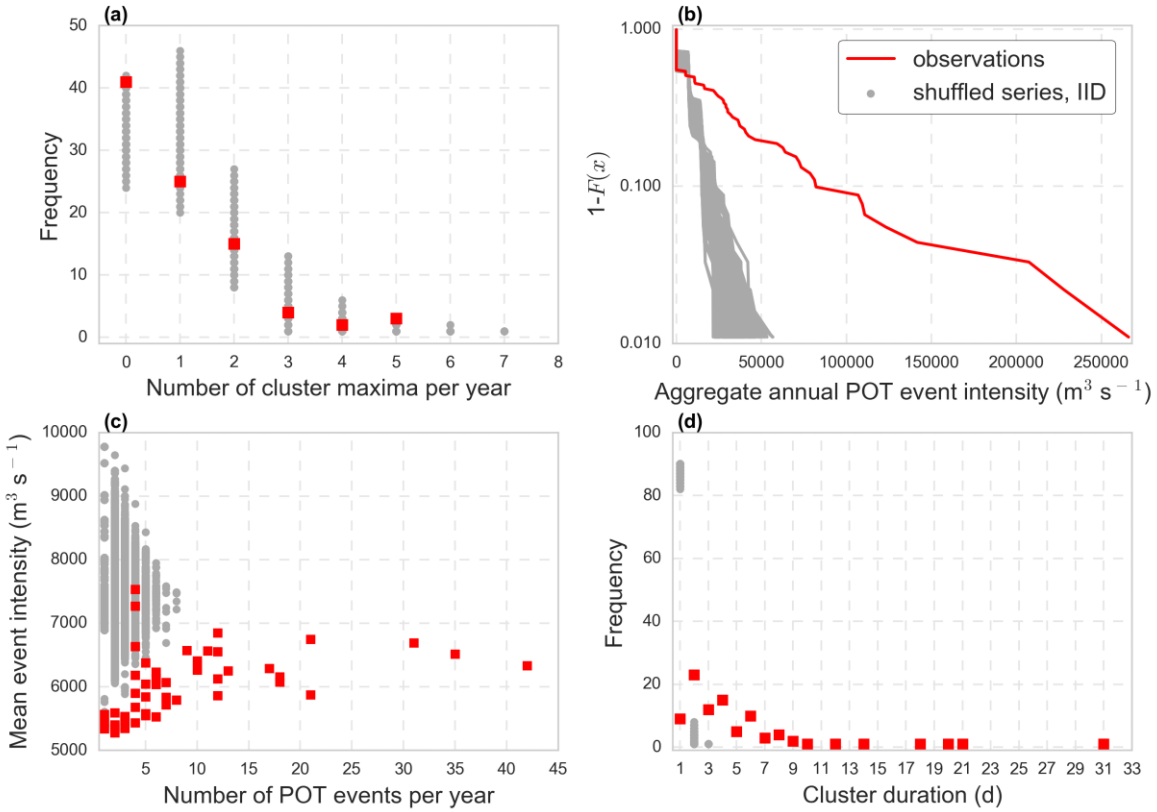


Figure 7.9 Plots of the following functionals of POT events from the original Po streamflow series (90 years) and 1000 shuffled timeseries (IID): (a) frequency distribution of the number of cluster maxima per year, (b) exceedance probability of the aggregate annual POT event intensity, (c) number of POT events per year versus their mean annual intensity, and (d) frequency distribution of cluster duration.

Next, we test the extent to which the observed extremal patterns can be reproduced by variants of HK-type models accounting for the process second-order (climacogram-based) multi-scaling properties (including joint moments) and the distributional properties up to the 4th moment (Dimitriadis and Koutsoyiannis, 2018; Koutsoyiannis, 2016). First, it is interesting to note the different scaling regimes between the two series (Fig.7.10). The Bologna rainfall series exhibits two distinct scaling regimes (two slopes in the climacogram); a weaker dependence structure in short-time scales and an intensified one in time-scales greater than 3 years. On the contrary, the climacogram of the Po streamflow series exhibits a strong curvature in short timescales (less than three months), indicative of short-term dependence, while approaches a HK behaviour at greater scales. The latter is similar to the first scaling regime of the rainfall process. Since these series belong to the same hydrological region, it could be postulated that the observed scaling patterns reflect the propagation of dependence from the rainfall to the streamflow process. At shorter time scales, the rainfall dependence dynamics are intensified through the catchment storage mechanism appearing as 'river memory' (see also Chapter 5), whereas in greater time-scales, the streamflow scaling behaviour approaches the first scaling regime of the rainfall process. Yet the Po record is not long enough to assess whether this assumption is supported at greater scales as well.

The fitted climacogram models are also shown in Fig. 7.10. The streamflow second-order scaling is captured by a GHK model, with three parameters, $H = 0.61$, $q = 35.149$, $\lambda = 1.027$, while the rainfall process is captured with a FHK model having an additional parameter to account for the change of the scaling regime, $H = 0.92$, $q = 1.12$, $\lambda = 0.013$, $\lambda' = 646.2$. A complete presentation of other modelling options is provided in Koutsoyiannis (2020b).

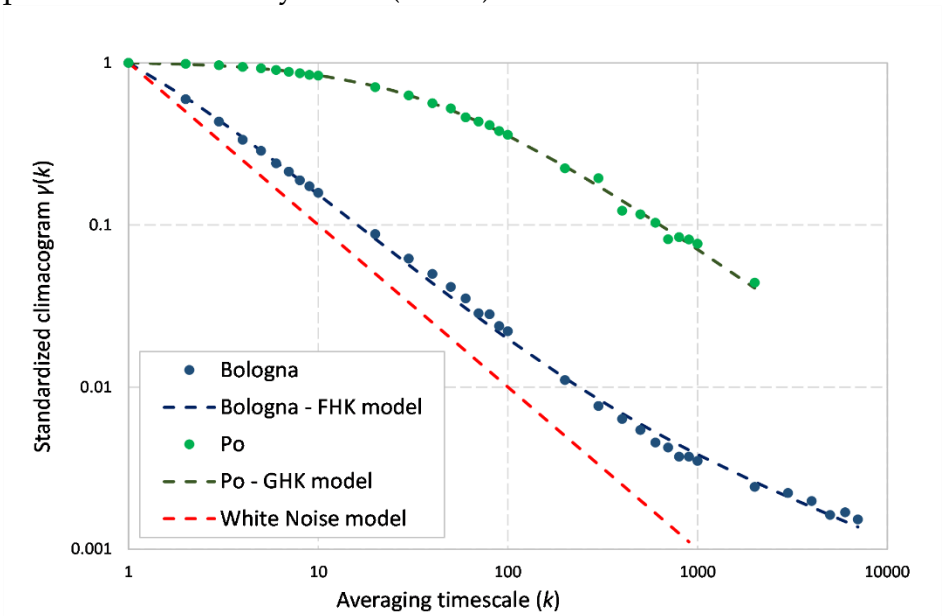


Figure 7.10 Standardized climacograms of the Bologna rainfall and Po streamflow along with the fitted climacogram models and the theoretical climacogram of a White Noise process.

Results on the reproduction of the observed extremal properties by the synthetic series are shown in Fig. 7.11-7.12. Overall, results suggest that both models, albeit formally calibrated only on the first four moments and the second-order scaling behaviour, prove successful in capturing observed patterns of the extremal behaviour as well. In Fig. 7.12 it is seen that the GHK model satisfactorily reproduces the positive association between the number of POT events per year and their intensity. A few discrepancies are yet observed. In the Bologna series, the cluster duration is underestimated, while for both the rainfall and the streamflow series, the exceedance probabilities of the aggregate annual intensities are also slightly underestimated in the tail region of the synthetic series. The former suggests that a stronger short-term dependence structure could be more appropriate for the Bologna series, while the latter could be remedied by the inclusion of higher-order moments in the model's calibration scheme. In any case, the above results suggest that it is possible to capture extremal patterns by preserving only essential properties of the parent process, i.e. it may suffice to model the second-order scaling behaviour along with a certain number of moments.

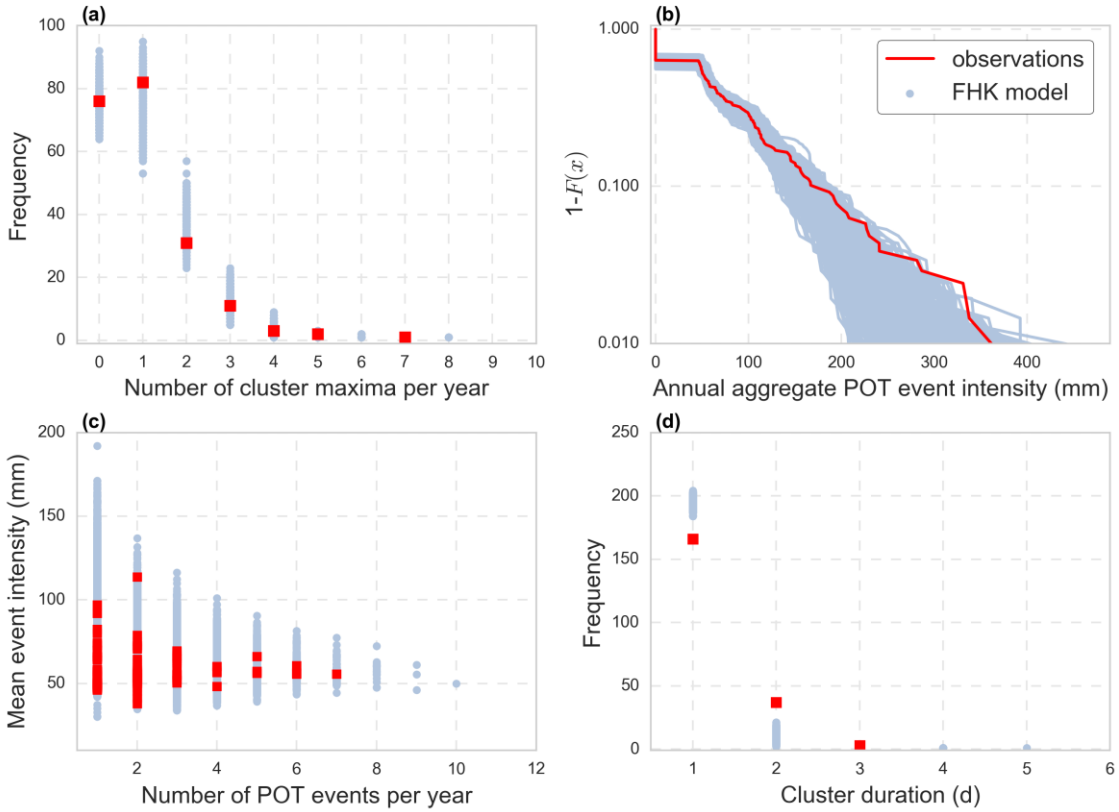


Figure 7.11 Plots of the following functionals of POT events from the original Bologna rainfall series (206 years) and 1000 synthetic timeseries from FHK model: (a) frequency distribution of the number of cluster maxima per year, (b) exceedance probability distribution of the aggregate annual POT event intensity, (c) number of POT events per year versus their mean annual intensity, and (d) frequency distribution of cluster duration.

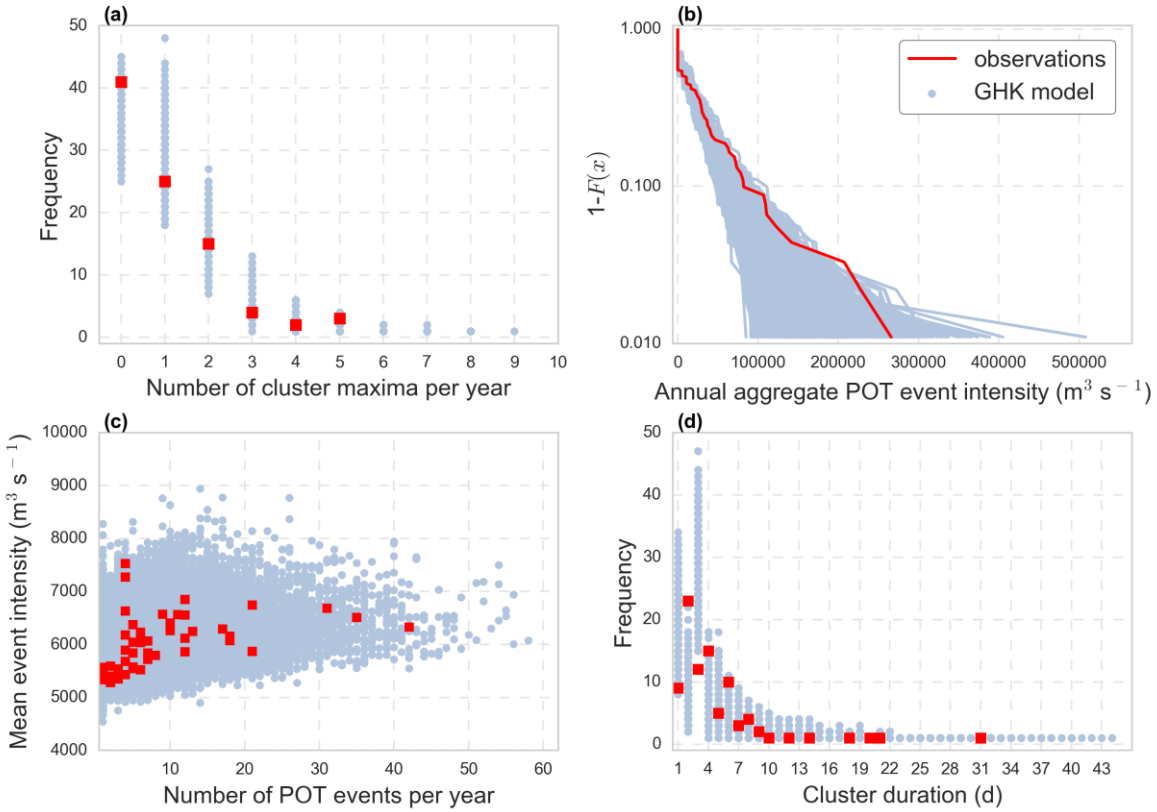


Figure 7.12 Plots of the following functionals of POT events from the original Po streamflow series (90 years) and 1000 synthetic series from a GHK model: (a) frequency distribution of the number of cluster maxima per year, (b) exceedance probability distribution of the aggregate annual POT event intensity, (c) number of POT events per year versus their mean annual intensity, and (d) frequency distribution of cluster duration.

7.5 Conclusions and outlook

In this Chapter, the distributional properties of extremes from persistent processes are studied with a twofold goal: (a) to test the sub-asymptotic appropriateness of extreme value theory and conceptualize the effects of persistence on extremal properties, and (b) assess the reproduction of the latter by HK-type modelling of the parent process.

Regarding the first goal, it is shown that the sub-asymptotic performance of the Fréchet distribution for the extremes of persistent process is mainly a matter of the fitting algorithm. In this respect, the maximum likelihood method yields poor fitting results, due to its reliance on the independence assumption, but on the contrary, results from the weighted-least-squares method are very good even for strong persistence and small block lengths. On the other hand, it is questionable whether the extremal index theory (Section 2.3.2) is well-suited for persistent processes. In fact, the extremal index theory is developed for processes producing only short-term clustering, with limited long-range dependence, and therefore, it is reasonable for results to not hold for cases of persistence. As a matter of fact, contrary to what the theory suggests for dependent processes, the scale parameter, related to the variability of extremes, increases as a result of the increased variability of the persistent process compared to an IID process. Still in agreement to the theory, dependence decreases the location parameter of the block maxima distribution, yet it does not affect the shape parameter, as theoretically expected.

In terms of block maxima analysis, dependence manifests itself by 'hiding' the extreme-generating potential of the process. Compared to an IID process, in an HK process, it is much more likely to observe maxima of lower magnitude at a given scale. On the contrary, the POT analysis exposes both the short-term clustering of extremes as well as the long-term clustering studied in Chapter 6. In this Chapter, the focus is placed on the extremal properties on the annual scale. It is found that POT for persistent processes have a bilateral character compared to IID processes; absence of POT events in a year is more likely, yet in the case occurrence of extremes is triggered, a higher cluster duration and greater intensity thereof should be expected.

Regarding flood risk estimation, a direct consequence of the increase in duration is the increase in the period of time an area is inundated (Dimitriadis and Koutsoyiannis, 2020). The increase in the aggregate intensity of extreme events in a given scale also increases the collective risk, which is relevant for insurance practices against natural catastrophes (Serinaldi and Kilsby, 2016b; Goulianou et al., 2019; Manolis et al., 2020; Papoulakos et al., 2020). On a higher level, it is argued that the understanding of temporal risk dynamics is important to any field concerned with risk planning and preparedness for natural disasters. Furthermore, as estimation uncertainty is dominant for persistent and heavy-tailed processes, a priori awareness of their dynamics is essential in order to cautiously interpret empirical evidence and avoid a false perception of the true extremal properties. This makes the record length issue for characterizing extremes even more important for dependent processes (Koutsoyiannis and Montanari, 2007).

The extremal properties of long-term real-world series showed clear deviations from the IID case and consistency with HK-type of models reproducing the first four moments and the second-order scaling behaviour. Interestingly, the dependence patterns of the two series, rainfall at Bologna and streamflow in the Po River, differed substantially despite belonging to the same hydrological region. The rainfall series characterized by long-term persistence, but weak short-term dependence, showed less short-term clustering at the annual scale compared to the streamflow process, for which the intra-annual clustering patterns were very pronounced. The fitted HK models successfully captured the variability of the observed extremal patterns, even though seasonality was not explicitly modelled, but only partially retained through the second-order scaling.

Overall, the HK framework constitutes both an advantageous framework in terms of explaining the temporal variability of extremes and a promising direction in terms of joint modelling of lower and higher-order properties. More research is required with respect to the latter as improvements are expected from inclusion of higher moments in the modelling scheme, using *k*-moments estimation (Koutsoyiannis, 2019c).

8. PROJECTIONS OF FUTURE RAINFALL EXTREMES

Non-stationarity approaches have been increasingly popular in hydrology, reflecting scientific concerns regarding intensification of the water cycle due to global warming. A considerable share of relevant studies is dominated by the practice of identifying linear trends in data through in-sample analysis. In this Chapter, the problem of trend identification is reframed using the out-of-sample predictive performance of trends as the reference point for model selection. A systematic methodological framework is devised in which linear trends are compared to simpler mean models, based on their performance in predicting climatic-scale (30-year) annual rainfall indices, i.e. maxima, totals, wet-day average and probability dry, from long-term daily records. Analysis of empirical records spanning over 150 years of daily data suggests that future long-term variability is better captured using local mean models rather than trends. In line with theoretical findings for persistent processes, it is shown that prediction-wise, simple is preferable to trendy.

8.1 Introduction

> *"A trend is a trend is a trend / But the question is, will it bend?*
> *Will it alter its course / Through some unforeseen force*
> *And come to a premature end?"*
> (Sir Alec Cairncross, 1969, signing as "*Stein Age Forecaster*")

In the past decades there has been a plethora of trend analyses in rainfall studies (Bunting et al., 1976; Haylock and Nicholls, 2000, 2000; Rotstayn and Lohmann, 2002; Modarres and da Silva, 2007; Ntegeka and Willems, 2008; Kumar et al., 2010), and it could be argued that relevant studies are still on the rise (e.g. Biasutti, 2019; Degefu et al., 2019; Folton et al., 2019; Khan et al., 2019; Papalexiou and Montanari, 2019; Quadros et al., 2019; Rahimi and Fatemi, 2019). A quantitative analysis of the relevant literature is provided in Section 8.3.1. This boom of trend studies has had various scopes, most of which are related to global warming assessment (IPCC, 2013). These include historic climate variability quantification, attribution to deterministic drivers, projections to the future and impact assessments (e.g. Kumar et al., 2010; Parmesan and Yohe, 2003; Biasutti, 2013; Rotstayn and Lohmann, 2002). Arguably what is common in the majority of trend studies, even when not explicitly stated, is the expectation for a monotonically changing future, which as a result, has initiated a growing discourse on the appropriate modelling approach.

In climatology and hydrology, there has been an ongoing debate between stationary vs nonstationary methods, with the former representing a well-established hydrological practice (Montanari and Koutsoyiannis, 2014; Koutsoyiannis and Montanari, 2015) and the latter reflecting recent attempts of the scientific community to find a new way to respond to change and uncertainty under the anthropogenic climate change scenario (Milly et al., 2008; Craig, 2010; Milly et al., 2015). Yet deterministic trend modelling has been examined —and mostly criticized, on different grounds, namely with respect to empirical evidence (McKitrick and Christy, 2019; Cohn and Lins, 2005), theoretical consistency (Koutsoyiannis and Montanari, 2015),

modelling efficiency (Montanari and Koutsoyiannis, 2014), and meaningfulness of the results (Serinaldi et al., 2018, 2020). It has also been argued that the concepts of change and uncertainty are already well-represented within the stationarity framework (Koutsoyiannis and Montanari, 2007; Serinaldi and Kilsby, 2018b). In this research, we examine the trend modelling framework from a new perspective, through the evaluation of its out-of-sample modelling qualities, namely, its predictive powers for a given record.

For this purpose, we introduce a validation framework for the evaluation of the results, adding simpler, mean models in the pool of candidates, and basing the reasoning of model selection on the statistical out-of-sample performance of the models. While split-sample techniques (Klemeš, 1986) and multi-model approaches (Georgakakos et al., 2004; Duan et al., 2007) are certainly not new in hydrology, they are usually disregarded as concepts in the field of trend modelling, where the research question typically revolves around explanatory performance, mostly by means of in-sample measures, as hypothesis testing (Shmueli, 2010). In this work, we extend the simple split-sample validation by introducing a moving window calibration and validation approach that progressively scans each record by sliding windows of climatic-length, i.e. 30 years according to the common climate definition (IPCC, 2013). In this manner, we obtain a sample of estimates of the models' predictive performance, instead of a single value.

By shifting the focus to the predictive modelling of linear trend, this analysis seeks to answer the following key questions: (a) how well are the rainfall statistics of the most recent climatic period predicted by the linear trend calibrated to the prior 30-year period? and (b) how do the statistics of the predictive performance of linear trends compare to the ones derived from application of simple mean models?

The first question is driven by the omnipresent scientific concerns regarding intensification of extremes due to global warming during the last decades (e.g. Houghton et al., 1991; Parmesan and Yohe, 2003; Oreskes, 2004; Solomon et al., 2007; McCarl et al., 2008; Moss et al., 2010; Craig, 2010; Pachauri et al., 2014; Kellogg, 2019). According to the fifth (latest) IPCC assessment (IPCC, 2013), the expected intensification mechanism suggests a 6%–7% increase of the global water vapour per °C of warming, followed by a 1% to 3% increase in global mean precipitation. Recently, the physical assumptions behind these estimates have been questioned and revisited in light of global datasets (Koutsoyiannis, 2020a), while the evaluation of hydrological impacts from increased greenhouse emissions remains an open research subject with often conflicting evidence (e.g. Hirsch and Ryberg, 2012; Mallakpour and Villarini, 2015; Blöschl et al., 2019). Therefore, the first examination of predictability is consciously biased in favour of a model capturing the variability of the most recent period of data.

The second question introduces the abovementioned methodological framework for validating model predictions, which is applied to the empirical long-term rainfall records as well as to synthetic series produced in order to mimic the natural long-term variability of the rainfall process. A discussion on the relevance of the framework in light of potential deterministic changes is also provided.

8.2 Dataset

Our dataset is an update of the previous long-term dataset explored in Iliopoulou et al. (2018a) of long rainfall records surpassing 150 years of daily values. It includes the 60 longest available daily rainfall records collected from global datasets, i.e. the Global Historical Climatology Network Daily database (Menne et al., 2012), the European Climate Assessment and Dataset (Klein Tank et al., 2002), as well as third parties listed in the Appendix A (Table A1). The geographic location of the rain gauges is shown in Figure 6.1. The length of the timeseries provides rare insights into long-term rainfall variability and enables the statistical evaluation of the predictive performance of linear trends from multiple time windows.

8.3 Overview of literature approaches

8.3.1 A quantitative review on rainfall trends

The aim of this literature review is to evaluate the academic interest in trends of rainfall variables by means of a quantitative analysis of research papers appearing in Google Scholar. We base this analysis on the quantification of the occurrence of associated words in Google Scholar using Python code developed by Strobel (2018), omitting results related to citations and patents. This analysis was performed on 21/10/2019 and in order to refer to full calendar years it contains results published till the end of 2018.

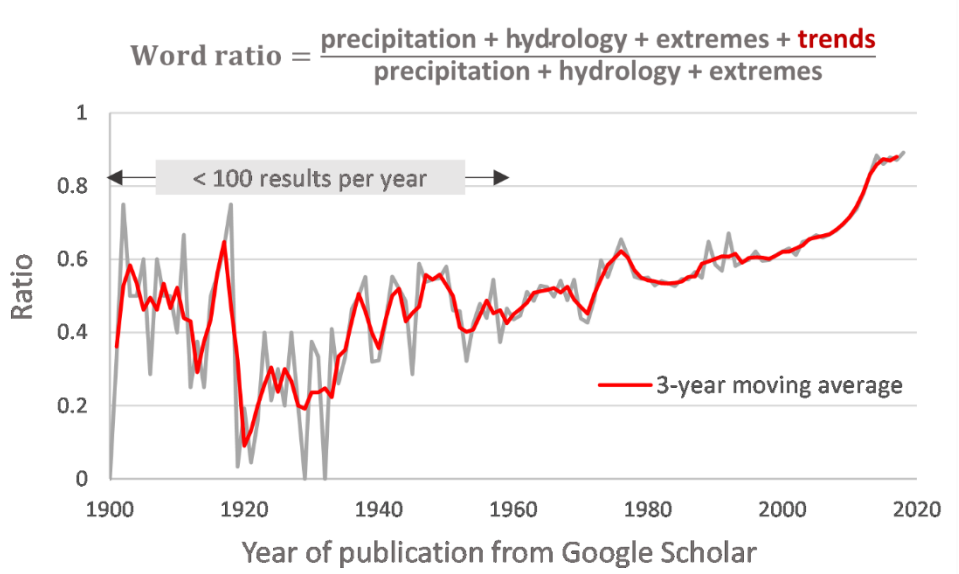


Figure 8.1 Temporal evolution along with three-year moving average of the ratio of the occurrence of the word 'trends' in Scholar items containing the words 'precipitation', 'hydrology' and 'extremes'.

In Fig. 8.1, we show the temporal evolution of the ratio of appearance of the word 'trends' in items also containing the complete list of words ['precipitation', 'hydrology', 'extremes']. Results have been randomly varying from the beginning till the mid 20th century, when there were less than 100 results per year fulfilling the criteria of containing the list in the denominator of the ratio. It can be seen though that approximately from the 1960 and later on there has been an increasing trend in relevant publications containing the word 'trends,' reaching 89% in 2018. Obviously,

results belonging to a different context than the one assumed might have been calculated as well but we assume their effect to be analogous both in the nominator and the denominator of the ratio, thus not significantly affecting the conclusion.

To further refine our search to more technical papers explicitly referring to rainfall trends we define the following search terms. Word combination A is the full list ['precipitation|rainfall trends', 'precipitation|rainfall data|records'], where the symbol | refers to 'or', and word combinations inside '' should be found together, i.e. one possible combination is the list ['precipitation trends', 'rainfall data']. Word combination B is an extension of word combination A that also includes the word 'projections', while word combination C is an extension of word combination A also including the word sequence 'linear trend|trends|model|regression'. The absolute numbers of the results are shown in Fig. 8.2a, while in Fig.8.2b we show their relative ratio. Expectedly, the total number of studies containing rainfall trends are rising, however this is not surprising in terms of absolute numbers, considering the increasing availability of papers in Scholar over the years. However, the use of the word 'projections' appears to be increasing in relative terms as well. The relative use of word combination C, related to the linear trend, has slightly increased too over the years, stabilizing over the past 5-year period to approximately half of the related publications (Fig.8.2b).

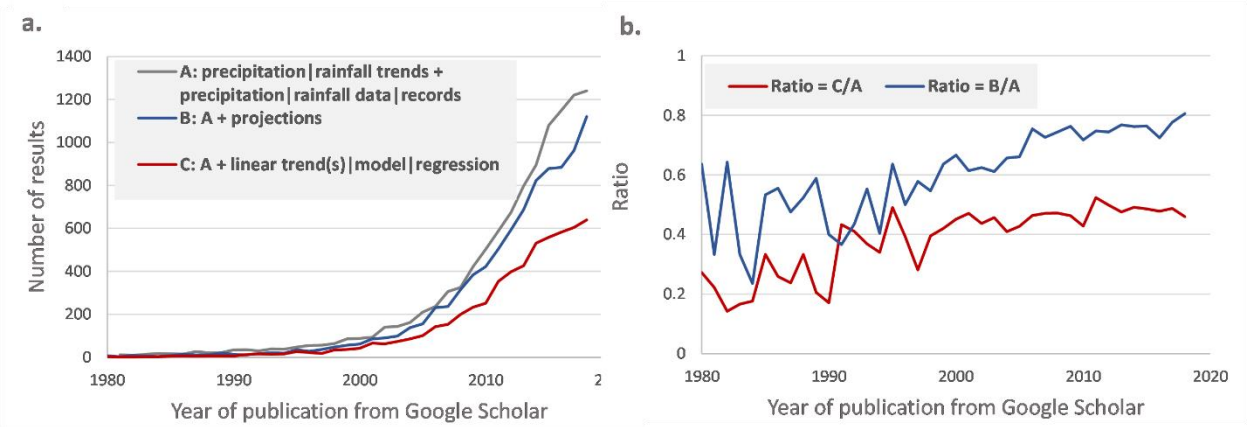


Figure 8.2 (a) Temporal evolution of the occurrence of the word combinations A, B and C and their relative ratio (b).

As a final refinement, we consider words appearing only in the title of papers, which should limit the results to strictly related papers. Results are shown in Fig. 8.3. The standard term that is contained in every result is 'rainfall|precipitation' followed by the appearance, anywhere in the title, of the single terms, trends|trend, variability, change|changes, and non-stationary|non-stationarity|nonstationary|nonstationarity. Note that we consider also plural terms where applicable, as well as possible differences in spelling, while this time, we do not require words to be found in a specific order as in the previous in-text search (for instance, it could be "**trends** in **rainfall**..." or "**rainfall trends** in the.."). We do not compute ratios over the items containing in their title the words 'rainfall|precipitation' because these terms alone are too generic, and can be found in a variety of studies, a significant part of which are only loosely related to hydrology (e.g. physics, chemistry, radar technologies etc.). Instead, to provide a more relevant reference point for comparison, we use two words semantically 'uncharged' with the trend concept, which are however widely used in

combination with the standard terms, namely the words 'model' and 'distribution' (e.g. "a **rainfall model**..." or "the **distribution** of the ... **precipitation**").

Apparently, the conceptually more inclusive terms 'changes' and 'variability' are ranking first in the related search terms, with the explicit use of the word 'trend(s)' ranking third, yielding consistently over the last ten years above 200 results per year (288 in 2018, as per results appearing on Google Scholar on 21/10/2019). Terms related to non-stationarity are slowly rising over the past ten years (39 in-title results in 2018), while being close to zero before 2000. It is interesting to note the evolution of the use of terms explicitly associated with the temporal properties of rainfall compared to the terms more related to marginal properties ('distribution'), or being more of a general use, perhaps implying both properties ('model'). The mere use of the word 'trend(s)' has exceeded the use of an all-times classic word for rainfall, i.e. distribution, which clearly shows a certain shift in academic interest. Likewise, the ever higher-scoring word 'model' has been outnumbered in the past three years by the word 'change(s)'.

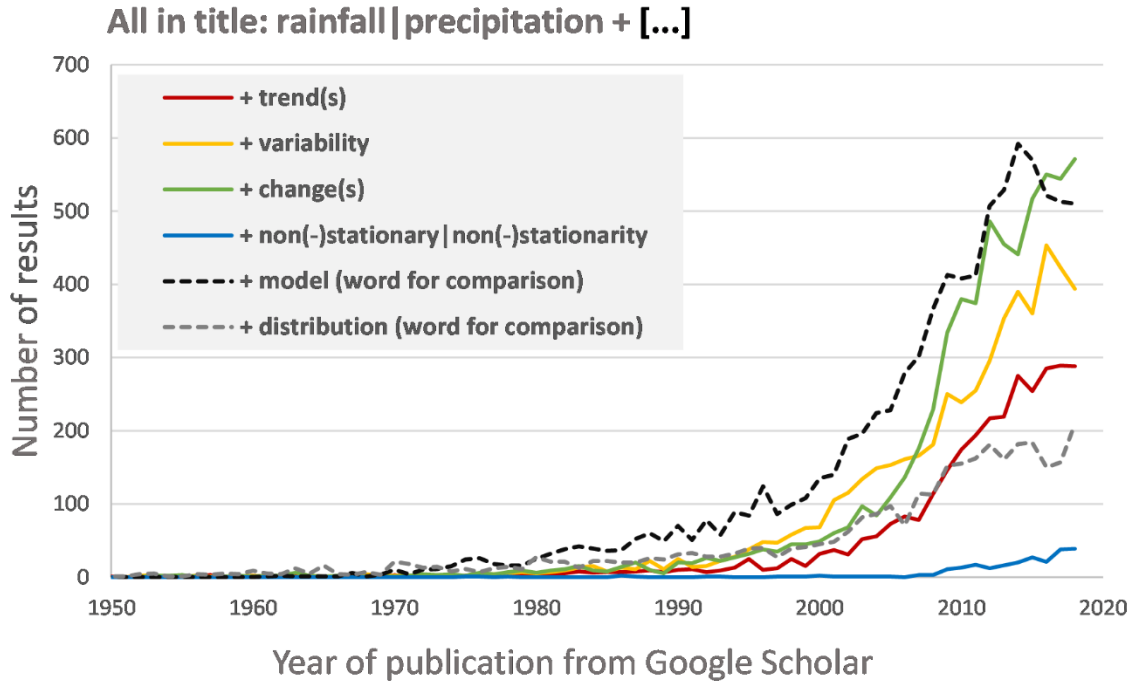


Figure 8.3 Temporal evolution of the occurrence of the word combinations in titles of Scholar items.

In conjunction, these results suggest that over the last two decades, there has been a rising scientific interest in the temporal properties of rainfall and their future evolution, with 'trends' taking up a considerable share of this emerging focus.

8.3.2 From explanatory trends to out-of-sample performance

It is well-known that studying the explanatory power of trends in hydroclimatic data is a very active research field, as confirmed by the above literature analysis. Before discussing literature modelling strategies for trends, it is imperative to define the meaning of a trend per se. Although 'trends' are frequently used as a synonym of temporal 'changes' (Fig. 8.3 provides a quantitative analysis on the use of both words) and their notion has sometimes been extended to encompass stochastic stationary

models (Fatichi et al., 2009; Chandler and Scott, 2011), the general idea behind the trend concept, is that the expected value of a response variable $\underline{y}$ is specified as a deterministic function of time t, $\mathrm{E}\left[\underline{y}\right] = f(t)$. The function f may take different forms —the linear model being only the first one adopted, and the most widely used. Indeed, this definition of a trend can be traced back to the development of the field of econometrics in the early 20th century, when 'secular' trends, meaning long-term trends, were deemed to be a component of financial timeseries, along with seasonal variation, cycles and residual elements (Persons, 1922; Mitchell, 1930). Decomposition of a timeseries into components, one of them being a trend, continued to dominate the econometrics literature, although even at early times certain critiques were raised (Slutsky, 1927).

The most established technique to evaluate fitted trends is statistical hypothesis testing, i.e. a statistical inference technique that estimates the probability of an outcome as far from what is expected as the observed under the assumption that the null hypothesis is true (Gauch, 2003). The latter is known as the p-value and is compared to predefined significance levels, in order to reject or not the null hypothesis. This is a scientific method for model evaluation, which has been in part misused. For instance, its misuse in hydrology has been showcased by seminal studies (e.g. Cohn and Lins, 2005; Koutsoyiannis and Montanari, 2007; Serinaldi et al., 2018) which have established the fact that for hydrological, non IID data the null hypothesis, which tacitly contains independence, is a priori wrong, and its rejection, if correctly interpreted, should point out to the wrong independence assumption. Still, the common practice has been to misinterpret outcomes in favour of trends. Part of the statistician community argues against the concept of significance testing (Nuzzo, 2014; Wasserstein and Lazar, 2016; Amrhein and Greenland, 2018; Trafimow et al., 2018; Wasserstein et al., 2019), with the main critique summarized in the statement of the American Statistical Association that "*the widespread use of 'statistical significance' (generally interpreted as '$p \leq 0.05$') as a license for making a claim of a scientific finding (or implied truth) leads to considerable distortion of the scientific process*" (Wasserstein and Lazar, 2016). Other inference techniques for assessing the plausibility of changes under an a priori assumed model are also used, most notably change point analysis (Hinkley, 1970), which attempts to identify points of abrupt changes in the data. This approach too, is very sensitive on a priori hypotheses about the expected degree of variability in the data (a brief discussion on the issue in provided in Chandler and Scott, 2011).

With a stronger focus on modelling power rather than confirmatory analysis, model selection criteria have been developed arising from Akaike's work (Akaike, 1969). Akaike has contributed to the introduction of information theory into model selection criteria (Akaike, 1974) which are now established worldwide in model inference (Anderson and Burnham, 2004) and are increasingly adopted in hydrology as well (e.g. Ye et al., 2008; Laio et al., 2009; Iliopoulou et al., 2018a). Information criteria are useful in that they try to achieve a better out-of-sample performance by prompting for parsimony when fitting the model to the calibration set. There is a vast literature on the asymptotic equivalence of information criteria and out-of-sample prediction measures under specific conditions (Stone, 1977; Shibata, 1980; Wei, 1992; Inoue and Kilian, 2006), which typically though imply large record lengths.

A discourse regarding the relative powers of the abovementioned 'in-sample' measures compared to the assessment of predictive or out-of-sample performance is active in numerous scientific fields (Breiman, 2001; Stein, 2002; Inoue and Kilian, 2006; Yarkoni and Westfall, 2017; Shmueli, 2010), while in fact, it has been argued that the distinction between the two approaches might only arise due to the different objectives of each study (Gauch, 2003; Inoue and Kilian, 2005). Obviously, predictive modelling dominates in operational fields concerned with short-term prediction, as numerical weather prediction (Lorenc, 1986), and in such domains, it is widely acknowledged that the model yielding the best predictions, in non-stochastic terms, is not necessarily the 'true' one (Shmueli, 2010).

The premise of this work is that while explanatory performance of trends has been thoroughly explored in hydrological studies (e.g. Chandler and Scott (2011) provide a comprehensive review on the matter), much less attention has been given to the predictive performance of trend modelling. A simple explanation might lie in the fact that in many environmental studies trends have been employed as descriptors of changes or causal effects, and less as models for predictions, in spite of the fact that they strongly communicate expectations for the future by suggesting causal mechanisms (e.g. Fig. A2 on the combined use of the word 'trends' and 'projections'). The second reason could be related to the scarcity of long-term environmental data for out-of-sample validation. Therefore, our aim is to assess the relevance of long-term trend modelling in terms of point prediction, not examining elements of stochastic prediction and categorically, not engaging in the identification of a 'true' model for the data. We deem that this shift in point-of-view may provide contrasting insights to current literature with respect to the relevance of trends for operational applications.

8.4 Methodology

8.4.1 Out-of-sample validation schemes

Cross-validation techniques are a systematic way to assess predictive power (Stone, 1974; Simonoff, 2012). The procedure typically entails multiple runs of validation schemes on random partitions of the original dataset and summarizes the model skill from the sample of all validation scores. Standard cross-validation is not straightforward to apply for timeseries data where the order of the data must be respected. Instead the use of a 'holdout' set for validation is frequently applied, e.g. in hydrology this is done by reserving some data for validation, while the rest are used for calibration (Klemeš, 1986). We consider an alternative approach respecting the data order, by performing calibration and validation in moving-window partitions of the original dataset, that constantly shift forward in time till the end of the record is reached. This approach is known as 'walk-forward' analysis in the field of econometrics (Kirkpatrick II and Dahlquist, 2010), and it is advantageous in that instead of a single measure of out-of-sample performance obtained by the 'split-sample' approach, a sample of values is obtained, which can be statistically analysed. Further, it compensates for hindsight bias providing realistic estimates of historical predictability of changes by a given model. The statistics of a model's past performance can be considered a proxy of its future performance.

8.4.2 Static calibration and validation

We apply this type of analysis to the rainfall records by formulating two distinct calibration-validation schemes, which are illustrated in Fig. 8.4. In the first scheme (Fig.8.4a), we evaluate the models' performance in capturing the variability of the recent 30-year period of each station based on calibration on the prior 30-year period. By this 'static validation' scheme we intend to evaluate whether extremes have changed in a consistent manner in the second half of the 20th century, as they are commonly assumed. We also examine the performance of the models in backward validation, i.e. in predicting observations occurring before the calibration period (Fig. 8.4a). In order to maximize the exploitation of the length of each record, we apply this evaluation to the most recent period of each station, even if the final dates of all records do not coincide. We favour separate treatment of each station, since in this case our focus is placed on the operational exploitation of records for predictive purposes and less on a summary of the results for a specific time period. However, the majority of the records span the whole 20th century, and extend beyond, with a few exceptions that are mentioned in Table A1. In a second examination, we directly evaluate changes in the predictive performance of each model throughout the past 110 years up to 2009. Specifically, we compare the prediction errors of each model for the following climatic periods: 1900–1929 (calibration period 1870–1899), 1930–1959 (calibration period 1900–1929), 1960–1989 (calibration period 1930–1959), and 1980–2009 (calibration period 1950–1979). The end year (2009) of the last period (overlapping with the previous one by 10 years) is selected in order to maximize the number of stations having predictions for all four periods. This results to 52 stations for the AM and 51 for the AT, WDAV and PD indices.

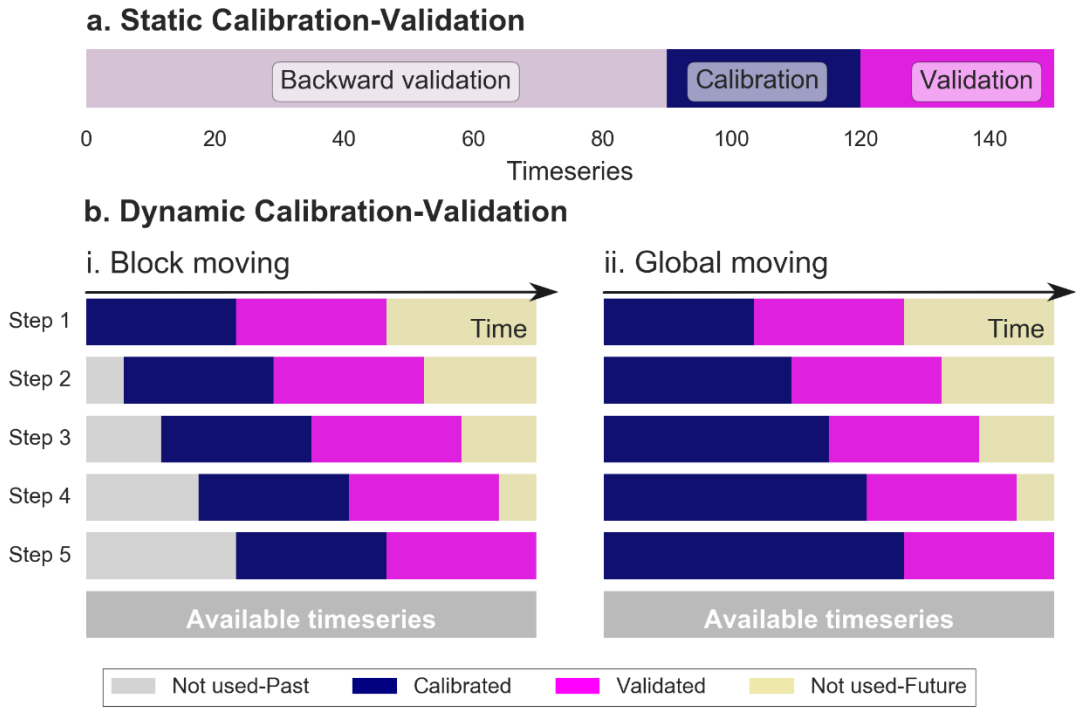


Figure 8.4 Explanatory sketch showing the two calibration and validation schemes (a. Static and b. Dynamic) for an example station.

8.4.3 Dynamic calibration and validation

The second scheme (Fig.8.4b) focuses on the historical performance of the models by the 'dynamic' (else, 'walk-forward') validation scheme introduced before. It assumes a hypothetical observer moving in time and making predictions for the future 30-year period updating the models as access to new information progressively becomes available. We formulate two different schemes for making these predictions. In the first, which we call block-moving calibration and validation, the models are calibrated on 30-year periods and validated by the next 'unobserved' 30 years, and this procedure is repeated by rolling the calibration and validation origin in time (Fig.8.4bi). New information is gradually taking the place of the past information, which is discarded by the 30-year sliding windows. The start of the first moving-window coincides with the start of each station, while the start of the last calibration moving-window is 59 years prior to the end of the station, so that 30 years of validation data remain available. This last validation window is the recent 30-year window that is exploited for validation in the static scheme (Fig. 8.4a). The second scheme of the dynamic calibration-validation, which we call global-moving, validates the models using sliding 30-year periods, exactly as in the prior scheme, but calibrates the models on the whole available record, that is known at each time step to the observer. Therefore, the origin of the calibration window remains stable, but the window gradually extends in length as more data are assimilated into the model, while no data are discarded (Fig.8.4bii). This scheme explores the potential of employing all available information to make a prediction for the future. Since the validation periods are the same in both schemes, results between the two can be directly compared.

For the evaluation of the candidate models we estimate the Root Mean Square Error, a standard and established metric of goodness of fit (Sharma et al., 2019). The RMSE is defined as the square root of the mean square error of the predicted values $\hat{x}_i$ with respect to the observed x_i:

$$\mathrm{RMSE} = \left(\frac{\sum_{i=1}^{n} (\hat{x}_i - x_i)^2}{n} \right)^{1/2} \tag{48}$$

where n is the length of the data. We present the sample RMSE distribution of the models for each station and we summarize the results by computing the average RMSE for each station and its standard deviation. For the longest uninterrupted record of the station, we present a comprehensive analysis including the temporal evolution of the errors.

8.4.4 Predictive models

Let $\underline{x}_i$ be a stochastic process in discrete time i, i.e. a collection of random variables $\underline{x}_i$, and $x := (x_1, \dots, x_n)$ a single realization (observation) of the latter, i.e. a timeseries. We assume that in time $i \leq n$ the hypothetical observer makes a forecast based on a subset of the historical information. Namely from the entire available information that we have (the observed series $(x_1, \dots, x_n)$) we assume that the hypothetical observer knows only the subseries $x = (x_1, \dots, x_i)$.

To predict the unobserved periods, past or future, we employ two model structures. The first is the typical linear trend model, encompassing two parameters, a slope b and an intercept a, whose mean μ is a deterministic linear function of time t:

$$\mu(t) = a + bt \tag{49}$$

The trend model is fitted via least-squares regression. Robust regression techniques are also explored, namely median quantile regression (Koenker and Hallock, 2001) and the Theil-Sen slope estimation (Sen, 1968; Theil, 1992), but they did not yield better predictions, and hence, the least-squares approach, which is also more rigorous in theoretical terms (e.g. Papoulis, 1990), was retained. For details on the application and discussion of the results, the reader is referred to the analysis presented in Appendix B.3.

The second model considered is the mean model, including only one parameter, the mean of the calibration period, extrapolated to the unobserved periods:

$$\mu(t) = a \tag{50}$$

According to the followed calibration scheme, fitted to block-moving (local) 30 years or to all the known (global) period, the trend model is termed local trend (L-Trend) and global trend (G-Trend), respectively, and likewise, the mean model, is termed local mean (L-Mean) and global mean (G-Mean). In the local models, the period $[i - 59, i - 30]$ is used for calibration and the $[i - 29, i]$ for validation, while in the global models, the period $[1, i - 30]$ is used for calibration and the $[i - 29, i]$ period for validation as in the former scheme. We note that these two seemingly simplistic predictive models, i.e. the linear model fitted with least-squares and the local average, can be found in a variety of theoretical results in statistical sciences, for instance use of (temporally) local data constitutes a central concept in the *k*-nearest neighbours technique, as discussed in Hastie et al. (2005), as well as in local regression as discussed in Chandler and Scott (2011).

8.4.5 Selected indices of rainfall extremes and quality control

We examine four statistical indices of rainfall: annual maxima (AM), annual totals (AT), annual wet-day average rainfall (WDAV) and probability dry (PD) also computed at the annual scale. As wet, we consider any day with rainfall surpassing the threshold of 1 mm, while values below this threshold are counted as dry days taken into account for the PD estimation. We employ the following criteria for missing values. For the annual maxima we use a methodology proposed by Papalexiou and Koutsoyiannis (2013), according to which an annual maximum in a year with missing values is not accepted if (a) it belongs to the lowest 40% of the annual maxima values and (b) 30% or more of the observations for that year are missing. For the rest of the indices, we do not compute the yearly index in years with more than 15% of missing values. In general, most records have low percentages of missing values (Table A1), which in most cases are clustered in the beginning of the records. A few records have consecutive missing periods which might imply a change of instrumentation or relocation of the gauge. To avoid possible artefacts in trend estimation in static validation (in backward validation) that may arise from such cases, we analyse periods containing less than 5% of consecutive missing values of the yearly indices. For the

dynamic calibration and validation scheme, we fit the models only if there exist at least 27 valid indices in each of the 30-year periods of calibration and validation.

8.4.6 Predictability of climatic changes under natural variability

In order to understand the predictive performance of the considered models under typical conditions of natural variability, we run similar experiments with synthetic timeseries reproducing increasing degrees of persistence. We recall that persistence, also known as Hurst-Kolmogorov dynamics, is associated with enhanced natural variability at all scales (Koutsoyiannis, 2003), which in turn implies increased unpredictability at large time horizons, with some potential for predictability at short time steps due to the presence of temporal clustering (Dimitriadis et al., 2016). This provides a scientifically relevant comparison to the empirical data as rainfall series are known to exhibit mild to moderate degree of persistence (e.g. Iliopoulou et al., 2018b; Iliopoulou and Koutsoyiannis, 2019). Moreover, segments of persistent series resemble trends and can easily be misinterpreted as such (Cohn and Lins, 2005).

Therefore, we examine both the comparative predictive performance of the four models for persistent processes, where long-term changes are the rule (Serinaldi and Kilsby, 2018a), and the effect of available record length on the quality of the model predictions. The latter becomes relevant in the global-moving scheme, in which the calibration period varies in length.

8.5 Models' performance in static validation

Results from the performance of the local mean and local trend models on the last 30 years of each station, as well as on the years preceding the 30-year calibration, are shown in Figure 8.5 for all studied indices.

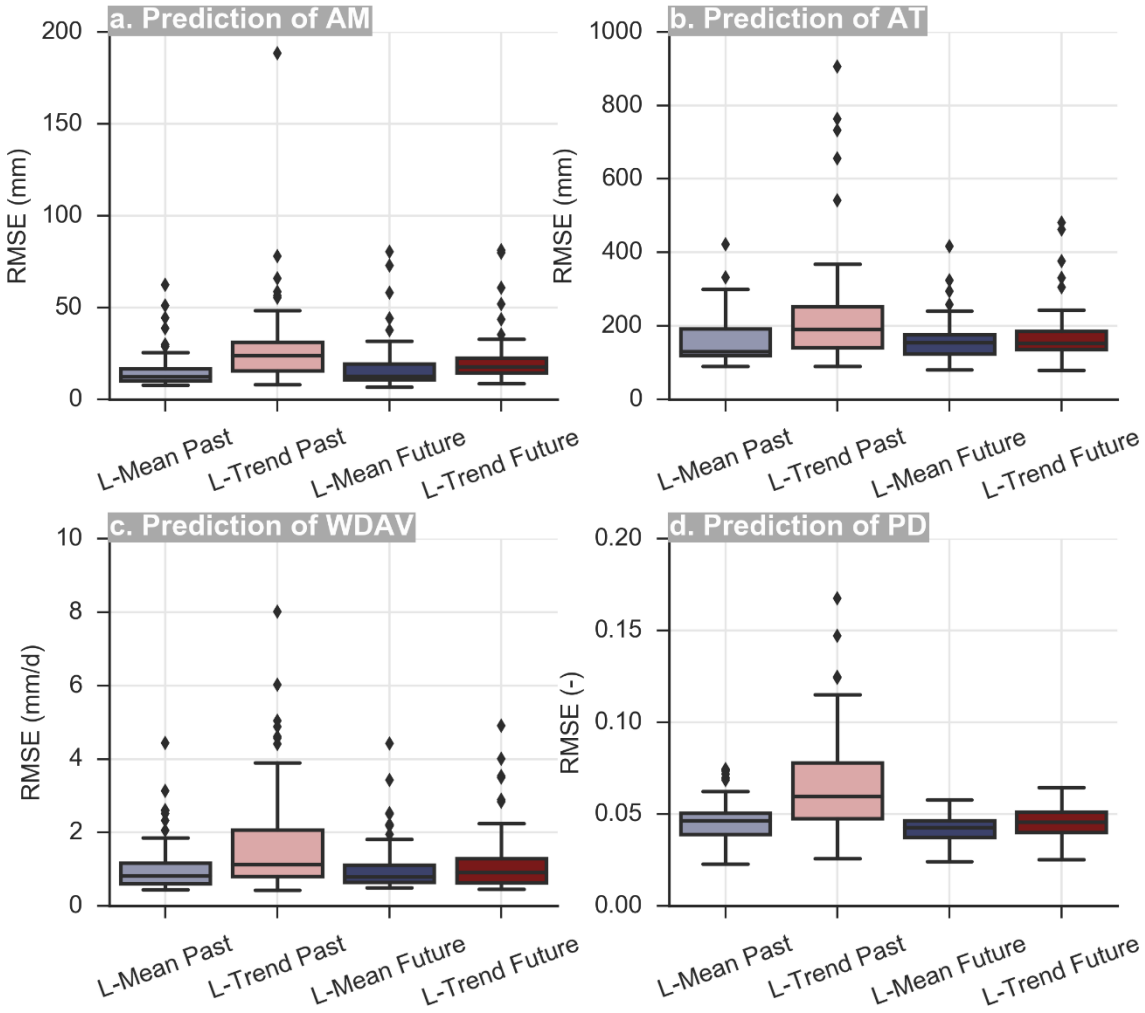


Figure 8.5 Boxplots of the RMSE distribution from the static validation application to all stations, for the local mean (L-Mean) and local trend (L-Trend) models, for all rainfall indices. The band inside the box reports the median of the distribution, the lower and upper ends of the box represent the 1st and 3rd quartiles, respectively, and the whiskers extend to the most extreme value within 1.5 IQR (interquartile range) from the box ends; outliers are plotted as points.

The local mean model performs on average better than the local trend model for all indices in capturing their most recent changes of extremes, while the performance of the local trend deteriorates considerably with respect to hindcasting the past. Interestingly, the larger discrepancies of the trends —both in future and past validation periods, are encountered in the annual maxima, followed by probability dry. In most of the opposite cases, of trends showing a better performance, the fitted slope is very mild, thus hardly differing from the local mean. A visual examination of the plots of the 60 long-term stations, provided in the Appendix B.3 (Fig. B.5-B.8), suggests a positive answer to the opening question, providing empirical evidence that climatic trends fluctuate and in fact, abruptly reverse.

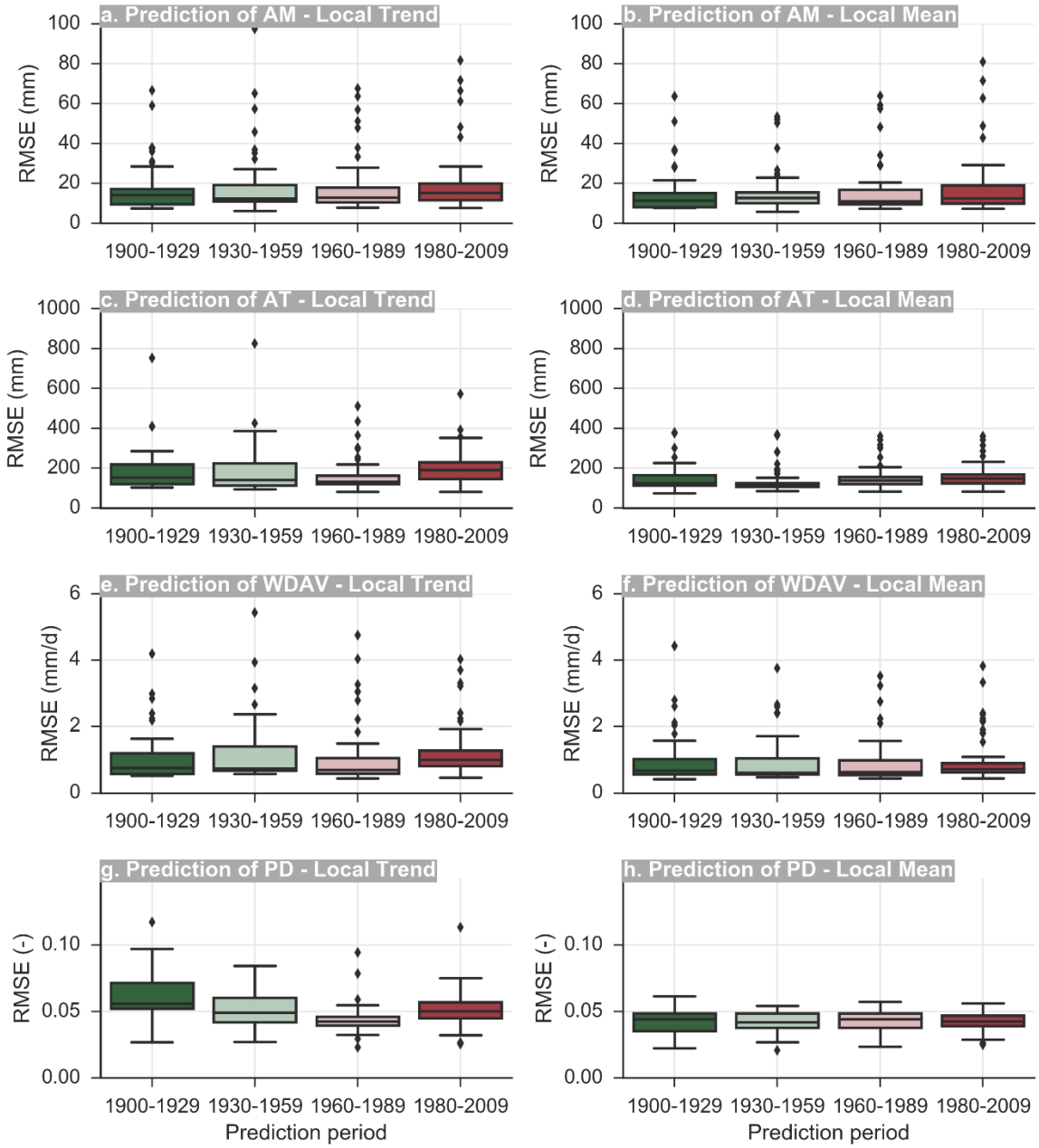


Figure 8.6 Boxplots of the RMSE distribution from the static validation application to the stations with data in all four prediction periods, 1900-1929, 1930-1959, 1960-1989, 1980-2009, for the local mean (L-Mean) and local trend (L-Trend) models, for all rainfall indices. For the boxplots' properties description, see Figure 8.4.

In order to gain further insights into temporal changes of predictability, we compare the predictive performance of each model (L-Mean, L-Trend) for four distinct climatic periods, covering the past 110 years up to year 2009. It is observed (Fig. 8.6) that the error distribution of the L-Trend model does not present pronounced temporal differences for the indices among these periods, with the exception of PD which shows a larger, yet not consistent, variability over these periods. Among the four periods, the L-Trend model performed best in the prediction of the 1960–1989 period, based on calibration on 1930–1959, a period which however does not include the decades of pronounced increase in greenhouse emissions (from the 60s and thereafter). The

predictive performance of trends on the latest period is not markedly different from the previous periods, if not it is slightly worse for some indices, e.g. the AT. A particular pattern is neither observed for the L-Mean. As it will be discussed next, these results seem to be well-within the range of the statistical variability of the predictive skill of each model, evaluated from the whole record. Finally, in this examination as well, the L-Mean model proves superior to the L-Trend (only one or two exceptions are seen).

8.6 Moving-window validation of predictive performance

In this section, we explore the predictive qualities of the models by delving into the statistical analysis of the whole record, considering the models from the global-moving calibration as well, namely, the global trend and the global mean.

8.6.1 An examination of one of the longest records

As an illustration of the application of the methodology, we first explore the longest uninterrupted station of our dataset, i.e. the Prague station in Czech Republic (211 years), shown in Figure 8.7. The models' error evolution pattern is reflective of their performance. For the majority of time, the mean models are at the lower front of the errors, with the local mean model showing slightly superior performance. The local trend model results in higher errors and its predictions may quickly deteriorate, taking longer to converge to the mean models' predictions in areas of lower errors (Fig. 8.7). This is attributed to the fact that the trend model projects to the future sensitive features of the calibration period, i.e. extreme observations or 'trendy' behaviour, which do not have a high chance to survive the end of the calibration period. The more parsimonious structure of the mean model encapsulates minimal but robust knowledge of the process behaviour, which is more likely to characterize its future evolution as well. In the absence of an underlying global trend and as the sample grows larger, the global trend model converges to the predictions of the mean models, but its performance remains slightly inferior even towards the end of the record.

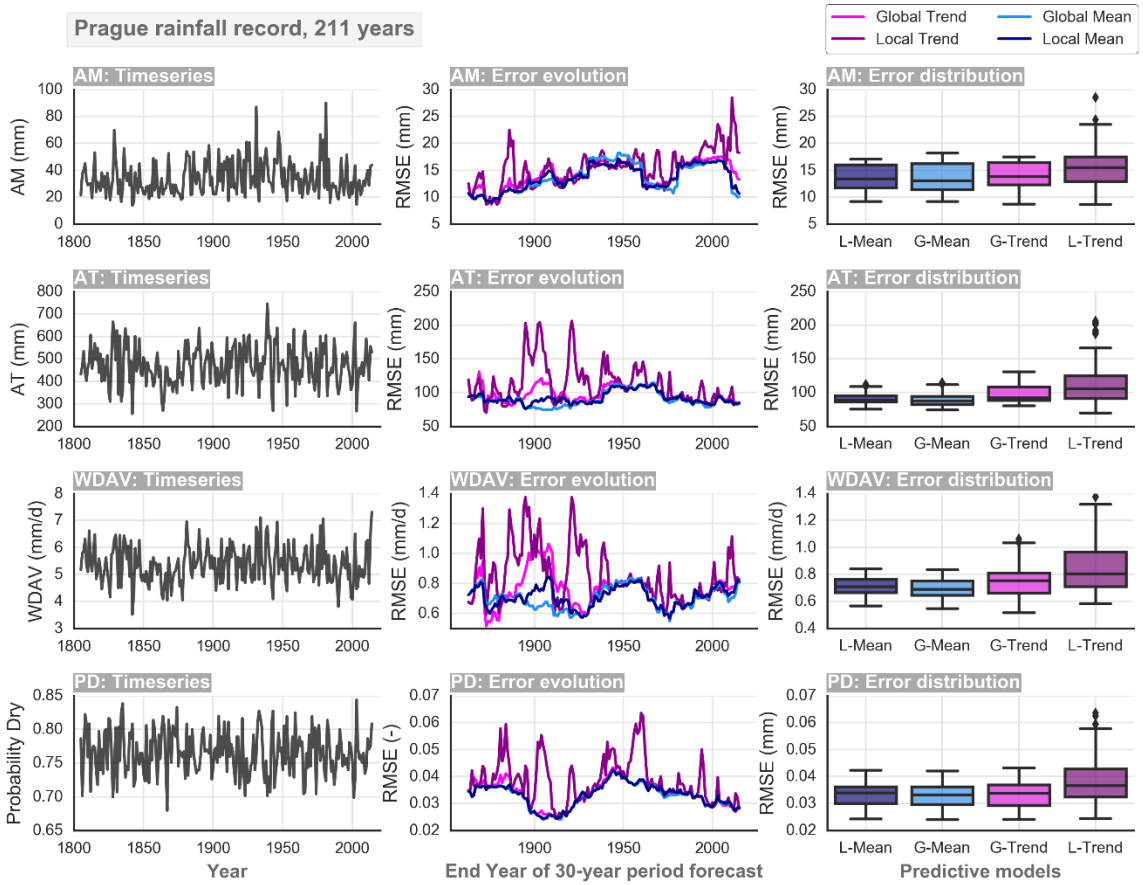


Figure 8.7 Case study of the rainfall station in Prague. Timeseries of annual maxima, annual totals, annual wet-day average and annual probability dry, error evolution and distribution of the prediction RMSE for the four prediction models, global and local trend, and global and local mean.

8.6.2 Application to all records

Figures 8.8-8.11 show the empirical distributions of the models' predictive RMSE for each rainfall index and for all 60 stations. For most stations the local mean and global mean models have the lower probabilities of exceeding high errors, contrary to the local trend model whose error distribution is clearly shifted to the right, in the higher error area. The distribution of the predictive RMSE of global trend model is located in between the two, showing in general a better behaviour than the local trend.

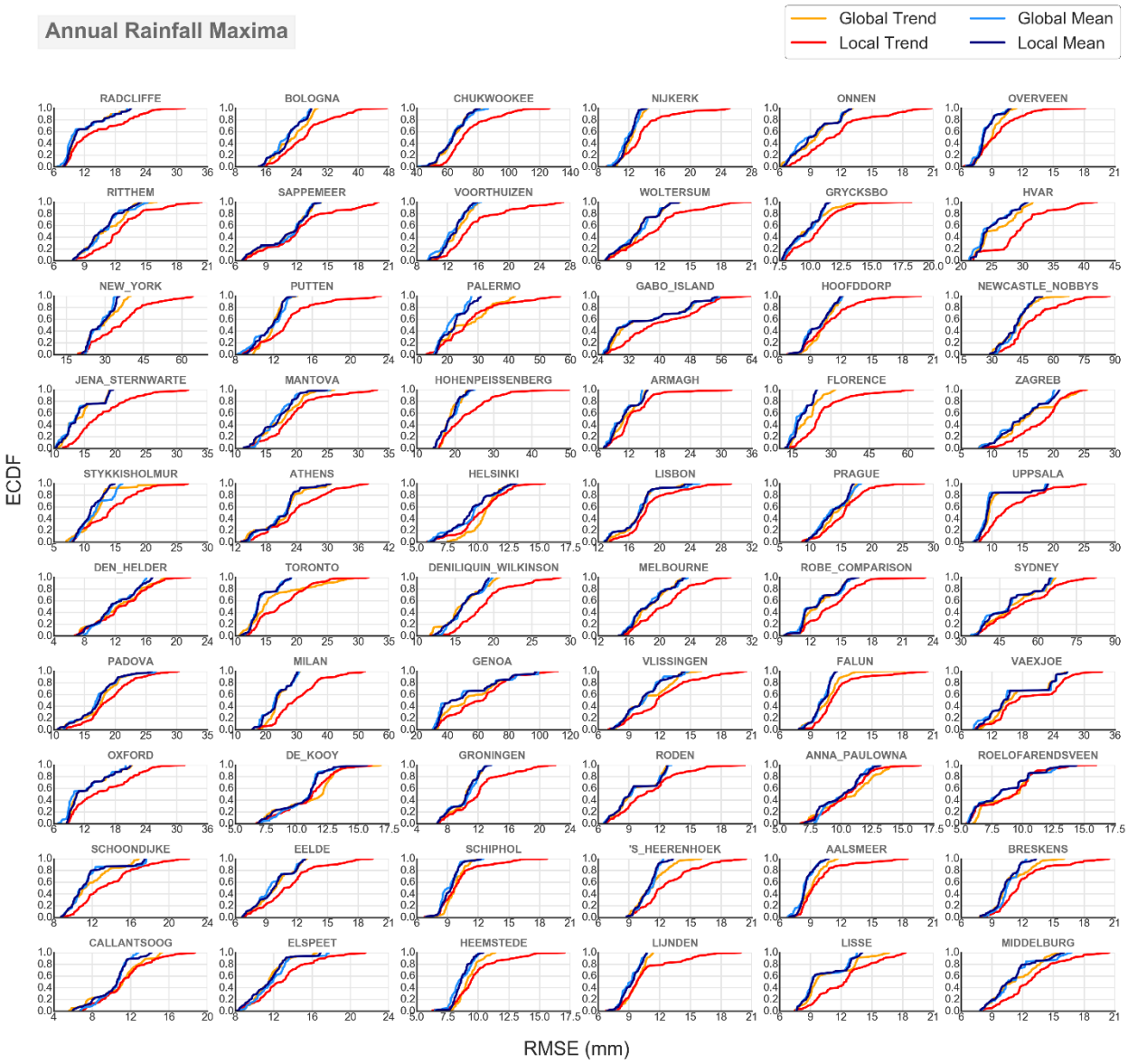


Figure 8.8 Empirical cumulative distribution function (ECDF) for the prediction RMSE of annual maxima for the local trend, the global trend, the global mean and the local mean model for the 60 stations.

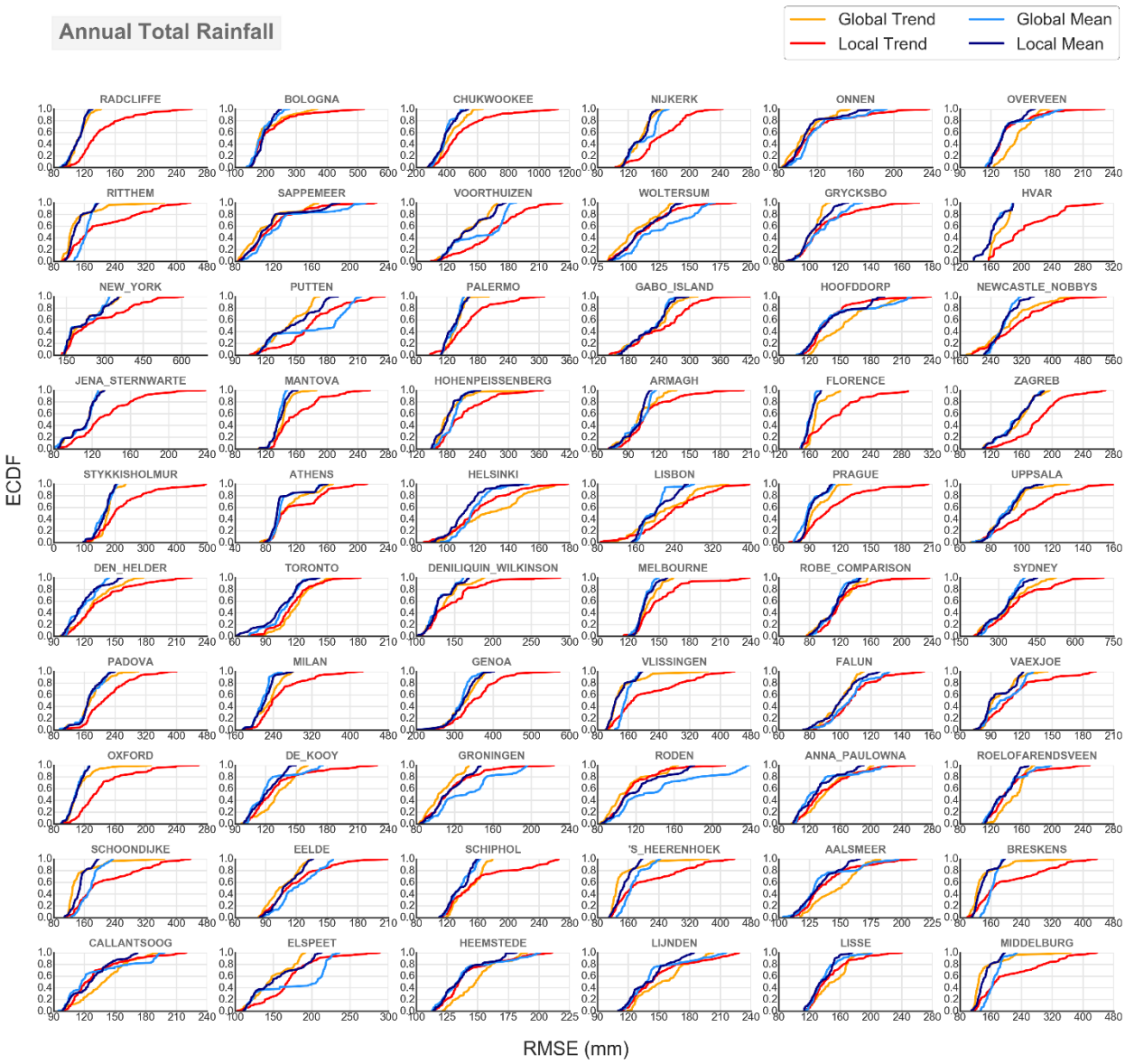


Figure 8.9 Empirical cumulative distribution function (ECDF) for the prediction RMSE of annual totals for the local trend, the global trend, the global mean and the local mean model for the 60 stations.

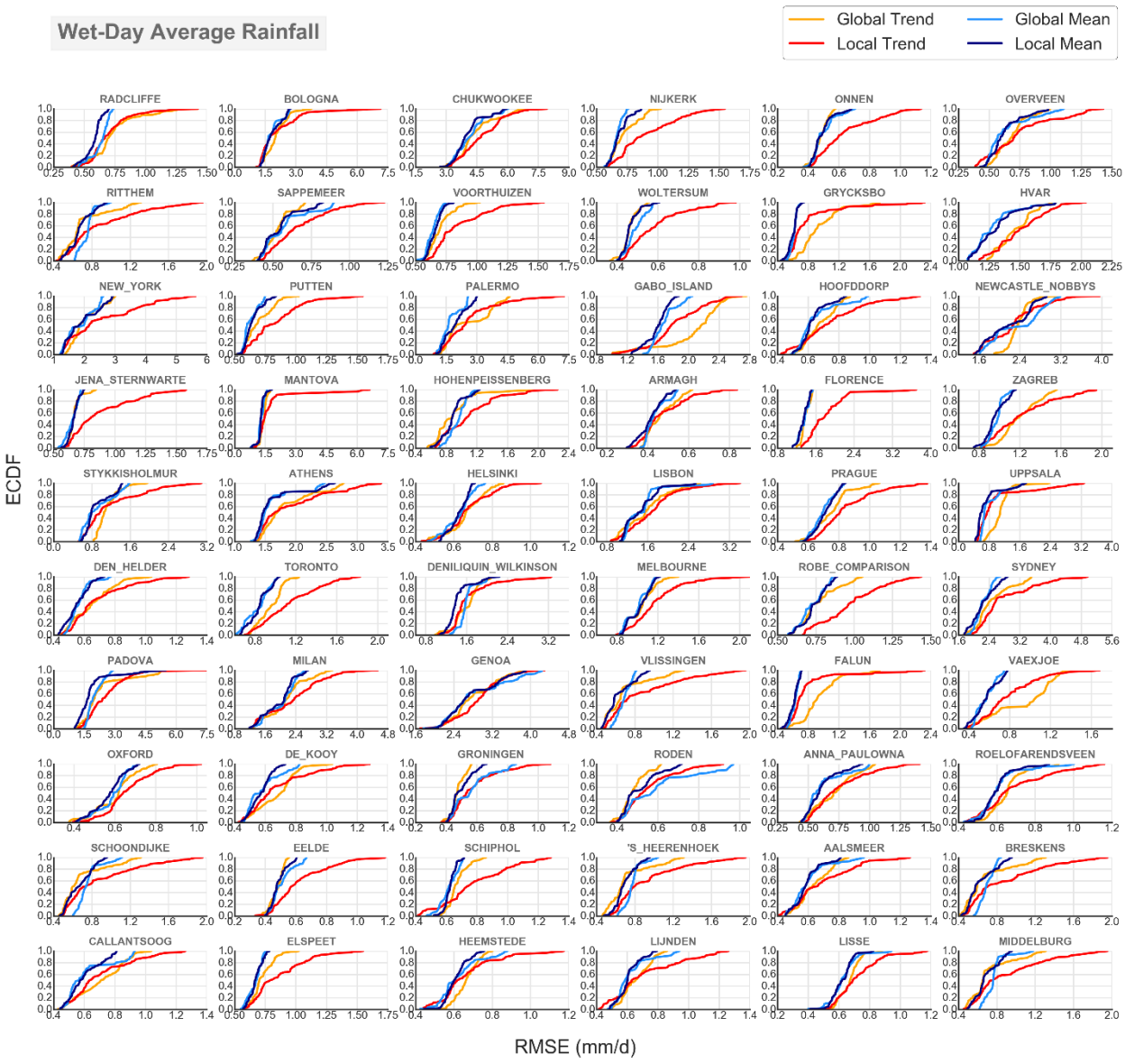


Figure 8.10 Empirical cumulative distribution function (ECDF) for the prediction RMSE of wet-day average rainfall for the local trend, the global trend, the global mean and the local mean model for the 60 stations

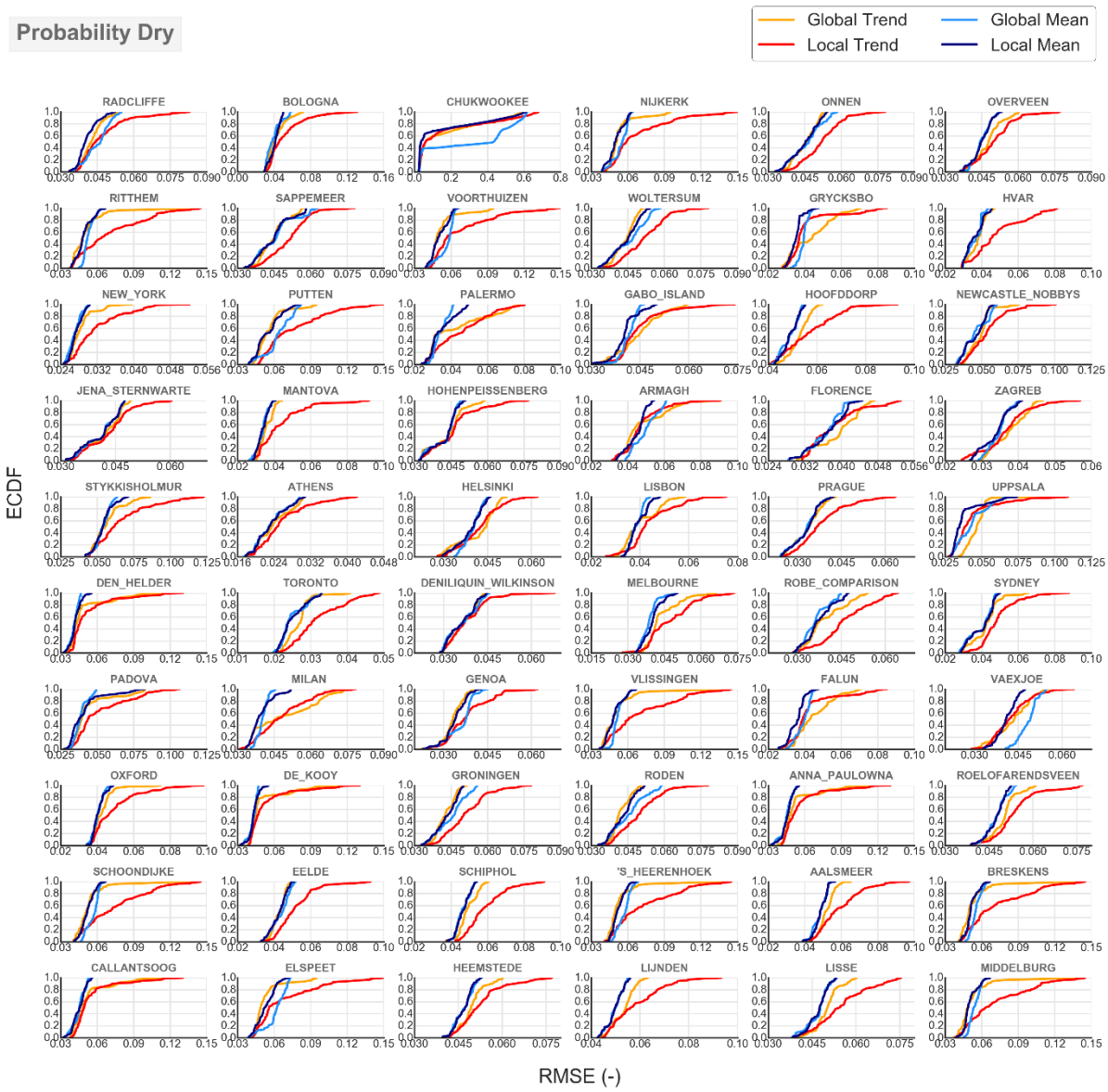


Figure 8.11 Empirical cumulative distribution function (ECDF) for the prediction RMSE of probability dry for the local trend, the global trend, the global mean and the local mean model for the 60 stations.

A summary of the distributional properties of the prediction RMSE of all stations shown in Fig. 8.8-8.11, is provided in Fig. 8.12, in terms of the average and the standard deviation of the RMSE distribution of each station. The average values of the latter also summarized in Table 8.1. Accordingly, the models' performance can be ranked from best to worst as follows: (1) local mean, (2) global mean, (3) global trend and (4) local trend. The local mean model marginally outperforms the global mean with respect to the average RMSE, yet in terms of the standard deviation of the RMSE distribution (Fig. 8.12b, d, f, h), it is evident that the local mean model prevails showing smaller standard deviation of prediction errors, and thus more reliable performance. In this case, the linear trend model shows markedly inferior performance.

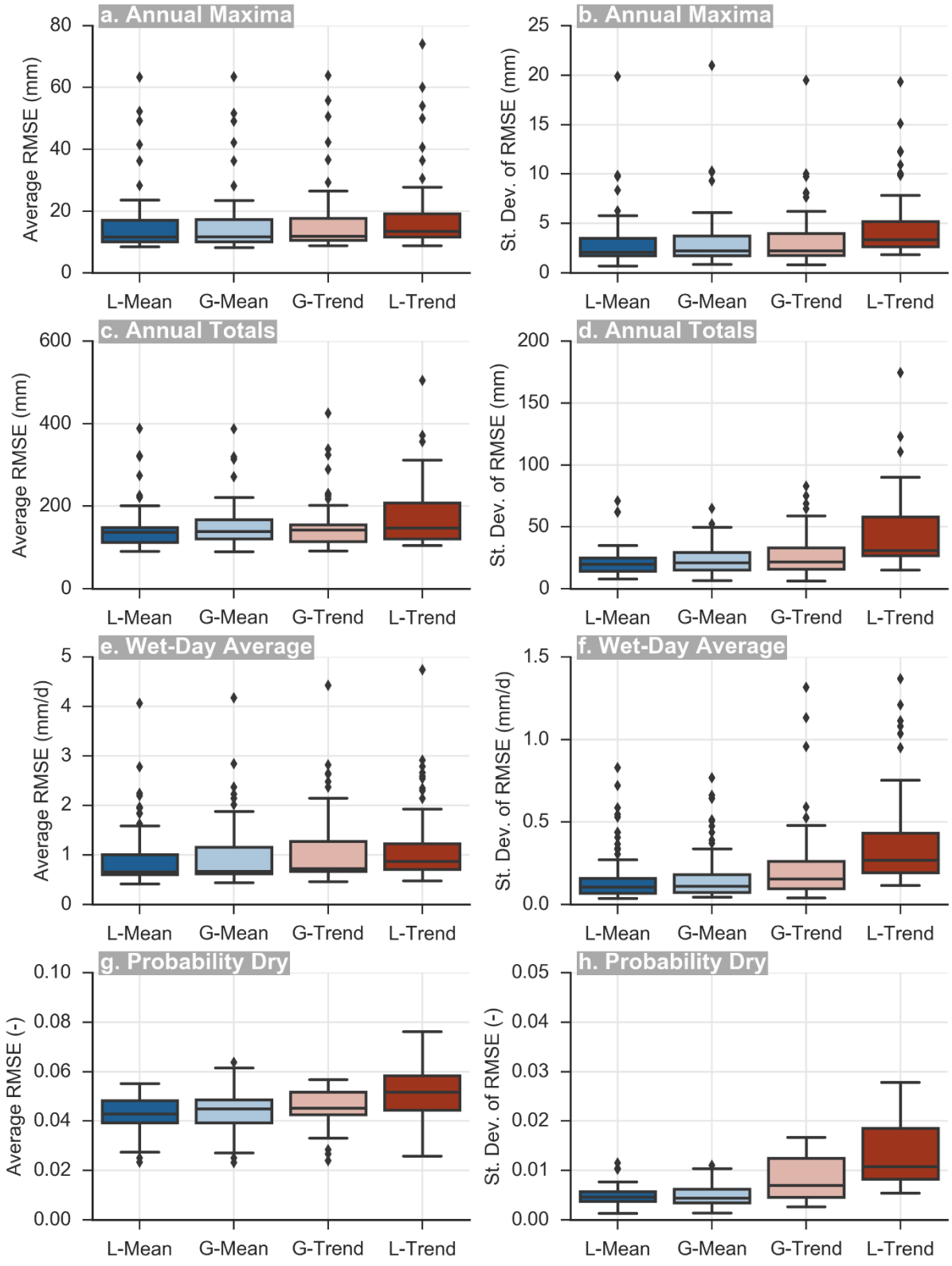


Figure 8.12 Boxplots of the average prediction RMSE and standard deviation of RMSE as estimated for each station from moving window application of the local (L-) mean, global (G-) mean and local (L-) and global (G-) trend for all the indices. For the boxplots' properties description see Figure 8.4.

Table 8.1 Averages of the average RMSE and the standard deviation of RMSE of the four models (local (L-) mean, global (G-) mean, local (L-) trend and global (G-) trend) from all stations and for all four indices, as shown in Figure 8.12.

	Annual Maxima (mm)				Annual Totals (mm)			
	L-mean	G-mean	G-trend	L-trend	L-mean	G-mean	G-trend	L-trend
Average RMSE	16.00	16.05	16.73	18.76	149.07	154.18	154.77	174.7
St. Dev. RMSE	3.04	3.13	3.37	4.74	21.52	23.02	27.4	45.45
	Wet-Day Average (mm/d)				**Probability Dry (-)**			
	L-mean	G-mean	G-trend	L-trend	L-mean	G-mean	G-trend	L-trend
Average RMSE	0.98	1.01	1.11	1.2	0.04	0.05	0.05	0.05
St. Dev. RMSE	0.18	0.18	0.27	0.39	0.01	0.01	0.01	0.02

8.7 Models' performance under natural variability: an experiment with synthetic series

Following the rationale outlined in Section 8.4.6, the goal of this experiment is to test the performance of the predictive models in conditions of enhanced structured uncertainty, characterized by changes at all scales and 'trend-like' behaviour for small periods. As the latter are distinctive features of persistent processes (Koutsoyiannis, 2002), we produce five long-term timeseries from a standard normal distribution with length N = 10 000 that reproduce HK dynamics, using the SMA algorithm (Koutsoyiannis, 2000; Dimitriadis and Koutsoyiannis, 2018). The series are generated with increasing degree of persistence, quantified through the Hurst parameter H, from mild persistence H = 0.6 to very strong H = 0.99. In order to explore the impact of record length we also examine smaller segments of the same timeseries of lengths N = 100 and N = 1000. Because smaller segments are impacted by larger estimation uncertainty, we plot the average ECDF of the prediction RMSE estimated from non-overlapping segments extracted from the original timeseries of length N = 10 000. Therefore, the N = 100 plots correspond to the average of 100 timeseries of length 100, derived from the 10 000 series. Likewise, the N = 1000 series are the average of 10 timeseries of length 1000. The plots of the ECDF distribution (Fig.8.13) of the prediction RMSE for the four predictive models are produced employing the same dynamic validation schemes applied for the real-world stations.

The contrasting performance of the two local models is observed here as well; local features are better exploited by the mean rather than the trend model, irrespective of the record size. The latter becomes important when the global models are

considered. In the absence of a global underlying trend, the increased variability encountered in small calibration periods ($N = 100$) leads the global trend model to bad predictions. When the trend model is calibrated from larger series, the trend component is smoothed out, and therefore, the prediction performance approaches the one from the mean models. Regarding the competition between global and local mean, it appears that it is a function of both the record length and degree of persistence. For large record lengths and $H > 0.7$, the local mean model prevails, while for small record lengths and medium persistence, the two are comparable. In persistent process, where clustering arises, local information is likely to be more relevant for prediction, yet for long-term prediction as is the case here, 'local' may need to extend a few steps back in the past, which for small record lengths could be within the reach of the calibration period employed for the global mean model. Obviously though, results from the global model become less relevant when the sample is large and therefore global information extends too far in the past. A thorough treatment of the theoretical basis and practical formulation of local mean models in relation to the persistence properties of the parent process is given by Koutsoyiannis (2020b).

We note that the behaviour observed in the $N = 100$ plots is qualitatively consistent with the one observed from the rainfall records. Moreover, indices known for their persistence properties, such as annual totals (Iliopoulou et al., 2018b; Tyralis et al., 2018) and probability dry (Koutsoyiannis, 2006) show a slight preference for the local mean model. In other cases where persistence is less manifested, as in annual maxima (Iliopoulou and Koutsoyiannis, 2019), the performance of the global and the local mean model in terms of the average RMSE are indistinguishable (Fig. 8.12); the variance of the errors still being smaller for the latter.

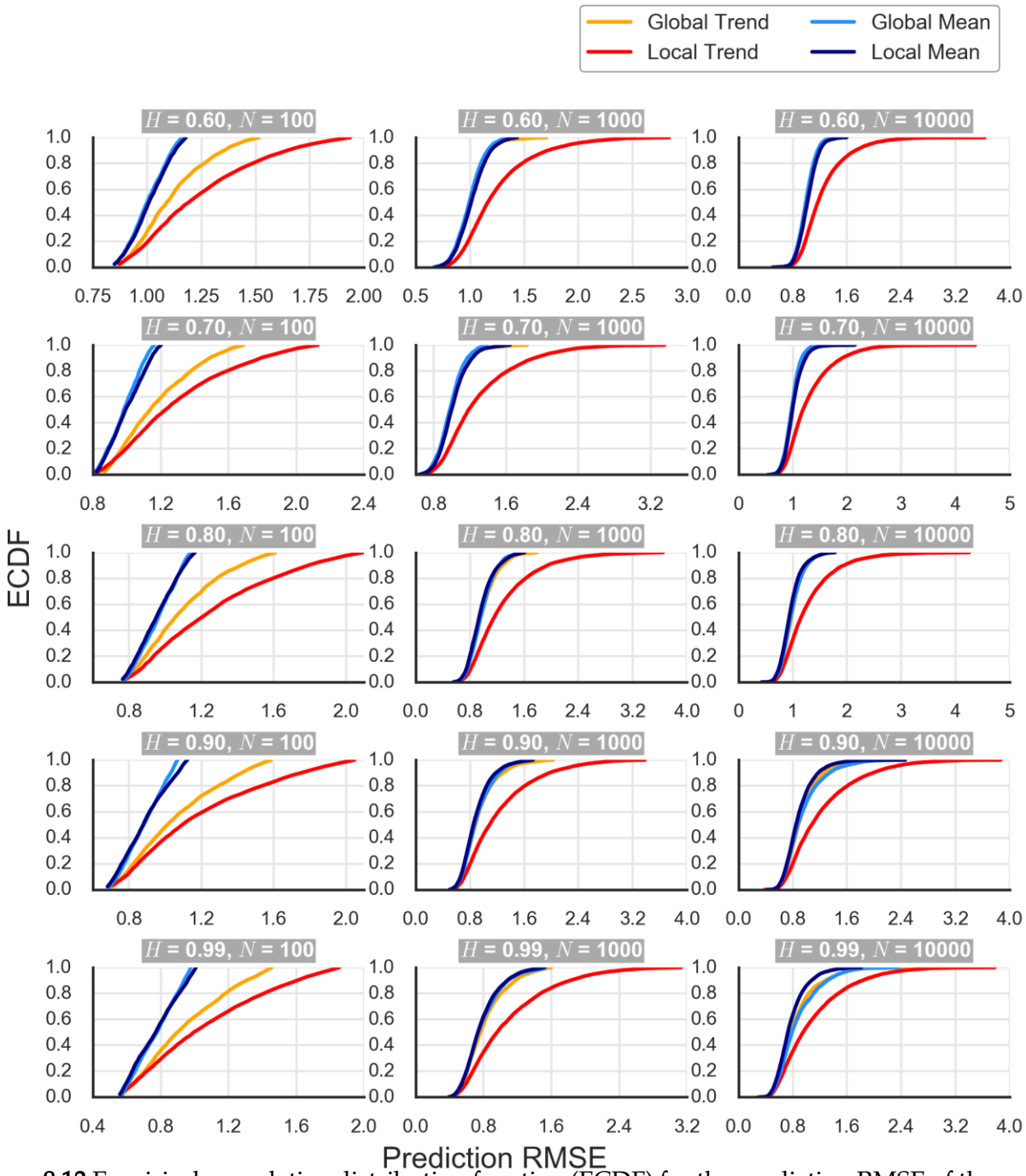


Figure 8.13 Empirical cumulative distribution function (ECDF) for the prediction RMSE of the HK timeseries resulting from application of the local trend, the global trend, the global mean and the local mean model, for segments of the original timeseries with increasing sample size, N =100, 1000, 10 000 (original). The ECDF for the first two lengths are the averages as computed from 100 and 10 non-overlapping segments of the 10 000 values.

8.8 Discussion

8.8.1 On parsimony and predictive accuracy

In the above controlled experiment, where the generating mechanism of the data is known, it is evident that among the four 'false' models, the local mean yields the most accurate predictions in terms of RMSE, using in-sample data more efficiently by means

of its single parameter. The increase in predictive accuracy and statistical efficiency is tightly associated with the notion of parsimony, which is a dual criterion measuring the model's fit to the data as well its simplicity (Gauch, 2003). In these terms, the local mean model is deemed to be a parsimonious model, since it fits the out-of-sample data either better or at least equally well to the more complicated trend model.

The reason behind the sometimes interchangeable use of the words parsimony and simplicity is a certain tendency of simple models to make reliable predictions, which among other approaches as information criteria discussed in Section 8.3.1, is also incorporated as a concept in Bayesian analysis assigning higher prior probabilities to simpler models, and a posteriori favouring the simpler model (Berger and Bernardo, 1992; Berger and Pericchi, 1996; Gauch, 2003 and references therein). More recent developments from the Bayesian standpoint include constructing penalized complexity priors (Simpson et al., 2017), while the concept informs variable selection in linear regression though various techniques as the Lasso and ridge regression (Tibshirani, 1996). Another demonstration of the relation between predictive accuracy and simplicity is the possibly better predictive performance in terms of mean square error of simpler, yet misspecified models, compared to the ones derived from the correctly structured model (Hocking, 1976); for instance, Wu et al. (2007) provided a set of conditions for which this holds true in the case of linear models. Therefore, theoretical arguments are in favour of simpler predictive models, all the more so in the case of natural processes characterized by a great degree of variability, for which our understanding is limited. A comprehensive discussion on the connection of simplicity to wider epistemological and philosophical principles is provided in Gauch (2003).

8.8.2 On alternative climatic predictors of rainfall

It is beyond the scope of the paper to formulate and suggest a good climatic prediction method for rainfall. Having shown however that past climatic trends of rainfall are not useful predictors of its future evolution, it is tempting to reflect on a common alternative option for long-term prediction, namely the use of large-scale climatic oscillations. The latter are considered a potential source of decadal climatic predictability (Latif et al., 2006). The predictive skill arising from the use of a climatic oscillation as a covariate for prediction relies upon two factors; existence of significant correlation of rainfall with large-scale climatic oscillations, and reliable predictability of the latter. On the over-decadal climatic scale examined here fulfilment of both conditions is challenging. There is an increasing number of studies relating climatic oscillations to decadal rainfall, but both the type of the correlated oscillation and the specification of the correlation (type, lagged response), are region-specific (e.g. Krichak et al., 2002; Scaife et al., 2008; Lee and Ouarda, 2010; Sun et al., 2015; Krishnamurthy and Krishnamurthy, 2016; Nalley et al., 2019). Therefore, with respect to multi-sites analyses, the identification of robust response patterns of decadal rainfall to climatic oscillations constitutes a nontrivial research subject. Even more challenging is the predictability of the climatic oscillations themselves on the 30-year scale. For instance, it is only during the last 5 years, that prediction of the North Atlantic Oscillation (NAO) has become skillful on the seasonal scale, and at the moment research efforts are directed towards predictability on beyond annual scales (Scaife et

al., 2014; Smith et al., 2016). While some progress has been reported in terms of the decadal predictability of climatic oscillations related to the NAO, as the Atlantic Multi-decadal Oscillation (AMO), predictability of the actual values of the NAO beyond the seasonal scale remains very limited (Smith et al., 2016; Yeager and Robson, 2017). A relevant case study by Lee and Quarda (2010) concluded that predictions of decadal streamflow extremes using the NAO as a covariate were impacted by large uncertainty to the point of almost being non-informative. Although a promising research subject, it appears that in the best case, there is still way to go before attaining hydrologically relevant climatic predictions based on climatic oscillations, at least to the degree that this is becoming possible at the seasonal scale for some regions (e.g. Scaife et al., 2014). Yet the case that this proves to be infeasible cannot be excluded (Koutsoyiannis, 2010).

8.8.3 Can a stationary framework be compatible with a deterministic forcing?

A question that often arises is the relevance of past predictability under the hypothesis of a climate impacted by monotonic anthropogenic forcing, not existing in the past. In this case, it could be argued that the examination of the predictive performance in the past in which stationarity is implicitly assumed, is an irrelevant approach as the past might no longer representative be of the future. As a first remark, it is worth recalling that change is not synonymous to non-stationarity, while in the presence of uncertainty in every real-world system, the choice of a stationary versus a non-stationary model is done in terms of modelling convenience rather than based on the existence (or co-existence) of deterministic drivers (Montanari and Koutsoyiannis, 2014; Koutsoyiannis and Montanari, 2015b). De Luca et al. (2019) yet shed further light on this misconception by the following experiment. They show that artificially imposed trends —of the projected magnitude of climate scenarios, on the parameters of a sub-hourly rainfall generator regarding bursts intensity, duration, and number of occurrences, were masked on coarser temporal scales and as a result, they could be adequately modelled by a stationary extreme value model. This suggests that the presence of deterministic drivers in a system does not disfavour stationary modelling. For there is the possibility that even systematic changes may not be manifested at the scales of interest to the degree that they warrant a more complicated representation for the future. Hence, the examination of a stationary framework is justified also in the presence of monotonic and accelerating forcing, as it aligns with the abovementioned principle of parsimonious modelling. Therefore, the question shifts from the existence or not of deterministic drivers, to evaluation of the degree to which observed changes require a more complicated modelling. In our case, it is assumed that the past is still representative *enough* for the future in order to achieve a similar degree of predictability by the given models, which is not falsified by the examination of the recent period. The entire question however relies on a simplistic view of complex systems, i.e. that just one factor (or the change thereof) suffices to determine the system's future evolution. In our view, this is not a logically consistent framework for dealing with complex systems.

8.9 Summary and conclusions

Under the popular assumption of intensification of the water cycle due to global warming, a considerable deal of contemporary research in hydrology revolves around the study of temporal changes of extremes, with the application of trend analyses being on the rise during the past two decades. While the explanatory analysis of trends has dominated the relevant studies, assessment of the predictive skill of trend models has not been equally assessed, despite the apparent significance of such a task for risk planning. This research reframes the problem of trend evaluation, as a model selection problem oriented towards identifying the model with the best predictive qualities in deterministic terms, which is neither equivalent to the 'true' model nor to the model better at explaining the in-sample data.

For this purpose, we introduce a systematic framework for evaluating projections of trends by means of comparing the prediction RMSE to the one obtained from simpler mean models. We perform a variation of cross-validation, also known as walk-forward analysis, devising two distinct calibration and validation schemes (Fig. 8.4). In block-moving calibration we fit the linear trend and mean models to 30 years of data (local trend and local mean) and we validate the results based on the outcome of their predictions for the next 30 years. This procedure is repeated using sliding windows till the end of the record is met. In global-moving calibration, we fit the models to all the known period (global trend and global mean), assuming that in the beginning, one knows only the first 30 years, and progressively the calibration period grows larger. In this case too, we evaluate the outcome of the predictions of the models for the next 30 years, therefore the projections of the four models can be compared in terms of the statistics of their empirical distribution of errors.

The models compete in predicting the out-of-sample behaviour of four rainfall indices: annual maxima, annual totals, annual wet-day average rainfall and probability dry at the annual scale, as estimated from a unique dataset comprising the 60 longest rainfall records surpassing 150 years of daily data. Results show that models rank from best to worst as follows: local mean, global mean, global trend and local trend. A separate examination of the latest 30-year period for each station confirmed the above rank of the models as well. The temporal changes in the prediction error distribution among four fixed climatic periods, common for all stations covering 110 years up to 2009, are also investigated. Fluctuations of predictability do occur among the climatic periods, yet no increase in predictability is achieved by the local trend model for the latest period (1980–2009), compared to earlier periods. Results from both analyses show that future rainfall variability is on average better predicted by mean models, since local trend models identify features of the process that are unlikely to survive the end of the calibration period, either being extreme observations, or 'trend-like' behaviour. These features are smoothed out in longer segments, which is the reason behind the better performance of global trends. Robust regression techniques were also employed for the calibration of local trends but perhaps not surprisingly, did not improve the out-of-sample predictions (see discussion in Appendix B.3).

In an attempt to reproduce the observed behaviour, we generate long-term timeseries exhibiting long-term persistence or HK dynamics (Koutsoyiannis, 2011b; O'Connell et al., 2016; Dimitriadis, 2017), and carry out the same analysis. Persistent processes show enhanced variability and a user unfamiliar with their properties may

misinterpret segments of their timeseries as trends, which perhaps explains why trend claims have been that common lately. Results from the synthetic records show qualitative similarities with the ones from empirical rainfall records, known to exhibit persistence, depending on the scale and studied index (Koutsoyiannis, 2006; Markonis and Koutsoyiannis, 2016; Iliopoulou et al., 2018b; Iliopoulou and Koutsoyiannis, 2019). The local and global mean outperform the local trend model for all degrees of persistence and sample sizes, while for small record lengths ($N = 100$) the performance of the global trend model is notably inferior too. Local and global mean models hardly show differences for medium degrees of persistence, but the local mean prevails for strong persistence.

From a systematic investigation of long-term rainfall records, corroborated by simulation results, we have verified that local trends have poor out-of-sample performance, being outperformed in their predictions by simpler models, as the local mean. This empirical finding suggests that the large inherent variability present in the rainfall process makes the practice of extrapolating local features in the long-term future dubious, especially when the complexity of the latter increases. This in turn questions the theoretical and practical relevance of projections of rainfall trends and the grounds of the related abundant publications.

9. SYNOPSIS AND OUTLOOK

9.1 Summary of scientific background and motivation

Scientific interest in hydrological extremes has historically been at the center of hydrology and engineering studies. At present, amid growing climate change concerns, this interest has reached all-time high levels (see Section 8.3.1). What is more, during the past decades, the increasing availability of examples of catastrophic events and related engineering disasters has called into question traditional risk perception and modelling approaches. As a result, there is no shortage of scientific efforts to explain the variability observed in rainfall and streamflow extremes.

In most studies, the assumption of independence of extremes is omnipresent, although its validity has been challenged for the parent hydrological processes since the mid 20th century (Hurst, 1951). Yet the process's extremes are still widely treated as independent random variables, while it has become common practice to view any deviations from independence as signals of deterministic drivers. As a result of a priori resorting to the independence assumption, the modelling focus is dominated by the study of the marginal distribution of extremes overlooking their temporal variability and dependence properties. At the same time, the scarce studies that deal with dependent extremes mostly employ methods of statistics that treat extremal dependence as a singular behaviour, decoupled from the variability of the parent process.

The central objective of this thesis is to investigate and model the temporal dynamics of extremes under the framework of stochastics, without employing the IID assumptions of statistics that are unlikely to be tenable in real-word conditions. To this aim, a rare dataset of long-term observational records is compiled. The goal is to integrate the understanding and modelling of the temporal dynamics of extremes from seasonal to climatic scales, to that of the parent hydrological process, as an inherent part of its variability. Such an approach provides novel insights into the dynamics of hydrological extremes that may enhance risk perception and inform related mitigation practices. The relevant contributions are discussed below.

9.2 The main contributions

The magnitude of extremes is determined by the marginal distribution of their parent process, yet their temporal distribution that critically affects our perception of them, is also controlled by the joint properties of the process. As the assumption of independence dominates the study of extremes, the latter are seldom studied. This thesis contributes to the stochastic characterization and modelling of the temporal dynamics of daily rainfall and streamflow extremes at three scales: (a) the seasonal, (b) the annual, and (c) the climatic. The respective contributions are presented below.

9.2.1 On seasonal dynamics

Chapters 4 and 5 focus on seasonal dynamics in the rainfall and streamflow extremes. In particular, Chapter 4 deals with the change in the distributional properties of

seasonal rainfall extremes, and Chapter 5 investigates seasonal dependence of streamflow extremes. The main contributions are:

(a) An objective methodology is proposed to perform season identification in extreme daily rainfall and model the resulting extreme properties in each season. The framework employs the Akaike information criterion to resolve the problem of subjectively selecting an optimal number of extreme rainfall seasons and their monthly partition.

(b) The effect of seasonality on extreme rainfall properties is discerned. Seasonality affects the central tendency of rainfall maxima, being manifested by a change in the scale and location parameters of the seasonal extreme value distributions. On the other hand, the shape of their probability distribution and its tail do not substantially vary from season to season. Therefore, a pooled estimation of the shape parameter of seasonal and annual extremes is suggested to reduce uncertainty.

(c) Estimation uncertainty in fitting seasonal-annual maxima distributions is shown to be relevant even for long-term rainfall records. In this respect, extreme-oriented fitting methods, namely weighted-least squares, are proposed to resolve inconsistencies that may arise from an independent fitting of the extreme value distributions to seasonal and annual extremes.

(d) Rivers in Europe are shown to exhibit persistent features at the seasonal timescale, manifested as correlation between preceding average flows and anteceding seasonal 'extreme' flows, i.e. peak flows in high flow season, and average flows in low flow season, respectively. This correlation can be explored to increase seasonal predictability which is generally higher for low flows, but may be significant for high flows as well depending on the climatic region and catchment properties.

(e) Seasonal streamflow predictability is found to be enhanced in less humid climatic regimes and catchments dominated by baseflow and characterized by slower response times.

9.2.2 On long-term persistence dynamics

Chapters 2-3 and 6-7 deal with long-term extremal dynamics stemming from presence of persistence, i.e. HK dynamics in the parent process. In particular, Chapter 2 reviews the existing theory and modelling practices for dependent extremes, Chapters 3 revisits the case for persistence in the annual rainfall process, while Chapters 6-7 deal with propagation of persistence to the properties of the extremes. Specifically, Chapters 6 examines the effects of persistence on the multi-scale occurrences of extremes, whereas Chapter 7 investigates its manifestations in extreme value modelling. The respective contributions are:

(a) The presence of HK dynamics in the annual rainfall process is validated using a global rainfall dataset (1265 stations). Persistence is quantified through a common Hurst parameter equal to $H \approx 0.6$. Approximately 2.5% of the stations show even stronger dependence that cannot be explained by the common H parameter. Annual rainfall correlations are low but deviate from independence,

while the decay of the correlation structure is slower than predicted by a Markovian process.

(b) A rare dataset of long-term daily rainfall records surpassing 150 years is compiled to gain insights into rainfall historical variability. The abovementioned second-order scaling behaviour is supported by evidence from this dataset as well, and further the presence of a second weaker scaling regime at shorter time-scales (of the order of months to few years), is revealed.

(c) It is shown that extremes tend to 'hide' the persistence of the parent process, often falsely signalling independence. Furthermore, persistence and heavy tails have synergistic effects on the temporal properties of extremes. As a matter of fact, persistence of non-Gaussian extremes cannot be retrieved solely by second-order characterizations, such as the Hurst parameter and the dispersion index. The latter are only relevant for Gaussian processes, thus of limited interest to the studied processes.

(d) A new probabilistic index is formulated to reveal extremal long-term clustering via the multi-scale probability of not exceeding a threshold, termed the NEPvS index (Non-Exceedance Probability vs Scale). A related two-parameter model is introduced which captures the scaling behaviour of extreme event occurrences for processes exhibiting a range of second-order and marginal properties, including strong persistence and heavy tails.

(e) Evidence of persistence diminishes as the threshold increases and thus, the examination of lower thresholds is essential for retrieving it. As the threshold increases, the behaviour of extremes may falsely suggest independence of the parent process.

(f) The index brings forward the central manifestation of persistence in extremes, i.e. the increase of the probability of non-occurrence. This means that prolonged periods of absence of extreme events are more probable for persistent processes than for IID ones.

(g) On the other hand, persistence also alters the conditional properties of extremes at a given scale (as in annual POT events) by producing short-term clustering. Compared to an IID process, dependent extremes show an increase in duration, and hence, in aggregate intensity, and are characterized by positive association between their number and intensity. Thus, their temporal dynamics are more challenging to hydrological design and risk management than those of an IID process.

(h) Extreme value theory under dependence is reviewed and open questions pertaining to persistent processes are approached through Monte Carlo simulations. It is found that the GEV distribution remains a good sub-asymptotic model for block maxima even for strongly persistent processes. In such cases, similar to the IID case, convergence to the Fréchet distribution is still much faster than to the Gumbel, while fitting to the higher-quantile region improves by the weighted least-squares method.

(i) The extremal index theory formulated for extremes exhibiting local dependence is found to be only partially relevant for persistent processes, characterized by both short- and long-range dependence. In agreement to the theory, in a persistent process the probability of exceedance of extremes is lower compared

to the IID process which acts as the upper bound of the exceedance probability distribution. Persistence is reflected in the parameterization of the GEV distribution in the following way: (i) the shape parameter remains unaltered, as it is a property of the marginal distribution, (ii) the location parameter is lower, as a result of clustering, yet (iii) the scale parameter increases —contrariwise to extremal index theory, as extremes inherit part of the increased variability of the parent persistent process.

(j) The investigation of temporal properties of rainfall extremes from long-term records exposes departures from the IID behaviour which are shown to be consistent to their persistence structure. Evidence of stronger dependence is found for streamflow extremes.

(k) HK-type models calibrated only the first four moments and the second-order scaling behaviour of the process, show promising results in capturing both short-term and long-term clustering patterns of rainfall and streamflow extremes.

9.2.3 On future projections of climatic rainfall dynamics

Chapter 8 examines the empirical and theoretical grounds for the increasing body of literature dealing with rainfall trends and their projections to the future. The respective contributions are:

(a) A prediction-oriented framework is introduced for the evaluation of trends, formulated as a variant of cross-validation suited for analysis of timeseries. The framework allows bypassing the caveats of 'statistical-significance' methods, by directly considering predictive skill of trends instead of their explanatory power. It also enables a statistical assessment of hindsight bias in terms of the ability to foresee climatic trends.

(b) The predictive performance of trend models is compared to the one of simpler mean models which shows that the process's mean is on average a better predictor of the climatic behaviour of rainfall indices (annual totals, maxima, average and probability dry). The superior performance of the mean model is also the case for the most recent climatic period.

(c) It is further shown that persistence favours prediction based on the recent past, i.e. based on the local mean, rather than the entire past, i.e. based on the global mean, in spite of the latter being closer to the true mean. The same holds true for the empirical rainfall records.

9.3 Directions for further research

At present the interest in the temporal variability of rainfall extremes and flood events is rising as the management of hydroclimatic risk is considered one of the most prominent challenges for the scientific community. This research investigated the temporal variability of hydrological extremes harnessing rare evidence from long-term empirical records. Such evidence was in favour of temporal behaviours differing from the ones of IID processes, and whose modelling invokes a stochastic approach. A set of probabilistic frameworks and stochastic tools was developed based on the

idea of characterizing and modelling the observed extremal variability using inherent features of the parent process, highlighting, in particular, the role of the second-order properties. Further research is needed to improve the understanding of variability and dependence dynamics in hydrological extremes and identify links to engineering design and risk management practices. There are several ways in which this research can be extended on both fronts.

In the first place, the globally increasing availability of hydrological data invites more studies on identifying dependence structures of extremes occurring at multiple spatio-temporal scales. In this respect, the probabilistic NEPvS index for temporal clustering could be applied to identify long-term clustering in other types of hydroclimatic extremes as well, such as floods and droughts (e.g. Zoukos et al., 2018). Moreover, the index can be easily extended to the study of multi-variate extremes in order to characterize multi-scale tail-dependence among different processes. This is essential to the study of spatial hydrological extremes, e.g. pertaining to modelling of joint flooding, and could also be of use to the emerging research field of compound events, i.e. of extreme impacts caused by joint dependent occurrences of less extreme events (Zscheischler et al., 2018).

More research is required to improve the sub-asymptotic modelling of real-world extremes by parsimonious modelling of their parent process. It is essential to extend the investigation of HK-type stochastic models performed herein (Dimitriadis and Koutsoyiannis, 2018) and identify the properties that are the most essential to an efficient reproduction of extremal variability, besides the identified second-order behaviour. In this respect, the recently proposed unbiased estimators of high order moments, known as *k*-moments (Koutsoyiannis, 2019c), offer an alternative way to deal with the uncertainty involved in characterizing the distribution's tail. More insights into the temporal variability of hydroclimatic extremes are expected from their application (Glynis et al., 2020).

Furthermore, there is vast research potential in exploiting the existing physical understanding of the rainfall-generating process to improve seasonal prediction of rainfall and streamflow extremes. There is evidence that weather types and rainfall-producing mechanisms affect the spatial distribution and probability of occurrence of extreme rainfall (Mamassis and Koutsoyiannis, 1996; Mamassis, 1997). On this basis, the formulated frameworks on extreme rainfall and flood seasonality (Iliopoulou et al. 2018; 2019) could be applied to regions dominated by specific weather types and rainfall-producing mechanisms to investigate causal links. The methodology could then be refined to probabilistically update seasonal predictions of extreme events driven by specific weather-types.

In addition, there are still ample grounds for bridging the gap between research in hydrological dependence and practical applications in the wider fields of engineering and finance. For instance, it is less acknowledged that temporal dependence in hydrological extremes may be manifested in the temporal distribution of insurance claims, affecting the risk management practices of the insurance and re-insurance sectors (Serinaldi and Kilsby, 2016b; Papoulakos et al, 2020). In this respect, the observed patterns of extremal clustering could be linked to the temporal variability of actual insurance claims, revisiting the hydrological basis of current insurance practices against hydrological hazards.

More straightforward effects of dependence and temporal patterns of extremes may be sought on various aspects of hydrological design, including the estimation of ombrian (IDF) curves, probabilistic flood mapping, and the estimation of return period and probability of failure of engineering works subjected to water-related uncertainty, among others (Koutsoyiannis, 2020b; Dimitriadis and Koutsoyiannis, 2020; Roy et al. 2018; Serinaldi, 2015; Volpi et al. 2015). The present work calling into question the practice of using projections of rainfall trends for long-term planning (Iliopoulou and Koutsoyiannis, 2020) could also be extended to other hydroclimatic processes, such as floods, seeking for robust alternatives in view of high climatic unpredictability.

Last but not least, a promising avenue to achieve an improvement in holistic risk perception and mitigation is to investigate stochasticity in the evolution of societal vulnerability to hydrological extremes. Diverse types of spatial information on the human-water interface are becoming increasingly available and drive research dealing with human vulnerability indices and the spatio-temporal evolution thereof (Ceola et al., 2014; Sargentis et al., 2020). In this respect, combining historical evidence on the temporal variability of hydrological extremes to that of human vulnerability may provide original insights into the evolution of hydrological risk.

REFERENCES

Aguilar, C., Montanari, A., Polo, M.-J., 2017. Real-time updating of the flood frequency distribution through data assimilation. Hydrology and Earth System Sciences 21, 3687–3700. https://doi.org/10.5194/hess-21-3687-2017

Akaike, H., 1969. Fitting autoregressive models for prediction. Annals of the institute of Statistical Mathematics 21, 243–247.

Akaike, H., 1973. Information Theory and an Extension of the Maximum Likelihood Principle", in B. Petrov and B. Csake (eds), Second International Symposium on Information Theory, Akademai Kiado, Budapest.

Akaike, H., 1974. A new look at the statistical model identification. IEEE transactions on automatic control 19, 716–723.

Allamano, P., Laio, F., Claps, P., 2011. Effects of disregarding seasonality on the distribution of hydrological extremes. Hydrology and Earth System Sciences 15, 3207–3215.

Amrhein, V., Greenland, S., 2018. Remove, rather than redefine, statistical significance. Nature Human Behaviour 2, 4.

Ancona-Navarrete, M.A., Tawn, J.A., 2000. A comparison of methods for estimating the extremal index. Extremes 3, 5–38.

Anderson, D.R., Burnham, K., 2004. Model selection and multi-model inference. Second. NY: Springer-Verlag 63.

Ault, T. R., J. E. Cole, J. T. Overpeck, G. T. Pederson, S. St. George, B. Otto-Bliesner, C. A. Woodhouse, and C. Deser (2013), The continuum of hydroclimate variability in western North America during the last millennium, Journal of Climate, 26(16), 5863-5878.

Balkema, A.A., de Haan, L., 1974. Residual life time at great age. The Annals of probability 792–804.

Baratti, E., Montanari, A., Castellarin, A., Salinas, J.L., Viglione, A., Bezzi, A., 2012. Estimating the flood frequency distribution at seasonal and annual time scales. Hydrology and Earth System Sciences 16, 4651–4660.

Bárdossy, A., Pegram, G.G.S., 2009. Copula based multisite model for daily precipitation simulation. Hydrology and Earth System Sciences 13, 2299–2314.

Barredo, J.I., 2007. Major flood disasters in Europe: 1950–2005. Natural Hazards 42, 125–148.

Beirlant, J., Goegebeur, Y., Segers, J., Teugels, J.L., 2006. Statistics of extremes: theory and applications. John Wiley & Sons.

Beran, J., 1994. Statistics for long-memory processes. CRC press.

Berger, J.O., Bernardo, J.M., 1992. On the development of the reference prior method. Bayesian statistics 4, 35–60.

Berger, J.O., Pericchi, L.R., 1996. The intrinsic Bayes factor for model selection and prediction. Journal of the American Statistical Association 91, 109–122.

Berman, S.M., 1964. Limit theorems for the maximum term in stationary sequences. The Annals of Mathematical Statistics 502–516.

Biasutti, M., 2013. Forced Sahel rainfall trends in the CMIP5 archive. Journal of Geophysical Research: Atmospheres 118, 1613–1623.

Biasutti, M., 2019. Rainfall trends in the African Sahel: Characteristics, processes, and causes. Wiley Interdisciplinary Reviews: Climate Change e591.

Bierkens, M.F.P., van Beek, L.P.H., 2009. Seasonal Predictability of European Discharge: NAO and Hydrological Response Time. Journal of Hydrometeorology 10, 953–968. https://doi.org/10.1175/2009JHM1034.1

Birkhoff, G.D., 1931. Proof of the ergodic theorem. Proceedings of the National Academy of Sciences 17, 656–660.

Blöschl, G., Hall, J., Viglione, A., Perdigão, R.A., Parajka, J., Merz, B., Lun, D., Arheimer, B., Aronica, G.T., Bilibashi, A., 2019. Changing climate both increases and decreases European river floods. Nature 573, 108–111.
Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. Journal of econometrics 31, 307–327.
Bortot, P., Gaetan, C., 2014. A latent process model for temporal extremes. Scandinavian Journal of Statistics 41, 606–621.
Bortot, P., Gaetan, C., 2016. Latent process modelling of threshold exceedances in hourly rainfall series. Journal of agricultural, biological, and environmental statistics 21, 531–547.
Bortot, P., Tawn, J.A., 1998. Models for the extremes of Markov chains. Biometrika 85, 851–867.
Bowers, M.C., Tung, W.W., Gao, J.B., 2012. On the distributions of seasonal river flows: lognormal or power law? Water Resources Research 48.
Box, G.E.P., Jenkins, G.M., 1970. Time series analysis: forecasting and control, 1976. ISBN: 0-8162-1104-3.
Breidt, F.J., Crato, N., De Lima, P., 1998. The detection and estimation of long memory in stochastic volatility. Journal of econometrics 83, 325–348.
Breiman, L., 2001. Statistical modelling: The two cultures (with comments and a rejoinder by the author). Statistical science 16, 199–231.
Brunner, M.I., Bárdossy, A., Furrer, R., 2019. Stochastic simulation of streamflow time series using phase randomization. Hydrology and Earth System Sciences 23, 3175–3187.
Buishand, T.A., 1978. Some remarks on the use of daily rainfall models. Journal of Hydrology 36, 295–308.
Buishand, T.A., Demaré, G.R., 1990. Estimation of the annual maximum distribution from samples of maxima in separate seasons. Stochastic Hydrology and Hydraulics 4, 89–103.
Bunde, A., U. Büntgen, J. Ludescher, J. Luterbacher, and H. von Storch (2013), Is there memory in precipitation?, Nature Climate Change, 3(3), 174-175.
Bunting, A., Dennett, M.D., Elston, J., Milford, J.R., 1976. Rainfall trends in the west African Sahel. Quarterly Journal of the Royal Meteorological Society 102, 59–64.
Burnham, K.P., Anderson, D.R., 2002. Information and likelihood theory: a basis for model selection and inference. Model selection and multimodel inference: a practical information-theoretic approach 49–97.
Burnham, K.P., Anderson, D.R., 2004. Multimodel inference understanding AIC and BIC in model selection. Sociological methods & research 33, 261–304.
Cairncross, A., 1969. Economic forecasting. The Economic Journal 79, 797–812.
Ceola, S., Laio, F., Montanari, A., 2014. Satellite nighttime lights reveal increasing human exposure to floods worldwide. Geophysical Research Letters 41, 7184–7190.
Cervi, F., Blöschl, G., Corsini, A., Borgatti, L., Montanari, A., 2017. Perennial springs provide information to predict low flows in mountain basins. Hydrological Sciences Journal 62, 2469–2481.
Chandler, R., Scott, M., 2011. Statistical methods for trend detection and analysis in the environmental sciences. John Wiley & Sons.
Chen, D., Chen, H.W., 2013. Using the Köppen classification to quantify climate variation and change: An example for 1901–2010. Environmental Development 6, 69–79. https://doi.org/10.1016/j.envdev.2013.03.007
Chen, L., Guo, S., Yan, B., Liu, P., Fang, B., 2010a. A new seasonal design flood method based on bivariate joint distribution of flood magnitude and date of occurrence. Hydrological Sciences Journal 55, 1264–1280. https://doi.org/10.1080/02626667.2010.520564
Chen, L., Guo, S., Yan, B., Liu, P., Fang, B., 2010b. A new seasonal design flood method based on bivariate joint distribution of flood magnitude and date of occurrence.

Hydrological Sciences Journal–Journal des Sciences Hydrologiques 55, 1264–1280.
Chen, L., Singh, V.P., Guo, S., Fang, B., Liu, P., 2013. A new method for identification of flood seasons using directional statistics. Hydrological Sciences Journal 58, 28–40.
Chernick, M.R., 1981. A limit theorem for the maximum of autoregressive processes with uniform marginal distributions. The Annals of Probability 9, 145–149.
Chiew, F.H.S., Zhou, S.L., McMahon, T.A., 2003. Use of seasonal streamflow forecasts in water resources management. Journal of Hydrology 270, 135–144. https://doi.org/10.1016/S0022-1694(02)00292-5
Cohn, T.A., Lins, H.F., 2005. Nature's style: Naturally trendy. Geophysical Research Letters 32.
Coles, S., Bawa, J., Trenner, L., Dorazio, P., 2001. An introduction to statistical modelling of extreme values. Springer.
Coles, S., Heffernan, J., Tawn, J., 1999. Dependence measures for extreme value analyses. Extremes 2, 339–365.
Coles, S., Pericchi, L.R., Sisson, S., 2003. A fully probabilistic approach to extreme rainfall modelling. Journal of Hydrology 273, 35–50.
Conover, W.J., 1971. Practical Nonparametric Statistics. New York: John Fliley and Sons. Inc.
Conover, W.J., 1980. Practical nonparametric statistics, 2n edition, John Wiley and Sons, New York.
Cowpertwait, P.S., 1998. A Poisson-cluster model of rainfall: some high-order moments and extreme values. Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences 454, 885–898.
Cox, D.R., Isham, V., 1980. Point processes. CRC Press.
Craig, R.K., 2010. Stationarity is dead-long live transformation: five principles for climate change adaptation law. Harv. Envtl. L. Rev. 34, 9.
Cunderlik, J.M., Burn, D.H., 2002. Analysis of the linkage between rain and flood regime and its application to regional flood frequency estimation. Journal of Hydrology 261, 115–131.
Cunderlik, J.M., Ouarda, T.B., Bobée, B., 2004a. Determination of flood seasonality from hydrological records/Détermination de la saisonnalité des crues à partir de séries hydrologiques. Hydrological Sciences Journal 49.
Cunderlik, J.M., Ouarda, T.B., Bobée, B., 2004b. On the objective identification of flood seasons. Water Resources Research 40.
Davis, R.A., Resnick, S.I., 1989. Basic properties and prediction of max-ARMA processes. Advances in applied probability 21, 781–803.
Davison, A.C., Padoan, S.A., Ribatet, M., 2012. Statistical Modelling of Spatial Extremes. Statist. Sci. 27, 161–186. https://doi.org/10.1214/11-STS376
de Haan, L., 1984. A spectral representation for max-stable processes. The annals of probability 12, 1194–1204.
de Haan, L., 1994. Extreme value statistics, in: Extreme Value Theory and Applications. Springer, pp. 93–122.
de Haan, L., Resnick, S., Rootzén, H., de Vries, C., 1989. Extremal behaviour of solutions to a stochastic difference equation, with applications to ARCH processes. Stochastic Processes and their Applications 213–224.
De Luca, D.L., Petroselli, A., Galasso, L., 2019. Modelling climate changes with stationary models: is it possible or is it a paradox?, in: International Conference on Numerical Computations: Theory and Algorithms. Springer, pp. 84–96.
De Martonne, E.M., 1926. L'indice d'aridité. Bulletin de l'Association de géographes français 3, 3–5.
De Michele, C., 2019. Advances in Deriving the Exact Distribution of Maximum Annual Daily Precipitation. Water 11, 2322.

De Michele, C., Salvadori, G., 2003. A generalized Pareto intensity-duration model of storm rainfall exploiting 2-copulas. Journal of Geophysical Research: Atmospheres 108.
Degefu, M.A., Alamirew, T., Zeleke, G., Bewket, W., 2019. Detection of trends in hydrological extremes for Ethiopian watersheds, 1975–2010. Regional Environmental Change 1–11.
Dhakal, N., Jain, S., Gray, A., Dandy, M., Stancioff, E., 2015. Nonstationarity in seasonality of extreme precipitation: A nonparametric circular statistical approach and its application. Water Resources Research 51, 4499–4515.
Dijk, A.I., Peña-Arancibia, J.L., Wood, E.F., Sheffield, J., Beck, H.E., 2013. Global analysis of seasonal streamflow predictability using an ensemble prediction system and observations from 6192 small catchments worldwide. Water Resources Research 49, 2729–2746.
Dimitriadis, P., 2017. Hurst-Kolmogorov dynamics in hydrometeorological processes and in the microscale of turbulence, PhD thesis, Department of Water Resources and Environmental Engineering – National Technical University of Athens.
Dimitriadis, P., Koutsoyiannis, D., 2015. Climacogram versus autocovariance and power spectrum in stochastic modelling for Markovian and Hurst–Kolmogorov processes. Stochastic environmental research and risk assessment 29, 1649–1669.
Dimitriadis, P., Koutsoyiannis, D., Tzouka, K., 2016. Predictability in dice motion: how does it differ from hydro-meteorological processes? Hydrological Sciences Journal 61, 1611–1622.
Dimitriadis, P., Koutsoyiannis, D., 2018. Stochastic synthesis approximating any process dependence and distribution. Stoch Environ Res Risk Assess 32, 1493–1515. https://doi.org/10.1007/s00477-018-1540-2
Dimitriadis, P., Koutsoyiannis, D., 2020. The mode of the climacogram estimator for a Gaussian Hurst-Kolmogorov process. Journal of Hydroinformatics 22, 160–169.
Duan, Q., Ajami, N.K., Gao, X., Sorooshian, S., 2007. Multi-model ensemble hydrologic prediction using Bayesian model averaging. Advances in Water Resources 30, 1371–1386.
Durrans, S.R., Eiffe, M.A., Thomas Jr, W.O., Goranflo, H.M., 2003. Joint seasonal/annual flood frequency analysis. Journal of Hydrologic Engineering 8, 181–189.
Efstratiadis, A., Koussis, A.D., Koutsoyiannis, D., Mamassis, N., 2014. Flood design recipes vs. reality: can predictions for ungauged basins be trusted? Natural Hazards and Earth System Sciences 14, 1417–1428.
Eichner, J.F., Kantelhardt, J.W., Bunde, A., Havlin, S., 2011. The statistics of return intervals, maxima, and centennial events under the influence of long-term correlations, in: In Extremis. Springer, pp. 2–43.
Embrechts, P., 2009. Copulas: A personal view. Journal of Risk and Insurance 76, 639–650.
Embrechts, P., Kluppelberg, C., Mikosch, T., 1999. Modelling extremal events. British actuarial journal 5, 465–465.
Embrechts, P., Klüppelberg, C., Mikosch, T., 2013. Modelling extremal events: for insurance and finance. Springer Science & Business Media.
Embrechts, P., McNeil, A., Straumann, D., 2002. Correlation and dependence in risk management: properties and pitfalls. Risk management: value at risk and beyond 1, 176–223.
Embrechts, P., Samorodnitsky, G., Dacorogna, M.M., Muller, U.A., 1996. How heavy are the tails of a stationary HARCH (k) process? a study of the moments. Cornell University Operations Research and Industrial Engineering.

Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica: Journal of the Econometric Society 987–1007.
Fang, B., Guo, S., Wang, S., Liu, P., Xiao, Y., 2007. Non-identical models for seasonal flood frequency analysis. Hydrological Sciences Journal 52, 974–991. https://doi.org/10.1623/hysj.52.5.974
Fatichi, S., Barbosa, S.M., Caporali, E., Silva, M.E., 2009. Deterministic versus stochastic trends: Detection and challenges. Journal of Geophysical Research: Atmospheres 114.
Favre, A.-C., El Adlouni, S., Perreault, L., Thiémonge, N., Bobée, B., 2004. Multivariate hydrological frequency analysis using copulas. Water resources research 40.
FEMA, Guidance for Flood Risk Analysis and Mapping, 2016. https://www.fema.gov/media-library-data/1481729383463-f7e8bfce74b8f86f97a4bf03776d5f58/Coastal_Flood_Frequency_and_Extreme_Value_Analysis_Guidance_Nov_2016.pdf?fbclid=IwAR0jg7eBebLyUVBKsYuovz5bcJoYSrk-Tl5UWI8Eetv0wPog3R-Q4Iuk4AM
Ferro, C.A., Segers, J., 2003. Inference for clusters of extreme values. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 65, 545–556.
Finkenstadt, B., Rootzén, H., 2003. Extreme values in finance, telecommunications, and the environment. CRC Press.
Fisher, R.A., Tippett, L.H.C., 1928. Limiting forms of the frequency distribution of the largest or smallest member of a sample, in: Mathematical Proceedings of the Cambridge Philosophical Society. Cambridge University Press, pp. 180–190.
Folton, N., Martin, E., Arnaud, P., L'Hermite, P., Tolsa, M., 2019. A 50-year analysis of hydrological trends and processes in a Mediterranean catchment. Hydrology and Earth System Sciences 23, 2699–2714.
Fraedrich, K., and C. Larnder, 1993, Scaling regimes of composite rainfall time series, Tellus A, 45(4), 289-298.
Fraedrich, K., and R. Blender, 2003, Scaling of atmosphere and ocean temperature correlations in observations and climate models, Physical Review Letters, 90(10), 108501.
Fréchet, M., 1927. Sur la loi de probabilité de l'écart maximum. Ann. Soc. Math. Polon. 6, 93–116.
Fréchet, M., 1951. Sur les tableaux de corrélation dont les marges sont données. Ann. Univ. Lyon, 3^ e serie, Sciences, Sect. A 14, 53–77.
Fuller, W.E., 1914. Flood Flows—Trans. A. SCE 565.
Gabriel, K.R., 1971. The biplot graphic display of matrices with application to principal component analysis. Biometrika 58, 453–467.
Galambos, J., 1994. Extreme value theory for applications, in: Extreme Value Theory and Applications. Springer, pp. 1–14.
Gauch Jr, H.G.., 2003. Scientific method in practice. Cambridge University Press.
Genest, C., MacKay, R.J., 1986. Copules archimédiennes et families de lois bidimensionnelles dont les marges sont données. Canadian Journal of Statistics 14, 145–159.
Georgakakos, K.P., Seo, D.-J., Gupta, H., Schaake, J., Butts, M.B., 2004. Towards the characterization of streamflow simulation uncertainty through multimodel ensembles. Journal of Hydrology, The Distributed Model Intercomparison Project (DMIP) 298, 222–241. https://doi.org/10.1016/j.jhydrol.2004.03.037
Giraitis, L., Robinson, P.M., Surgailis, D., 2000. A model for long memory conditional heteroscedasticity. Annals of Applied Probability 1002–1024.
Glynis, K.-G., Iliopoulou, T., Dimitriadis, P., Koutsoyiannis, D., 2020. Stochastic investigation of daily air temperature extremes from a global ground station network, Stochastic Environmental Research and Risk Assessment, in review.
Gnedenko, B., 1943. Sur la distribution limite du terme maximum d'une serie aleatoire. Annals of mathematics 423–453.

Gneiting, T., Schlather, M., 2004. Stochastic models that separate fractal dimension and the Hurst effect. SIAM review 46, 269–282.
Gneiting, T., Ševčíková, H., Percival, D.B., 2012. Estimators of fractal dimension: Assessing the roughness of time series and spatial data. Statistical Science 247–277.
Golian, S., Saghafian, B., Maknoon, R., 2010. Derivation of Probabilistic Thresholds of Spatially Distributed Rainfall for Flood Forecasting. Water Resour Manage 24, 3547–3559. https://doi.org/10.1007/s11269-010-9619-7
Goulianou, T., Papoulakos, K., Iliopoulou, T., Dimitriadis, P., Koutsoyiannis, D., 2019. Stochastic characteristics of flood impacts for agricultural insurance practices. European Geosciences Union General Assembly 21.
Gower, J.C., Hand, D.J., 1995. Biplots. CRC Press.
Granger, C.W., Joyeux, R., 1980. An introduction to long-memory time series models and fractional differencing. Journal of time series analysis 1, 15–29.
Grimaldi, S., Serinaldi, F., 2006. Asymmetric copula in multivariate flood frequency analysis. Advances in Water Resources 29, 1155–1167.
Gringorten, I.L., 1966. A stochastic model of the frequency and duration of weather events. Journal of Applied Meteorology 5, 606–624.
Gudmundsson, L., Tallaksen, L.M., Stahl, K., Fleig, A.K., 2011. Low-frequency variability of European runoff. Hydrology and Earth System Sciences 15, 2853.
Gumbel, E.J., 1941. The return period of flood flows. The annals of mathematical statistics 12, 163–190.
Gumbel, E.J., 1958. Statistics of Extremes, Columbia Univ. Press, New York 201.
Gupta, V.K., Waymire, E.C., 1993. A statistical analysis of mesoscale rainfall as a random cascade. Journal of Applied Meteorology 32, 251–267.
Gustard, A., Demuth, S., others, 2009. Manual on low-flow estimation and prediction. Opera.
Hall, J., Arheimer, B., Borga, M., Brázdil, R., Claps, P., Kiss, A., Kjeldsen, T.R., Kriauciuniene, J., Kundzewicz, Z.W., Lang, M., 2014. Understanding flood regime changes in Europe: A state of the art assessment. Hydrology and earth system sciences 18, 2735–2772.
Hall, P., Peng, L., Yao, Q., 2002. Moving-maximum models for extrema of time series. Journal of Statistical Planning and Inference 103, 51–63.
Hamed, K.H., 2008. Trend detection in hydrologic data: the Mann–Kendall trend test under the scaling hypothesis. Journal of hydrology 349, 350–363.
Harman, C.J., Troch, P.A., Sivapalan, M., 2011. Functional model of water balance variability at the catchment scale: 2. Elasticity of fast and slow runoff components to precipitation change in the continental United States. Water Resources Research 47.
Haslett, J., and A. E. Raftery, 1989, Space-time modelling with long-memory dependence: Assessing Ireland's wind power resource, Applied Statistics, 1-50.
Hastie, T., Tibshirani, R., Friedman, J., Franklin, J., 2005. The elements of statistical learning: data mining, inference and prediction. The Mathematical Intelligencer 27, 83–85.
Haylock, M., Nicholls, N., 2000. Trends in extreme rainfall indices for an updated high quality data set for Australia, 1910–1998. International Journal of Climatology: A Journal of the Royal Meteorological Society 20, 1533–1541.
Hinkley, D.V., 1970. Inference about the change-point in a sequence of random variables.
Hirpa, F.A., Gebremichael, M., Over, T.M., 2010. River flow fluctuation analysis: Effect of watershed area. Water Resources Research 46.
Hirsch, R.M., Ryberg, K.R., 2012. Has the magnitude of floods across the USA changed with global CO2 levels? Hydrological Sciences Journal 57, 1–9.
Hirschboeck, K.K., 1988. Flood hydroclimatology. Flood geomorphology 27, 49.

Hocking, R.R., 1976. A Biometrics invited paper. The analysis and selection of variables in linear regression. Biometrics 32, 1–49.
Hoeffding, W., 1940. Scale-invariant correlation measures for discontinuous distributions, In NI Fisher and PK Sen (eds.), The Collected Works of Wassily Hoeffding. New York: Springer-Verlag.
Hosking, J.R.M., 1981. Fractional differencing. Biometrika 68 165–176. Mathematical Reviews (MathSciNet): MR614953 Zentralblatt MATH 464.
Houghton, J.T., Jenkins, G.J., Ephraums, J.J., 1991. Climate change.
Hurst, H.E., 1951. Long-term storage capacity of reservoirs. Trans. Amer. Soc. Civil Eng. 116, 770–808.
Ibragimov, R., Lentzas, G., 2008. Copulas and long memory. Harvard Institute of Economic Research Discussion Paper.
Iglehart, L.D., 1969. Diffusion approximations in collective risk theory. Journal of Applied Probability 6, 285–292.
Iliopoulou, T. and Koutsoyiannis, D., 2020. Projecting the future of rainfall extremes: better classic than trendy. Journal of Hydrology, p.125005.
Iliopoulou, T., Koutsoyiannis, D., 2019. Revealing hidden persistence in maximum rainfall records. Hydrological Sciences Journal 64, 1673–1689.
Iliopoulou, T., Aguilar, C., Arheimer, B., Bermúdez, M., Bezak, N., Ficchì, A., Koutsoyiannis, D., Parajka, J., Polo, M.J., Thirel, G. and Montanari, A., 2019. A large sample analysis of European rivers on seasonal river flow correlation and its physical drivers.
Iliopoulou, T., Koutsoyiannis, D., Montanari, A., 2018a. Characterizing and modelling seasonality in extreme rainfall. Water Resources Research 54, 6242–6258.
Iliopoulou, T., Papalexiou, S.M., Markonis, Y., Koutsoyiannis, D., 2018b. Revisiting long-range dependence in annual precipitation. Journal of Hydrology 556, 891–900.
Inoue, A., Kilian, L., 2005. In-sample or out-of-sample tests of predictability: Which one should we use? Econometric Reviews 23, 371–402.
Inoue, A., Kilian, L., 2006. On the selection of forecasting models. Journal of Econometrics 130, 273–306.
IPCC: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK and New York, NY, 1535 pp. http://www.climatechange2013.org/report/ (accessed 2020-02-14), 2013.
Jenkinson, A.F., 1955. The frequency distribution of the annual maximum (or minimum) values of meteorological elements. Quarterly Journal of the Royal Meteorological Society 81, 158–171.
Jhun, J.G., Moon, B.K., 1997. Restorations and analyses of rainfall amount observed by Chukwookee. J. Korean Meteor. Soc 33, 691–707.
Joe, H., 2014. Dependence modelling with copulas. Chapman and Hall/CRC.
Jolliffe, I., 2002. Principal component analysis. Wiley Online Library.
Juncosa, M.L., 1949. The asymptotic behaviour of the minimum in a sequence of random variables. Duke Mathematical Journal 16, 609–618.
Kaczmarska, J., Isham, V., Onof, C., 2014. Point process models for fine-resolution rainfall. Hydrological Sciences Journal 59, 1972–1991.
Kantelhardt, J. W., E. Koscielny-Bunde, D. Rybski, P. Braun, A. Bunde, and S. Havlin , 2006. Long-term persistence and multifractality of precipitation and river runoff records, Journal of Geophysical Research: Atmospheres (1984–2012), 111(D1).
Katz, R.W., Parlange, M.B., 1998. Overdispersion phenomenon in stochastic modelling of precipitation. Journal of Climate 11, 591–601.

Kellermann, P., Schönberger, C., Thieken, A.H., 2019. Large-scale application of the flood damage model RAilway Infrastructure Loss (RAIL).
Kellogg, W.W., 2019. Climate change and society: consequences of increasing atmospheric carbon dioxide. Routledge.
Kelly, K.S., Krzysztofowicz, R., 1997. A bivariate meta-Gaussian density for use in hydrology. Stochastic Hydrology and hydraulics 11, 17–31.
Kendall, M.G., 1938. A New Measure of Rank Correlation. Biometrika 30, 81–93. https://doi.org/10.2307/2332226
Kesten, H., 1973. Random difference equations and renewal theory for products of random matrices. Acta Mathematica 131, 207–248.
Khan, N., Pour, S.H., Shahid, S., Ismail, T., Ahmed, K., Chung, E.-S., Nawaz, N., Wang, X., 2019. Spatial distribution of secular trends in rainfall indices of Peninsular Malaysia in the presence of long-term persistence. Meteorological Applications.
Khintchine, A., 1933. Zu birkhoffs lösung des ergodenproblems. Mathematische Annalen 107, 485–488.
Kim, D., Olivera, F., Cho, H., 2013. Effect of the inter-annual variability of rainfall statistics on stochastically generated rainfall time series: part 1. Impact on peak and extreme rainfall values. Stochastic environmental research and risk assessment 27, 1601–1610.
Kim, D., Onof, C., 2020. A stochastic rainfall model that can reproduce important rainfall properties across the timescales from several minutes to a decade. Journal of Hydrology 589, 125150. https://doi.org/10.1016/j.jhydrol.2020.125150
Kirkpatrick II, C.D., Dahlquist, J.A., 2010. Technical analysis: the complete resource for financial market technicians. FT press.
Klein Tank, A.M.G., Wijngaard, J.B., Können, G.P., Böhm, R., Demarée, G., Gocheva, A., Mileta, M., Pashiardis, S., Hejkrlik, L., Kern-Hansen, C., 2002. Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment. International journal of climatology 22, 1441–1453.
Klemeš, V., 1986. Operational testing of hydrological simulation models. Hydrological Sciences Journal 31, 13–24.
Klüppelberg, C., Lindner, A., Maller, R., 2004. A continuous-time GARCH process driven by a Lévy process: stationarity and second-order behaviour. Journal of Applied Probability 41, 601–622.
Koenker, R., Hallock, K.F., 2001. Quantile regression. Journal of economic perspectives 15, 143–156.
Kolmogorov, A.N., 1931. Ueber die analytischen Methoden in der Wahrscheinlichkeitstheorie. Math. Ann. 104, 415–458. (English translation: On analytical methods in probability theory, In: Kolmogorov, A.N., Selected Works of A. N. Kolmogorov - Volume 2, Probability Theory and Mathematical Statistics, ed. By A.N. Shiryayev, Kluwer, Dordrecht, The Netherlands, 62-108, 1992).
Kolmogorov, A.N., 1933. Grundbegriffe der Wahrscheinlichkeitsrechnung, Ergebnisse der Math. (2), Berlin (2nd English Edition: Foundations of the Theory of Probability, 84 pp. Chelsea Publishing Company, New York, 1956).
Kolmogorov, A.N., 1940. Wiener spirals and some other interesting curves in a Hilbert space. Dokl. Akad. Nauk SSSR, 26, 115-118. (English edition: Kolmogorov, A.N., 1991, Selected Works of A. N. Kolmogorov - Volume 1, Mathematics and Mechanics, ed. by Tikhomirov, V.M., Kluwer, Dordrecht, The Netherlands, 324-326).
Kolmogorov, A.N., 1947. Statistical theory of oscillations with continuous spectrum. Collected papers on the 30th anniversary of the Great October Socialist Revolution, Vol. 1, Akad. Nauk SSSR, Moscow-Leningrad, 242-252. (English edition: Kolmogorov, A.N., 1992. Selected Works of A. N. Kolmogorov

Volume 2, Probability Theory and Mathematical Statistics, ed. by Shiryayev, A.N., Kluwer, Dordrecht, The Netherlands, 321-330).
Koskinas, A., Tegos, A., Tsira, P., Dimitriadis, P., Iliopoulou, T., Papanicolaou, P., Koutsoyiannis, D., Williamson, T., 2019. Insights into the oroville dam 2017 spillway incident. Geosciences 9, 37.
Kossieris, P., Makropoulos, C., Onof, C., Koutsoyiannis, D., 2018. A rainfall disaggregation scheme for sub-hourly time scales: Coupling a Bartlett-Lewis based model with adjusting procedures. Journal of Hydrology 556, 980–992.
Kottegoda, N.T., Rosso, R., 2008. Applied statistics for civil and environmental engineers. Blackwell Malden, MA.
Kottek, M., Grieser, J., Beck, C., Rudolf, B., Rubel, F., 2006. World Map of the Köppen-Geiger climate classification updated. Meteorologische Zeitschrift 259–263. https://doi.org/10.1127/0941-2948/2006/0130
Kotz, S., Nadarajah, S., 2000. Extreme value distributions: theory and applications. World Scientific.
Koutsoyiannis, D., 1999. A probabilistic view of Hershfield's method for estimating probable maximum precipitation. Water resources research 35, 1313–1322.
Koutsoyiannis, D., 2000. A generalized mathematical framework for stochastic simulation and forecast of hydrologic time series. Water Resources Research 36, 1519–1533.
Koutsoyiannis, D., 2002. The Hurst phenomenon and fractional Gaussian noise made easy. Hydrological Sciences Journal 47, 573–595.
Koutsoyiannis, D., 2003. Climate change, the Hurst phenomenon, and hydrological statistics. Hydrological Sciences Journal 48, 3–24.
Koutsoyiannis, D., 2004a. Statistics of extremes and estimation of extreme rainfall: II. Empirical investigation of long rainfall records/Statistiques de valeurs extrêmes et estimation de précipitations extrêmes: II. Recherche empirique sur de longues séries de précipitations. Hydrological sciences journal 49.
Koutsoyiannis, D., 2004b. Statistics of extremes and estimation of extreme rainfall: I. Theoretical investigation/Statistiques de valeurs extrêmes et estimation de précipitations extrêmes: I. Recherche théorique. Hydrological sciences journal 49.
Koutsoyiannis, D., 2005. Uncertainty, entropy, scaling and hydrological stochastics. 2. Time dependence of hydrological processes and time scaling/Incertitude, entropie, effet d'échelle et propriétés stochastiques hydrologiques. 2. Dépendance temporelle des processus hydrologiques et échelle temporelle. Hydrological Sciences Journal 50.
Koutsoyiannis, D., 2006. An entropic-stochastic representation of rainfall intermittency: The origin of clustering and persistence. Water Resources Research 42.
Koutsoyiannis, D., 2010. HESS Opinions" A random walk on water". Hydrology and Earth System Sciences 14, 585–601.
Koutsoyiannis, D., 2011a. Hurst–Kolmogorov dynamics as a result of extremal entropy production. Physica A: Statistical Mechanics and its Applications 390, 1424–1432.
Koutsoyiannis, D., 2011b. Hurst-Kolmogorov Dynamics and Uncertainty. JAWRA Journal of the American Water Resources Association 47, 481–495.
Koutsoyiannis, D., 2014. Random musings on stochastics (Lorenz Lecture), in: AGU 2014 Fall Meeting.
Koutsoyiannis, D., 2016. Generic and parsimonious stochastic modelling for hydrology and beyond. Hydrological Sciences Journal 61, 225–244.
Koutsoyiannis, D., 2017. Entropy production in stochastics. Entropy 19, 581.
Koutsoyiannis, D., 2019a. Simple stochastic simulation of time irreversible and reversible processes. Hydrological Sciences Journal.

Koutsoyiannis, D., 2019b. Time's arrow in stochastic characterization and simulation of atmospheric and hydrological processes. Hydrological Sciences Journal.
Koutsoyiannis, D., 2019c. Knowable moments for high-order stochastic characterization and modelling of hydrological processes. Hydrological Sciences Journal 0, 1–15. https://doi.org/10.1080/02626667.2018.1556794
Koutsoyiannis, D., 2020a. Revisiting global hydrological cycle: Is it intensifying? Hydrology and Earth Systems Science, 24, 3899–3932, https://doi.org/10.5194/hess-24-3899-2020, 2020.
Koutsoyiannis, D., 2020b. Stochastics of Hydroclimatic Extremes - A Cool Look at Risk National Technical University of Athens.
Koutsoyiannis, D., and A. Montanari, 2007, Statistical analysis of hydroclimatic time series: Uncertainty and insights, Water Resources Research, 43(5).
Koutsoyiannis, D., Montanari, A., 2015a. Negligent killing of scientific concepts: the stationarity case. Hydrological Sciences Journal 60, 1174–1183.
Koutsoyiannis, D., Dimitriadis, P., Lombardo, F., Stevens, S., 2018. From fractals to stochastics: Seeking theoretical consistency in analysis of geophysical data, in: Advances in Nonlinear Geosciences. Springer, pp. 237–278.
Koutsoyiannis, D., Kozonis, D., Manetas, A., 1998. A mathematical framework for studying rainfall intensity-duration-frequency relationships. Journal of Hydrology 206, 118–135.
Koutsoyiannis, D., Mamassis, N., Efstratiadis, A., Zarkadoulas, N., Markonis, Y., 2012. Floods in Greece, in: Kundzewicz, Z.W. (Ed.), Changes of Flood Risk in Europe. pp. 238–256.
Koutsoyiannis, D., Manetas, A., 1996. Simple disaggregation by accurate adjusting procedures. Water Resources Research 32, 2105–2117.
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. Hydrological Sciences Journal 53, 142–164.
Krichak, S.O., Kishcha, P., Alpert, P., 2002. Decadal trends of main Eurasian oscillations and the Eastern Mediterranean precipitation. Theoretical and Applied Climatology 72, 209–220.
Krishnamurthy, L., Krishnamurthy, V., 2016. Teleconnections of Indian monsoon rainfall with AMO and Atlantic tripole. Climate dynamics 46, 2269–2285.
Kumar, V., Jain, S.K., Singh, Y., 2010. Analysis of long-term rainfall trends in India. Hydrological Sciences Journal–Journal des Sciences Hydrologiques 55, 484–496.
Kutiel, H., Trigo, R.M., 2014. The rainfall regime in Lisbon in the last 150 years. Theoretical and applied climatology 118, 387–403.
Laio, F., Di Baldassarre, G., Montanari, A., 2009. Model selection techniques for the frequency analysis of hydrological extremes. Water Resources Research 45.
Lang, M., Ouarda, T., Bobée, B., 1999. Towards operational guidelines for over-threshold modelling. Journal of hydrology 225, 103–117.
Langousis, A., Veneziano, D., 2007. Intensity-duration-frequency curves from scaling representations of rainfall. Water Resources Research 43.
Latif, M., Collins, M., Pohlmann, H., Keenlyside, N., 2006. A review of predictability studies of Atlantic sector climate on decadal time scales. Journal of Climate 19, 5971–5987.
Laux, P., Vogl, S., Qiu, W., Knoche, H.R., Kunstmann, H., 2011. Copula-based statistical refinement of precipitation in RCM simulations over complex terrain. Hydrology and Earth System Sciences 15, 2401–2419.
Laux, P., Wagner, S., Wagner, A., Jacobeit, J., Bárdossy, A., Kunstmann, H., 2009. Modelling daily precipitation features in the Volta Basin of West Africa. International Journal of Climatology: A Journal of the Royal Meteorological Society 29, 937–954.

Lavergnat, J., 2016. On the generation of colored non-Gaussian time sequences.
Leadbetter, M.R., 1974. On extreme values in stationary sequences. Probability theory and related fields 28, 289–303.
Leadbetter, M.R., 1983. Extremes and local dependence in stationary sequences. Probability Theory and Related Fields 65, 291–306.
Leadbetter, M.R., Lindgren, G., Rootzén, H., 2012. Extremes and related properties of random sequences and processes. Springer Science & Business Media.
Leadbetter, M.R., Rootzen, H., 1988. Extremal theory for stochastic processes. The Annals of Probability 431–478.
Lecce, S.A., 2000. Seasonality of flooding in North Carolina. Southeastern Geographer 40, 168–175.
Ledford, A.W., Tawn, J.A., 1996. Statistics for near independence in multivariate extreme values. Biometrika 83, 169–187.
Ledford, A.W., Tawn, J.A., 1997. Modelling dependence within joint tail regions. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 59, 475–499.
Ledford, A.W., Tawn, J.A., 2003. Diagnostics for dependence within time series extremes. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 65, 521–543.
Lee, D., Ward, P., Block, P., 2015. Defining high-flow seasons using temporal streamflow patterns from a global model. Hydrology and Earth System Sciences 19, 4689.
Lee, J.-J., Kwon, H.-H., Kim, T.-W., 2012. Spatio-temporal analysis of extreme precipitation regimes across South Korea and its application to regionalization. Journal of hydro-environment research 6, 101–110.
Lee, T., Ouarda, T., 2010. Long-term prediction of precipitation and hydrologic extremes with nonstationary oscillation processes. Journal of Geophysical Research: Atmospheres 115.
Lee, T., Salas, J.D., 2011. Copula-based stochastic simulation of hydrological data applied to Nile River flows. Hydrology Research 42, 318–330.
Li, J., Thyer, M., Lambert, M., Kuzera, G., Metcalfe, A., 2016. Incorporating seasonality into event-based joint probability methods for predicting flood frequency: A hybrid causative event approach. Journal of Hydrology 533, 40–52. https://doi.org/10.1016/j.jhydrol.2015.11.038
Lombardo, F., Napolitano, F., Russo, F., Koutsoyiannis, D., 2019. On the exact distribution of correlated extremes in hydrology. Water Resources Research.
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. Hydrology and Earth System Sciences 18, 243–255.
Lorenc, A.C., 1986. Analysis methods for numerical weather prediction. Quarterly Journal of the Royal Meteorological Society 112, 1177–1194.
Loynes, R.M., 1965. Extreme values in uniformly mixing stationary stochastic processes. The Annals of Mathematical Statistics 36, 993–999.
Mackey, M.C., 2011. Time's arrow: The origins of thermodynamic behaviour. Courier Corporation.
MacQueen, J., others, 1967. Some methods for classification and analysis of multivariate observations, in: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability. Oakland, CA, USA., pp. 281–297.
Mahanama, S., Livneh, B., Koster, R., Lettenmaier, D., Reichle, R., 2011. Soil Moisture, Snow, and Seasonal Streamflow Forecasts in the United States. J. Hydrometeor. 13, 189–203. https://doi.org/10.1175/JHM-D-11-046.1
Mamassis, 1997, Rainfall analysis by weather type, PhD thesis, Department of Water Resources, Hydraulic and Maritime Engineering – National Technical University of Athens, Athens.

Mamassis, N. and Koutsoyiannis, D., 1996. Influence of atmospheric circulation types on space-time distribution of intense rainfall. Journal of Geophysical Research: Atmospheres, 101(D21), pp.26267-26276.
Mandelbrot, B., and J. Wallis, 1969, Some long-run properties of geophysical records, Water resources research, 5(2), 321-340.
Mandelbrot, B.B., 1971. A fast fractional Gaussian noise generator. Water Resources Research 7, 543–553.
Mandelbrot, B.B., 1974. Intermittent turbulence in self-similar cascades: divergence of high moments and dimension of the carrier. Journal of Fluid Mechanics 62, 331–358.
Mandelbrot, B.B., 1983. The fractal geometry of nature. WH freeman New York.
Manolis, G.T., Papoulakos, K., Iliopoulou, T., Dimitriadis, P., Tsaknias, D., Koutsoyiannis, D., 2020. Clustering mechanisms of flood occurrence; modelling and relevance to insurance practices. https://doi.org/10.5194/egusphere-egu2020-9357
Marani, M., 2003. On the correlation structure of continuous and discrete point rainfall. Water Resources Research 39.
Marani, M., Zanetti, S., 2015. Long-term oscillations in rainfall extremes in a 268 year daily time series. Water Resources Research 51, 639–647.
Markonis, Y. , 2015. Stochastic Investigation of Large-Scale Hydroclimatic Correlations over the Mediterranean, PhD thesis, Department of Water Resources and Environmental Engineering – National Technical University of Athens.
Markonis, Y., and D. Koutsoyiannis, 2013. Climatic variability over time scales spanning nine orders of magnitude: Connecting Milankovitch cycles with Hurst–Kolmogorov dynamics, Surveys in Geophysics, 34(2), 181-207.
Markonis, Y., and D. Koutsoyiannis, 2015. Scale-dependence of persistence in precipitation records., Nature Climate Change, DOI:10.1038/NCLIMATE2894.
Markonis, Y., Koutsoyiannis, D., 2016. Scale-dependence of persistence in precipitation records. Nature Climate Change 6, 399.
Markonis, Y., Moustakis, Y., Nasika, C., Sychova, P., Dimitriadis, P., Hanel, M., Máca, P., Papalexiou, S.M., 2018. Global estimation of long-term persistence in annual river runoff. Advances in Water Resources 113, 1–12. https://doi.org/10.1016/j.advwatres.2018.01.003
Marshak, A., Davis, A., Cahalan, R., Wiscombe, W., 1994. Bounded cascade models as nonstationary multifractals. Physical Review E 49, 55.
Mascaro, G., 2018. On the distributions of annual and seasonal daily rainfall extremes in central Arizona and their spatial variability. Journal of Hydrology 559, 266–281. https://doi.org/10.1016/j.jhydrol.2018.02.011
Mathai, J., Mujumdar, P.P., 2019. Multisite Daily Streamflow Simulation With Time Irreversibility. Water Resources Research 55, 9334–9350.
McCarl, B.A., Villavicencio, X., Wu, X., 2008. Climate change and future analysis: is stationarity dying? American Journal of Agricultural Economics 90, 1241–1247.
McKitrick, R., Christy, J., 2019. Assessing Changes in US Regional Precipitation on Multiple Time Scales. Journal of Hydrology 124074.
Menabde, M., Harris, D., Seed, A., Austin, G., Stow, D., 1997. Multiscaling properties of rainfall and bounded random cascades. Water Resources Research 33, 2823–2830.
Menne, M.J., Durre, I., Vose, R.S., Gleason, B.E., Houston, T.G., 2012. An Overview of the Global Historical Climatology Network-Daily Database. J. Atmos. Oceanic Technol. 29, 897–910. https://doi.org/10.1175/JTECH-D-11-00103.1
Merz, B., Nguyen, V.D., Vorogushyn, S., 2016. Temporal clustering of floods in Germany: Do flood-rich and flood-poor periods exist? Journal of Hydrology 541, 824–838.
Mikosch, T., 2006. Copulas: Tales and facts–rejoinder. Extremes 9, 55–62.

Mikosch, T., Starica, C., 2002. Long-range dependence effects and ARCH modelling. Theory and applications of long-range dependence 439–459.
Milly, P.C., Betancourt, J., Falkenmark, M., Hirsch, R.M., Kundzewicz, Z.W., Lettenmaier, D.P., Stouffer, R.J., 2008. Stationarity is dead: Whither water management? Science 319, 573–574.
Milly, P.C., Betancourt, J., Falkenmark, M., Hirsch, R.M., Kundzewicz, Z.W., Lettenmaier, D.P., Stouffer, R.J., Dettinger, M.D., Krysanova, V., 2015. On critiques of "Stationarity is dead: Whither water management?" Water Resources Research 51, 7785–7789.
Mimikou, M., Koutsoyiannis, D., 1995. Extreme floods in Greece: The case of 1994, in: US-ITALY Research Workshop on the Hydrometeorology, Impacts, and Management of Extreme Floods, Perugia, Italy.
Mitchell, W.C., 1930. Business cycles: the problems and its setting Business cycles: The problem and its setting. National Bureau of Economic Research, New York.
Mittal, Y., Ylvisaker, D., 1975. Limit distributions for the maxima of stationary Gaussian processes. Stochastic Processes and their Applications 3, 1–18.
von Mises, R., 1936. La distribution de la plus grande de n valuers. Rev. math. Union interbalcanique 1, 141–160.
Modarres, R., da Silva, V. de P.R., 2007. Rainfall trends in arid and semi-arid regions of Iran. Journal of arid environments 70, 344–355.
Molini, A., Katul, G.G., Porporato, A., 2009. Revisiting rainfall clustering and intermittency across different climatic regimes. Water resources research 45.
Montanari, A., 2003. Long-range dependence in hydrology. Theory and applications of long-range dependence 461–472.
Montanari, A., 2012. Hydrology of the Po River: looking for changing patterns in river discharge. Hydrology and Earth System Sciences 16, 3739–3747.
Montanari, A., Brath, A., 2004. A stochastic approach for assessing the uncertainty of rainfall-runoff simulations. Water Resources Research 40.
Montanari, A., Koutsoyiannis, D., 2014. Modelling and mitigating natural hazards: Stationarity is immortal! Water Resources Research 50, 9748–9756.
Moss, R.H., Edmonds, J.A., Hibbard, K.A., Manning, M.R., Rose, S.K., Van Vuuren, D.P., Carter, T.R., Emori, S., Kainuma, M., Kram, T., 2010. The next generation of scenarios for climate change research and assessment. Nature 463, 747.
Mudelsee, M., 2007. Long memory of rivers from spatial aggregation. Water Resources Research 43.
Nalley, D., Adamowski, J., Biswas, A., Gharabaghi, B., Hu, W., 2019. A multiscale and multivariate analysis of precipitation and streamflow variability in relation to ENSO, NAO and PDO. Journal of Hydrology 574, 288–307.
Newell, G.F., 1964. Asymptotic extremes for m-dependent random variables. The Annals of Mathematical Statistics 35, 1322–1325.
Newman, M.E., 2005. Power laws, Pareto distributions and Zipf's law. Contemporary physics 46, 323–351.
Ntegeka, V., Willems, P., 2008. Trends and multidecadal oscillations in rainfall extremes, based on a more than 100-year time series of 10 min rainfall intensities at Uccle, Belgium. Water Resources Research 44.
Ntigkakis, C., Markopoulos-Sarikas, G., Dimitriadis, P., Iliopoulou, T., Efstratiadis, A., Koukouvinos, A., Koussis, A.D., Mazi, K., Katsanos, D., Koutsoyiannis, D., 2018. Hydrological investigation of the catastrophic flood event in Mandra, Western Attica, in: EGU General Assembly Conference Abstracts. p. 17591.
Nuzzo, R., 2014. Scientific method: statistical errors. Nature News 506, 150.
O'Brien, G.L., 1974. The maximum term of uniformly mixing stationary processes. Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete 30, 57–63.
O'Brien, G.L., Torfs, P.J., Vervaat, W., 1990. Stationary self-similar extremal processes. Probability theory and related fields 87, 97–119.

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). Hydrological Sciences Journal 61, 1571–1590.
Onof, C., Chandler, R.E., Kakou, A., Northrop, P., Wheater, H.S., Isham, V., 2000. Rainfall modelling using Poisson-cluster processes: a review of developments. Stochastic Environmental Research and Risk Assessment 14, 384–411.
Onof, C., Wang, L.-P., 2019. Modelling rainfall with a Bartlett-Lewis process: New developments. Hydrology and Earth System Sciences Discussions 1–40. https://doi.org/10.5194/hess-2019-406
Oreskes, N., 2004. The scientific consensus on climate change. Science 306, 1686–1686.
Over, T.M., Gupta, V.K., 1994. Statistical analysis of mesoscale rainfall: Dependence of a random cascade generator on large-scale forcing. Journal of Applied Meteorology 33, 1526–1542.
Pachauri, R.K., Allen, M.R., Barros, V.R., Broome, J., Cramer, W., Christ, R., Church, J.A., Clarke, L., Dahe, Q., Dasgupta, P., 2014. Climate change 2014: synthesis report. Contribution of Working Groups I. II and III to the fifth assessment report of the Intergovernmental Panel on Climate Change 151.
Papalexiou, S. M., D. Koutsoyiannis, and A. Montanari, 2011. Can a simple stochastic model generate rich patterns of rainfall events?, Journal of hydrology, 411(3), 279-289.
Papalexiou, S.M., 2018. Unified theory for stochastic modelling of hydroclimatic processes: Preserving marginal distributions, correlation structures, and intermittency. Advances in water resources 115, 234–252.
Papalexiou, S. M., and D. Koutsoyiannis, 2012. Entropy based derivation of probability distributions: A case study to daily rainfall, Advances in Water Resources, 45, 51-57.
Papalexiou, S.M., Koutsoyiannis, D., 2013. Battle of extreme value distributions: A global survey on extreme daily rainfall. Water Resources Research 49, 187–201.
Papalexiou, S.M., Koutsoyiannis, D., 2016. A global survey on the seasonal variation of the marginal distribution of daily precipitation. Advances in Water Resources 94, 131–145.
Papalexiou, S.M., Montanari, A., 2019. Global and Regional Increase of Precipitation Extremes under Global Warming. Water Resources Research.
Papoulakos, K., Iliopoulou, T., Dimitriadis, P., Tsaknias, D., Koutsoyiannis, D., 2020. Investigating the impacts of clustering of floods on insurance practices; a spatiotemporal analysis in the USA. https://doi.org/10.5194/egusphere-egu2020-8667
Papoulis, A., 1991. Probability, Random Variables, and Stochastic Processes, 3rd ed. McGraw-Hill, New York.
Parajka, J., Blaschke, A.P., Blöschl, G., Haslinger, K., Hepp, G., Laaha, G., Schöner, W., Trautvetter, H., Viglione, A., Zessner, M., 2016. Uncertainty contributions to low-flow projections in Austria. Hydrol. Earth Syst. Sci. 20, 2085–2101. https://doi.org/10.5194/hess-20-2085-2016
Parajka, J., Kohnová, S., Bálint, G., Barbuc, M., Borga, M., Claps, P., Cheval, S., Dumitrescu, A., Gaume, E., Hlavčová, K., others, 2010a. Seasonal characteristics of flood regimes across the Alpine–Carpathian range. Journal of hydrology 394, 78–89.
Parajka, J., Kohnová, S., Bálint, G., Barbuc, M., Borga, M., Claps, P., Cheval, S., Dumitrescu, A., Gaume, E., Hlavčová, K., others, 2010b. Seasonal characteristics of flood regimes across the Alpine–Carpathian range. Journal of hydrology 394, 78–89.
Parajka, J., Kohnová, S., Merz, R., Szolgay, J., Hlavčová, K., Blöschl, G., 2009. Comparative analysis of the seasonality of hydrological characteristics in Slovakia and Austria/Analyse comparative de la saisonnalité de

caractéristiques hydrologiques en Slovaquie et en Autriche. Hydrological Sciences Journal 54, 456–473.
Park, J., Onof, C., Kim, D., 2019. A hybrid stochastic rainfall model that reproduces some important rainfall characteristics at hourly to yearly timescales. Hydrology and Earth System Sciences 23, 989–1014.
Parmesan, C., Yohe, G., 2003. A globally coherent fingerprint of climate change impacts across natural systems. Nature 421, 37.
Paschalis, A., Molnar, P., Fatichi, S., Burlando, P., 2014. On temporal stochastic modelling of precipitation, nesting models across scales. Advances in water resources 63, 152–166.
Pelletier, J. D., 1998. The power spectral density of atmospheric temperature from time scales of 10− 2 to 10 6 yr, Earth and planetary science letters, 158(3), 157-164.
Pelletier, J. D., and D. L. Turcotte, 1997. Long-range persistence in climatological and hydrological time series: analysis, modelling and application to drought hazard assessment, Journal of Hydrology, 203(1), 198-208.
Persons, W.M., 1922. Measuring and Forecasting General Business Conditions. American institute of finance.
Pickands III, J., 1975. Statistical inference using extreme order statistics. the Annals of Statistics 3, 119–131.
Pizarro, A., Manfreda, S., Tubaldi, E., 2020. The science behind scour at bridge foundations: A review. Water 12, 374.
Potter, K. W., 1979. Annual precipitation in the northeast United States: Long memory, short memory, or no memory?, Water Resources Research, 15(2), 340-346.
Poveda, G. , 2011. Mixed memory,(non) Hurst effect, and maximum entropy of rainfall in the tropical Andes, Advances in Water Resources, 34(2), 243-256.
Pryor, S.C., Schoof, J.T., 2008. Changes in the seasonality of precipitation over the contiguous USA. J. Geophys. Res. 113, D21108. https://doi.org/10.1029/2008JD010251
Quadros, L.E. de, Mello, E.L. de, Gomes, B.M., Araujo, F.C., 2019. Rainfall trends for the State of Paraná: present and future climate. Revista Ambiente & Água 14.
Rahimi, M., Fatemi, S.S., n.d. Mean versus Extreme Precipitation Trends in Iran over the Period 1960–2017. Pure and Applied Geophysics 1–19.
Ravbar, N., 2013. VARIABILITY OF GROUNDWATER FLOW AND TRANSPORT PROCESSES IN KARST UNDER DIFFERENT HYDROLOGIC CONDITIONS/SPREMENLJIVOST PRETAKANJA VODA IN PRENOSA SNOVI V KRASU OB RAZLICNIH HIDROLOSKIH POGOJIH. Acta Carsologica 42, 327.
Resnick, S.I., 2007. Heavy-tail phenomena: probabilistic and statistical modelling. Springer Science & Business Media.
Ribatet, M., Ouarda, T.B., Sauquet, E., Gresillon, J.-M., 2009. Modelling all exceedances above a threshold using an extremal dependence structure: Inferences on several flood characteristics. Water Resources Research 45.
Rodriguez-Iturbe, I., De Power, B.F., Valdes, J.B., 1987a. Rectangular pulses point process models for rainfall: analysis of empirical data. Journal of Geophysical Research: Atmospheres 92, 9645–9656.
Rodriguez-Iturbe, Ignacio, Cox, D.R., Isham, V., 1987b. Some models for rainfall based on stochastic point processes. Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences 410, 269–288.
Rootzen, H., 1986. Extreme value theory for moving average processes. The Annals of Probability 14, 612–652.
Rootzén, H., 1988. Maxima and exceedances of stationary Markov chains. Advances in applied probability 20, 371–390.

Rosenblatt, M., 1956. A central limit theorem and a strong mixing condition. Proceedings of the National Academy of Sciences of the United States of America 42, 43.
Rotstayn, L.D., Lohmann, U., 2002. Tropical rainfall trends and the indirect aerosol effect. Journal of Climate 15, 2103–2116.
Roy, T., Dimitriadis, P. Iliopoulou, T., Koutsoyiannis, D., Wood. E. F. "Effects of Hurst-Kolmogorov Dynamics in Intensity-Duration-Frequency Curves." In AGU Fall Meeting Abstracts. 2018
Rust, H.W., Maraun, D., Osborn, T.J., 2009. Modelling seasonality in extreme precipitation. The European Physical Journal-Special Topics 174, 99–111.
Salvadori, G., De Michele, C., 2006. Statistical characterization of temporal structure of storms. Advances in Water Resources 29, 827–842.
Salvadori, G., De Michele, C., 2010. Multivariate multiparameter extreme value models and return periods: A copula approach. Water resources research 46.
Sansom, J., 1998. A hidden Markov model for rainfall using breakpoint data. Journal of Climate 11, 42–53.
Santer, B.D., Wigley, T.M.L., Boyle, J.S., Gaffen, D.J., Hnilo, J.J., Nychka, D., Parker, D.E., Taylor, K.E., 2000. Statistical significance of trends and trend differences in layer-average atmospheric temperature time series. Journal of Geophysical Research: Atmospheres 105, 7337–7356.
Sargentis, G.-F, Iliopoulou, T., Sigourou, S., Dimitriadis, P., Koutsoyiannis, D., 2020. Evolution of clustering quantified by a stochastic method-cases studies on natural and human social structures. Sustainability, in review.
Scaife, A.A., Arribas, A., Blockley, E., Brookshaw, A., Clark, R.T., Dunstone, N., Eade, R., Fereday, D., Folland, C.K., Gordon, M., 2014. Skillful long-range prediction of European and North American winters. Geophysical Research Letters 41, 2514–2519.
Scaife, A.A., Folland, C.K., Alexander, L.V., Moberg, A., Knight, J.R., 2008. European climate extremes and the North Atlantic Oscillation. Journal of Climate 21, 72–83.
Schertzer, D., Lovejoy, S., 1987. Physical modelling and analysis of rain and clouds by anisotropic scaling multiplicative processes. Journal of Geophysical Research: Atmospheres 92, 9693–9714.
Sen, P.K., 1968. Estimates of the regression coefficient based on Kendall's tau. Journal of the American statistical association 63, 1379–1389.
Seneviratne, S.I., Koster, R.D., Guo, Z., Dirmeyer, P.A., Kowalczyk, E., Lawrence, D., Liu, P., Mocko, D., Lu, C.-H., Oleson, K.W., others, 2006. Soil moisture memory in AGCM simulations: analysis of global land–atmosphere coupling experiment (GLACE) data. Journal of Hydrometeorology 7, 1090–1112.
Serinaldi, F., 2009a. Copula-based mixed models for bivariate rainfall data: an empirical study in regression perspective. Stochastic environmental research and risk assessment 23, 677–693.
Serinaldi, F., 2009b. A multisite daily rainfall generator driven by bivariate copula-based mixed distributions. Journal of Geophysical Research: Atmospheres 114.
Serinaldi, F., 2010. Multifractality, imperfect scaling and hydrological properties of rainfall time series simulated by continuous universal multifractal and discrete random cascade models. Nonlinear Processes in Geophysics 17, 697–714.
Serinaldi, F., 2013. On the relationship between the index of dispersion and Allan factor and their power for testing the Poisson assumption. Stochastic environmental research and risk assessment 27, 1773–1782.
Serinaldi, F., 2015. Dismissing return periods!. Stochastic environmental research and risk assessment, 29(4), pp.1179-1189.
Serinaldi, F., Chebana, F., Kilsby, C.G., 2020. Dissecting innovative trend analysis. Stochastic Environmental Research and Risk Assessment 1–22.

Serinaldi, F., Grimaldi, S., 2007. Fully nested 3-copula: procedure and application on hydrological data. Journal of Hydrologic Engineering 12, 420–430.
Serinaldi, F., Kilsby, C.G., 2013. On the sampling distribution of Allan factor estimator for a homogeneous Poisson process and its use to test inhomogeneities at multiple scales. Physica A: Statistical Mechanics and its Applications 392, 1080–1089.
Serinaldi, F., Kilsby, C.G., 2014. Rainfall extremes: Toward reconciliation after the battle of distributions. Water resources research 50, 336–352.
Serinaldi, F., Kilsby, C.G., 2016a. Irreversibility and complex network behaviour of stream flow fluctuations. Physica A: Statistical Mechanics and its Applications 450, 585–600. https://doi.org/10.1016/j.physa.2016.01.043
Serinaldi, F., Kilsby, C.G., 2016b. Understanding persistence to avoid underestimation of collective flood risk. Water 8, 152.
Serinaldi, F., Kilsby, C.G., 2018a. Unsurprising Surprises: The Frequency of Record-breaking and Overthreshold Hydrological Extremes Under Spatial and Temporal Dependence. Water Resources Research 54, 6460–6487.
Serinaldi, F., Kilsby, C.G., Lombardo, F., 2018b. Untenable nonstationarity: An assessment of the fitness for purpose of trend tests in hydrology. Advances in Water Resources 111, 132–155.
Serinaldi, F., Lombardo, F., 2020. Probability distribution of waiting time of the *k*th extreme event under serial dependence. Journal of Hydrologic Engineering.
Serra-Llobet, A., Tàbara, J.D., Sauri, D., 2013. The Tous dam disaster of 1982 and the origins of integrated flood risk management in Spain. Natural hazards 65, 1981–1998.
Sharma, P.N., Shmueli, G., Sarstedt, M., Danks, N., Ray, S., 2019. Prediction-oriented model selection in partial least squares path modelling. Decision Sciences.
Shepard, N., 1996. Statistical Aspects of ARCH Models and Stochastic Volatility. Obozrenie Prikl. Prom. Mat 3, 764–826.
Shibata, R., 1980. Asymptotically efficient selection of the order of the model for estimating parameters of a linear process. The annals of statistics 147–164.
Shmueli, G., 2010. To explain or to predict? Statistical science 25, 289–310.
Sibuya, M., 1960. Bivariate extreme statistics. Annals of the Institute of Statistical Mathematics 11, 195–210.
Simonoff, J.S., 2012. Smoothing methods in statistics. Springer Science & Business Media.
Simpson, D., Rue, H., Riebler, A., Martins, T.G., Sørbye, S.H., 2017. Penalising model component complexity: A principled, practical approach to constructing priors. Statistical science 32, 1–28.
Sivapalan, M., Blöschl, G., Merz, R., Gutknecht, D., 2005. Linking flood frequency to long-term water balance: Incorporating effects of seasonality. Water Resources Research 41.
Sklar, M., 1959. Fonctions de repartition an dimensions et leurs marges. Publ. inst. statist. univ. Paris 8, 229–231.
Slutsky, E.E., 1927. Slozhenie sluchainykh prichin, kak istochnik tsiklicheskikh protsessov. Voprosy kon''yunktury 3, 34–64.
Smith, D.M., Scaife, A.A., Eade, R., Knight, J.R., 2016. Seasonal to decadal prediction of the winter North Atlantic Oscillation: emerging capability and future prospects. Quarterly Journal of the Royal Meteorological Society 142, 611–617.
Smith, R.L., Tawn, J.A., Coles, S.G., 1997. Markov chain models for threshold exceedances. Biometrika 84, 249–268.
Solomon, S., Qin, D., Manning, M., Averyt, K., Marquis, M., 2007. Climate change 2007-the physical science basis: Working group I contribution to the fourth assessment report of the IPCC. Cambridge university press.
Spearman, C., 1904. The Proof and Measurement of Association between Two Things. The American Journal of Psychology 15, 72–101. https://doi.org/10.2307/1412159

Srikanthan, R., McMahon, T.A., 2001. Stochastic generation of annual, monthly and daily climate data: A review. Hydrology and Earth System Sciences Discussions 5, 653–670.
Stedinger, J.R., Griffis, V.W., 2008. Flood frequency analysis in the United States: Time to update. American Society of Civil Engineers.
Stein, R.M., 2002. Benchmarking default prediction models: Pitfalls and remedies in model validation. Moody's KMV, New York 20305.
Steirou, E., 2011. Investigation of methods for hydroclimatic data homogenization, National Technical University of Athens, Athens.
Stephenson, D. B., V. Pavan, and R. Bojariu, 2000. Is the North Atlantic Oscillation a random walk?, International Journal of Climatology, 20(1), 1-18.
Stone, M., 1974. Cross-validatory choice and assessment of statistical predictions. Journal of the Royal Statistical Society: Series B (Methodological) 36, 111–133.
Stone, M., 1977. An asymptotic equivalence of choice of model by cross-validation and Akaike's criterion. Journal of the Royal Statistical Society: Series B (Methodological) 39, 44–47.
Sun, C., Li, J., Feng, J., Xie, F., 2015. A decadal-scale teleconnection between the North Atlantic Oscillation and subtropical eastern Australian rainfall. Journal of Climate 28, 1074–1092.
Svensson, C., 2016. Seasonal river flow forecasts for the United Kingdom using persistence and historical analogues. Hydrological Sciences Journal 61, 19–35. https://doi.org/10.1080/02626667.2014.992788
Szolgayova, E., Laaha, G., Blöschl, G., Bucher, C., 2014. Factors influencing long range dependence in streamflow of European rivers. Hydrological Processes 28, 1573–1586.
Tegos, A., Tyralis, H., Koutsoyiannis, D., Hamed, K., 2017. An R function for the estimation of trend significance under the scaling hypothesis-application in PET parametric annual time series. Open Water Journal 4, 6.
Telesca, L., Cuomo, V., Lapenna, V., Macchiato, M., 2002. On the methods to identify clustering properties in sequences of seismic time-occurrences. Journal of seismology 6, 125–134.
Theil, H., 1992. A rank-invariant method of linear and polynomial regression analysis, in: Henri Theil's Contributions to Economics and Econometrics. Springer, pp. 345–381.
Thurner, S., Lowen, S.B., Feurstein, M.C., Heneghan, C., Feichtinger, H.G., Teich, M.C., 1997. Analysis, synthesis, and estimation of fractal-rate stochastic point processes. Fractals 5, 565–595.
Thyer, M., Kuczera, G., 2000. Modelling long-term persistence in hydroclimatic time series using a hidden state Markov Model. Water resources research 36, 3301–3310.
Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society: Series B (Methodological) 58, 267–288.
Todorovic, P., Woolhiser, D.A., 1975. A stochastic model of n-day precipitation. Journal of Applied Meteorology 14, 17–24.
Trafimow, D., Amrhein, V., Areshenkoff, C.N., Barrera-Causil, C.J., Beh, E.J., Bilgiç, Y.K., Bono, R., Bradley, M.T., Briggs, W.M., Cepeda-Freyre, H.A., 2018. Manipulating the alpha level cannot cure significance testing. Frontiers in Psychology 9.
Tsoukalas, I., 2018. Modelling and simulation of non-Gaussian stochastic processes for optimization of water-systems under uncertainty, PhD thesis, Department of Water Resources and Environmental Engineering – National Technical University of Athens.
Tsoukalas, I., Makropoulos, C., Koutsoyiannis, D., 2018. Simulation of Stochastic Processes Exhibiting Any-Range Dependence and Arbitrary Marginal Distributions. Water Resources Research 54, 9484–9513.

Tye, M.R., Blenkinsop, S., Fowler, H.J., Stephenson, D.B., Kilsby, C.G., 2016. Simulating multimodal seasonality in extreme daily precipitation occurrence. Journal of Hydrology 537, 117–129.
Tye, M.R., Katz, R.W., Rajagopalan, B., 2018. Climate change or climate regimes? Examining multi-annual variations in the frequency of precipitation extremes over the Argentine Pampas. Climate Dynamics 1–16.
Tyralis, H., and D. Koutsoyiannis, 2011. Simultaneous estimation of the parameters of the Hurst–Kolmogorov stochastic process, Stochastic Environmental Research and Risk Assessment, 25(1), 21-33.
Tyralis, H., Dimitriadis, P., Koutsoyiannis, D., O'Connell, P.E., Tzouka, K., Iliopoulou, T., 2018. On the long-range dependence properties of annual precipitation using a global network of instrumental measurements. Advances in Water Resources 111, 301–318.
Tyralis, H., Koutsoyiannis, D., 2011. Simultaneous estimation of the parameters of the Hurst–Kolmogorov stochastic process. Stochastic Environmental Research and Risk Assessment 25, 21–33.
Tyralis, H., Langousis, A., 2019. Estimation of intensity–duration–frequency curves using max-stable processes. Stochastic Environmental Research and Risk Assessment 33, 239–252. https://doi.org/10.1007/s00477-018-1577-2
Vandenberghe, S., Verhoest, N.E.C., De Baets, B., 2010. Fitting bivariate copulas to the dependence structure between storm characteristics: A detailed analysis based on 105 year 10 min rainfall. Water resources research 46.
Veneziano, D., Furcolo, P., Iacobellis, V., 2006a. Imperfect scaling of time and space–time rainfall. Journal of Hydrology 322, 105–119.
Veneziano, D., Langousis, A., 2010. Scaling and fractals in hydrology, in: Advances in Data-Based Approaches for Hydrologic Modelling and Forecasting. World Scientific, pp. 107–243.
Veneziano, D., Langousis, A., Furcolo, P., 2006b. Multifractality and rainfall extremes: A review. Water resources research 42.
Verhoest, N.E., Vandenberghe, S., Cabus, P., Onof, C., Meca-Figueras, T., Jameleddine, S., 2010. Are stochastic point rainfall models able to preserve extreme flood statistics? Hydrological processes 24, 3439–3445.
Villarini, G., 2012. Analyses of annual and seasonal maximum daily rainfall accumulations for Ukraine, Moldova, and Romania. Int. J. Climatol. 32, 2213–2226. https://doi.org/10.1002/joc.3394
Villarini, G., Serinaldi, F., Krajewski, W.F., 2008. Modelling radar-rainfall estimation uncertainties using parametric and non-parametric approaches. Advances in Water Resources 31, 1674–1686.
Vitolo, R., Stephenson, D.B., Cook, I.M., Mitchell-Wallace, K., 2009. Serial clustering of intense European storms. Meteorologische Zeitschrift 18, 411–424.
Volpi, E., Fiori, A., Grimaldi, S., Lombardo, F., Koutsoyiannis, D., 2019. Save hydrological observations! Return period estimation without data decimation. Journal of Hydrology 571, 782–792.
Volpi, E., Fiori, A., Grimaldi, S., Lombardo, F., Koutsoyiannis, D., 2015. One hundred years of return period: Strengths and limitations. Water Resources Research 51, 8570–8585.
Walker, G.T., 1931. On periodicity in series of related terms. Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character 131, 518–532.
Wang, Q.J., Robertson, D.E., Chiew, F.H.S., 2009. A Bayesian joint probability modelling approach for seasonal forecasting of streamflows at multiple sites. Water Resources Research 45.
Wardhana, K., Hadipriono, F.C., 2003. Analysis of recent bridge failures in the United States. Journal of performance of constructed facilities 17, 144–150.

Wasserstein, R.L., Lazar, N.A., 2016. The ASA Statement on p-Values: Context, Process, and Purpose. The American Statistician 70, 129–133. https://doi.org/10.1080/00031305.2016.1154108
Wasserstein, R.L., Schirm, A.L., Lazar, N.A., 2019. Moving to a world beyond "$p<$ 0.05." Taylor & Francis.
Watson, G.S., 1954. Extreme values in samples from $ m $-dependent stationary stochastic processes. The Annals of Mathematical Statistics 25, 798–800.
Waymire, E.D., Gupta, V.K., 1981. The mathematical structure of rainfall representations: 1. A review of the stochastic rainfall models. Water resources research 17, 1261–1272.
Wei, C.-Z., 1992. On predictive least squares principles. The Annals of Statistics 20, 1–42.
Whittle, P., 1952. Tests of fit in time series. Biometrika 39, 309–318.
Whittle, P., 1953. The analysis of multiple stationary time series. Journal of the Royal Statistical Society: Series B (Methodological) 15, 125–139.
Wilks, D.S., 1999. Interannual variability and extreme-value characteristics of several stochastic daily precipitation models. Agricultural and Forest Meteorology 93, 153–169.
Willems, P., 2013. Adjustment of extreme rainfall statistics accounting for multidecadal climate oscillations. Journal of hydrology 490, 126–133.
Wold, H., 1938. A study in the analysis of stationary time series (PhD Thesis). Almqvist & Wiksell.
Wold, H.O., 1948. On prediction in stationary time series. The Annals of Mathematical Statistics 19, 558–567.
Wu, H., Qian, H., 2017. Innovative trend analysis of annual and seasonal rainfall and extreme values in Shaanxi, China, since the 1950s. Int. J. Climatol. 37, 2582–2592. https://doi.org/10.1002/joc.4866
Wu, S., Harris, T.J., McAuley, K.B., 2007. The use of simplified or misspecified models: Linear case. The Canadian Journal of Chemical Engineering 85, 386–398.
Yarkoni, T., Westfall, J., 2017. Choosing prediction over explanation in psychology: Lessons from machine learning. Perspectives on Psychological Science 12, 1100–1122.
Ye, M., Meyer, P.D., Neuman, S.P., 2008. On model selection criteria in multimodel analysis. Water Resources Research 44. https://doi.org/10.1029/2008WR006803
Yeager, S.G., Robson, J.I., 2017. Recent progress in understanding and predicting Atlantic decadal climate variability. Current Climate Change Reports 3, 112–127.
Yossef, N.C., Winsemius, H., Weerts, A., Beek, R., Bierkens, M.F., 2013. Skill of a global seasonal streamflow forecasting system, relative roles of initial conditions and meteorological forcing. Water Resources Research 49, 4687–4699.
Yule, G.U., 1927. VII. On a method of investigating periodicities disturbed series, with special reference to Wolfer's sunspot numbers. Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character 226, 267–298.
Zhai, Y., Y. Guo, J. Zhou, N. Guo, J. Wang, and Y. Teng (2014), The spatio-temporal variability of annual precipitation and its local impact factors during 1724–2010 in Beijing, China, Hydrological Processes, 28(4), 2192-2201.
Zhang, L., Singh, V.P., 2007. Bivariate rainfall frequency distributions using Archimedean copulas. Journal of Hydrology 332, 93–109.
Zhang, Q., Zhou, Y., Singh, V.P., Chen, X., 2012. The influence of dam and lakes on the Yangtze River streamflow: long-range correlation and complexity analyses. Hydrological Processes 26, 436–444.
Zscheischler, J., Westra, S., Van Den Hurk, B.J., Seneviratne, S.I., Ward, P.J., Pitman, A., AghaKouchak, A., Bresch, D.N., Leonard, M., Wahl, T. and Zhang, X., 2018.

Future climate risk from compound events. Nature Climate Change, 8(6), pp.469-477.
Zoukos, A., Iliopoulou, T. Dimitriadis, P., Koutsoyiannis, D. Global investigation of the multi-scale probabilistic behaviour of dry spells from rainfall records, European Geosciences Union General Assembly 2018, Geophysical Research Abstracts, Vol. 20, Vienna, EGU2018-17966-1, doi:10.13140/RG.2.2.13555.78886, European Geosciences Union, 2018.

Appendix A

THE LONG-TERM RAINFALL RECORDS

A.1 History and acknowledgments for dataset compilation

The long-term rainfall dataset was manually compiled with the aim to assemble the longest-term rainfall records available at the time. The first version was compiled in 2015 and included 27 rainfall records of over 150 years of daily data, which were explored in the analysis of Chapter 4. The dataset was updated in 2018 to its latest version including 60 rainfall records, which were used in the analyses of Chapters 6-8.

We greatly thank the Radcliffe Meteorological Station, the Icelandic Meteorological Office (Trausti Jónsson), the Czech Hydrometeorological Institute, the Finnish Meteorological Institute, the National Observatory of Athens, the Department of Earth Sciences of the Uppsala University and the Regional Hydrologic Service of the Tuscany Region (servizio.idrologico@regione.toscana.it) for providing the required data for each region respectively. We are also grateful to Professor Ricardo Machado Trigo (University of Lisbon) for providing the Lisbon timeseries, to Professor Marco Marani (University of Padua) for providing the Padua timeseries and to Professor Joo-Heon Lee (Joongbu University) for providing the Seoul timeseries. All the above data were freely provided after contacting the acknowledged sources. The remaining timeseries are publicly available by the data providers in the ECA&D project (http://www.ecad.eu), and in the GHCN-Daily database (https://data.noaa.gov/dataset/global-historical-climatology-network-daily-ghcn-daily-version-3).

Table A1 contains the essential information on the long-term rainfall stations, including name, geographic coordinates, record length and the respective data sources. A map depicting the location of the gauges is included in Chapter 6 (Fig. 6.5), while summary statistics for the rainfall records are provided in Table 6.2. The longest record in the dataset is the Padova rainfall data (289 years) followed by the Chuk-woo-kee rainfall data from Seoul (241) years.

Table A.1 Properties (name, source, latitude, longitude, start year, end year, record length and missing values percentage) of the 60 longest stations used in the analysis sorted by decreasing length. For the global datasets, the European Climate Assessment dataset (ECA; http://www.ecad.eu) and the Global Historical Climatology Network Daily database (GHCND; https://data.noaa.gov/dataset/global-historical-climatology-network-daily-ghcn-daily-version-3), the station identifier is also reported. Asterisks (*) in the "end year" column denote data that have been continued from a second source. The country of each station is abbreviated in parentheses aside its name.

Name	Source	Lat	Lon	Start year	End year	Record length	Missing %
PADOVA (IT)	Marani and Zanetti (2015)	45.87	11.53	1725	2013	289	5.04
CHUK-WOO-KEE, SEOUL (KR)	Jhun and Moon (1997) and Korea Meteorological Agency	37.53	127.02	1777	2017*	241	0.00

Name	Source	Lat	Lon	Start year	End year	Record length	Missing %
HOHENPEISSENBERG (DE)	ECA: 48 HOHENPEISSENBERG DE	47.80	11.01	1781	2017	237	25.56
PALERMO (IT)	GHCND:ITE00105250	38.11	13.35	1797	2008	212	17.16
PRAGUE (CZ)	Czech Hydrometeorological Institute	50.05	14.25	1804	2014	211	0.20
BOLOGNA (IT)	GHCND:ITE00100550 and Dext3r of ARPA Emilia Romagna, Rete di monitoraggio RIRER (http://www.smr.arpa.emr.it/dext3r/)	44.50	11.35	1813	2018 *	206	0.00
JENA STERNWARTE GM (DE)	GHCND:GM000004204	50.93	11.58	1826	2015	190	5.47
RADCLIFFE (UK)	Radcliffe Meteorological Station (Burt and Howden, 2011)	51.76	-1.26	1827	2014	188	0.05
UPPSALA (SE)	Department of Earth Sciences of the Uppsala University	59.86	17.63	1836	2014	179	0.02
TORONTO (CA)	GHCND:CA006158350	43.67	-79.40	1840	2015	176	5.97
GENOA (IT)	GHCND:ITE00100552	44.41	8.93	1833	2008	176	0.00
ONNEN (NL)	ECA :2491 ONNEN NL	53.15	6.67	1846	2018	173	1.10
SAPPEMEER (NL)	ECA:2507 SAPPEMEER NL	53.17	6.73	1846	2018	173	1.10
WOLTERSUM (NL)	ECA:2553 WOLTERSUM NL	53.27	6.72	1846	2018	173	1.14
GRONINGEN (NL)	ECA:147 GRONINGEN NL	53.18	6.60	1846	2018	173	1.10
RODEN (NL)	ECA:516 RODEN NL	53.15	6.43	1846	2018	173	1.10
EELDE (NL)	ECA:164 EELDE NL	53.12	6.58	1846	2018	173	1.10
HELSINKI (FI)	Finnish Meteorological Institute	60.17	24.93	1845	2015	171	0.33
MANTOVA (IT)	GHCND:ITE00100553	45.16	10.80	1840	2008	169	5.75
DEN_HELDER (NL)	ECA:146 DEN_HELDER NL	52.93	4.75	1850	2018	169	1.13
DE_KOOY (NL)	ECA:145 DE_KOOY NL	52.92	4.78	1850	2018	169	1.13
ANNA_PAULOWNA (NL)	ECA:521 ANNA_PAULOWNA NL	52.87	4.83	1850	2018	169	1.13
CALLANTSOOG (NL)	ECA:2382 CALLANTSOOG NL	52.85	4.70	1850	2018	169	1.13
RITTHEM (NL)	ECA:2503 RITTHEM NL	51.47	3.62	1854	2018	165	1.16
VLISSINGEN (NL)	ECA:166 VLISSINGEN NL	51.44	3.60	1854	2018	165	1.16
SCHOONDIJKE (NL)	ECA:572 SCHOONDIJKE NL	51.35	3.55	1854	2018	165	1.16
'S_HEERENHOEK (NL)	ECA:2350 'S_HEERENHOEK NL	51.47	3.77	1854	2018	165	1.16
BRESKENS (NL)	ECA:2377 BRESKENS NL	51.40	3.55	1854	2018	165	1.16
MIDDELBURG (NL)	ECA:2474 MIDDELBURG NL	51.48	3.60	1854	2018	165	1.16
ARMAGH (UK)	GHCND:UK000047811	54.35	-6.65	1838	2001	164	0.26
OXFORD (UK)	GHCND:UK000056225	51.77	-1.27	1853	2015	163	0.42
HVAR (HR)	ECA:1686 HVAR HR	43.17	16.45	1857	2018	162	7.74
MELBOURNE REGIONAL OFFICE (AS)	GHCND:ASN00086071	-37.81	144.97	1855	2015	161	1.29
STYKKISHOLMUR (IS)	Icelandic Meteorological Office	65.08	-22.73	1856	2015	160	1.00

Name	Source	Lat	Lon	Start year	End year	Record length	Missing %
GRYCKSBO_D (SE)	ECA:6456 GRYCKSBO_D SE	60.69	15.49	1860	2018	159	0.62
FALUN (SE)	GHCND:SW000010537	60.62	15.62	1860	2018	159	0.89
VAEXJOE (SE)	GHCND:SWE00100003	56.87	14.80	1860	2018	159	4.13
FLORENCE (IT)	Regional Hydrologic Service of the Tuscany Region	43.80	11.20	1822	1979	158	2.00
SYDNEY OBSERVATORY HILL (AS)	GHCND:ASN00066062	-33.86	151.21	1858	2015	158	0.48
DENILIQUIN WILKINSON ST (AS)	GHCND:ASN00074128	-35.53	144.95	1858	2014	157	1.37
ZAGREB GRIC (HR)	GHCND:HR000142360	45.82	15.98	1860	2015	156	1.54
ROBE COMPARISON (AS)	GHCND:ASN00026026	-37.16	139.76	1860	2015	156	3.66
GABO ISLAND LIGHTHOUSE (AS)	GHCND:ASN00084016	-37.57	149.92	1864	2018	155	3.36
NEWCASTLE NOBBYS SIGNAL STATIO (AS)	GHCND:ASN00061055	-32.92	151.80	1862	2015	154	2.55
OVERVEEN (NL)	ECA:2497 OVERVEEN NL	52.40	4.60	1866	2018	153	1.25
HOOFDDORP (NL)	ECA:151 HOOFDDORP NL	52.32	4.70	1866	2018	153	1.25
ROELOFARENDSVEEN (NL)	ECA:540 ROELOFARENDSVEEN NL	52.22	4.62	1866	2018	153	1.29
SCHIPHOL (NL)	ECA:593 SCHIPHOL NL	52.32	4.79	1866	2018	153	1.25
AALSMEER (NL)	ECA:2351 AALSMEER NL	52.27	4.77	1866	2018	153	1.25
HEEMSTEDE (NL)	ECA:2430 HEEMSTEDE NL	52.35	4.63	1866	2018	153	1.25
LIJNDEN_(NH) (NL)	ECA:2466 LIJNDEN_(NH) NL	52.35	4.75	1866	2018	153	1.25
LISSE (NL)	ECA:2467 LISSE NL	52.27	4.55	1866	2018	153	1.29
NIJKERK (NL)	ECA:2484 NIJKERK NL	52.23	5.47	1867	2018	152	0.75
VOORTHUIZEN (NL)	ECA:2542 VOORTHUIZEN N	52.18	5.62	1867	2018	152	0.75
PUTTEN_(GLD) (NL)	ECA: 551 PUTTEN_(GLD) NL	5.62	14.00	1867	2018	152	0.75
ATHENS (GR)	National Observatory of Athens	37.97	23.72	1863	2014	152	0.66
ELSPEET (NL)	ECA:2404 ELSPEET NL	52.28	5.78	1867	2018	152	0.75
LISBON (PT)	Kutiel and Trigo (2014)	39.20	-9.25	1863	2013	151	1.06
MILAN (IT)	GHCND:ITE00100554	45.47	9.19	1858	2008	151	0.12
NEW_YORK_CNTRL_PK_TWR (US)	GHCND: USW00094728	40.78	-73.97	1869	2018	150	0.51

Appendix B

SUPPLEMENTARY MATERIAL TO CHAPTERS 5, 6 & 8

B.1 Supplement to Chapter 5

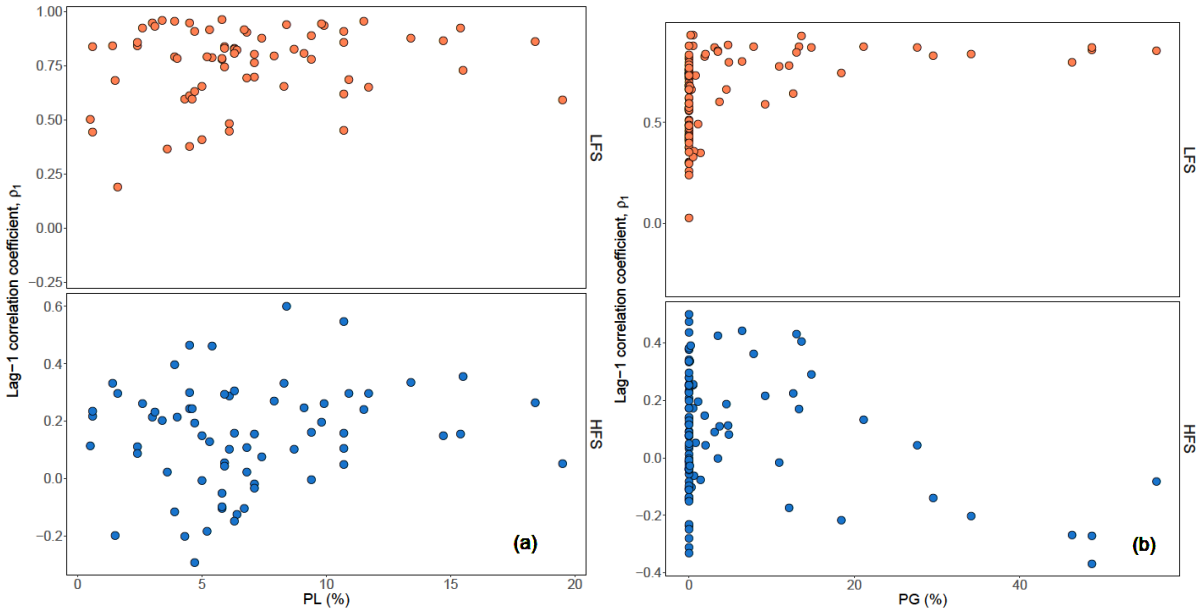


Figure B.1 Scatterplots of lag-1 HFS (bottom) and LFS (top) streamflow correlations versus percentage of lakes PL of the Swedish catchments (a) and percentage of glaciers PG of the Austrian catchments (b).

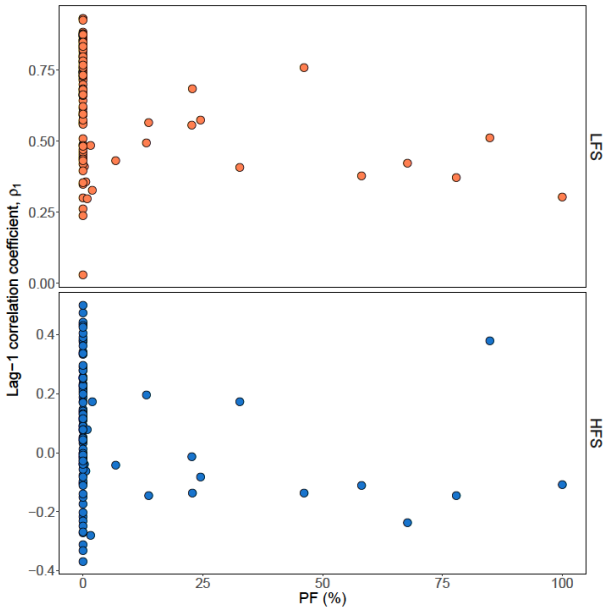


Figure B.2 Scatterplots of lag-1 correlation vs percentage of flysch area coverage PF for HFS (bottom) and LFS (top) analysis for the Austrian catchments.

B.2 Supplement to Chapter 6

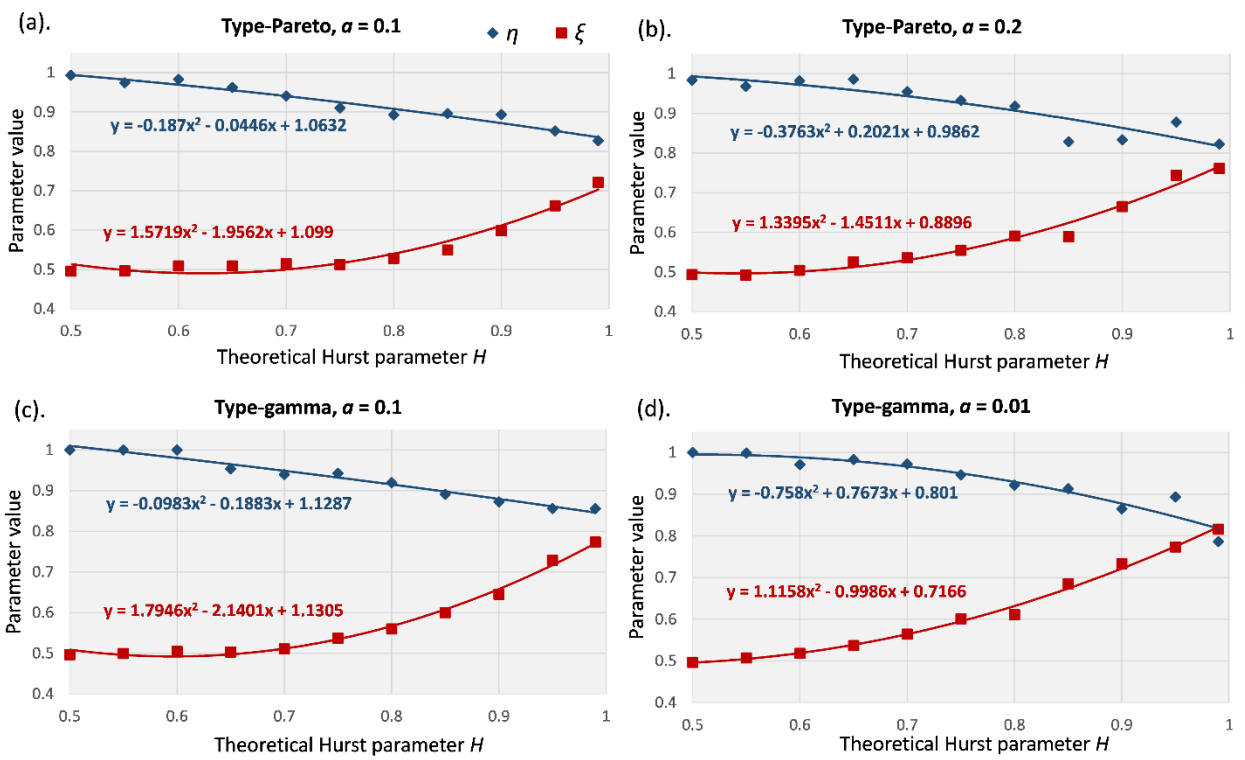


Figure B.3 Plots of η and ξ parameters versus the H parameter and polynomial fitting for the (a) type-Pareto with α=0.1, (b) type-Pareto with α=0.2, (c) type-gamma with α=0.1 and (d) type-gamma with α=0.01.

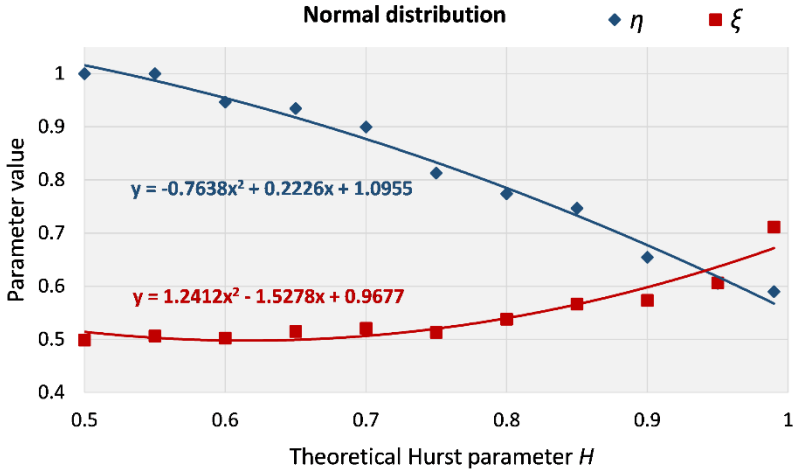


Figure B.4 Plots of η and ξ parameters versus the H parameter and polynomial fitting for the normal distribution.

B.3 Supplement to Chapter 8

RAINFALL TRENDS PERFORMANCE IN RECENT CLIMATIC PERIOD

In Fig B.5-B.8, we illustrate the static validation scheme showing results from the projections of the local trend and the local mean model for all rainfall indices.

Figure B.5 Local trend vs the local mean in projecting annual maxima for the 60 longest rainfall stations.

Figure B.6 Local trend vs the local mean in projecting annual totals for the 60 longest rainfall stations.

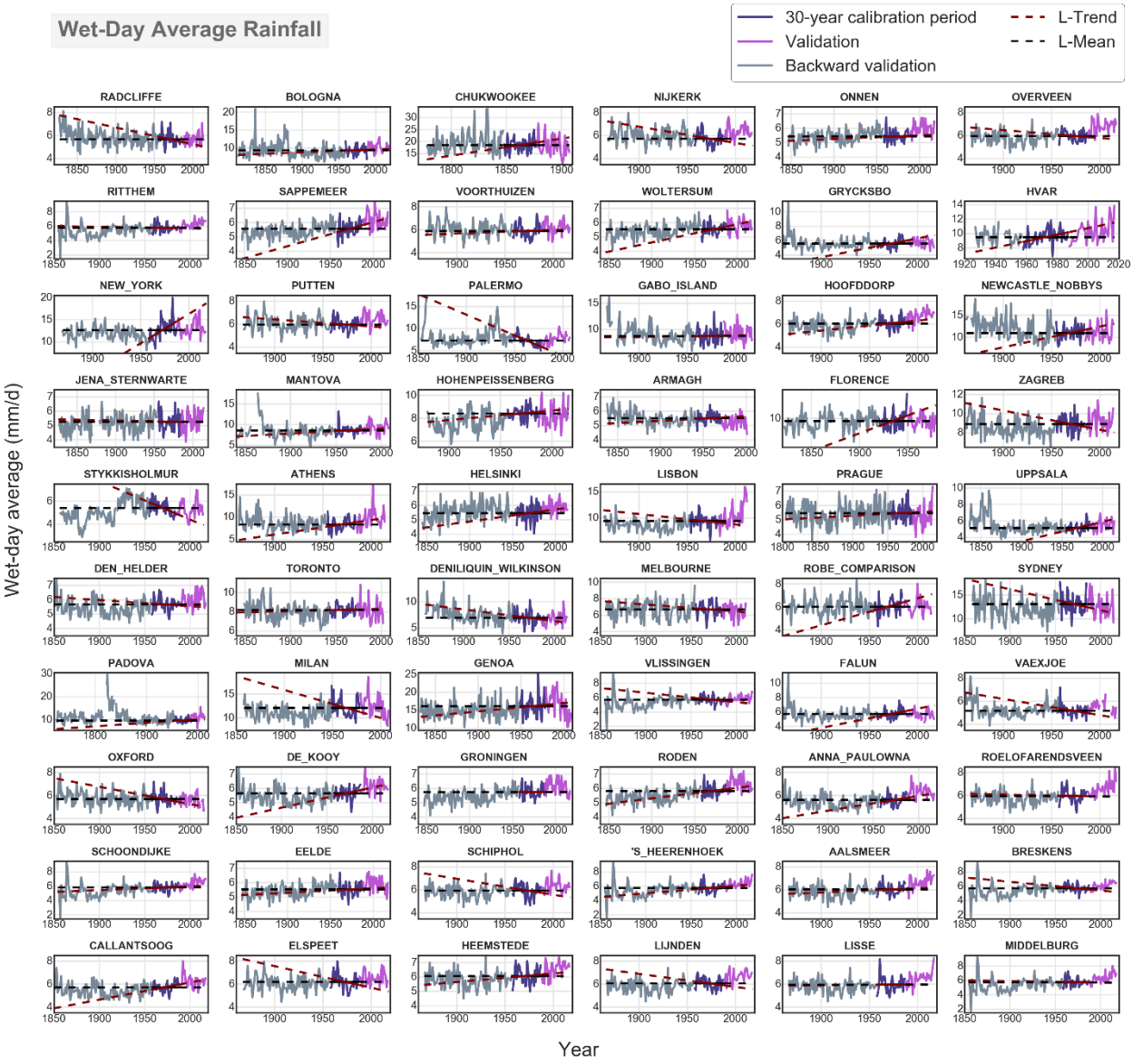


Figure B.7 Local trend vs the local mean in projecting wet-day average rainfall for the 60 longest rainfall stations.

Figure B.8 Local trend vs the local mean in projecting probability dry for the 60 longest rainfall stations.

FITTING ALGORITHMS: LEAST-SQUARES VS ROBUST REGRESSION

We explore the effect of the linear trend definition and fitting algorithm on the results of the local trends, as trends in small segments are expected to be more sensitive to the choice of the fitting algorithm (Santer et al., 2000). The first algorithm is the widely used ordinary least-square estimation (OLS), which fits equation (48) to the data, by minimizing the sum of the squares of the differences between the observed data and the predictions of the linear model. Secondly, two alternative trend calibration approaches are explored that place less weight on influential observations (“outliers”) and thus belong to the range of ‘robust regression’ techniques. The first is the least absolute deviations (LAD) method, which estimates the regression coefficients by minimising the sum of absolute deviations of the predicted from the observed values, and is a special case of quantile regression, fitting the trend line to the median of the

observations, rather than the mean (Chandler and Scott, 2011). The second is the non-parametric method of Theil-Sen slope estimation (Sen, 1968; Theil, 1992), which estimates the slope b of the linear model as the median of the pairwise slopes of all sample points. Among the different approaches that exist for the intercept coefficient, we follow Conover (1980) and estimated the intercept as $a = y_{0.5} - bx_{0.5}$, where $y_{0.5}$ and $x_{0.5}$ are the sample medians.

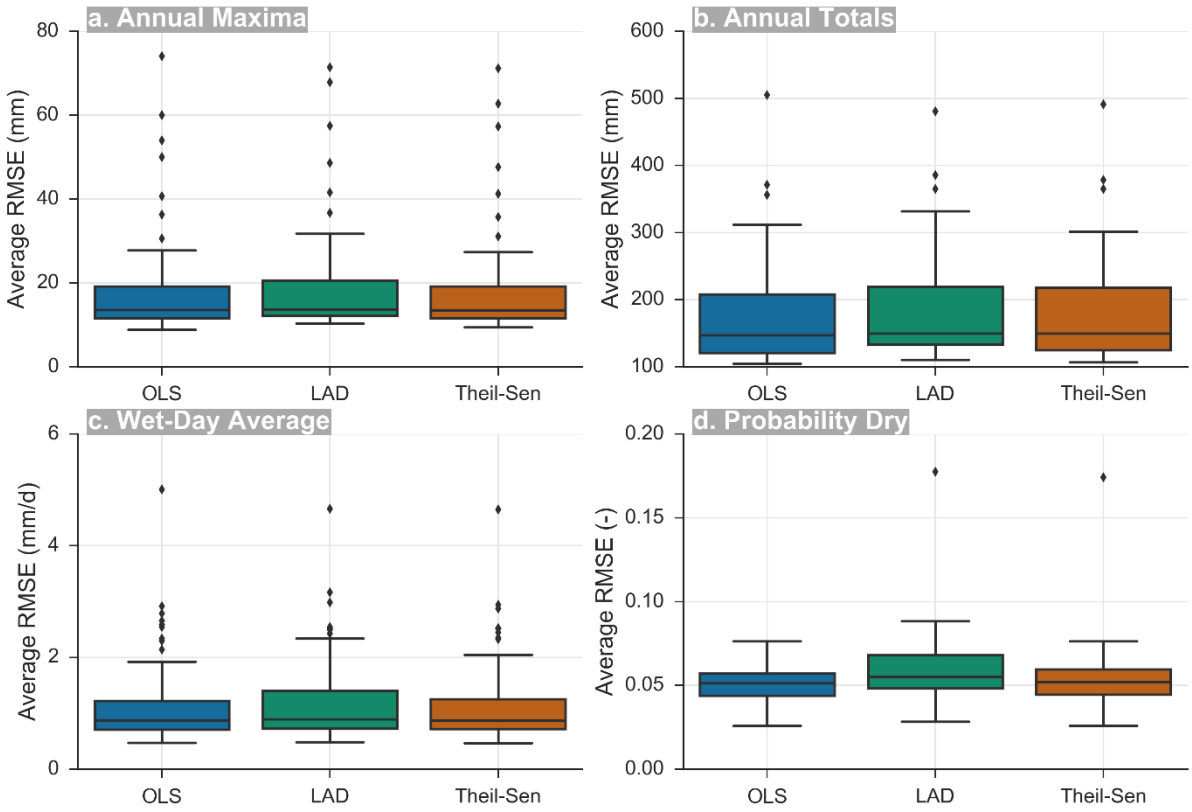


Figure B.9 Boxplots of the average prediction RMSE as estimated for each station from moving window validation of the local trend using Least Squares regression (LS), least absolute deviation regression (LAD) and the Theil-Sen regression. For the boxplots' properties description see Figure 8.5.

LIST OF PUBLICATIONS

Publications related to Thesis:

Publications in peer-reviewed journals:

1. **T. Iliopoulou**, and D. Koutsoyiannis, Projecting the future of rainfall extremes: better classic than trendy, *Journal of Hydrology*, 588, doi:10.1016/j.jhydrol.2020.125005, 2020.
2. **T. Iliopoulou**, and D. Koutsoyiannis, Revealing hidden persistence in maximum rainfall records, *Hydrological Sciences Journal*, 64 (14), 1673–1689, doi:10.1080/02626667.2019.1657578, 2019.
3. **T. Iliopoulou**, C. Aguilar , B. Arheimer, M. Bermúdez, N. Bezak, A. Ficchi, D. Koutsoyiannis, J. Parajka, M. J. Polo, G. Thirel, and A. Montanari, A large sample analysis of European rivers on seasonal river flow correlation and its physical drivers, *Hydrology and Earth System Sciences*, 23, 73–91, doi:10.5194/hess-23-73-2019, 2019.
4. **T. Iliopoulou**, D. Koutsoyiannis, and A. Montanari, Characterizing and modelling seasonality in extreme rainfall, *Water Resources Research*, 54 (9), 6242–6258, doi:10.1029/2018WR023360, 2018.
5. **T. Iliopoulou**, S.M. Papalexiou, Y. Markonis, and D. Koutsoyiannis, Revisiting long-range dependence in annual precipitation, *Journal of Hydrology*, 556, 891–900, doi:10.1016/j.jhydrol.2016.04.015, 2018.

Conference publications and presentations with evaluation of abstract:

6. **T. Iliopoulou**, and D. Koutsoyiannis: Rainfall trends in hindsight and in foresight, EGU General Assembly 2020, Online, 4–8 May 2020, https://doi.org/10.5194/egusphere-egu2020-8753, 2020
7. K. Papoulakos, **T. Iliopoulou**, P. Dimitriadis, D. Tsaknias, and D. Koutsoyiannis, Investigating the impacts of clustering of floods on insurance practices; a spatiotemporal analysis in the USA, EGU General Assembly 2020, Online, 4–8 May 2020, https://doi.org/10.5194/egusphere-egu2020-8667, 2020.
8. G. T. Manolis, K. Papoulakos, **T. Iliopoulou**, P. Dimitriadis, D. Tsaknias, and D. Koutsoyiannis, Clustering mechanisms of flood occurrence; modelling and relevance to insurance practices, EGU General Assembly 2020, Online, 4–8 May 2020, https://doi.org/10.5194/egusphere-egu2020-9357, 2020.
9. **T. Iliopoulou**, and D. Koutsoyiannis, Comparing estimators for inferring dependence from records of hydrological extremes, *European Geosciences Union General Assembly 2019, Geophysical Research Abstracts, Vol. 21*, Vienna, EGU2019-9621, European Geosciences Union, 2019.
10. T. Goulianou, K. Papoulakos, **T. Iliopoulou**, P. Dimitriadis, and D. Koutsoyiannis, Stochastic characteristics of flood impacts for agricultural insurance practices, *European Geosciences Union General Assembly 2019, Geophysical Research Abstracts, Vol. 21*, Vienna, EGU2019-5891, European Geosciences Union, 2019.

11. **T. Iliopoulou**, and D. Koutsoyiannis, A probabilistic index based on a two-state process to quantify clustering of rainfall extremes, *European Geosciences Union General Assembly 2018, Geophysical Research Abstracts, Vol. 20*, Vienna, EGU2018-9777, European Geosciences Union, 2018.
12. T. Roy, P. Dimitriadis, **T. Iliopoulou**, D. Koutsoyiannis, and E. F. Wood. "Effects of Hurst-Kolmogorov Dynamics in Intensity-Duration-Frequency Curves." In *AGU Fall Meeting Abstracts*. 2018.
13. Zoukos, **T. Iliopoulou**, P. Dimitriadis, and D. Koutsoyiannis, Global investigation of the multi-scale probabilistic behaviour of dry spells from rainfall records, *European Geosciences Union General Assembly 2018, Geophysical Research Abstracts, Vol. 20*, Vienna, EGU2018-17966-1, doi:10.13140/RG.2.2.13555.78886, European Geosciences Union, 2018.
14. V. Skoura, P. Dimitriadis, **T. Iliopoulou**, M. Crok, and D. Koutsoyiannis, A trendy analysis for the identification of extremal changes and trends in hydroclimatic processes; application to global precipitation, *European Geosciences Union General Assembly 2018, Geophysical Research Abstracts, Vol. 20*, Vienna, EGU2018-17889-1, European Geosciences Union, 2018.
15. **T. Iliopoulou**, A. Montanari, and D. Koutsoyiannis, Estimating seasonality in river flows, *European Geosciences Union General Assembly 2018, Geophysical Research Abstracts, Vol. 20*, Vienna, EGU2018-12772, European Geosciences Union, 2018.
16. P. Dimitriadis, **T. Iliopoulou**, and D. Koutsoyiannis, Simulating precipitation at a fine time scale using a single continuous-state distribution, *European Geosciences Union General Assembly 2018, Geophysical Research Abstracts, Vol. 20*, Vienna, EGU2018-18614, European Geosciences Union, 2018.
17. **T. Iliopoulou**, C. Aguilar , B. Arheimer, M. Bermúdez, N. Bezak, A. Ficchi, D. Koutsoyiannis, J. Parajka, M. J. Polo, G. Thirel, and A. Montanari, Investigating the physical basis of river memory and application to flood frequency prediction, *European Geosciences Union General Assembly 2017, Geophysical Research Abstracts, Vol. 19*, Vienna, European Geosciences Union, 2017.
18. **T. Iliopoulou**, and D. Koutsoyiannis, Investigating links between Long-Range Dependence in mean rainfall and clustering of extreme rainfall, *European Geosciences Union General Assembly 2017, Geophysical Research Abstracts, Vol. 19*, Vienna, EGU2017-9890-1, doi:10.13140/RG.2.2.25992.21763, European Geosciences Union, 2017.
19. **T. Iliopoulou**, E. Zorzetto, M. Marani, A. Montanari, and D. Koutsoyiannis. "Long rainfall records and investigation of their extremes and long-term properties." In *EGU General Assembly Conference Abstracts*, vol. 18. 2016.

Publications within wider research field:

Publications in peer-reviewed journals:

20. G.-F. Sargentis, **T. Iliopoulou**, S. Sigourou, P. Dimitriadis, D. Koutsoyiannis, Evolution of Clustering Quantified by a Stochastic Method—Case Studies on

Natural and Human Social Structures. Sustainability, 12(19), p.7972. *Sustainability*, 2020.

21. K.-G. Glynis, **T. Iliopoulou**, P. Dimitriadis, D. Koutsoyiannis, Stochastic investigation of daily air temperature extremes from a global ground station network, *Stochastic Environmental Research and Risk Assessment*, in review, 2020.
22. G.-F. Sargentis, P. Dimitriadis, R. Ioannidis, **T. Iliopoulou**, E. Frangedaki, and D. Koutsoyiannis, Optimal utilization of water resources for local communities in mainland Greece (case study of Karyes, Peloponnese), *Procedia Manufacturing*, 44, 253–260, doi:10.1016/j.promfg.2020.02.229, 2020.
23. R. Ioannidis, **T. Iliopoulou**, C. Iliopoulou, L. Katikas, A. Petsou, M.-E. Merakou, M.-E. Asimomiti, N. Pelekanos, G. Koudouris, P. Dimitriadis, C. Plati, E. Vlahogianni, K. Kepaptsoglou, N. Mamassis, and D. Koutsoyiannis, Solar-powered bus route: introducing renewable energy into a university campus transport system, *Advances in Geosciences*, 49, doi:10.5194/adgeo-49-215-2019, 2019.
24. G.-F. Sargentis, P. Dimitriadis, R. Ioannidis, **T. Iliopoulou**, and D. Koutsoyiannis, Stochastic evaluation of landscapes transformed by renewable energy installations and civil works, *Energies*, 12 (4), 2817, doi:10.3390/en12142817, 2019.
25. Koskinas, A. Tegos, P. Tsira, P. Dimitriadis, **T. Iliopoulou**, P. Papanicolaou, D. Koutsoyiannis, and T. Williamson, Insights into the Oroville Dam 2017 spillway incident, *Geosciences*, 9 (37), doi:10.3390/geosciences9010037, 2019.
26. G. Koudouris, P. Dimitriadis, **T. Iliopoulou**, N. Mamassis, and D. Koutsoyiannis, A stochastic model for the hourly solar radiation process for application in renewable resources management, *Advances in Geosciences*, 45, 139–145, doi:10.5194/adgeo-45-139-2018, 2018.
27. Klousakou, M. Chalakatevaki, P. Dimitriadis, **T. Iliopoulou**, R. Ioannidis, G. Karakatsanis, A. Efstratiadis, N. Mamassis, R. Tomani, E. Chardavellas, and D. Koutsoyiannis, A preliminary stochastic analysis of the uncertainty of natural processes related to renewable energy resources, *Advances in Geosciences*, 45, 193–199, doi:10.5194/adgeo-45-193-2018, 2018.
28. Tyralis, P. Dimitriadis, D. Koutsoyiannis, P.E. O'Connell, K. Tzouka, and **T. Iliopoulou**, On the long-range dependence properties of annual precipitation using a global network of instrumental measurements, *Advances in Water Resources*, 111, 301–318, doi:10.1016/j.advwatres.2017.11.010, 2018.
29. Koudouris, P. Dimitriadis, **T. Iliopoulou**, N. Mamassis, and D. Koutsoyiannis, Investigation on the stochastic nature of the solar radiation process, *Energy Procedia*, 125, 398–404, 2017.
30. M. Chalakatevaki, P. Stamou, S. Karali, V. Daniil, P. Dimitriadis, K. Tzouka, **T. Iliopoulou**, D. Koutsoyiannis, P. Papanicolaou, and N. Mamassis, Creating the electric energy mix in a non-connected island, *Energy Procedia*, 125, 425–434, doi:10.1016/j.egypro.2017.08.089, 2017.
31. Papoulakos, G. Pollakis, Y. Moustakis, A. Markopoulos, **T. Iliopoulou**, P. Dimitriadis, D. Koutsoyiannis, and A. Efstratiadis, Simulation of water-energy

fluxes through small-scale reservoir systems under limited data availability, *Energy Procedia*, 125, 405–414, doi:10.1016/j.egypro.2017.08.078, 2017.

Book chapters and fully evaluated conference publications

32. N. Mamassis, A. Efstratiadis, P. Dimitriadis, **T. Iliopoulou**, R. Ioannidis, and D. Koutsoyiannis, Water and Energy, Handbook of Water Resources Management: Discourses, Concepts and Examples, edited by J. Bogardi, K. D. Wasantha, R. R. P. Nandalal, R. van Nooyen, and A. Bhaduri, Chapter 20, Springer Nature, Switzerland, 2020, (in press).

Conference publications and presentations with evaluation of abstract

33. V. Kourakos, **T. Iliopoulou**, P. Dimitriadis, D. Koutsoyiannis, V. Kaleris, and A. Langousis, Investigation of marginal distribution and dependence structure of simulated streamflow by a rainfall-runoff model, EGU General Assembly 2020, Online, 4–8 May 2020, https://doi.org/10.5194/egusphere-egu2020-11402, 2020.
34. K.-G. Glynis, **T. Iliopoulou**, P. Dimitriadis, and D. Koutsoyiannis, Investigating the tail behaviour of surface temperature on global scale using K- moments, European Geosciences Union General Assembly 2019, Geophysical Research Abstracts, Vol. 21, Vienna, European Geosciences Union, 2019.
35. S. Vavoulogiannis, N. Ioannidis, **T. Iliopoulou**, P. Dimitriadis, and D. Koutsoyiannis, Stochastic investigation of rainfall and runoff series from a large hydrometeorological dataset, European Geosciences Union General Assembly 2019, Geophysical Research Abstracts, Vol. 21, Vienna, European Geosciences Union, 2019.
36. M. Karataraki, A. Thanasko, K. Printziou, G. Koudouris, R. Ioannidis, **T. Iliopoulou**, P. Dimitriadis, C. Plati, and D. Koutsoyiannis, Campus solar roads: a feasibility analysis, European Geosciences Union General Assembly 2019, Geophysical Research Abstracts, Vol. 21, Vienna, EGU2019-15648-2, European Geosciences Union, 2019.
37. M.-E. Asimomiti, N. Pelekanos, P. Dimitriadis, **T. Iliopoulou**, E. Vlahogianni, and D. Koutsoyiannis, Campus solar roads: Stochastic modelling of passenger demand, European Geosciences Union General Assembly 2019, Geophysical Research Abstracts, Vol. 21, Vienna, EGU2019-10585, European Geosciences Union, 2019.
38. Petsou, M.-E. Merakou, **T. Iliopoulou**, C. Iliopoulou, P. Dimitriadis, R. Ioannidis, K. Kepaptsoglou, and D. Koutsoyiannis, Campus solar roads: Optimization of solar panel and electric charging station location for university bus route, European Geosciences Union General Assembly 2019, Geophysical Research Abstracts, Vol. 21, Vienna, EGU2019-10832, European Geosciences Union, 2019.
39. R. Ioannidis, P. Dimitriadis, G.-F. Sargentis, E. Frangedaki, **T. Iliopoulou**, and D. Koutsoyiannis, Stochastic similarities between hydrometeorogical and art processes for optimizing architecture and landscape aesthetic parameters, European Geosciences Union General Assembly 2019, Geophysical Research Abstracts, Vol. 21, Vienna, EGU2019-11403, European Geosciences Union, 2019.

40. M. Chalakatevaki, E. Klousakou, P. Dimitriadis, **T. Iliopoulou**, and D. Koutsoyiannis, Stochastic investigation of the uncertainty of hydrometeorological processes by means of the climacogram, European Geosciences Union General Assembly 2018, Geophysical Research Abstracts, Vol. 20, Vienna, EGU2018-17714-1, European Geosciences Union, 2018.

41. S. Sigourou, P. Dimitriadis, **T. Iliopoulou**, R. Ioannidis, A. Skopeliti, K. Sakellari, G. Karakatsanis, L. Tsoulos, and D. Koutsoyiannis, Statistical and stochastic comparison of climate change vs. urbanization, European Geosciences Union General Assembly 2018, Geophysical Research Abstracts, Vol. 20, Vienna, EGU2018-18608-2, European Geosciences Union, 2018.

42. P. Dimitriadis, H. Tyralis, **T. Iliopoulou**, K. Tzouka, Y. Markonis, N. Mamassis, and D. Koutsoyiannis, A climacogram estimator adjusted for timeseries length; application to key hydrometeorological processes by the Köppen-Geiger classification, European Geosciences Union General Assembly 2018, Geophysical Research Abstracts, Vol. 20, Vienna, EGU2018-17832, European Geosciences Union, 2018.

43. Gkolemis, P. Dimitriadis, G. Karakatsanis, **T. Iliopoulou**, and D. Koutsoyiannis, A stochastic investigation of the intermittent behaviour of wind; application to renewable energy resources management, European Geosciences Union General Assembly 2018, Geophysical Research Abstracts, Vol. 20, Vienna, EGU2018-15979-3, European Geosciences Union, 2018.

44. M. Nezi, P. Dimitriadis, A. Pizarro, **T. Iliopoulou**, and D. Koutsoyiannis, Stochastic investigation of the streamflow process adjusted for human impact, European Geosciences Union General Assembly 2018, Geophysical Research Abstracts, Vol. 20, Vienna, EGU2018-17473-1, European Geosciences Union, 2018.

45. G. Koudouris, P. Dimitriadis, **T. Iliopoulou**, G. Karakatsanis, and D. Koutsoyiannis, A stochastic model for hourly solar radiation process applied in renewable resources management, European Geosciences Union General Assembly 2018, Geophysical Research Abstracts, Vol. 20, Vienna, EGU2018-16275-2, European Geosciences Union, 2018.

46. E. Klousakou, M. Chalakatevaki, R. Tomani, P. Dimitriadis, A. Efstratiadis, **T. Iliopoulou**, R. Ioannidis, N. Mamassis, and D. Koutsoyiannis, Stochastic investigation of the uncertainty of atmospheric processes related to renewable energy resources, European Geosciences Union General Assembly 2018, Geophysical Research Abstracts, Vol. 20, Vienna, EGU2018-16982-2, European Geosciences Union, 2018.

47. P. Dimitriadis, **T. Iliopoulou**, H. Tyralis, and D. Koutsoyiannis, Identifying the dependence structure of a process through pooled timeseries analysis, IAHS Scientific Assembly 2017, Port Elizabeth, South Africa, IAHS Press, Wallingford – International Association of Hydrological Sciences, 2017.

48. V. Daniil, G. Pouliasis, E. Zacharopoulou, E. Demetriou, G. Manou, M. Chalakatevaki, I. Parara, C. Georganta, P. Stamou, S. Karali, E. Hadjimitsis, G. Koudouris, E. Moschos, D. Roussis, K. Papoulakos, A. Koskinas, G. Pollakis, N. Gournari, K. Sakellari, Y. Moustakis, N. Mamassis, A. Efstratiadis, H. Tyralis, P. Dimitriadis, **T. Iliopoulou**, G. Karakatsanis, K. Tzouka, I. Deligiannis, V. Tsoukala,

P. Papanicolaou, and D. Koutsoyiannis, The uncertainty of atmospheric processes in planning a hybrid renewable energy system for a non-connected island, European Geosciences Union General Assembly 2017, Geophysical Research Abstracts, Vol. 19, Vienna, EGU2017-16781-4, doi:10.13140/RG.2.2.29610.62406, European Geosciences Union, 2017.

49. P. Stamou, S. Karali, M. Chalakatevaki, V. Daniil, K. Tzouka, P. Dimitriadis, **T. Iliopoulou**, P. Papanicolaou, D. Koutsoyiannis, and N. Mamassis, Creating the electric energy mix of a non-connected Aegean island, European Geosciences Union General Assembly 2017, Geophysical Research Abstracts, Vol. 19, Vienna, EGU2017-10130-10, doi:10.13140/RG.2.2.36537.77927, European Geosciences Union, 2017.
50. G. Koudouris, P. Dimitriadis, **T. Iliopoulou**, N. Mamassis, and D. Koutsoyiannis, Investigation of the stochastic nature of solar radiation for renewable resources management, European Geosciences Union General Assembly 2017, Geophysical Research Abstracts, Vol. 19, Vienna, EGU2017-10189-4, doi:10.13140/RG.2.2.16215.06564, European Geosciences Union, 2017.
51. K. Papoulakos, G. Pollakis, Y. Moustakis, A. Markopoulos, **T. Iliopoulou**, P. Dimitriadis, D. Koutsoyiannis, and A. Efstratiadis, Simulation of water-energy fluxes through small-scale reservoir systems under limited data availability, European Geosciences Union General Assembly 2017, Geophysical Research Abstracts, Vol. 19, Vienna, 19, EGU2017-10334-4, European Geosciences Union, 2017.
52. P. Dimitriadis, Y. Markonis, **T. Iliopoulou**, E. Feloni, N. Gournari, I. Deligiannis, P. Kastis, C. Nasika, E. Lerias, Y. Moustakis, A. Petsiou, A. Sotiriadou, A. Markopoulos, V. Tyrogiannis, and D. Koutsoyiannis, Stochastic similarities between hydroclimatic processes for variability characterization, European Geosciences Union General Assembly 2016, Geophysical Research Abstracts, Vol. 18, Vienna, European Geosciences Union, 2016.
53. E. Lerias, A. Kalamioti, P. Dimitriadis, Y. Markonis, **T. Iliopoulou**, and D. Koutsoyiannis, Stochastic investigation of temperature process for climatic variability identification, European Geosciences Union General Assembly 2016, Geophysical Research Abstracts, Vol. 18, Vienna, EGU2016-14828-3, European Geosciences Union, 2016.
54. Sotiriadou, A. Petsiou, E. Feloni, P. Kastis, **T. Iliopoulou**, Y. Markonis, H. Tyralis, P. Dimitriadis, and D. Koutsoyiannis, Stochastic investigation of precipitation process for climatic variability identification, European Geosciences Union General Assembly 2016, Geophysical Research Abstracts, Vol. 18, Vienna, EGU2016-15137-5, doi:10.13140/RG.2.2.28955.46881, European Geosciences Union, 2016.