



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΠΟΛΙΤΙΚΩΝ ΜΗΧΑΝΙΚΩΝ

**ΤΟΜΕΑΣ ΥΔΑΤΙΚΩΝ ΠΟΡΩΝ ΚΑΙ
ΠΕΡΙΒΑΛΛΟΝΤΟΣ**

ΣΤΑΤΙΣΤΙΚΗ ΑΝΑΛΥΣΗ ΠΛΗΜΜΥΡΙΚΩΝ ΠΑΡΟΧΩΝ ΣΕ ΠΑΓΚΟΣΜΙΑ ΚΛΙΜΑΚΑ



Διπλωματική Εργασία

Μαρκόπουλος Απόστολος

**Επιβλέπων Καθηγητής: Κουτσογιάννης Δημήτρης, Καθηγητής
Ε.Μ.Π.**

Αθήνα, Ιούλιος 2017

ΕΥΧΑΡΙΣΤΙΕΣ

Ολοκληρώνοντας τη διπλωματική μου εργασία, θα ήθελα πρώτα και κύρια να ευχαριστήσω θερμά τον επιβλέποντα καθηγητή μου, κ. Κουτσογιάννη Δημήτρη, για τις χρήσιμες συμβουλές του, τον χρόνο αλλά και την καθοριστική καθοδήγησή του καθ' όλη τη διάρκεια εκπόνησης της παρούσας εργασίας.

Επίσης, θα ήταν μεγάλη παράλειψη να μην ευχαριστήσω και τον υποψήφιο διδάκτορα, Δημητριάδη Παναγιώτη, του οποίου η σημαντικότερη βοήθεια και ο χρόνος που μου αφιέρωσε κατέστησαν δυνατή την ολοκλήρωση της διπλωματικής μου.

Ευχαριστώ, επίσης, όλους εκείνους του ανθρώπους που με τον τρόπο τους έπαιξαν καταλυτικό ρόλο, όχι μόνο για να μην χάσω το κουράγιο και την θέλησή μου αυτούς τους μήνες, αλλά με έκαναν να πιστέψω στην αξία αυτής της εργασίας. Μάριε, Κική, Βαγγέλη, Γιάννη, Άγγελε, Αλέκο και Μάγια (και σίγουρα πολλοί άλλοι), η στήριξή σας όχι μόνο αυτή την περίοδο αλλά σε όλη τη διάρκεια της φοιτητικής μου ζωής ήταν ανεκτίμητη.

Τέλος, οφείλω να ευχαριστήσω την οικογένειά μου η οποία μου προσέφερε απεριόριστη ψυχολογική υποστήριξη και συμπαράσταση όλα αυτά τα χρόνια και δεν σταμάτησε ποτέ να πιστεύει σε μένα. Η προσπάθεια να μοιάσω έστω και στο ελάχιστο στους γονείς μου και στην αγαπημένη μου αδερφή, είναι αυτή που μου έδινε και θα μου δίνει κίνητρο στα πάντα. Δεν θα μπορούσα όμως να μην αφιερώσω τη διπλωματική αυτή εργασία, στην γιαγιά μου Αθηνά (που έφυγε κατά την διάρκεια εκπόνησής της) και η οποία δεν σταμάτησε να μου λέει «να πάρω το χαρτί».

Απόστολος Μαρκόπουλος

Ιούλιος 2017

ΠΕΡΙΕΧΟΜΕΝΑ

ΕΥΧΑΡΙΣΤΙΕΣ	iii
ΠΕΡΙΕΧΟΜΕΝΑ.....	v
ΠΕΡΙΛΗΨΗ	ix
ABSTRACT	xi
1 ΕΙΣΑΓΩΓΗ.....	1
1.1 Γενικά.....	1
1.2 Αντικείμενο της εργασίας.....	2
1.3 Διάρθρωση εργασίας.....	3
2 ΕΙΣΑΓΩΓΗ ΣΤΗΝ ΣΤΑΤΙΣΤΙΚΗ ΑΝΑΛΥΣΗ ΥΔΡΟΛΟΓΙΚΩΝ ΔΕΔΟΜΕΝΩΝ	5
2.1 Προκαταρκτική Ανάλυση Δεδομένων	5
2.2 Επιλογή δεδομένων με σκοπό την ανάλυση τάσεων και μεταβολών	6
2.3 Αναλύσεις τάσεων παρατηρημένων πλημμυρικών επεισοδίων στο πλαίσιο ενός κλίματος που αλλάζει	7
3 ΤΥΠΙΚΗ ΣΤΑΤΙΣΤΙΚΗ ΚΑΙ ΧΡΟΝΙΚΗ ΕΞΑΡΤΗΣΗ.....	11
3.1 Η έννοια της απλής στοχαστικής ανέλιξης.....	11
3.1.1 Ανέλιξη στη συναθροισμένη κλίμακα.....	13

3.2	Απλά στοχαστικά μοντέλα.....	14
3.2.1	Το μοντέλο του Λευκού θορύβου	14
3.2.2	Το μοντέλο AR(1).....	14
3.3	Υπόθεση της ανεξαρτησίας, η εξαίρεση παρά ο κανόνας	16
3.3.1	Το φαινόμενο της μακροπρόθεσμης εμμονής	19
3.3.2	Μοντέλα διατήρησης της εμμονής και μέθοδοι εκτίμησης του συντελεστή Hurst	20
3.3.3	Η επιρροή της εμμονής στο υπολογισμό των στατιστικών παραμέτρων.....	21
3.3.4	Το Mann-Kendall trend test υπό το πρίσμα της δυναμικής Hurst-Kolmogorov	23
4	ΠΡΟΣΑΡΜΟΓΗ ΣΥΝΑΡΤΗΣΕΩΝ ΚΑΤΑΝΟΜΗΣ ΣΕ ΔΕΔΟΜΕΝΑ ΑΚΡΑΙΩΝ ΤΙΜΩΝ	27
4.1	Εισαγωγή στη θεωρία ακραίων τιμών.....	27
4.2	Συναρτήσεις Κατανομών Πιθανότητας Ακραίων Τιμών και Μέθοδοι εκτίμησης των παραμέτρων τους	29
4.2.1	Η Γενικευμένη Ακραίων Τιμών- ΓΑΤ	29
4.2.2	Η Λογαριθμοκανονική Κατανομή	30
4.2.3	Η Κατανομή Γάμμα δύο παραμέτρων.....	30

4.2.4	Η κατανομή Γάμμα τριών παραμέτρων (Pearson III)	30
4.2.5	Η Κατανομή Log-Pearson III	31
4.2.6	Μέθοδοι εκτίμησης των παραμέτρων των κατανομών.....	32
4.3	Κριτήρια επιλογής της καταλληλότερης κατανομής.....	34
5	ΕΦΑΡΜΟΓΗ ΣΕ ΠΡΑΓΜΑΤΙΚΑ ΔΕΔΟΜΕΝΑ	39
5.1	Περιγραφή Βάσης Δεδομένων	39
5.2	Επιλογή κατάλληλων δεδομένων προς επεξεργασία	43
5.3	Στατιστική επεξεργασία των σταθμών.....	47
5.4	Ανθρώπινος παράγοντας.....	54
5.5	Ανάλυση τάσεων με τη μέθοδο της γραμμικής παλινδρόμησης	62
5.6	Ανάλυση τάσεων με το Mann-Kendall trend test	67
5.7	Προσαρμογή Επιλεγμένων Συναρτήσεων Κατανομής	83
6	ΣΥΜΠΕΡΑΣΜΑΤΑ.....	89
6.1	Σύνοψη μελέτης και συμπεράσματα.....	89
6.2	Προτάσεις για μελλοντική έρευνα.....	92
7	ΒΙΒΛΙΟΓΡΑΦΙΚΕΣ ΑΝΑΦΟΡΕΣ	93

ΠΕΡΙΛΗΨΗ

Η στατιστική ανάλυση των πλημμυρικών παροχών σε παγκόσμια ή/και τοπική κλίμακα συγκεντρώνει τα τελευταία χρόνια όλο και μεγαλύτερο ενδιαφέρον στους επιστημονικούς κύκλους, ειδικά στα πλαίσια ενός κλίματος που αλλάζει. Για τη συγκεκριμένη στατιστική ανάλυση χρησιμοποιήθηκαν δεδομένα μεγίστων στιγμιαίων παροχών από τον Παγκόσμιο Κατάλογο Μεγίστων Παρατηρημένων Πλημμυρών (World's Catalogue of Maximum Observed Floods, IAHS, 2003). Επιλέχθηκαν 125 σταθμοί παγκοσμίως, βάσει κριτηρίων ποιότητας και πληρότητας των καταγραφών τους. Δόθηκε ιδιαίτερο βάρος στην ανάλυση τάσεων και για το σκοπό αυτό πραγματοποιήθηκε εκτεταμένη έρευνα σε κάθε σταθμό ξεχωριστά για την ύπαρξη ταμειυτήρων ανάντη των σταθμών μέτρησης. Οι μέθοδοι που χρησιμοποιήθηκαν για την ανάλυση αυτή, ήταν η μέθοδος της γραμμικής παλινδρόμησης και το τεστ ανάλυσης τάσεων Mann-Kendall. Στους σταθμούς που παρατηρήθηκε ισχυρή δομή συσχέτισης το Mann-Kendall τεστ τροποποιήθηκε με μια συγκεκριμένη μεθοδολογία που προτείνεται. Για τον προσδιορισμό της δομής συσχέτισης των δεδομένων υπολογίστηκαν οι συντελεστές αυτοσυσχέτισης και οι συντελεστές Hurst. Οι συντελεστές μακροπρόθεσμης εμμονής υπολογίστηκαν με τη μέθοδο της συναθροισμένης διασποράς (κλιμακόγραμμα). Στα δεδομένα προσαρμόστηκαν τέσσερις συναρτήσεις κατανομής πιθανότητας της οικογένειας ακραίων τιμών και η καταλληλότητα της προσαρμογής τους ελέγχθηκε από τρία διαφορετικά κριτήρια, το Κριτήριο Πληροφορίας Akaike (AIC), το Μπεϋζιανό Κριτήριο Πληροφορίας (BIC) και το κριτήριο Anderson-Darling τροποποιημένο (ADC).

ABSTRACT

Major scientific interest is raised in statistical analysis of large floods on global or/and regional scale, especially in the framework of a changing climate. A global data base of flood discharges (World's Catalogue of Maximum Observed Floods, IAHS, 2013) is processed. Based on length and quality criteria, 125 stations worldwide are chosen for the analysis. Emphasis was given to trend detection, and therefore a thorough research was carried out to each station individually to locate major dams and reservoirs upstream of the stations. For the trend analysis two methods have been applied, linear regression and Mann-Kendall trend test. For stations that were strongly correlated, the Mann-Kendall trend test was modified with a specific methodology that is being proposed. Autocorrelation and Hurst parameters were used to quantify the correlation structure of the data set. Hurst parameter was calculated by applying the aggregated variance method. Four different probability distribution functions, belonging to the extreme value family, were fitted to all timeseries. The goodness-of-fit test that were used, were the Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC) and a modification of the well-known Anderson-Darling Criterion (ADC).

1 ΕΙΣΑΓΩΓΗ

1.1 Γενικά

Μεγάλες και καταστροφικές πλημμύρες λαμβάνουν χώρα κάθε χρόνο παγκοσμίως. Αξιοσημείωτα πλημμυρικά επεισόδια έπληξαν το καλοκαίρι του 2010 περιοχές του Πακιστάν, της Ινδίας και της Κίνας, την Κολομβία από τον Οκτώβριο μέχρι το Δεκέμβριο του 2010 και την Αυστραλία κατά τη διάρκεια του νότιου καλοκαιριού μεταξύ 2010 και 2011. Οι οικονομικές απώλειες στην Κίνα λόγω των προαναφερθέντων πλημμυρών εκτιμήθηκαν στα 51 δισεκατομμύρια δολάρια. Το 2010 και πάλι, προκλήθηκαν περίπου 2000 θάνατοι λόγω πλημμυρών κατά τη διάρκεια των μουσώνων στο Πακιστάν. Το 2011 ακραία πλημμυρικά επεισόδια αναφέρθηκαν στη Μοζαμβίκη, τη Ναμίμπια, τη Βραζιλία, την Κολομβία, το Μεξικό και τις ΗΠΑ με τις ανθρώπινες απώλειες να ξεπερνούν τις 50 σε κάθε χώρα από ένα σύνολο 1000 αθροιστικά και με τεράστιες υλικές καταστροφές.

Το Κέντρο Επιδημιολογίας Καταστροφών (Centre for the Epidemiology of Disasters), από το οποία πάρθηκαν τα παραπάνω στοιχεία, αναφέρει αύξηση στις καταστροφές και τις απώλειες λόγω πλημμυρών στο πέρασμα του χρόνου. Παρόλα αυτά, μέρος αυτής της αύξησης μπορεί να αποδοθεί σε βελτιώσεις ως προς τα μέσα που αναφέρονται πλέον τέτοια γεγονότα, στην παγκόσμια πληθυσμιακή αύξηση αλλά και στη συνεχώς αυξανόμενη αστικοποίηση περιοχών ανά τον κόσμο.

Οι παράγοντες που επηρεάζουν το μηχανισμό γένεσης των πλημμυρικών επεισοδίων είναι πάρα πολλοί. Κυρίαρχο ρόλο παίζει η βροχόπτωση ως προς την ένταση, τη διάρκεια, τη χρονική τοποθέτηση του συμβάντος, κ.α. Επιπρόσθετα, οι συνθήκες που επικρατούν στη λεκάνη απορροής έχουν ιδιαίτερη σημασία, όπως η εδαφική υγρασία και η υδατοπερατότητα, η βαθμός αστικοποίησης, η ύπαρξη φραγμάτων και ταμιευτήρων, κλπ. Έντονα και μεγάλης διάρκειας φαινόμενα βροχόπτωσης αποτελούν τη βασική αιτία μιας

πλημμύρας κυρίως σε μεγάλης έκτασης λεκάνες απορροής, ενώ σε μικρής έκτασης λεκάνες ακόμα και έντονες αλλά μικρής διάρκειας πλημμύρες έχει παρατηρηθεί ότι οδήγησαν σε ακραία πλημμυρικά επεισόδια. Η σχέση, επομένως, της βροχής και της απορροής σε μία λεκάνη απορροής είναι μια περίπλοκη διαδικασία που απαιτεί μελέτη σε βάθος και ακρίβεια.

Λαμβάνοντας υπόψη τις επιπτώσεις σε υλικές και ανθρώπινες απώλειες, τα τελευταία χρόνια η ανάλυση τέτοιων φαινομένων παρουσιάζει ιδιαίτερο ενδιαφέρον ανάμεσα στους επιστημονικούς κύκλους. Τέλος, η σημαντικότητα αυτών των μελετών αποκτά ακόμα πιο ιδιαίτερο χαρακτήρα όταν οι μεγάλες πλημμύρες αποδίδονται από μερίδα της επιστημονικής κοινότητας σε μια «αφύσικη» και «καταστροφική» κλιματική αλλαγή.

1.2 Αντικείμενο της εργασίας

Στα πλαίσια της παρούσας εργασίας εξετάζονται δεδομένα πλημμυρικών παροχών από 371 σταθμούς παγκοσμίως που έχουν ληφθεί από τον Παγκόσμιο Κατάλογο Πλημμυρικών Παροχών (World's Catalogue of Maximum Observed Floods, IAHS, 2003). Διατίθενται χρονοσειρές οι οποίες παρέχουν τη μέγιστη καταγεγραμμένη στιγμιαία παροχή ανά έτος για κάθε σταθμό. Αναλύονται ως προς την ποιότητα τους σε συνάρτηση με το πλήθος των κενών τους τιμών και το μήκος των καταγραφών τους. Έτσι, επιλέγονται 125 σταθμοί οι οποίοι εξετάζονται στατιστικά και δίνεται ιδιαίτερο βάρος στην ανάλυση τάσεων, η οποία προσεγγίζεται με δύο διαφορετικές μεθόδους. Για την καλύτερη αξιοπιστία και ακρίβεια των αποτελεσμάτων, προηγείται της ανάλυσης τάσεων μία ενδεδειγμένη έρευνα σε κάθε έναν από τους επιλεγμένους σταθμούς με σκοπό τον αποκλεισμό των ποταμών εκείνων που ανάντη των σημείων μέτρησης υπάρχουν ταμειυτήρες που έχουν επηρεάσει τη δίαίτά τους. Για την εφαρμογή των μεθόδων ανάλυσης τάσεων στους εναπομείναντες, «καθαρούς» από ανθρώπινες επεμβάσεις, σταθμούς δόθηκε ιδιαίτερη προσοχή στη δομή συσχέτισης των δεδομένων, μέσω του συντελεστή αυτοσυσχέτισης και του συντελεστή μακροπρόθεσμης

εμμονής. Τέλος, επιχειρείται και η προσαρμογή τεσσάρων κατανομών πιθανότητας στα δεδομένα και η καταλληλότητα της προσαρμογής τους ελέγχεται μέσω τριών διαφορετικών κριτηρίων.

1.3 Διάρθρωση εργασίας

Η εργασία αποτελείται από επτά διαφορετικά κεφάλαια, συμπεριλαμβανομένης της εισαγωγής.

Στο δεύτερο κεφάλαιο αναφέρονται οι βασικές μέθοδοι που χρησιμοποιούνται στην ανάλυση στατιστικών δεδομένων και γίνεται μια σύντομη αναφορά σε παρόμοιες μελέτες.

Στο τρίτο κεφάλαιο δίνονται οι βασικές έννοιες της στοχαστικής υδρολογίας και εξηγούνται διάφορες έννοιες οι οποίες χρησιμοποιούνται στη συνέχεια στην ανάλυση. Το ενδιαφέρον εστιάζεται στις επιπτώσεις που υπάρχουν στις στατιστικές εκτιμήσεις παραμέτρων μέσω της κλασικής στατιστικής όταν αναλύονται δεδομένα που παραβιάζουν την υπόθεση της ανεξαρτησίας.

Στο τέταρτο κεφάλαιο ο αναγνώστης εισάγεται στη Θεωρία Ακραίων Τιμών και παρουσιάζονται μέσω της μαθηματικής τους προσέγγισης οι κατανομές εκείνες που συνήθως προσαρμόζονται σε δεδομένα ακραίων τιμών. Επίσης αναφέρονται τα κριτήρια που χρησιμοποιήθηκαν στη μελέτη για την καταλληλότητα προσαρμογής των κατανομών.

Στο πέμπτο κεφάλαιο παρουσιάζεται η μεθοδολογία που εφαρμόστηκε στα πραγματικά δεδομένα και τα αποτελέσματα όλων των αναλύσεων που έγιναν.

Στο έκτο κεφάλαιο γίνεται η σύνοψη της εργασίας και εξάγονται τα τελικά συμπεράσματα καθώς και προτάσεις για μελλοντική έρευνα.

Στο έβδομο κεφάλαιο δίνεται η λίστα με όλες τις βιβλιογραφικές αναφορές.

2 ΕΙΣΑΓΩΓΗ ΣΤΗΝ ΣΤΑΤΙΣΤΙΚΗ ΑΝΑΛΥΣΗ ΥΔΡΟΛΟΓΙΚΩΝ ΔΕΔΟΜΕΝΩΝ

Στο παρόν κεφάλαιο γίνεται αναφορά στους τρόπους και στις βασικές μεθοδολογίες που συχνότερα υιοθετούνται από τους μελετητές για την εξαγωγή αξιόπιστων αποτελεσμάτων μιας στατιστικής ανάλυσης υδρολογικών δεδομένων. Πρέπει να γίνει σαφές πως καθοριστικό ρόλο, πέρα από τα εργαλεία που θα χρησιμοποιηθούν κατά την ανάλυση, παίζουν αδιαμφισβήτητα η επιλογή των κατάλληλων δεδομένων από τον μελετητή, ο τρόπος συλλογής τους αλλά και η ποιότητα στις μεθόδους καταγραφής τους. Τέλος, γίνεται μια συνοπτική αναφορά σε προηγούμενες μελέτες στατιστικής ανάλυσης παροχών ποταμών.

2.1 Προκαταρκτική Ανάλυση Δεδομένων

Η προκαταρκτική ανάλυση δεδομένων αποτελεί μία προ-επεξεργασία της βάσης δεδομένων που διατίθεται. Αποτελεί βασικό στοιχείο σε οποιαδήποτε ανάλυση δεδομένων και περιλαμβάνει την εξαγωγή γραφημάτων για την καλύτερη διερεύνηση, κατανόηση και παρουσίαση των δεδομένων. Μία μελέτη με επίκεντρο την ανάλυση τάσεων που δεν περιλαμβάνει μία τέτοια ενδελεχή διερεύνηση των δεδομένων δεν μπορεί να θεωρηθεί πλήρης. Το πρώτο βήμα είναι η εξέταση των ακατέργαστων δεδομένων με σκοπό την αναγνώριση ιδιαίτερων χαρακτηριστικών, όπως προβλήματα των μετρήσεων (αδικαιολόγητα υψηλές τιμές, μεγάλα κενά στις μετρήσεις, κ.α.), εμφανή χρονικά μοτίβα (π.χ. εποχικότητα) και χωρικά μοτίβα. Μία τέτοια πρωταρχική ανάλυση παίζει επίσης σημαντικό ρόλο στην τεκμηρίωση ή όχι βασικών υποθέσεων της κλασικής στατιστικής, όπως η υπόθεση της ανεξαρτησίας των δεδομένων ή η αναγνώριση στατιστικών κατανομών που μπορούν να αναπαραστήσουν τα δεδομένα. Όπως προαναφέρθηκε μία τέτοια προ-επεξεργασία μπορεί να βοηθήσει τον εκάστοτε μελετητή να εντοπίσει προβληματικές τιμές

του δείγματος, όμως σε καμία περίπτωση δεν υπονοείται ότι μπορεί να αντικαταστήσει τους ελέγχους ποιότητας των δεδομένων. Η πλέον κατάλληλη και αξιόπιστη διερεύνηση των δεδομένων περιλαμβάνει τη δημιουργία και τη μελέτη διαγραμμάτων με σκοπό να αναδείξουν τα βασικά χαρακτηριστικά των δεδομένων, αλλά και να υποδείξουν περαιτέρω διαγράμματα που ενδεχομένως χρειάζονται για την κυρίως ανάλυση. Συνηθισμένοι τύποι διαγραμμάτων που είναι χρήσιμα κυρίως στην στατιστική/στοχαστική υδρολογία είναι τα ιστογράμματα, τα διαγράμματα πιθανότητας (P-P plots/Q-Q plots), τα διαγράμματα χρονοσειρών, τα αυτοσυσχετογράμματα κ.α. Περισσότερες λεπτομέρειες και εφαρμογές για την προ-επεξεργασία υδρολογικών δεδομένων δίνονται από τους Grubb and Robson(2000) και Tufte(1983).

2.2 Επιλογή δεδομένων με σκοπό την ανάλυση τάσεων και μεταβολών

Η ανάλυση τάσεων ή οποιωνδήποτε άλλων μεταβολών σε υδρολογικά δεδομένα αποτελεί ένα σημαντικό και δύσκολο εγχείρημα, που όμως τυχαίνει ολοένα και μεγαλύτερου ενδιαφέροντος τις τελευταίες δεκαετίες. Η επιλογή της κατάλληλης βάσης δεδομένων αποτελεί ίσως τον πιο ουσιαστικό παράγοντα. Τα δεδομένα οφείλουν να έχουν «υποστεί» ελέγχους ποιότητας προτού γίνει οποιαδήποτε άλλη ανάλυση. Πιο συγκεκριμένα, κάποια παραδείγματα προβλημάτων που μπορούν να προκαλέσουν μεταβολή στις διαθέσιμες χρονοσειρές είναι: τυπογραφικά σφάλματα, δυσλειτουργίες στα όργανα μέτρησης, αλλαγές στις μεθόδους και στους τρόπους μέτρησης ή ακόμα και στην τοποθεσία μέτρησης. Το ζήτημα της αναγνώρισης μιας μεταβολής σε δεδομένα πλημμυρικών παροχών είναι ιδιαίτερος περίπλοκος, αν αναλογιστεί κανείς ότι η γένεση της απορροής είναι αποτέλεσμα πολλών διαφορετικών παραγόντων, όπως η κατακρήμνιση, η έκταση της λεκάνης απορροής, οι απώλειες λόγω εξάτμισης, κ.α. Όλοι οι παραπάνω παράγοντες υποδεικνύουν την ειδική αντιμετώπιση που απαιτείται από τον εκάστοτε μελετητή για την επιλογή δεδομένων και σταθμών με απώτερο σκοπό την αναζήτηση αλλαγών και μεταβολών.

Για την αναγνώριση μεταβολών σε ένα σετ δεδομένων συνήθως χρησιμοποιούνται τα τεστ στατιστικού ελέγχου. Η εφαρμογή ενός τεστ στατιστικού ελέγχου απαιτεί των καθορισμό μιας μηδενικής και μιας εναλλακτικής υπόθεσης, οι οποίες αποτελούν διατυπώσεις ικανές να περιγράψουν το αντικείμενο που ερευνάται από το τεστ. Για παράδειγμα, για να ελεγχθεί η ύπαρξη τάσεων στη μέση τιμή μιας χρονοσειράς η μηδενική υπόθεση(H_0) θα ήταν η «μη ύπαρξης τάσης» και η εναλλακτική υπόθεση(H_1) ότι «η μέση τιμή είτε αυξάνεται είτε μειώνεται». Η διαδικασία που ακολουθείται είναι να υποτεθεί αρχικά ότι η μηδενική υπόθεση ισχύει και να ελεγχθεί στη συνέχεια εάν με βάση τα δεδομένα αυτή η υπόθεση επιβεβαιώνεται. Για τη σύγκριση μεταξύ της μηδενικής και της εναλλακτικής υπόθεσης χρησιμοποιείται, κάποια στατιστική παράμετρος και αξιολογείται το επίπεδο σημαντικότητας αυτής της παραμέτρου που έχει εξαχθεί από τα διαθέσιμα δεδομένα. Στην ουσία, αυτή η στατιστική παράμετρος είναι μια αριθμητική τιμή υπολογισμένη από τα δεδομένα που είναι ικανή να υποδείξει την διαφορά μεταξύ των δύο υποθέσεων. Ένα χαρακτηριστικό παράδειγμα μιας τέτοιας στατιστικής παραμέτρου είναι η κλίση της ευθείας που προσαρμόστηκε σε κάποια δεδομένα με την διαδικασία της γραμμικής παλινδρόμησης. Εάν δεν υπάρχει κάποια τάση (μηδενική υπόθεση) στη μέση τιμή των δεδομένων, τότε η κλίση της ευθείας αυτής πρέπει να είναι κοντά στο μηδέν. Εάν υπάρχει κάποια αξιοσημείωτη μεταβολή(εναλλακτική υπόθεση) στη μέση τιμή των δεδομένων, τότε η κλίση της ευθείας θα πρέπει να αποκλίνει αρκετά από το μηδέν, με τις θετικές τιμές να υποδεικνύουν αυξητικές τάσεις ενώ οι αρνητικές τιμές να υποδεικνύουν μειωτικές τάσεις.

2.3 Αναλύσεις τάσεων παρατηρημένων πλημμυρικών επεισοδίων στο πλαίσιο ενός κλίματος που αλλάζει

Σφοδρότατες πλημμύρες στην Ευρώπη και την Βόρεια Αμερική κατά τη διάρκεια του πρόσφατου παρελθόντος (Kunkel et al.,1994; CEH Wallingford/Met Office, 2001; Marsh & Bradford, 2003; Saurí et al.,2003; Kundzewicz, 2004) έφεραν στην επιφάνεια το ερώτημα για το εάν είναι ή όχι αποτέλεσμα ενός μεταβαλλόμενου κλίματος. Αποτελέσματα υδρολογικών

μοντέλων που ως εισόδους (inputs) χρησιμοποιούν προσομοιωμένα δεδομένα, συχνά υποδεικνύουν την αύξηση σε μέγεθος των παροχών των ποταμών στα πλαίσια ενός μελλοντικά πιο θερμού κλίματος (Miller & Russell, 1992; Nijssen et al., 2001; Reynard et al., 2001; Milly et al., 2002).

Η Διακυβερνητική Συνδιάσκεψη για την Κλιματική Αλλαγή (IPCC – Intergovernmental Panel on Climate Change το 2001 καταλήγει στο ότι ένα μεγάλο μέρος παρατηρήσεων μας παρέχει μία συγκεντρωτική εικόνα για έναν πλανήτη που ολοένα και θερμαίνεται. Στοιχεία που έχουν εξαχθεί από παρατηρήσεις υποδεικνύουν ότι είναι πιθανό ακραία φαινόμενα βροχόπτωσης να έχουν αυξηθεί στα βόρεια γεωγραφικά πλάτη της γης, ενώ η συχνότητα και η ένταση των ξηρασιών έχει ενταθεί σε κάποιες περιοχές της Ασίας και της Αφρικής.

Οι επιδράσεις μιας παγκόσμιας υπερθέρμανσης υποστηρίζεται (Trenberth, 1998) ότι δεν έχει μόνο σαν αποτέλεσμα υψηλότερες θερμοκρασίες στον πλανήτη αλλά επιδρά και στον υδρολογικό κύκλο. Έτσι, υψηλότερες θερμοκρασίες έχουν σαν αποτέλεσμα μεγαλύτερα ποσοστά εξατμίζοντος νερού αλλά και μεγαλύτερη χωρητικότητα υγρασίας στην ατμόσφαιρα. Κατά συνέπεια, εμφανίζονται εντονότερα φαινόμενα βροχόπτωσης και κατ' επέκταση και εντονότερα πλημμυρικά επεισόδια.

Η ανάλυση τάσεων όπως έχει προαναφερθεί μπορεί να αποτελέσει ένα χρήσιμο εργαλείο για την επιβεβαίωση ή μη των παραπάνω ισχυρισμών για όλο και εντονότερες πλημμύρες, αλλά με δεδομένο ότι οι αναλύσεις αυτές θα γίνονται σε πραγματικά παρατηρημένα δεδομένα. Παρακάτω παρατίθενται κάποια παραδείγματα μελετών που έχουν γίνει με επίκεντρο την ανάλυση τάσεων παροχών ποταμών, σε παγκόσμια ή/και τοπική κλίμακα. Η παράθεση αυτών των παραδειγμάτων δεν γίνεται με σκοπό την απόρριψη των υποθέσεων για πιο έντονες πλημμύρες στο μέλλον, αλλά περισσότερο για να αναδειχθεί η πολυπλοκότητα του φαινομένου αλλά και η ποικιλία των αποτελεσμάτων για τις διάφορες περιοχές του πλανήτη. Εξαιτίας των χρήσεων γης, την κατασκευή φραγμάτων και άλλων

ανθρώπων επεμβάσεων είναι πολύ δύσκολο να βρεθεί ένα σετ δεδομένων που θα μπορεί να θεωρηθεί ότι αποτυπώνει μια ανεπηρέαστη φυσική διεργασία. Παρόλα αυτά, αποτελέσματα μελετών στις οποίες χρησιμοποιήθηκαν μέσες ημερήσιες παροχές έχουν ως εξής(επίπεδο σημαντικότητας στο 95% για όλες τις μελέτες):

1. Υπάρχουν κάποια στοιχεία για μειωτικές τάσεις στα ποτάμια του δυτικού Καναδά(Adamowski & Bocci,2001; Burn & Hag Elnur, 2002; Kundzewicz et al., 2004)
2. Στη μεγαλύτερη έκταση των ΗΠΑ υπάρχουν λίγες στατιστικά σημαντικές τάσεις και όπου υπάρχουν, ποικίλλουν ως προς το πρόσημο (Douglas et al., 2000; Lins & Slack, 1999; Kundzewicz et al., 2004)
3. Στην Ευρώπη υπάρχουν ορισμένες ενδείξεις για θετικές τάσεις στη Βόρεια Σκανδιναβία(Lindström & Bergström, 2003, 2004; Kundzewicz et al., 2004), αλλά από την άλλη σε τοπικό επίπεδο δεν βρέθηκαν στατιστικά σημαντικές τάσεις στα ποτάμια του Ηνωμένου Βασιλείου(Robson et al., 1998)
4. Στην μελέτη των Kundzewicz et al.(2004) παρά το γεγονός ότι το ένα τέταρτο των παρατηρήσεων στην κεντρική Ευρώπη παρουσιάζουν στατιστικά σημαντικές τάσεις, είναι άλλες θετικές και άλλες αρνητικές.

Είναι εμφανής λοιπόν η πολυπλοκότητα του φαινομένου και των παραγόντων που συμβάλλουν στη δημιουργία της απορροής ενός ποταμού αλλά και η ποικιλία που συναντάται ανά γεωγραφική θέση. Με άλλα λόγια, θα ήταν τουλάχιστον παράτολμο, με τα δεδομένα που υπάρχουν μέχρι σήμερα, να γίνεται λόγος για έναν παγκόσμιο κανόνα αύξησης των πλημμυρών.

3 ΤΥΠΙΚΗ ΣΤΑΤΙΣΤΙΚΗ ΚΑΙ ΧΡΟΝΙΚΗ ΕΞΑΡΤΗΣΗ

Στο κεφάλαιο αυτό παρουσιάζονται κάποιες από τις βασικές έννοιες της στοχαστικής υδρολογίας και περιγράφονται οι μέθοδοι εκείνες που χρησιμοποιούνται στη συνέχεια στην ανάλυση. Εκτός αυτού, γίνεται μια σύντομη αναφορά και σε εναλλακτικές μεθόδους που έχουν χρησιμοποιηθεί για ανάλογες αναλύσεις δεδομένων με τις αντίστοιχες βιβλιογραφικές αναφορές.

3.1 Η έννοια της απλής στοχαστικής ανέλιξης

Στοχαστική ανέλιξη X_i ονομάζεται ένα σύνολο τυχαίων μεταβλητών σε ακολουθία $\{X_1, X_2, X_3, \dots, X_i\}$ όπου i είναι παράμετρος που παίρνει τιμές από ένα σύνολο T το οποίο συνήθως παριστάνει τον χρόνο. Η υλοποίηση μιας στοχαστικής ανέλιξης, δηλαδή ένα σύνολο παρατηρήσεων $x(i)$ της $X(i)$ για μεταβαλλόμενο χρόνο i ονομάζεται χρονοσειρά. Η στοχαστική ανέλιξη χαρακτηρίζεται από ένα σύνολο στατιστικών παραμέτρων με κυριότερες τις εξής (Koutsoyiannis, 2013):

- Μέση τιμή:

$$\mu := E[X_i] \quad (1)$$

- Τυπική απόκλιση:

$$\sigma := \sqrt{Var[X_i]} = \sqrt{E[(X_i - \mu)^2]} \quad (2)$$

- Αυτοσυνδιασπορά:

$$\gamma_j = Cov[X_i, X_{i+j}] = E[(X_i - \mu)(X_{i+j} - \mu)] \quad (3)$$

- Συντελεστής αυτοσυσχέτισης:

$$\rho(i, j) = \frac{Cov[X_i, X_{i+j}]}{\sqrt{Var[X_i]Var[X_{i+j}]}} \quad (4)$$

Ο συντελεστής αυτοσυσχέτισης (4) ανήκει στο διάστημα[-1,1] με την τιμή 0 να δηλώνει πλήρως ασυσχέτιστη ανέλιξη και τις τιμές {-1,1} να δηλώνουν απόλυτη αρνητική ή θετική συσχέτιση, αντίστοιχα. Η τάξη της αυτοσυσχέτισης που μελετάται στις υδρολογικές εφαρμογές αφορά την κλίμακα ενδιαφέροντος. Για παράδειγμα, για την μελέτη ετήσιας απορροής ενός ποταμού η χρονική κλίμακα συνηθίζεται να είναι από ένα ή και περισσότερα έτη. Αντίστοιχα, προσομοίωση βροχοπτώσεων σε μια μικρή κλίμακα θα σήμαινε τιμές για την κλίμακα ενδιαφέροντος από μερικά λεπτά έως και ώρες (Koutsoyiannis, 2002).

Γενικά οι στατιστικές παράμετροι μιας στοχαστικής ανέλιξης, όπως η μέση τιμή και η τυπική απόκλιση, μπορεί να μεταβάλλονται με το χρόνο. Ιδιαίτερο ενδιαφέρον, όμως, παρουσιάζουν οι στοχαστικές εκείνες ανελιξεις στις οποίες δεν παρατηρείται κάποια μεταβολή των στατιστικών τους χαρακτηριστικών με την πάροδο του χρόνου. Μια στοχαστική ανέλιξη λέγεται *στάσιμη με την ευρεία έννοια*, όταν η μέση τιμή της είναι σταθερή και η αυτοσυνδιασπορά της εξαρτάται μόνο από τη διαφορά του χρόνου. Μια στάσιμη στοχαστική ανέλιξη είναι *εργοδική* αν κάθε παράμετρος της κατανομής μπορεί να προσδιοριστεί από μια απλή δειγματοσυνάρτηση της ανέλιξης. Πολύ συχνά γίνεται η παραδοχή ότι οι παραπάνω ιδιότητες ισχύουν, διευκολύνοντας έτσι την προσομοίωση τέτοιων φαινομένων. Στην πραγματικότητα χωρίς την στασιμότητα δεν θα μπορούσαμε να συνδέσουμε τα δεδομένα που μετρήθηκαν στο παρελθόν με το τρέχον, αλλά και η εργοδικότητα μας εξασφαλίζει ότι ο χρονικός μέσος όρος που υπολογίστηκε από το δείγμα τείνει προς τη θεωρητική μέση τιμή της κάθε μεταβλητής (Κουτσογιάννης, 1997).

3.1.1 Ανέλιξη στη συναθροισμένη κλίμακα

Έστω k η τιμή της χρονικής κλίμακας όπου ο k είναι ακέραιος θετικός αριθμός. Η συναθροισμένη ανέλιξη στη κλίμακα k ορίζεται ως εξής :

$$Z_i^{(k)} := \sum_{l=(i-1)k+1}^{ik} X_l \quad (5)$$

Οπότε για $k=1$, $Z_i^{(1)} = X_i$, για $k=2$, $Z_1^{(2)} = X_1 + X_2$, $Z_2^{(2)} = X_3 + X_4$, κ.ο.κ. .

Οι στατιστικές ιδιότητες της συναθροισμένης ανέλιξης προκύπτουν από τις αντίστοιχες της X_i . Συγκεκριμένα:

- Μέση τιμή:

$$E[Z_i^{(k)}] = k\mu \quad (6)$$

- Αυτοσυνδιασπορά:

$$\gamma_j^{(k)} = \text{cov}[Z_i^{(k)}, Z_{i+j}^{(k)}] = \sum_{l=1}^k \sum_{m=jk+1}^{(j+1)k} \gamma_{m-l} \quad (7)$$

Όπου $j=0, \pm 1, \pm 2, \dots$

- Αυτοσυσχέτιση:

$$\rho_j^{(k)} := \text{Corr}[Z_i^{(k)}, Z_{i+j}^{(k)}] = \frac{\gamma_j^{(k)}}{\gamma_0^{(k)}} \quad (8)$$

Όπου $j=0, 1, 2, \dots$

3.2 Απλά στοχαστικά μοντέλα

Στην ενότητα αυτή περιγράφονται δύο βασικά στοχαστικά μοντέλα καθώς και οι μεταξύ τους αποκλίσεις.

3.2.1 Το μοντέλο του Λευκού Θορύβου

Στο μοντέλο του Λευκού Θορύβου διαφορετικές ανελίξεις X_i είναι ανεξάρτητες τυχαίες μεταβλητές έτσι ώστε να ισχύει για $j \neq 0$ $\gamma_j = 0$. Τα βασικά στατιστικά χαρακτηριστικά της συναθροισμένης ανελίξης θα είναι :

- Μέση τιμή:

$$E[Z_i^{(k)}] = kE[X_i] \quad (9)$$

- Αυτοσυνδιασπορά:

$$\gamma_0^{(k)} := \text{Var}[Z_i^{(k)}] = k\gamma_0 \quad (10)$$

- Αυτοσυσχέτιση:

$$\gamma_j^{(k)} = 0, \rho_j^{(k)} = 0 \quad (11)$$

3.2.2 Το μοντέλο AR(1)

Αποτελεί το πιο απλό στοχαστικό μοντέλο με την διατήρηση κάποιας εξάρτησης. Η εκτίμηση της κατάστασης (i) μιας στοχαστικής ανελίξης γίνεται με βάση την αμέσως προηγούμενη κατάσταση ($i-1$). Η διαδικασία αυτή φαίνεται και από την εξίσωση (12), όπου υποθέτοντας μια στοχαστική ανελίξη X_i , το μοντέλο στη βασική κλίμακα εκφράζεται ως εξής :

$$X_i = \rho X_{i-1} + V_i \quad (12)$$

Όπου ρ είναι ο συντελεστής αυτοσυσχέτισης 1^{ης} τάξης και V_i η ακολουθία λευκού θορύβου με μέση τιμή $(1-\rho)\mu$ και διασπορά $(1-\rho^2)\gamma_0$. Με δεδομένο ότι η εξάρτηση της X_i από τη μεταβλητή X_{i-1} επαρκεί για να προσδιοριστεί πλήρως η εξάρτηση του μέλλοντος η διαδικασία χαρακτηρίζεται ως Μαρκοβιανή. Επίσης, η συσχέτισή της θα είναι :

$$\rho_j := \text{corr}[X_i, X_{i+j}] \quad (13)$$

Έχει δειχθεί από τον (Koutsoyiannis, 2002) ότι τα βασικά χαρακτηριστικά της συναθροισμένης ανέλιξης θα είναι :

- Διασπορά :

$$Y_j^{(k)} := \gamma_0 \frac{\kappa(1-\rho^2) - 2\rho(1-\rho^k)}{(1-\rho)^2} \quad j \geq 1 \quad (14)$$

- Αυτοσυνδιασπορά :

$$Y_j^{(k)} := \gamma_0 \frac{\rho^{kj-k+1} (1-\rho^k)^2}{(1-\rho)^2} \quad j \geq 1 \quad (15)$$

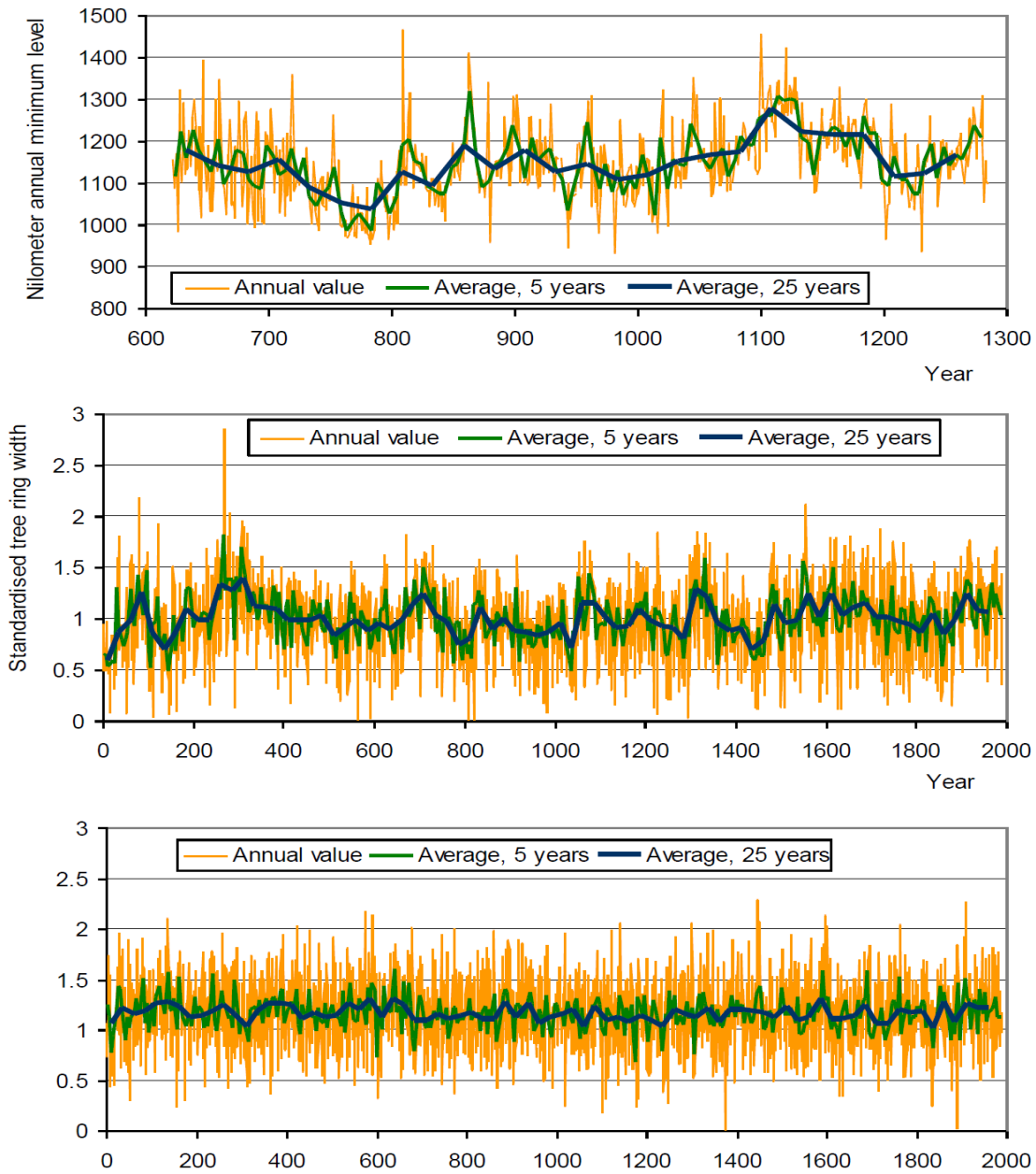
- Αυτοσυσχέτιση :

$$\rho_j^{(k)} = \rho_1^{(k)} \rho^{k(j-1)} \quad (16)$$

όπου $\rho_1^{(k)} = \frac{\rho(1-\rho^k)^2}{\kappa(1-\rho^2) - 2\rho(1-\rho^k)} \quad j \geq 1$

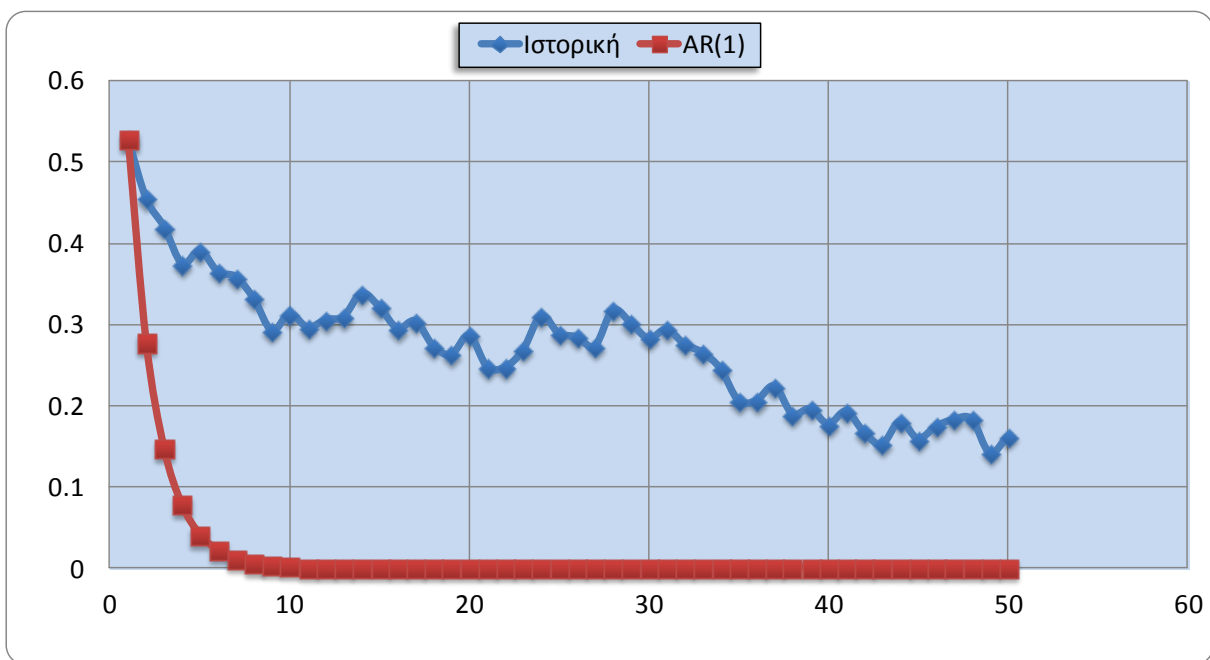
3.3 Υπόθεση της ανεξαρτησίας, η εξαίρεση παρά ο κανόνας

Ήδη από τις αρχές του προηγούμενου αιώνα στατιστικοί και φυσικοί επιστήμονες είχαν παρατηρήσει αποκλίσεις των απλών στοχαστικών μοντέλων του λευκού θορύβου και των Μαρκοβιανών ανελίξεων στη δομή της αυτοσυσχέτισης. Η αποτυχία των απλών μοντέλων αναφοράς να περιγράψουν με επιτυχία την πραγματικότητα αναδείκνυε πως η, κατά τα άλλα πολύ βολική για την κλασική στατιστική, υπόθεση της ανεξαρτησίας των δεδομένων ήταν περισσότερο μια εξαίρεση παρά ο κανόνας. Ο Hurst (1951) μελέτησε την χρονοσειρά της ετήσιας ελάχιστης στάθμης του Νείλου για την οποία υπήρχαν διαθέσιμες καταγραφές από το 622 έως το 1284 μ.Χ. (663 καταγραφές) μετρημένες στο «Νειλόμετρο» στο νησί Ρόντα κοντά στο Κάιρο (Toussoun, 1925; Beran, 1994) και παρατήρησε την τάση «ομαδοποίησης» ομοειδών παρατηρήσεων. Η χρονοσειρά των δεδομένων του ποταμού Νείλου αποτελεί μία από τις πιο μελετημένες χρονοσειρές εξαιτίας της συμβολής της στην ανακάλυψη του φαινομένου της μακροπρόθεσμης εμμονής. Στο σχήμα 3.1 της επόμενης σελίδας φαίνεται η χρονοσειρά της ετήσιας στάθμης του Νείλου και οι συναθροισμένες ανελίξεις για $k=5$ και $k=25$, μία χρονοσειρά τυποποιημένου πάχους δακτυλίων από μια παλαιοκλιματική μελέτη στο Mammoth Creek, Utah (ΗΠΑ) καθώς και μία χρονοσειρά λευκού θορύβου με τα ίδια στατιστικά χαρακτηριστικά για σύγκριση. Παρατηρούμε ότι οι διακυμάνσεις της συναθροισμένης ανέλιξης ειδικά για $k = 25$ είναι πολύ μεγαλύτερες στις ιστορικές χρονοσειρές παρά στη χρονοσειρά λευκού θορύβου με τα ίδια στατιστικά χαρακτηριστικά. Αξίζει να σημειωθεί επίσης ότι υπάρχουν τμήματα («παράθυρα») των παρατηρημένων χρονοσειρών που μπορούν να ερμηνευτούν ως ντετερμινιστική τάση ανόδου ή καθόδου, όμως η συνολική εικόνα οδηγεί στην παρατήρηση ότι πρόκειται μάλλον για τυχαίες διακυμάνσεις σε πολλαπλές κλίμακες.



Σχήμα 3.1 Διάγραμμα της χρονοσειράς της ετήσιας στάθμης του ποταμού Νείλου, της χρονοσειράς των τυποποιημένων δεντροδακτυλίων της περιοχής Mammoth, Utah και μίας χρονοσειρά λευκού θορύβου με τα ίδια στατιστικά χαρακτηριστικά. *Πηγή: Koutsoyiannis, 2002*

Μία άλλη χαρακτηριστική περίπτωση απόκλισης από τα απλά μοντέλα είναι η παρατήρηση της δομής αυτοσυσχέτισης που παρουσιάζει το δείγμα των παλαιοκλιματικών θερμοκρασιακών ανωμαλιών του Βόρειου Ημισφαιρίου κατά την περίοδο 1000-1991 (Jones et al., 1998). Όπως φαίνεται στο Σχήμα 3.2 το μοντέλο AR(1) αποτυγχάνει να περιγράψει ικανοποιητικά την δομή αυτοσυσχέτισης της παρατηρημένης χρονοσειράς, φθίνει εκθετικά και πρακτικά για υστέρηση μεγαλύτερη από 7 μηδενίζεται. Το συγκεκριμένο φαινόμενο έκτοτε έχει παρατηρηθεί σε πολλούς επιστημονικούς κλάδους. Ιδιαίτερο ενδιαφέρον και για το αντικείμενο της παρούσας εργασίας αλλά και από υδρολογικής άποψης παρουσιάζουν οι μελέτες που έχουν γίνει για την ύπαρξη του φαινομένου σε γεωφυσικά φαινόμενα όπως είναι η απορροή των ποταμών (Eltahir 1996; Puckridge, Walker and Costelloe 2000; Koutsoyiannis 2003; Blender and Fraedrich 2006; Koscielny-Bunde et al. 2006; Daniels 2007)



Σχήμα 3.2 Διάγραμμα της αυτοσυσχέτισης σε συνάρτηση με την τιμή υστέρησης για μια ιστορική χρονοσειρά και σε σύγκριση με την αντίστοιχη θεωρητική ενός μοντέλου AR(1)

3.3.1 Το φαινόμενο της μακροπρόθεσμης εμμονής

Το φαινόμενο της μακροπρόθεσμης εμμονής είχε παρατηρηθεί από διάφορους επιστήμονες πολύ πριν γίνει εφικτό να περιγραφεί μαθηματικά. Είχε παρατηρηθεί ότι οι αυτοσυσχετίσεις απομακρυσμένων γεγονότων στον χρόνο έτειναν προς το μηδέν με πολύ αργότερο ρυθμό από αυτόν που θα περίμενε κανείς με βάση την υπόθεση της ανεξαρτησίας ή τις αλυσίδες Markov. Η ανάλυση των ιστορικών στοιχείων της ελάχιστης στάθμης του Νείλου ήταν εκείνη που οδήγησε στην ανακάλυψη του φαινομένου από τον Hurst (1951) αλλά και στην μαθηματική του διατύπωση. Ο Hurst κατασκεύασε ένα διάγραμμα, γνωστό ως μέθοδος R/S (Rescaled Adjusted Range), με σκοπό την κανονικοποίηση της ροής του ποταμού και τα αποτελέσματά του έρχονταν σε αντίθεση με τα αναμενόμενα των έως τότε γνωστών δομών Markov. Είναι αξιοσημείωτο ότι περίπου 10 χρόνια πριν, είχε διαπιστωθεί η ίδια στατιστική συμπεριφορά από τον Ρώσο μαθηματικό Kolmogorov (1941) και ως συνέπεια πέρα από φαινόμενο Hurst, συχνά αναφέρεται και ως δυναμική Hurst-Kolmogorov (HK Dynamics).

Σύμφωνα με τον Koutsoyiannis(2002) η μελέτη της συμπεριφοράς της τυπικής απόκλισης στη συναθροισμένη κλίμακα μπορεί να χρησιμεύσει ως ένας απλούστερος ορισμός της μακροπρόθεσμης εμμονής-σε αντιδιαστολή με τους δύο μαθηματικούς ορισμούς που δίνει ο Beran(1994). Ως γνωστόν, στην κλασική στατιστική ισχύει:

$$\sigma^{(k)} = \frac{\sigma}{\sqrt{k}} \quad (17)$$

όπου $\sigma^{(k)}$ η τυπική απόκλιση της συναθροισμένης ανάλυσης $X_i^{(k)}$ στην κλίμακα συνάθροισης k . Σε χρονοσειρές που παρουσιάζουν μακροπρόθεσμη εμμονή παρατηρείται ένας γενικότερος νόμος :

$$\sigma^{(k)} = \frac{\sigma}{k^{1-H}} \quad (18)$$

όπου H είναι η λεγόμενη σταθερά Hurst που για θετικά συσχετισμένες ανελίξεις λαμβάνει τιμές από 0.5-1. Η τιμή 0.5 υποδηλώνει ανεξαρτησία των δεδομένων ενώ οι τιμές <0.5 υποδηλώνουν αντι-εμμονή.

3.3.2 Μοντέλα διατήρησης της εμμονής και μέθοδοι εκτίμησης του συντελεστή Hurst

Το βασικότερο μοντέλο διατήρησης της εμμονής είναι οι ανελίξεις απλής ομοιοθεσίας, όπως ο κλασματικός γκαουσιανός θόρυβος (FGN- Fractional Gaussian Noise). Αυτού του είδους οι ανελίξεις όπως προαναφέρθηκε είχαν ήδη εισαχθεί θεωρητικά από τον Κοιμογορον (1941), αλλά οι επιστήμονες της εποχής αγνοούσαν την σημασία τους πριν οι Mandelbrot and Wallis (1968) τις εισαγάγουν στα κείμενα τους.

Ο κλασματικός γκαουσιανός θόρυβος περιγράφεται μαθηματικά ως εξής (Koutsoyiannis, 2002):

$$Z_i^{(k)} - k\mu =_d \left(\frac{k}{l}\right)^H (Z_j^{(l)} - l\mu) \quad (19)$$

Στην παραπάνω σχέση το σύμβολο $=_d$ δείχνει την ισότητα στην από κοινού κατανομή των δύο μελών.

Παρακάτω αναφέρονται επιγραμματικά κάποιες από τις βασικότερες μεθόδους για τον υπολογισμό του συντελεστή Hurst. Για περαιτέρω ανάλυση και σύγκριση υπάρχουν αρκετές αναφορές που μπορεί να ανατρέξει ο αναγνώστης (Tyralis and Koutsoyiannis, 2011; Weron 2002; William Rea et al. 2009; W. Rea et al. 2012).

- Η μέθοδος R/S
- Η μέθοδος της συναθροισμένης διασποράς

- Η μέθοδος των ελαχίστων τετραγώνων βασισμένη στην τυπική απόκλιση (LSSD- Least Squares based on Standard Deviation)
- Η μέθοδος των απολύτων τιμών της συναθροισμένης σειράς
- Η μέθοδος του Higuchi
- Το περιοδόγραμμα
- Η μέθοδος του Peng
- Οι εκτιμήτριες του Whittle
- Η Κυματική Ανάλυση (Wavelet Analysis)

3.3.3 Η επιρροή της εμμονής στο υπολογισμό των στατιστικών παραμέτρων

Μία από τις σημαντικότερες επιπτώσεις του φαινομένου της μακροπρόθεσμης εμμονής αφορά την τυπική στατιστική εκτίμηση. Πιο συγκεκριμένα, η δυναμική Hurst-Kolmogorov επιφέρει αρνητική μεροληψία στην εκτίμηση παραμέτρων διασποράς όπως είναι η τυπική απόκλιση. Ωστόσο, είναι δυνατή η εκτίμηση της αβεβαιότητας των παραμέτρων λαμβάνοντας υπόψη την δυναμική Hurst – Kolmogorov (Koutsoyiannis 2003; Koutsoyiannis and Montanari 2007; Koutsoyiannis 2011a).

Η εκτίμηση της μέσης τιμής $\bar{X} := \frac{1}{n} \sum_{i=1}^n X_i$ αποτελεί μια αμερόληπτη εκτιμήτρια ανεξάρτητα από το είδος της ανέλιξης, δηλαδή $E[\bar{X}] = \mu$. Σύμφωνα με όσα γνωρίζουμε από την κλασική στατιστική, η διασπορά της θα είναι:

$$\mathit{Var}[\bar{X}] = \frac{\sigma^2}{n} \quad (20)$$

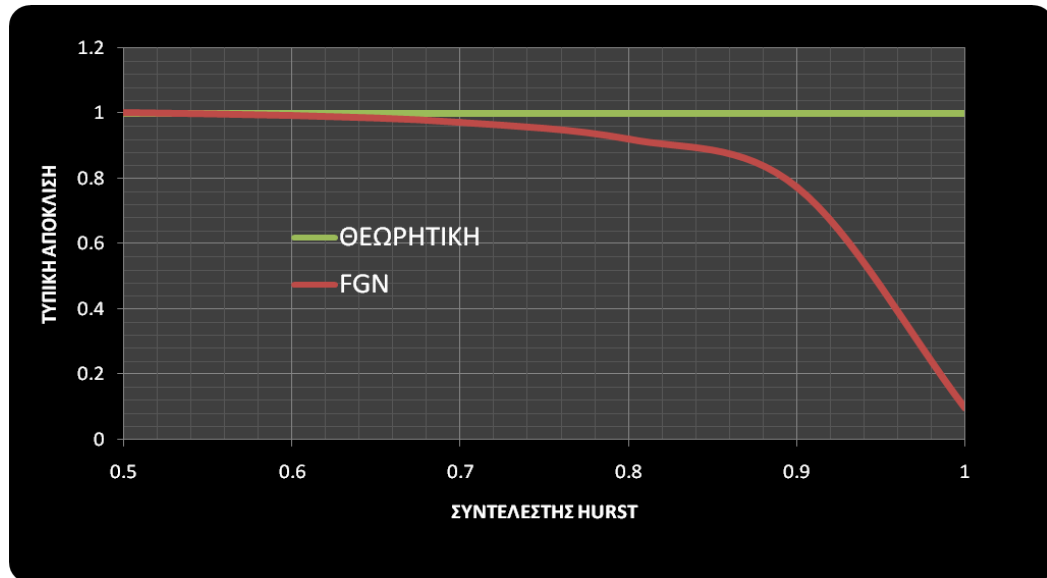
Υπό την προϋπόθεση όμως της ύπαρξης εμμονής σύμφωνα με τους Beran(1994) και Adenstedt(1974) θα ισχύει:

$$Var[\bar{X}] = \frac{\sigma^2}{n^{2-2H}} \quad (21)$$

Είναι προφανές ότι για $H=0.5$ οι (20) και (21) δίνουν το ίδιο τυπικό σφάλμα. Ακόμα, σύμφωνα με τους Koutsoyiannis and Montanari (2007) το ισοδύναμο μήκος μιας χρονοσειράς (n') με εμμονή ώστε να ισούται το τυπικό της σφάλμα με αυτό το (n) που προκύπτει σύμφωνα με την κλασική στατιστική εκτίμηση είναι $n' = n^{2(1-H)}$. Έτσι παρά το γεγονός ότι η μέση τιμή του δείγματος αποτελεί μια αμερόληπτη εκτιμήτρια της μέσης τιμής σε κάθε περίπτωση, η διασπορά της, που εκφράζει την αβεβαιότητα στην εκτίμησή της, είναι πολύ μεγαλύτερη στην περίπτωση ύπαρξης εμμονής. Παρεμφερείς επιπτώσεις παρατηρούμε και στην τυπική εκτιμήτρια της διασποράς, που υπό την υπόθεση της κλασικής στατιστικής δίνεται από την σχέση $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$. Λαμβάνοντας όμως υπόψη το φαινόμενο της μακροπρόθεσμης εμμονής, σύμφωνα με τον Beran(1994), για κάποια γνωστή τιμή του συντελεστή Hurst η τυπική εκτιμήτρια της διασποράς πλέον θα είναι:

$$S^2 = \frac{1}{n-n^{2H-1}} \sum_{i=1}^n (X_i - \bar{X})^2 \quad (22)$$

Στο Σχήμα 3.3 φαίνεται το αποτέλεσμα μιας προσομοίωσης κατά την οποία παρήχθησαν 10,000 συνθετικές χρονοσειρές μήκους $n = 100$ ετών και διατηρούσαν τον συντελεστή Hurst για μία Γκαουσιανή κατανομή. Παρατηρείται ότι όσο αυξάνεται η τιμή του H , αυξάνεται και η αρνητική μεροληψία στην εκτίμηση της τυπικής απόκλισης. Ειδικότερα για τιμές $H > 0.8$ η κλασική εκτιμήτρια υποεκτιμά σημαντικά την τυπική απόκλιση



Σχήμα 3.3 Διάγραμμα της τυπική απόκλισης σε συνάρτηση με την τιμή του συντελεστή Hurst για ένα FGN μοντέλο και σε σύγκριση με την θεωρητική της τιμή.

3.3.4 Το Mann-Kendall trend test υπό το πρίσμα της δυναμικής Hurst-Kolmogorov

Υπάρχουν πολλοί εναλλακτικοί τρόποι ώστε να ελέγξει ένας μελετητής την ύπαρξη τάσεων ή άλλου τύπου αλλαγές στα υδρολογικά δεδομένα. Στην παρούσα εργασία δίνεται έμφαση σε μία συγκεκριμένη κατηγορία αυτών των τεστ στατιστικού ελέγχου, τα οποία επειδή οι στατιστικές τους παράμετροι εξαρτώνται από την κλάση των δεδομένων και όχι από τα ίδια τα δεδομένα, θεωρείται ότι το αποτέλεσμά τους είναι ανεξάρτητο της κατανομής που ακολουθούν στην πραγματικότητα τα δεδομένα προς ανάλυση (distribution-free methods).

Μία παρατήρηση έχει κλάση N εάν είναι η N -οστή μεγαλύτερη τιμή στο σετ των παρατηρημένων δεδομένων. Τα περισσότερα τεστ που εμπίπτουν σε αυτή την κατηγορία υποθέτουν ότι τα δεδομένα είναι ανεξάρτητα και με την ίδια κατανομή πιθανότητας. Μία υπόθεση που όπως αναλύθηκε στην ενότητα 3.3.3 μπορεί να οδηγήσει σε λανθασμένες στατιστικές εκτιμήσεις ένα δεν ληφθεί υπόψη η δομή συσχέτισης των δεδομένων.

Το Mann-Kendall trend test (Mann, 1945; Kendall, 1975) βασίζεται στις κλάσεις των παρατηρήσεων μιας χρονοσειράς και στην κατανομή τους μέσα στο χρόνο. Για μία χρονοσειρά $X = \{x_1, x_2, \dots, x_n\}$ η στατιστική παράμετρος ορίζεται ως εξής:

$$S = \sum_{i < j} a_{ij} \quad (23)$$

όπου

$$a_{ij} = \text{sign}(x_j - x_i) = \text{sign}(R_j - R_i) = \begin{cases} 1 & x_i < x_j \\ 0 & x_i = x_j \\ -1 & x_i > x_j \end{cases} \quad (24)$$

και R_i, R_j είναι οι κλάσεις των παρατηρήσεων x_i και x_j αντίστοιχα. Από την εξίσωση (24) γίνεται αντιληπτό ότι η στατιστική παράμετρος της εξίσωσης (23) εξαρτάται από τις κλάσεις των παρατηρήσεων παρά από τις τιμές των παρατηρήσεων και κατ' επέκταση θεωρείται ότι η στατιστική σημαντικότητα του τεστ μένει ανεπηρέαστη από την πραγματική κατανομή των δεδομένων. Η παρατήρηση αυτή έρχεται σε αντίθεση με τα παραμετρικά τεστ τάσεων, τα οποία υποθέτουν ότι τα δεδομένα ακολουθούν την κανονική κατανομή και η ισχύς των οποίων μπορεί εύκολα να μειωθεί στην περίπτωση ασύμμετρων δεδομένων (Yue et al., 2002b).

Υπό το πρίσμα της κλασικής στατιστικής (ανεξαρτησία δεδομένων) σύμφωνα με τον Kendall (1975) η μέση τιμή και η διασπορά της παραμέτρου S είναι :

$$E[S] = 0 \quad (25)$$

$$V_0[S] = n(n-1)(2n+5)/18 \quad (26)$$

όπου n είναι το πλήθος των παρατηρήσεων. Επίσης ο Kendall (1975) αποδεικνύει ότι η κατανομή της S τείνει προς την κανονικότητα καθώς το n παίρνει μεγαλύτερες τιμές. Η στατιστική σημαντικότητα της τάσης μπορεί να ελεγχθεί, συγκρίνοντας την πιθανότητα

εμφάνιση της κανονικοποιημένης μεταβλητής u της εξίσωσης (27) με το επιλεγμένο από τον μελετητή επίπεδο σημαντικότητας (significance level):

$$u = \begin{cases} (S - 1)/\sqrt{V_0(S)} & S > 0 \\ 0 & S = 0 \\ (S + 1)/\sqrt{V_0(S)} & S < 0 \end{cases} \quad (27)$$

Λαμβάνοντας υπόψη το φαινόμενο της μακροπρόθεσμης εμμονής, αδιαμφισβήτητα απαιτείται τροποποίηση της εκτιμήτριας της διασποράς της στατιστικής παραμέτρου S . Για τιμές του H ίσες ή μικρότερες του 0.5 ο μελετητής μπορεί να κάνει χρήση της κλασικής στατιστικής χωρίς να παραβιάζεται η απαραίτητη υπόθεση της ανεξαρτησίας του Mann-Kendall trend test.

Στη θεωρία ελέγχου υποθέσεων, ο έλεγχος μιας υπόθεσης είναι μια δισήμαντη στατιστική διαδικασία που οδηγεί είτε στην απόρριψη είτε στη αποδοχή (ακριβέστερα: μη απόρριψη) της υπόθεσης. Υπάρχει πληθώρα αναφορών που γίνεται λεπτομερέστερη παρουσίαση της παραπάνω θεωρίας (Paroulis, 1990 ; Freund et al., 1988; Hirsch et al., 1993). Το σημείο εκκίνησης για ένα τεστ ελέγχου υποθέσεων είναι η διατύπωση της «μηδενικής» και της «εναλλακτικής» υπόθεσης. Στην περίπτωση της ανάλυσης τάσεων, ως μηδενική λαμβάνεται η υπόθεση «μη ύπαρξης τάσης» (Kundzewicz and Robson, 2004) και ελέγχεται εάν τα δεδομένα μας υπακούουν σε αυτήν την υπόθεση. Το ποσοστό απόρριψης της μηδενικής υπόθεσης όταν στην πραγματικότητα δεν υπάρχει κάποια τάση είναι γνωστό ως *Σφάλμα Τύπου I* και θα έπρεπε να ισούται με το επιλεγμένο επίπεδο σημαντικότητας. Όπως έχει δειχθεί όμως (Koutsoyiannis 2006; Koutsoyiannis 2003; Hamed and Rao 1998) υπό το πρίσμα της μακροπρόθεσμης εμμονής, ή γενικότερα στις περιπτώσεις που τα δεδομένα παρουσιάζουν ισχυρή δομή συσχέτισης για γεγονότα απομακρυσμένα στο χρόνο, η κλασική στατιστική υποεκτιμά σημαντικά τη διασπορά της εξίσωσης (26), με αποτέλεσμα την αύξηση του *Σφάλματος Τύπου I*. Αυτό έχει σαν αποτέλεσμα την απόρριψη σε πολύ

μεγαλύτερα ποσοστά της μηδενικής υπόθεσης και κατ' επέκταση σε εσφαλμένα μεγαλύτερο αριθμό στατιστικά σημαντικών τάσεων.

Για την αντιμετώπιση, των παραπάνω προβληματικών επιπτώσεων ο Hamed (2008) προτείνει μία συγκεκριμένη αναλυτική λύση για την τροποποίηση του τεστ, ώστε να λαμβάνεται υπόψη η ύπαρξη εμμονής των δεδομένων στον υπολογισμό της διασποράς που θα χρησιμοποιηθεί εν τέλει στην εξίσωση (27). Στην παρούσα εργασία, αποφεύγεται η μαθηματικά περίπλοκη λύση του Hamed και υιοθετείται η πολύ πιο άμεση προσέγγιση του Koutsoyiannis (2003; 2006) μέσω FGN προσομοιώσεων για τον υπολογισμό της διασποράς της παραμέτρου S . Τα αποτελέσματα αυτής της διαδικασίας παρουσιάζονται αναλυτικά στο Κεφάλαιο 5 καθώς και η επιβεβαίωση των παραπάνω αναφερθέντων προβλημάτων μέσω της εφαρμογής σε πραγματικά δεδομένα.

4 ΠΡΟΣΑΡΜΟΓΗ ΣΥΝΑΡΤΗΣΕΩΝ ΚΑΤΑΝΟΜΗΣ ΣΕ ΔΕΔΟΜΕΝΑ ΑΚΡΑΙΩΝ ΤΙΜΩΝ

Στο παρόν κεφάλαιο γίνεται μια συνοπτική παρουσίαση της θεωρίας ακραίων τιμών , γίνεται αναφορά στις κατανομές εκείνες που θεωρούνται οι πλέον κατάλληλες για προσαρμογή σε δεδομένα ακραίων τιμών καθώς και σε ορισμένα τεστ που χρησιμοποιούνται ώστε να ελεγχθεί η καταλληλότητα της προσαρμογής τους. Δίνονται οι μαθηματικοί τους ορισμοί, καθώς και βιβλιογραφικές αναφορές για την προσαρμογή συναρτήσεων κατανομής πιθανότητας.

4.1 Εισαγωγή στη θεωρία ακραίων τιμών

Σύμφωνα με τη Θεωρία Ακραίων Τιμών, για δεδομένο αριθμό n ανεξάρτητων και με την ίδια κατανομή τυχαίων μεταβλητών, η μεγαλύτερη από αυτές $X = \max\{Y_1, Y_2, \dots, Y_n\}$ ακολουθεί συνάρτηση πιθανότητας:

$$H_n(x) = (F(x))^n \quad (28)$$

αφού:

$$H_n(x) = P(X \leq x) = P[\max(Y_1, Y_2, \dots, Y_n) \leq x] = P(Y_1 \leq x, Y_2 \leq x, \dots, Y_n \leq x) = [F(x)]^n$$

όπου $F(x) = P(Y_i \leq x)$ είναι η συνάρτηση πυκνότητας πιθανότητας των Y_i και αναφέρεται συνήθως ως μητρική κατανομή (parent distribution). Εάν το n δεν είναι σταθερό, αλλά μπορεί να θεωρηθεί ως πραγματοποίηση μιας τυχαίας μεταβλητής με Poisson κατανομή και μέση τιμή ν , τότε η κατανομή της μεταβλητής X γίνεται (Todorovic and Zelenhasic 1970; Rossi et al., 1984):

$$H'_v(x) = \exp\{-v[1 - F(x)]\} \quad (29)$$

Σύμφωνα όμως με τον Koutsoyiannis (2004):

$$\begin{aligned} \ln(F(x)[F(x)]^n) &= n \ln\{1 - [1 - F(x)]\} = n\{-[1 - F(x)] - [1 - F(x)]^2 - \dots\} \\ &\approx -n[1 - F(x)] \end{aligned}$$

Άρα για μεγάλα n ή $F(x)$, έχουμε $H_n(x) \approx H'_n(x)$ και μάλιστα με αριθμητική αναζήτηση προκύπτει (Koutsoyiannis, 2004) ότι ακόμα και για μικρά n η διαφορά τους δεν είναι στατιστικά σημαντική.

Στις υδρολογικές, λοιπόν, εφαρμογές που αφορούν την κατανομή των ετήσιων μεγίστων τιμών πλημμύρας μπορεί να υποτεθεί ότι ο αριθμός των τιμών Y_i δεν είναι σταθερός και άρα το μοντέλο Poisson να θεωρηθεί κατάλληλο. Δεδομένης και της μικρής διαφοράς των σχέσεων (28) και (29) μπορούμε να συμπεράνουμε ότι είναι κατάλληλο μοντέλο για κάθε υδρολογική εφαρμογή. Στην στατιστική υδρολογία, όμως, αντί αυτών των σχέσεων έχουν χρησιμοποιηθεί ασυμπτωτικές ή οριακές τιμές ακραίων τιμών, οι οποίες λαμβάνονται από τις παραπάνω σχέσεις όταν το n τείνει στο άπειρο.

Ο Gumbel (1958) βασιζόμενος στο πρωτοποριακό έργο των Fréchet (1927), Fisher and Tippett (1928) και Gnedenko (1941) διαμόρφωσε την θεωρία των ακραίων τιμών. Σύμφωνα με αυτήν, καθώς το n τείνει προς το άπειρο, η $H_n(x)$ συγκλίνει σε μία εκ των τριών προαναφερθέντων ασυμπτωτικών κατανομών. Το ίδιο μπορεί να ισχύσει και για την $H'_v(x)$ καθώς το v τείνει αντιστοίχως προς το άπειρο. Σύμφωνα με τον Jenkinson (1955,1969) και οι τρεις μπορούν να περιγραφούν με μία κοινή μαθηματική έκφραση η οποία έγινε γνωστή ως Γενικευμένη Ακραίων Τιμών- ΓΑΤ):

$$H(x) = \exp\left\{-\left[1 + \kappa\left(\frac{x}{\lambda} - \psi\right)\right]^{-1/\kappa}\right\} \quad \kappa x \geq \kappa\lambda\left(\psi - \frac{1}{\kappa}\right) \quad (30)$$

Όπου ψ , $\lambda > 0$ και κ είναι παράμετροι θέσης, κλίμακας και σχήματος, αντίστοιχα. Αξίζει να σημειωθεί ότι ο Leadbetter (1974) έδειξε ότι η παραπάνω σχέση δεν ισχύει μόνο για ανεξάρτητες τυχαίες μεταβλητές αλλά και για εξαρτημένες τυχαίες μεταβλητές με δεδομένο ότι δεν υπάρχει μεγάλη εξάρτηση για μεγάλα ποσοστά υπέρβασης.

4.2 Συναρτήσεις Κατανομών Πιθανότητας Ακραίων Τιμών και Μέθοδοι εκτίμησης των παραμέτρων τους

4.2.1 Η Γενικευμένη Ακραίων Τιμών- ΓΑΤ

Όπως άλλωστε δηλώνει και οι ονομασία της, η ΓΑΤ κατανομή είναι ευρέως διαδεδομένη για την χρήση της στην μελέτη ακραίων τιμών (ελαχίστων ή μεγίστων). Για τις διάφορες τιμές του κ από την εξίσωση (30) προκύπτουν οι τρεις ασυμπτωτικές κατανομές:

- Για $\kappa=0$ προκύπτει η κατανομή μεγίστων τύπου I (Κατανομή Gumbel ή EV1). Σε αυτή την περίπτωση με απλούς μαθηματικούς υπολογισμούς παίρνει τη μορφή:

$$H(x) = \exp\left\{-\exp\left[-\frac{x-\varepsilon}{\lambda}\right]\right\} \quad -\infty \leq x \leq +\infty \quad (31)$$

- Για $\kappa > 0$ προκύπτει η κατανομή μεγίστων τύπου II (Κατανομή Frechet ή EV2) έχοντας την εξής μορφή:

$$H(x) = \exp\left\{-\left(\frac{\varepsilon}{x}\right)^{1/\kappa}\right\} \quad x \geq 0 \quad (32)$$

- Για $\kappa < 0$ προκύπτει η κατανομή μεγίστων τύπου III (EV3). Η περίπτωση αυτή βέβαια δεν έχει κάποιο πρακτικό ενδιαφέρον από την σκοπιά της υδρολογίας μιας και αναφέρεται σε μεταβλητές φραγμένες στο 0 από τα δεξιά.

4.2.2 Η Λογαριθμοκανονική Κατανομή

Η Λογαριθμοκανονική Κατανομή δύο παραμέτρων προκύπτει από την Κανονική Κατανομή με τον μετασχηματισμό:

$$y = \ln x \leftrightarrow x = e^y \quad (33)$$

Έτσι η μεταβλητή X ακολουθεί λογαριθμοκανονική κατανομή δύο παραμέτρων αν η μεταβλητή Y ακολουθεί κανονική κατανομή $N(\mu_Y, \sigma_Y)$. Παρακάτω δίνεται η συνάρτηση πυκνότητας πιθανότητας της λογαριθμοκανονικής κατανομής δύο παραμέτρων :

$$f_X(x) = \frac{1}{x\sqrt{2\pi}\sigma_Y} e^{-\frac{1}{2}\left(\frac{\ln x - \mu_Y}{\sigma_Y}\right)^2} \quad x > 0 \quad (34)$$

Όπου μ_Y και $\sigma_Y > 0$ είναι οι παράμετροι κλίμακας και σχήματος αντίστοιχα.

4.2.3 Η Κατανομή Γάμμα δύο παραμέτρων

Η κατανομή γάμμα δύο παραμέτρων είναι γενίκευση την εκθετικής κατανομής, με παράμετρο κλίμακας λ και παράμετρο σχήματος κ . Είναι θετικά ασύμμετρη, ορίζεται μόνο για θετικές τιμές της μεταβλητής και είναι μία από τις πιο διαδεδομένες κατανομές της τεχνικής υδρολογίας. Η συνάρτηση πυκνότητας πιθανότητας θα είναι:

$$f(x) = \frac{\lambda^\kappa}{\Gamma(\kappa)} x^{\kappa-1} e^{-\lambda x} \quad (35)$$

4.2.4 Η κατανομή Γάμμα τριών παραμέτρων (Pearson III)

Με την προσθήκη μίας παραμέτρου θέσης στην κατανομή γάμμα δύο παραμέτρων, παίρνουμε την κατανομή γάμμα τριών παραμέτρων που είναι γνωστή ως κατανομή Pearson τύπου III. Η παράμετρος θέσης c , η οποία αποτελεί το κάτω όριο της μεταβλητής, επιτρέπει την καλύτερη προσαρμογή της κατανομής στα δεδομένα. Οι ιδιότητες της κατανομής είναι

παρόμοιες με αυτές της κατανομής γάμμα δύο παραμέτρων και σύμφωνα με τη μέθοδο των ροπών η παράμετρος c θα είναι (Κουτσογιάννης, 1997):

$$c = \bar{x} - \frac{\kappa}{\lambda} \quad (36)$$

Αντίστοιχα με την κατανομή γάμμα δυο παραμέτρων η συνάρτηση πυκνότητας πιθανότητας της κατανομής Pearson III θα είναι:

$$f(x) = \frac{\lambda^\kappa}{\Gamma(\kappa)} (x - c)^{\kappa-1} e^{-\lambda(x-c)} \quad x > c \quad (37)$$

Όπου c είναι η παράμετρος θέσης, $\kappa > 0$ η παράμετρος σχήματος και $\lambda > 0$ η παράμετρος κλίμακας.

4.2.5 Η Κατανομή Log-Pearson III

Η κατανομή Log-Pearson III προκύπτει από την κατανομή Pearson III και τον μετασχηματισμό:

$$y = \ln x \leftrightarrow x = e^y \quad (38)$$

Άρα λέμε ότι μια μεταβλητή X ακολουθεί κατανομή Log-Pearson III αν η μεταβλητή Y ακολουθεί κατανομή Pearson III. Η συνάρτηση πυκνότητας πιθανότητας για την κατανομή Log-Pearson III θα είναι :

$$f(x) = \frac{\lambda^\kappa}{x\Gamma(\kappa)} (\ln x - c)^{\kappa-1} e^{-\lambda(\ln x - c)}, \quad x > e^c \quad (39)$$

Ένα από τα βασικότερα χαρακτηριστικά της κατανομής αυτής είναι ότι μπορεί να έχει πολύ μεγάλο συντελεστή ασυμμετρίας. Σε αυτή της την ιδιότητα οφείλεται και η ευρεία διάδοσή της στην τεχνική υδρολογία μαζί με τις κατανομές ακρότατων που αναφέρθηκαν.

4.2.6 Μέθοδοι εκτίμησης των παραμέτρων των κατανομών

Έστω τυχαία μεταβλητή X με γνωστό τύπο συνάρτησης κατανομής, ο οποίος περιέχει τις άγνωστες παραμέτρους $\theta_1, \theta_2, \dots, \theta_r$. Ως συνέπεια, η πυκνότητα πιθανότητας της X είναι μια συνάρτηση $f_X(x; \theta_1, \theta_2, \dots, \theta_r)$. Η εκτίμηση αυτών των παραμέτρων με βάση το δείγμα μπορεί να γίνει με διάφορους τρόπους. Παρακάτω παρατίθενται οι δύο πιο γνωστές μεθοδολογίες εκτίμησης των παραμέτρων μιας κατανομής. Η στατιστική έχει αναπτύξει και άλλες γενικές μεθόδους υπολογισμού των παραμέτρων κατανομών, όπως τη μέθοδο της μέγιστης εντροπίας ή τη μέθοδο των L-ροπών. (Singh and Rajagopal (1986) και Stedinger et al. (1993))

- Μέθοδος των ροπών

Η μέθοδος των ροπών βασίζεται στην εξίσωση των θεωρητικών ροπών της κατανομής της X με τις αντίστοιχες δειγματικές εκτιμήσεις των ροπών. Άρα αν r είναι ο αριθμός των άγνωστων παραμέτρων της κατανομής, γράφουμε r εξισώσεις της μορφής :

$$m_X^{(k)} = \hat{m}_X^{(k)} \quad k = 1, 2, \dots, r \quad (40)$$

Όπου τα $m_X^{(k)}$ είναι οι θεωρητικές ροπές περί την αρχή οι οποίες είναι συναρτήσεις των άγνωστων παραμέτρων και δίνονται από τη σχέση :

$$m_X^{(k)} = \int_{-\infty}^{+\infty} x^k f_X(x, \theta_1, \theta_2, \dots, \theta_r) dx \quad (41)$$

Ενώ τα $\hat{m}_X^{(k)}$ είναι οι αριθμητικές εκτιμήσεις που υπολογίζονται από το δείγμα σύμφωνα με την :

$$\hat{m}_X^{(k)} = \frac{1}{n} \sum_{i=1}^n x_i^k \quad (42)$$

Καταστρώνοντας και στη συνέχεια επιλύοντας τις r εξισώσεις, υπολογίζονται οι άγνωστες παράμετροι $\vartheta_1, \vartheta_2, \dots, \vartheta_r$.

- Μέθοδος της Μέγιστης Πιθανοφάνειας

Έστω η τυχαία μεταβλητή X με πυκνότητα πιθανότητας $f_X(x; \vartheta_1, \vartheta_2, \dots, \vartheta_r)$ όπου $\theta_1, \theta_2, \dots, \theta_r$ παράμετροι και δείγμα X_1, X_2, \dots, X_n της μεταβλητής. Λόγω της ανεξαρτησίας των μεταβλητών X_1, X_2, \dots, X_n η από κοινού συνάρτηση πυκνότητας πιθανότητας τους είναι :

$$f_{X_1, \dots, X_n}(x_1, \dots, x_n, \theta_1, \dots, \theta_r) = \prod_{i=1}^n f_X(x_i, \theta_1, \dots, \theta_r) \quad (43)$$

Για δεδομένες τιμές των τυχαίων μεταβλητών ίσες με τις παρατηρήσεις x_1, \dots, x_n , η $f_{X_1, \dots, X_n}()$ είναι συνάρτηση των παραμέτρων $\vartheta_1, \vartheta_2, \dots, \vartheta_r$ και λέγεται *συνάρτηση πιθανοφάνειας αυτών των παραμέτρων*.

Η μέθοδος της μέγιστης πιθανοφάνειας εκτιμά τις παραμέτρους $\theta_1, \theta_2, \dots, \theta_r$ με τέτοιο τρόπο ώστε η συνάρτηση πιθανοφάνειας να γίνει μέγιστη. Στην περίπτωση αυτή θα ισχύει :

$$\frac{\partial f_{X_1, \dots, X_n}(x_1, \dots, x_n, \theta_1, \dots, \theta_r)}{\partial \theta_k} = 0 \quad k = 1, \dots, r \quad (44)$$

Από τις r εξισώσεις τις παραπάνω μορφής προκύπτουν οι r άγνωστες παράμετροι. Επειδή όμως ο χειρισμός των εξισώσεων είναι πολύπλοκος, αντί της μεγιστοποίησης της συνάρτησης πιθανοφάνειας, επιχειρείται η μεγιστοποίηση του λογαρίθμου της:

$$L(x_1, \dots, x_n, \theta_1, \dots, \theta_r) = \ln f_{X_1, \dots, X_n}(x_1, \dots, x_n, \theta_1, \dots, \theta_r) = \prod_{i=1}^n \ln f_X(x_i, \theta_1, \dots, \theta_r) \quad (45)$$

Η συνάρτηση $L()$ λέγεται λογαριθμική συνάρτηση πιθανοφάνειας. Για να είναι μέγιστη προφανώς θα πρέπει να ισχύει:

$$\frac{\partial L(x_1, \dots, x_n, \theta_1, \dots, \theta_r)}{\partial \theta_k} = 0 \quad k = 1, \dots, r \quad (46)$$

Από τη λύση αυτών των r εξισώσεων προκύπτουν οι r άγνωστες παράμετροι.

4.3 Κριτήρια επιλογής της καταλληλότερης κατανομής

Για τον έλεγχο της καταλληλότητας της προσαρμογής της εκάστοτε κατανομής υπάρχουν δύο προσεγγίσεις: είτε τα τεστ στατιστικού ελέγχου (goodness-of-fit tests) είτε κριτήρια επιλογής της κατάλληλης κατανομής έναντι άλλων υποψήφιας κατανομών (Strupczewski et al., 2001a). Παρά την ύπαρξη πληθώρας άρθρων και δημοσιεύσεων σε επιστημονικά περιοδικά σχετικά με το ποιες είναι οι κατάλληλες κατανομές για την μοντελοποίηση παρατηρήσεων ακραίων τιμών, δυσανάλογη είναι η συζήτηση σχετικά με τα κριτήρια τα ίδια που μπορεί ένας μελετητής να χρησιμοποιήσει για να ελέγξει την προσαρμογή μιας κατανομής σε παρατηρημένα δεδομένα. Η πιο συνηθισμένη διαδικασία που ακολουθείται από τους στατιστικούς είναι να υπόκεινται οι επιλεγμένες συναρτήσεις πιθανότητας σε τεστ στατιστικού ελέγχου (π.χ. Τεστ χ^2 , Kolmogorov- Smirnov, Cramer von Mises, κ.α.): Οι κατανομές εκείνες που γίνονται αποδεκτές σε ένα καθορισμένο επίπεδο σημαντικότητας, επιλέγονται ως οι κατάλληλες (Stedinger et al. , 1992). Η συγκεκριμένη διαδικασία παρουσιάζει ορισμένους περιορισμούς, μιας και τα αποτελέσματα εξαρτώνται σε πολύ μεγάλο βαθμό από την (υποκειμενική) επιλογή του επιπέδου σημαντικότητας κατά τον έλεγχο αλλά υστερούν ενδεχομένως και σε σαφήνεια αφού περισσότερες από μία κατανομές μπορούν να γίνουν αποδεκτές για το ίδιο επίπεδο σημαντικότητας (Burnham and Anderson, 2002). Η ασάφεια αυτή μπορεί να ξεπεραστεί στα τεστ στατιστικού ελέγχου, αν ο μελετητής επιλέξει την κατανομή εκείνη που «δεν απορρίπτει» την μηδενική υπόθεση με το μεγαλύτερο/(ισχυρότερο) επίπεδο σημαντικότητας (Stephens, 1986).

Τα κριτήρια επιλογής που χρησιμοποιούνται στην παρούσα εργασία μπορούν να θεωρηθούν μια εναλλακτική τακτική απέναντι στα τεστ στατιστικού ελέγχου τα οποία

παρουσιάζουν τις προβληματικές που προαναφέρθηκαν. Τα βασικότερα πλεονεκτήματά τους είναι η αντικειμενικότητα (δεν υπάρχει ανάγκη για τον ορισμό ενός επιπέδου εμπιστοσύνης) αλλά και η σαφήνεια (μία και μόνο μία κατανομή επιλέγεται ως η πλέον κατάλληλα προσαρμοσμένη) (Laiο et al., 2009).

Έστω ένα δείγμα n δεδομένων, $D = \{x_1, x_2, \dots, x_n\}$, ταξινομημένων από το μικρότερο στο μεγαλύτερο τα οποία παρήχθησαν από μία άγνωστη μητρική κατανομή $f(x)$. Για την αναπαράσταση των δεδομένων χρησιμοποιούνται N_m μοντέλα M_j , $j=1,2,\dots,N_m$. Οι συναρτήσεις πιθανότητας M_j έχουν παραμέτρους που έχουν υπολογιστεί από το δείγμα D και αναζητείται το μοντέλο M_{opt} εκείνο που θα αναπαριστά καλύτερα τα δεδομένα. Ουσιαστικά, οι τεχνικές που χρησιμοποιούνται ποσοτικοποιούν την διαφορά μεταξύ την μητρικής συνάρτησης και του μοντέλου προς εξέταση, με την επιπρόσθετη όμως δυσκολία ότι σχεδόν ποτέ η μητρική συνάρτηση δεν είναι γνωστή και έτσι η διαφορά αυτή δεν υπολογίζεται αλλά εκτιμάται.

Τρία είναι τα κριτήρια που χρησιμοποιούνται στην παρούσα εργασία: το κριτήριο πληροφορίας Akaike (*Akaike Information Criterion – AIC*), το Μπεϋζιανό κριτήριο πληροφορίας (*Bayesian Information Criterion - BIC*) και μία παραλλαγή του *Anderson-Darling* κριτηρίου (ADC).

➤ *Το Κριτήριο Πληροφορίας Akaike*

Το κριτήριο πληροφορίας Akaike (Akaike, 1973) είναι βασισμένο στη χρήση της Kullback-Leibler πληροφορίας που μετράει τη διαφορά μεταξύ της μητρικής συνάρτησης $f(x)$ και του επιλεγμένου μοντέλου $M_j = g_j(x, \theta)$. Για το τυχαίο j μοντέλο υπολογίζουμε :

$$AIC_j = -2 \ln(L_j(\hat{\theta})) + 2p_j \quad (47)$$

Όπου $L_j(\hat{\theta}) = \prod_{i=1}^n g_j(x, \hat{\theta})$ είναι η συνάρτηση πιθανοφάνειας, οι παράμετροι είναι υπολογισμένες με τη μέθοδο της μέγιστης πιθανοφάνειας και p_j είναι ο αριθμός των εκτιμημένων παραμέτρων του j μοντέλου (Linhardt and Zucchini, 1986). Μετά τον υπολογισμό των AIC_j όλων των υπό εξέταση j μοντέλων επιλέγεται ως καταλληλότερο εκείνο με την μικρότερη υπολογισμένη πληροφορία AIC_{min} .

➤ Το Μπεϋζιανό Κριτήριο Πληροφορίας

Το Μπεϋζιανό Κριτήριο Πληροφορίας εισήχθη πρώτη φορά από τον Schwarz (1978). Ακολουθώντας την ίδια λογική με το AIC, για το τυχαίο j μοντέλο υπολογίζουμε :

$$BIC_j = -2 \ln(L_j(\hat{\theta})) + \ln(n)p_j \quad (48)$$

Όπως και προηγουμένως μετά τον υπολογισμό των BIC_j , επιλέγουμε το μοντέλο εκείνο με την μικρότερη υπολογισμένη πληροφορία BIC_{min} . Παρατηρούμε ότι παρά το τελείως διαφορετικό θεωρητικό πλαίσιο από το οποίο προέρχονται τα δύο κριτήρια καταλήγουν σε παρόμοιες εξισώσεις.

➤ Το κριτήριο Anderson-Darling

Αν και τα δύο προαναφερθέντα κριτήρια επιλογής μοντέλου χρησιμοποιούνται ως τεχνικές σε πολλούς κλάδους, δεν δημιουργήθηκαν για να εφαρμόζονται σε μικρά μήκη δείγματος και σε θετικά ασύμμετρα δεδομένα, παραδοχές που συναντώνται αρκετά συχνά στην τεχνική υδρολογία. Για αυτό το λόγο, επιλέγεται και το Anderson-Darling κριτήριο το οποίο έχει επιδείξει πολύ ικανοποιητική συμπεριφορά όταν εφαρμόζεται σε υδρολογικά δεδομένα (Onoz and Bayazit, 1995; Lai, 2004; Viglione et al., 2007). Το συγκεκριμένο κριτήριο έχει την μορφή :

$$ADC_j = 0.0403 + 0.116 \left(\frac{\Delta_{AD,j} - \xi_j}{\beta_j} \right)^{\frac{n_j}{0.861}} \text{ αν } 1.2\xi_j \leq \Delta_{AD,j} \quad (49)$$

και

$$ADC_j = \left[0.0403 + 0.116 \left(\frac{0.2\xi_j}{\beta_j} \right)^{\frac{n_j}{0.861}} \right] \frac{\Delta_{AD,j} - 0.2\xi_j}{\xi_j} \text{ αν } 1.2\xi_j > \Delta_{AD,j} \quad (50)$$

όπου $\Delta_{AD,j} = \Delta_{AD} = [g_j(x, \tilde{\theta}), f_n(x)]$ είναι η διαφορά μεταξύ μητρικής συνάρτησης και του εξεταζόμενου μοντέλου, οι β_j, ξ_j, n_j είναι παράμετροι που εξαρτώνται από την επιλογή του μοντέλου και έχουν πινακοποιηθεί από τον Laiο (2004) για επτά συγκεκριμένες κατανομές που χρησιμοποιούνται συνήθως για προσαρμογή σε δεδομένα ακραίων τιμών. Όπως και προηγουμένως, επιλέγεται το μοντέλο εκείνο που παρουσιάζει την μικρότερη πληροφορία ADC_{min} . Ουσιαστικά, η εφαρμογή του κριτηρίου είναι ισοδύναμη με το να υπολογίσουμε τις τιμές p του αντίστοιχου τεστ στατιστικού ελέγχου αλλά να διαλέγουμε μόνο την κατανομή εκείνη που «περνά» το επίπεδο σημαντικότητας με την μεγαλύτερη τιμή του p .

5 ΕΦΑΡΜΟΓΗ ΣΕ ΠΡΑΓΜΑΤΙΚΑ ΔΕΔΟΜΕΝΑ

5.1 Περιγραφή Βάσης Δεδομένων

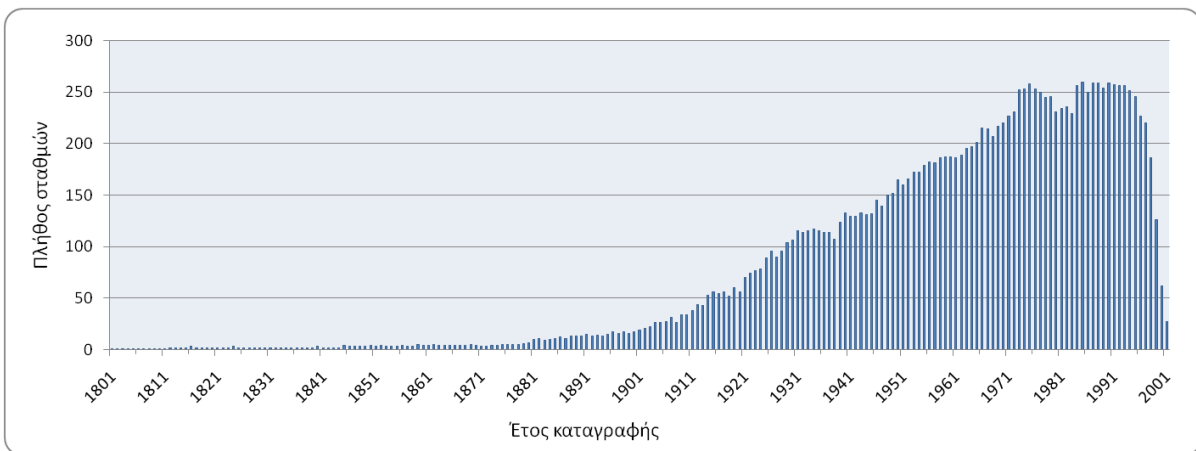
Τα δεδομένα που μελετήθηκαν προέρχονται από τον Παγκόσμιο Κατάλογο Μεγίστων Παρατηρημένων Πλημμυρών της Διεθνούς Ένωσης Υδρολογικών Επιστημών (World Catalogue of Maximum Observed Floods - International Association Of Hydrological Sciences). Ο κατάλογος παρέχει πληροφορίες πλημμυρικών επεισοδίων από 168 χώρες ανά τον κόσμο οι οποίες διατίθενται σε μορφή τριών διαφορετικών πινάκων. Το περιεχόμενο των πινάκων αυτών είναι το εξής:

- Πίνακας (1) : Οι τοποθεσίες και η κωδική ονομασία των σταθμών μέτρησης καθώς και περαιτέρω πληροφορίες (όπου είναι διαθέσιμες) σχετικά με την γεωμορφολογία και το κλιματικό καθεστώς της περιοχής (π.χ. έκταση της λεκάνης απορροής, μέση τιμή βροχόπτωσης κατά την περίοδο καταγραφής, κλίση εδάφους της λεκάνης απορροής, φυτοκάλυψη)
- Πίνακας (2) : Στοιχεία των μεγίστων πλημμυρικών επεισοδίων ανά σταθμό παρατήρησης του Πίνακα (1), όπως και συμπληρωματικές πληροφορίες για τις συνθήκες του συγκεκριμένου πλημμυρικού επεισοδίου (π.χ. διάρκεια βροχόπτωσης)
- Πίνακας (3): Αναλυτικές χρονοσειρές από επιλεγμένους σταθμούς των στιγμιαίων μεγίστων παροχών ανά έτος

Οι Πίνακες (1) και (2) διαθέτουν στοιχεία για πάνω από 1500 σταθμούς παγκοσμίως ενώ οι χρονοσειρές του Πίνακα (3) προέρχονται από 371 σταθμούς, οι οποίοι βέβαια περιλαμβάνονται στους πρώτους δύο πίνακες. Στο σύνολο των 371 διαθέσιμων χρονοσειρών, το μήκος των παρατηρήσεων εκτείνεται από 8 έως και 246 χρόνια, με μετρήσεις από τον 18^ο αιώνα μέχρι το 2001. Οι μέθοδοι εκτίμησης των μεγίστων παροχών ήταν είτε από μόνιμους σταθμούς μέτρησης, είτε από σημεία όπου η εκ των υστέρων

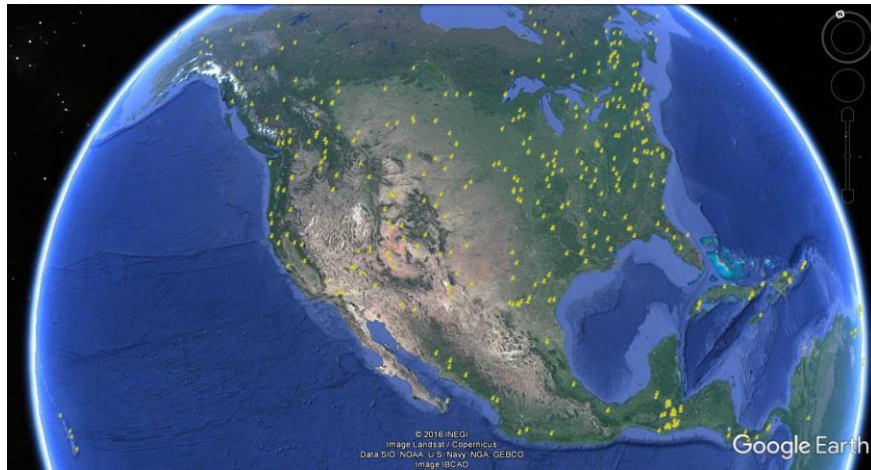
εκτίμηση της πλημμύρας ήταν δυνατή, είτε από σημεία όπου το μέγιστο επίπεδο της πλημμύρας ήταν ορατό. Για τα περισσότερα δείγματα γίνεται αξιολόγηση της πηγής προέλευσης με τη χρήση «σημαιών ποιότητας».

Η συγκρότηση της συγκεκριμένης βάσης δεδομένων είναι το αποτέλεσμα της χρόνιας συνεργασίας της Διεθνούς Ένωσης Υδρολογικών Επιστημών (IAHS), της UNESCO και του Παγκόσμιου Μετεωρολογικού Οργανισμού (WMO) καθώς και των επαφών της IAHS με τους εθνικούς εκπροσώπους του Παγκόσμιου Υδρολογικού Προγράμματος (IHP). Η καταγραφή πλημμυρικών παροχών είναι αδιαμφισβήτητα μία πολύ δύσκολη διαδικασία. Το ποσοστό αβεβαιότητας που δίνεται από του συντάκτες της συγκεκριμένης βάσης δεδομένων φτάνει το 10% και αναφέρεται στη μεροληψία που προκύπτει από τις ιστορικές αλλαγές στα όργανα και τις μεθόδους μέτρησης. Η χρονική εξέλιξη του δικτύου των σταθμών είναι επίσης ιδιαίτερα έντονη. Χαρακτηριστικό είναι ότι πριν το 1900 συνολικά οι διαθέσιμοι σταθμοί αποτελούσαν το 6,5% του μεγίστου αριθμού των σταθμών κατά τον επόμενο αιώνα, όπως φαίνεται και από το Σχήμα 5.1.



Σχήμα 5.1 Πλήθος σταθμών για κάθε έτος καταγραφής

Στο Σχήμα 5.2 της επόμενης σελίδας φαίνονται κάποιες απεικονίσεις της γεωγραφικής κατανομής των σταθμών πάνω στη γη με χρήση του προγράμματος Google Earth. Η πυκνότητα του δικτύου των σταθμών είναι γενικά μεγαλύτερη στις Η.Π.Α. και την κεντρική Ευρώπη και μικρότερη σε περιοχές όπως η Αφρική, η Λατινική Αμερική και η Ασία.



Σχήμα 5.2 Γεωγραφική κατανομή σταθμών

5.2 Επιλογή κατάλληλων δεδομένων προς επεξεργασία

Για την επιθυμητή στατιστική ανάλυση, την εξαγωγή αξιόπιστων αποτελεσμάτων, την ανάλυση τάσεων αλλά και τον υπολογισμό στατιστικών παραμέτρων που επηρεάζονται έντονα από τη δομή συσχέτισης των χρονοσειρών (όπως π.χ., ο συντελεστής Hurst) τα κατάλληλα δεδομένα επιλέχθηκαν ως εξής:

- Επιλέχθηκαν, αποκλειστικά, σταθμοί με τουλάχιστον 50 καταγραφές στις χρονοσειρές
- Διορθώθηκαν όλες οι προβληματικές τιμές δηλαδή οι αδικαιολόγητα πολύ μεγάλες και αντικαταστάθηκαν με κενό. Ένδειξη για τον εντοπισμό των προβληματικών αυτών τιμών αποτέλεσε η πολύ μεγάλη τιμή της κύρτωσης της χρονοσειράς.

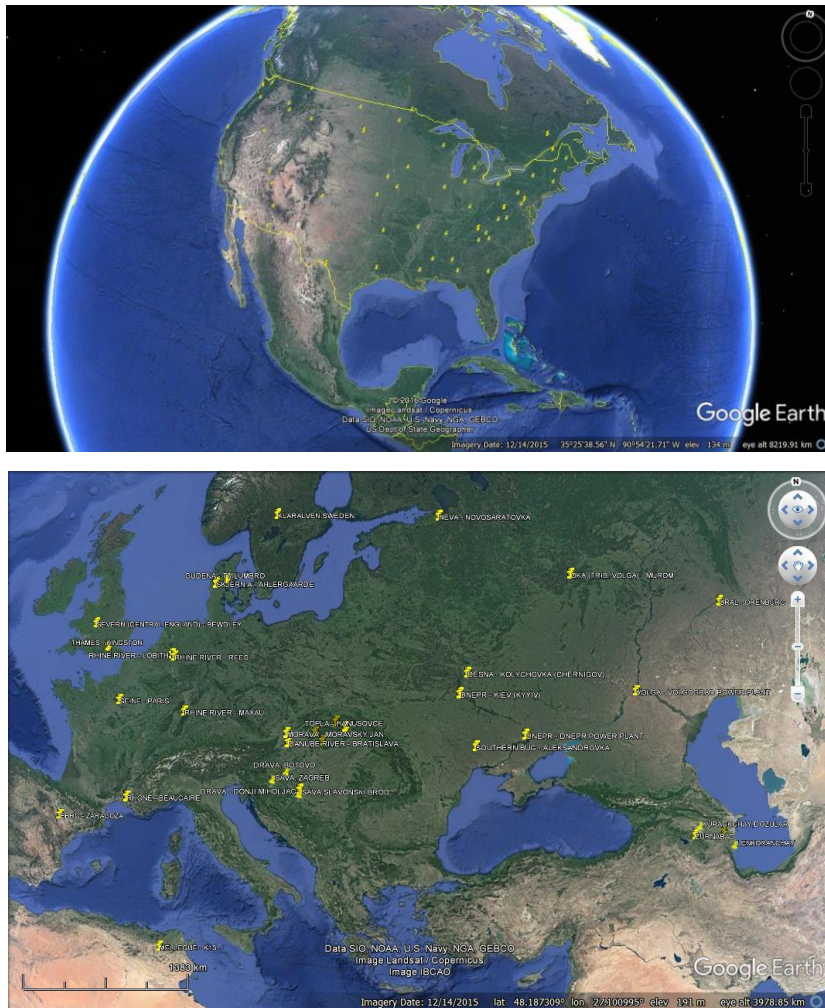
Έτσι προέκυψαν 125 σταθμοί από 26 διαφορετικές χώρες του κόσμου. Στον Πίνακα 5.1 φαίνεται η κατανομή των επιλεγμένων σταθμών ανά ήπειρο.

Πίνακας 5.1 Αριθμός των 125 επιλεγμένων σταθμών ανά ήπειρο και σε σχέση με τα έτη καταγραφής

ΑΡΙΘΜΟΣ ΕΤΩΝ ΚΑΤΑΓΡΑΦΗΣ	50-60	61-70	71-80	81-90	91-100	>100	ΣΥΝΟΛΟ
ΕΥΡΩΠΗ	4	12	13	9	5	5	48
ΑΦΡΙΚΗ	3	2	4				9
ΑΣΙΑ	6	2	2				10
ΩΚΕΑΝΙΑ			1				1
Β. ΑΜΕΡΙΚΗ	12	12	17	10	3	2	56
Ν. ΑΜΕΡΙΚΗ		1					1
ΣΥΝΟΛΟ ΣΤΑΘΜΩΝ	25	29	37	18	9	7	125

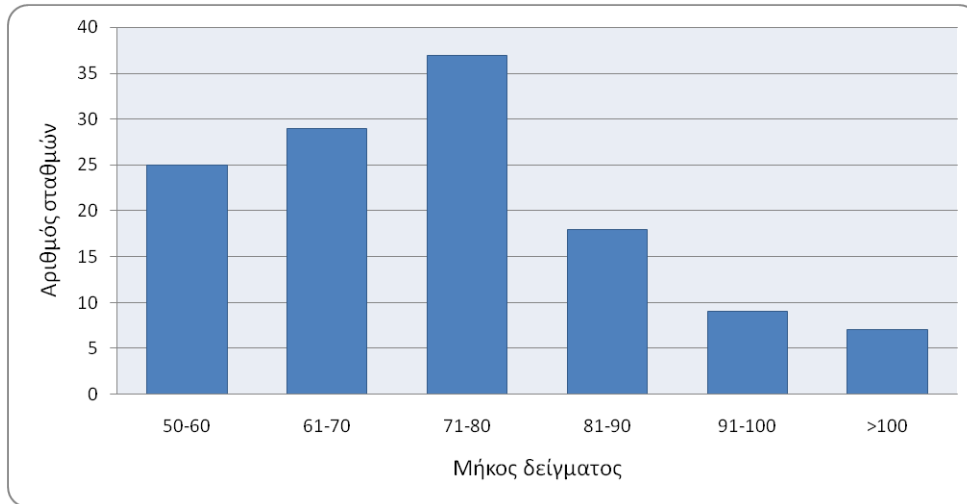
Είναι φανερό ότι και μετά την αρχική επιλογή οι σταθμοί της Ευρώπης και της Β. Αμερικής υπερτερούν αριθμητικά. Ο μέσος όρος των ετών καταγραφής είναι 76 έτη και ο σταθμός με τις περισσότερες μετρήσεις είναι αυτός στο Παρίσι της Γαλλίας στον ποταμό Σηκουάνα με συνολικά 146 έτη καταγραφής. Μόνο επτά ποτάμια διαθέτουν καταγραφές με περισσότερα από εκατό χρόνια και 33 ποτάμια από τα επιλεγμένα 125 διαθέτουν πλήρεις καταγραφές.

Στο Σχήμα 5.3 δίνονται απεικονίσεις από το Google Earth με τη θέση των 125 επιλεγμένων σταθμών προς διερεύνηση πάνω στη γη.



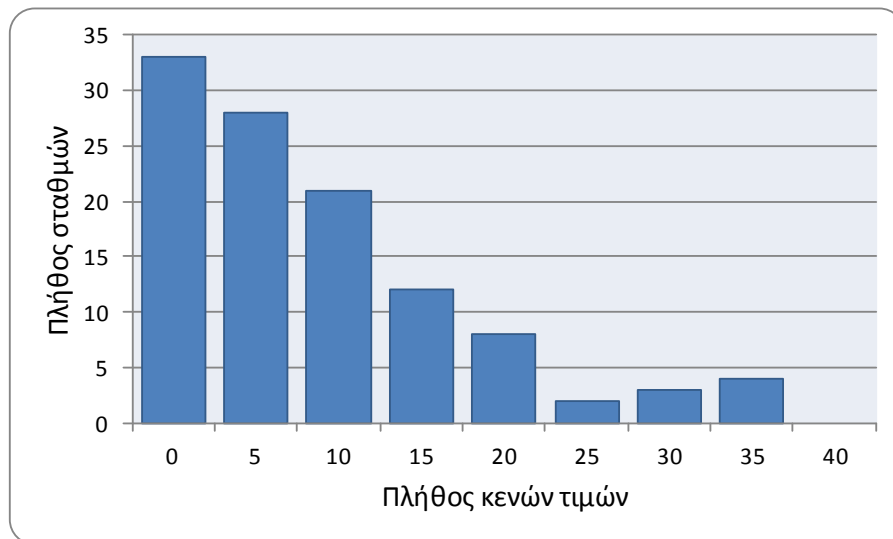
Σχήμα 5.3 Γεωγραφική κατανομή των 125 επιλεγμένων σταθμών

Στο Σχήμα 5.4 φαίνεται ότι η μέση τιμή των καταγραφών είναι 76 έτη, η διάμεσός τους είναι 72 έτη, η τυπική απόκλιση 23.5 έτη και διατίθενται συνολικά 9846 καταγραφές.

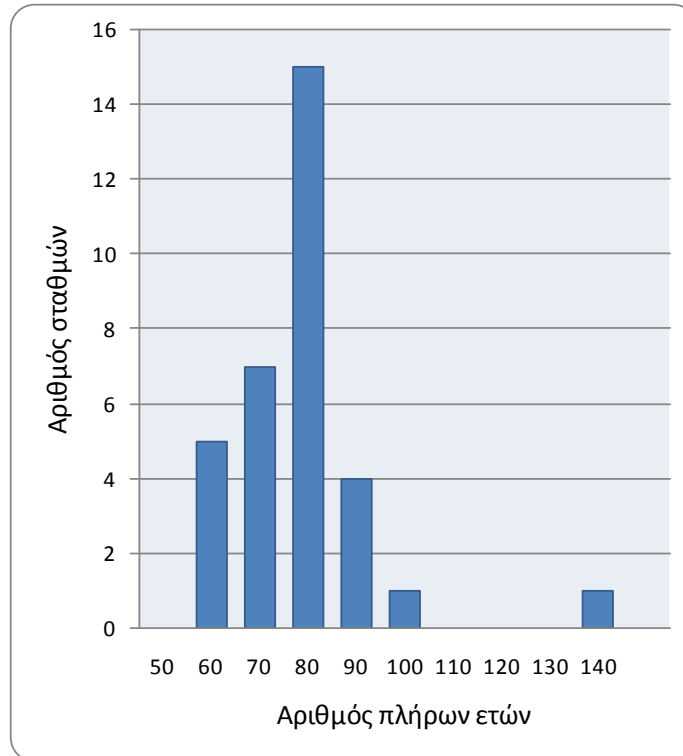


Σχήμα 5.4 Κατανομή σταθμών σε σχέση με το μήκος δείγματος

Η μέση τιμή των κενών στις χρονοσειρές είναι 7.1 κενές τιμές, η διάμεσος τους είναι 4, η τυπική απόκλιση 8.5 σε συνολικά 789 κενές τιμές, όπως φαίνεται και στο Σχήμα 5.5.



Σχήμα 5.5 Κατανομή των κενών τιμών



Σχήμα 5.6 Κατανομή πλήρων ετών

Στο Σχήμα 5.6 φαίνεται το ιστόγραμμα της κατανομής των 33 σταθμών με πλήρεις καταγραφές, όπου η μέση τιμή είναι 75 έτη, η διάμεσος 75, η τυπική απόκλιση 15.2 και οι πλήρεις καταγραφές συνολικά είναι 2480.

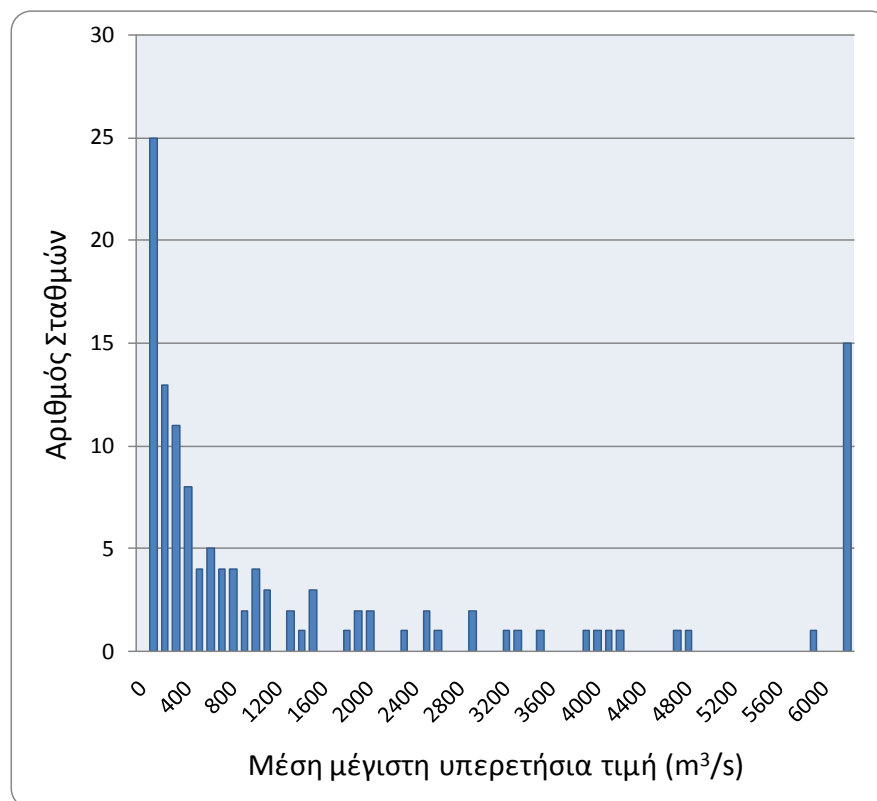
5.3 Στατιστική επεξεργασία των σταθμών

Παρακάτω δίνονται διάφορα γραφήματα και πίνακες που παρουσιάζουν το στατιστικό προφίλ των 125 σταθμών που πληρούσαν τα προαναφερθέντα κριτήρια. Για κάθε σταθμό υπολογίστηκε η μέση τιμή του δείγματος, η μέση μέγιστη υπερετήσια τιμή (m^3/s). Υπολογίστηκαν επίσης η τυπική απόκλιση, η διάμεσος και ο συντελεστής ασυμμετρίας του δείγματος των ετήσιων μεγίστων παροχών, καθώς και οι ακραίες παρατηρήσεις του δείγματος, δηλαδή η ελάχιστη και η μέγιστη τιμή. Για κάθε ένα από αυτά τα μεγέθη έγινε στατιστική ανάλυση σε επίπεδο δείγματος σταθμών, που περιλαμβάνει τον υπολογισμό της μέσης τιμής τους, της τυπικής απόκλισης, της ελάχιστης τιμής, των βασικών ποσοστημορίων και της μέγιστης τιμής. Τα αποτελέσματα της ανάλυσης φαίνονται στον Πίνακα 5.2 ενώ οι κατανομές των χαρακτηριστικών στατιστικών μεγεθών φαίνονται στα Σχήματα 5.7-5.10.

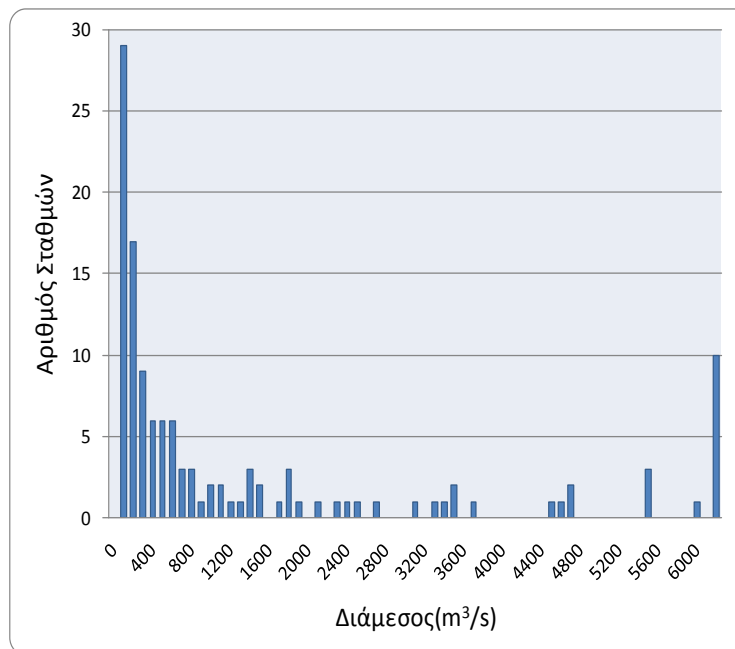
Πίνακας 5.2 Στατιστικά χαρακτηριστικά στατιστικών μεγεθών του δείγματος

	ΜΕΣΗ ΤΙΜΗ	ΤΥΠΙΚΗ ΑΠΟΚΛΙΣΗ	ΔΙΑΜΕΣΟΣ	ΣΥΝΤΕΛΕΣΤΗΣ ΑΣΥΜΜΕΤΡΙΑΣ	ΕΛΑΧΙΣΤΗ ΤΙΜΗ	ΜΕΓΙΣΤΗ ΤΙΜΗ
ΜΕΣΗ ΤΙΜΗ	2468.09	1127.04	2236.30	1.58	1092.24	5829.16
ΤΥΠΙΚΗ ΑΠΟΚΛΙΣΗ	6514.39	2083.68	6471.62	0.98	4437.52	11043.89
ΕΛΑΧΙΣΤΗ ΤΙΜΗ	9.58	6.00	7.50	-0.11	0.62	32.00
Q _{2.5}	18.23	8.27	10.81	0.25	1.34	55.78
Q ₂₅	144.54	95.78	115.00	0.83	20.70	501.00
Q ₅₀	525.31	342.39	419.00	1.49	93.00	1750.00
Q ₇₅	1952.80	1090.51	1750.00	2.20	521.00	5070.00
Q _{97.5}	10551.84	7750.69	10288.00	3.99	6350.00	32741.00
ΜΕΓΙΣΤΗ ΤΙΜΗ	58560.25	11983.99	59280.00	5.02	44510.00	81110.00

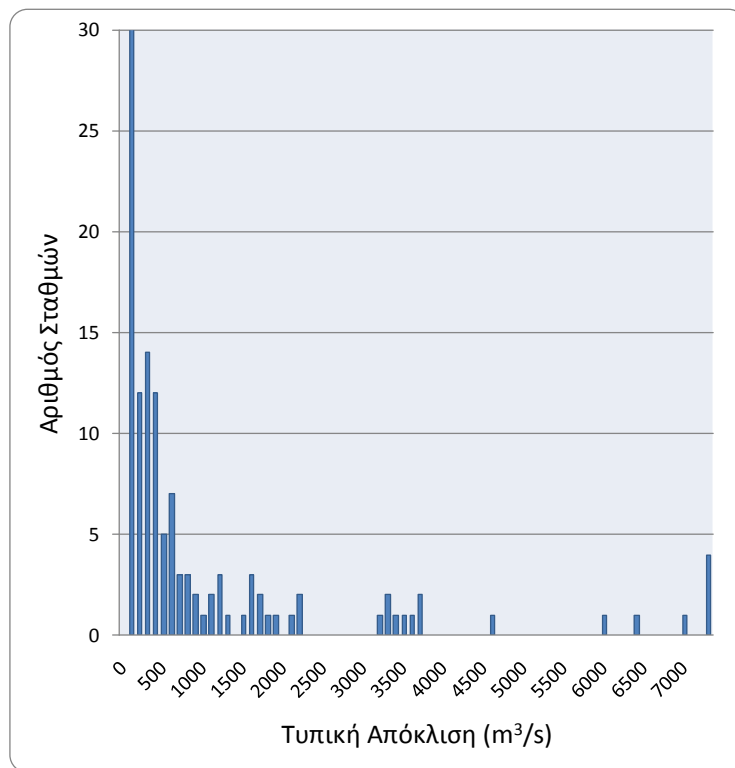
Στον Πίνακα 5.2 φαίνεται ότι η μέση τιμή της μέγιστης υπερετήσιας τιμής είναι 2468.1 m³/s. Η τυπική της απόκλιση με τιμή 6514.14 m³/s, είναι ενδεικτική των μεγάλων διαφορών που εντοπίζονται σε παγκόσμια κλίμακα, όπως επίσης και το γεγονός πως το 95% των τιμών βρίσκεται στο διάστημα 18.23-10551.84 m³/s. Η διάμεσος αλλά και τα υπόλοιπα στατιστικά χαρακτηριστικά είναι αρκετά κοντά στις εκτιμήσεις με τις αντίστοιχες της μέσης τιμής. Επίσης, το μεγάλο εύρος της ελάχιστης και της μέγιστης τιμής είναι ενδεικτικό της μεγάλης διασποράς που χαρακτηρίζει το φαινόμενο σε παγκόσμια κλίμακα. Σημειώνεται ακόμα, ότι μια τιμή μέγιστης παροχής για κάποιο έτος δεν αποτελεί απαραίτητα πλημμυρική παροχή με τη έννοια ενός καταστροφικού φαινομένου, αλλά είναι η μεγαλύτερη στιγμιαία παροχή που μετρήθηκε κατά το συγκεκριμένο έτος, γεγονός που δικαιολογεί ακόμα καλύτερα τις μεγάλες αποκλίσεις στα παραπάνω στατιστικά χαρακτηριστικά.



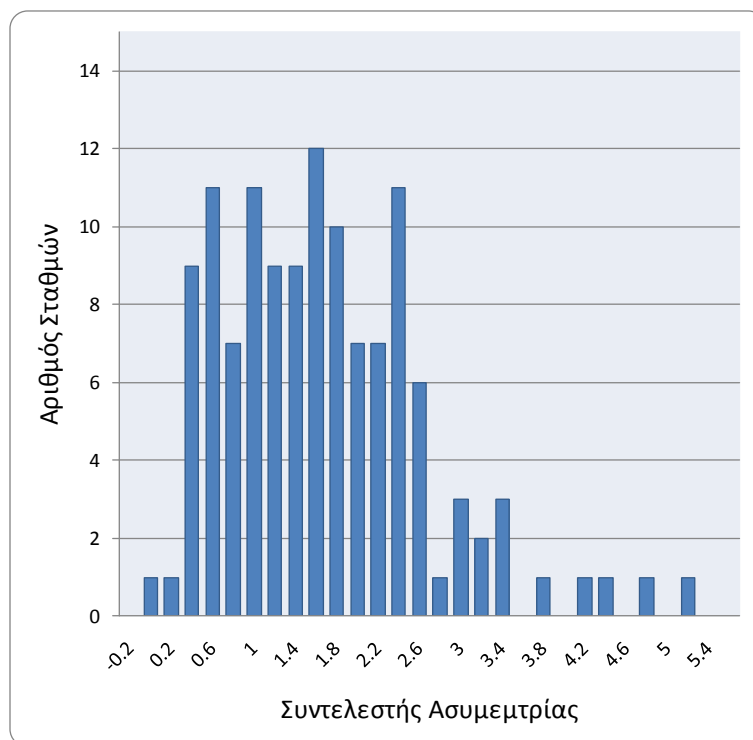
Σχήμα 5.7 Κατανομή μέγιστης υπερετήσιας τιμής



Σχήμα 5.8 Κατανομή διαμέσου



Σχήμα 5.9 Κατανομή τυπικής απόκλισης

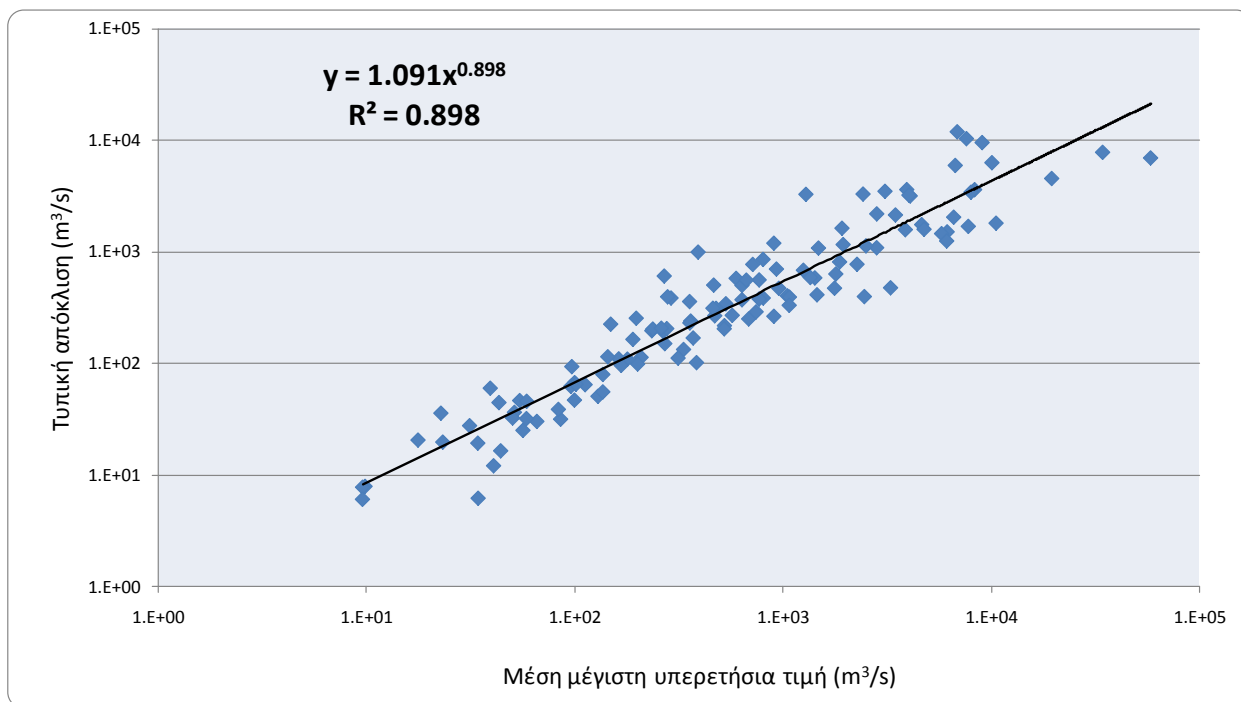


Σχήμα 5.10 Κατανομή του συντελεστή ασυμμετρίας

Στα παραπάνω γραφήματα φαίνονται πιο εποπτικά κάποια από τα συμπεράσματα, όπως το γεγονός ότι η κατανομή της μέση τιμής πλησιάζει την κατανομή της διαμέσου. Η θετική ασυμμετρία που διέπει τις μέγιστες ετήσιες παροχές των σταθμών είναι αναμενόμενη, όπως συμβαίνει και σε άλλα ακραία φαινόμενα (δεδομένα μεγίστων βροχοπτώσεων, ξηρασιών, κ.α.). Παρατηρούμε επίσης ότι στο σύνολο των 125 σταθμών, 15 σταθμοί παρουσιάζουν μέσες τιμές πάνω από τα 6000 m³/s ενώ 25 έχουν μέση τιμή μικρότερη ή ίση των 100 m³/s. Συγκεκριμένα, η μέγιστη μέση υπερετήσια τιμή παρατηρήθηκε στο σταθμό Kihasa που βρίσκεται στο Κονγκό, στον ποταμό Ζαΐρ.

Η τυπική απόκλιση είναι σχεδόν για τους μισούς σταθμούς μικρότερη ή ίση των 300 m³/s ενώ σε άλλες περιοχές παρατηρείται τυπική απόκλιση μεγαλύτερη των 5000 m³/s. Η μεγαλύτερη τυπική απόκλιση μετρήθηκε στον σταθμό Scotia των Η.Π.Α. στον ποταμό Eel.

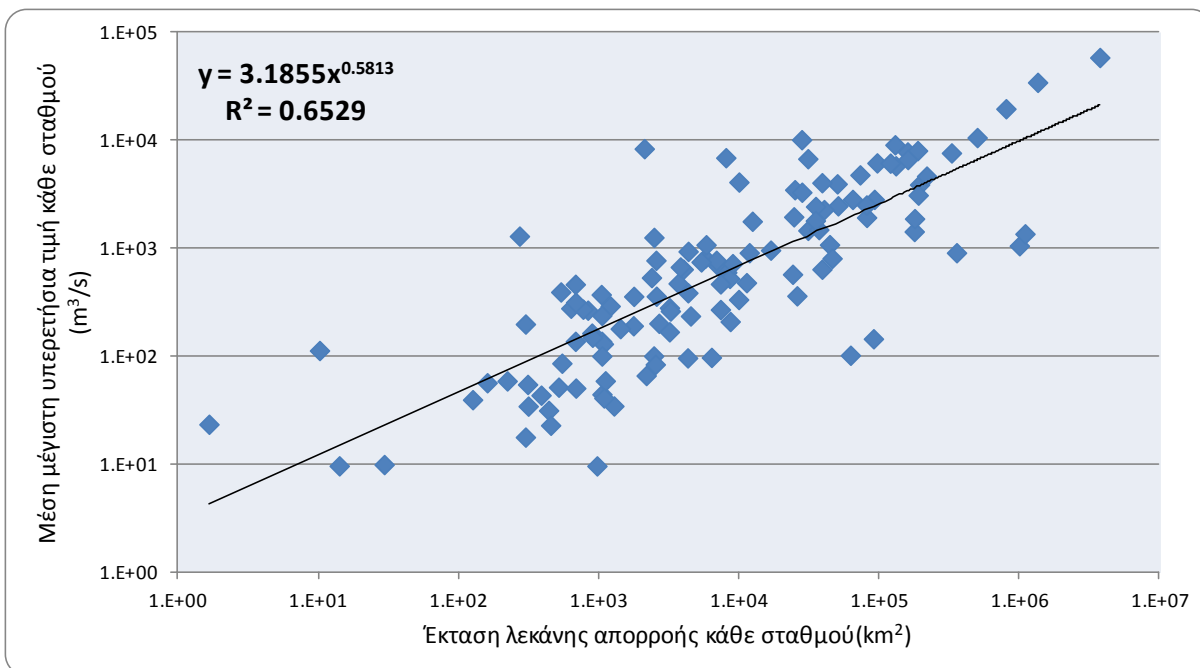
Πιο συγκεκριμένα, φαίνεται ότι και η κατανομή της τυπικής απόκλισης είναι αρκετά κοντά με αυτή της μέσης τιμής, κάτι που είναι λογικό όπως φαίνεται και από το Σχήμα 5.11 όπου απεικονίζεται η μέση μέγιστη υπερετήσια τιμή έναντι της τυπικής της απόκλισης. Παρόμοια σχέση παρουσιάζουν και η μέση ετήσια παροχή ενός ποταμού με την τυπική της απόκλιση (Macmahon et al., 2007).



Σχήμα 5.11 Διάγραμμα της μέσης μέγιστης υπερετήσιας τιμής και της τυπικής της απόκλισης

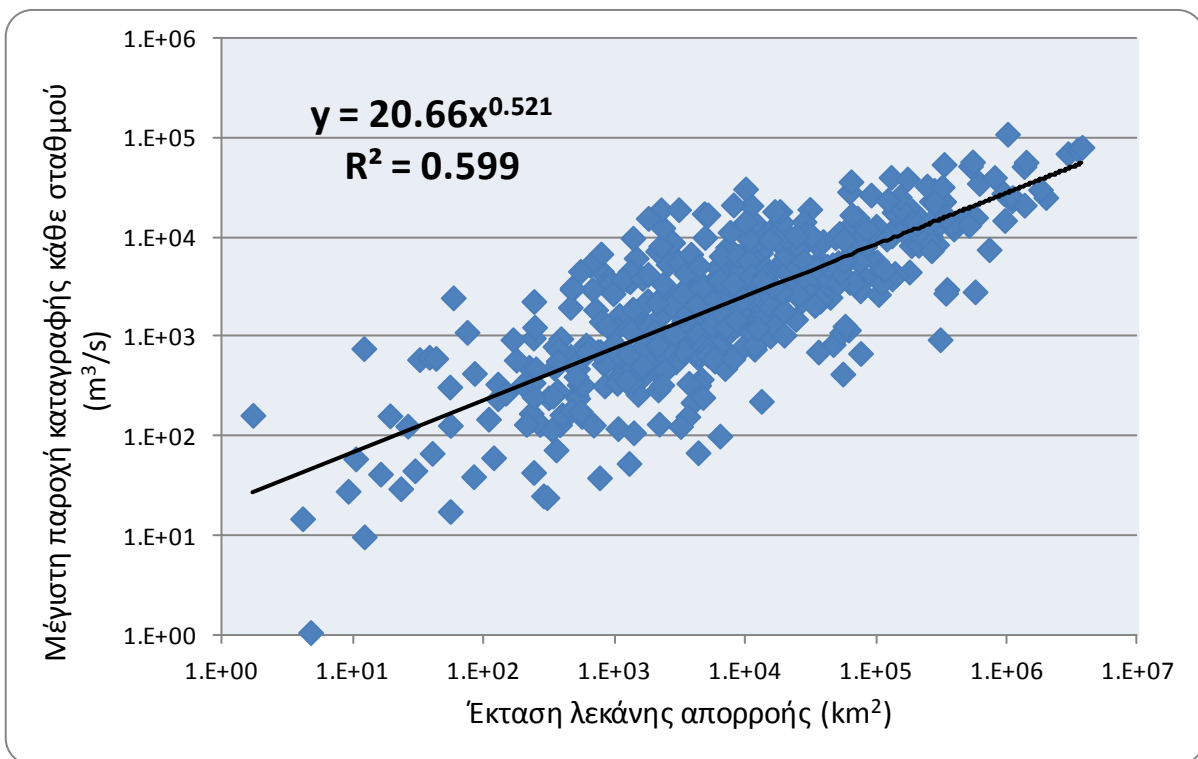
Η μέση τιμή των μεγίστων παροχών σε κάθε σταθμό και η τυπική τους απόκλιση συνδέονται με μια σχέση δύναμης και μάλιστα η στατιστική αβεβαιότητα είναι ιδιαίτερα χαμηλή.

Εξετάστηκε επίσης, η σχέση που μπορεί να υπάρχει ανάμεσα στο μέγεθος της μέσης μέγιστης υπερετήσιας παροχής και στην έκταση της λεκάνης απορροής. Πράγματι, όπως φαίνεται και στο Σχήμα 5.12 συνδέονται με σχέση δύναμης και με μία σχετικά ικανοποιητική στατιστική αβεβαιότητα.



Σχήμα 5.12 Διάγραμμα της έκτασης της λεκάνης απορροής συναρτήσει της μέσης μέγιστης υπερετήσιας τιμής κάθε σταθμού

Πέρα από τη σχέση της μέσης μέγιστης υπερετήσιας τιμής και της λεκάνης απορροής εξετάστηκε και η σχέση της δεύτερης με την μέγιστη τιμή της κάθε χρονοσειράς. Όπως αναφέρεται στο υποκεφάλαιο 5.1 ο Πίνακας (2) της βάσης δεδομένων παρέχει πληροφορίες για την μέγιστη καταγεγραμμένη πλημμύρα για πάνω από 1500 σταθμούς, μεταξύ των οποίων βρίσκονται και οι 125 σταθμοί με τις διαθέσιμες χρονοσειρές. Επιλέχτηκαν και από τον Πίνακα (2) οι σταθμοί με τουλάχιστον 50 χρόνια μετρήσεων και η σχέση της μέγιστης των μεγίστων παροχών με τη λεκάνη απορροής φαίνεται στο Σχήμα 5.13. Ο αριθμός των σταθμών του Πίνακα (2) που πληρούν το παραπάνω κριτήριο είναι συνολικά 481.



Σχήμα 5.13 Διάγραμμα της έκτασης της λεκάνης απορροής συναρτήσει της μέγιστης παρατηρημένης παροχής του κάθε σταθμού

Όπως και προηγουμένως με τη μέση μέγιστη τιμή, φαίνεται ότι και εδώ τα δύο μεγέθη συνδέονται με σχέση δύναμης και μάλιστα με παραπλήσια στατιστική αβεβαιότητα. Εκτός αυτού, η ετεροσυσχέτιση μεταξύ των δύο μεγεθών στις δύο τελευταίες περιπτώσεις υπολογίστηκε αντίστοιχα 0.89 και 0.72, ένδειξη όχι μόνο για την ισχυρή τους συσχέτιση αλλά και για την μεγαλύτερη αβεβαιότητα που διέπει τη μέγιστη τιμή των μεγίστων σε κάθε σταθμό.

Τέλος, λαμβάνοντας υπόψη την σχέση της λεκάνης απορροής με την παροχή αλλά και την σχέση της μέσης τιμής με την τυπική απόκλιση, όπως παρουσιάστηκε στην προηγούμενη ενότητα, έγινε αδιαστατοποίηση των δεδομένων, διαιρώντας την κάθε τιμή των χρονοσειρών με την μέση τους τιμή. Έτσι καταλήξαμε σε αδιάστατες χρονοσειρές με μέση

τιμή 1, απαλλαγμένες από τον παράγοντα «λεκάνη απορροής» μιας και ενέχεται στον υπολογισμό της μέσης τιμής της χρονοσειράς.

5.4 Ανθρώπινος παράγοντας

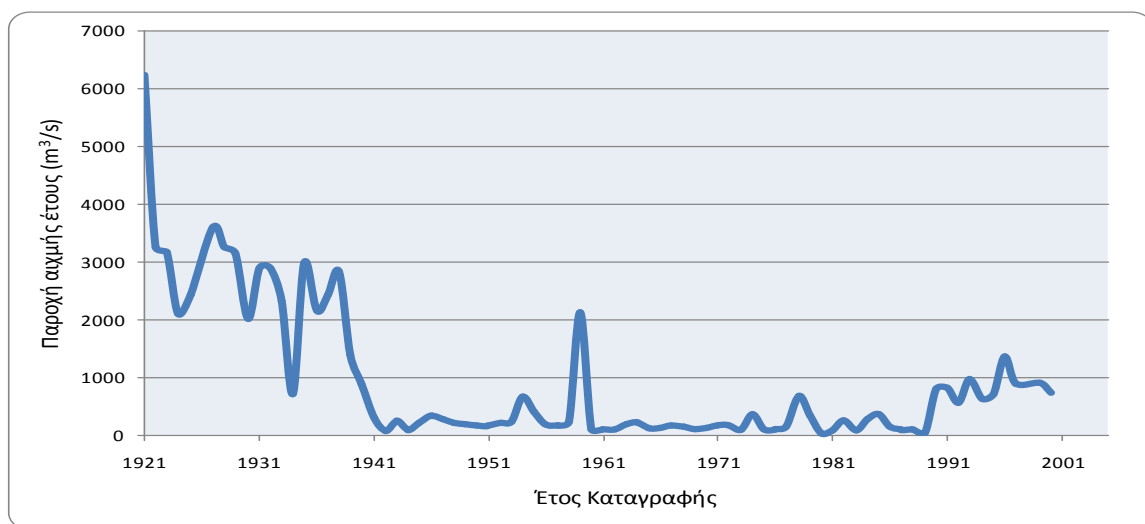
Τον τελευταίο μισό αιώνα τα ποσοστά θνησιμότητας καθώς και οι οικονομικές απώλειες που σχετίζονται με τις πλημμύρες έχουν αυξηθεί δραματικά. Ενδεικτικά αναφέρεται ότι οι πλημμύρες στις Η.Π.Α. προκαλούν τις περισσότερες καταστροφές σε σχέση με οποιοδήποτε άλλο ακραίο καιρικό φαινόμενο (Mallakrour and Villarini, 2015;). Έχει έτσι εμφανιστεί ένα παγκόσμιο και συνεχώς αυξανόμενο ενδιαφέρον για την ανάγκη αναγνώρισης των αιτιών που προκαλούν τέτοια φαινόμενα. Η ανάλυση για την ύπαρξη αυξητικών ή μειωτικών τάσεων στα πλημμυρικά επεισόδια αποτελεί την πρωταρχική ιδέα για την επιβεβαίωση ή μη των παραπάνω ανησυχιών.

Πριν όμως κανείς προχωρήσει σε μια τέτοια ανάλυση, θα πρέπει να λάβει υπόψη του την ενδεχόμενη επιρροή των ανθρώπινων παρεμβάσεων στις λεκάνες απορροής των ποταμών και στις ευρύτερες περιοχές, όπως μπορεί να είναι η ύπαρξη ταμιευτήρων ή/και μεγάλων υδροηλεκτρικών έργων.

Για παράδειγμα, αναφέρονται οι παράγοντες που με βάση την Ευρωπαϊκή Οδηγία του 2000 (2000/06/EC) καθιστούν ένα ποτάμι «βαρέως τροποποιημένο» (*heavily modified river basin*):

- Μεταβολή στο ρου του ποταμού
- Ύπαρξη δραστηριοτήτων αποθήκευσης υδάτων για ύδρευση, άρδευση, παραγωγή ενέργειας
- Αντιπλημμυρικά έργα, αποστράγγιση του εδάφους
- Άλλες δραστηριότητες εξίσου σημαντικές που επηρεάζουν την βιώσιμη ανθρώπινη ανάπτυξη

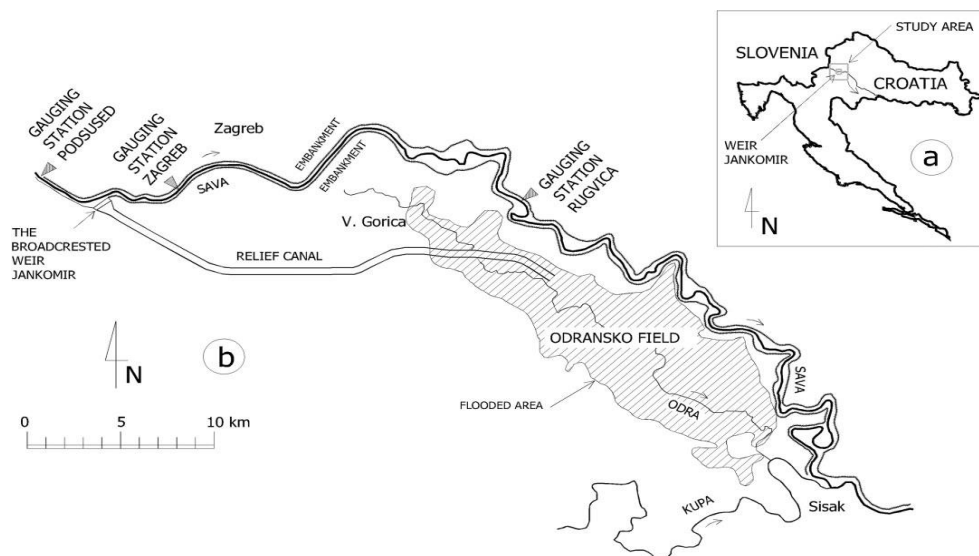
Με βάση τα παραπάνω αλλά και με δεδομένο ότι επεξεργαζόμαστε μέγιστες ετήσιες παροχές, επιλέχθηκε να γίνει μία λεπτομερής διερεύνηση στα 125 διαθέσιμα ποτάμια για την ύπαρξη μεγάλων ταμιευτήρων ανάντη των σταθμών μέτρησης. Το αποτέλεσμα είναι μία καινούρια βάση δεδομένων η οποία περιλαμβάνει μόνο τις χρονοσειρές που είναι απαλλαγμένες από την ύπαρξη ταμιευτήρων ανάντη των σταθμών μέτρησης και που μπορεί να επηρεάζαν έντονα την διαίτα των ποταμών. Η συγκεκριμένη ενδεδειγμένη αναζήτηση ταμιευτήρων πραγματοποιήθηκε με τη βοήθεια των χαρτών της Google, τις διαθέσιμες πληροφορίες από τη βάση δεδομένων για τις συντεταγμένες των σταθμών και είχε ως αποτέλεσμα να αποκλειστούν 12 σταθμοί. Στις περιπτώσεις όπου δεν ήταν διαθέσιμες οι συντεταγμένες των σταθμών, χρησιμοποιήθηκαν οι πληροφορίες της βάσης δεδομένων του Global Runoff Data Centre (GRDC, Koblenz, Germany) και για τους σταθμούς των Η.Π.Α. τα δεδομένα από την United States Geological Survey (USGS, USA). Ένα χαρακτηριστικό παράδειγμα για την επιρροή της διαίτας ενός ποταμού λόγω ύπαρξης ταμιευτήρα φαίνεται στο Σχήμα 5.13, όπου είναι εμφανής η μεταβολή στις παροχές αιχμής μετά το 1940 οπότε και κατασκευάστηκε φράγμα ανάντη του σταθμού μέτρησης.



Σχήμα 5.14 Χρονοσειρά του ποταμού Κολοράντο στις Η.Π.Α.

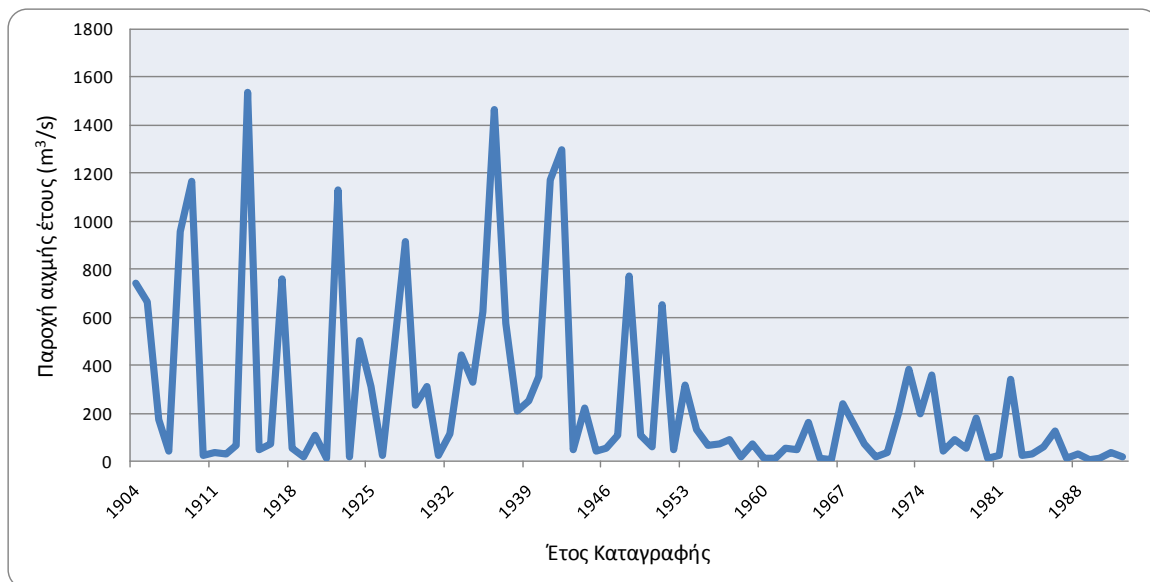
Ακολουθούν πληροφορίες σχετικά με τους 12 ποταμούς που τελικά αποκλείστηκαν από την ανάλυση:

1. Ποταμός Αράς – Σταθμός Garadonly: Ο Αράς είναι ποταμός της δυτικής Ασίας που διαρρέει εδάφη της Τουρκίας, της Αρμενίας, του Ιράν και του Αζερμπαϊτζάν. Έχει συνολικό μήκος 1.072 km και είναι παραπόταμος του ποταμού Κούρα, μαζί με τον οποίο αποτελούν τα δύο βασικότερα ποτάμια της περιοχής του Καυκάσου. Υπάρχουν δύο μεγάλα φράγματα ανάντη του σταθμού μέτρησης που χρησιμοποιούνται για αρδευτικούς, υδρευτικούς αλλά και για σκοπούς παραγωγής ηλεκτρικής ενέργειας. Το μεγαλύτερο από αυτά, το Aras Dam, κατασκευάστηκε μεταξύ του 1963-1970.
2. Ποταμός Σάβος – Σταθμός Μπροντ : Ο Σάβος είναι ποταμός της Σλοβενίας, Κροατίας, Βοσνίας και Ερζεγοβίνης και της Σερβίας, δεξιός παραπόταμος του Δούναβη. Έχει μήκος 940 km και είναι το πιο μακρύ ποτάμι στη νοτιοανατολική Ευρώπη. Ο συγκεκριμένος σταθμός βρίσκεται στην Κροατία έξω από την κωμόπολη Μπροντ και βόρεια αυτού κατασκευάστηκε το 1970 υδατοφράκτης (Σχήμα 5.15) που ακολουθείται από κανάλι ανακούφισης με σκοπό την αντιπλημμυρική προστασία.



Σχήμα 5.15 Θέση του κατασκευασμένου υδατοφράκτη στον ποταμό Σάβο και του σταθμού μέτρησης

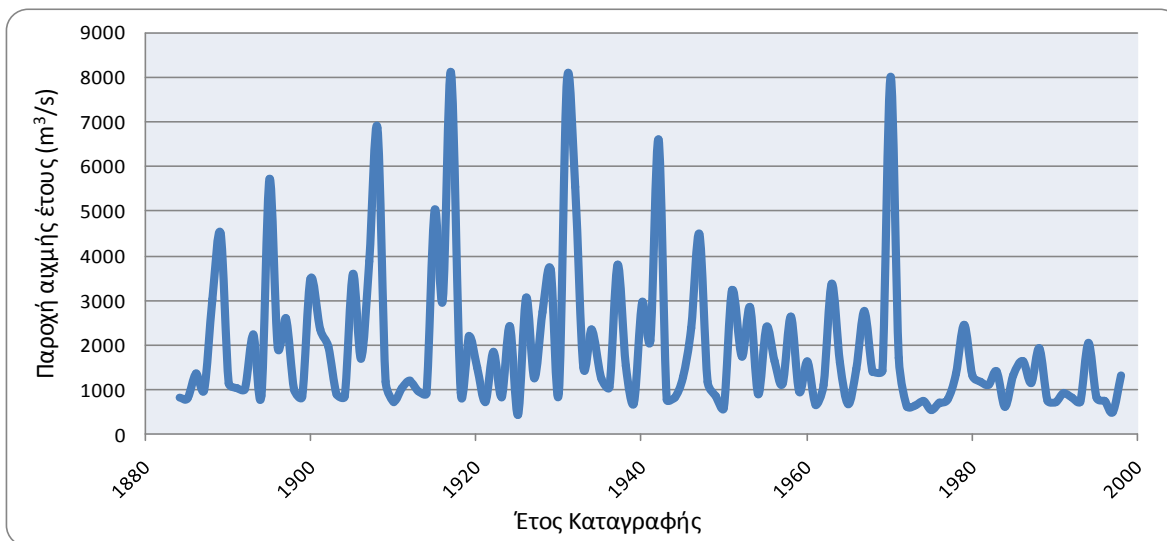
3. Ποταμός Σάιρ – Μαλάουι : Αποτελεί το σημαντικότερο ποταμό του Μαλάουι, έχοντας μήκος πάνω από 400 km. Το 1970 κατασκευάστηκε φράγμα (Nkula Dam) ανάντη του σταθμού μέτρησης εξυπηρετώντας κυρίως σκοπούς παραγωγής ηλεκτρικής ενέργειας αλλά και περιορισμένης κάλυψης αρδευτικών αναγκών.
4. Ποταμός Μαγκάλιες –Σταθμός Hartbeespoort : Ο συγκεκριμένος σταθμός βρίσκεται στη Νότια Αφρική, ακριβώς κατάντη από το φράγμα Hartbeespoort το οποίο κατασκευάστηκε αρχικά το 1920 και ανακατασκευάστηκε το 1969, έχει ολική χωρητικότητα 195 εκ. m³, είναι τοξωτό και χρησιμοποιείται κυρίως για αρδευτικούς σκοπούς αλλά και για την κάλυψη των υδρευτικών αναγκών των βιομηχανιών της περιοχής. Είναι εμφανής η επιρροή του ταμιευτήρα στην παροχή αιχμής όπως φαίνεται και στο Σχήμα 5.16 της χρονοσειράς του ποταμού.



Σχήμα 5.16 Χρονοσειρά ποταμού Μαγκάλιες στη Νότια Αφρική στον σταθμό Hartbeespoort

5. Ποταμός Βάαλ – Σταθμός Σταντερτόουν : Ο ποταμός Βάαλ βρίσκεται στη Νότια Αφρική και αποτελεί έναν από τους βασικότερους παραποτάμους του Οράγγη. Έχει συνολικό μήκος 1.120 km και λεκάνη απορροής περί τα 400.000 km². Το 1981 ανάντη του σταθμού μέτρησης κατασκευάστηκε ταμιευτήρας συνολικής χωρητικότητας 350 εκ. m³ για την εξυπηρέτηση των υδρευτικών αναγκών των βιομηχανιών της περιοχής.
6. Ποταμός Οράγγης – Σταθμός Vluytjies Kraal : Ο ποταμός Οράγγης είναι ο μακρύτερος ποταμός της Νοτίου Αφρικής. Ρέει προς τα δυτικά και εκβάλλει στον Ατλαντικό Ωκεανό διανύοντας περίπου 2.2 km. Η λεκάνη απορροής του έχει έκταση 973.000 km². Ανάντη του σταθμού μέτρησης βρίσκονται δύο μεγάλα φράγματα , το Van der Kloof Dam και το Gariep Dam, τα οποία κατασκευάστηκαν το 1977 και το 1971 αντιστοίχως.
7. Ποταμός Κλαράλβεν – Σταθμός Edsforsen : Ο συγκεκριμένος ποταμός βρίσκεται στο δυτικό τμήμα της Σουηδίας, έχει συνολικό μήκος 460 km και η έκταση της λεκάνης απορροής του είναι περίπου 12.000 km². Το 1962 μπήκε σε λειτουργία το φράγμα Holjes Dam με σκοπό την παραγωγή υδροηλεκτρικής ενέργειας.

8. Ποταμός Ντέσνα – Σταθμός Σχερνίγκοβ : Ο συγκεκριμένος ποταμός διασχίζει την Ρωσία και την Ουκρανία, αποτελεί παραπόταμο του ποταμού Δνείπερου, έχει συνολικό μήκος 1.130 km και η λεκάνη απορροής εκτείνεται περί τα 90.000 km². Το 1972 κατασκευάζεται για την παραγωγή υδροηλεκτρικής ενέργειας το φράγμα Καριν , το οποίο έχει χωρητικότητα 3 km³, που όπως φαίνεται από στο Σχήμα 5.17 επηρέασε αισθητά τις ετήσιες παροχές αιχμής.

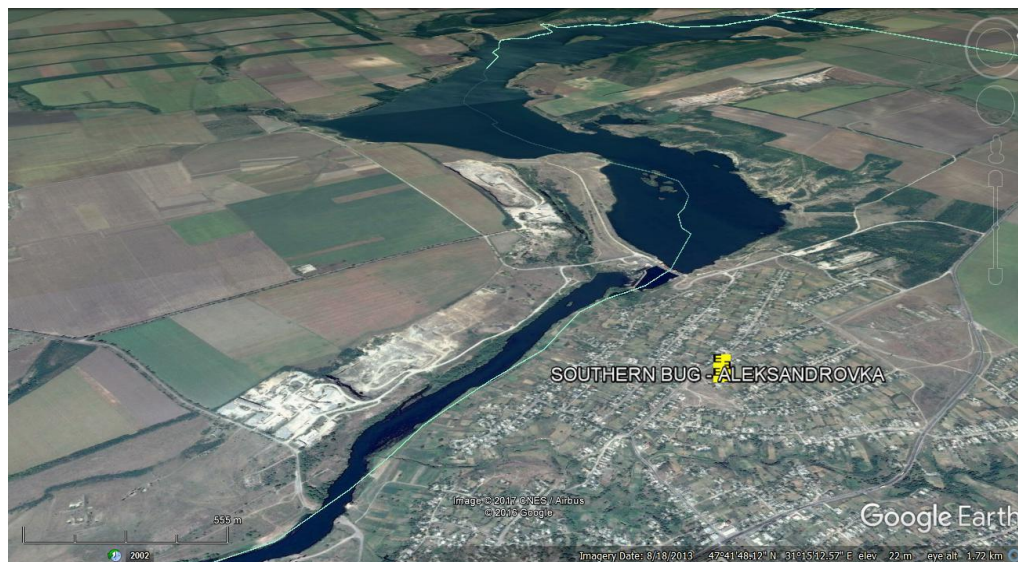


Σχήμα 5.17 Χρονοσειρά ποταμού Ντέσνα στον σταθμό Σχερνίγκοβ στην Ουκρανία

9. Ποταμός Δνείπερος – Σταθμός Κιέβου : Ο ποταμός Δνείπερος αποτελεί το τρίτο μεγαλύτερο ποτάμι στην Ευρώπη μετά τον Βόλγα και τον Δούναβη. Έχει συνολικό μήκος 2.285 km, διασχίζει την Ρωσία, την Ουκρανία και την Λευκορωσία και η λεκάνη απορροής του προσεγγίζει τα 505.000 km². Πηγάζει από τα υψώματα Βαλντάι στην κεντρική Ρωσία δυτικά της Μόσχας σε μια περιοχή σκεπασμένη από έλη και τύρφη και στη συνέχεια ρέει προς τα νότια για να καταλήξει στον Εύξεινο Πόντο. Ανάντη του σταθμού μέτρησης βρίσκεται το Kyen Dam, κατασκευασμένο το 1964.

10. Ποταμός Southern Bug – Σταθμός Αλεξαντρόφσκα : Ο ποταμός αυτός είναι ο δεύτερος μεγαλύτερος στην Ουκρανία, έχει συνολικό μήκος 806 km και η λεκάνη απορροής

του έκταση ίση με 46200 km². Όπως φαίνεται και στο Σχήμα 5.18 ανάντη του σταθμού μέτρησης υπάρχει ταμιευτήρας που καλύπτει τις αρδευτικές ανάγκες της ευρύτερης περιοχής.



Σχήμα 5.18 Σταθμός Southern Bug στην Ουκρανία και η θέση του ταμιευτήρα ανάντη του σταθμού

11. Ποταμός Κολοράντο – Σταθμός Γκραντ Κάνυον : Ο ποταμός Κολοράντο είναι από τους κυριότερους ποταμούς που ρέουν στις Η.Π.Α. και στο βόρειο τμήμα του Μεξικού. Έχει συνολικό μήκος 2.330 km και λεκάνη απορροής 637,137 km². Ο συγκεκριμένος σταθμός μας παρέχει μετρήσεις από έναν παραπόταμο του Κολοράντο αλλά σχεδόν όλοι (και ο συγκεκριμένος παραπόταμος) ελέγχονται από το φράγμα Χούβερ (Hoover Dam). Τα φράγματα που υπάρχουν κατά μήκος του Κολοράντο είναι 14 και χρησιμοποιούνται για παραγωγή υδροηλεκτρικής ενέργειας, άρδευση αλλά και την ύδρευση βιομηχανιών και παραπάνω από 35 εκ. κατοίκων. Το κυριότερο από αυτά είναι το Χούβερ, με συνολική χωρητικότητα πάνω από 44 εκ. m³ κατασκευασμένο περί το 1940, του οποίου όπως έχει προαναφερθεί είναι εμφανής η επιρροή του στις παροχές αιχμής που μας δίνονται για τον συγκεκριμένο σταθμό.



Σχήμα 5.19 Το φράγμα Hoover στον ρου του ποταμού Κολοράντο στις Η.Π.Α.

12. Ποταμός Yellowstone – Σταθμός Σύδνεϋ, ΗΠΑ : Ο ποταμός αυτός είναι παραπόταμος του ποταμού Μισούρι, έχει μήκος 1,000 km και περίπου 180,000 km² έκταση λεκάνης απορροής. Το 1953 κατασκευάστηκε το φράγμα Γκάρισον , το οποίο όπως φαίνεται να έχει επηρεάσει τις τιμές των παροχών αιχμής.

5.5 Ανάλυση τάσεων με τη μέθοδο της γραμμικής παλινδρόμησης

Όπως αναφέρθηκε και στο υποκεφάλαιο 5.4, πλέον αποτελεί παγκόσμια ανησυχία οι ολοένα και αυξανόμενες οικονομικές και ανθρώπινες απώλειες που σχετίζονται με ακραία πλημμυρικά επεισόδια. Σε ένα κλίμα που είναι κοινώς αποδεκτό ότι αλλάζει, τεράστια μερίδα της παγκόσμιας επιστημονικής κοινότητας στηρίζει τα επιχειρήματά της στο ότι οι αυξανόμενες βροχοπτώσεις θα οδηγήσουν στην αύξηση, σε συχνότητα και μέγεθος, των πλημμυρικών επεισοδίων *παγκοσμίως* (Milly et al., 2002). Στα πλαίσια επομένως αυτής της «καταστροφικής» κλιματικής αλλαγής επιχειρείται στην παρούσα εργασία μια εκτενής διερεύνηση στην ύπαρξη ή μη *στατιστικά σημαντικών* τάσεων στις ετήσιες παροχές αιχμής των 113 σταθμών που επιλέχθηκαν σύμφωνα με την διερεύνηση που αναλύεται στο υποκεφάλαιο 5.4.

Αρχικά, τα διαθέσιμα δεδομένα μας έχουν ήδη επιλεγθεί κατά τέτοιο τρόπο ώστε να υπάρχουν τουλάχιστον 50 καταγραφές σε κάθε χρονοσειρά, προϋπόθεση που είναι εξέχουσας σημασίας για την ανάλυση τάσεων (Kundzewicz and Robson, 2004).

Στη συνέχεια σε κάθε έναν από τους 113 σταθμούς υπολογίστηκε η γραμμή τάσης ανά 15ετία, σύμφωνα με την παρακάτω εξίσωση της τυπικής γραμμικής παλινδρόμησης :

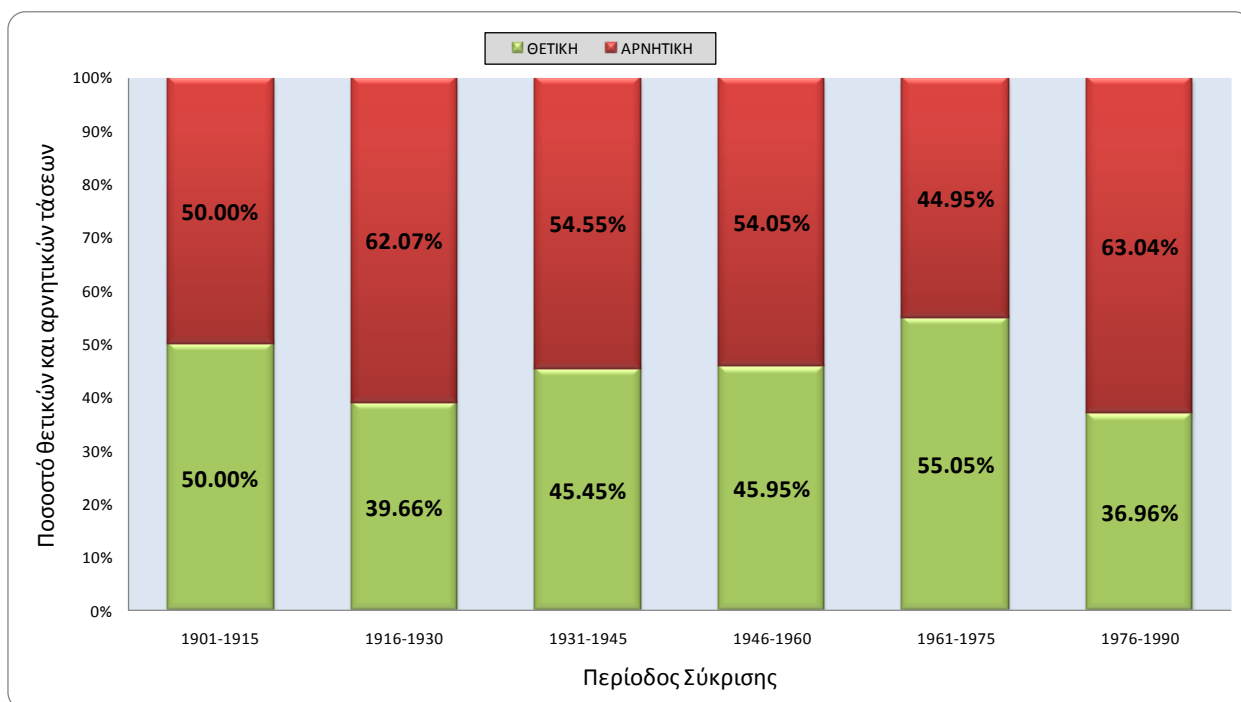
$$m = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (51)$$

Για τον αξιόπιστο υπολογισμό της κάθε τάσης τέθηκε ως κριτήριο το να υπάρχουν τουλάχιστον 10 διαθέσιμες καταγραφές «εσωτερικά» της κάθε 15ετίας. Λαμβάνοντας επίσης υπόψη ότι πριν το 1900 ο μέγιστος αριθμός των σταθμών ανά τον κόσμο ήταν το 6,5% του μεγίστου αριθμού σταθμών του επόμενου αιώνα, η ανάλυση έγινε για τη χρονιά 1901 και μετέπειτα.

Στον Πίνακα 5.3 παρουσιάζεται ο αριθμός των σταθμών που είχαν τουλάχιστον 10 μετρήσεις ανά 15ετία και Σχήμα 5.20 φαίνεται και το ποσοστό των θετικών και αρνητικών κλίσεων ανά 15ετία.

Πίνακας 5.3 Σύνολο σταθμών με τουλάχιστον 10 μετρήσεις ανά 15ετία

ΠΕΡΙΟΔΟΣ ΜΕΤΡΗΣΗΣ	1901-1915	1916-1930	1931-1945	1946-1960	1961-1975	1976-1990
ΣΥΝΟΛΟ ΣΤΑΘΜΩΝ	22	58	88	111	109	92

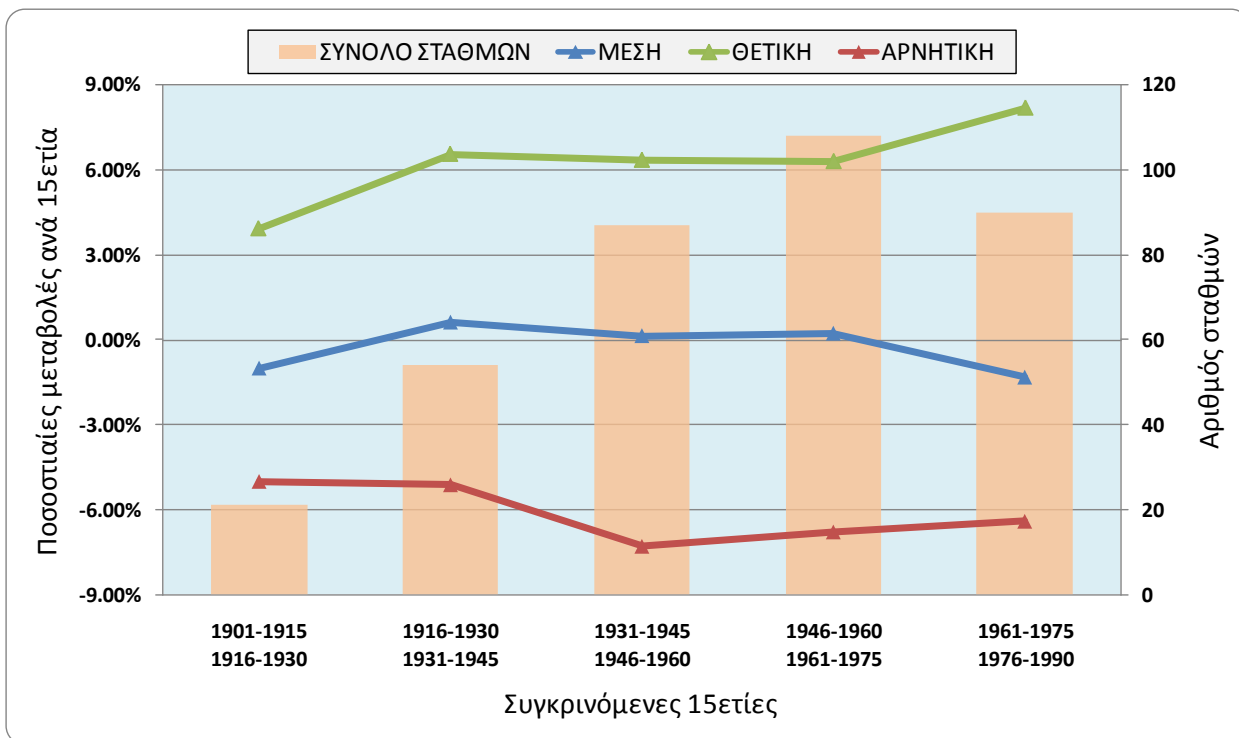


Σχήμα 5.20 Ποσοστά θετικών και αρνητικών τάσεων ανά 15ετία σύγκρισης

Με βάση το τελευταίο σχήμα, είναι αξιοσημείωτο πως ειδικά στην πιο πρόσφατη 15ετία, στο σύνολο των 92 ποταμών για τους οποίους υπολογίστηκε η κλίση της γραμμής τάσης, είναι λιγότεροι αυτοί με θετική κλίση παρά με αρνητική.

Στη συνέχεια της ανάλυσης, έγινε προσπάθεια να συγκριθούν το σύνολο των τάσεων σε μια 15ετία με αυτές της επόμενης. Πιο αναλυτικά, σε όποιο σταθμό ήταν διαθέσιμες οι τάσεις δύο διαδοχικών 15ετιών, υπολογίστηκε η κλίση αυτών των δύο σημείων. Αν δηλαδή σε έναν

σταθμό την 15ετία 1901-1915 έχει υπολογιστεί μια τάση της τάξεως του +1% και την επόμενη (1916-1930) είναι -2%, τότε ως αποτέλεσμα θα λάβουμε μια μεταβολή της τάξεως του -3%. Τα αποτελέσματα αυτής της διαδικασίας φαίνονται στο Σχήμα 5.21



Σχήμα 5.21 Ποσοστιαίες μεταβολές τάσεων μεταξύ δύο διαδοχικών 15ετιών και ο αριθμός των σταθμών που συγκρίνονται

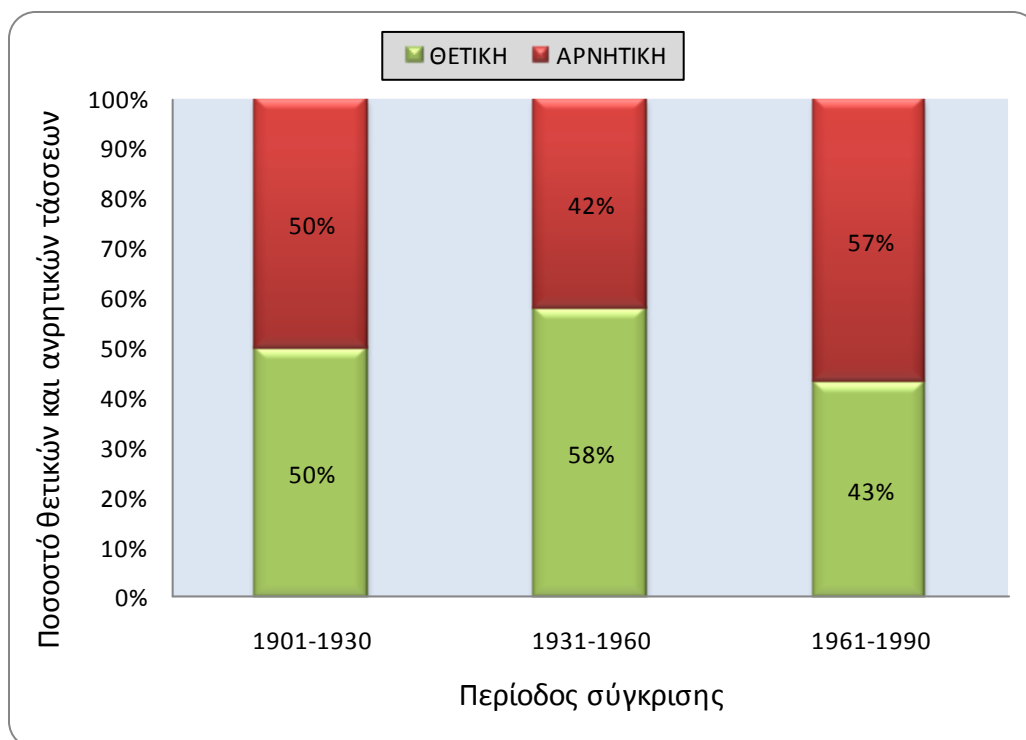
Ως «ΜΕΣΗ» έχει χαρακτηριστεί ο μέσος όρος των μεταβολών όλων των σταθμών, ως «ΘΕΤΙΚΗ» λήφθηκε ο μέσος όρος μόνο εκείνων των σταθμών που παρουσίαζαν μεταβολή στις τάσεις τους μεγαλύτερη ή ίση από +2.5% ενώ ως «ΑΡΝΗΤΙΚΗ» λήφθηκε ο μέσος όρος των σταθμών εκείνων με ποσοστιαία μεταβολή μικρότερη ή ίση από -2.5%. Όπως φαίνεται, δεν παρατηρείται κάποια αξιοσημείωτη αλλαγή στις τάσεις των ποταμών με τη μέθοδο που ακολουθήθηκε, πόσω μάλλον αν αναλογιστεί κανείς ότι το όριο του 2.5% που λήφθηκε για τον χαρακτηρισμό θετικών ή αρνητικών μεταβολών, ειδικά για ετήσιες παροχές αιχμής, είναι αρκετά συντηρητικό. Εδώ πρέπει να σημειωθεί ότι παρά την αύξηση των «θετικών»

μεταβολών μεταξύ των περιόδων 1961-1975 και 1976-1990 παρατηρείται μείωση στη «ΜΕΣΗ» μεταβολή. Αυτό συμβαίνει διότι στις δύο τελευταίες περιόδους σύγκρισης ('61 -'75 με '76 -'90) οι σταθμοί που εμπίπτουν στην κατηγορία «ΘΕΤΙΚΗ» είναι 16 ενώ αυτοί της «ΑΡΝΗΤΙΚΗΣ» είναι συνολικά 39 και είναι λογικό να επηρεάζουν πολύ περισσότερο τη «ΜΕΣΗ» τιμή του συνόλου.

Προσθετικά με την παραπάνω ανάλυση υπολογίστηκαν και οι τάσεις όλων των σταθμών ανά 30ετία, με το σκεπτικό ότι τα τριάντα χρόνια αποτελούν ακόμη πιο αξιόπιστη επιλογή για τον υπολογισμό τάσεων. Σε αυτή την περίπτωση τέθηκε ως κριτήριο να υπάρχουν τουλάχιστον 20 μετρήσεις «εσωτερικά» κάθε 30ετίας, σε αντιδιαστολή με το όριο των 10 μετρήσεων στην προηγούμενη περίπτωση. Τα αποτελέσματα φαίνονται στον Πίνακα 5.4 και στο Σχήμα 5.22.

Πίνακας 5.4 Αριθμός σταθμών ανά 30ετία σύγκρισης

ΠΕΡΙΟΔΟΣ ΜΕΤΡΗΣΗΣ	1901-1930	1931-1960	1961-1990
ΣΥΝΟΛΟ ΣΤΑΘΜΩΝ	22	100	104



Σχήμα 5.22 Ποσοστό θετικών και αρνητικών τάσεων ανά 30ετία σύγκρισης

Όπως και στην προηγούμενη περίπτωση, παρατηρείται και εδώ ότι στην πιο πρόσφατη περίοδο υπολογισμού των τάσεων, οι σταθμοί με αρνητικές τάσεις υπερτερούν αριθμητικά αυτών με θετικές στο σύνολο των 104 σταθμών.

5.6 Ανάλυση τάσεων με το Mann-Kendall trend test

Στην προηγούμενη περίπτωση, η διαδικασία υπολογισμού των τάσεων των χρονοσειρών και των ποσοσטיαίων μεταβολών τους μπορεί να θεωρηθεί μεροληπτική από τη στιγμή που οι τάσεις προκύπτουν από τα συγκεκριμένα «παράθυρα» 15ετίας και 30ετίας. Για το λόγο αυτό, η ανάλυση τάσεων στη συγκεκριμένη βάση δεδομένων προσεγγίστηκε και από μια διαφορετική σκοπιά, η οποία βασίστηκε στην εφαρμογή του Mann-Kendall trend test (Mann 1945, Kendall 1975). Το θεωρητικό υπόβαθρο γύρω από το συγκεκριμένο τεστ περιγράφεται αναλυτικότερα στο Κεφάλαιο 3. Ένα από τα βασικότερα πλεονεκτήματα του τεστ, πέρα από την μαθηματική του απλότητα, είναι ότι εξαρτάται αποκλειστικά από τις κλάσεις των καταγραφών μιας χρονοσειράς και ως συνέπεια το αποτέλεσμά του θεωρείται μία παράμετρος που δεν επηρεάζεται από την πραγματική κατανομή των δεδομένων.

Μία από τις βασικότερες, παρ' όλα αυτά, προϋποθέσεις για την εφαρμογή τεστ στατιστικού ελέγχου, όπως και το συγκεκριμένο, είναι η ανεξαρτησία των δεδομένων. Παρακάτω μελετάται η δομή συσχέτισης που παρουσιάζουν τα δεδομένα, με βάση το συντελεστή αυτοσυσχέτισης αλλά και το συντελεστή μακροπρόθεσμης εμμονής.

5.6.1 Έλεγχος δομής συσχέτισης δεδομένων

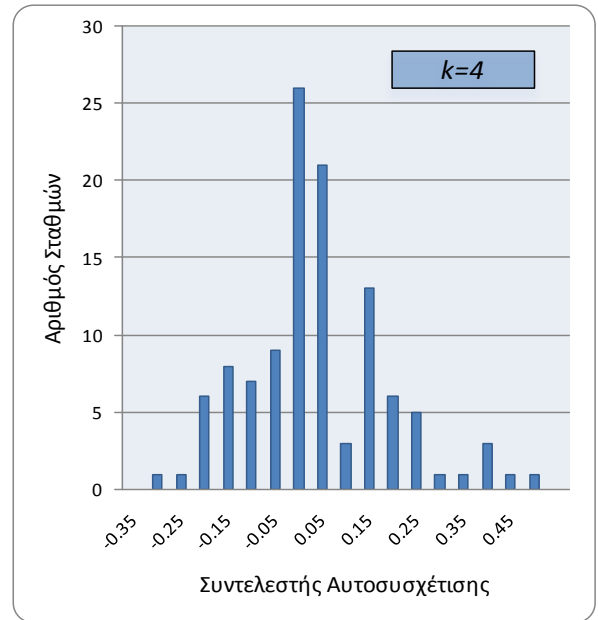
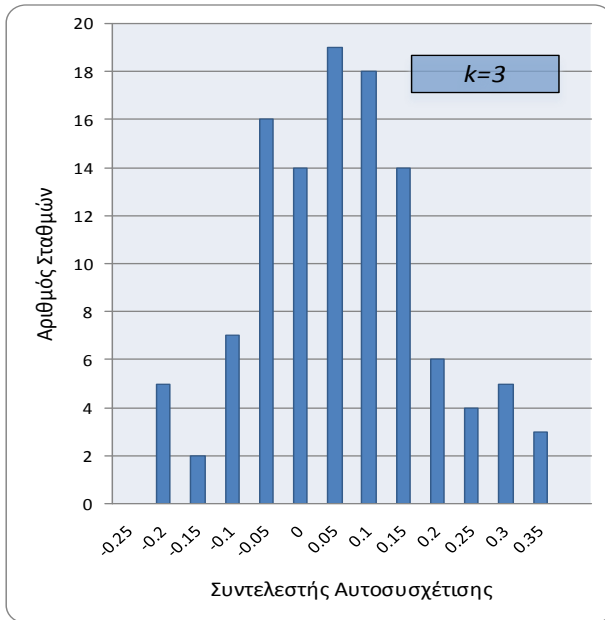
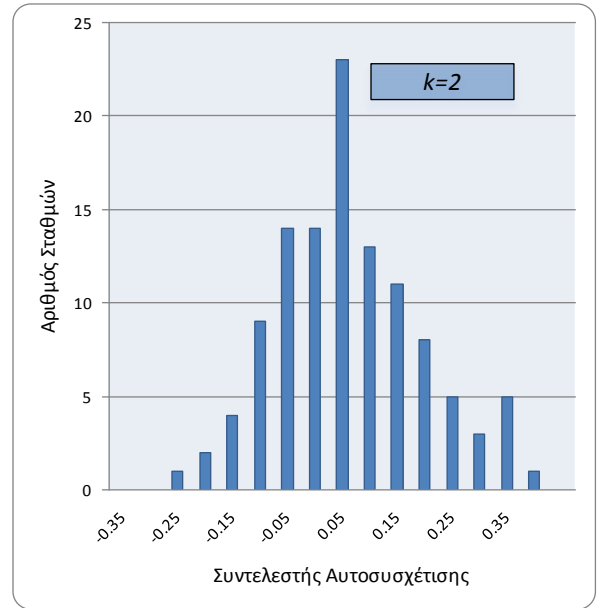
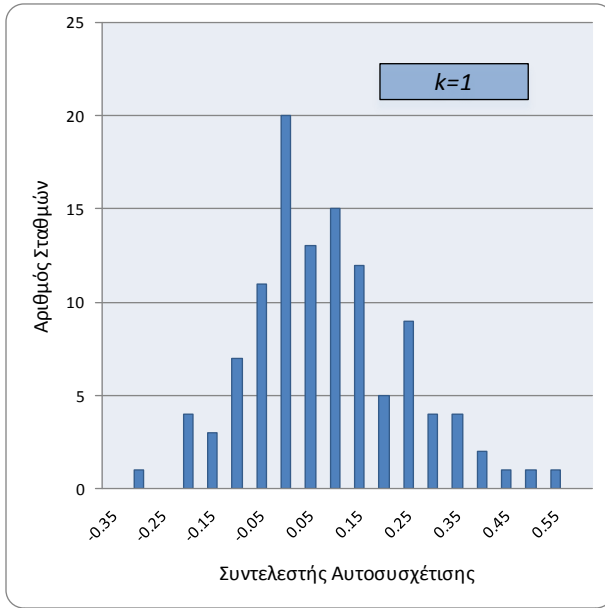
➤ Συντελεστής αυτοσυσχέτισης

Στον Πίνακα 5.5 φαίνονται τα βασικά στατιστικά χαρακτηριστικά του συντελεστή αυτοσυσχέτισης όλων των σταθμών για διάφορες τιμές της υστέρησης k . Παρατηρείται ότι για μεγαλύτερες τιμές της υστέρησης k , η μέση τιμή του συντελεστή αυτοσυσχέτισης προσεγγίζει το 0. Παρεμφερή συμπεριφορά ακολουθεί η διάμεσος. Η τυπική απόκλιση είναι περίπου η ίδια για τα τέσσερα μεγέθη αλλά και η μέγιστη τιμή που παρατηρείται ακόμα και για υστέρηση $k=1$ είναι 0.503.

Πίνακας 5.5 Στατιστικά χαρακτηριστικά των συντελεστών αυτοσυσχέτισης

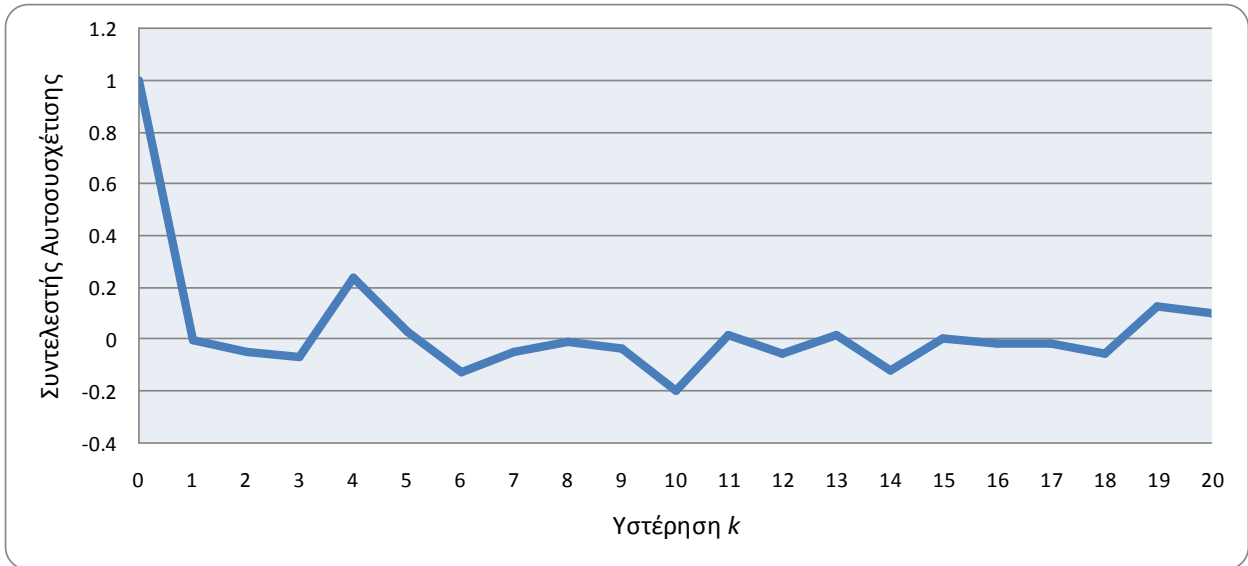
ΥΣΤΕΡΗΣΗ	ΜΕΣΗ ΤΙΜΗ	ΤΥΠΙΚΗ ΑΠΟΚΛΙΣΗ	ΔΙΑΜΕΣΟΣ	$Q_{2,5}$	Q_{75}	$Q_{97,5}$	ΕΛΑΧΙΣΤΗ ΤΙΜΗ	ΜΕΓΙΣΤΗ ΤΙΜΗ
1	0.059	0.156	0.042	-0.219	0.139	0.396	-0.318	0.503
2	0.040	0.135	0.035	-0.200	0.113	0.329	-0.296	0.393
3	0.036	0.124	0.036	-0.219	0.109	0.300	-0.238	0.331
4	0.016	0.151	-0.003	-0.232	0.115	0.369	-0.332	0.477

Μια ακόμα πιο ξεκάθαρη εικόνα μπορούμε να πάρουμε από το Σχήμα 5.23, όπου απεικονίζονται τα ιστογράμματα της κατανομής των συντελεστών αυτοσυσχέτισης για διάφορες τιμές της υστέρησης. Πιο συγκεκριμένα, 91 ποτάμια από τα 113 έχουν συντελεστή αυτοσυσχέτισης για υστέρηση 1 μικρότερο ή ίσο από 0.2 και μάλιστα το 95% των δεδομένων βρίσκεται μεταξύ του -0,2 και 0.396. Αξίζει πάντως να σημειωθεί ότι η εκτίμηση της αυτοσυσχέτισης ενέχει σημαντική αρνητική μεροληψία και άρα οι πραγματικές τιμές αναμένονται μεγαλύτερες από αυτές που εκτιμήθηκαν.



Σχήμα 5.23 Ιστογράμματα κατανομής συντελεστών αυτοσυσχετίσης για διάφορες τιμές της υστέρησης k

Στο Σχήμα 5.24 δίνεται ένα διάγραμμα του συντελεστή αυτοσυσχέτισης στον ποταμό Gudenaα στη Δανία συναρτήσει της υστέρησης k , όπου φαίνεται καθαρά η απότομη πτώση κοντά στο 0 της αυτοσυσχέτισης της χρονοσειράς για υστερήσεις μεγαλύτερες του 1.



Σχήμα 5.23 Διάγραμμα του συντελεστή αυτοσυσχέτισης συναρτήσει της υστέρησης k στον ποταμό Gudenaα στην Δανία

Με μία πρώτη εποπτική ματιά, φαίνεται ότι κανείς θα μπορούσε να υποθέσει πως τα δεδομένα μας δεν παρουσιάζουν κάποια ισχυρή δομή αυτοσυσχέτισης, κάτι που είναι και διαισθητικά λογικό στην περίπτωση ετήσιων παροχών αιχμής. Ακόμα όμως δεν παύουν να υπάρχουν ποτάμια που τουλάχιστον για τις πρώτες τιμές της κλάσης k διατηρούν σχετικά υψηλές τιμές αυτοσυσχέτισης.

➤ Ο συντελεστής μακροπρόθεσμης εμμονής

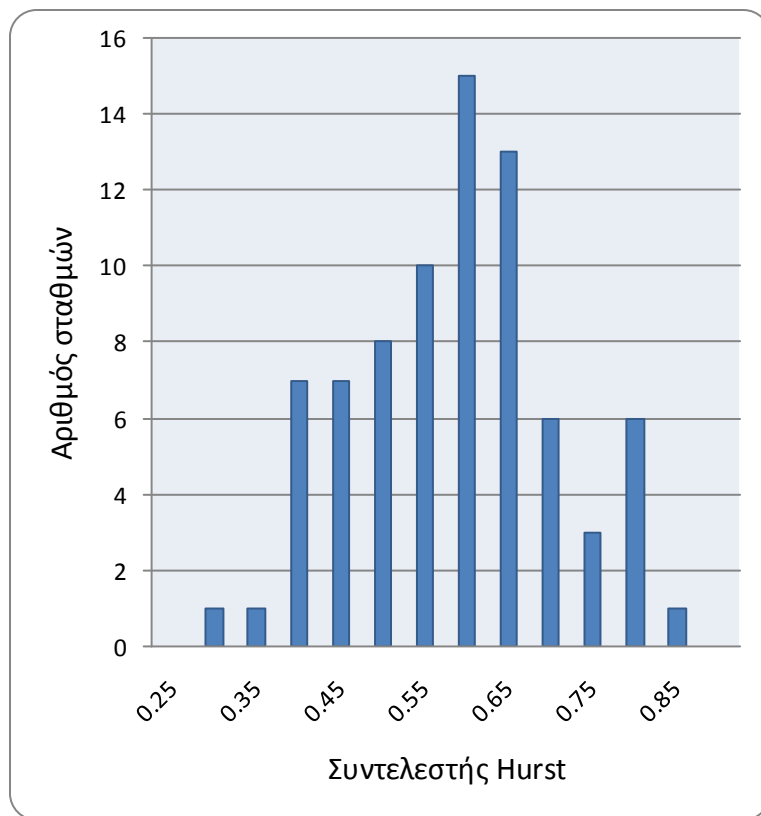
Μία διαφορετική προσέγγιση που μπορεί να περιγράψει τη δομή συσχέτισης των δεδομένων μας είναι ο υπολογισμός του συντελεστή μακροπρόθεσμης εμμονής (συντελεστής Hurst). Περισσότερες λεπτομέρειες για το φαινόμενο της μακροπρόθεσμης εμμονής αλλά και τους τρόπους υπολογισμού του συντελεστή Hurst γίνεται στο Κεφάλαιο 3. Στην προκειμένη περίπτωση, χρησιμοποιήθηκε η καμπύλη τους συναθροισμένης διασποράς (Koutsoyiannis, 2002) για τον υπολογισμό του H και έγινε εκ νέου επιλογή των κατάλληλων δεδομένων με βάση τις κενές τους τιμές (Slater, 2016) ως εξής :

- Το σύνολο των κενών τιμών της χρονοσειράς δεν θα πρέπει να είναι μεγαλύτερο από το 10% του συνόλου των καταγραφών
- Οι διαδοχικές κενές τιμές δεν θα πρέπει να ξεπερνάνε σε αριθμό τις 3
- Τροποποιήθηκαν χρονοσειρές, εξαλείφοντας τις τιμές εκείνες που εμφανίζονταν μετά ή πριν από μεγάλα κενά (π.χ. 10 διαδοχικές κενές καταχωρήσεις) αλλά το μήκος τους εξακολουθούσε να είναι πάνω από 50

Με τον τρόπο αυτό επιλέχθηκαν 78 ποταμοί από το σύνολο των 113 για τους οποίους υπολογίστηκε ο συντελεστής Hurst με τη μέθοδο της συναθροισμένης διασποράς και εν συνεχεία εφαρμόστηκε το Mann-Kendall trend test με ορισμένες τροποποιήσεις που θα αναλυθούν στη συνέχεια. Όσο αναφορά των συντελεστή Hurst, φαίνονται στον Πίνακα 5.6 κάποια στοιχεία του στατιστικού του προφίλ καθώς και στο σχήμα 5.24 απεικονίζεται το ιστόγραμμα της κατανομής του.

Πίνακας 5.6 Βασικά στατιστικά χαρακτηριστικά του συντελεστή Hurst

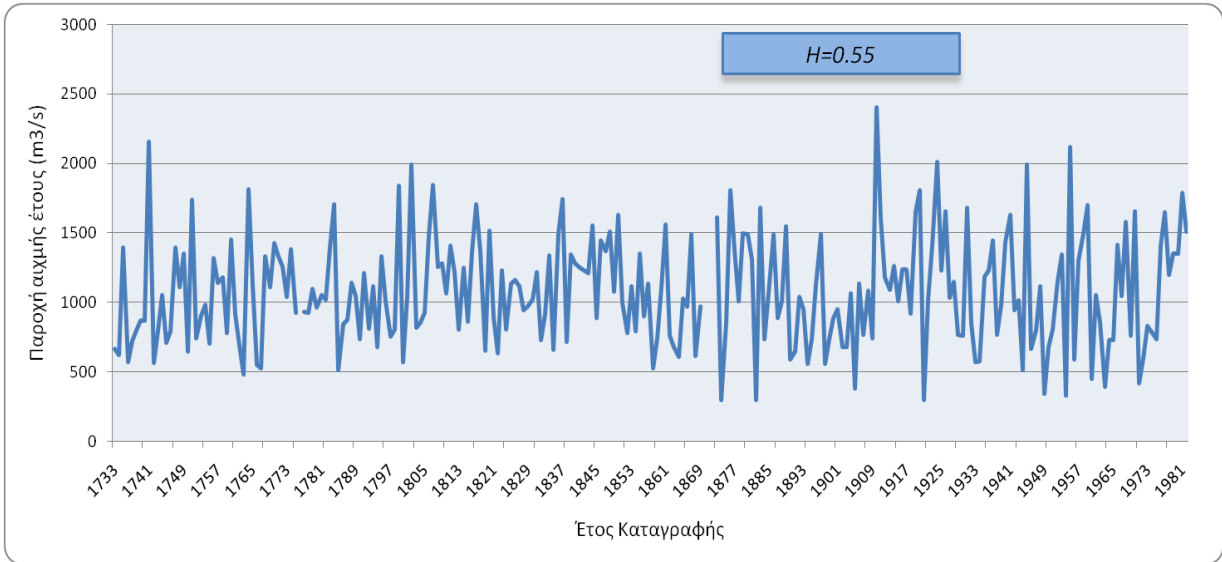
ΕΛΑΧΙΣΤΗ ΤΙΜΗ	Q _{2,5}	Q ₂₅	ΔΙΑΜΕΣΟΣ	Q ₇₅	Q _{97,5}	ΜΕΓΙΣΤΗ ΤΙΜΗ	ΜΕΣΗ ΤΙΜΗ	ΤΥΠΙΚΗ ΑΠΟΚΛΙΣΗ
0.26	0.35	0.48	0.57	0.62	0.79	0.84	0.56	0.12



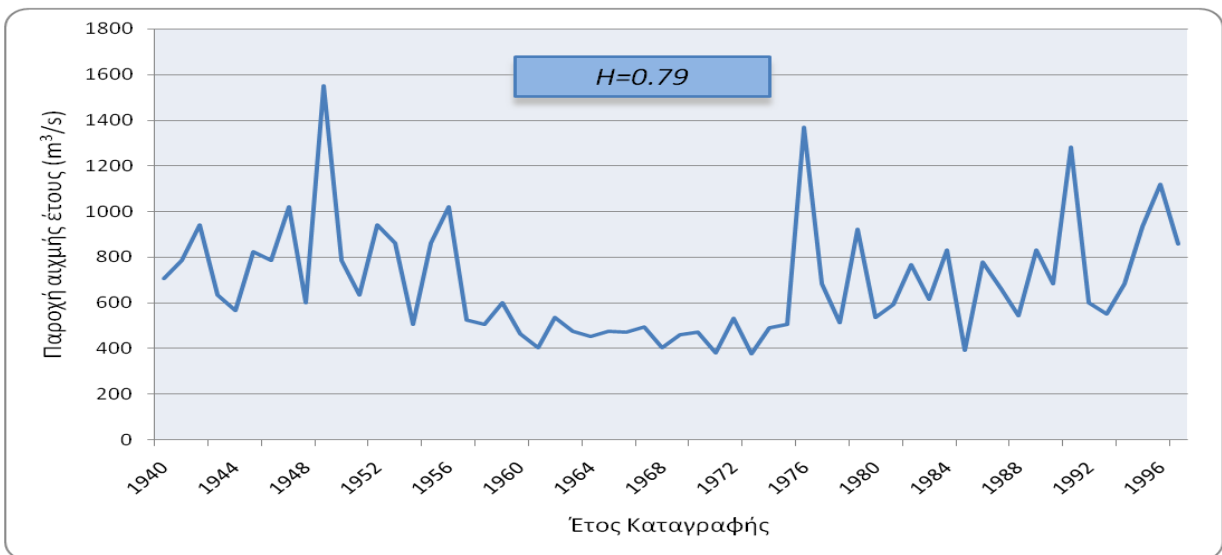
Σχήμα 5.24 Ιστόγραμμα κατανομής του H της μεθόδου συναθροισμένης διασποράς

Η μέση τιμή των συντελεστών Hurst του δείγματος είναι 0.56, η διάμεσος 0.57, η τυπική απόκλιση 0.12. Η ελάχιστη τιμή που παρατηρήθηκε είναι ίση με 0.26 ενώ η μέγιστη είναι 0.84. Όπως φαίνεται υπάρχουν ενδείξεις ύπαρξης μακροπρόθεσμης εμμονής. Μόνο όμως το 25% των σταθμών παρουσιάζει H μεγαλύτερο ή ίσο από 0.62, δηλαδή σημαντική δομή αυτοσυσχέτισης. Πολύ ισχυρή δομή αυτοσυσχέτισης, $H \geq 0.79$ παρουσιάζει το 2.5% των σταθμών. Για διάστημα εμπιστοσύνης 95% οι τιμές κυμαίνονται από 0.35-0.79, γεγονός που καταδεικνύει ότι υπάρχει σημαντική στατιστική αβεβαιότητα στην εκτίμηση. Ακόμα, παρατηρείται ότι ένα ποσοστό της τάξης του 25% παρουσιάζει χρονική ανεξαρτησία ως προς τα φαινόμενο της μακροπρόθεσμης εμμονής.

Στα Σχήματα 5.25 και 5.26 παρουσιάζονται ενδεικτικά διαγράμματα των χρονοσειρών χαρακτηριστικών σταθμών με χαμηλό και υψηλό συντελεστή Hurst αντίστοιχα.

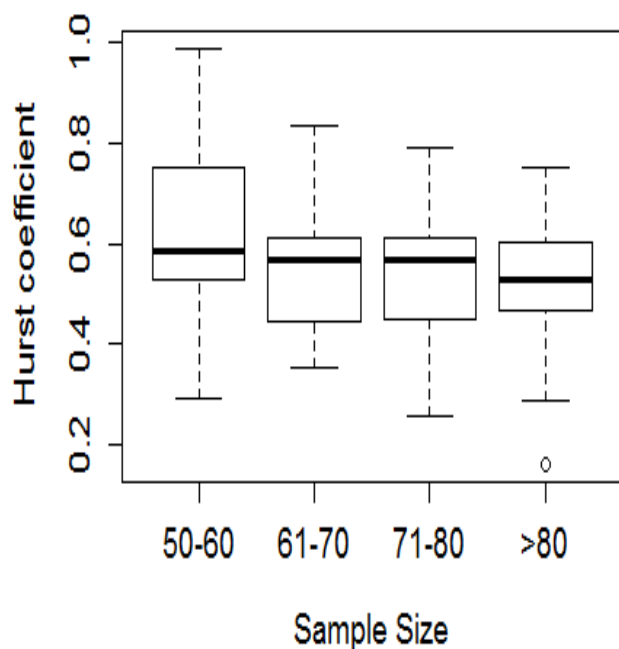


Σχήμα 5.25 Διάγραμμα χρονοσειράς παροχής αιχμής σταθμού με χαμηλό συντελεστή Hurst

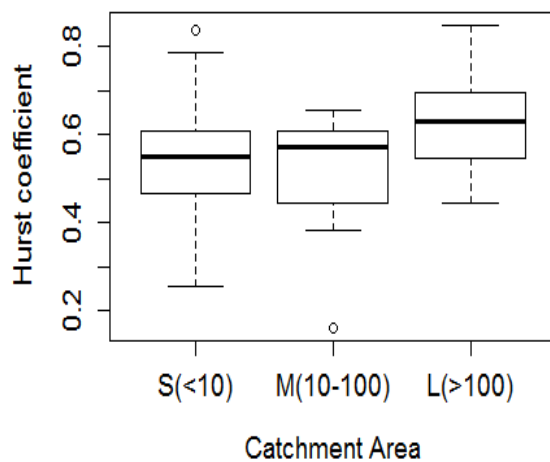


Σχήμα 5.26 Διάγραμμα χρονοσειράς παροχής αιχμής σταθμού με υψηλό συντελεστή Hurst

Εξετάστηκε ακόμα η ενδεχόμενη συσχέτιση του συντελεστή Hurst με το μήκος δείγματος αλλά και με τη λεκάνη απορροής. Τα αποτελέσματα παρουσιάζονται στα επόμενα δύο γραφήματα.



Σχήμα 5.27 Box-plots του συντελεστή Hurst για τα διάφορα μήκη δείγματος



Σχήμα 5.28 Box-plots του συντελεστή Hurst για διάφορες τιμές των λεκανών απορροής

Πιο συγκεκριμένα, για τα παραπάνω box-plots(Σχήμα 5.27 και 5.28), σημειώνεται ότι οι διάφορες λεκάνες απορροής χωρίστηκαν σε τρεις κατηγορίες ανάλογα με την έκτασή τους σε km^2 , όπου μικρές (S) χαρακτηρίστηκαν οι λεκάνες με έκταση μικρότερη ή ίση των $10 \cdot 10^3 \text{ km}^2$, ως μεσαίες (M) αυτές με έκταση μεταξύ $10 \cdot 10^3$ και $100 \cdot 10^3 \text{ km}^2$ και μεγάλες (L) όσες

είναι πάνω από $100 \cdot 10^3 \text{ km}^2$. Φαίνεται πως υπάρχει συσχέτιση του συντελεστή μακροπρόθεσμης εμμόνης με το μέγεθος της λεκάνης απορροής και μάλιστα φαίνεται ο συντελεστής να αυξάνεται για μεγαλύτερες λεκάνες απορροής σε έκταση (Hirra et al, 2010 , Szolgaγona et al., 2014). Από την άλλη μεριά το Σχήμα 5.28 δεν δείχνει να υπάρχει κάποια ισχυρή συσχέτιση με το μήκος δείγματος και εκτός αυτού παρατηρείται ότι για μικρότερα μήκη δείγματος (κατηγορία 50-60) η διασπορά του συντελεστή είναι αισθητά μεγαλύτερη.

5.6.2 Εφαρμογή του Mann- Kendall trend test

Σύμφωνα με τους Mann(1945) και Kendall(1975) το συγκεκριμένο τεστ βασίζεται στη συσχέτιση μεταξύ των κλάσεων μιας χρονοσειράς και την θέση τους μέσα στο χρόνο. Για μια χρονοσειρά $X=\{x_1, x_2, x_3, \dots, x_n\}$ η στατιστική παράμετρος S δίνεται από την σχέση (52):

$$S = \sum_{i < j} a_{ij} \quad (52)$$

Όπου

$$a_{ij} = \text{sign}(x_j - x_i) = \text{sign}(R_j - R_i) = \begin{cases} 1 & x_i < x_j \\ 0 & x_i = x_j \\ -1 & x_i > x_j \end{cases} \quad (53)$$

και τα R_i και R_j είναι οι κλάσεις των παρατηρήσεων x_i και x_j αντίστοιχα. Η μέση τιμή και η διασπορά της στατιστικής παραμέτρου S , υπό την προϋπόθεση ότι είναι ασυσχέτιστα τα δεδομένα, δίνονται από τις σχέσεις (54) και (55) αντιστοίχως.

$$E[S] = 0 \quad (54)$$

$$V_0[S] = n(n-1)(2n+5)/18 \quad (55)$$

Η στατιστική σημαντικότητα των υπολογισμένων τάσεων (S) μπορεί να ελεγχθεί συγκρίνοντας την πιθανότητα εμφάνισης της κανονικοποιημένης μεταβλητής u της εξίσωσης (56) με το επιθυμητό επίπεδο στατιστικής σημαντικότητας:

$$u = \begin{cases} (S-1)/\sqrt{V_0(S)} & S > 0 \\ 0 & S = 0 \\ (S+1)/\sqrt{V_0(S)} & S < 0 \end{cases} \quad (56)$$

Στο συγκεκριμένο τεστ στατιστικού ελέγχου, ως μηδενική υπόθεση λαμβάνεται «η μη ύπαρξη τάσης» με βάση όσα αναφέρονται και στο Κεφάλαιο 3. Επίσης όπως αναλύεται

επίσης στο Κεφάλαιο 3, τιμές του Hurst μεγαλύτερες από 0.5 μπορούν να επηρεάσουν τον υπολογισμό άλλων στατιστικών παραμέτρων. Πέρα από τον Koutsoyiannis (2003) που υποστήριξε ότι υπάρχει ανάγκη στην υδροκλιματική έρευνα για την εφαρμογή των κλασικών στατιστικών, που υποθέτουν ανεξάρτητα και με την ίδια κατανομή δεδομένα, να λαμβάνεται υπόψη το φαινόμενο της μακροπρόθεσμης εμμονής και άλλες έρευνες έχουν αναδείξει αυτή την αναγκαιότητα (Beran and Feng, 2002; Kantelhardt et al., 2002; Craigmile et al., 2004a, 2004b). Για να αναδειχθεί η επιρροή του φαινομένου της μακροπρόθεσμης εμμονής στην ανάλυση τάσεων, προσομοιώθηκαν FGN χρονοσειρές για διαφορετικό συντελεστή Hurst και εφαρμόστηκε σε κάθε περίπτωση το Mann-Kendall test. Ο Πίνακας 5.7 δείχνει το ποσοστό των δειγμάτων που απορρίφθηκε η μηδενική υπόθεση με βάση το Mann-Kendall test για επίπεδα σημαντικότητας 10% και 5% , σε 10,000 χρονοσειρές με μήκος δείγματος $n=100$ και τιμές του H από 0.5 έως και 0.9 ανά περίπτωση.

Πίνακας 5.7 Ποσοστό δειγμάτων που απορρίπτεται η μηδενική υπόθεση (Σφάλμα τύπου I) για συνθετικές FGN χρονοσειρές μήκους $n=100$ έτη και τιμές του H από 0.5 έως 0.9

	ΠΟΣΟΣΤΟ ΑΠΟΡΡΙΨΗΣ ΣΤΟ ΕΠΙΠΕΔΟ 10%	ΠΟΣΟΣΤΟ ΑΠΟΡΡΙΨΗΣ ΣΤΟ ΕΠΙΠΕΔΟ 5%
H=0.5	9.7	4.6
H=0.6	23.0	15.2
H=0.7	37.8	29.1
H=0.8	50.5	42.6
H=0.9	63.6	56.7

Είναι ξεκάθαρο ότι τιμές του H μεγαλύτερες του 0.5 έχουν ως αποτέλεσμα μεγαλύτερα ποσοστά απόρριψης της μηδενικής υπόθεσης από ότι θα περιμέναμε. Για παράδειγμα, για $H=0.7$ το ποσοστό απόρριψης (Σφάλμα τύπου I, βλ. Κεφάλαιο 3) είναι κοντά στο 40% για επίπεδο σημαντικότητας 10% (4 φορές πάνω από το αναμενόμενο) και 30% στο επίπεδο 5% (6 φορές πάνω από το αναμενόμενο). Για ακόμα μεγαλύτερες τιμές του συντελεστή Hurst, τα ποσοστά απόρριψης είναι ακόμα μεγαλύτερα, το οποίο θα έχει σαν αποτέλεσμα την

λανθασμένη αναγνώριση τάσεων ως στατιστικά σημαντικών όταν βέβαια αυτές δεν υπάρχουν.

Ο Hamed(2008), όπως έχει προαναφερθεί για την διόρθωση της αρνητικής επιρροής του φαινομένου της μακροπρόθεσμης εμμονής στις εκτιμήσεις του τεστ προτείνει την τροποποίησή του, έτσι ώστε να λαμβάνεται υπόψη και η συγκεκριμένη δομή συσχέτισης των εκάστοτε δεδομένων. Στην παρούσα εργασία και κατά αντιστοιχία με τον Koutsoyiannis (2002), για τα ποτάμια εκείνα που παρουσίασαν συντελεστή Hurst μεγαλύτερο ή ίσο από 0.6, προσομοιώθηκαν 1000 συνθετικές FGN χρονοσειρές διατηρώντας τον Hurst της εκάστοτε παρατηρημένης χρονοσειράς. Με τη βοήθεια των προσομοιωμένων χρονοσειρών υπολογίστηκε η νέα τιμή της διασποράς της στατιστικής παραμέτρου του Mann-Kendall test και αυτή η τιμή τελικά χρησιμοποιήθηκε στην εξίσωση (56) για τον υπολογισμό του u .

Στον Πίνακα 5.8 φαίνονται τα αποτελέσματα της ανάλυσης των τάσεων για τα 78 διαθέσιμα ποτάμια, εκ των οποίων τα 29 είχαν $H \geq 0.6$ και σε αυτά υπολογίστηκε με στοχαστικό τρόπο η τιμή της διασποράς της στατιστικής παραμέτρου S .

Πίνακας 5.8 Αποτελέσματα ανάλυσης τάσεων σύμφωνα με το Mann-Kendall trend test

ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΝΑΛΥΣΗΣ ΤΑΣΕΩΝ-MANN-KENDALL TEST			
	10%	5%	2.50%
ΘΕΤΙΚΕΣ	11	5	1
ΑΡΝΗΤΙΚΕΣ	22	13	7

Όπως ήταν αναμενόμενο οι στατιστικά σημαντικές τάσεις είναι αισθητά λιγότερες όσο το επίπεδο σημαντικότητας επιλέγεται πιο μικρό (πιο αυστηρό). Αξίζει να σημειωθεί ότι σε κάθε περίπτωση οι αρνητικές/μειωτικές τάσεις είναι περισσότερες από τις θετικές, όπως άλλωστε είχε δειχθεί στην ενότητα 4.5 για την πιο πρόσφατη χρονική περίοδο.

Στη συνέχεια στον Πίνακα 5.9 παρουσιάζονται τα αποτελέσματα της ανάλυσης τάσεων σε περίπτωση που δεν είχε ληφθεί υπόψη το φαινόμενο της μακροπρόθεσμης εμμονής και άρα δεν είχε τροποποιηθεί η τιμή της διασποράς στις περιπτώσεις αυτές.

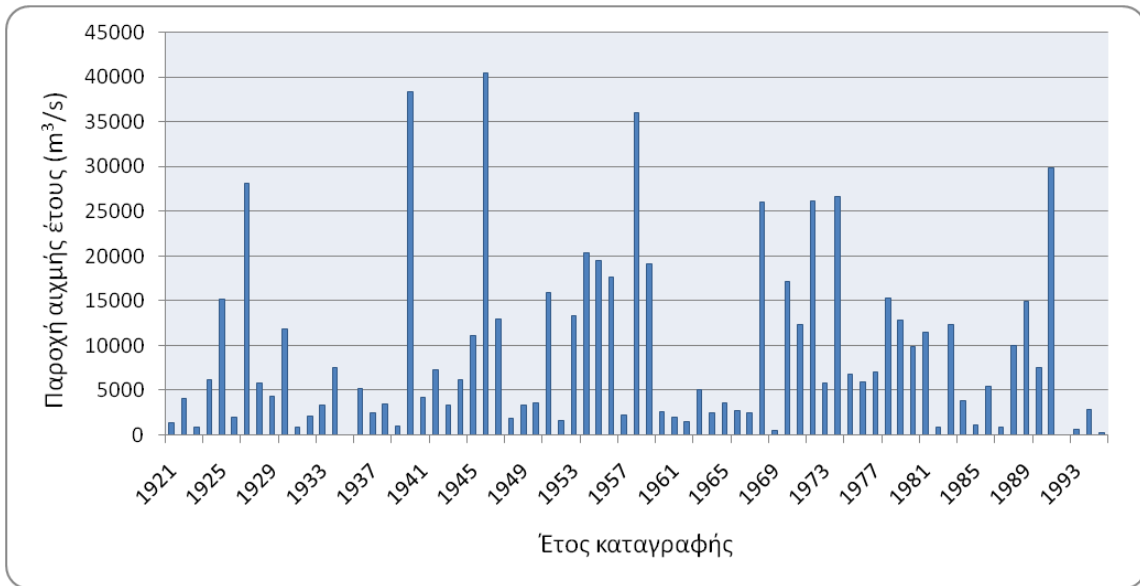
Πίνακας 5.9 Αποτελέσματα του Mann-Kendall trend test για διάφορα επίπεδα σημαντικότητας χωρίς να ληφθεί υπόψη η δομή συσχέτισης των δεδομένων

ΑΠΟΤΕΛΕΣΜΑΤΑ ΤΟΥ MANN-KENDALL TEST ΧΩΡΙΣ ΚΑΜΙΑ ΤΡΟΠΟΠΟΙΗΣΗ ΣΤΟΝ ΥΠΟΛΟΓΙΣΜΟ ΤΗΣ ΔΙΑΣΠΟΡΑΣ			
	10%	5%	2.50%
ΘΕΤΙΚΕΣ	11	9	4
ΑΡΝΗΤΙΚΕΣ	23	13	10

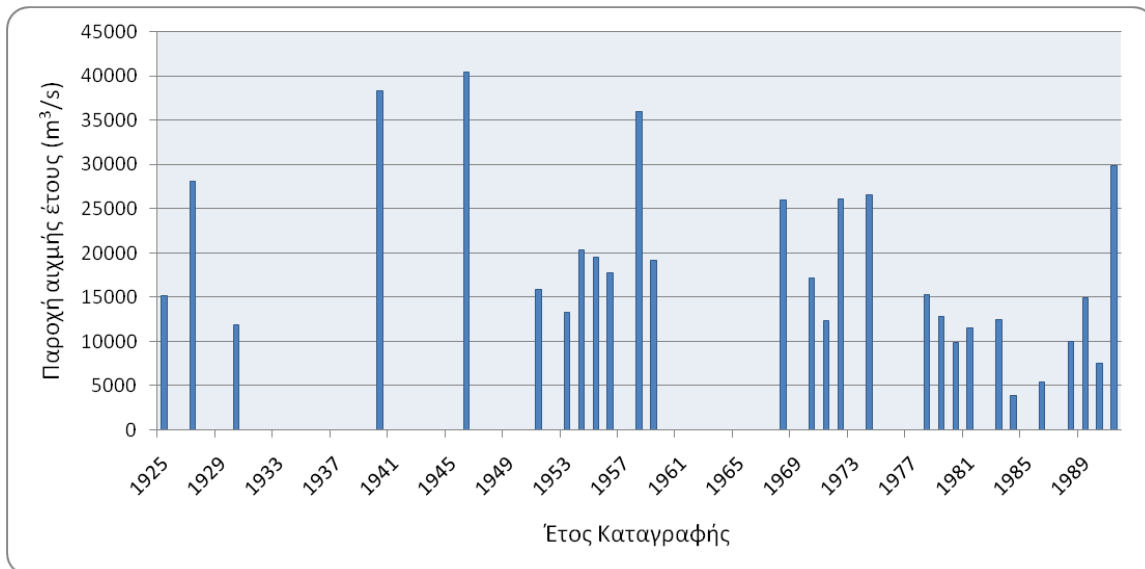
Είναι σαφές πως οι φαινομενικά στατιστικά σημαντικές τάσεις είναι αρκετά περισσότερες σε αριθμό σε αυτήν την περίπτωση. Μάλιστα, είναι άξιο αναφοράς το γεγονός πως η διαφοροποίηση είναι εντονότερη για επίπεδο σημαντικότητας 2.5%.

5.6.3 Λιγότερο έντονες πλημμύρες, αλλά ενδεχομένως πιο συχνές

Με βάση την παραπάνω ανάλυση τάσεων στη συγκεκριμένη βάση δεδομένων, γίνεται σαφές ότι όχι μόνο δεν επιβεβαιώνεται η ύπαρξη ενός παγκόσμιου κανόνα για αυξητικές τάσεις στις παροχές αιχμής των ποταμών αλλά και το γεγονός ότι οι αρνητικές τάσεις που ορίζονται ως στατιστικά σημαντικές είναι περισσότερες από τις αυξητικές. Μία άλλη παράμετρος που εξετάστηκε πέρα από το μέγεθος των πλημμυρικών παροχών, είναι και η συχνότητα με την οποία εμφανίζονται (Hirsch and Archfield, 2015). Μία από τις πιο ευρέως διαδεδομένες μεθοδολογίες για την μελέτη της συχνότητας εμφάνισης ακραίων φαινομένων σε παρατηρημένες χρονοσειρές είναι η επεξεργασία των χρονοσειρών με τέτοιο τρόπο ώστε να κατακρατούνται οι τιμές εκείνες που είναι πάνω από ένα όριο που έχει επιλέξει ο μελετητής (POT- Peak Over Threshold). Σε αυτή την περίπτωση επιλέχθηκε ως όριο η μέση τιμή της κάθε δεκαετίας. Οι παροχές που ήταν μεγαλύτερες από αυτή την τιμή σε κάθε δεκαετία παρέμειναν στην χρονοσειρά ενώ οι υπόλοιπες εξαλείφθηκαν. Στα σχήματα 5.29 και 5.30 φαίνονται δύο γραφήματα μιας χρονοσειράς πριν και μετά την συγκεκριμένη ανάλυση.



Σχήμα 5.29 Διάγραμμα χρονοσειράς με τη μορφή ιστογράμματος πριν την POT επεξεργασία.



Σχήμα 5.30 Διάγραμμα χρονοσειράς του σχήματος 5.30 με τη μορφή ιστογράμματος μετά την POT επεξεργασία.

Στη συγκεκριμένη περίπτωση μάλιστα παρατηρείται ότι τα πλημμυρικά επεισόδια που ξεπέρασαν το όριο που τέθηκε είναι περισσότερα στις πιο πρόσφατες χρονιές.

Μετρήθηκε ότι από το 1955 και παρελθοντικά, συνολικά υπάρχουν 4231 μετρήσεις ενώ από το 1956 και έπειτα οι μετρήσεις είναι 4327. Μετά την εφαρμογή της παραπάνω μεθοδολογίας σε όλες τις χρονοσειρές, το σύνολο των τιμών που κατακρατήθηκαν για την περίοδο πριν το 1955 είναι 1799 ενώ μετά το 1956 είναι 1529. Συγκεντρωτικά τα αποτελέσματα φαίνονται στον Πίνακα (5.10), όπου φαίνεται πως δεν υπάρχει κάποια ένδειξη για πιο συχνά πλημμυρικά επεισόδια την πιο πρόσφατη περίοδο

Πίνακας 5.10 Συγκεντρωτικά αποτελέσματα της ΡΟΤ ανάλυσης

	ΣΥΝΟΛΟ ΜΕΤΡΗΣΕΩΝ	ΣΥΝΟΛΟ ΚΑΤΑΚΡΑΤΗΜΕΝΩΝ ΜΕΤΡΗΣΕΩΝ	ΠΟΣΟΣΤΟ ΚΑΤΑΚΡΑΤΗΣΗΣ
ΜΕΤΡΗΣΕΙΣ ΠΡΙΝ ΤΟ 1955	4231	1799	43%
ΜΕΤΡΗΣΕΙΣ ΜΕΤΑ ΤΟ 1956	4327	1529	35%

5.7 Προσαρμογή Επιλεγμένων Συναρτήσεων Κατανομής

Ένα σημαντικό πρόβλημα στη διαχείριση των υδατικών πόρων αποτελεί η εκτίμηση του μεγέθους ενός πλημμυρικού επεισοδίου για μία συγκεκριμένη περίοδο επαναφοράς. Κατά την ανάλυση συχνότητας πλημμυρών συχνά ο μελετητής υποθέτει μία συνάρτηση κατανομής πιθανότητας, η οποία προσαρμόζεται στα διαθέσιμα δεδομένα με σκοπό την εκτίμηση του μεγέθους πλημμυρικών επεισοδίων που αντιστοιχούν σε περιόδους επαναφοράς πολύ μεγαλύτερες από το συνολικό μήκος δείγματος. Η επιλογή της κατανομής προς προσαρμογή εμπίπτει συνήθως στα υποκειμενικά κριτήρια του μελετητή ανάμεσα στις γνωστές κατανομές πιθανότητας που θεωρούνται οι πλέον κατάλληλες για δεδομένα με θετική ασυμμετρία (όπως η περίπτωση ετήσιων πλημμυρικών παροχών). Σε συνέχεια των όσων αναφέρονται για την προσαρμογή συναρτήσεων κατανομής στο Κεφάλαιο 4, στην ενότητα αυτή επιλέγονται τέσσερις κατανομές οι οποίες προσαρμόζονται στα δεδομένα μας. Συγκεκριμένα οι κατανομές που επιλέχθηκαν προς προσαρμογή είναι :

- Λογαριθμοκανονική κατανομή (LN)
- Κατανομή γάμμα (GAM)
- Κατανομή Log-Pearson III (LP3)
- Γενικευμένη Κατανομή Ακραίων Τιμών (GEV)

Οι συγκεκριμένες επιλογές των κατανομών βασίστηκαν από τη μία στη βιβλιογραφία που τις αναφέρει στην οικογένεια των κατανομών που είναι οι πλέον κατάλληλες για προσαρμογή σε δεδομένα ακραίων τιμών αλλά από την άλλη και με σκοπό την σύγκριση δι-παραμετρικών (οι δύο πρώτες κατανομές) και τρι-παραμετρικών (οι δύο επόμενες κατανομές) κατανομών ως προς την καταλληλότητα της προσαρμογής τους.

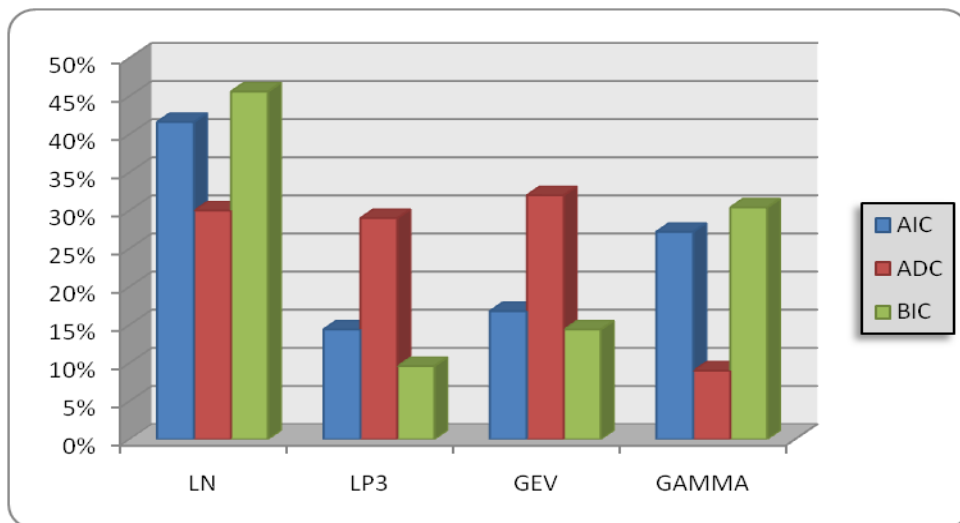
Όσο αναφορά, τα τεστ που πραγματοποιήθηκαν για την καταλληλότητα προσαρμογής των κατανομών αυτά είναι τα Akaike Information Criterion (AIC), το Bayesian Information

Criterion (BIC) και μια παραλλαγή του Anderson-Darling κριτηρίου (ADC). Τα συγκεκριμένα κριτήρια προτιμήθηκαν σε σχέση με τα κλασικά και πιο ευρέως διαδεδομένα τεστ καταλληλότητας προσαρμογής, διότι υπερτερούν ως προς την *επιλογή* και τη *διάκριση* μιας συνάρτησης κατανομής έναντι κάποιας άλλης που εξετάζεται.

Είναι ιδιαίτερα σημαντικό να σημειωθεί πως ο σκοπός της διαδικασίας αυτής είναι η αναγνώριση της θεωρητικής εκείνης κατανομής που βρίσκεται *πιο κοντά* στην κατανομή που αναπαρήγαγε στην πραγματικότητα τα δεδομένα (η λεγόμενη μητρική κατανομή). Αδιαμφισβήτητα, κανείς δεν πρέπει να κάνει την παρανόηση ότι τα δεδομένα παρήχθησαν κατά αυτόν τον τρόπο: *Το μοντέλο που επιλέγεται, αποτελεί μια βολική αναπαράσταση μιας πολύ πιο πολύπλοκης φυσικής διεργασίας.*

Οι παράμετροι για την προσαρμογή της κάθε κατανομής υπολογίστηκαν με τη Μέθοδο της Μέγιστης Πιθανοφάνειας. Γίνεται εκτενέστερη αναφορά στις διάφορες μεθόδους που μπορούν να εφαρμοστούν για τον υπολογισμό των παραμέτρων μιας συνάρτησης κατανομής στο Κεφάλαιο 4.

Στη συνέχεια ελέγχθηκε η καταλληλότητα προσαρμογής της κάθε κατανομής έναντι των υπολοίπων από κάθε κριτήριο ξεχωριστά. Τα συγκεντρωτικά αποτελέσματα φαίνονται στο Σχήμα 5.31, όπου απεικονίζονται σε μορφή ιστογράμματος το ποσοστό των φορών που η κάθε κατανομή επιλέχτηκε ως καταλληλότερη από το κάθε κριτήριο.



Σχήμα 5.31 Ποσοστό επιλογών του κάθε κριτηρίου ανάμεσα στις τέσσερις εξεταζόμενες κατανομές

Είναι αξιοσημείωτο ότι όλες οι κατανομές επιλέγονται σε αρκετά ικανοποιητικό ποσοστό ως καταλληλότερες από τουλάχιστον ένα κριτήριο. Από την άλλη μεριά είναι εμφανής η «τάση» των κριτηρίων AIC και BIC προς την επιλογή κατανομών με τις λιγότερες παραμέτρους.

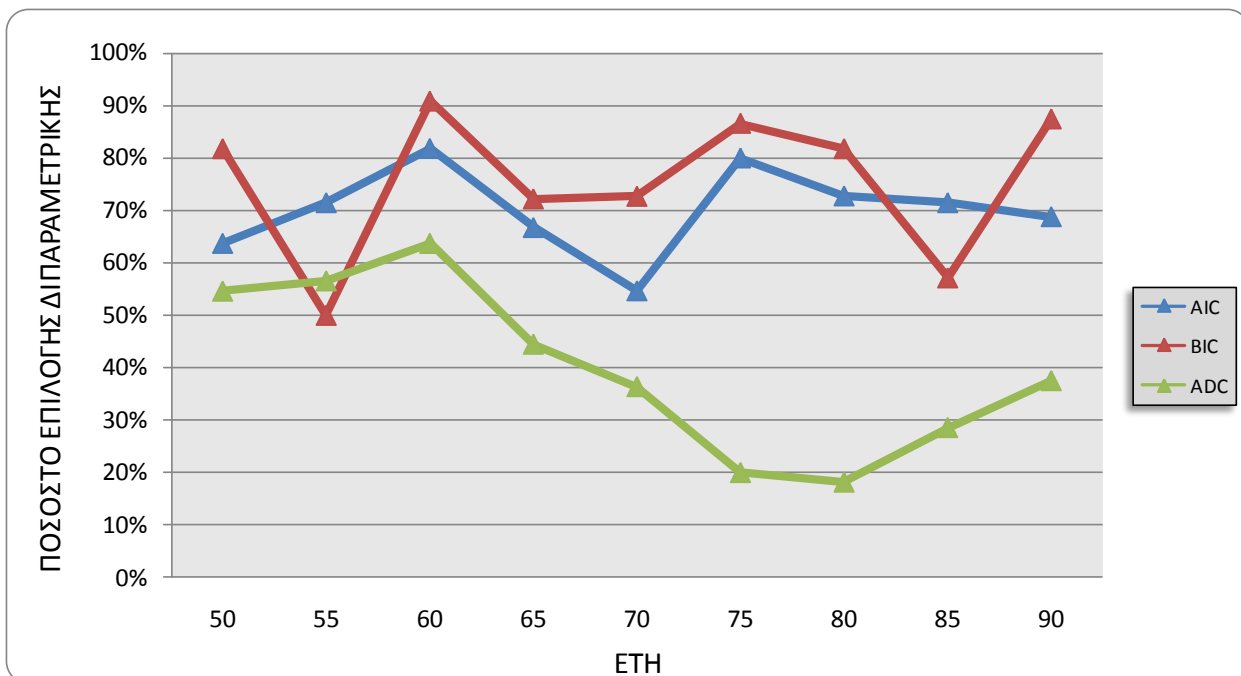
Ενδιαφέρον παρουσιάζει και η σύγκριση των κριτηρίων μεταξύ τους ως προς το είδος της επιλογής τους. Στον πίνακα 5.11 φαίνονται τα ποσοστά των φορών εκείνων όπου δύο κριτήρια κάνουν την ίδια επιλογή για τον ίδιο ποταμό. Τα AIC και BIC στο 93% των περιπτώσεων κάνουν ακριβώς την ίδια επιλογή, γεγονός που είναι λογικό κρίνοντας και από τις εξισώσεις (47) και (48). Επίσης στις μισές περιπτώσεις το κριτήριο ADC επιλέγει ως καταλληλότερα προσαρμοσμένη κατανομή διαφορετική από αυτή που επιλέγουν τα άλλα δύο κριτήρια.

Πίνακας 5.11 Ποσοστά φορών που δύο κριτήρια επιλέγουν ως καταλληλότερη την ίδια κατανομή

ΣΥΓΚΡΙΣΗ ΕΠΙΛΟΓΩΝ		
AIC-BIC	AIC-ADC	ADC-BIC
93%	50%	44%

Το στοιχείο αυτό αναδεικνύει τις δυσκολίες που εγκυμονούν ακόμα και στην επιλογή του κριτηρίου του ίδιου που θα κρίνει την καταλληλότητα της προσαρμογής μιας κατανομής.

Ενδιαφέροντα συμπεράσματα μπορούν να εξαχθούν αν μελετήσει κανείς την επιλογή των κριτηρίων ως προς το πλήθος των παραμέτρων των κατανομών που επέλεξαν αλλά και την ενδεχόμενη συσχέτιση της επιλογής αυτής με το μήκος δείγματος. Για το σκοπό αυτό διαιρέθηκε το δείγμα μας σε εννέα κλάσεις, με βάση το μήκος του. Τα όρια των κλάσεων είναι τα {50,56,61,66,71,76,81,86,91}. Στο Σχήμα 5.32 απεικονίζονται οι φορές επί τοις εκατό σε κάθε κλάση που κάποιο κριτήριο επέλεξε διπαραμετρική κατανομή (δηλαδή ή την λογαριθμική ή την γάμμα). Είναι φανερό η παρεμφερής συμπεριφορά των AIC και BIC και μάλιστα επιλέγουν στα ίδια περίπου ποσοστά για όλα τα μήκη δείγματος κυρίως τις διπαραμετρικές κατανομές. Από την άλλη μεριά, το ADC επιλέγει εν γένει σε μικρότερα ποσοστά κάποια από τις διπαραμετρικές κατανομές, αλλά εδώ φαίνεται ότι όσο το μήκος δείγματος μεγαλώνει οι διπαραμετρικές κατανομές επιλέγονται ακόμα πιο περιορισμένα. (Laio et al. 2009)



Σχήμα 5.32 Ποσοστά επιλογής μιας δι-παραμετρικής κατανομής από κάθε κριτήριο συναρτήσεως του μήκους δείγματος

Η προσαρμογή των συναρτήσεων κατανομής που επιλέχθηκαν, έδειξε ότι οι συγκεκριμένες κατανομές είναι ικανοποιητικά κατάλληλες για την μοντελοποίηση ακραίων τιμών. Επίσης εξίσου σημαντικό είναι και το στοιχείο ότι η λογαριθμοκανονική κατανομή επιλέγεται από όλα τα κριτήρια σε αρκετά μεγάλα ποσοστά. Σημειώνεται ακόμα ότι το κριτήριο BIC είναι αυτό που στις περισσότερες των περιπτώσεων αναδεικνύει ως καταλληλότερα προσαρμοσμένη κατανομή κάποια με δύο παραμέτρους παρά με τρεις (αρχή της φειδωλότητας).

Παρά τα πολύ συγκεκριμένα αποτελέσματα των παραπάνω μεθόδων στην προσαρμογή κατανομών, το ποιο κριτήριο πρέπει να υιοθετηθεί για τον έλεγχο καταλληλότητας προσαρμογής κατανομών σε πραγματικά υδρολογικά δεδομένα είναι ένα ερώτημα που παραμένει αναπάντητο. Κατά τον γράφοντα, μια αρκετά λειτουργική στρατηγική θα ήταν η

χρήση του AIC (ή του BIC, αφού παρέχει παρεμφερή πληροφορία) σε συνδυασμό με το ADC. Στην περίπτωση που τα δύο κριτήρια κάνουν την ίδια επιλογή, η συγκεκριμένη κατανομή μπορεί να θεωρηθεί κατάλληλα προσαρμοσμένη και ότι αναπαριστά ικανοποιητικά την φυσική πραγματικότητα. Τέλος, ο συνδυασμός των παραπάνω κριτηρίων σε σχέση με τα κλασικά τεστ καταλληλότητας της προσαρμογής θα μπορούσε να είναι ένα ισχυρό εργαλείο. Στην περίπτωση που μια κατανομή επιλεγμένη από κάποιο από τα παραπάνω κριτήρια γίνεται αποδεκτή σε κάποιο επίπεδο σημαντικότητας από ένα τεστ στατιστικού ελέγχου, ο μελετητής μπορεί να θεωρήσει την κατανομή αυτή ως κατάλληλα προσαρμοσμένη.

6 ΣΥΜΠΕΡΑΣΜΑΤΑ

6.1 Σύνοψη μελέτης και συμπεράσματα

Σκοπός της παρούσας εργασίας ήταν η στατιστική ανάλυση και επεξεργασία μιας βάσης δεδομένων πλημμυρικών παροχών σε παγκόσμια κλίμακα. Δόθηκε ιδιαίτερη βαρύτητα στην διερεύνηση και στην ανάλυση τάσεων στις παροχές αιχμής της συγκεκριμένης βάσης δεδομένων αλλά και προσαρμόστηκαν τέσσερις διαφορετικές κατανομές, ενώ επίσης ελέγχθηκε και με διάφορους τρόπους η καταλληλότητα της προσαρμογής των κατανομών αυτών. Χρησιμοποιήθηκαν δεδομένα από τον *Παγκόσμιο Κατάλογο Μεγίστων Παρατηρημένων Πλημμυρών (World Catalogue of Maximum Observed Floods, 2003)* και συγκεκριμένα 125 σταθμοί που πληρούσαν ορισμένα κριτήρια ποιότητας και πληρότητας καταγραφών. Το δείγμα των σταθμών αναλύθηκε αρχικά ως προς τα στατιστικά χαρακτηριστικά της παροχής αιχμής (μέση τιμή, διάμεσος, τυπική απόκλιση, ασυμμετρία, ακραίες τιμές) και παρατηρήθηκε η μεγάλη διασπορά του φαινομένου σε παγκόσμια κλίμακα αλλά και η θετική ασυμμετρία των διαθέσιμων χρονοσειρών. Τα αποτελέσματα της ανάλυσης αυτής παρουσιάζονται σε διαγράμματα και πίνακες στο κεφάλαιο 5.

Για την εξασφάλιση της ποιότητας και της αξιοπιστίας της διαδικασίας ανάλυσης τάσεων αποφασίστηκε η αναζήτηση ανθρώπινων παρεμβάσεων ανάντη των σταθμών μέτρησης και πιο συγκεκριμένα, η αναζήτηση μεγάλων ταμιευτήρων που θα μπορούσαν να έχουν επηρεάσει τη δίαιτα των ποταμών. Μέσω αυτής της διαδικασίας απορρίφθηκαν συνολικά 12 σταθμοί. Η αναζήτηση τάσεων στους εναπομείναντες ποταμούς προσεγγίστηκε στη συνέχεια με δύο διαφορετικές μεθόδους. Η πρώτη βασίστηκε ουσιαστικά στον υπολογισμό των κλίσεων των ευθειών που προσαρμόστηκαν σε όλες τις χρονοσειρές με τη μέθοδο της γραμμικής παλινδρόμησης. Η ανάλυση του συνόλου των σταθμών με αυτή τη μέθοδο έδειξε ότι δεν υπάρχουν ισχυρές ενδείξεις για σημαντικές μεταβολές στις παροχές αιχμής των ποταμών αλλά και μάλιστα ότι οι ποταμοί που μπορεί να έχουν μειωτικές τάσεις είναι

περισσότεροι σε αριθμό από ότι οι σταθμοί με αυξητικές τάσεις. Η δεύτερη μέθοδος βασίστηκε στο Mann-Kendall trend test όπου για την κατάλληλη εφαρμογή του δόθηκε ιδιαίτερη βαρύτητα στο δομή συσχέτισης των δεδομένων. Η μελέτη αυτή με γνώμονα τον συντελεστή συσχέτισης έδειξε πως οι παροχές αιχμής των ποταμών δεν παρουσιάζουν κάποια ισχυρή δομή συσχέτισης σε παγκόσμιο επίπεδο. Ακόμα, προσεγγίστηκε η συσχέτιση των δεδομένων και υπό το πρίσμα της μακροπρόθεσμης εμμονής. Ο συντελεστής Hurst υπολογίστηκε με τη μέθοδο του κλιμακογράμματος (συναθροισμένης διασποράς) αφού πρώτα επιλέχθηκαν εκ νέου οι κατάλληλοι σταθμοί με βάση τη δομή των κενών των χρονοσειρών. Ο υπολογισμός τους συντελεστή Hurst για τους 78 κατάλληλους σταθμούς που απέμειναν, υπέδειξε ότι σε 29 σταθμούς υπάρχουν ενδείξεις μακροπρόθεσμης εμμονής ($H > 0.6$). Παρόλα αυτά, η κατανομή του συντελεστή στο σύνολο των δεδομένων είναι τέτοια ώστε μόνο το 25% των δεδομένων να παρουσιάζει $H > 0.62$ και ακόμα μικρότερο ποσοστό (2.5%) να παρουσιάζει $H > 0.79$. Στη συνέχεια μέσω του λογισμικού R, εφαρμόστηκε χωρίς καμία τροποποίηση το Mann-Kendall trend test στα δεδομένα εκείνα που θεωρήθηκαν ασυσχέτιστα ($H < 0.6$) αλλά και σε εκείνα με κάποια δομή συσχέτισης, με τη μόνη διαφοροποίηση ότι η διασπορά της στατιστικής παραμέτρου του τεστ υπολογίστηκε με στοχαστικό τρόπο διατηρώντας σε κάθε περίπτωση το συντελεστή Hurst της παρατηρημένης χρονοσειράς. Τα τελικά αποτελέσματα για τα διάφορα επίπεδα σημαντικότητας υπέδειξαν την ύπαρξη στατιστικά σημαντικών τάσεων σε περιορισμένο αριθμό σταθμών και μάλιστα σε κάθε περίπτωση και πάλι (όπως και με την πρώτη μέθοδο), αυτοί που παρουσίασαν μειωτικές τάσεις ήταν περισσότεροι σε σύγκριση με αυτούς με μειωτικές τάσεις.

Σχετικά με την προσαρμογή κατανομών στα δεδομένα επιλέχθηκαν τέσσερις διαφορετικές κατανομές, η λογαριθμοκανονική κατανομή, η κατανομή γάμμα, η κατανομή log-Pearson III και η Γενικευμένη Ακραίων Τιμών. Βασικά κριτήρια για την επιλογή προς προσαρμογή αυτών των κατανομών ήταν η καταλληλότητα τους βάσει της διαθέσιμης βιβλιογραφίας και

από την άλλη η δυνατότητα σύγκρισης δι-παραμετρικών και τρι-παραμετρικών κατανομών. Η μέθοδος που υιοθετήθηκε για τον υπολογισμό των παραμέτρων των κατανομών ήταν η Μέθοδος της Μέγιστης Πιθανοφάνειας, η οποία αναλύεται στο Κεφάλαιο 4. Τέλος, ελέγχθηκε η καταλληλότητα της προσαρμογής αυτών των κατανομών από τρία διαφορετικά κριτήρια, το Κριτήριο Πληροφορίας Akaike (AIC), το Μπεϋζιανό Κριτήριο Πληροφορίας (BIC) και μία παραλλαγή του κριτηρίου Anderson-Darling(ADC). Η διαφορά αυτών των κριτηρίων σε σχέση με τα κλασικά τεστ ελέγχου καταλληλότητας προσαρμογής μιας κατανομής είναι κυρίως η δυνατότητα που δίνουν στον μελετητή για την διάκριση ως καταλληλότερης μίας και μόνο κατανομής έναντι υπολοίπων, γεγονός που αποτέλεσε και το βασικό κριτήριο για την επιλογή αυτών των κριτηρίων στην παρούσα εργασία.

Από τη διαδικασία αυτή προέκυψε ότι και οι τέσσερις επιλεγμένες κατανομές επιλέχθηκαν σε σχετικά ικανοποιητικά ποσοστά. Η λογαριθμοκανονική κατανομή ήταν αυτή που τουλάχιστον από τα κριτήρια AIC και BIC επιλέχθηκε στις περισσότερες περιπτώσεις, ενώ δείχθηκε επίσης ότι τα δύο αυτά κριτήρια στην πλειονότητα των περιπτώσεων έκαναν ακριβώς της ίδια επιλογή. Παρατηρήθηκε επίσης, ότι τα κριτήρια AIC και BIC σε σύγκριση με το ADC επέλεξαν ως καταλληλότερη μια διπαραμετρική κατανομή(αρχή της φειδωλότητας) σε μεγαλύτερα ποσοστά και μάλιστα ασχέτως από το μήκος του διαθέσιμου δείγματος. Από την άλλη μεριά, το ADC έτεινε προς την επιλογή τριπαραμετρικών κατανομών για όλο και μεγαλύτερο μήκος δείγματος.

6.2 Προτάσεις για μελλοντική έρευνα

Στη βιβλιογραφία υπάρχουν πολλές έρευνες και μελέτες που αναλύουν ακραία καιρικά φαινόμενα και ειδικότερα πλημμυρικές παροχές, με επίκεντρο την ανάλυση τάσεων. Στο μέλλον θα μπορούσε να δοθεί έμφαση στους διάφορους τρόπους και τα εναλλακτικά τεστ στατιστικού ελέγχου που θα μπορούσαν να είχαν εφαρμοστεί στη συγκεκριμένη βάση δεδομένων, εξάγοντας σημαντικά συμπεράσματα για την αξιοπιστία τους και τις διαφοροποιήσεις τους. Επίσης, παρουσιάζει ενδιαφέρον η επανάληψη των αναλύσεων που χρησιμοποιήθηκαν με διαφορετικά κριτήρια επιλογής των σταθμών για την επίδραση του μήκους δείγματος στις στατιστικές εκτιμήσεις των παραμέτρων.

Εκτός αυτού όμως, εξέχουσας σημασίας θα ήταν η εφαρμογή των παραπάνω αναλύσεων σε βάσεις δεδομένων που έχουν επεξεργαστεί άλλες μελέτες, με την προϋπόθεση ότι έχουν ελεγχθεί όλοι οι σταθμοί για την ύπαρξη ταμιευτήρων ανάντη των σταθμών μέτρησης αλλά και έχει ληφθεί υπόψη η όποια δομή συσχέτισης των δεδομένων που ενδεχομένως παραβιάζει τις υποθέσεις της κλασικής στατιστικής. Διαδικασία που αναμφισβήτητα είναι χρονοβόρα και επίπονη αλλά με πολύ πιο αξιόπιστα αποτελέσματα.

Τέλος θα είχε ειδική σημασία η επανάληψη της ανάλυσης και των υπολογισμών σε βάθος χρόνου, όταν θα είναι διαθέσιμα πολύ μεγαλύτερα μήκη καταγραφών του δείγματος και η σύγκριση των εξαχθέντων αποτελεσμάτων.

7 ΒΙΒΛΙΟΓΡΑΦΙΚΕΣ ΑΝΑΦΟΡΕΣ

Adamowski, K. & Bocci, C. (2001) Geostatistical regional trend detection in river flow data. *Hydrol. Processes* 15, 3331–3341

Adenstedt, R. K. (1974) On large sample estimation for the mean of a stationary random sequence. *Ann. Statist.* 2, 1095–1107

Akaike, H. (1973), Information theory and an extension of the maximum likelihood principle, in *Second International Symposium on Information Theory*, edited by B. N. Petrov and F. Csaki, pp. 267–281, Acad. Kiado, Budapest

Beran, J. (1994) *Statistics for Long-Memory Processes*, vol. 61 of *Monographs on Statistics and Applied Probability*,

Beran, J., and Y. Feng (2002), SEMIFAR models—A semiparametric approach to modelling trends, *Comput. Stat. Data Anal.*, 40(2), 393–419.

Blender, Richard, and Klaus Fraedrich. 2006. 'Long-term Memory of the Hydrological Cycle and River Runoffs in China in a High-resolution Climate Model'. *International Journal of Climatology* 26 (12) (October 1): 1547–1565. doi:10.1002/joc.1325

Burn, D. H. & Hag Elnur, M. A. (2002) Detection of hydrologic trends and variability. *J. Hydrol.* 255, 107–122.

Burnham, K. P., and D. R. Anderson (2002), *Model Selection and Multimodel Inference*, 2nd ed., Springer, New York

CEH Wallingford/Met Office (2001) To what degree can the October/November 2000 flood events be attributed to climate change? Technical report to Defra, Project FD2304, March 2001. CEH Wallingford, UK.

Chapman and Hall, New York, USA.

Craigmile, P. F., P. Guttorp, and D. B. Percival (2004a), Wavelet-based parameter estimation for trend contaminated fractionally differenced processes, NRCSE Tech. Rep. Ser. 77, Univ. of Wash., Seattle.

Craigmile, P. F., P. Guttorp, and D. B. Percival (2004b), Trend assessment in a long memory dependence model using the discrete wavelet transform, *Environmetrics*, 15(4), 35–313.

Daniels, J. Michael. 2007. 'Flood Hydrology of the North Platte River Headwaters in Relation to Precipitation Variability'. *Journal of Hydrology* 344 (1–2) (September 30): 70–81. doi:10.1016/j.jhydrol.2007.06.020

Douglas, E. M., Vogel, R. M. & Kroll, C. N. (2000) Trends in floods and low flows in the United States: impact of spatial correlation. *J. Hydrol.* 240, 90–105

Eltahir, E.A.B. 1996. 'El Nino and the Natural Variability in the Flow of the Nile River'. *Water Resources Research* 32 (1): 131–137.

European Directive 2000/06/EC

Fisher, R., & Tippett, L. 1928. Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proceedings of the Cambridge Philosophical Society*, Vol 24, 180–190

Fréchet, M. (1927) Sur la loi de probabilité de l'écart maximum. Ann. de la Soc. Polonaise de Math., Cracow 6, 93–117

Freund, J. E., F. J. Williams, and B. M. Perles, Elementary Business Statistics, The Modern Approach, Prentice-Hall, 1988

Gnedenko, B. 1941. Limit theorems for the maximal term of a variational series. Doklady Akad. Nauk SSSR, Moscow, 32, 37 (in Russian)

GRDC (Global Runoff Data Centre) (2003) Dataset of 21 river discharge time series from the Global Runoff Data Centre GRDC, Koblenz, Germany. <http://grdc.bafg.de>

Grubb, H. & Robson, A. (2000) Exploratory/visual analysis. In: *Detecting Trend and Other Changes in Hydrological Data* (ed. by Z. W. Kundzewicz & A. Robson), 19–47. World Climate Programme—Water, World Climate Data and Monitoring Programme, WCDMP-45, WMO/TD no. 1013. World Meteorological Organization, Geneva, Switzerland

Gumbel, Emil Julius. 1958. Statistics of extremes. New York: Columbia University Press

Hamed, K. (2008) Trend Detection in Hydrologic Data: The Mann-Kendall Trend Test under the Scaling Hypothesis. *Journal of Hydrology*, 349, 350-363

Hamed, K.H., Rao, A.R., 1998. A modified Mann–Kendall trend test for autocorrelated data. *J. Hydrol.* 204, 182–196

Herschey R., World Catalogue of Maximum Observed Floods, IAHS Press, 2003

Hirpa, Feyera A., Mekonnen Gebremichael, and Thomas M. Over. "River flow fluctuation analysis: Effect of watershed area." *Water Resources Research* 46.12 (2010)

Hirsch, R. M., D. R. Helsel, T. A. Cohn, and E. J. Gilroy, Statistical analysis of hydrologic data, in Handbook of Hydrology, edited by D. R. Maidment, McGraw-Hill, 1993

Hurst, H. E. (1951) Long term storage capacities of reservoirs. Trans. ASCE 116, 776–808

IPCC (Intergovernmental Panel on Climate Change) (2001) *Climate Change 2001: Synthesis Report*. Third Assessment Report of the IPCC. Cambridge University Press, Cambridge, UK

Jenkinson, A. F. (1955) The frequency distribution of the annual maximum (or minimum) value of meteorological elements. *Quart. J. Roy. Met.. Soc.* **81**, 158–171.

Jenkinson, A. F. (1969) Estimation of maximum floods. Tech. Note 98, Ch. 5, 183–257. World Meteorological Organization, Geneva, Switzerland.

Jones, P. D., K. R. Briffa, T. P. Barnett, and S. F. B. Tett. 1998. 'Millennial Temperature Reconstructions'. IGBP PAGES/world Data center-A for Paleoclimatology Data Contribution Series 39

Kendall, M.G., 1975. Rank Correlation Methods. Griffin, London.

Kolmogorov, A. 1941. 'The Local Structure of Turbulence in Incompressible Viscous Fluid for Very Large Reynolds' Numbers'. In *Akademiia Nauk SSSR Doklady*, 30:301–305

Koscielny-Bunde, E., J. W. Kantelhardt, P. Braun, A. Bunde, and S. Havlin. 2006. 'Long-term Persistence and Multifractality of River Runoff Records: Detrended Fluctuation Studies'. *Journal of Hydrology* 322 (1): 120–137.

Koutsoyiannis, D. 2003. 'Climate Change, the Hurst Phenomenon, and Hydrological Statistics'. *Hydrological Sciences Journal* 48 (1): 3–24

Koutsoyiannis, D. 2004. Statistics of Extremes and Estimation of Extreme Rainfall, 1, Theoretical Investigation. *Hydrological Sciences Journal* 49 (4): 575–590

Koutsoyiannis, D. 2011a. Hurst-Kolmogorov Dynamics and Uncertainty. *JAWRA Journal of the American Water Resources Association* 47 (3): 481–495

Koutsoyiannis, D. Encolpion of stochastics: Fundamentals of stochastic processes, [doi:10.13140/RG.2.2.10956.82564](https://doi.org/10.13140/RG.2.2.10956.82564), Department of Water Resources and Environmental Engineering – National Technical University of Athens, Athens, 2013.

Koutsoyiannis, D. Nonstationarity versus scaling in hydrology, *Journal of Hydrology*, 324, 239–254, 2006

Koutsoyiannis, D. The Hurst phenomenon and fractional Gaussian noise made easy, *Hydrological Sciences Journal*, 47 (4), 573–595, [doi:10.1080/02626660209492961](https://doi.org/10.1080/02626660209492961), 2002.

Koutsoyiannis, D., and A. Montanari. 2007. ‘Statistical Analysis of Hydroclimatic Time Series: Uncertainty and Insights’. *Water Resources Research* 43 (5): W05429

Kundzewicz ZW, Robson AJ Change detection in river flow records—review of methodology. *Hydrological Science Journal* 49(1):7–19, 2004

Kundzewicz, Z. W. (2004) Searching for change in hydrological data—Editorial. *Hydrol. Sci. J.* 49(1), 3–6.

Kunkel, K. E., Changnon, S. A. & Angel, J. R. (1994) Climatic aspects of the 1993 upper Mississippi river basin flood. *Bull. Am. Met. Soc.* 75, 811–822.

Laio, F. (2004), Cramer– von Mises and Anderson-Darling goodness of fit tests for extreme value distributions with unknown parameters, *Water Resour. Res.*, 40, W09308, doi:10.1029/2004WR003204

Laio, F., G. Di Baldassarre, and A. Montanari (2009), Model selection techniques for the frequency analysis of hydrological extremes, *Water Resour. Res.*, 45, W07416, doi:10.1029/2007WR006666

Leadbetter M. R. (1974) On extreme values in stationary sequences. *Z. Wahrscheinlichkeitstheorie u. Verwandte Gebiete* 28, 289–303.

Lindström, G. & Bergström, S. (2003) Long-term variation in runoff and temperature in Sweden. In: *Water Resources Systems—Water Availability and Global Change* (Proc. Sapporo Symp., July 2003) (ed. by S. Franks, G. Blöschl, M. Kumagai, K. Musiake & D. Rosbjerg). IAHS Publ. 280, IAHS Press, Wallingford, UK

Lindström, G. & Bergström, S. (2004) Runoff trends in Sweden 1807–2002. *Hydrol. Sci. J.* 49(1), 69–83.

Linhart, H., and W. Zucchini (1986), *Model Selection*, John Wiley, Hoboken, N. J.

Lins, H. F. & Slack, J. R. (1999) Streamflow trends in the United States. *Geophys. Res. Lett.* 26, 227–230.

Mallakpour I, Villarini G. 2015. The changing nature of flooding across the central United States. *Nat. Clim. Change* 5: 250–254, doi:10.1038/NCLIMATE2516

Mandelbrot, B.B. and J. Wallis, 1968, Noah, Joseph and operational hydrology, *Water Resources Research* 4, 909-918.

- Mann, H.B., 1945. Nonparametric tests against trend. *Econometrica* 13, 245–259
- Marsh, T. J. & Bradford, R. B. (2003) The floods of August 2002 in central Europe. *Weather* 58, 168.
- Miller, J. R. & Russell, G. L. (1992) The impact of global warming on river runoff. *J. Geophys. Res. D* 97, 2757–2764.
- Milly, P. C. D., Wetherald, R. T., Dunne, K. A. & Delworth, T. L. (2002) Increasing risk of great floods in a changing climate. *Nature* 415, 514–517.
- Nijssen, B., O'Donnell, G. M., Hamlet, A. F. & Lettenmaier, D. P. (2001) Hydrologic sensitivity of global rivers to climate change. *Climatic Change* 50, 143–175.
- Onoz, B., and M. Bayazit (1995), Best-fit distribution of largest available flood samples, *J. Hydrol.*, 167, 195 – 204, doi:10.1016/0022-1694(94)02633-M.
- Papoulis, A., Probability and Statistics, Prentice-Hall, 1990
- Puckridge, J.T., K.F. Walker, and J.F. Costelloe. 2000. 'Hydrological Persistence and the Ecology of Dryland Rivers'. *River Research and Applications* 16 (5): 385–402
- Rea, William, Les Oxley, Marco Reale, and Jennifer Brown. 2009. 'Estimators for Long Range Dependence: An Empirical Study'. arXiv Preprint arXiv:0901.0762. <http://arxiv.org/abs/0901.0762>.
- Reynard, N. S., Prudhomme, C. & Crooks, S. M. (2001) The flood characteristics of large UK rivers: Potential effects of changing climate and land use. *Climatic Change* 48, 343–359
- Robson, A. J., Jones, T. K., Reed, D. W. & Bayliss, A. C. (1998) A study of national trend and variation in UK floods. *Int. J. Climatol.* 18, 165–182.

Rossi, F., Fiorentino, M., & Versace, P. (1984). Two-component extreme value distribution for flood frequency analysis. *Water Resour. Res.*, 20(7), 847–856.

Saurí, D., Milego, R., Canalís, A., Ripoll, A. & Kleeschulte, S. (2003) Mapping the impacts of recent natural disasters and technological accidents in Europe. Environmental issue report no. 35, European Environment Agency, Copenhagen, Denmark.

Schwarz, G. (1978), Estimating the dimension of a model, *Ann. Stat.*, 6, 461– 464, doi:10.1214/aos/1176344136.

Singh, V. P., and A. K. Rajagopal, A new method of parameter estimation for hydrologic frequency analysis, *Hydrological Science and Technology*, 2(3) 33-44, 1986

Slater L , Villarini G . On the impact of gaps on trend detection in extreme streamflow time series. *Int J Climatol* 2016. doi:[10.1002/joc.4954](https://doi.org/10.1002/joc.4954).

Stedinger, J. R., R. M. Vogel, and E. Foufula-Georgiou (1992), Frequency analysis of extreme events, in *Handbook of Hydrology*, edited by R. Maidment, chap. 18, pp. 18.1– 18.66, McGraw-Hill, New York

Stedinger, J. R., Vogel, R. M., & Foufoula-Georgiou, E. (1993). Frequency Analysis of Extreme Events, Chapter 18 in *Handbook of Hydrology*. Maidment, DR, McGraw-Hill

Stephens, M. A. (1986), Tests based on EDF statistics, in *Goodness-of-Fit Techniques*, edited by R. B. D'Agostino and A. M. Stephens, pp. 97–194, Marcel Dekker, New York

Strupczewski, W. G., V. P. Singh, and W. Feluch (2001), Non-stationary approach to at-site flood frequency modeling: I. Maximum likelihood estimation, *J. Hydrol.*, 248, 123 – 142, doi:10.1016/S0022- 1694(01)00397-3

Szolgayová, E., et al. "Factors influencing long range dependence in streamflow of European rivers." *Hydrological Processes* 28.4 (2014): 1573-1586.

Todorovic, P., and E. Zelenhasic. 1970. A Stochastic Model for Flood Analysis. *Water Resources Research* 6 (6): 1641.

Toussoun, O. (1925) *Mémoire sur l'histoire du Nil*. In: *Mémoires a l'Institut d'Egypte*, vol. 18, 366–404

Trenberth K. E. (1998) Atmospheric moisture residence times and cycling: Implications for rainfall rates and climate change. *Climatic Change* 39, 667–694.

Tufte E. R. (1983) *The Visual Display of Quantitative Information*. Graphics Press, Cheshire, Connecticut, USA

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. DOI:10.1007/s00477-010-0408-x

Viglione A., F. Laio, and P. Claps (2007), A comparison of homogeneity tests for regional frequency analysis, *Water Resour. Res.*, 43, W03428, doi:10.1029/2006WR005095

Weron Rafal. 2002. 'Estimating Long-range Dependence: Finite Sample Properties and Confidence Intervals'. *Physica A: Statistical Mechanics and Its Applications* 312 (1): 285–299

Yue S., Pilon P., Cavadias G., 2002b. Power of the Mann–Kendall and Spearman's rho tests for detecting monotonic trends in hydrological series. *J. Hydrol.* 259, 254–271

Δ. Κουτσογιάννης, Στατιστική Υδρολογία, Έκδοση 4, doi:10.13140/RG.2.1.5118.2325, Εθνικό Μετσόβιο Πολυτεχνείο, Αθήνα, 1997

