



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ  
ΣΧΟΛΗ ΠΟΛΙΤΙΚΩΝ ΜΗΧΑΝΙΚΩΝ  
ΤΟΜΕΑΣ ΥΔΑΤΙΚΩΝ ΠΟΡΩΝ ΚΑΙ ΠΕΡΙΒΑΛΛΟΝΤΟΣ

ΣΤΑΤΙΣΤΙΚΗ ΑΝΑΛΥΣΗ ΚΑΙ ΥΔΡΟΛΟΓΙΚΗ  
ΕΡΜΗΝΕΙΑ ΤΗΣ ΔΙΑΙΤΑΣ ΤΩΝ ΗΜΕΡΗΣΙΩΝ  
ΠΑΡΟΧΩΝ ΣΕ ΜΕΓΑΛΟ ΔΕΙΓΜΑ  
ΠΟΤΑΜΩΝ ΤΗΣ ΕΥΡΩΠΗΣ

Διπλωματική Εργασία

Προμπονάς Εμμανουήλ

Επιβλέπων : Ανδρέας Ευστρατιάδης, Επίκουρος Καθηγητής Ε.Μ.Π.

Αθήνα, Νοέμβριος 2021

# Περιεχόμενα

- 1 Εισαγωγή
- 2 Βάση Δεδομένων
- 3 Στατιστικά Εργαλεία
- 4 Υπολογιστικά Εργαλεία
- 5 Αποτελέσματα Ανάλυσης
- 6 Συμπεράσματα

# ΕΙΣΑΓΩΓΗ (1/2)



## Έναυσμα σύλληψης της εργασίας

Ο κλάδος της υδρολογίας στηρίζεται στην μελέτη- παρατήρηση των φυσικών-υδρολογικών διεργασιών μέσω της επεξεργασίας υπάρχουσών μετρήσεων

Οι μετρήσεις λοιπόν έχουν μια άμεση και προφανή χρησιμότητα αφού η μελέτη των φυσικών διεργασιών-φαινομένων, η κατανόηση τους και η γνώση της χρονικής τους εξέλιξης βασίζεται στη δημιουργία χρονοσειρών με μετρήσεις και στην επεξεργασία αυτών των χρονοσειρών

Ένα από τα συνηθέστερα προβλήματα το οποίο καλείται να αντιμετωπίσει όποιος ασχολείται με την επιστήμη της υδρολογίας και συγκεκριμένα με τις ημερήσιες παροχές ποταμών είναι η έλλειψη (συστηματική ή μη) χρονοσειρών μετρήσεων

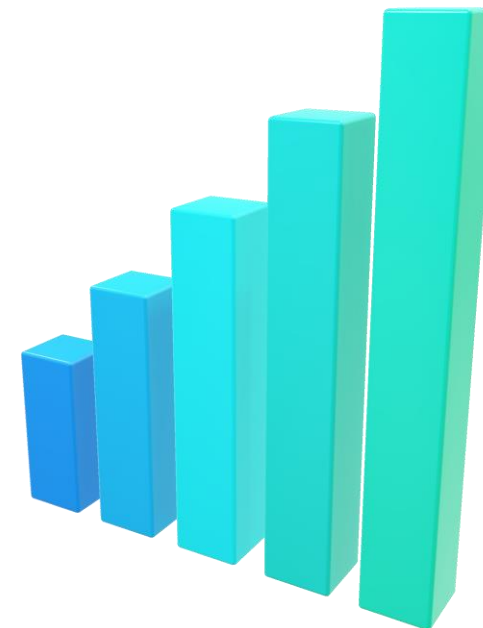


## ΕΙΣΑΓΩΓΗ (2/2)

Σκοπός της Διπλωματικής εργασίας



Κρίνεται λοιπόν απαραίτητη η **στατιστική ανάλυση** των υπάρχοντων δεδομένων ώστε να αποκτήσουμε μια γενική εικόνα για την στατιστική συμπεριφορά των ημερήσιων παροχών και ενδεχομένως να ανιχνεύσουμε κάποια ομαδοποίηση συμπεριφορών ανάλογα με τα κλιματικά χαρακτηριστικά κάθε περιοχής.



# Περιεχόμενα

## 2 Βάση Δεδομένων

# ΒΑΣΗ ΔΕΔΟΜΕΝΩΝ (1/3)

## Βάση Δεδομένων – Κριτήρια Επιλογής Σταθμών

---

### Global Historical Climatology Network, GHCN-Daily της N.O.A.A.



Το GHCN-Daily αποτελεί μια ενσωματωμένη βάση δεδομένων που περιέχει ημερήσια κλιματικά αρχεία ελεγμένης ποιότητας. Περιέχει δεδομένα από περίπου 75.000 σταθμούς από 180 χώρες. Το μήκος των αρχείων καθώς και η χρονική περίοδος των καταγραφών ποικίλουν από λιγότερο από ένα χρόνο μέχρι περισσότερο από 175 χρόνια.

### Κριτήρια επιλογής σταθμών

- μήκος εγγραφής πάνω από 50 χρόνια
- οι τιμές που λείπουν να είναι λιγότερες από 20%

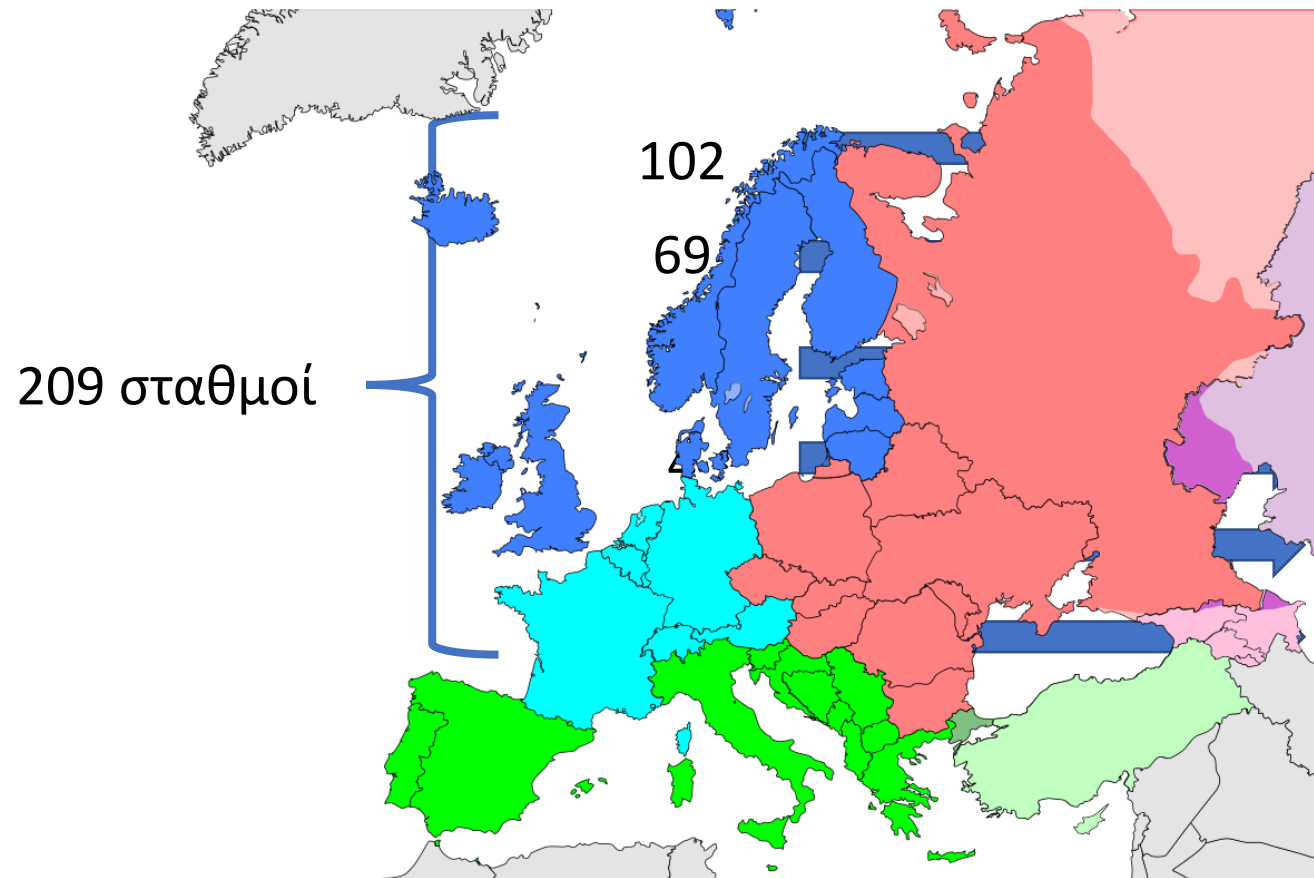
# ΒΑΣΗ ΔΕΔΟΜΕΝΩΝ (2/3)

Περιοχή μελέτης - Πλήθος σταθμών

Περιοχή μελέτης



Ευρώπη



Αυστρία

Σουηδία

Σλοβενία

Γαλλία

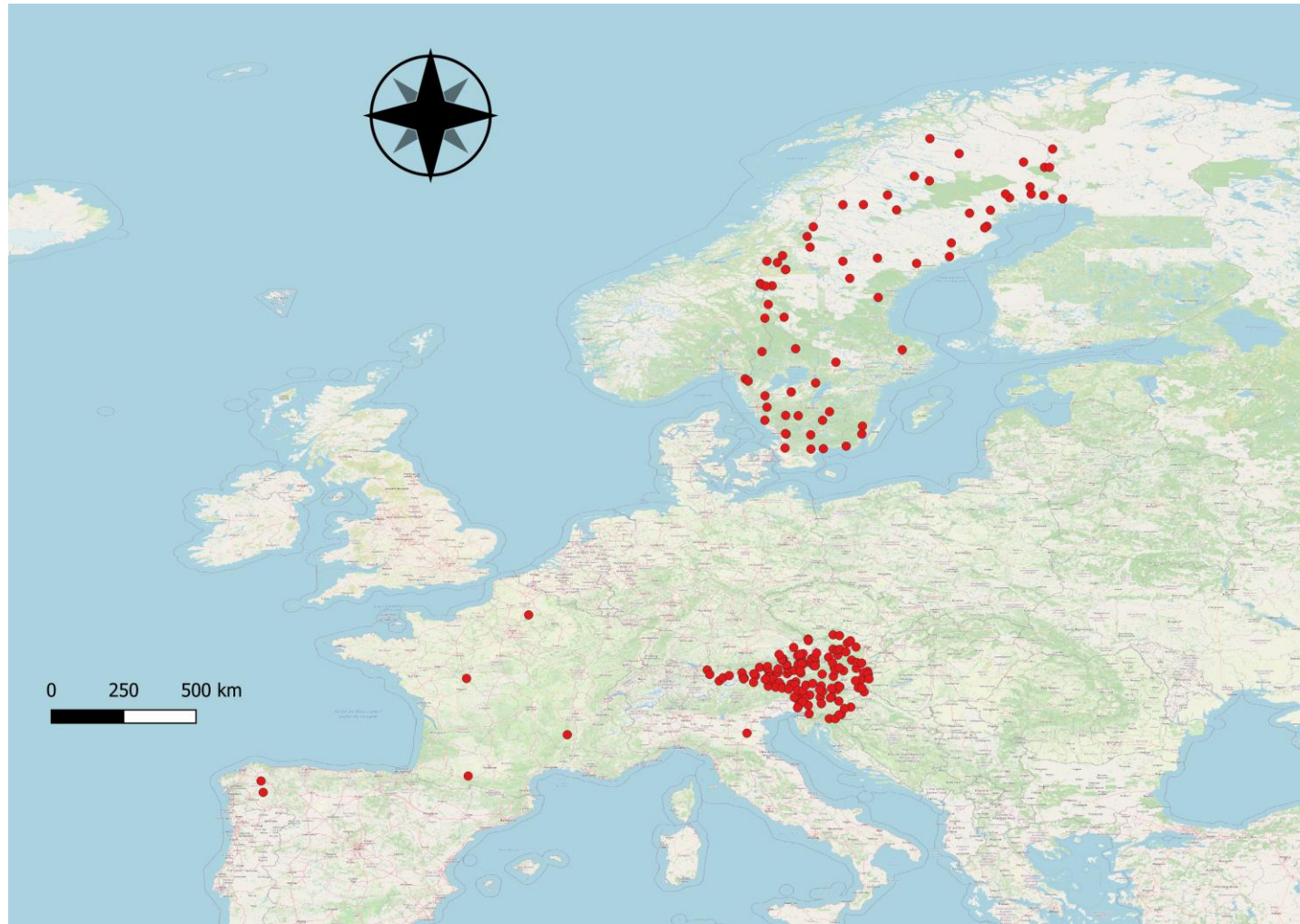
Ισπανία

Ιταλία



## ΒΑΣΗ ΔΕΔΟΜΕΝΩΝ (3/3)

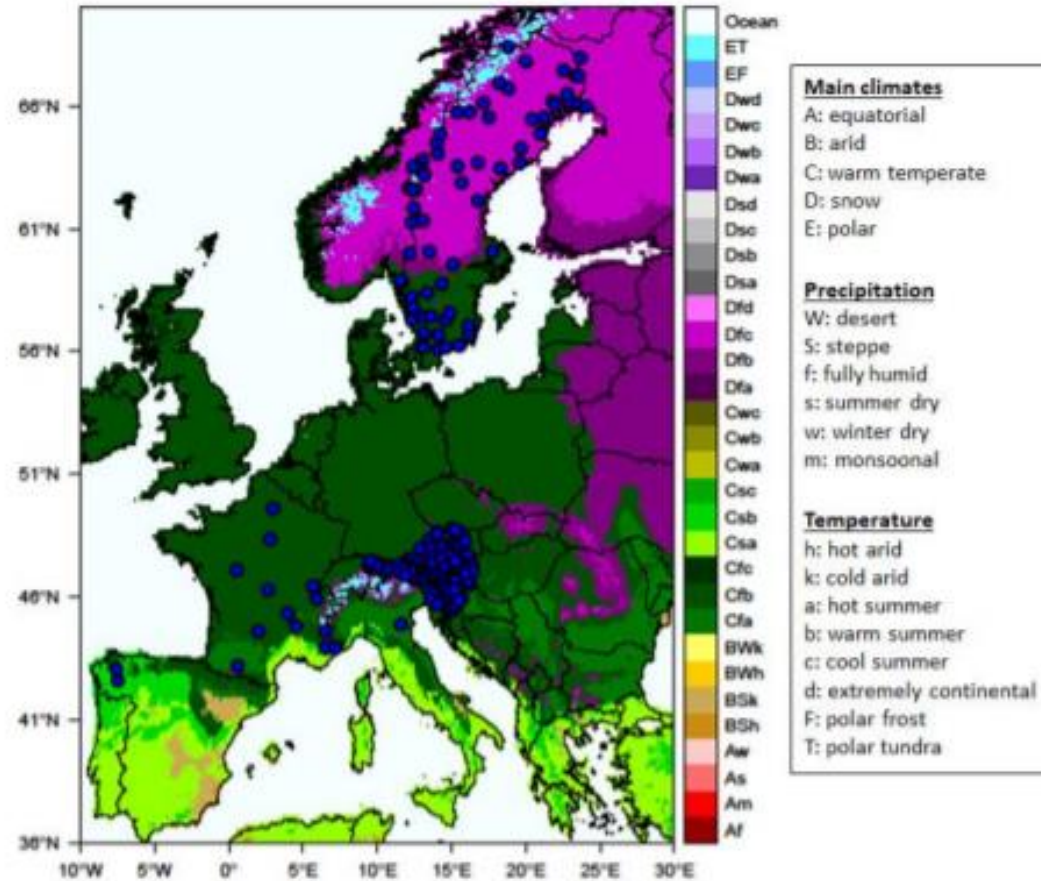
Γεωγραφική κατανομή σταθμών που χρησιμοποιήθηκαν στη μελέτη





# ΒΑΣΗ ΔΕΔΟΜΕΝΩΝ (3/3)

Κλιματικό καθεστώς σταθμών που χρησιμοποιήθηκαν στη μελέτη



Ενημερωμένος κλιματικός χάρτης Köppen – Geiger

# Περιεχόμενα

3

Στατιστικά Εργαλεία

# Κατανομές Μελέτης (1/3)

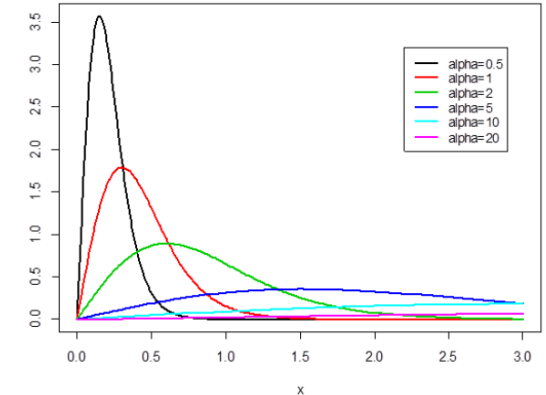
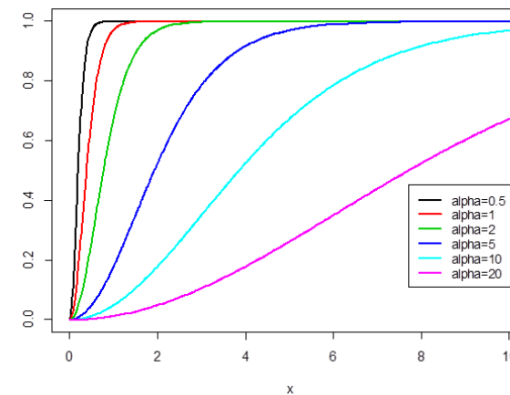


Παρουσίαση CDF (Αθροιστική Συνάρτηση Κατανομής) και PDF (Συνάρτηση Πυκνότητας Πιθανότητας)

## ➤ Pareto Burr Feller (PBF)

$$\text{CDF: } F(x|\lambda, \zeta, \xi) = 1 - \left(1 + \xi\zeta \left(\frac{x}{\lambda}\right)^\zeta\right)^{-\frac{1}{\xi\zeta}}, \quad x \geq 0$$

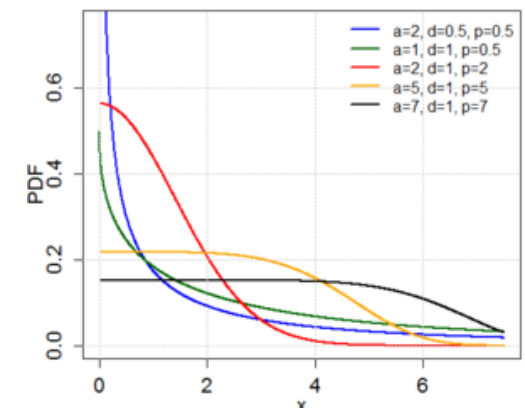
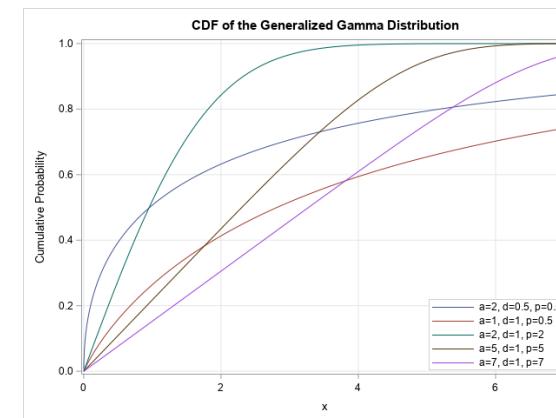
$$\text{PDF: } f(x|\lambda, \zeta, \xi) = \frac{1}{\lambda\xi} \left(\frac{x}{\lambda}\right)^{\zeta-1} \left(1 + \left(\frac{x}{\lambda}\right)^\zeta\right)^{-\frac{1}{\xi\zeta}-1}, \quad x \geq 0$$



## ➤ Generalized Gamma (GG)

$$\text{CDF: } F(x|b, a_1, a_2) = 1 - \frac{\Gamma\left[\frac{a_1}{a_2}, \left(\frac{x}{b}\right)^{a_2}\right]}{\Gamma\left(\frac{a_1}{a_2}\right)}, \quad x \geq 0$$

$$\text{PDF: } f(x|b, a_1, a_2) = \frac{a_2}{b\Gamma(a_1/a_2)} \left(\frac{x}{b}\right)^{a_1-1} \exp\left[-\left(\frac{x}{b}\right)^{a_2}\right], \quad x \geq 0$$



## Κατανομές Μελέτης (2/3)

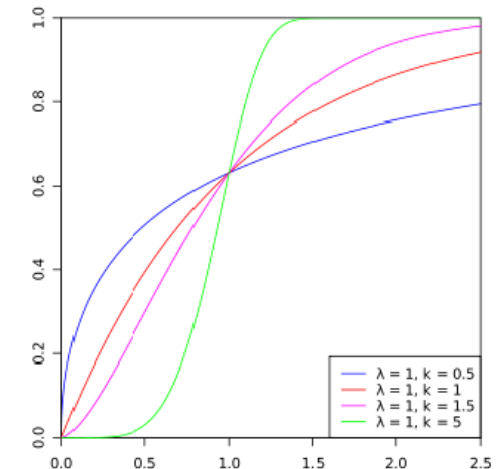
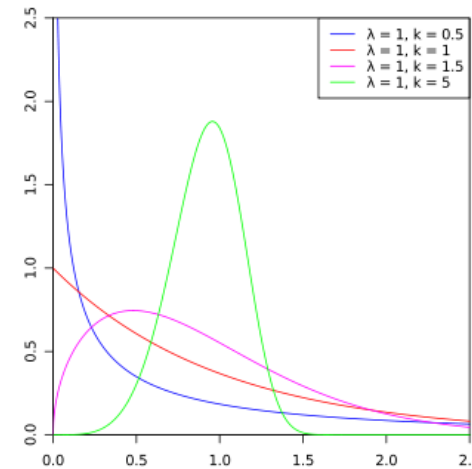


Παρουσίαση CDF (Αθροιστική Συνάρτηση Κατανομής) και PDF (Συνάρτηση Πυκνότητας Πιθανότητας)

### ➤ Weibull

$$\text{CDF: } F(x|a, b) = 1 - e^{-\left(\frac{x}{a}\right)^b}, x > 0, a > 0, b > 0$$

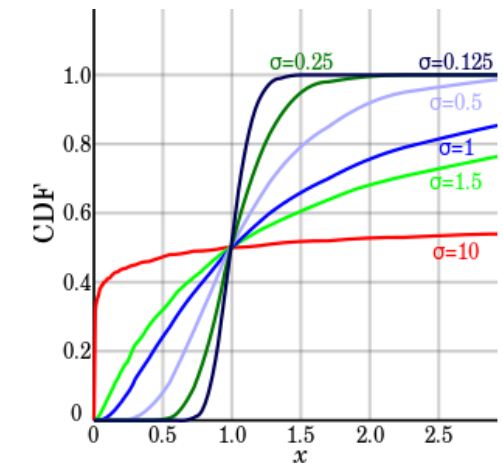
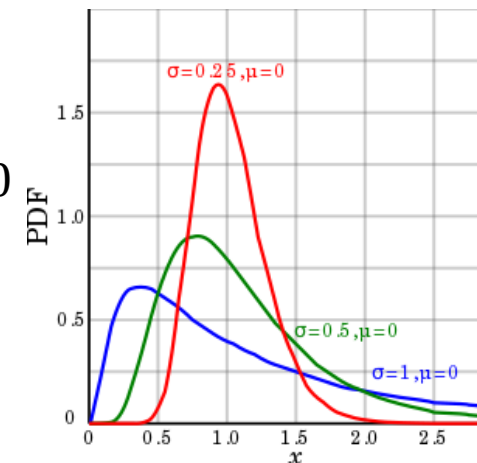
$$\text{PDF: } f(x|a, b) = \frac{b}{a} \left(\frac{x}{a}\right)^{b-1} e^{-\left(\frac{x}{a}\right)^b}, x > 0, a > 0, b > 0$$



### ➤ Lognormal

$$\text{CDF: } F(x|\mu, \sigma) = \frac{1}{2} + \frac{1}{2} \operatorname{erf} \left[ \frac{\ln(x) - \mu}{\sqrt{2}\sigma} \right], x > 0, -\infty < \mu < +\infty, \sigma > 0$$

$$\text{PDF: } f(x|\mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln(x) - \mu)^2}{2\sigma^2}}, x > 0, -\infty < \mu < +\infty, \sigma > 0$$



## Κατανομές Μελέτης (3/3)

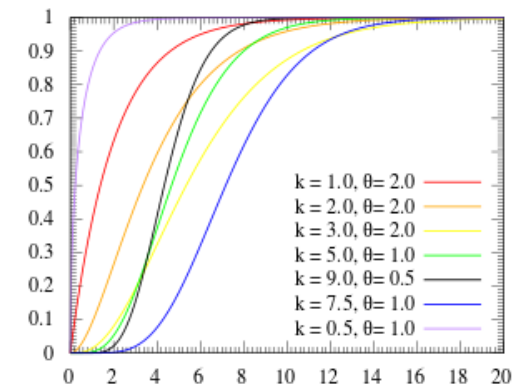
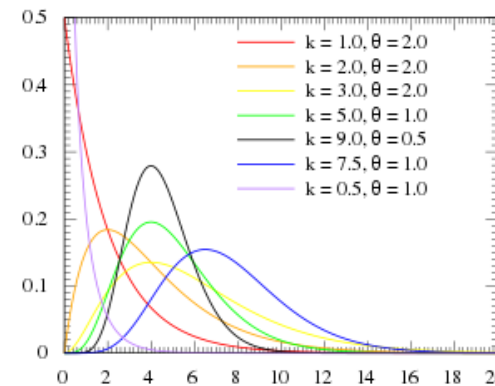


Παρουσίαση CDF (Αθροιστική Συνάρτηση Κατανομής) και PDF (Συνάρτηση Πυκνότητας Πιθανότητας)

### ➤ Gamma

$$\text{CDF: } F(x|k, \theta) = \frac{1}{\Gamma(k)} \gamma\left(k, \frac{x}{\theta}\right), x > 0, k > 0, \theta > 0$$

$$\text{PDF: } f(x|k, \theta) = \frac{1}{\Gamma(k)\theta^k} x^{k-1} e^{-\frac{x}{\theta}}, x > 0, k > 0, \theta > 0$$



## L – Ροπές (1/3)



### Ορισμός L - Ροπών

- Οι L-ροπές είναι μία σειρά παραμέτρων, ανάλογες με τις κλασσικές ροπές, που μπορούν να περιγράψουν τα στατιστικά χαρακτηριστικά του δείγματος (θέση – μέση τιμή, μεταβλητότητα, ασυμμετρία, κύρτωση κλπ.).
- Εισήχθησαν από τον Hosking (1990).
- Παράγονται από γραμμικούς συνδυασμούς των τιμών ενός ταξινομημένου δείγματος (από αυτό το γεγονός προέρχεται και ο χαρακτηρισμός L, δηλαδή Linear - γραμμικές).

$$\lambda_r = r^{-1} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} EX_{r-k:r} \quad \text{όπου:}$$

- $X_{k:n}$  είναι η  $k^{\text{η}}$  στατιστική τάξη από ένα ανεξάρτητο δείγμα μεγέθους  $n$  από μία κατανομή  $X$
- $E$  είναι η αναμενόμενη τιμή

## L – Ροπές (2/3)



Χρήση L - Ροπών

$\lambda_1 = EX$		Μέση Τιμή (Ροπή πρώτης τάξης)
$\lambda_2 = (EX_{2:2} - XE_{1:2})/2$		Μεταβλητότητα
$\lambda_3 = (EX_{3:3} - 2XE_{2:3} + EX_{1:3})/3$		Ασσυμετρία
$\lambda_4 = (EX_{4:4} - 3XE_{3:4} + 3EX_{2:4} - EX_{1:4})/4$		Κύρτωση

Αδιάστατοι  
Συντελεστές

$$\tau_2 = \frac{\lambda_2}{\lambda_1}$$



Συντελεστής Μεταβλητότητας

$$\tau_3 = \frac{\lambda_3}{\lambda_2}$$



Συντελεστής Ασσυμετρίας

$$\tau_4 = \frac{\lambda_4}{\lambda_2}$$



Συντελεστής Κύρτωσης

## L – Ροπές (3/3)



### Κριτήρια Επιλογής L - Ροπών

---

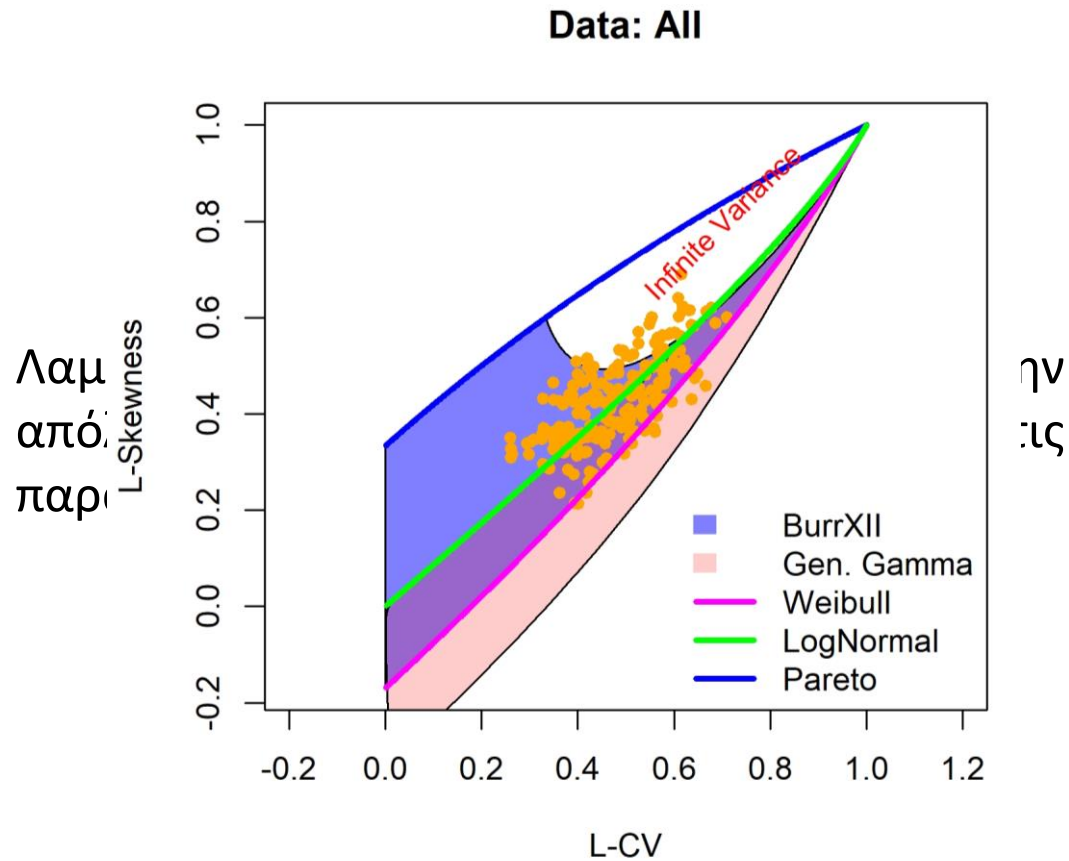
- 1) L – ροπές υψηλότερης τάξης, εκτιμώνται ακόμη και αν μόνο η πρώτη κλασσική ροπή (μέσος όρος) είναι πεπερασμένη.
- 2) Οι L – Ροπές δεν επηρεάζονται σημαντικά (είναι λιγότερο ευαίσθητες) από τις ακραίες τιμές και το μήκος του δείγματος.



# Δείκτης Αξιολόγησης (1/1)



Ο δείκτης που χρησιμοποιήθηκε για τον έλεγχο των κατανομών



## L – ratio diagramms

Τα διαγράμματα λόγου ροπής παρέχουν μια γραφική σύγκριση μεταξύ των στατιστικών δειγμάτων, που υπολογίζονται από τα δεδομένα της βάσης και των θεωρητικών, όπως δίδονται από τους τύπους των υποψήφιων παραμετρικών κατανομών. Η εγγύτητα των στατιστικών δειγμάτων στον θεωρητικό τόπο (δηλ. Σημείο, καμπύλη ή περιοχή) μιας συγκεκριμένης κατανομής παρέχει μια ένδειξη της καταλληλότητας αυτής της κατανομής για να περιγράψει καλύτερα τις καταγραφές δεδομένων.

# Περιεχόμενα

## 4 Υπολογιστικά Εργαλεία

# ΥΠΟΛΟΓΙΣΤΙΚΑ ΕΡΓΑΛΕΙΑ (1/3)



## Εργαλείο Ανάλυσης Δεδομένων – Στατιστικό Πακέτο R

- Η γλώσσα R είναι μια γλώσσα προγραμματισμού ανοικτού κώδικα συνοδευόμενη από ένα περιβάλλον για επεξεργασία δεδομένων, υπολογισμών και γραφημάτων.
- Τα κύρια δομικά στοιχεία της είναι οι υπάρχουσες συναρτήσεις (functions), που εκτελούνται μέσω λογικών τελεστών και τελεστών σύγκρισης, και τα πακέτα (packages) εκτέλεσης προσχεδιασμένων σειρών συναρτήσεων.
- Δυνατό σημείο της R είναι η ευκολία με την οποία παράγει καλά σχεδιασμένα διαγράμματα οπτικής απεικόνισης των Στατιστικών και Μαθηματικών Διαδικασιών που εκτελούνται.
- Ραγδαία ανάπτυξη της R σε όλες τις επιστήμες που οφείλεται στην ικανότητα που έχει, αφενός να διαχειρίζεται μεγάλο όγκο δεδομένων και αφετέρου να αναλύει και να επεξεργάζεται δεδομένα με μεγάλη ταχύτητα.

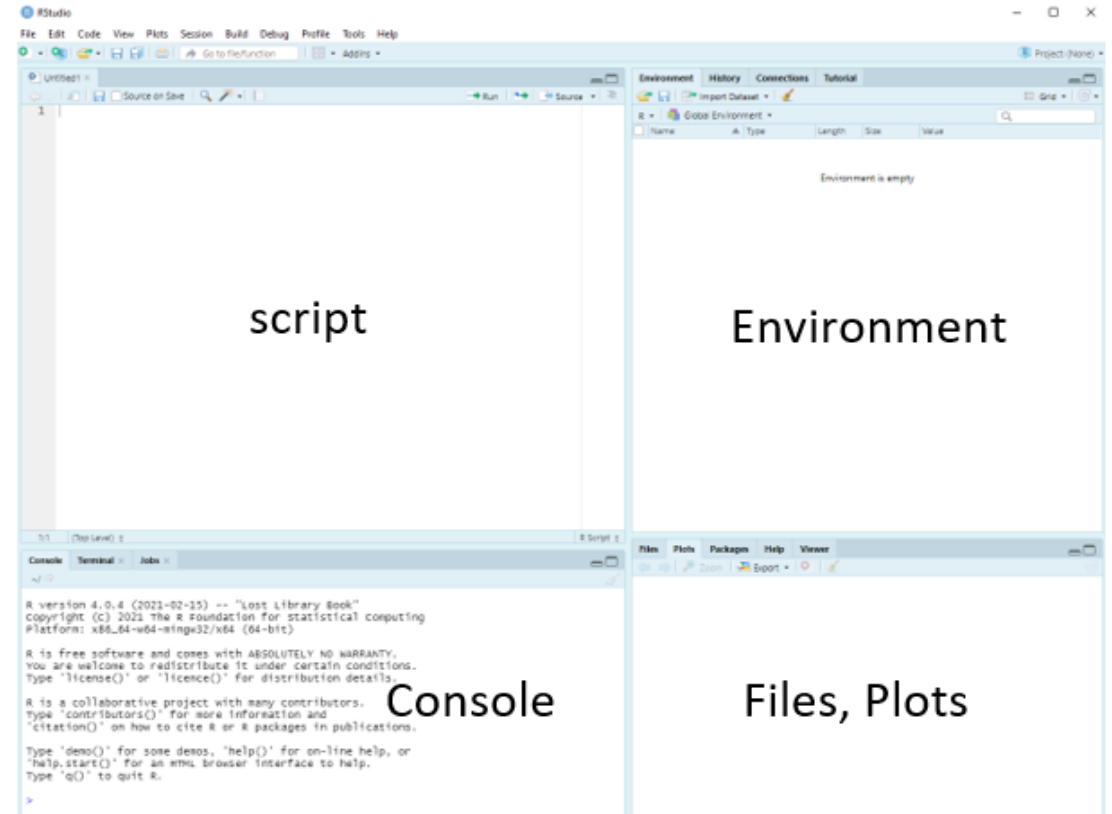


# ΥΠΟΛΟΓΙΣΤΙΚΑ ΕΡΓΑΛΕΙΑ (2/3)



## Προγραμματιστικό περιβάλλον R-Studio

- Το πρόγραμμα Rstudio είναι ένα ελεύθερο προγραμματιστικό περιβάλλον για την γλώσσα R, το οποίο δίνει τη δυνατότητα γραφής και εκτέλεσης του κώδικα, καθώς και εξαγωγής αποτελεσμάτων και γραφημάτων.

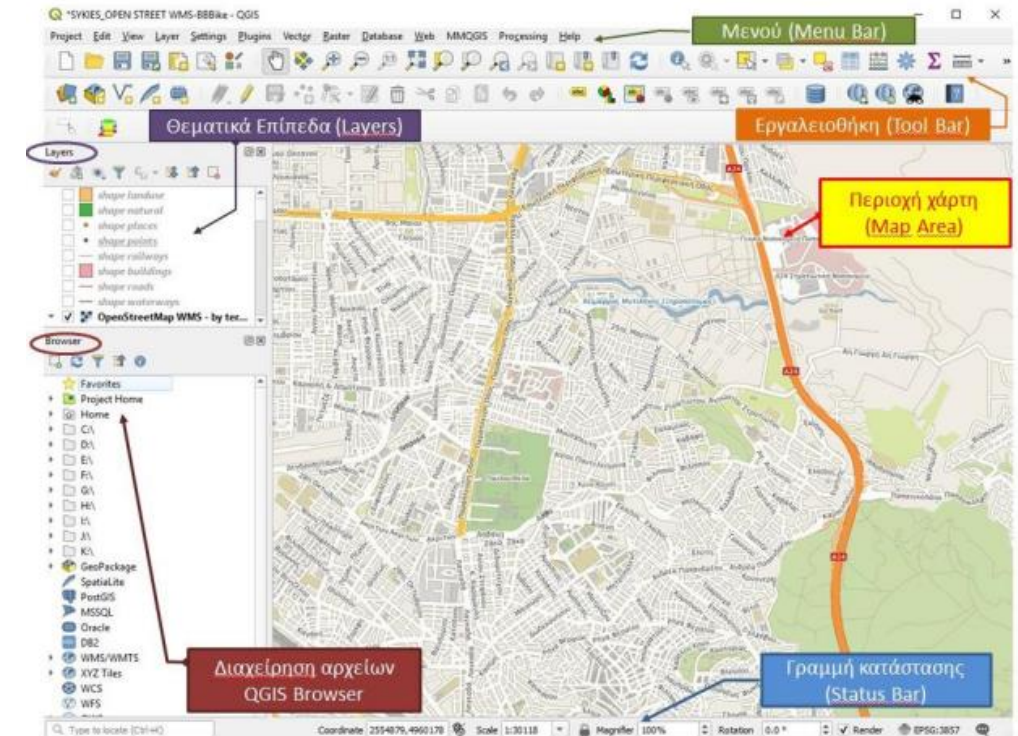


# ΥΠΟΛΟΓΙΣΤΙΚΑ ΕΡΓΑΛΕΙΑ (3/3)



## Ελεύθερο λογισμικό QGIS

- Το QGIS είναι ένα φιλικό ανοιχτού κώδικα (open source) Σύστημα Γεωγραφικών Πληροφοριών (GIS) που υποστηρίζει την προβολή, την επεξεργασία και την ανάλυση των γεωχωρικών δεδομένων όπως και την σύνθεση και εξαγωγή γραφικών χαρτών.

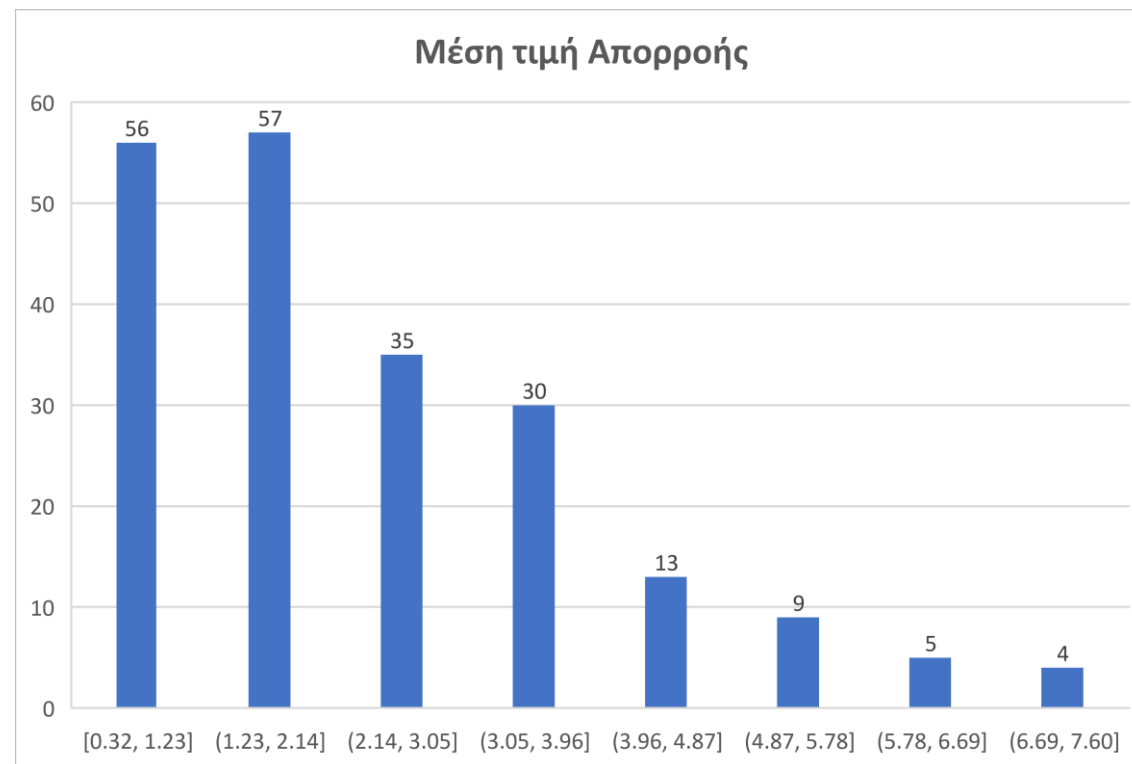
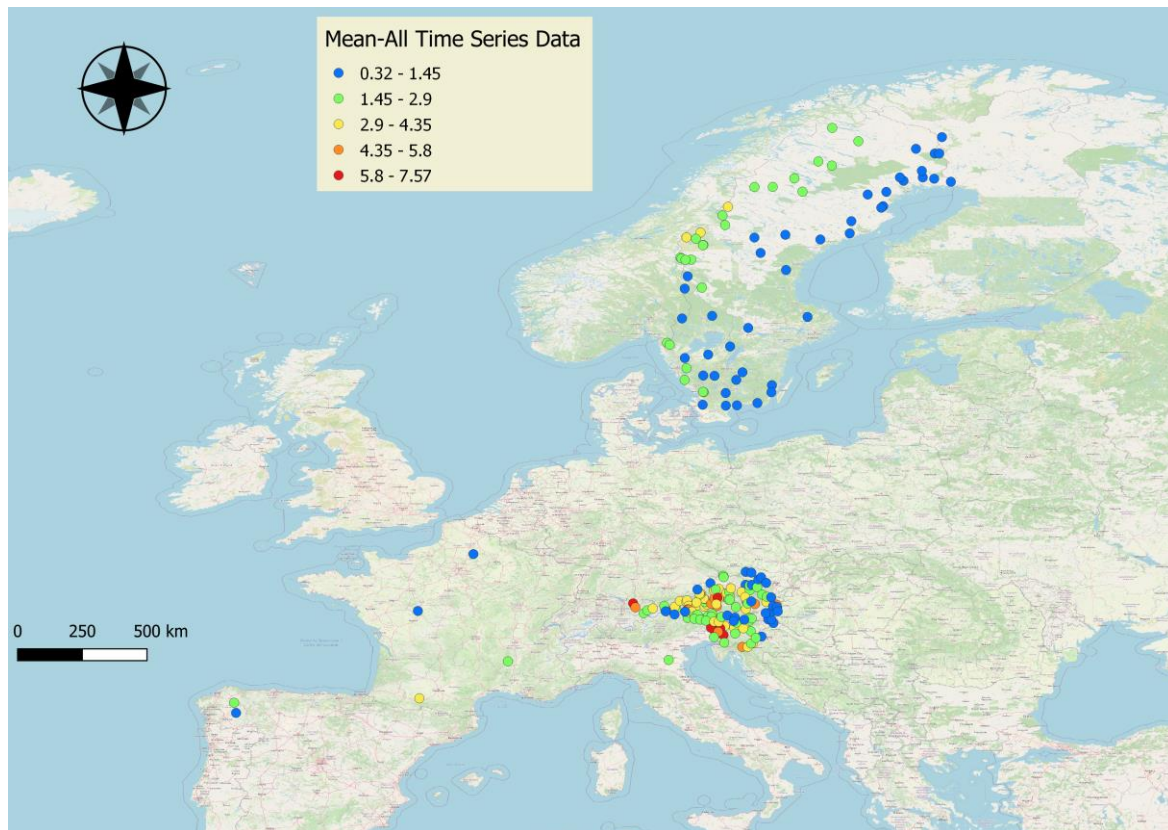


# Περιεχόμενα

## 5 Αποτελέσματα Ανάλυσης

# ΣΤΑΤΙΣΤΙΚΑ ΜΕΓΕΘΗ (1/3)

## Μέση τιμή Ημερήσιας Απορροής (Mean)

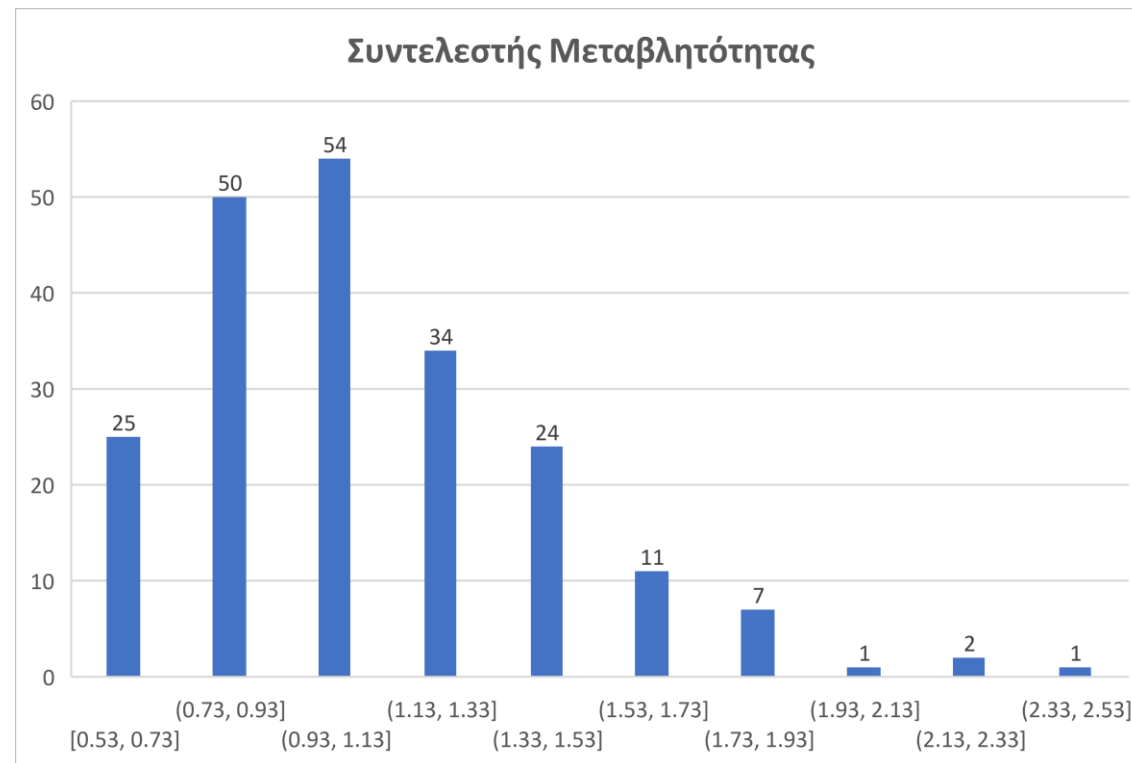
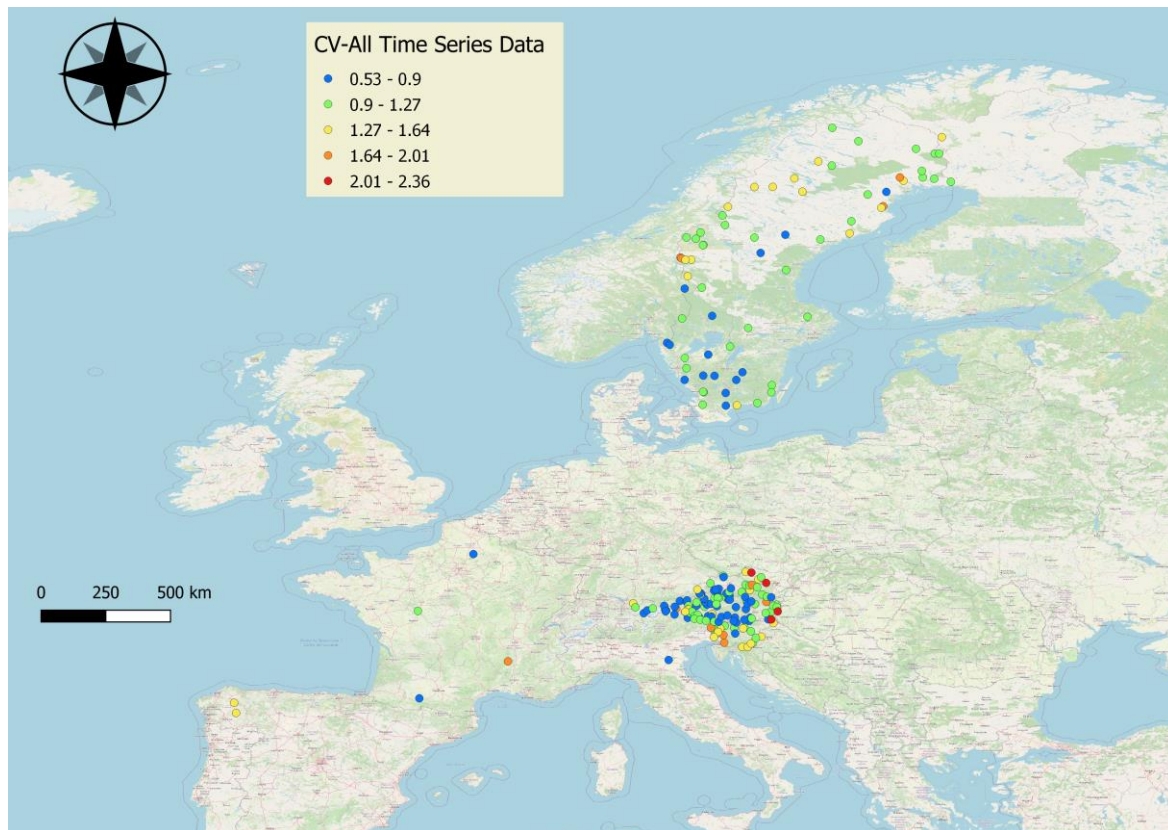


Μεγάλο εύρος υψών μέσης απορροής ετήσια 115mm – 2725mm

# ΣΤΑΤΙΣΤΙΚΑ ΜΕΓΕΘΗ (2/3)



## Συντελεστής Μεταβλητότητας (CV)

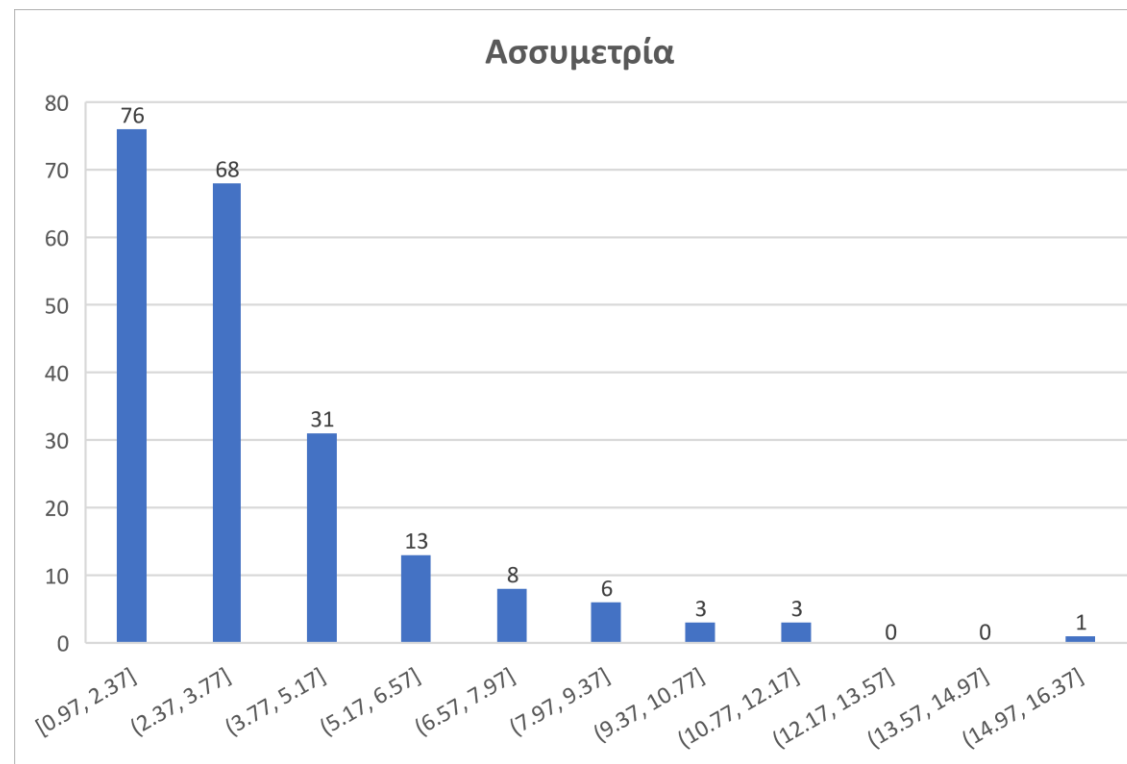
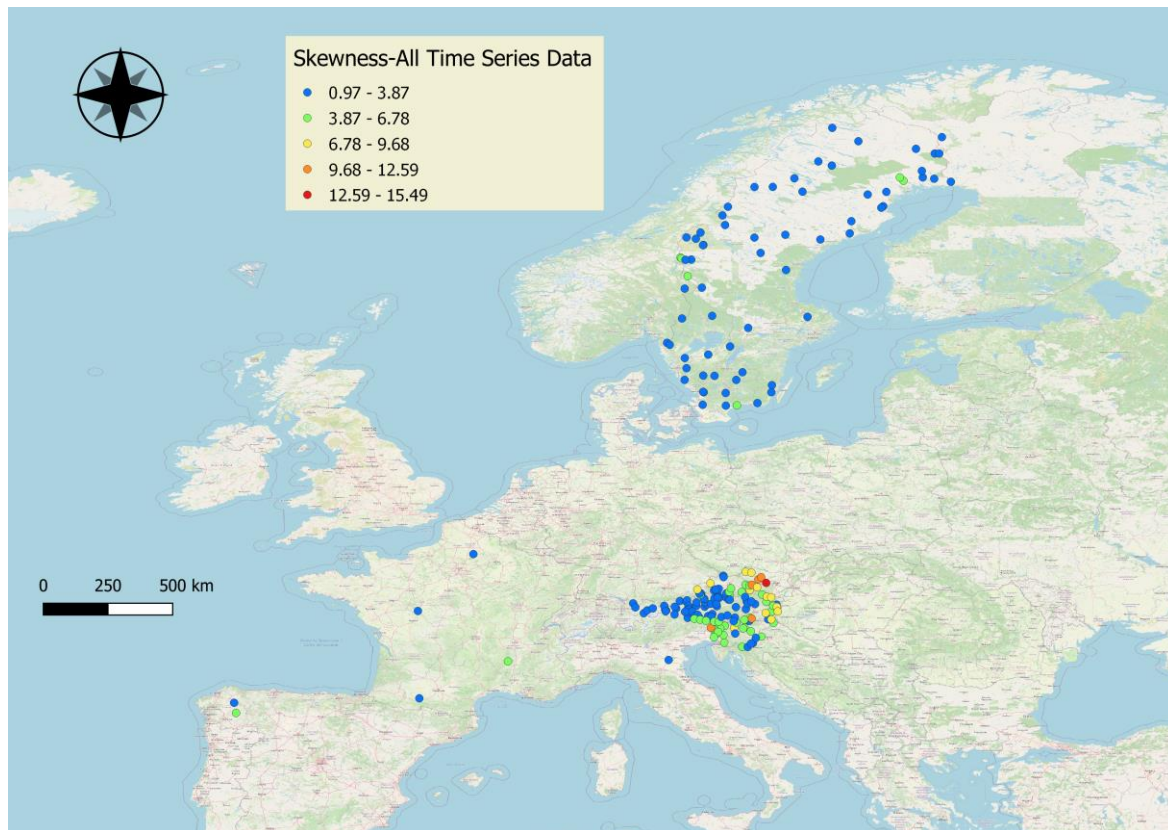




# ΣΤΑΤΙΣΤΙΚΑ ΜΕΓΕΘΗ (3/3)

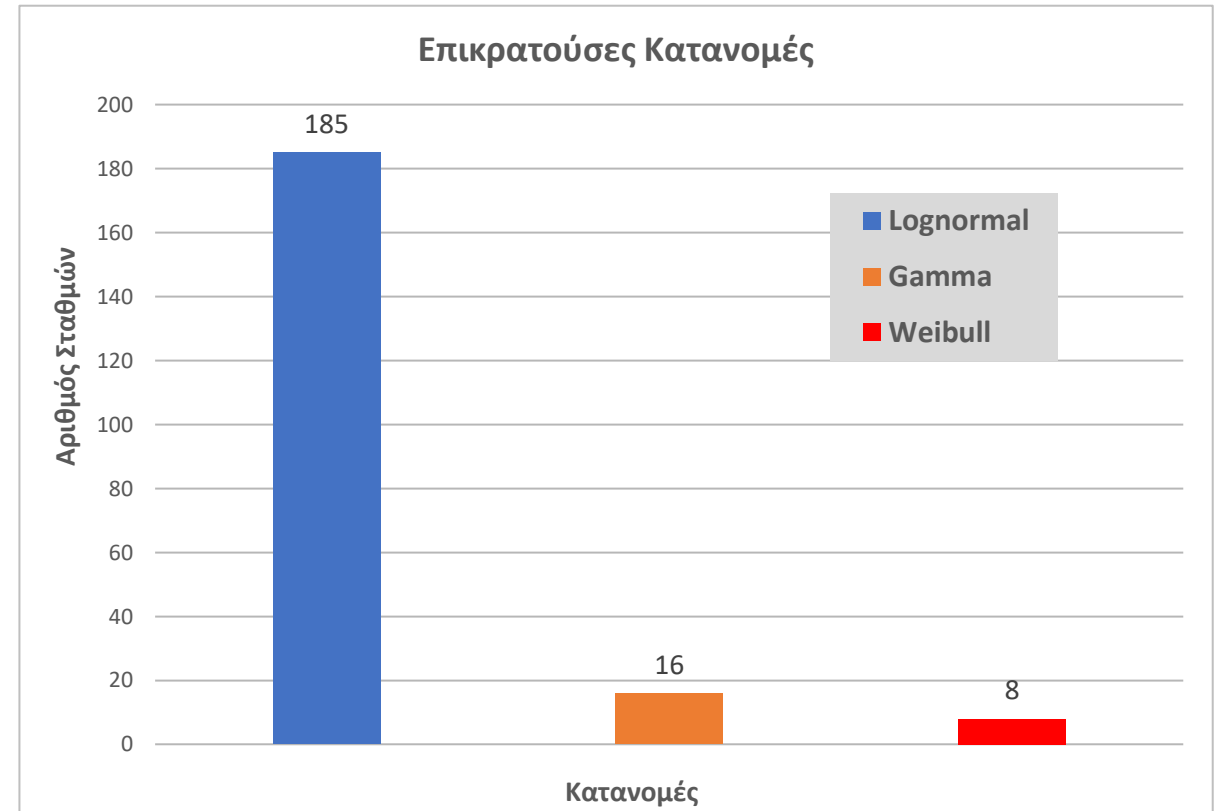
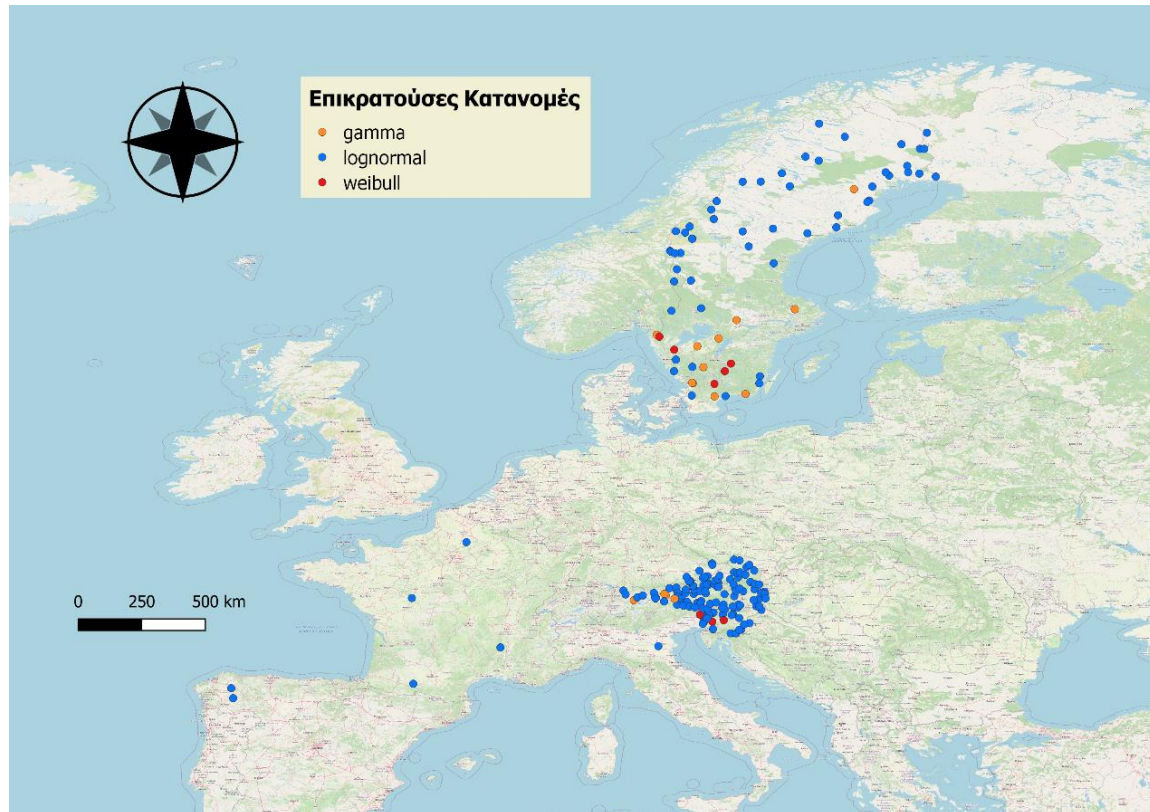


## Ασυμμετρία (skewness)



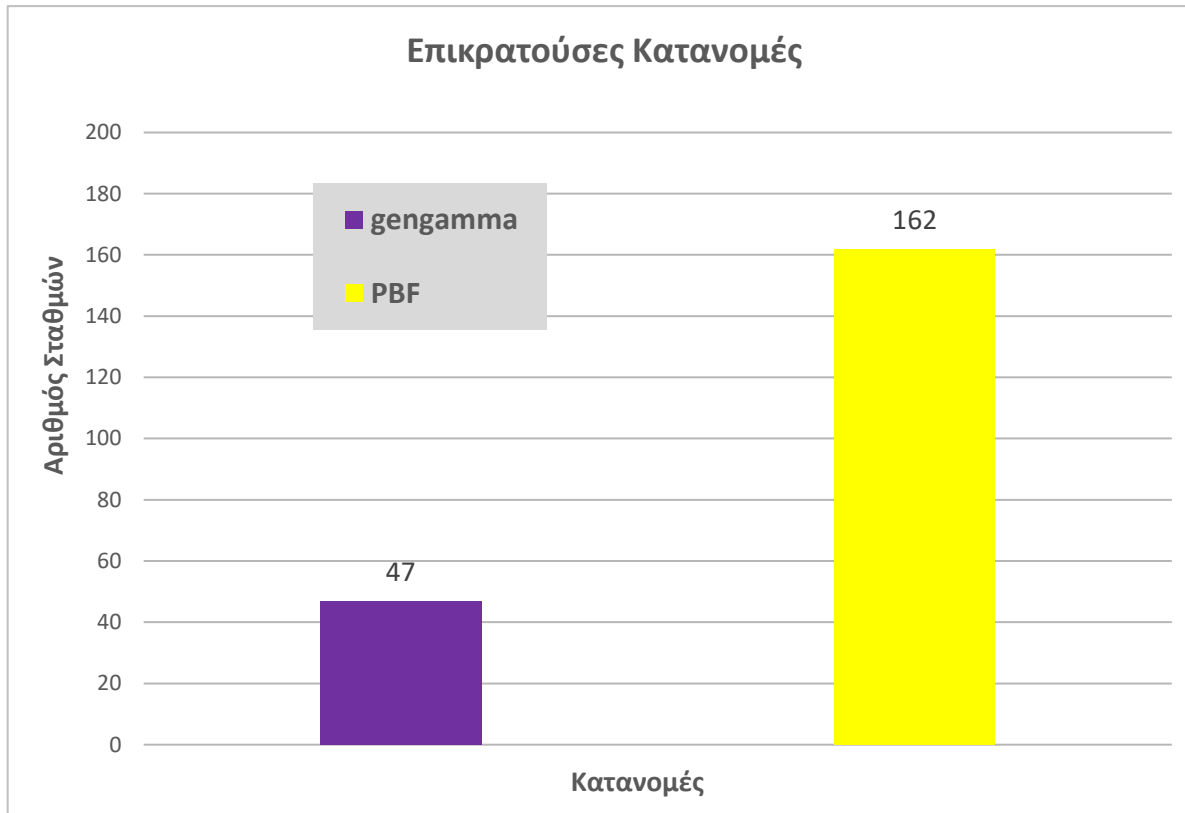
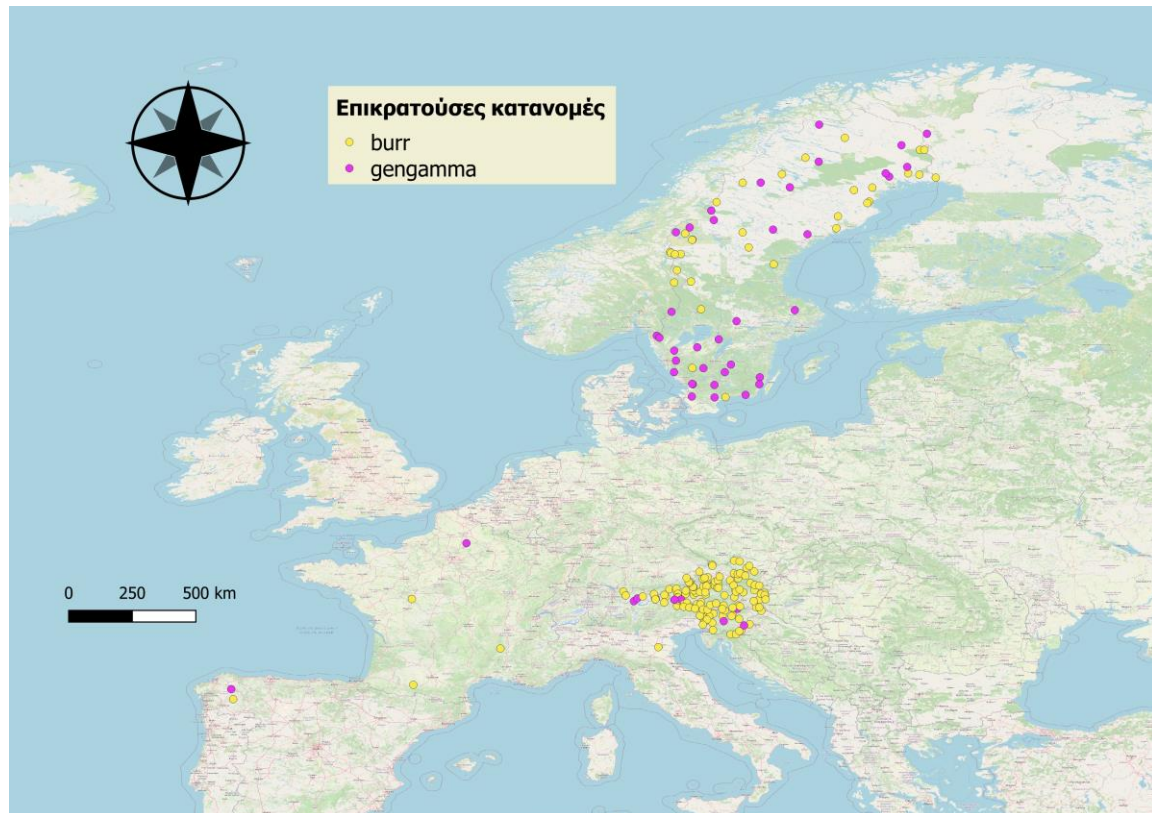
# ΠΡΟΣΑΡΜΟΓΗ ΚΑΤΑΝΟΜΩΝ (1/4)

Προσαρμογή Κατανομών Lognormal, Gamma, Weibull



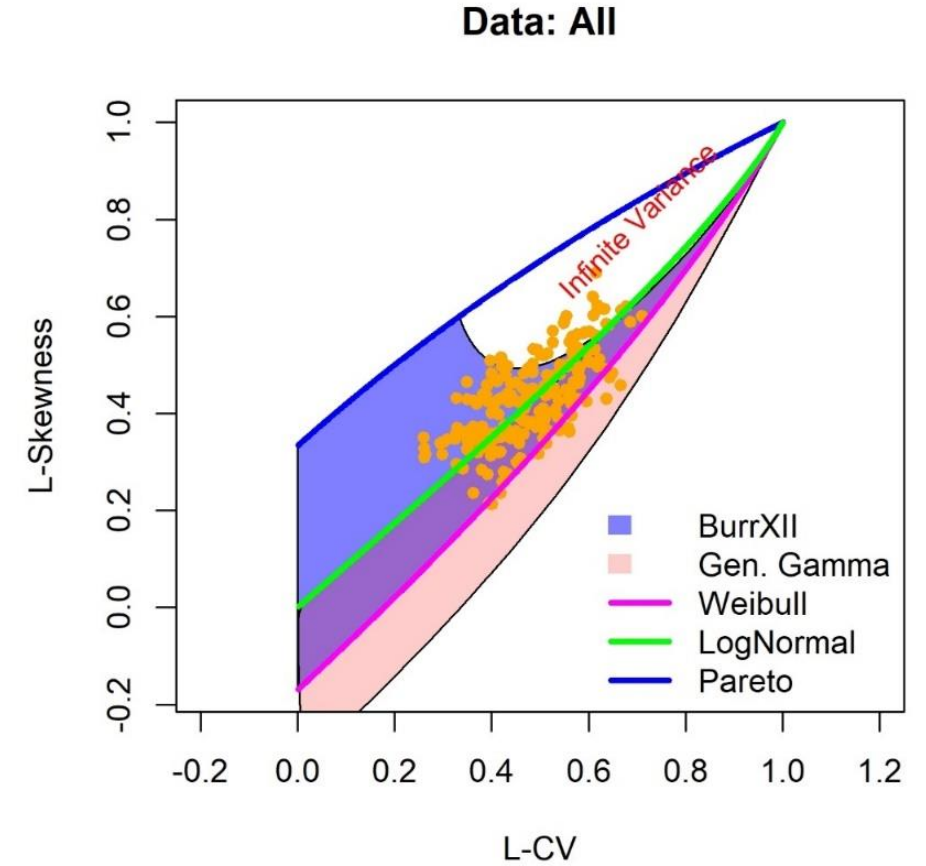
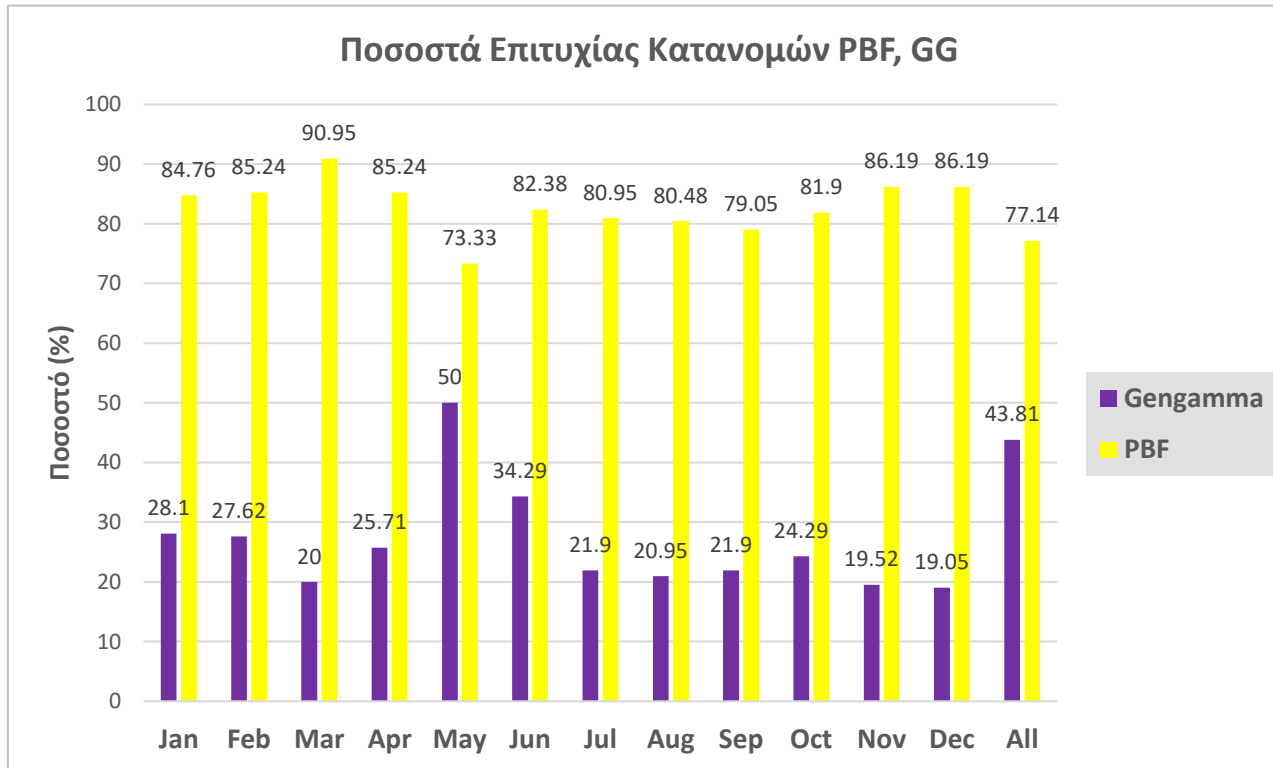
# ΠΡΟΣΑΡΜΟΓΗ ΚΑΤΑΝΟΜΩΝ (2/4)

Προσαρμογή Κατανομών PBF, Generalized Gamma



# ΠΡΟΣΑΡΜΟΓΗ ΚΑΤΑΝΟΜΩΝ (3/4)

Διαγράμματα Λόγου Ροπής (L - ratio diagrams)

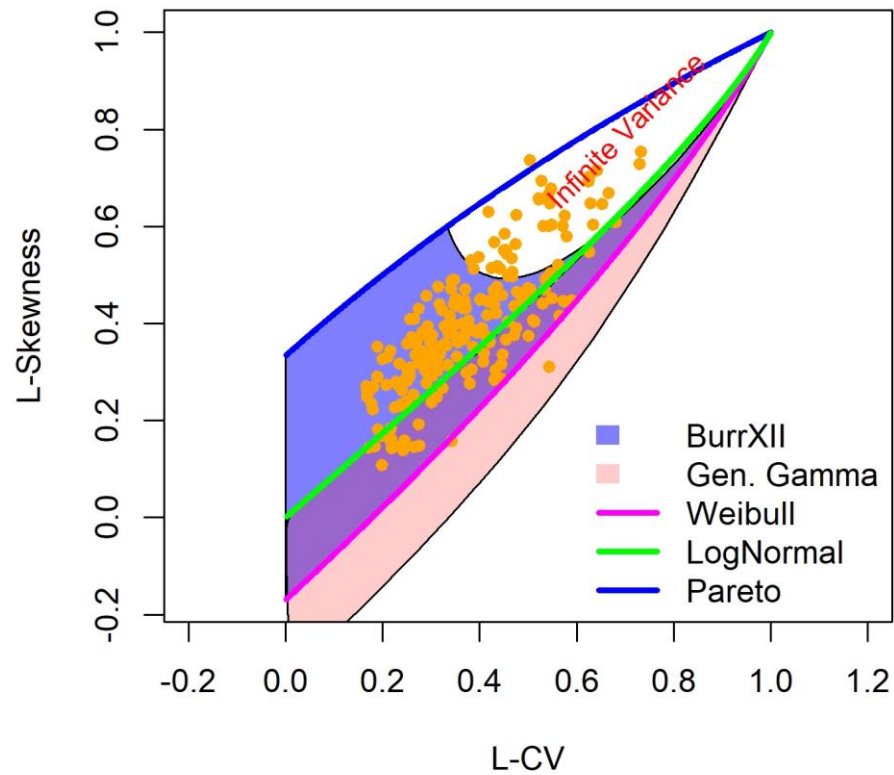


# ΠΡΟΣΑΡΜΟΓΗ ΚΑΤΑΝΟΜΩΝ (4/4)

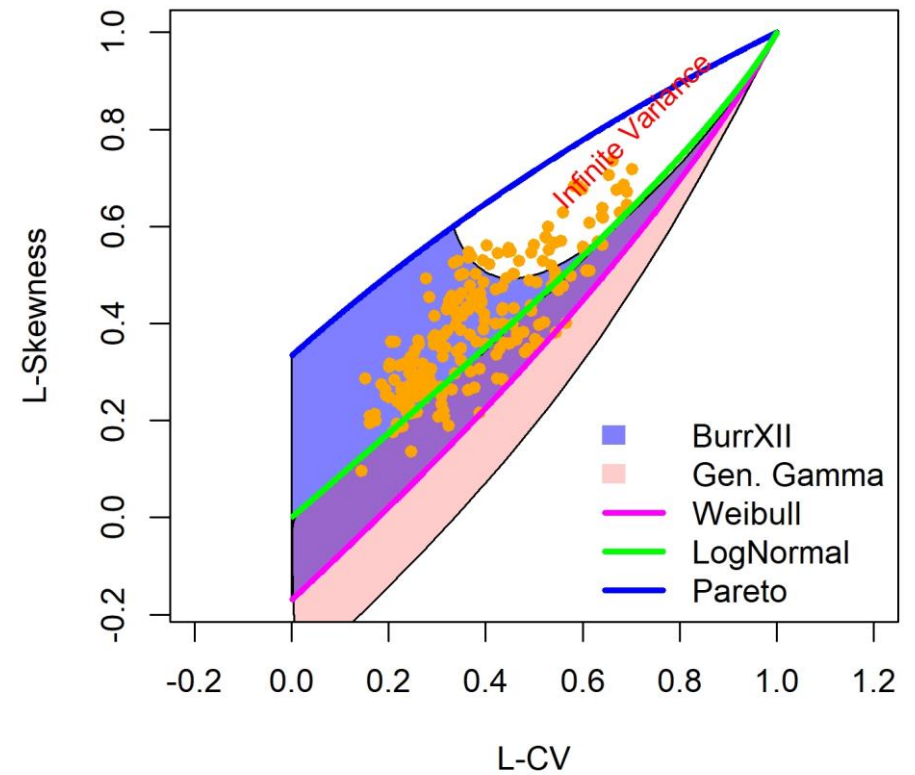
Διαγράμματα Λόγου Ροπής (L - ratio diagrams)



Data: Aug



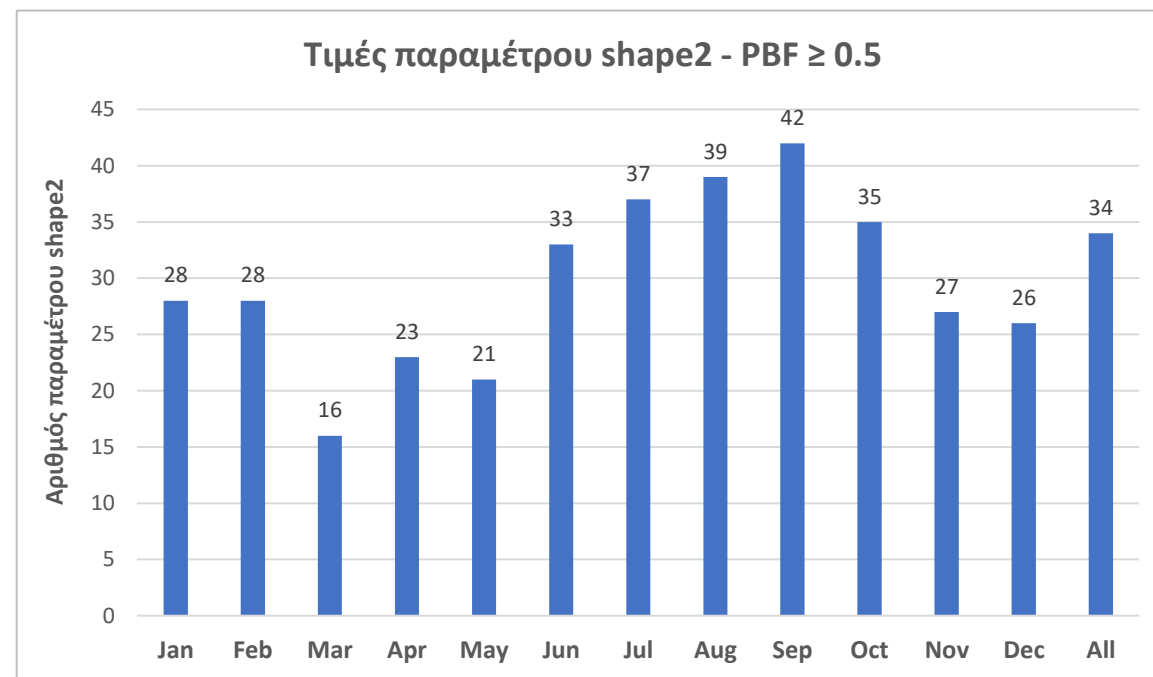
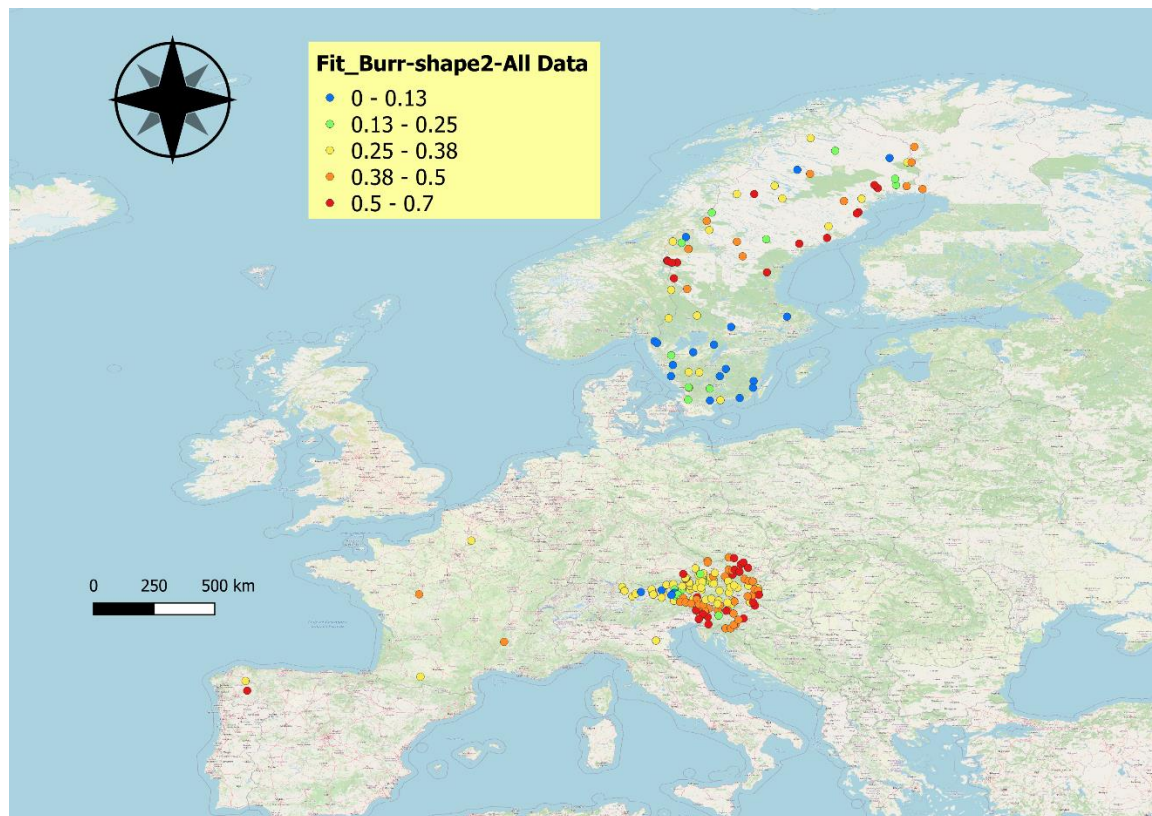
Data: Sep



# ΕΚΤΙΜΗΣΗ ΠΑΡΑΜΕΤΡΩΝ (1/1)



## Εκτίμηση παραμέτρων κατανομής PBF



# Περιεχόμενα

6 Συμπεράσματα

## ΣΥΜΠΕΡΑΣΜΑΤΑ



- Από τη σύγκριση των δι-παραμετρικών κατανομών με κριτήριο το μέσο Τετραγωνικό Σφάλμα επικράτησε η κατανομή lognormal με ποσοστό 88.5%.
- Από τη σύγκριση των τρι-παραμετρικών κατανομών με κριτήριο το μέσο Τετραγωνικό Σφάλμα επικράτησε η κατανομή PBF με ποσοστό 77.5%.
- Από τα διαγράμματα λόγου ροπής στην κλίμακα L προέκυψε ότι η κατανομή PBF προσαρμόζεται καλύτερα στα δεδομένα με ποσοστό 77%.
- Σε σταθμούς στην Αυστρία και στη Σλοβενία παρατηρούνται υψηλές τιμές μέσης απορροής λόγω κλιματικού καθεστώσ Cfb ενώ στη Σκανδιναβία παρατηρούνται συστηματικά χαμηλές τιμές απορροής.
- Μεγαλύτερες τιμές CV παρατηρούνται στην εδαφική επικράτεια της Αυστρίας και συγκεκριμένα κοντά στην Βιέννη αλλά και σε κάποιους σταθμούς στην περιοχή της Σλοβενίας.



# ΣΥΜΠΕΡΑΣΜΑΤΑ



## Προτάσεις για μελλοντική έρευνα

---

- Η εύρεση χρονοσειρών με μήκος παρατηρήσεων μεγαλύτερο των 60 ετών. Αυτό προβλέπει την επέκταση της βάσης δεδομένων GCHN-Daily με χρονοσειρές που θα καλύπτουν μεγαλύτερο χρονικό εύρος παρατηρήσεων ή ακόμα της προμήθειας δεδομένων από διαφορετικές πηγές.
- Η επέκταση των δειγμάτων μας σε περισσότερες περιοχές της Ευρώπης με άλλα κλιματικά χαρακτηριστικά, καθώς οι αναλύσεις αφορούν αποκλειστικά σε παρατηρήσεις κυρίως στην Κεντρική και Βόρεια Ευρώπη.
- Η αξιοποίηση των αποτελεσμάτων τέτοιων στατιστικών αναλύσεων σε έργα υδρολογικού σχεδιασμού, όπως τα Μικρά Υδροηλεκτρικά Έργα (Μ.Υ.Η.Ε.) με στόχο τον σωστό και ακριβή σχεδιασμό τους καθώς και σε Μελέτες Περιβαλλοντικών Επιπτώσεων (ΜΠΕ).
- Ενδιαφέρον θα ήταν και η εξαγωγή κάποιων περιοχικών σχέσεων ανάλογα με τα φυσιογραφικά χαρακτηριστικά κάθε περιοχής.
- Η επανάληψη της στατιστικής ανάλυσης και των υπολογισμών σε βάθος χρόνου, όταν θα είναι διαθέσιμα πολύ μεγαλύτερα μήκη καταγραφών του δείγματος και η σύγκριση των εξαχθέντων αποτελεσμάτων.



1. Ioannis Tsoukalas, Christos Markopoulos, and Demetris Koutsoyiannis, 2018, Simulation of Stochastic Processes Exhibiting Any-Range Dependence and Arbitrary Marginal Distributions, Department of Water Resources, School of Civil Engineering, National Technical University of Athens, Zographou, Greece
2. Koutsoyiannis Dimitris, Statistical Hydrology, Edition 4, 312 pages, National Technical University of Athens, Athens, 1997
3. Koutsoyiannis Demetris, 2017, Article: Entropy Production in Stochastics, Department of Water Resources, School of Civil Engineering, National Technical University of Athens, 15780 Athina, Greece
4. Panagiotis Kossieris, and Christos Markopoulos, 2017, Article: Exploring the Statistical and Distributional Properties of Residential Water Demand at Fine Time Scales, Department of Water Resources, School of Civil Engineering, National Technical University of Athens, Heron Polytechniou 5, GR-157 80 Zographou, Greece
5. T. Iliopoulou, October 2020, Stochastic investigation of hydrological extremes: influence of temporal variability and dependence, PhD thesis, 237 pages, Department of Water Resources and Environmental Engineering – National Technical University of Athens
6. Κοκολάκης Γ. - Φουσκάκης Δ., 2009, Στατιστική Θεωρία & Εφαρμογές (σελ. 370), Εκδόσεις Συμεών, Αθήνα, ΕΜΠ
7. Κουτσογιάννης, Δ., και Θ. Ξανθόπουλος, Τεχνική Υδρολογία, Έκδοση 3, (418 σελ.), Εθνικό Μετσόβιο Πολυτεχνείο, Αθήνα, 1999
8. Κουτσογιάννης, Δ., 1996. Στατιστική Υδρολογία, Έκδοση 3, Εθνικό Μετσόβιο Πολυτεχνείο, Αθήνα, σελ. 76.
9. Κουτσογιάννης Δ., 2013. «Σημειώσεις Στοχαστικών Μεθόδων στους Υδατικούς Πόρους», Επανεκδοση, Εθνικό Μετσόβιο Πολυτεχνείο, Αθήνα.
10. Μαρία Νέζη, Πολυμεταβλητή στατιστική ανάλυση ακραίων βροχοπτώσεων και παροχών σε δείγμα 400 λεκανών απορροής των ΗΠΑ από τη βάση δεδομένων MOPEX, Διπλωματική Εργασία, 97 σελίδες, Εθνικό Μετσόβιο Πολυτεχνείο, Σχολή Πολιτικών Μηχανικών, Τομέας Υδατικών Πόρων και Περιβάλλοντος
11. Μιμίκου Μ.Α., Ε.Α. Μπαλτάς, Τεχνική Υδρολογία, 5η Έκδοση, (416 σελ.), Εθνικό Μετσόβιο Πολυτεχνείο, Αθήνα, 2012
12. Φουσκάκης Δ., 2013, Ανάλυση Δεδομένων με Χρήση της R (σελ. 504), Εκδόσεις Τσότρας Αθήνα, ΕΜΠ

## Ειδικές Ευχαριστίες...



- ✓ Ιωάννης Τσουκαλάς (Πολιτικός Μηχανικός, MSc, Δρ. Μηχανικός)
- ✓ Παναγιώτης Κοσσιέρης (Πολιτικός Μηχανικός, MSc, Δρ. Μηχανικός)
- ✓ Τζωρτζίνα Σακκή (Πολιτικός Μηχανικός, Υποψ. Δρ. Μηχανικός)
- ✓ Άνυ Ηλιοπούλου (Πολιτικός Μηχανικός, MSc, Δρ. Μηχανικός)

ΤΕΛΟΣ ΠΑΡΟΥΣΙΑΣΗΣ

ΕΥΧΑΡΙΣΤΩ ΠΟΛΥ ΓΙΑ ΤΗΝ  
ΠΡΟΣΟΧΗ ΣΑΣ !

