



Τμήμα Πολιτικών Μηχανικών
Πανεπιστήμιο Θεσσαλίας

ΥΔΡΟΛΟΓΙΚΗ ΠΡΟΣΟΜΟΙΩΣΗ ΚΑΙ ΠΡΟΓΝΩΣΗ

Ενότητα 2: Στατιστική και πιθανοτική ανάλυση
ακραίων υδρολογικών τιμών

2.2. Περιοχική Ανάλυση Ακραίων Τιμών

Καθ. Αθανάσιος Λουκάς

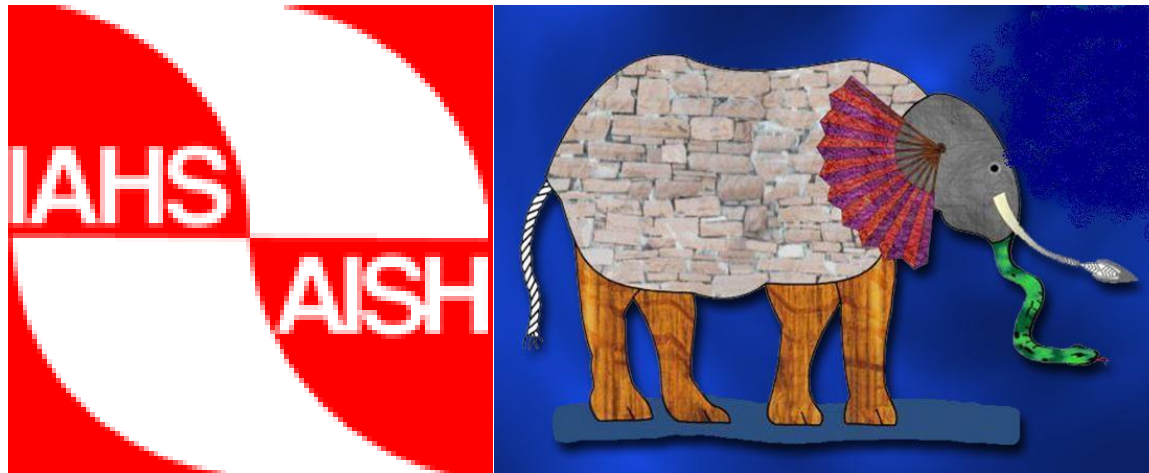
Εργαστήριο Υδρολογίας και Ανάλυσης Υδατικών Συστημάτων

Τμήμα Πολιτικών Μηχανικών

Πολυτεχνική Σχολή

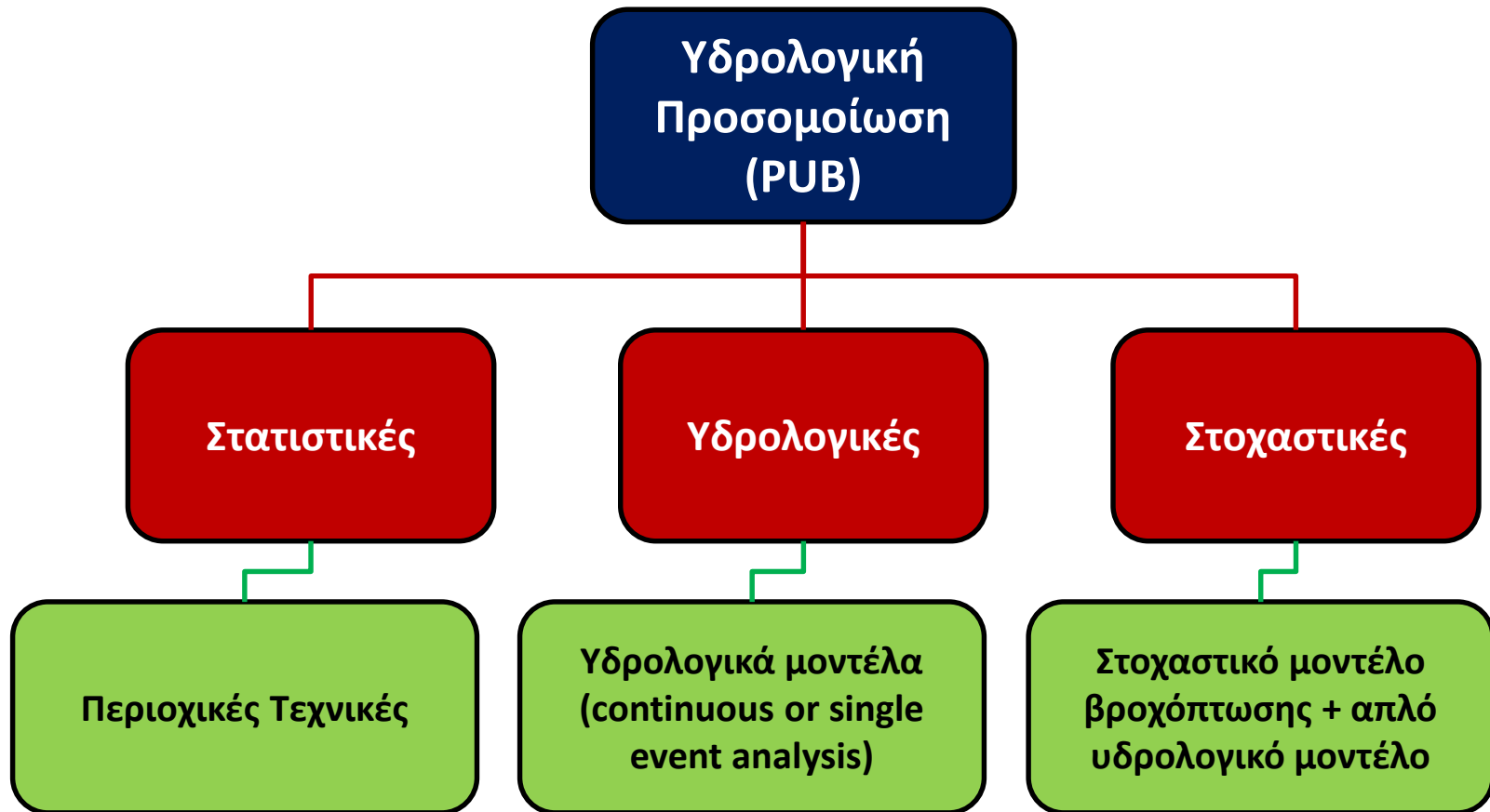
ΜΕΘΟΔΟΙ ΥΔΡΟΛΟΓΙΚΗΣ ΠΡΟΣΟΜΟΙΩΣΗΣ ΠΛΗΜΜΥΡΩΝ

- **Λεκάνες απορροής χωρίς υδρομετρήσεις**
 - → **ungauged watersheds**
- **IAHS Decade (2003-2012) on Prediction in Ungauged Basins (PUB)** (Hrachowitz et al., 2013)
 - <http://iahs.info/pub/index.php>

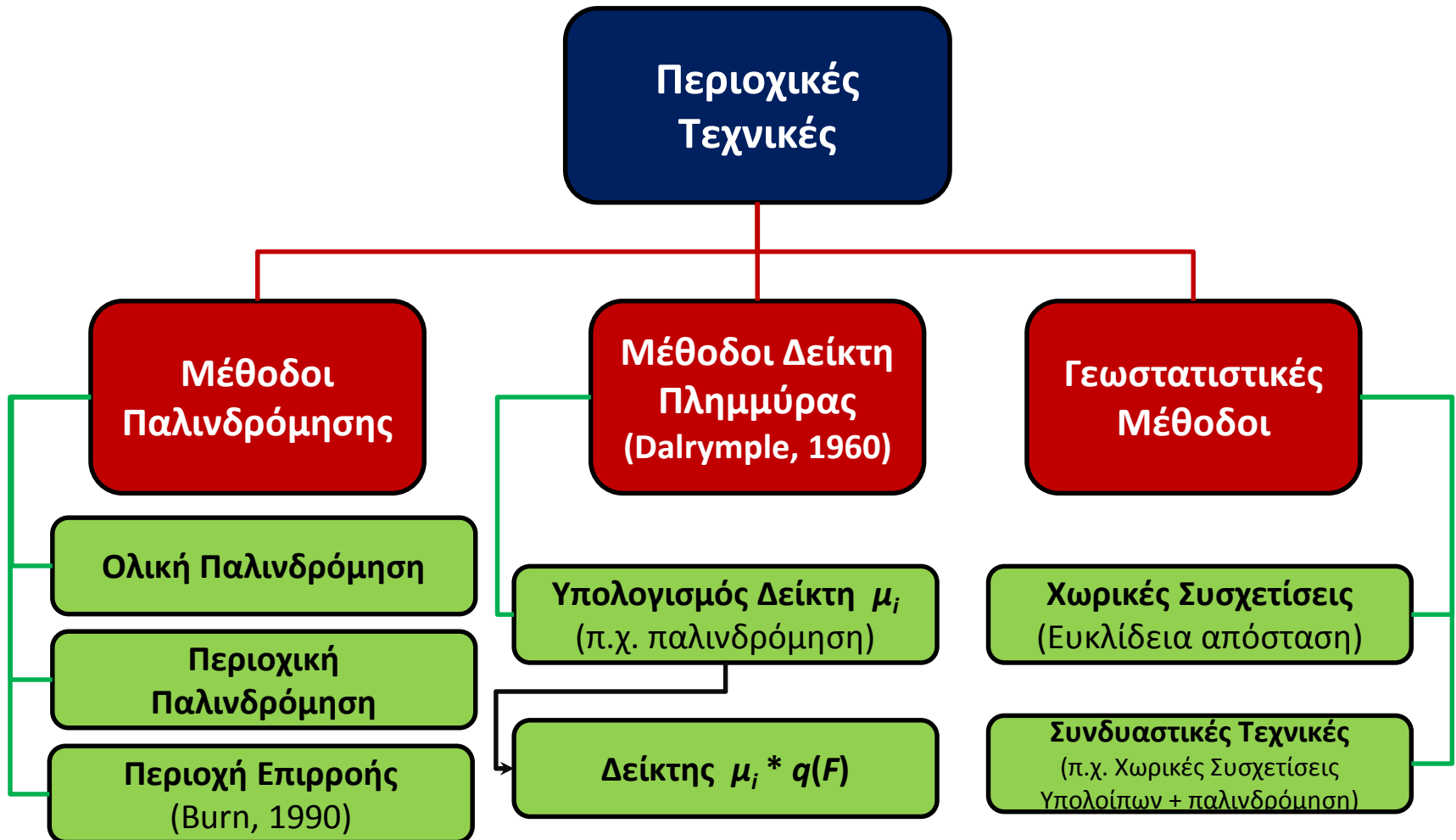


Hrachowitz, M., Savenije, H.H.G., Blöschl, G., McDonnell, J.J., Sivapalan, M., Pomeroy, J.W., Arheimer, B., Blume, T., Clark, M.P., Ehret, U., Fenicia, F., Freer, J.E., Gelfan, A., Gupta, H.V., Hughes, D.A., Hut, R.W., Montanari, A., Pande, S., Tetzlaff, D., Troch, P.A., Uhlenbrook, S., Wagener, T., Winsemius, H.C., Woods, R.A., Zehe, E., Cudennec, C.: A decade of Predictions in Ungauged Basins (PUB) - a review, *Hydrolog. Sci. J.* 58(6), 1198-1255, 2013.

ΜΕΘΟΔΟΙ ΥΔΡΟΛΟΓΙΚΗΣ ΠΡΟΣΟΜΟΙΩΣΗΣ ΠΛΗΜΜΥΡΩΝ



ΣΤΑΤΙΣΤΙΚΕΣ ΜΕΘΟΔΟΙ ΥΔΡΟΛΟΓΙΚΗΣ ΑΝΑΛΥΣΗΣ ΠΛΗΜΜΥΡΩΝ



Burn, D. H.: An appraisal of the “region of influence” approach to flood frequency analysis, Hydrolog. Sci. J., 35, 149–165, 1990.

Dalrymple, T.: Flood frequency analysis, Water Supply Paper 1543A, US Geological Survey, Reston, Virginia, 1960.

ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΑΚΡΑΙΩΝ ΤΙΜΩΝ

- Εισαγωγή στη περιοχική ανάλυση συχνότητας ακραίων τιμών
- Επιλογή του ακραίου δείγματος και αρχική επιλογή «περιοχών» εφαρμογής
- Υπολογισμός μέτρων ασυμφωνίας και ανομοιογένειας/ετερογένειας κατά Hosking & Wallis (1997)
- Εφαρμογή της μεθόδου του δείκτη πλημμύρας (Index Flood method)
- Επιλογή της κατάλληλης περιοχικής κατανομής ακραίων τιμών
- Περιοχική Ανάλυση Συχνότητας Πλημμυρών (Regional Flood Frequency Analysis)
- Παραδείγματα – Εφαρμογές Περιοχικής Ανάλυσης Συχνότητας

Hosking, J.R.M. and Wallis, J.R. «Regional Frequency Analysis : An Approach based on L-Moments», Cambridge University Press, 238 pp., 1997.



ΕΙΣΑΓΩΓΗ

Ο προσδιορισμός των συναρτήσεων κατανομής των ακραίων τιμών των υδρομετεωρολογικών και υδρολογικών μεταβλητών είναι ιδιαίτερα σημαντικός για:

Την εκτίμηση των πλημμυρών

Την εκτίμηση της πλημμυρικής παροχής

Το σχεδιασμό υδροτεχνικών έργων

Καθίσταται ιδιαίτερα δύσκολος σε περίπτωση που οι διαθέσιμες χρονοσειρές είναι **σύντομης διάρκειας**. Περιπλέκεται τόσο ο προσδιορισμός της κατάλληλης συνάρτησης κατανομής, όσο και η εκτίμηση των παραμέτρων της επιλεγμένης συνάρτησης.

Το γεγονός αυτό έχει σημαντική επίπτωση **στην αβεβαιότητα πρόβλεψης** που ενυπάρχει στην εκτίμηση του επιπέδου επαναφοράς, οδηγώντας σε μεγάλο εύρος διαστημάτων εμπιστοσύνης που πολλαπλασιάζεται με την αύξηση της περιόδου επαναφοράς.

Η μείωση της αβεβαιότητας στις μακροπρόθεσμες εκτιμήσεις των υδρολογικών μεταβλητών (ύψος και ένταση βροχόπτωσης, πλημμυρική παροχή) αποτελεί το βασικό κίνητρο.



ΕΙΣΑΓΩΓΗ

Κύριος στόχος των μεθόδων που βασίζονται στην περιοχική ανάλυση:

Η επίτευξη πιο ικανών εκτιμητριών των ποσοστιαίων σημείων των μελετώμενων μεγεθών σε μια δεδομένη θέση

Διαθέσιμες πληροφορίες από διάφορες θέσεις ή σταθμούς μέτρησης με κοινά χαρακτηριστικά στοιχεία



ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΑΚΡΑΙΩΝ ΤΙΜΩΝ

Περιοχική Ανάλυση Συχνότητας (Regional Frequency Analysis - RFA)

- Λόγω έλλειψης δεδομένων, οι περιοχικές εκτιμήσεις σε σχέση με τη σημειακή εκτίμηση είναι προτιμητέες στα άκρα των θεωρητικών κατανομών (distribution tails - extremes).
 - Ακριβέστερες εκτιμήσεις
- “Χρησιμοποιεί χώρο αντί για χρόνο” με δεδομένα από παρακείμενους σταθμούς για την εκτίμηση ποσοστημορίων (quantiles) των εξεταζόμενων μεταβλητών (π.χ. παροχή) σε ένα δοσμένο σημείο
- Θεωρητικές κατανομές σε σημεία χωρίς μετρήσεις (π.χ. λεκάνες απορροής) μπορούν εύκολα να υπολογιστούν
- Η περιοχή μελέτης πρέπει να είναι ομοιογενής (Stedinger et al. 1993)
- Πολλές εφαρμογές περιοχικής εκτίμησης (U.S., Canada, UK, Africa, Belgium, κ.α.)

Stedinger, J.R., Vogel, R.M. & Foufoula-Georgiou, E., 1993. Frequency analysis of extreme events. Chapter 18. In: Handbook of Applied Hydrology, Maidment, D.A. (Eds.), McGraw-Hill, New York.



ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΑΚΡΑΙΩΝ ΤΙΜΩΝ

- Η **Περιοχική Ανάλυση Συχνότητας (Regional Frequency Analysis)** καθίσταται απαραίτητη όταν τα διαθέσιμα δεδομένα μιας θέσης/σταθμού είναι ανεπαρκή για μια αξιόπιστη εκτίμηση των ποσοστιαίων σημείων μιας μεταβλητής.
- Μια «**περιοχή**» πρέπει να ελέγχεται στατιστικά όσον αφορά στην ομοιογένεια των στοιχείων της (Stedinger et al., 1993).
 - Στην πραγματικότητα καταδεικνύεται μια ομάδα σταθμών που ελέγχονται σε σχέση με την ομοιογένειά τους, προκειμένου να σχηματίσουν την μονάδα της «περιοχής» που απαιτεί η Περιοχική Ανάλυση Συχνότητας.
- Ο όρος «**περιοχή**» εμπεριέχει ένα σύνολο γειτονικών σταθμών, ωστόσο η γεωγραφική εγγύτητα δεν είναι απαραίτητως ένας δείκτης ομοιογένειας των κατανομών συχνότητας της μεταβλητής στις διάφορες θέσεις.
 - Όλα τα διαθέσιμα στοιχεία στους διάφορους σταθμούς αυτής της περιοχής αναλύονται για να προσδιορίσουν τα χαρακτηριστικά της κατανομής συχνότητας της περιοχής (π.χ. για ανάλυση πλημμυρικής παροχής χαρακτηριστικά όπως υψόμετρο, έκταση, κλίση, τύπος εδάφους, μέση ετήσια βροχόπτωση χρησιμοποιούνται)



ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΑΚΡΑΙΩΝ ΤΙΜΩΝ

Η μέθοδος του Δείκτη Πλημμύρας (Index Flood method)

- Οι μετρήσεις (παρατηρήσεις) σε μία θέση/σταθμό μίας ομοιογενούς περιοχής είναι ομοιόμορφα κατανεμημένες (identically distributed)
- Οι μετρήσεις (παρατηρήσεις) σε μία θέση/σταθμό είναι ανεξάρτητες μεταξύ τους (serially independent values)
- Οι μετρήσεις (παρατηρήσεις) σε διαφορετικές θέσεις/σταθμούς είναι ανεξάρτητες μεταξύ τους (independent values)
- Οι θεωρητικές κατανομές σε διαφορετικές θέσεις/σταθμούς είναι ίδιες εκτός από ένα συντελεστή κλίμακας (scale factor)
- Ο μαθηματικός τύπος της καμπύλης περιοχικής αύξησης ή ανάπτυξης (regional growth curve) έχει δομηθεί σωστά



ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΑΚΡΑΙΩΝ ΤΙΜΩΝ

Βήματα Περιοχής Ανάλυσης (Hosking and Wallis, 1997):

- ❑ Προεπεξεργασία και έλεγχος συνέπειας των δεδομένων (screening of the data)
- ❑ Έλεγχος της περιοχικής ομοιογένειας
- ❑ Κανονικοποίηση (normalization) των ετήσιων μεγίστων τιμών της υπό-εξέταση μεταβλητής (π.χ. παροχή) με τη χρήση ενός δείκτη πλημμύρας (index measure for flood/storm)
- ❑ Προσαρμογή θεωρητικής κατανομής (π.χ. GEV, EVI) σε κάθε δείγμα και υπολογισμός παραμέτρων κατανομής
- ❑ Συνδυασμός μεθόδων υπολογισμού των παραμέτρων για περιοχικές εκτιμήσεις των παραμέτρων
- ❑ Χρήση των περιοχικών κατανομών (που υπολογίστηκαν προηγουμένως) με το δείκτη πλημμύρας για υπολογισμό των σημειακών κατανομών



ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΑΚΡΑΙΩΝ ΤΙΜΩΝ

L-Ροπές και Περιοχική Ανάλυση Συχνότητας (L-moments and RFA)

Οι L – Ροπές μπορούν να εφαρμοστούν σε τέσσερα στάδια της περιοχικής εκτίμησης συχνότητας:

- Προεπεξεργασία και έλεγχος συνέπειας των δεδομένων
- Έλεγχος και αναγνώριση ομοιογενών περιοχών
- Προσαρμογή θεωρητικής κατανομής και υπολογισμός παραμέτρων κατανομής ομοιογενών περιοχών

- The k-th L-moment is denoted λ_k

- The L-moment ratios are:

- 1) The L-coefficient of variation (L-CV) λ_2/λ_1

- 2) The L-skewness $\tau_3 = \lambda_3/\lambda_2$

- 3) The L-kurtosis $\tau_4 = \lambda_4/\lambda_2$



ΑΡΧΙΚΗ ΕΠΙΛΟΓΗ «ΠΕΡΙΟΧΩΝ» ΕΦΑΡΜΟΓΗΣ

Η μέθοδος των L-Ροπών
(Hosking, 1990)

Οι L-ροπές ορίζονται ως ο γραμμικός
συνδυασμός των αναμενόμενων
τιμών των διατεταγμένων
στατιστικών

- Οι L-ροπές μπορούν να εκφραστούν ως συνάρτηση των πιθανοτικά σταθμισμένων ροπών.
- Οι πιθανοτικά σταθμισμένες ροπές μιας τυχαίας μεταβλητής X με αθροιστική συνάρτηση πυκνότητας πιθανότητας $F(x)$, ορίζονται ως :

$$\beta_r = E\{X[F_X(x)]^r\}$$

- Οι πρώτες τέσσερις L-ροπές εκφρασμένες με όρους πιθανοτικά σταθμισμένων ροπών δίνονται από τις σχέσεις:

$$\lambda_1 = \beta_0, \lambda_2 = 2\beta_1 - \beta_0, \lambda_3 = 6\beta_2 - 6\beta_1 + \beta_0, \lambda_4 = 20\beta_3 - 30\beta_2 + 12\beta_1 - \beta_0$$

- Οι αναλογικές L-ροπές, ο L-συντελεστής διακύμανσης (L-CV (t)), η L-ασυμμετρία (L-skewness (t_3)) και η L-κύρτωση (L-kurtosis (t_4)):

$$t = \lambda_2 / \lambda_1$$

$$t_3 = \lambda_3 / \lambda_2$$

$$t_4 = \lambda_4 / \lambda_2$$

ΕΛΕΓΧΟΣ ΟΜΟΙΟΓΕΝΕΙΑΣ (GUMBEL, 1958)

□ Αφού προσδιοριστεί μια εύλογη «περιοχή» θα πρέπει να υπολογιστεί το μέτρο ομοιογένειας σύμφωνα με τα κριτήρια και τους περιορισμούς του **Gumbel (1958)**

- Ο έλεγχος ομοιογένειας (Gumbel, 1958) εφαρμόζεται σε όλους τους σταθμούς/λεκάνες απορροής που έχουν μετρήσεις παροχής
 - Οι σταθμοί που δεν πληρούν το κριτήριο ομοιογένειας αφαιρούνται από την ανάλυση
 - **Υπόθεση:** η κατανομή **EVI** είναι κατάλληλη για την εκτίμηση των πλημμυρικών παροχών (π.χ. μέγιστες ημερήσιες παροχές ανά έτος) σε όλη την περιοχή
- **Βήματα Υπολογισμού:**
 - Υπολογισμός αναλογίας πλημμύρας με περίοδο επαναφοράς 10 έτη $Q_{T=10}$ προς τη μέση ετήσια πλημμύρα (mean annual flood) για όλους τους σταθμούς και υπολογισμός μέσης τιμής της αναλογίας για όλη την περιοχή
 - Η Μέση Ετήσια πλημμύρα για κάθε σταθμό πολλαπλασιάζεται με την περιοχική μέση αναλογία ώστε να υπολογιστεί μία περιοχική εκτίμηση πλημμύρας με περίοδο επαναφοράς 10 έτη $Q_{regT=10}$. Η περίοδος επαναφοράς T_{10} της περιοχικής πλημμύρας βρίσκεται από την εφαρμογή της EVI για τον κάθε σταθμό

Gumbel, E.J., 1958. "Statistics of Extremes." Columbia University Press. New York, New York.



ΕΛΕΓΧΟΣ ΟΜΟΙΟΓΕΝΕΙΑΣ (GUMBEL, 1958)

□ Αφού προσδιοριστεί μια εύλογη «περιοχή» θα πρέπει να υπολογιστεί το μέτρο ομοιογένειας σύμφωνα με τα κριτήρια και τους περιορισμούς του **Gumbel (1958)**

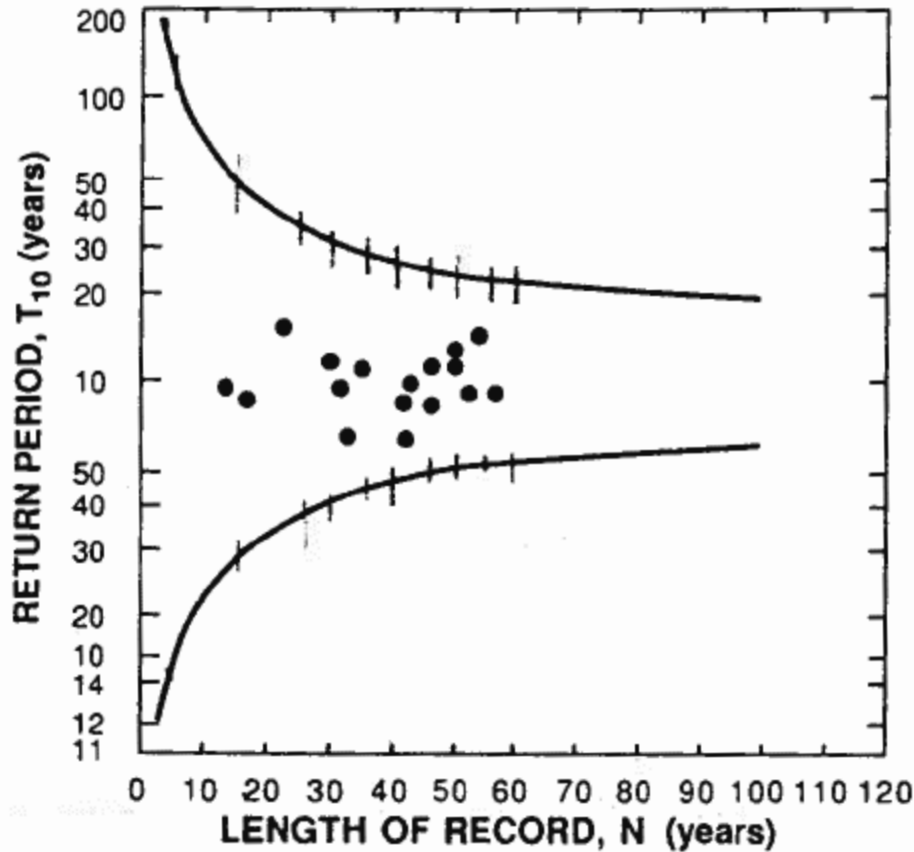
• Βήματα Υπολογισμού:

- Η υπολογισμένη περίοδος επαναφοράς για κάθε σταθμό τοποθετείται σε διάγραμμα σε σχέση με την χρονική περίοδο των διαθέσιμων δεδομένων για κάθε σταθμό, N σε έτη (station length of record)
- Τα 95% διαστήματα εμπιστοσύνης υπολογίζονται προσεγγιστικά με τη σχέση:
$$y_u = 2.25 \pm \frac{6.33}{\sqrt{N}}$$
 όπου y_u η ανοιγμένη μεταβλητή και ίση με:
$$y_u = -\ln \left[-\ln \left(1 - \frac{1}{T_u} \right) \right], T_u$$
 η υπολογισμένη περίοδος επαναφοράς
- Οι σταθμοί που βρίσκονται πάνω και κάτω από τα 95% διαστήματα εμπιστοσύνης αποκλείονται από την περιοχική ανάλυση.
- Όσοι σταθμοί βρίσκονται μέσα στα διαστήματα εμπιστοσύνης αποτελούν ομοιογενή περιοχή και σε αυτούς εφαρμόζονται οι τεχνικές της περιοχικής ανάλυσης πλημμύρας



ΕΛΕΓΧΟΣ ΟΜΟΙΟΓΕΝΕΙΑΣ (GUMBEL, 1958)

Τεστ Ομοιογένειας



Όρια της T_{10} για EVI τεστ ομοιογένειας

Μέγεθος δείγματος N (έτη)	Κάτω Όριο T_L (έτη)	Άνω Όριο T_u (έτη)
10	1.8	70
20	2.8	40
50	4.4	24
100	5.6	18



ΜΕΤΡΑ ΑΣΥΜΦΩΝΙΑΣ ΚΑΙ ΑΝΟΜΟΙΟΓΕΝΕΙΑΣ

□ Αφού προσδιοριστεί μια εύλογη «περιοχή» θα πρέπει να υπολογιστούν τα μέτρα ασυμφωνίας (discordancy measure) και ανομοιογένειας/ετερογένειας (heterogeneity measure) σύμφωνα με τα κριτήρια και τους περιορισμούς των **Hosking & Wallis (1997)**.

Μέτρο Ασυμφωνίας
 D_i

$$D_i = \frac{1}{3} N (u_i - \bar{u})^T A^{-1} (u_i - \bar{u})$$

$$\text{με } A = \sum_{i=1}^N (u_i - \bar{u})(u_i - \bar{u})^T$$

$$\text{και } \bar{u} = \frac{1}{N} \sum_{i=1}^N u_i$$

N : ο αριθμός των θέσεων της περιοχής

u : το διάνυσμα των αναλογικών L-ροπών (t , t_3 , t_4) για τη θέση i .

❖ Προσεγγίζεται από τις L-ροπές των δεδομένων των θέσεων της κάθε περιοχής.

❖ Είναι μια μοναδική στατιστική ποσότητα, βασισμένη στη διαφορά των αναλογικών L-ροπών μιας συγκεκριμένης θέσης από τις αντίστοιχες μιας ομοιογενούς περιοχής.

❖ Ένας σταθμός θεωρείται ότι βρίσκεται σε ασυμφωνία με τους υπόλοιπους μιας «περιοχής» εάν το μέτρο D_i είναι μεγαλύτερο από μια κρίσιμη τιμή που εξαρτάται από τον αριθμό των σταθμών της περιοχής.

❖ Μπορεί να χρησιμοποιηθεί και για να προσδιορίσει εσφαλμένες τιμές.

ΜΕΤΡΑ ΑΣΥΜΦΩΝΙΑΣ ΚΑΙ ΑΝΟΜΟΙΟΓΕΝΕΙΑΣ

**Μέτρο Ανομοιογένειας -
Ετερογένειας H**

✓ Η εκτίμηση των L-ροπών μιας ομοιογενούς περιοχής γίνεται με τη μέθοδο Monte Carlo, χρησιμοποιώντας χρονοσειρές ίσου μήκους με αυτό των διαθέσιμων δεδομένων.

✓ Αν $H < 1$: Η περιοχή είναι “αποδεκτά ομοιογενής”

✓ Αν $1 \leq H < 2$: Η περιοχή είναι “πιθανώς ετερογενής”

✓ Αν $H \geq 2$: Η περιοχή είναι “σίγουρα ετερογενής”

❖ Αναφέρεται στη μεταβλητότητα των L-ροπών των δεδομένων των σταθμών/θέσεων συγκριτικά με την αντίστοιχη μεταβλητότητα που θα αναμενόταν για μια ομοιογενή περιοχή.

❖ $H(1)$: Η τυπική απόκλιση των συντελεστών t των θέσεων της περιοχής, τοποθετώντας σ'αυτούς συντελεστές βαρύτητας σε συνάρτηση με το μήκος των χρονοσειρών.

❖ $H(2)$: Η μέση απόσταση των τοπικών συντεταγμένων των θέσεων από μια μέση περιοχική τιμή σε ένα διάγραμμα των t και t_3 .

❖ $H(3)$: Η μέση απόσταση των τοπικών συντεταγμένων των θέσεων από μια μέση περιοχική τιμή σε ένα διάγραμμα των t_3 και t_4 .



ΜΕΤΡΑ ΑΣΥΜΦΩΝΙΑΣ ΚΑΙ ΑΝΟΜΟΙΟΓΕΝΕΙΑΣ

Μέτρο Ανομοιογένειας -Ετερογένειας H

H(1)



$$H(1) = \frac{V_1 - \mu_V}{\sigma_V}$$

H(2)



$$H(2) = \frac{V_2 - \mu_V}{\sigma_V}$$

H(3)



$$H(3) = \frac{V_3 - \mu_V}{\sigma_V}$$

$$V_1 = \sqrt{\frac{\sum_{i=1}^N n_i (t^{(i)} - t^R)^2}{\sum_{i=1}^N n_i}}$$

$$t^R = \frac{\sum_{i=1}^N n_i t^{(i)}}{\sum_{i=1}^N n_i}$$

$$V_2 = \frac{\sum_{i=1}^N n_i \sqrt{(t^{(i)} - t^R)^2 + (t_3^{(i)} - t_3^R)^2}}{\sum_{i=1}^N n_i}$$

$$t_3^R = \frac{\sum_{i=1}^N n_i t_3^{(i)}}{\sum_{i=1}^N n_i}$$

$$V_3 = \frac{\sum_{i=1}^N n_i \sqrt{(t_3^{(i)} - t_3^R)^2 + (t_4^{(i)} - t_4^R)^2}}{\sum_{i=1}^N n_i}$$

$$t_4^R = \frac{\sum_{i=1}^N n_i t_4^{(i)}}{\sum_{i=1}^N n_i}$$

- μ_V και σ_V είναι η μέση τιμή και η τυπική απόκλιση των προσομοιωμένων τιμών των V_1 , V_2 και V_3 , αντίστοιχα



ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΠΛΗΜΜΥΡΑΣ

- **Μέθοδοι Περιοχικής Ανάλυσης Συχνότητας Πλημμύρας (Regional Flood Frequency Analysis Methods)**
 - **Μέθοδος Δείκτη Πλημμύρας (Index Flood Method)**
 - **Μέθοδος Παλινδρόμησης των Ποσοστημορίων (π.χ. Q_{100}) (Method of Direct Regression of Quantiles) (DRQ)**
 - **Μέθοδος Παλινδρόμησης των Παραμέτρων της Κατανομής (Method of Regression for Distribution Parameters (RDP))**
- Όλα τα διαθέσιμα (φυσιογραφικά, γεωμορφολογικά και υδρομετεωρολογικά) στοιχεία στους διάφορους σταθμούς αυτής της περιοχής αναλύονται για να προσδιορίσουν τα χαρακτηριστικά της κατανομής συχνότητας της περιοχής (π.χ. για ανάλυση πλημμυρικής παροχής χαρακτηριστικά όπως υψόμετρο, έκταση, κλίση, σχήμα λεκάνης απορροής, τύπος εδάφους, χρήσεις γης, ημερήσια ακραία βροχόπτωση, μέση ετήσια βροχόπτωση χρησιμοποιούνται)

Watt, W.E., Lathem, K.W., Neill, C.R., Richards, T.L., and Rousselle, J., (eds). «Hydrology of Floods in Canada-A Guide to Planning and Design», National Research Council of Canada, Ottawa, Ontario, 1989.



Η ΜΕΘΟΔΟΣ ΤΟΥ ΔΕΙΚΤΗ ΠΛΗΜΜΥΡΑΣ

Παραδοχές της μεθόδου

- ⇒ Τα δεδομένα της μεταβλητής σε N σταθμούς/θέσεις μιας ομοιογενούς περιοχής κατανέμονται κατά πανομοιότυπο τρόπο.
- ⇒ Ενσωμάτωση ενός συντελεστή κλίμακας, του «**δείκτη πλημμύρας**», που προσδιορίζεται από τα δεδομένα του κάθε σταθμού.
- ⇒ Ο **δείκτης πλημμύρας** εΐθισται να συμπίπτει με τη μέση τιμή των υπό-μέλετη δεδομένων σε κάθε σταθμό.
- ⇒ Αν $Q(F)$ είναι η ποσοστιαία συνάρτηση στη θέση i με $0 < F < 1$ τη συνάρτησης κατανομής :
 - ο **δείκτης πλημμύρας** στη θέση i $Q(F) = \mu_i \cdot q(F)$ *αδιάστατη ποσοστιαία συνάρτηση*
- ⇒ Η αδιάστατη ποσοστιαία συνάρτηση $q(F)$ ονομάζεται καμπύλη περιοχικής αύξησης και είναι κοινή για όλες τις θέσεις μιας ομοιογενούς «περιοχής».



Η ΜΕΘΟΔΟΣ ΤΟΥ ΔΕΙΚΤΗ ΠΛΗΜΜΥΡΑΣ

Η καμπύλη περιοχικής αύξησης $q(F)$

⇒ Πραγματοποιείται προσαρμογή κατάλληλων συναρτήσεων κατανομής στα δεδομένα κάθε θέσης μιας ομοιογενούς «περιοχής» και υπολογίζονται οι παράμετροι της κατανομής αυτής, $\vartheta_k^{(i)}$ με τη μέθοδο των L-ροπών.

⇒ Οι εκτιμήτριες των παραμέτρων για το σύνολο της «περιοχής» είναι:

$$\hat{\theta}_k^R = \frac{\sum_{i=1}^N n_i \hat{\theta}_k^{(i)}}{\sum_{i=1}^N n_i}$$

N : Ο αριθμός των θέσεων κάθε «περιοχής»
 n_i : Το μήκος του ακραίου δείγματος στη θέση i

⇒ Οι εκτιμήτριες των «περιοχικών» παραμέτρων αντικαθίστανται στην $q(F)$ της συνάρτησης F : $q(F|\hat{\theta}_1^R, \dots, \hat{\theta}_k^R)$

⇒ Για τον προσδιορισμό μιας εύρωστης περιοχικής συνάρτησης κατανομής, χρησιμοποιείται το στατιστικό μέτρο Z (Z-statistic) των Hosking & Wallis (1997). Η προσαρμογή θεωρείται ικανοποιητική όταν: $|Z| \leq 1.64$



ΕΠΙΛΟΓΗ ΤΗΣ ΚΑΤΑΛΛΗΛΗΣ ΠΕΡΙΟΧΙΚΗΣ ΚΑΤΑΝΟΜΗΣ

ΑΚΡΑΙΩΝ ΤΙΜΩΝ

Έλεγχος πιθανών περιοχικών συναρτήσεων κατανομής

$$|Z| \leq 1.64$$

Η ασυμπτωτική κατανομή
GPD

➤ Εάν $x_{(1)}, x_{(2)}, \dots, x_{(k)}$ είναι ακραία γεγονότα, οι αντίστοιχες υπερβάσεις του ορίου $u : y_j = x_{(j)} - u$, με $j = 1, 2, \dots, k$ θα κατανέμονται:

$$F(y) = 1 - \left(1 + \frac{\xi \cdot y}{\sigma}\right)^{-1/\xi}$$

σ : παράμετρος κλίμακας

ξ : παράμετρος σχήματος

➤ Επίπεδο επαναφοράς m ετών:

$$q_m = u + \frac{\sigma}{\xi} [(m \zeta_u)^\xi - 1]$$

u : κατώτατο όριο , ζ_u : ποσοστό υπερβάσεων

➤ Οι εκτιμήσεις των επιπέδων επαναφοράς για δεδομένες περιόδους επαναφοράς στο σταθμό i προκύπτουν με εφαρμογή της σχέσης:

$$Q(F) = \mu_i \bullet q(F)$$

συνδυάζοντας την εκτίμηση του δείκτη πλημμύρας μ_i και τις εκτιμήσεις των αδιάστατων ποσοστιαίων σημείων που προκύπτουν από την $q(F)$.



ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΠΛΗΜΜΥΡΑΣ

Μέθοδος Δείκτης Πλημμύρας

1^ο Στάδιο. Ανάπτυξη Μεθόδου σε λεκάνες απορροής με μετρήσεις παροχής (π.χ. μέγιστες ημερήσιες παροχές ανά έτος)

- Παλινδρόμηση μεταξύ του Δείκτη Πλημμύρας – συνήθως η μέση ετήσια πλημμυρική παροχή – και των ανεξάρτητων φυσιολογικών/κλιματικών χαρακτηριστικών
 - Ανάπτυξη μοντέλου παλινδρόμησης: $Y_{\mu i} = a_0 + a_1X_1 + a_2X_2 + \dots + a_nX_n$
 - Υπολογισμός παραμέτρων παλινδρόμησης για γνωστά $Y_{\mu i}$ (εξαρτημένη μεταβλητή) και X_n (ανεξάρτητες μεταβλητές)
- Συσχέτιση Δείκτη Πλημμύρας για διάφορες περιόδους επαναφοράς
 - Εκτίμηση αδιάστατης καμπύλης περιοχικής αύξησης $q(F)$

2^ο Στάδιο. Εφαρμογή Μεθόδου σε λεκάνες απορροής χωρίς δεδομένα παροχής

- Εφαρμογή μοντέλου παλινδρόμησης για υπολογισμό $Y_{\mu i}$ (για γνωστά φυσιολογικά/κλιματικά χαρακτηριστικά της υπό-εξέτασης υδρολογικής λεκάνης και παράμετροι παλινδρόμησης όπως υπολογίστηκαν στο 1^ο Στάδιο)
- Εκτιμήσεις πλημμυρικών παροχών για δεδομένες περιόδους επαναφοράς
 $Q(F) = Y_{\mu i} * q(F)$



Μέθοδος Παλινδρόμησης των Ποσοστημορίων Πλημμύρας

1^ο Στάδιο. Ανάπτυξη Μεθόδου σε λεκάνες απορροής με μετρήσεις παροχής (π.χ. μέγιστες ημερήσιες παροχές ανά έτος)

- Εφαρμογή θεωρητικής κατανομής (π.χ. EVI, GEV) στα δεδομένα (μέγιστες ημερήσιες παροχές ανά έτος)
 - Υπολογισμός $Q(F)$ για διάφορες περιόδους επαναφοράς (π.χ. 5, 10, 50, 100 έτη) σε κάθε μία υδρολογική λεκάνη
- Παλινδρόμηση μεταξύ των υπολογισμένων $Q(F)$ (π.χ. για μία δοσμένη περίοδο επαναφοράς 100 έτη) και των ανεξάρτητων φυσιογραφικών/ κλιματικών χαρακτηριστικών
 - Ανάπτυξη μοντέλου παλινδρόμησης: $Q_{100} = a_0 + a_1X_1 + a_2X_2 + \dots + a_nX_n$
 - Υπολογισμός παραμέτρων παλινδρόμησης για γνωστά Q_{100} (εξαρτημένη μεταβλητή) και X_n (ανεξάρτητες μεταβλητές)
- Δημιουργία μοντέλων παλινδρόμησης για τις υπό-εξέταση περιόδους επαναφοράς (π.χ. 5 μοντέλα για 5 δοσμένες περιόδους επαναφοράς)
$$Q_T = a'_0 + a'_1X_1 + a'_2X_2 + \dots + a'_nX_n$$



ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΠΛΗΜΜΥΡΑΣ

Μέθοδος Παλινδρόμησης των Ποσοστημορίων Πλημμύρας

2^ο Στάδιο. Εφαρμογή Μεθόδου σε λεκάνες απορροής χωρίς δεδομένα παροχής

- Εφαρμογή μοντέλου παλινδρόμησης για υπολογισμό $Q_{T=100}$ (για γνωστά φυσιογραφικά/κλιματικά χαρακτηριστικά της υπό-εξέτασης υδρολογικής λεκάνης και παράμετροι παλινδρόμησης όπως υπολογίστηκαν στο 1^ο Στάδιο
 - $Q_{T=100}' = a_0 + a_1x_1 + a_2x_2 + \dots + a_nx_n$
- Εκτιμήσεις πλημμυρικών παροχών για δοσμένες περιόδους επαναφοράς με χρήση των μοντέλων παλινδρόμησης που αναπτύχθηκαν στο 1^ο Στάδιο



Μέθοδος Παλινδρόμησης των Παραμέτρων της Κατανομής

1^ο Στάδιο. Ανάπτυξη Μεθόδου σε λεκάνες απορροής με μετρήσεις παροχής (π.χ. μέγιστες ημερήσιες παροχές ανά έτος)

- Εφαρμογή θεωρητικής κατανομής (π.χ. EVI, GEV) στα δεδομένα (μέγιστες ημερήσιες παροχές ανά έτος)
 - Υπολογισμός των παραμέτρων της θεωρητικής κατανομής σε κάθε μία υδρολογική λεκάνη
 - Δύο (2) παράμετροι για EVI
 - Τρεις (3) παράμετροι για GEV, GPD
- Παλινδρόμηση μεταξύ των υπολογισμένων Παραμέτρων (παράμετρος θέσης (μ), κλίμακας (σ), σχήματος (ξ)) και των ανεξάρτητων φυσιογραφικών/κλιματικών χαρακτηριστικών
 - Ανάπτυξη μοντέλων παλινδρόμησης:
$$\mu = a_0 + a_1X_1 + a_2X_2 + \dots + a_nX_n$$
$$\sigma = a_0 + a_1X_1 + a_2X_2 + \dots + a_nX_n$$
$$\xi = a_0 + a_1X_1 + a_2X_2 + \dots + a_nX_n$$
 - Υπολογισμός παραμέτρων παλινδρόμησης για γνωστά μ , σ , ξ (εξαρτημένες μεταβλητές) και X_n (ανεξάρτητες μεταβλητές)



ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΠΛΗΜΜΥΡΑΣ

Μέθοδος Παλινδρόμησης των Παραμέτρων της Κατανομής

2^ο Στάδιο. Εφαρμογή Μεθόδου σε λεκάνες απορροής χωρίς δεδομένα παροχής

- Εφαρμογή μοντέλων παλινδρόμησης για υπολογισμό παραμέτρων [θέσης (μ), κλίμακας (σ), σχήματος (ξ)] θεωρητικής κατανομής με γνωστά φυσιογραφικά/κλιματικά χαρακτηριστικά της υπό-εξέτασης υδρολογικής λεκάνης και παράμετροι παλινδρόμησης όπως υπολογίστηκαν στο 1^ο Στάδιο

– Εφαρμογή μοντέλων παλινδρόμησης:

$$\mu' = a_0 + a_1x_1 + a_2x_2 + \dots + a_nx_n$$

$$\sigma' = a_0 + a_1x_1 + a_2x_2 + \dots + a_nx_n$$

$$\xi' = a_0 + a_1x_1 + a_2x_2 + \dots + a_nx_n$$

- Εκτιμήσεις πλημμυρικών παροχών για δοσμένες περιόδους επαναφοράς με χρήση της θεωρητικής κατανομής και τιμές παραμέτρων όπως υπολογίστηκαν στο 2^ο Στάδιο



ΠΑΡΑΔΕΙΓΜΑ ΕΦΑΡΜΟΓΗΣ

ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΑΚΡΑΙΩΝ ΤΙΜΩΝ

**Regional analysis involves the following steps
(Hosking and Wallis, 1997):**

- ❑ Screening of the data
- ❑ Testing of regional homogeneity
- ❑ Normalization of the annual maxima by an index (flood/storm) measure
- ❑ Fit a distribution (e.g. GEV) to each sample and obtain the parameters
- ❑ Combine estimates of the parameters to obtain regional estimates of parameters
- ❑ Scale regionalized distributions (estimated previously) by the index measures to get the local distributions



ΠΑΡΑΔΕΙΓΜΑ ΕΦΑΡΜΟΓΗΣ

ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΑΚΡΑΙΩΝ ΤΙΜΩΝ

1. Screening of the data

- ✓ Test for gross outliers, inconsistencies, shifts and trends.
- ✓ A **discordancy measure** by Hosking and Wallis (1997) is used to identify those sites from a given group of sites that are grossly discordant with the group as a whole.
- ✓ The discordancy measure is a single statistic, based on the difference between the L-moment ratios of a site and the average L-moment ratios of a group of similar sites. This statistic can also be used to identify erroneous data.
- ✓ The discordancy measure is of importance if $N \geq 7$.
- ✓ A site is declared to be discordant if the discordancy measure is larger than a critical value (depends on the number of sites of the sample)



ΠΑΡΑΔΕΙΓΜΑ ΕΦΑΡΜΟΓΗΣ

ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΑΚΡΑΙΩΝ ΤΙΜΩΝ

1. Screening of the data

✓ Suppose that there are N sites in the group. Let $\mathbf{u} = [\tau^{(i)}, \tau_3^{(i)}, \tau_4^{(i)}]^T$ be a vector containing the τ , τ_3 , and τ_4 values for site i .

✓ The unweighted group average is:

$$\bar{\mathbf{u}} = N^{-1} \sum_{i=1}^N \mathbf{u}_i$$

✓ Define the matrix:

$$\mathbf{A} = \sum_{i=1}^N (\mathbf{u}_i - \bar{\mathbf{u}})(\mathbf{u}_i - \bar{\mathbf{u}})^T.$$

✓ The discordancy measure D for site i is estimated:

$$D_i = \frac{1}{3} N (\mathbf{u}_i - \bar{\mathbf{u}})^T \mathbf{A}^{-1} (\mathbf{u}_i - \bar{\mathbf{u}}).$$



ΠΑΡΑΔΕΙΓΜΑ ΕΦΑΡΜΟΓΗΣ

ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΑΚΡΑΙΩΝ ΤΙΜΩΝ

✓ Declare a site i to be discordant if D_i is large. The definition of "large" depends on the number of sites in the group.

Critical values for the discordancy statistic D_i

Number of sites in region	Critical value	Number of sites in region	Critical value
5	1.333	10	2.491
6	1.648	11	2.632
7	1.917	12	2.757
8	2.140	13	2.869
9	2.329	14	2.971
		≥ 15	3

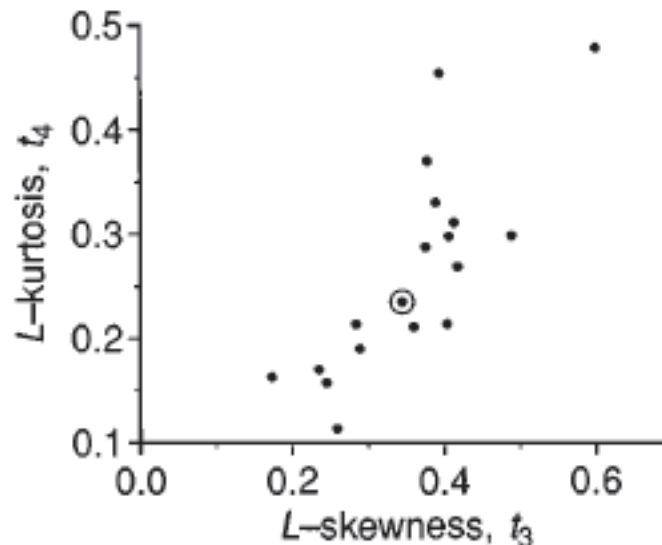
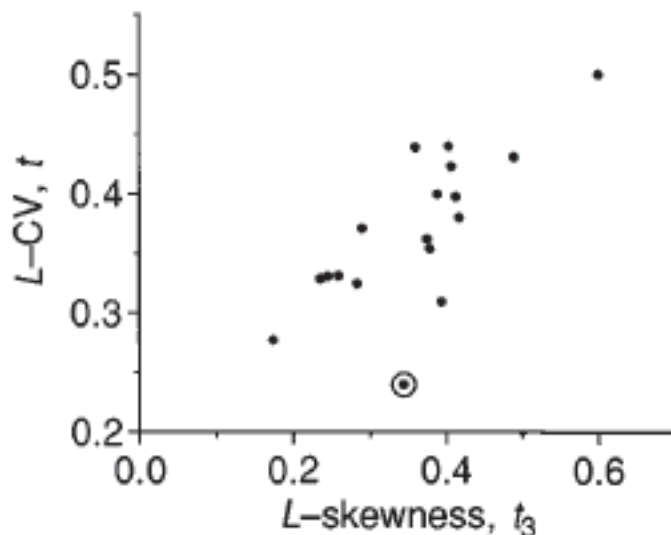


ΠΑΡΑΔΕΙΓΜΑ ΕΦΑΡΜΟΓΗΣ

ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΑΚΡΑΙΩΝ ΤΙΜΩΝ

Example

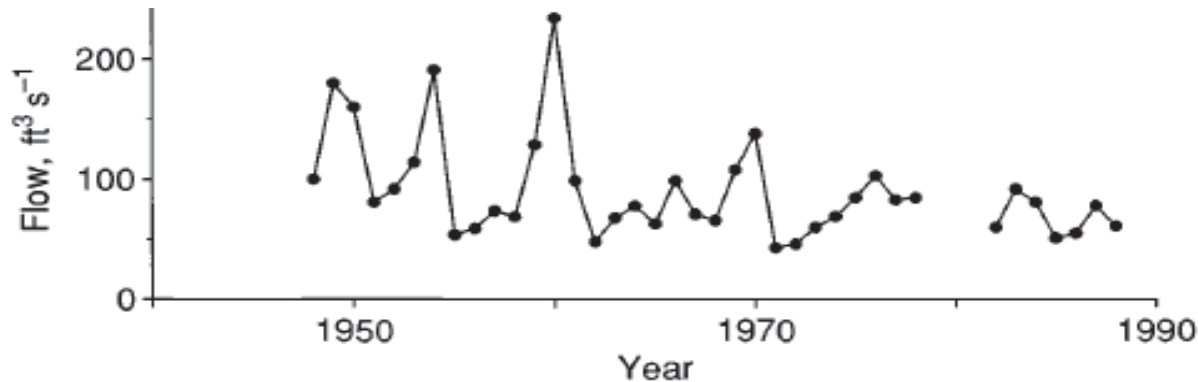
- Annual maximum streamflow data obtained from the U.S. Geological Survey (USGS). The 105 sites in the Region were assigned to six groups according to their drainage basin area. One group contains 18 sites with smallest drainage basin areas.
- The critical value of $D_i = 3$ is exceeded by only one site.



ΠΑΡΑΔΕΙΓΜΑ ΕΦΑΡΜΟΓΗΣ

ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΑΚΡΑΙΩΝ ΤΙΜΩΝ

Data from the discordant site



- Evidence of a downward trend over time.
- It is clear that the mean at this site is surprisingly low compared with other sites in the region with comparable drainage areas.
- There is, therefore, some reason to be suspicious of the data for this site.
- A final decision on whether to include this site in the analysis might require further investigation (whether the apparent trend can be attributed to random variation or whether it is a consequence of land-use changes or other man-induced effects that make the site an unreliable indicator of natural flow conditions)

ΠΑΡΑΔΕΙΓΜΑ ΕΦΑΡΜΟΓΗΣ

ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΑΚΡΑΙΩΝ ΤΙΜΩΝ

2. Identifying homogeneous regions

- ❖ **Cluster analysis** is a standard method of statistical multivariate analysis for dividing a data set into groups and has been successfully used to form regions for regional frequency analysis.
- ❖ A data vector is associated with each site, and sites are partitioned or aggregated into groups according to the similarity of their data vectors.
- ❖ Most clustering algorithms measure similarity by the *reciprocal of Euclidean distance* in a space of site characteristics. This distance measure is affected by the scale of measurement of the site characteristics, and in practice it is usual to rescale the site characteristics, so that they all have the same amount of variability.
- ❖ The output from the cluster analysis need not, and usually should not, be final.



ΠΑΡΑΔΕΙΓΜΑ ΕΦΑΡΜΟΓΗΣ

ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΑΚΡΑΙΩΝ ΤΙΜΩΝ

2. Identifying homogeneous regions

- ✓ An homogeneous region is formed by stations with frequency distributions that are identical, apart from a **station-specific scale factor**.
- ✓ A heterogeneity measure, called the **H-statistic**, is used to compare the between site variation of sample L-moments for a group of sites with what would be expected for a homogeneous region.
- ✓ A large positive value of the H-statistic indicate that the L-moment ratios are more dispersed than is consistent with the hypothesis of homogeneity.
- ✓ To estimate what would be expected for a homogeneous region, Monte Carlo simulations of a homogeneous region with sites having record lengths equal to those of the observed data are performed.



ΠΑΡΑΔΕΙΓΜΑ ΕΦΑΡΜΟΓΗΣ

ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΑΚΡΑΙΩΝ ΤΙΜΩΝ

2. Identifying homogeneous regions

- ✓ There are three measures of the H-statistic:
 - 1) $H(1)$ is the standard deviation of the at-site L-CVs, weighted according to record length.
 - 2) $H(2)$ is the average distance from the site coordinates to a regional average on a plot of L-CV versus L-Skewness
 - 3) $H(3)$ is the average distance from the site coordinates to the regional average on a plot of L-Skewness versus L-kurtosis
- ✓ If $H < 1 \rightarrow$ The region is “acceptably homogeneous”
- ✓ If $1 \leq H < 2 \rightarrow$ The region is “possibly heterogeneous”
- ✓ If $H \geq 2 \rightarrow$ The region is “definitely heterogeneous”



ΠΑΡΑΔΕΙΓΜΑ ΕΦΑΡΜΟΓΗΣ

ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΑΚΡΑΙΩΝ ΤΙΜΩΝ

2. Identifying homogeneous regions

- ✓ Hosking and Wallis, (1997) found that $H(2)$ is a weaker test of heterogeneity than $H(1)$.
- ✓ The heterogeneity measure $H(2)$ is used by the FEH (Institute of Hydrology, 1999):
 - If $2 < H(2) \leq 4 \rightarrow$ The region is “heterogeneous
 - If $H(2) > 4 \rightarrow$ The region is “strongly heterogeneous”



ΠΑΡΑΔΕΙΓΜΑ ΕΦΑΡΜΟΓΗΣ

ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΑΚΡΑΙΩΝ ΤΙΜΩΝ

➤ The Hosking and Wallis (1993) test for homogeneity

- Suppose that the region has N sites, with site i having record length n_i and sample L-moment ratios $\tau^{(i)}$ (L-CV), $\tau_3^{(i)}$ (L-skewness) and $\tau_4^{(i)}$ (L-kurtosis) of maximum k -day precipitation amounts.
- The test statistic $H(1)$ is:

$$H(1) = \frac{V_1 - \mu_V}{\sigma_V}$$

where

$$V_1 = \sqrt{\frac{\sum_{i=1}^N n_i (\tau^{(i)} - \tau^R)^2}{\sum_{i=1}^N n_i}} \quad \text{and} \quad \tau^R = \frac{\sum_{i=1}^N n_i \tau^{(i)}}{\sum_{i=1}^N n_i}$$

μ_V and σ_V are determined from simulations as the mean and standard deviations of the simulated values of V_1 .



ΠΑΡΑΔΕΙΓΜΑ ΕΦΑΡΜΟΓΗΣ

ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΑΚΡΑΙΩΝ ΤΙΜΩΝ

➤ The Hosking and Wallis (1993) test for homogeneity

- μ_V and σ_V are determined from simulations as the mean and standard deviations of the simulated values of V_1 .



500 realizations of a homogeneous region with N sites, each having a four-parameter kappa distribution (Hosking, 1994) with L-moment ratios equal to regional averages of sample L-moments τ^R , τ_3^R , τ_4^R and at-site mean equal to 1.

- The other two analogous tests are based:
 - H(2) : on the combination of L-CV (τ) and L-skewness (τ_3)
 - H(3) : on the combination of L-skewness (τ_3) and L-kurtosis (τ_4)



ΠΑΡΑΔΕΙΓΜΑ ΕΦΑΡΜΟΓΗΣ

ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΑΚΡΑΙΩΝ ΤΙΜΩΝ

➤ The Hosking and Wallis (1993) test for homogeneity

$$V_2 = \frac{\sum_{i=1}^N n_i \sqrt{(\tau^{(i)} - \tau^R)^2 + (\tau_3^{(i)} - \tau_3^R)^2}}{\sum_{i=1}^N n_i} \quad V_3 = \frac{\sum_{i=1}^N n_i \sqrt{(\tau_3^{(i)} - \tau_3^R)^2 + (\tau_4^{(i)} - \tau_4^R)^2}}{\sum_{i=1}^N n_i}$$

- μ_V and σ_V are determined from simulations as the mean and standard deviations of the simulated values of V_2 and V_3 , respectively.



500 realizations of a homogeneous region with N sites, each having a four-parameter kappa distribution (Hosking, 1994) with L-moment ratios equal to regional averages of sample L-moments τ^R , τ_3^R , τ_4^R and at-site mean equal to 1.



ΠΑΡΑΔΕΙΓΜΑ ΕΦΑΡΜΟΓΗΣ

ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΑΚΡΑΙΩΝ ΤΙΜΩΝ

➤ The Lu and Stedinger (1992) test for homogeneity

- Suppose that the region has N sites, with site i having record length n_i and sample L-moment ratios $\tau^{(i)}$ (L-CV), $\tau_3^{(i)}$ (L-skewness) and $\tau_4^{(i)}$ (L-kurtosis) of maximum k -day precipitation amounts.
- The test statistic χ^2_R is:

$$\chi^2_R = \sum_{i=1}^N \frac{(\xi_{0.9}^{(i)} - \xi_{0.9}^R)^2}{\text{Var} \xi_{0.9}^{(i)}}$$

where $\xi_{0.9}^R = \frac{\sum_{i=1}^N n_i \xi_{0.9}^{(i)}}{\sum_{i=1}^N n_i}$ and

$$\xi_{0.9}^{(i)} = 1 + \frac{\tau^{(i)}}{1-2^{-k}} \left(1 - \frac{(-\ln 0.90)^k}{\Gamma(1+k)}\right)$$

$$k = 7.8590C + 2.9554C^2$$

$$C = \frac{2}{\tau_3^{(i)} + 3} - \frac{\ln 2}{\ln 3}$$



ΠΑΡΑΔΕΙΓΜΑ ΕΦΑΡΜΟΓΗΣ

ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΑΚΡΑΙΩΝ ΤΙΜΩΝ

➤ The Lu and Stedinger (1992) test for homogeneity

- $Var\xi_{0.9}^{(i)}$ is determined from simulations as the variance of the 90% quantile



500 realizations of a homogeneous region with N sites, each having a three-parameter GEV distribution with L-moment ratios equal to $\tau^{(i)}$, $\tau_3^{(i)}$ and at-site mean equal to 1.

- ✓ If $\chi_R^2 < \chi_{0.95, N-1}^2$ the region is **homogeneous** at a 5% significance level
- ✓ If $\chi_R^2 \geq \chi_{0.95, N-1}^2$ the region is **heterogeneous** at a 5% significance level



ΠΑΡΑΔΕΙΓΜΑ ΕΦΑΡΜΟΓΗΣ

ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΑΚΡΑΙΩΝ ΤΙΜΩΝ

3. An index flood/storm estimation procedure

- The assumption that the N sites in a homogeneous region have an identical frequency distribution, apart from a site-specific scaling factor, termed the “**index flood**” in a streamflow analysis and the “**index storm**” in a precipitation analysis.
- The index variable is usually the mean of the site-specific data (Hosking and Wallis, 1997).
- Suppose that data are available at N sites, with site i having sample size n_i and observed data Q_{ij} , $j = 1, \dots, n_i$. Let $Q_i(F)$, $0 < F < 1$, be the quantile function of the frequency distribution at site.
- We may write :

$$Q_i(F) = \mu_i q(F) \longrightarrow \text{Regional growth curve, a dimensionless quantile function common to each site}$$

The index flood/storm, which can be taken to be the mean of the at-site frequency distribution



ΠΑΡΑΔΕΙΓΜΑ ΕΦΑΡΜΟΓΗΣ

ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΑΚΡΑΙΩΝ ΤΙΜΩΝ

3. An index flood/storm estimation procedure

- The dimensionless rescaled data $q_{ij} = Q_{ij}/\mu_i$, $j = 1, \dots, n_i$, $i = 1, \dots, N$, are the basis for estimating **the regional growth curve $q(F)$** , $0 < F < 1$.
- The mean of the regional frequency distribution is not an unknown parameter, because by taking μ_i to be the mean of the frequency distribution at site i we ensure that the regional frequency distribution has mean 1.
- A distribution function is fitted at the extremes of all sites. The parameters are estimated separately at each site, the site- i estimate of ϑ_k being denoted by $\hat{\theta}_k^{(i)}$
- The at-site estimates are combined to give regional estimates



ΠΑΡΑΔΕΙΓΜΑ ΕΦΑΡΜΟΓΗΣ

ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΑΚΡΑΙΩΝ ΤΙΜΩΝ

3. An index flood/storm estimation procedure

- The at-site estimates are combined to give regional estimates:

$$\hat{\theta}_k^R = \frac{\sum_{i=1}^N n_i \hat{\theta}_k^{(i)}}{\sum_{i=1}^N n_i}$$

- Substituting these estimates into $q(F)$ gives the estimated regional growth curve

$$q(F) = q(F | \hat{\theta}_1^R, \dots, \hat{\theta}_k^R)$$

- The quantile estimates at site i are obtained by combining the estimates of μ_i and $q(F)$:

$$\hat{Q}_i(F) = \hat{\mu}_i \hat{q}(F)$$



ΠΑΡΑΔΕΙΓΜΑ ΕΦΑΡΜΟΓΗΣ

ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΑΚΡΑΙΩΝ ΤΙΜΩΝ

3. An index flood/storm estimation procedure

- Some quantities that can take either positive or negative values, the frequency distributions in a region can be described:

$$Q(F) = \mu_i + \sigma_i q(F)$$

- An alternative approach to regional estimation is to model $\log Q$ rather than Q , basing the analysis on logarithmically transformed data.

$$\log Q_i(F) = \log \mu_i + \log q(F)$$



ΠΑΡΑΔΕΙΓΜΑ ΕΦΑΡΜΟΓΗΣ

ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΑΚΡΑΙΩΝ ΤΙΜΩΝ

4. Selection of a distribution - Estimation of parameters using L-moments

- The GEV distribution is used with CDF:

$$F_{\xi, \mu, \sigma}(x) = \exp\left(-\left(1 + \xi \frac{(x - \mu)}{\sigma}\right)^{-1/\xi}\right) \quad \text{with } 1 + \xi \frac{(x - \mu)}{\sigma} > 0, \quad \xi \neq 0,$$
$$F_{0, \mu, \sigma}(x) = \exp\left(-e^{-\frac{(x - \mu)}{\sigma}}\right) \quad \text{with } \xi = 0.$$

- The q th quantile estimate for $\xi \neq 0$ can be estimated:

$$x_q = \mu - \frac{\sigma}{\xi} [1 - (-\ln q)^{-\xi}]$$

- The shape parameter ξ can be estimated from the L-skewness ratio:

$$\tau_3 = -3 + \frac{2(1 - 3^\xi)}{(1 - 2^\xi)} \Rightarrow \xi = -7.8590C - 2.944C^2, \quad C = \frac{2}{(\tau_3 + 3)} - \frac{\ln 2}{\ln 3}$$



ΠΑΡΑΔΕΙΓΜΑ ΕΦΑΡΜΟΓΗΣ

ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΑΚΡΑΙΩΝ ΤΙΜΩΝ

4. Selection of a distribution - Estimation of parameters using L-moments

➤ On the shape parameter is estimated, the parameters σ and μ can be estimated:

$$\sigma = \frac{-\lambda_2 \xi}{(1 - 2^\xi) \Gamma(1 - \xi)}$$

$$\mu = \lambda_1 + \frac{\sigma}{\xi} [1 - \Gamma(1 - \xi)]$$



ΠΑΡΑΔΕΙΓΜΑ ΕΦΑΡΜΟΓΗΣ

ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΑΚΡΑΙΩΝ ΤΙΜΩΝ

4. Selection of a distribution

□ In some analyses there may be reason to consider other distributions. For example, annual maximum streamflow data may contain some values arising from floods caused by snowmelt and some from rainstorm events.

□ A mixture of two distributions may be entertained. The regional frequency distribution would be:

$$F(x) = pG_1(x) + (1 - p)G_2(x)$$

$G_1(x)$ and $G_2(x)$ are CDFs of data arising from two different causes and p is the proportion of the observations arising from the first cause.

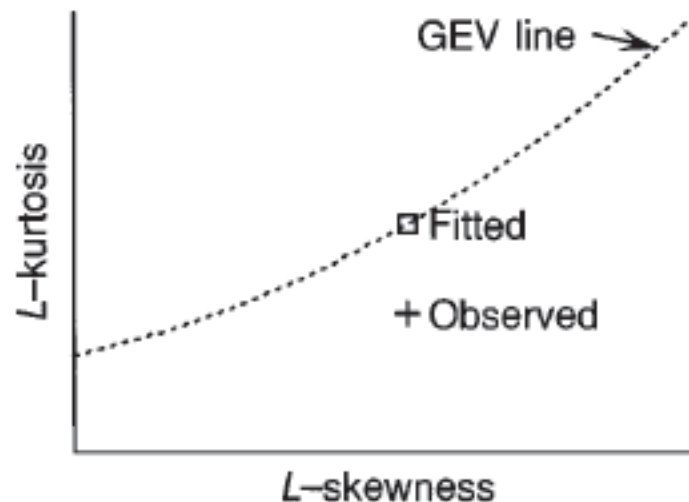


ΠΑΡΑΔΕΙΓΜΑ ΕΦΑΡΜΟΓΗΣ

ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΑΚΡΑΙΩΝ ΤΙΜΩΝ

5. Identifying a robust regional frequency distribution

- ✓ A number of goodness-of-fit techniques are available in the literature.
- ✓ The relation of L-kurtosis and L-Skewness for various distributions is compared to the corresponding relations obtained from the at-site and regional data.



ΠΑΡΑΔΕΙΓΜΑ ΕΦΑΡΜΟΓΗΣ

ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΑΚΡΑΙΩΝ ΤΙΜΩΝ

5. Identifying a robust regional frequency distribution

✓ The Z-statistic, introduced by Hosking and Wallis (1997) is used. This statistic was developed for three-parameter distributions and measures how well the theoretical L-kurtosis of the fitted distribution matches the regional L-kurtosis of the data. The fit is considered satisfactory if

$$|Z| \leq 1.64$$

✓ It may happen that none of the candidate distributions is accepted by the Z criterion. This sometimes occurs when the number of sites in the region or the at-site record lengths are large.

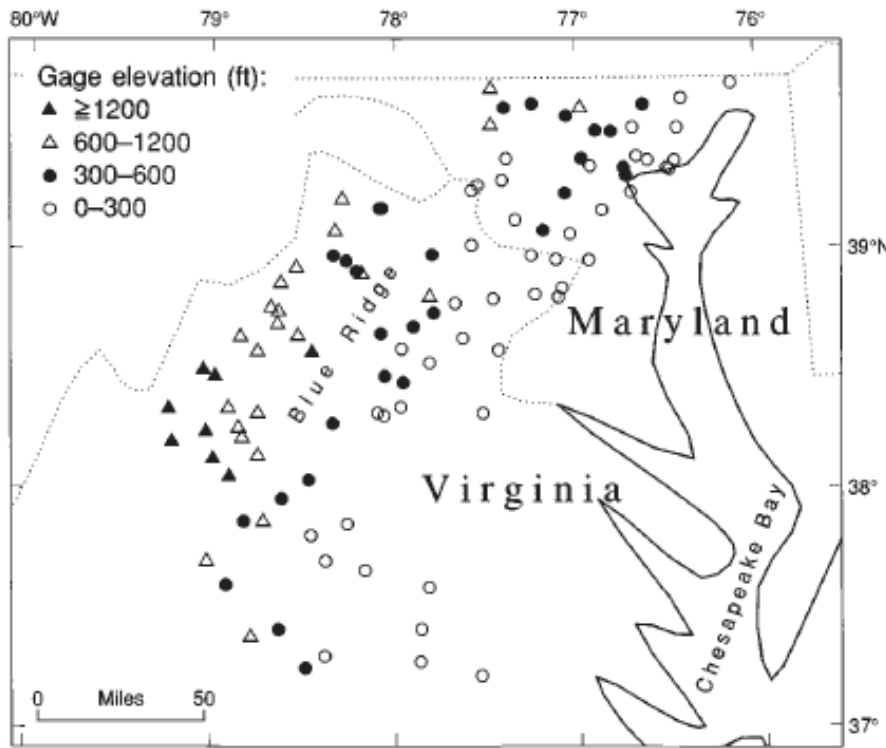


ΠΑΡΑΔΕΙΓΜΑ ΕΦΑΡΜΟΓΗΣ

ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΠΛΗΜΜΥΡΩΝ

Παράδειγμα Regional Flood Frequency Analysis

□ Smith (1992) analysed data on annual maximum streamflow at 104 gaging stations in the central Appalachia region of the United



□ Drainage areas of the basins cover a very wide range, from 0.3 mi² to 10,000 mi².

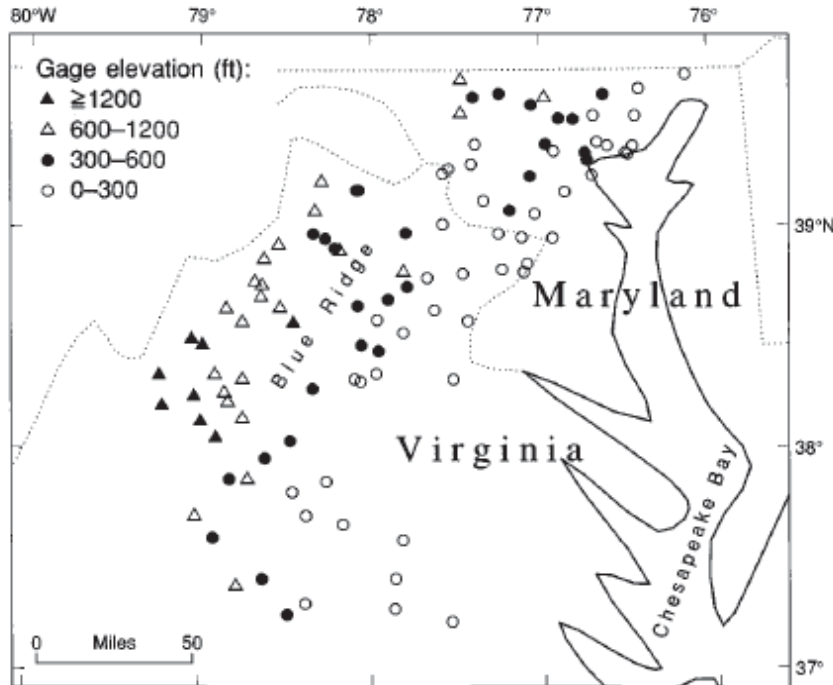
□ Elevations of the gaging sites range from 10 ft to 2,000 ft above sea level.



ΠΑΡΑΔΕΙΓΜΑ ΕΦΑΡΜΟΓΗΣ

ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΠΛΗΜΜΥΡΩΝ

Παράδειγμα Regional Flood Frequency Analysis



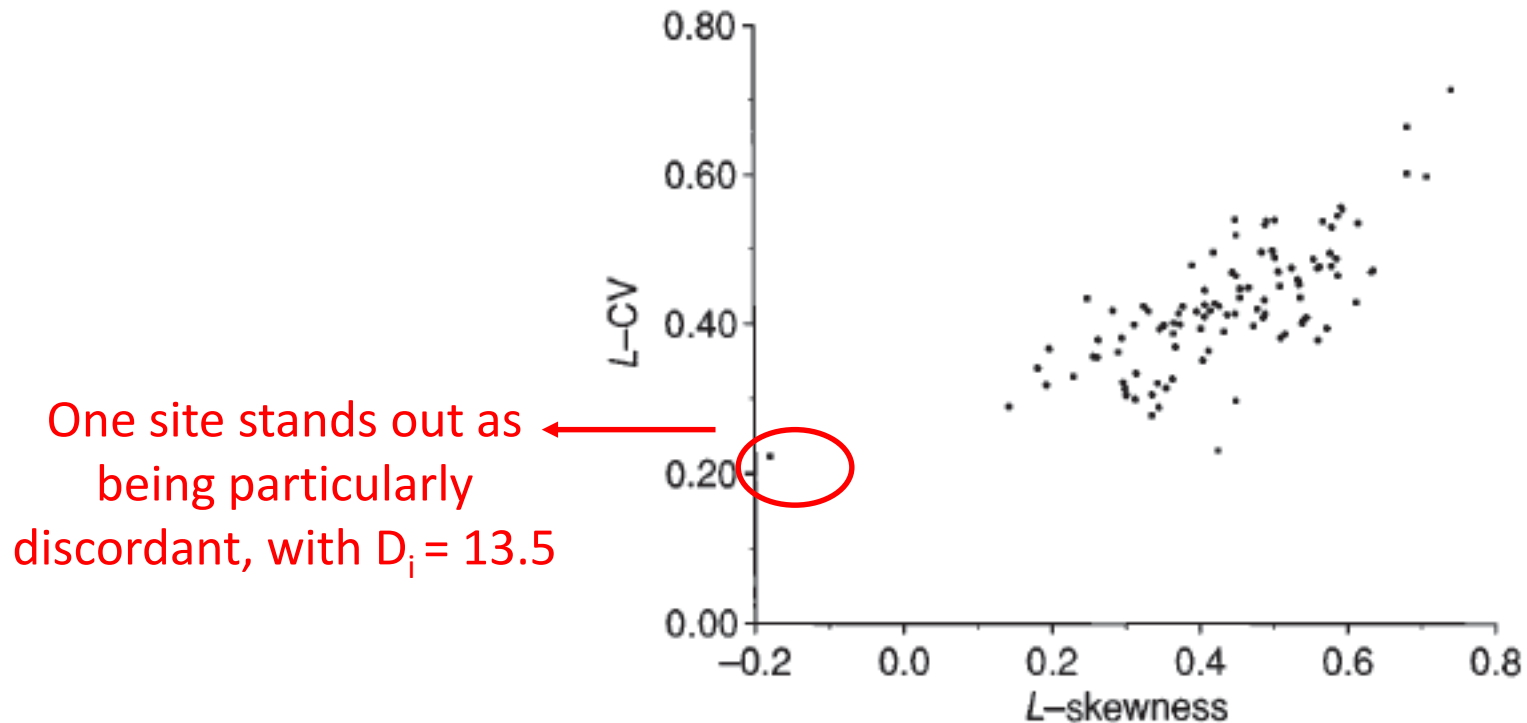
□ In addition to these variables, the data set contains the gage location (latitude and longitude) and the magnitude and date of the maximum instantaneous streamflow in each "water year," running from Oct. 1 through Sep. 30, for some or all of the years 1895-1991.

□ Some variables that might be expected to influence the frequency distribution, such as mean annual precipitation over the basin, the underlying geology of the basin, and the extent to which the basin is forested or urbanized, are not available.



ΠΑΡΑΔΕΙΓΜΑ ΕΦΑΡΜΟΓΗΣ ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΠΛΗΜΜΥΡΩΝ

Screening of the data



□ Treating the entire set of 104 sites as a single region, the discordancy statistic D_i was calculated for each site.

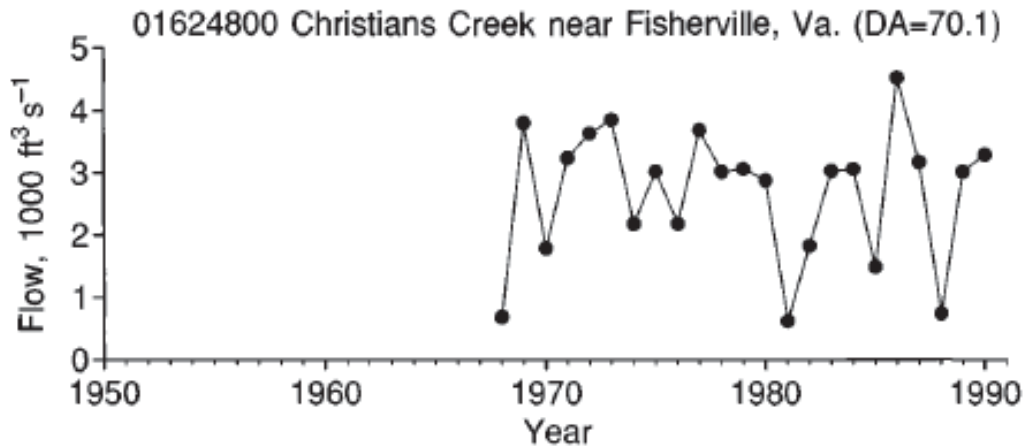


ΠΑΡΑΔΕΙΓΜΑ ΕΦΑΡΜΟΓΗΣ

ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΠΛΗΜΜΥΡΩΝ

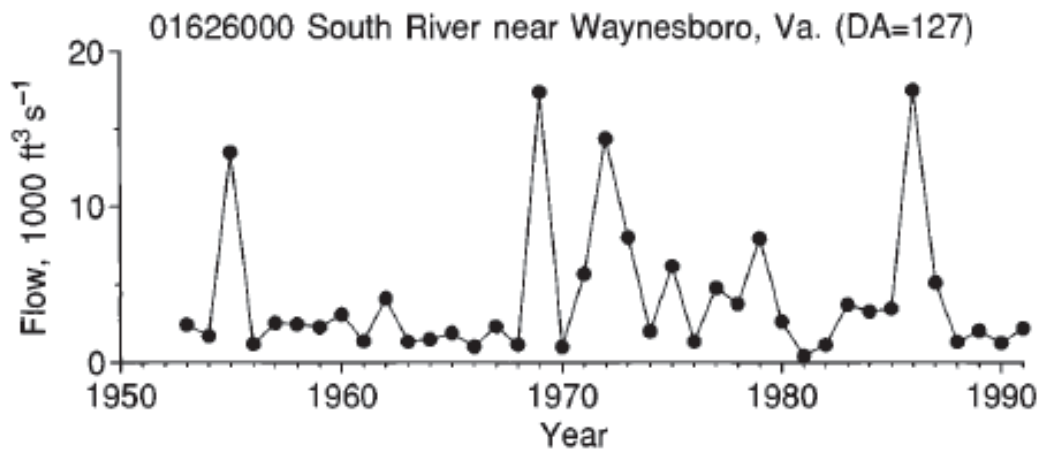
Screening of the data

Discordant site



- There is good qualitative agreement between the sites

Nearest neighbor



- The negative L-skewness occurs because the period of gaging included three years that had particularly low maximum flows



ΠΑΡΑΔΕΙΓΜΑ ΕΦΑΡΜΟΓΗΣ

ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΠΛΗΜΜΥΡΩΝ

Screening of the data


- The geology of the basin consists of fractured quartzite sandstone overlying dolomite, resulting in a Karstlike hydrology.
- The basin is very porous and flood response to even the most extreme storm events is slow.
- In consequence, the frequency distribution of annual maximum flows at the gaging site should be less skew than at sites with impervious geology and a fast response time to extreme storms.
- It is not clear whether this site is geologically so atypical of the rest of the data that it should be excluded from the analysis. We therefore decided to ***retain the site in the data set.***



ΠΑΡΑΔΕΙΓΜΑ ΕΦΑΡΜΟΓΗΣ

ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΠΛΗΜΜΥΡΩΝ

Formation of homogeneous regions

- Treating the entire set of 104 sites as a single region, the heterogeneity statistic was evaluated as $H = 2.08$.
 - The entire set is therefore not far from being homogeneous, or at worst "possibly heterogeneous".
 - **The frequency distribution is likely to depend on the drainage area of the basin.**
- 
- A reasonable starting point for regional frequency analysis would therefore be a subdivision of the set of sites, according to their drainage areas, into groups of not much more than 20.
 - The set of sites was accordingly divided into four groups of approximately equal size according to the sites' drainage areas



ΠΑΡΑΔΕΙΓΜΑ ΕΦΑΡΜΟΓΗΣ

ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΠΛΗΜΜΥΡΩΝ

Formation of homogeneous regions

Region	Range of area	Number of sites	Regional average		H
			t	t_3	
A	< 10	26	0.423	0.402	2.88
B	10–60	28	0.450	0.472	1.70
C	60–200	27	0.430	0.459	0.03
D	> 200	23	0.387	0.419	0.29

- The **heterogeneity measure H** indicates that homogeneous regions were achieved for sites with drainage area greater than 60 mi² but that the regions containing the smaller basins were heterogeneous.
- For the smaller basins, the frequency distribution of annual maximum streamflow is determined by more factors than drainage area alone.



ΠΑΡΑΔΕΙΓΜΑ ΕΦΑΡΜΟΓΗΣ

ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΠΛΗΜΜΥΡΩΝ

Formation of homogeneous regions

- The four available site characteristics, drainage basin area, gage elevation, gage latitude, and gage longitude, were used in a cluster analysis procedure.
- A logarithmic transformation to drainage basin area and a square root transform to gage elevation were applied to give a more symmetric distribution of the values of the site characteristics at the 104 sites.
- All four variables were then standardized by dividing by the standard deviation of their values at the 104 sites.

Site characteristic, X	Cluster variable, Y
Drainage basin area (mi^2)	$Y = \log(X) \times 3 / \text{s.d.}(\log X)$
Gage elevation (ft)	$Y = \sqrt{X} / \text{s.d.}(\sqrt{X})$
Gage latitude (deg)	$Y = X / \text{s.d.}(X)$
Gage longitude (deg)	$Y = X / \text{s.d.}(X)$

To give it an importance in the clustering procedure equal to that of the other variables together

Note: Here $\text{s.d.}(X)$ denotes the standard deviation of the 104 values of the site characteristic X .

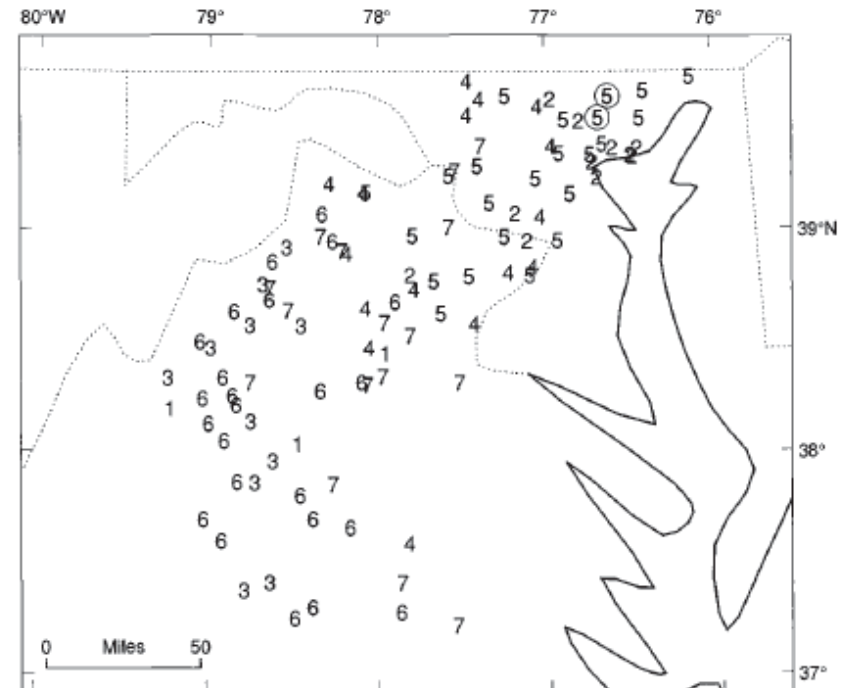
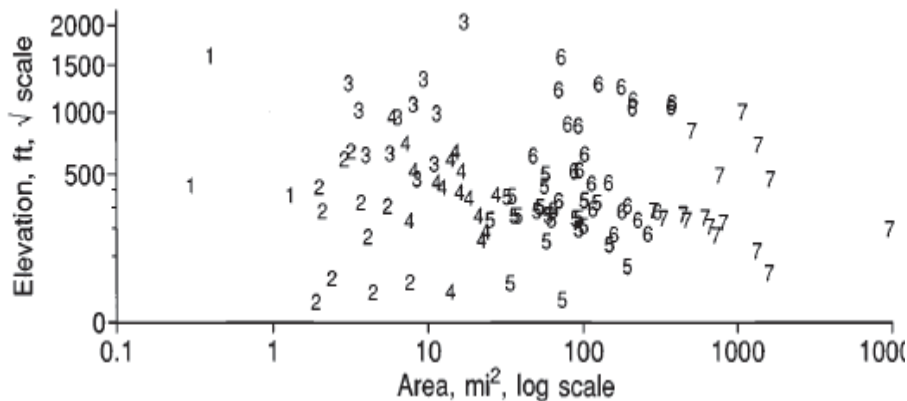


ΠΑΡΑΔΕΙΓΜΑ ΕΦΑΡΜΟΓΗΣ

ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΠΛΗΜΜΥΡΩΝ

Formation of homogeneous regions

- Cluster analysis was performed using Ward's method.
- The clusters obtained by Ward's method were adjusted using the *K*-means algorithm of Hartigan and Wong (1979). Seven clusters were identified.



ΠΑΡΑΔΕΙΓΜΑ ΕΦΑΡΜΟΓΗΣ

ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΠΛΗΜΜΥΡΩΝ

Formation of homogeneous regions

Region	Range of area	Number of sites	Regional average		H	Acceptable fit
			t	t_3		
1	0.3–1.3	3	0.348	0.331	–0.40	GLO GEV LN3 GPA PE3
2	1.9–7.6	11	0.379	0.363	1.65	GLO GEV LN3
3	3.1–17	11	0.498	0.471	1.62	GLO GEV LN3
4	5.9–28	16	0.450	0.430	1.69	GLO GEV LN3 GPA
5	25–190	23	0.424	0.491	0.24	none
6	47–380	24	0.438	0.442	–0.03	GLO GEV
6'	47–380	23	0.442	0.455	–0.71	GLO GEV
7	290–9600	16	0.376	0.420	0.87	GLO GEV

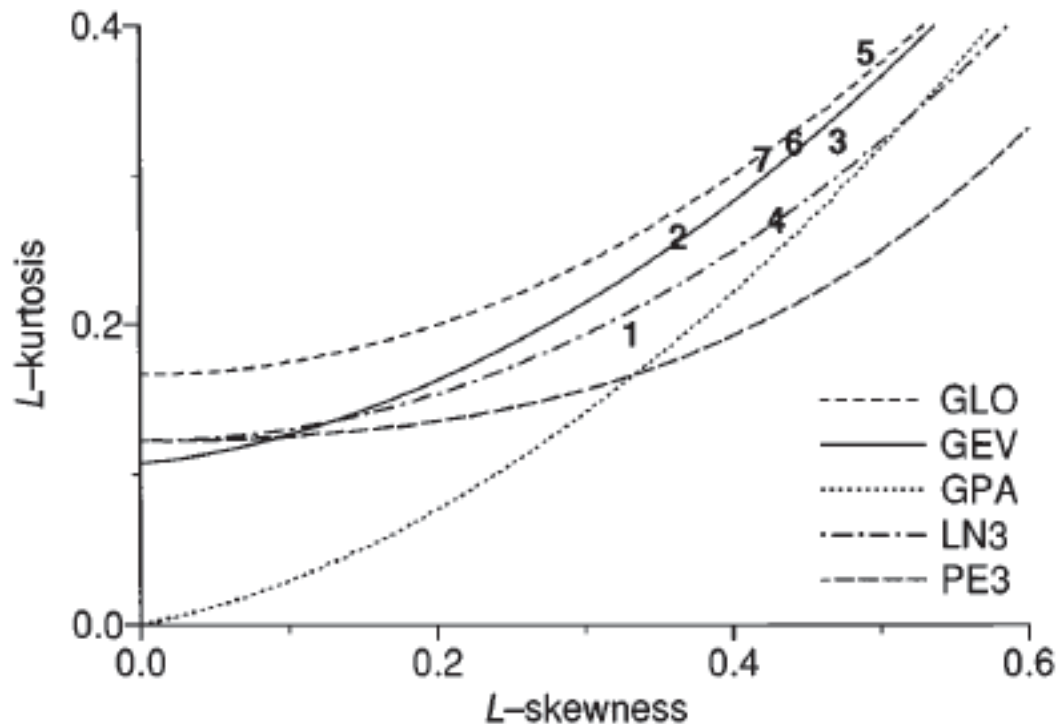
- Regions 2-4 are "possibly heterogeneous", with H values as high as 1.69, but should yield quantile estimates more accurate than those obtained from single-site frequency analysis.
- Region 6 contains the site with negative L-skewness, for which $D_i = 6.9$.



ΠΑΡΑΔΕΙΓΜΑ ΕΦΑΡΜΟΓΗΣ

ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΠΛΗΜΜΥΡΩΝ

Choice and estimation of the frequency distribution



If the use of a single distribution for each region is desired, the generalized logistic or generalized extreme-value would be a reasonable choice.



ΠΑΡΑΔΕΙΓΜΑ ΕΦΑΡΜΟΓΗΣ

ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΠΛΗΜΜΥΡΩΝ

Regional frequency distributions

Region	Nonexceedance probability									
	0.01	0.10	0.20	0.50	0.80	0.90	0.98	0.99	0.998	0.999
1	0.17	0.37	0.49	0.81	1.35	1.80	3.12	3.85	6.09	7.34
2	0.14	0.34	0.45	0.78	1.35	1.85	3.39	4.29	7.18	8.88
3	0.03	0.20	0.31	0.64	1.32	1.99	4.46	6.15	12.60	17.04
4	0.07	0.25	0.37	0.70	1.34	1.95	4.03	5.38	10.18	13.30
5	0.20	0.33	0.42	0.69	1.25	1.82	3.98	5.50	11.47	15.69
6	0.11	0.28	0.39	0.70	1.32	1.91	3.98	5.34	10.30	13.57
7	0.22	0.37	0.47	0.75	1.30	1.80	3.52	4.61	8.45	10.91

Apparent differences

- Regions 1 and 2 have relatively low upper-tail quantiles.
- Otherwise the upper-tail quantiles generally increase as the drainage area of the basin decreases. Region 4 is an exception.



ΠΑΡΑΔΕΙΓΜΑ ΕΦΑΡΜΟΓΗΣ

ΠΕΡΙΟΧΙΚΗ ΑΝΑΛΥΣΗ ΣΥΧΝΟΤΗΤΑΣ ΠΛΗΜΜΥΡΩΝ

Estimated quantiles of annual maximum streamflow, in cubic feet per second, for Blue Mount and Western Run

Method	Site	Nonexceedance probability				
		0.9	0.98	0.99	0.998	0.999
At-site	Blue Mount	4660	7570	9110	13600	16000
	Western Run	6430	15800	22990	54100	78000
Regional	Blue Mount	4960	10900	15000	31300	42800
	Western Run	6390	14000	19300	40300	55100

- The gauging sites are 7 mi apart and have similar elevations, 305 ft and 262 ft, respectively.
- The basins also have similar drainage areas, 52.9 mi² and 59.8 mi², respectively.
- Both sites have gauged records from 1945 to 1991



R packages for RFA

- ❑ RFA (Mathieu Ribatet, 2010)
- ❑ nsRFA (Alberto Viglione, 2010)
- ❑ ImomRFA (J. R. M. Hosking, 2010)

Package ImomRFA

- > library (Imom)
- > library(ImomRFA)



Βιβλιογραφία

- Μιμίκου, Μ.Α. «Τεχνολογία Υδατικών Πόρων», Εκδόσεις Παπασωτηρίου, 3^η Έκδοση, 2006.
- Hosking, J.R.M. and Wallis, J.R. «Regional Frequency Analysis : An Approach based on L-Moments», Cambridge University Press, 238 pp., 1997.
- Watt, W.E., Lathem, K.W., Neill, C.R., Richards, T.L., and Rousselle, J., (eds). «Hydrology of Floods in Canada-A Guide to Planning and Design», National Research Council of Canada, Ottawa, Ontario, 1989.



Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στο πλαίσιο του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «**Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Θεσσαλίας**» έχει χρηματοδοτήσει μόνο την αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.

