



**Γ' Κοινοτικό Πλαίσιο Στήριξης 2000-2006  
Επιχειρησιακό Πρόγραμμα Ανταγωνιστικότητα**

**NAMA  
ΕΜΠ  
ΔΕΥΑΚ  
ΑΕΙΦΟΡΙΚΗ  
MDS**

**ΟΛΟΚΛΗΡΩΜΕΝΗ ΔΙΑΧΕΙΡΙΣΗ ΥΔΑΤΙΚΩΝ ΣΥΣΤΗΜΑΤΩΝ ΣΕ  
ΣΥΖΕΥΞΗ ΜΕ ΕΞΕΛΙΓΜΕΝΟ ΥΠΟΛΟΓΙΣΤΙΚΟ ΣΥΣΤΗΜΑ**

**ΟΔΥΣΣΕΥΣ**

**Τεύχος 2**

**Περιγραφή συστήματος διαχείρισης και επεξεργασίας  
δεδομένων «Υδρογνώμων»**

*Σ. Κοζάνης, Α. Χριστοφίδης, και Α. Ευστρατιάδης*

**Αθήνα  
Σεπτέμβριος  
2005**



## Συνεργαζόμενοι φορείς



NAMA Σύμβουλοι Μηχανικοί και Μελετητές Α.Ε.



Εθνικό Μετσόβιο Πολυτεχνείο, Τομέας Υδατικών Πόρων  
Υδραυλικών και Θαλάσσιων Έργων



Δημοτική Επιχείρηση Ύδρευσης και Αποχέτευσης  
Καρδίτσας



Αειφορική Δωδεκανήσου Α.Ε.



Άδωνις Κοντός και ΣΙΑ Ε.Ε. (Marathon Data Systems)

## Ενότητα Εργασίας 2

Ανάπτυξη συστήματος διαχείρισης και επεξεργασίας  
δεδομένων

## Τεύχος 2

Περιγραφή συστήματος διαχείρισης και επεξεργασίας  
δεδομένων «Υδρογνώμων»



## ΠΕΡΙΛΗΨΗ

Ο «Υδρογνώμων» είναι ένα εργαλείο πληροφορικής για τη διαχείριση και ανάλυση της υδρολογικής πληροφορίας. Έχει κατασκευαστεί πάνω σε τυπική πλατφόρμα Windows, που βασίζεται σε αρχιτεκτονική client-server. Ο server είναι μια βάση δεδομένων όπου αποθηκεύονται τα υδρολογικά δεδομένα, ενώ διάφοροι σταθμοί εργασίας εκτελούν το πρόγραμμα, έχοντας πρόσβαση σε κοινή πληροφορία. Η ανάκτηση, επεξεργασία και οπτικοποίηση των δεδομένων υποστηρίζεται από πολυγλωσσικό γραφικό περιβάλλον εργασίας. Η διαχείριση των δεδομένων βασίζεται στη γεωγραφική οργάνωση οντοτήτων όπως μετρητικοί σταθμοί, λεκάνες απορροής και ταμιευτήρες. Σε κάθε οντότητα αντιστοιχούν χρονοσειρές, φυσικές ιδιότητες, υπολογιστικές παράμετροι, οπτικοακουστικό υλικό, κτλ. Η κύρια ενότητα της ανάλυσης υδρολογικών δεδομένων περιλαμβάνει εφαρμογές επεξεργασίας χρονοσειρών, όπως συνάθροιση και κανονικοποίηση χρονικού βήματος, παρεμβολή, ανάλυση παλινδρόμησης και συμπλήρωση ελλειπουσών τιμών, ελέγχους συνέπειας, φιλτράρισμα δεδομένων, οπτικοποίηση χρονοσειρών με χρήση γραφημάτων και πινάκων, κτλ. Το πρόγραμμα υποστηρίζει ακόμη εξειδικευμένες υδρολογικές εφαρμογές, στις οποίες περιλαμβάνονται μοντέλα εξατμοδιαπνοής, κατασκευή καμπυλών στάθμης-παροχής, έλεγχοι ομοιογένειας, ανάλυση υδατικού ισοζυγίου, κτλ. Η στατιστική ενότητα παρέχει εργαλεία για ανάλυση δειγμάτων, συναρτήσεις κατανομής, στατιστική πρόγνωση, προσομοίωση Monte-Carlo, ανάλυση ακραίων γεγονότων και κατασκευή όμβριων καμπυλών. Μια τελευταία ενότητα περιλαμβάνει ένα συγκεντρωτικό υδρολογικό μοντέλο, με εναλλακτικές διαμορφώσεις, που επιπλέον υποστηρίζεται από διαδικασίες αυτόματης βαθμονόμησης. Το παρόν είναι το τεύχος της θεωρητικής τεκμηρίωσης του συστήματος του «Υδρογνώμων».

## ABSTRACT

“Hydrognomon” is a software tool for the management and analysis of hydrological data. It is built on a standard Windows platform based on client-server architecture; a database server is holding hydrological data whereas several workstations are executing Hydrognomon, sharing common data. Data retrieval, processing and visualisation are supported by a multilingual Graphical User Interface. Data management is based on geographical organisation to entities such as measuring stations, river basins, and reservoirs. Each entity may possess time series, physical properties, calculation parameters, multimedia content, etc. The main part of hydrological data analysis consists of time series processing applications, such as time step aggregation and regularisation, interpolation, regression analysis and filling in of missing values, consistency tests, data filtering, graphical and tabular visualisation of time series, etc. The program supports also specific hydrological applications, including evapotranspiration modelling, stage-discharge analysis, homogeneity tests, water balance methods, etc. The statistical module provides tools for sampling analysis, distribution functions, statistical forecast, Monte-Carlo simulation, analysis of extreme events and construction of intensity-duration-frequency curves. A final module is a lumped hydrological model, with alternative configurations, also supported by automatic calibration facilities. This report is the scientific documentation of the “Hydrognomon” system.



## Πίνακας περιεχομένων

<b>1 Το σύστημα διαχείρισης και επεξεργασίας δεδομένων (χρονοσειρών) «Υδρογνώμων» 1</b>	
<b>2 Τυποποίηση δεδομένων χρονοσειρών..... 5</b>	
<b>2.1 Εισαγωγή..... 5</b>	
<b>2.2 Χρονικό βήμα..... 5</b>	
2.2.1 Χρήση υδρολογικού έτους – εποχιακά συναθροίσματα..... 6	
2.2.2 Αυστηρό χρονικό βήμα (Timestep strict) – σταθερή χρονική ολίσθηση ..... 7	
<b>2.3 Τυποποίηση μεταβλητών ..... 7</b>	
2.3.1 Κενές (null) εγγραφές – ελλείψεις ..... 9	
<b>2.4 Σημαιοθέτηση..... 9</b>	
<b>2.5 Χρήση γεωγραφικών οντοτήτων..... 10</b>	
2.5.1 Απεικόνιση γεωγραφικών οντοτήτων – γεωγραφικές επεξεργασίες..... 10	
<b>3 Έλεγχοι και αποκατάσταση της συνέπειας των δεδομένων..... 13</b>	
<b>3.1 Εισαγωγή..... 13</b>	
<b>3.2 Έλεγχοι ακραίων τιμών και χρονικής συνέπειας..... 13</b>	
3.2.1 Έλεγχος ακραίων τιμών ..... 13	
3.2.2 Έλεγχος χρονικής συνέπειας..... 14	
<b>3.3 Έλεγχος ομοιογένειας – ομογενοποίηση δεδομένων με την μέθοδο της διπλής αθροιστικής καμπύλης..... 14</b>	
3.3.1 Γραφική κατάρτιση της διπλής αθροιστικής καμπύλης ..... 15	
3.3.2 Αποκατάσταση της ομοιογένειας - ανόρθωση ..... 15	
3.3.3 Εντοπισμός εξωκείμενων ετησίων τιμών ..... 17	
<b>4 Χειρισμοί χρονοσειρών..... 19</b>	
<b>4.1 Εισαγωγή..... 19</b>	
<b>4.2 Κανονικοποίηση χρονικού βήματος (άρση χρονικών ολισθήσεων) ..... 19</b>	
4.2.1 Εναλλακτική μεθοδολογία ..... 21	
<b>4.3 Συνάθροιση – εξαγωγή χρονοσειρών μεγαλύτερου χρονικού βήματος (χρονική ολοκλήρωση)..... 22</b>	
<b>4.4 Γραμμικές πράξεις χρονοσειρών ..... 23</b>	
<b>4.5 Απεικόνιση – πλέγματα χρονοσειρών ..... 24</b>	
4.5.1 Προβολή ημερήσιων τιμών σε πίνακα για ένα έτος..... 24	
4.5.2 Προβολή των μηνιαίων τιμών σε πίνακα ..... 24	
4.5.3 Προβολή δύο ή περισσότερων χρονοσειρών σε συγκριτικό πίνακα (πλέγμα χρονοσειρών– timeseries grid)..... 25	
4.5.4 Δημιουργία διαγραμμάτων..... 26	
<b>5 Καμπύλες στάθμης – παροχής, παρεμβολές..... 27</b>	
<b>5.1 Εισαγωγή..... 27</b>	
<b>5.2 Καμπύλες στάθμης – όγκου επιφάνειας ταμιευτήρων ..... 27</b>	
<b>5.3 Εκτίμηση υπόγειων διαφυγών ταμιευτήρων ..... 27</b>	

<b>5.4</b>	<b>Εξαγωγή παροχής από μετρήσεις στάθμης και καμπύλες στάθμης - παροχής.....</b>	<b>28</b>
5.4.1	Η διόρθωση της στάθμης .....	30
5.4.2	Διόρθωση στάθμης για τα σφάλματα των σταθμηγράφων .....	32
5.4.3	Διόρθωση στάθμης για τις διαφορές κατά τις μετρήσεις παροχής.....	34
5.4.4	Η εξαγωγή των παροχών.....	35
<b>5.5</b>	<b>Κατάρτιση καμπυλών στάθμης – παροχής .....</b>	<b>36</b>
5.5.1	Ορισμός χρονικών περιόδων – εντοπισμός εξωκείμενων σημείων.....	36
5.5.2	Προσαρμογή καμπυλών στις υδρομετρήσεις.....	37
5.5.3	Εισαγωγή καμπυλών επέκτασης με εξισώσεις της υδραυλικής.....	38
<b>6</b>	<b>Μοντέλα εκτίμησης της εξατμοδιαπνοής.....</b>	<b>43</b>
<b>6.1</b>	<b>Εισαγωγή.....</b>	<b>43</b>
<b>6.2</b>	<b>Τα μοντέλα Penman / Penman - Monteith .....</b>	<b>43</b>
6.2.1	Υπολογισμοί φυσικών μεταβλητών .....	43
6.2.2	Το μοντέλο του Penman.....	47
6.2.3	Το μοντέλο των Penman – Monteith.....	47
6.2.4	Αριθμητική εφαρμογή – έλεγχος αποτελεσμάτων .....	47
<b>6.3</b>	<b>Τα μοντέλα Thornthwaite, Hargreaves και Blaney-Criddle .....</b>	<b>48</b>
6.3.1	Η μέθοδος Thornthwaite .....	48
6.3.2	Η μέθοδος Hargreaves.....	49
6.3.3	Η μέθοδος Blaney-Criddle .....	49
<b>6.4</b>	<b>Επέκταση και συμπλήρωση χρονοσειρών εξατμοδιαπνοής.....</b>	<b>50</b>
6.4.1	Μεθοδολογία προσαρμογής.....	50
6.4.2	Εφαρμογή.....	52
<b>7</b>	<b>Συσχέτιση – συμπλήρωση – επέκταση χρονοσειρών.....</b>	<b>53</b>
<b>7.1</b>	<b>Εισαγωγή.....</b>	<b>53</b>
<b>7.2</b>	<b>Μέθοδοι προσδιορισμού των παραμέτρων της παλινδρόμησης.....</b>	<b>53</b>
7.2.1	Μέση τιμή .....	53
7.2.2	Απλή – πολλαπλή γραμμική παλινδρόμηση .....	53
7.2.3	Ομογενής ευθεία .....	55
7.2.4	Οργανική συσχέτιση .....	56
7.2.5	Αυτοσυσχέτιση.....	56
7.2.6	Εποχικότητα.....	57
<b>7.3</b>	<b>Συμπλήρωση - επέκταση .....</b>	<b>57</b>
7.3.1	Εισαγωγή τυχαίου όρου .....	59
7.3.2	Αποκοπή στο μηδέν .....	60
<b>8</b>	<b>Υδατικά ισοζύγια.....</b>	<b>61</b>
<b>8.1</b>	<b>Εισαγωγή.....</b>	<b>61</b>
<b>8.2</b>	<b>Ισοζύγια υδάτινων φορέων .....</b>	<b>61</b>
8.2.1	Ισοζύγια ταμιευτήρων .....	61
8.2.2	Ισοζύγια υδραγωγείων .....	63
<b>8.3</b>	<b>Προσομοίωση διεργασιών λεκάνης – το μοντέλο «Ζυγός».....</b>	<b>63</b>
8.3.1	Περιγραφή του μοντέλου .....	63



8.3.2	Περιγραφή υπολογιστικής διαδικασίας.....	67
8.3.3	Βαθμονόμηση του μοντέλου.....	68
<b>9</b>	<b>Στατιστική επεξεργασία χρονοσειρών.....</b>	<b>73</b>
<b>9.1</b>	<b>Εισαγωγή.....</b>	<b>73</b>
<b>9.2</b>	<b>Εκτίμηση στατιστικών χαρακτηριστικών δείγματος.....</b>	<b>74</b>
9.2.1	Αντιμετώπιση των αρνητικών τιμών του δείγματος.....	74
9.2.2	Οι δειγματικές ροπές ως στατιστικοί δείκτες.....	75
9.2.3	Η χρήση των L-ροπών.....	76
<b>9.3</b>	<b>Οι εμπειρικές συναρτήσεις κατανομής.....</b>	<b>77</b>
<b>9.4</b>	<b>Οι θεωρητικές συναρτήσεις κατανομής.....</b>	<b>79</b>
9.4.1	Υπολογισμοί συναρτήσεων κατανομής.....	81
9.4.2	Στατιστική πρόγνωση.....	92
9.4.3	Μεθοδολογία εκτίμησης των παραμέτρων των θεωρητικών κατανομών.....	93
<b>9.5</b>	<b>Μετασχηματισμοί - απεικόνιση.....</b>	<b>102</b>
<b>9.6</b>	<b>Στατιστικοί έλεγχοι.....</b>	<b>103</b>
9.6.1	Η δοκιμή $\chi^2$ .....	103
9.6.2	Ο έλεγχος Kolmogorov-Smirnov.....	104
<b>9.7</b>	<b>Εκτίμηση των διαστημάτων εμπιστοσύνης μέσω στατιστικής προσομοίωσης (Monte – Carlo).....</b>	<b>105</b>
9.7.1	Γεννήτριες τυχαίων αριθμών.....	106
9.7.2	Εύρεση διαστημάτων εμπιστοσύνης του δείγματος.....	108
9.7.3	Εύρεση διαστημάτων εμπιστοσύνης του πληθυσμού.....	111
<b>10</b>	<b>Ανάλυση εξαιρετικών υδρολογικών γεγονότων – κατάρτιση όμβριων καμπυλών..</b>	<b>115</b>
<b>10.1</b>	<b>Εισαγωγή.....</b>	<b>115</b>
<b>10.2</b>	<b>Εκτίμηση μηνιαίων – ετήσιων μεγίστων μέσω εντάσεων.....</b>	<b>115</b>
<b>10.3</b>	<b>Έλεγχοι - αναγωγές.....</b>	<b>116</b>
10.3.1	Έλεγχος συνέπειας χρονοσειρών.....	116
10.3.2	Αναγωγές χρονικής ευκρίνειας – σειρών υπεράνω κατωφλίου.....	117
<b>10.4</b>	<b>Στατιστικές κατανομές.....</b>	<b>118</b>
<b>10.5</b>	<b>Κατάρτιση όμβριων καμπυλών με την συμβατική μέθοδο.....</b>	<b>120</b>
<b>10.6</b>	<b>Κατάρτιση όμβριων καμπυλών με την μέθοδο ενοποίησης διαρκειών.....</b>	<b>121</b>
10.6.1	Ενοποίηση διαρκειών – εκτίμηση των $\eta$ και $\theta$ .....	121
10.6.2	Προσδιορισμός των παραμέτρων της κατανομής.....	124
10.6.3	Τα όρια εμπιστοσύνης των όμβριων καμπυλών.....	125



## Πίνακας σχημάτων

Σχήμα 1.1: Σχηματική αναπαράσταση του Συστήματος Διαχείρισης και Επεξεργασίας Χρονοσειρών «Υδρογνώμων» (Kozanis, 2005).....	2
Σχήμα 3.1: Ανόρθωση του δείγματος βροχόπτωσης του σταθμού «Άνω Μπράλος» χρησιμοποιώντας ως σταθμό βάσης τον σταθμό «Αταλάντη».....	17
Σχήμα 3.2: Εντοπισμός εξωκείμενης τιμής μέσω της διπλής αθροιστικής καμπύλης.....	18
Σχήμα 4.1: Τεχνική άρσης χρονικών μετατοπίσεων.....	20
Σχήμα 4.2: Συνάθροιση ωριαίας χρονοσειράς μέσω των τιμών θερμοκρασίας σε ημερήσια, μηνιαία και ετήσια χρονοσειρά μέσω των τιμών. Σταθμός Πολυτεχνειούπολης Ζωγράφου.....	23
Σχήμα 5.1: Σχηματική αναπαράσταση της διόρθωσης Stout.....	31
Σχήμα 5.2: Διόρθωση Stout σύμφωνα με την Shaw (1994).....	31
Σχήμα 5.3: Διάφορες περιπτώσεις μεταβολής της διόρθωσης στάθμης.....	31
Σχήμα 5.4: Διαχωρισμός των υδρομετρήσεων σε ομοιογενείς περιόδους (με διαφορετικό χρωματισμό). Υδρομετρήσεις 1992-1998 από τον σταθμό «Πόρος Ρηγανίου» στον ποταμό Εύηνο...	36
Σχήμα 5.5: Μία εν δυνάμει καμπύλη στάθμης παροχής. Τα $X_i$ είναι δεδομένα και αναζητείται το βέλτιστο σετ $Y_i$ .....	37
Σχήμα 5.6: Χάραξη καμπυλών στάθμης παροχής. Καμπύλες με μία θλάση προσαρμοσμένες στον διαχωρισμό των περιόδων του Σχήματος 5.4.....	38
Σχήμα 5.7: Η υγρή διατομή σε σχέση με την διατομή του υδατορεύματος και την στάθμη του νερού.....	40
Σχήμα 5.8: Επιτρεπτές και προβληματικές διατομές για τον υπολογισμό των στοιχείων της υγρής διατομής.....	41
Σχήμα 5.9: Κατάρτιση καμπύλης στάθμης παροχής στο περιβάλλον του «Υδρογνώμων» για μία κλειστή διατομή. Χρήση σχέσης του Manning με $n=0.0015$ , $i=0.012$ .....	41
Σχήμα 6.1: Προσαρμογή του παραμετρικού μοντέλου στο δείγμα εξάτμισης του σταθμού του Ε.Μ.Π.....	52
Σχήμα 8.1: Σχηματική αναπαράσταση της γενικής μορφής ισοζυγίου ταμειυτήρα.....	61
Σχήμα 8.2: Σχηματική παράσταση αδρομερούς μοντέλου βροχόπτωσης – απορροής «Ζυγός».....	64
Σχήμα 8.3: Χρονοσειρές δυνητικής εξατμοδιαπνοής, βροχόπτωσης και αντλήσεων στην λεκάνη του Βοιωτικού Κηφισού.....	71
Σχήμα 8.4: Σύγκριση μεταξύ μετρημένης και υπολογισμένης απορροής στην λεκάνη του Βοιωτικού Κηφισού.....	71
Σχήμα 9.1: Εμπειρικές κατανομές Weibull, Blom, Cunnane, Gringorten. Ετήσια μέγιστα έντασης ωριαίων βροχοπτώσεων. Σταθμός «Πολυτεχνειούπολης» 1993-4 – 2003-4.....	78
Σχήμα 9.2: Κανονική, λογαριθμοκανονική και Galton κατανομή. Σημεία εμπειρικής πιθανότητας Weibull. Δείγμα: ετήσιες βροχοπτώσεις στον σταθμό «Αλίαρτος», 1907-8 – 2000-1.....	88
Σχήμα 9.3 Εκθετική, γάμα, Pearson III, Log-Pearson III κατανομή. Σημεία εμπειρικής πιθανότητας Weibull. Δείγμα: βροχοπτώσεις Απριλίου στον σταθμό «Αμφισσα», 1955 – 2000.....	89
Σχήμα 9.4 AT-1, AT-2, ΓAT κατανομές μεγίστων. Σημεία εμπειρικής πιθανότητας Weibull. Δείγμα: μέγιστες ετήσιες εντάσεις βροχής εξάωρου στον σταθμό «Ελληνικό», 1957-58 – 1986-87 (Πηγή Εθνική Μετεωρολογική Υπηρεσία).....	90
Σχήμα 9.5: Ελάχιστες μηνιαίες παροχές περιόδου Οκτώβριος – Μάρτιος. AT-1, AT-3, ΓAT κατανομές ελαχίστων. Μηνιαία παροχή Β. Κηφισού («Διώρυγα Καρδίτσας»), 1906-7 – 2003-4.....	91
Σχήμα 9.6: Όρια εμπιστοσύνης $x_U$ , $x_L$ , πειράματα $x_k$ και κατανομή τους περί του σημείου $x(u_d)$ . ....	109
Σχήμα 9.7: Διαστήματα εμπιστοσύνης 95% κανονικής κατανομής του δείγματος και του πληθυσμού, για το δείγμα: $x_i=i$ , $i=1..19$ , $n=19$ .....	112

Σχήμα 10.1: Όμβριες καμπύλες για περιόδους επαναφοράς $T=5, 10, 50, 500, 5000$ στο Ελληνικό, υπολογισμένες από την συνεπή ενιαία εξίσωση και συνάρτηση κατανομής AT-1 μεγίστων.....	124
Σχήμα 10.2: Συναρτήσεις κατανομής Gumbel (max) και Weibull (σημεία εμπειρικής κατανομής) για τον έλεγχο προσαρμογής της κατανομής .....	124
Σχήμα 10.3: Διαστήματα εμπιστοσύνης 95% της όμβριας καμπύλης για $T=50$ . Χρήση κατανομής AT-1 μεγίστων (Gumbel) στο δείγμα του Ελληνικού.....	125

## Πίνακας πινάκων

Πίνακας 2.1: Τύποι δεδομένων χρονοσειρών .....	5
Πίνακας 2.2: Τύποι χρονικού βήματος χρονοσειρών.....	6
Πίνακας 2.3: Τύποι μεταβλητής των χρονοσειρών .....	7
Πίνακας 2.4: Χρονικές στιγμές (timestamps) – χρονικά διαστήματα που αναφέρονται οι μεταβλητές .	8
Πίνακας 2.5: Σημαίες των δεδομένων των χρονοσειρών.....	9
Πίνακας 4.1: Μεθοδολογίες παραγωγής χρονοσειράς με μεγαλύτερο χρονικό βήμα.....	22
Πίνακας 4.2: Απεικόνιση ημερησίων τιμών χρονοσειράς για συγκεκριμένο έτος.....	24
Πίνακας 4.3: Απεικόνιση μηνιαίων τιμών χρονοσειράς .....	25
Πίνακας 6.1: Συγκρίσεις μεταξύ των αποτελεσμάτων των μεθόδων Penman, Penman Monteith. Δεδομένα από Κουτσογιάννη και Ξανθόπουλο (1999, σ. 210).....	48
Πίνακας 6.2: Σύγκριση αποτελεσμάτων για την μέθοδο υπολογισμού της εξατμοδιαπνοής του Thornthwaite. Δεδομένα από Κουτσογιάννη και Ξανθόπουλο (1999, σ. 228). .....	49
Πίνακας 8.1: Παράμετροι του αδρομερούς μοντέλου βροχόπτωσης – απορροής «Ζυγός».....	66
Πίνακας 8.2: Οι παράμετροι ελέγχου του εξελκτικού αλγόριθμο ανόπτωσης-απλόκου.....	70
Πίνακας 8.3: Ενδεικτική ρύθμιση μοντέλου «Ζυγός», όρια παραμέτρων και τελικές τιμές. Λεκάνη Βοιωτικού Κηφισού. ....	72
Πίνακας 9.1: Απλές (μεροληπτικές) εκτιμήτριες των δειγματικών ροπών μέχρι τρίτης τάξης .....	75
Πίνακας 9.2: Συντελεστές διόρθωσης μεροληψίας.....	76
Πίνακας 9.3: Οι εμπειρικές συναρτήσεις κατανομής (πιθανότητες μη-υπέρβασης ταξινομημένου δείγματος σε αύξουσα σειρά) που χρησιμοποιούνται στον «Υδρογνώμων» .....	78
Πίνακας 9.4: Οι συναρτήσεις κατανομής που περιλαμβάνονται στον «Υδρογνώμων» μαζί με τον αριθμό των παραμέτρων.....	80
Πίνακας 9.5: Πεδίο ορισμού της τυχαίας μεταβλητής ανά κατανομή .....	81
Πίνακας 9.6: Υπολογισμοί συναρτήσεων κατανομής με κλειστή ή προσεγγιστική λύση .....	82
Πίνακας 9.7: Επαλήθευση αλγορίθμων μέσω των τιμών της (τυποποιημένης) κανονικής κατανομής	84
Πίνακας 9.8: Επαλήθευση αλγορίθμων μέσω των τιμών της κατανομής Γ.....	87
Πίνακας 9.9: Παράδειγμα στατιστικής πρόγνωσης. Βροχοπτώσεις Οκτωβρίου από τον σταθμό «Αλιάρτος» 1907-8 – 2000-1. ....	93
Πίνακας 9.10: Μετασχηματισμοί πιθανοτήτων για την απεικόνιση σε «χαρτιά πιθανοτήτων» .....	102
Πίνακας 9.11: Ενδεικτικά αποτελέσματα από την εφαρμογή του ελέγχου $\chi^2$ στο δείγμα βροχοπτώσεων Οκτωβρίου του σταθμού «Αλιάρτος». Μέγεθος δείγματος $n=93$ , κλάσεων $k=11$ . .....	104
Πίνακας 9.12: Ενδεικτικά αποτελέσματα από την εφαρμογή του ελέγχου Kolmogorov - Smirnov στο δείγμα βροχοπτώσεων Οκτωβρίου του σταθμού «Αλιάρτος». Μέγεθος δείγματος $n=93$ . .....	105
Πίνακας 9.13: Σύγκριση ορίων εμπιστοσύνης που προκύπτουν με την μέθοδο της προσομοίωσης και της θεωρητικής τιμής για την μέση τιμή (κανονική κατανομή, $u_d=50\%$ , δείγμα $x_i=i$ , $i=1..19$ , $n=19$ ). .....	110
Πίνακας 9.14: Σύγκριση ορίων εμπιστοσύνης του πληθυσμού που προκύπτουν με την μέθοδο της προσομοίωσης και της θεωρητικής τιμής για την μέση τιμή (κανονική κατανομή, $u_d=50\%$ , δείγμα $x_i=i$ , $i=1..19$ , $n=19$ ). .....	112
Πίνακας 10.1: Τιμές του συντελεστή άρσης του σφάλματος διακριτοποίησης .....	117
Πίνακας 10.2: Οι κατανομές που χρησιμοποιούνται για την κατάρτιση των όμβριων καμπυλών, συναρτήσει της περιόδου επαναφοράς $T$ . Γενικευμένες εξισώσεις όμβριων καμπυλών που προκύπτουν από την διαίρεση με κάποιον όρο $b(d)$ ( $d$ : διάρκεια βροχόπτωσης). .....	119
Πίνακας 10.3: Κατανομές ΑΤ που χρησιμοποιούνται στις όμβριες καμπύλες, για (ανηγμένη) περίοδο επαναφοράς $T'$ που θα αντιστοιχούσε σε χρονοσειρά μεγίστων υπεράνω κατωφλίου $T'$ .....	120
Πίνακας 10.4: Καταγραμμένες ετήσιες μέγιστες εντάσεις βροχής στο σταθμό Ελληνικό (Πηγή Εθνική Μετεωρολογική Υπηρεσία).....	123



# 1 Το σύστημα διαχείρισης και επεξεργασίας δεδομένων (χρονοσειρών) «Υδρογνώμων»

---

Ο «Υδρογνώμων», το σύστημα δηλαδή πρόσβασης στη βάση δεδομένων, διαχείρισης των υδρολογικών, μετεωρολογικών και ποιοτικών δεδομένων και επεξεργασίας των χρονοσειρών είναι μία αυτόνομη εφαρμογή λογισμικού. Παρέχει τη δυνατότητα σύνδεσης με κοινές εφαρμογές όπως Συστήματα Γεωγραφικών Πληροφοριών (G.I.S.) και λογιστικά φύλλα (spreadsheets), που μπορούν να πραγματοποιούν επιπλέον επεξεργασίες ανάλογα με τις ανάγκες του χρήστη. Επιπλέον, το σύστημα μπορεί να προσπελαίνεται από το Διαδίκτυο μέσω συμβατικών εφαρμογών που θα προσπελαίνουν τη βάση απευθείας.

Στο παρόν τεύχος παρουσιάζονται οι μεθοδολογίες και οι αλγόριθμοι που ακολουθούνται για την διαχείριση και την επεξεργασία των δεδομένων και αποτελούν την επιστημονική τεκμηρίωση της εφαρμογής λογισμικού. Επιπλέον τίθενται οι προδιαγραφές τυποποίησης και απεικόνισης των δεδομένων. Η ανάπτυξη του συστήματος βασίστηκε:

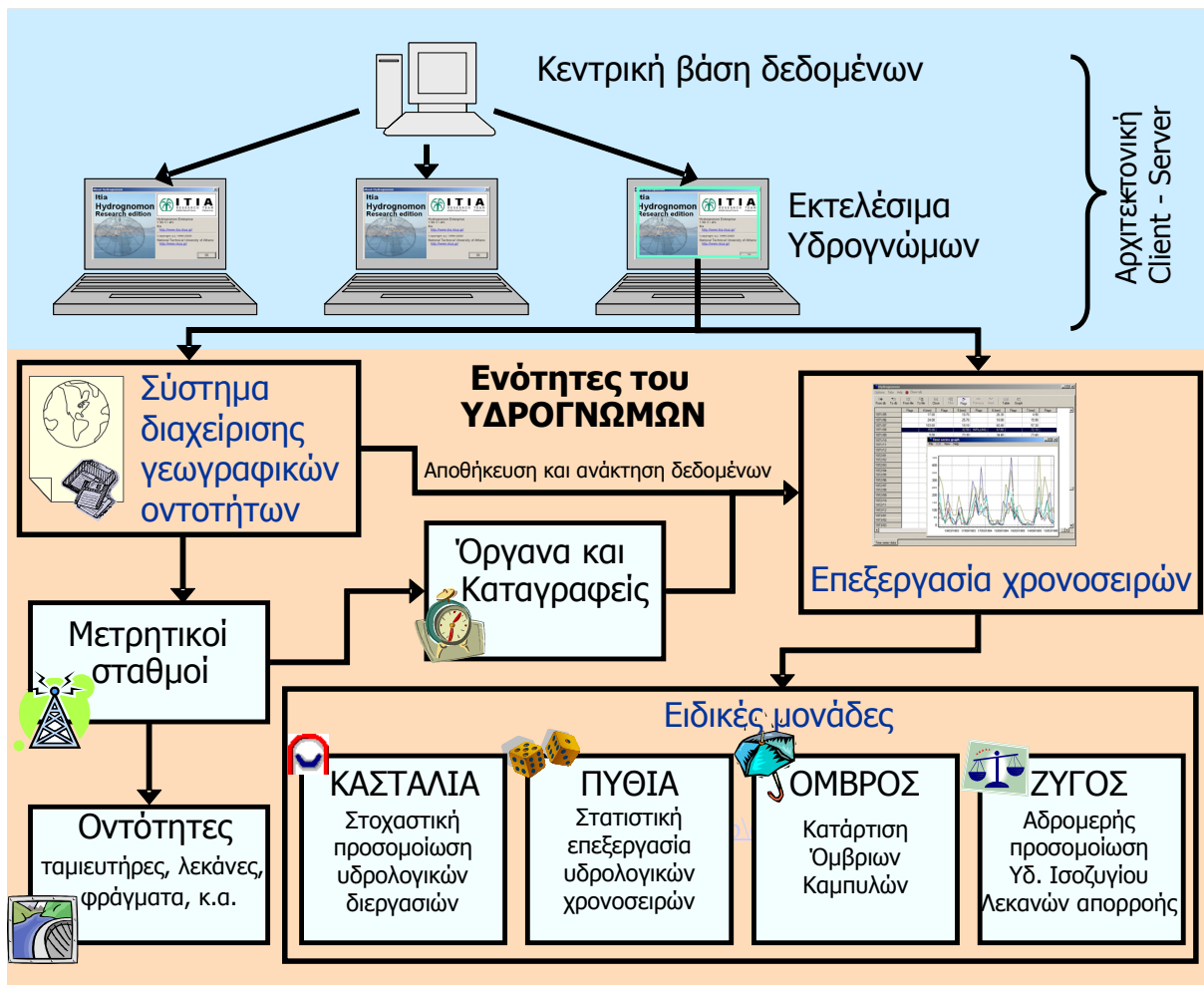
- Στην ερευνητική εμπειρία της ομάδας (<http://www.itia.ntua.gr/>)
- Στην εμπειρία από την ανάπτυξη παλαιότερης έκδοσης του λογισμικού (Χριστοφίδης, 1998, Χριστοφίδης και Κοζάνης, 2004), καθώς και από δοκιμαστικές εκδόσεις ανάπτυξης (Kozanis et al., 2005)
- Στον σχεδιασμό της βάσης δεδομένων (Χριστοφίδης, κ.α., 2005) καθώς και σε αξιοποίηση εμπειρίας και γνώσεων από τον παλαιότερο σχεδιασμό (Παπακώστας, 2004)
- Σε εκτεταμένη έρευνα αγοράς πάνω σε παρόμοια συστήματα (έως και Σεπτέμβριος 2005)
- Στην διεθνή βιβλιογραφία στα θέματα Υδρολογίας και Πληροφορικής.

Κατά τον σχεδιασμό του συστήματος, λαμβάνονται υπόψη οι αυξημένες ανάγκες διαχείρισης και επεξεργασίας των τηλεμετρικών δεδομένων, που αποτελούν εξαιρετικά μεγάλο όγκο πληροφοριών σε πολύ μικρό χρονικό βήμα. Για το λόγο αυτόν, τίθενται ιδιαίτερες προδιαγραφές, με αυστηρή μοντελοποίηση των χρονικών βημάτων και των διαδικασιών καθώς και κατάρτιση των κατάλληλων αλγορίθμων.

Το εν λόγω σύστημα δεν πραγματοποιεί προχωρημένες επεξεργασίες, όπως παραγωγή συνθετικών χρονοσειρών (η συγκεκριμένη λειτουργία παρέχεται από το υποσύστημα «Κασταλία» το οποίο ενσωματώνεται στον «Υδρογνώμων» ωστόσο η τεκμηρίωσή του περιλαμβάνεται σε ξεχωριστό τεύχος), αλλά απλές (τυπικές υδρολογικές) επεξεργασίες, οι οποίες λειτουργούν βοηθητικά. Οι σημαντικότερες λειτουργίες που επιτελεί το σύστημα είναι οι ακόλουθες:

- Μετατροπή χρονοσειρών σε σταθερό χρονικό βήμα (ως γνωστόν, οι πρωτογενείς χρονοσειρές έχουν κάποια σταθερότητα, αλλά συχνά παρουσιάζουν διαταραχές, η εξάλειψη των οποίων είναι αναγκαία για περαιτέρω επεξεργασία)
- Εξαγωγή χρονοσειρών μεγαλύτερου χρονικού βήματος (συνάθροιση), π.χ. ωριαίων από δεκάλεπτες, ημερήσιων από ωριαίες, μηνιαίων από ημερήσιες
- Τυπικοί έλεγχοι συνέπειας όπως ομοιογένειας, ακραίων τιμών και χρονικής συνέπειας
- Γραμμική παλινδρόμηση μεταξύ χρονοσειρών, πολλαπλή παλινδρόμηση, οργανική συσχέτιση και αυτοσυσχέτιση
- Υδατικά ισοζύγια: ταμειυτήρων, αγωγών καθώς και αδρομερές μοντέλο βροχής – απορροής

- Συμπλήρωση ελλειπουσών τιμών με χρήση της γραμμικής παλινδρόμησης, δυνατότητα εισαγωγή τυχαίου όρου για διατήρηση των στατιστικών χαρακτηριστικών. Επέκταση χρονοσειρών
- Γραμμικές πράξεις μεταξύ χρονοσειρών
- Κατάρτιση καμπυλών στάθμης – παροχής με στατιστικές μεθόδους και καμπυλών επέκτασης με χρησιμοποίηση υδραυλικών εξισώσεων
- Εξαγωγή χρονοσειρών παροχών από χρονοσειρές στάθμης, καθώς και χρονοσειρών όγκου και επιφανείας από χρονοσειρές στάθμης ταμιευτήρων και λιμνών
- Υπολογισμός εξάτμισης και δυνητικής εξατμοδιαπνοής με αναλυτικές ή ημιεμπειρικές μεθόδους
- Επέκταση δειγμάτων εξατμοδιαπνοής
- Εύρεση στατιστικών χαρακτηριστικών δείγματος χρονοσειράς, προσαρμογή στατιστικών παραμέτρων, στατιστικές προγνώσεις, στατιστικοί έλεγχοι και εύρεση διαστημάτων εμπιστοσύνης
- Ανάλυση χρονοσειρών εξαιρετικών βροχοπτώσεων - κατάρτιση ομβρίων καμπυλών με συνεπείς μεθοδολογίες.



Σχήμα 1.1: Σχηματική αναπαράσταση του Συστήματος Διαχείρισης και Επεξεργασίας Χρονοσειρών «Υδρογνώμων» (Kozanis, 2005)

Στο Σχήμα 1.1 απεικονίζονται τα υποσυστήματα που συνιστούν τον «Υδρογνώμων». Οι χρονοσειρές οργανώνονται βάσει του υποσυστήματος των γεωγραφικών οντοτήτων (υποσύνολο των οποίων είναι οι μετρητικοί σταθμοί). Οι μετρητικοί σταθμοί στην συνέχεια περιλαμβάνουν έναν αριθμό οργάνων



και καταγραφών. Η επεξεργασία των χρονοσειρών υποβοηθείται από ειδικά υποσυστήματα που έχουν ονομαστεί «Κασταλία» (υπεύθυνη για την στοχαστική προσομοίωση των υδρολογικών διεργασιών), «Πυθία» (υπεύθυνη για την στατιστική επεξεργασία των υδρολογικών χρονοσειρών), «Όμβρος» (υπεύθυνη για την κατάρτιση των όμβριων καμπυλών) και «Ζυγός» (υπεύθυνος για την αδρομερή προσομοίωση του υδατικού ισοζυγίου λεκανών απορροής). Η διαθεσιμότητα κάποιων από τα υποσυστήματα μπορεί να είναι περιορισμένη, ανάλογα με την άδεια χρήσης που διαθέτει ο χρήστης.

Τέλος αναφέρουμε την δυνατότητα αυτονόμησης του συστήματος από την σχεσιακή βάση δεδομένων αποθηκεύοντας τα δεδομένα των χρονοσειρών σε κοινά αρχεία κειμένου (ASCII files), τα οποία παρέχουν επιπλέον το πλεονέκτημα της εύκολης πρόσβασης και από τρίτα λογισμικά.



## 2 Τυποποίηση δεδομένων χρονοσειρών

### 2.1 Εισαγωγή

Η τυποποίηση των δεδομένων συσχετίζεται άμεσα με το μοντέλο διαχείρισης και αποθήκευσης που έχει υλοποιηθεί στην σχεσιακή βάση δεδομένων (Χριστοφίδης, κ.α., 2005) με την οποία συνεργάζεται ο «Υδρογνώμων». Η τυποποίηση υλοποιείται με τέτοιο τρόπο έτσι ώστε να ικανοποιούνται όλες οι ανάγκες που αφορούν την μέτρηση και αποθήκευση των πρωτογενών δεδομένων, την εξαγωγή επεξεργασμένων δεδομένων και κλιματικών χαρακτηριστικών, την γένεση συνθετικών χρονοσειρών, την κατάρτιση υδατικών ισοζυγίων καθώς και των υπόλοιπων συνιστωσών για την ορθολογική διαχείριση των υδατικών πόρων. Όπως θα δούμε στην συνέχεια το σημαντικότερο χαρακτηριστικό σύμφωνα με το οποίο γίνεται η τυποποίηση είναι το χρονικό βήμα. Άλλα σημαντικά χαρακτηριστικά τυποποίησης που θα παρουσιαστούν είναι ο τύπος της μεταβλητής, επιπλέον χαρακτηριστικά του χρονικού βήματος (ακανόνιστο – σταθερό χρονικό βήμα, χρήση υδρολογικού έτους). Τέλος παρουσιάζονται οι σημαίες που προσφέρονται για ειδικές επισημάνσεις στις τιμές των χρονοσειρών.

Οι βασικοί τύποι δεδομένων χρονοσειρών που αποθηκεύονται στην βάση δεδομένων (πρωτογενή, επεξεργασμένα και συνθετικά) παρουσιάζονται στον Πίνακα 2.1. Τα πρωτογενή δεδομένα διαχειρίζονται μόνο από το σύστημα του «Υδρογνώμων», τα επεξεργασμένα κυρίως από τον «Υδρογνώμων», ενώ τα συνθετικά από το υποσύστημα «Κασταλία» χρησιμοποιούνται δε σε σενάρια διαχείρισης του συστήματος του «Υδρονομέα». Τα συνθετικά δεδομένα έχουν επιπλέον την ιδιαιτερότητα των πολλαπλών προγνώσεων για την ίδια χρονοσειρά.

Πίνακας 2.1: Τύποι δεδομένων χρονοσειρών

Τύπος δεδομένων	Περιγραφή
Πρωτογενή	Οι μετρήσεις, αυτούσιες χωρίς να υποστούν καμία επεξεργασία, είτε αυτές γίνονται με συμβατικές μεθόδους είτε τηλεμετρικά – πλήρως αυτοματοποιημένα. Τα πρωτογενή δεδομένα ελέγχονται ως προς την συνέπειά τους και στην συνέχεια ανάγονται σε χρονοσειρές σταθερού χρονικού βήματος προτού υποστούν περαιτέρω επεξεργασίες.
Επεξεργασμένα	Προϊόντα επεξεργασίας των πρωτογενών δεδομένων, π.χ. συναθροισμένες χρονοσειρές, μέσες τιμές, χρονοσειρές από συμπλήρωση και αποκατάσταση της ομοιογένειας, κ.α. ή χρονοσειρές που προκύπτουν μέσω σύνθετων επεξεργασιών όπως από μοντέλα εκτίμησης της εξατμοδιαπνοής ή από υδατικά ισοζύγια.
Συνθετικά	Τα συνθετικά δεδομένα αναφέρονται είτε σε χρονοσειρές πρόγνωσης, δηλαδή σε τιμές ενός μεγέθους που αναφέρονται σε περιόδους που δεν υπάρχουν μετρήσεις π.χ. με χρήση του υποσυστήματος «Κασταλία» είτε σε χρονοσειρές σύνθεσης π.χ. από ενοποίηση δειγμάτων πολλών σταθμών σε ένα ενιαίο δείγμα.

### 2.2 Χρονικό βήμα

Το χρονικό βήμα  $T$  της χρονοσειράς, είναι το βασικότερο χαρακτηριστικό τυποποίησης. Έχει γίνει η επιλογή για το μικρότερο – απαραίτητο αριθμό χρονικών βημάτων που να καλύπτουν τις

περισσότερες περιπτώσεις επεξεργασίας – διαχείρισης που αντιμετωπίζουν τα συστήματα. Η ελαχιστοποίηση αυτή βοηθάει τόσο στην καλύτερη εποπτεία του χρήστη όσο και στην ευκολότερη επικοινωνία μεταξύ των διαφόρων συστημάτων και της βάσης δεδομένων. Τα χρονικά βήματα που υλοποιούνται συνοψίζονται στον Πίνακα .2.2 μαζί με τις δυνατές χρήσεις για κάθε ένα χωριστά.

Πίνακας 2.2: Τύποι χρονικού βήματος χρονοσειρών

Αγγλική ονομασία	Ελληνική ονομασία	Χρήση
Unknown ή Unspecified	Άγνωστο	Δεν έχει οριστεί χρονικό βήμα.
Ten-minute	Δεκάλεπτο	Πρωτογενή δεδομένα από τηλεμετρικούς σταθμούς ή από σάρωση των ταινιών καταγραφικών οργάνων.
Hourly	Ωριαίο	Πρωτογενή δεδομένα από καταγραφικά όργανα ή και συναθροισμένα δεδομένα π.χ. από δεκάλεπτα. Χρήση σε λεπτομερή μοντέλα π.χ. διόδευσης πλημμύρας.
Daily	Ημερήσιο	Πρωτογενή δεδομένα μετρημένα συνήθως με συμβατικές μεθόδους (π.χ. βροχόμετρα) ή επεξεργασμένες ή και συνθετικές χρονοσειρές. Χρήση σε λεπτομερή υδατικά ισοζύγια, μοντέλα εξατμοδιαπνοής, κ.α.
Monthly	Μηνιαίο	Επεξεργασμένες ή συνθετικές χρονοσειρές. Τυπικό χρονικό βήμα για κατάρτιση υδατικών ισοζυγίων, εκτίμηση υδατικών αναγκών, διαχείριση υδατικών πόρων εν γένει και εκτίμηση των κλιματικών χαρακτηριστικών.
Yearly ή Annual	Ετήσιο	Επεξεργασμένες ή συνθετικές χρονοσειρές. Χρήση στην διαχείριση υδατικών πόρων, στην στατιστική επεξεργασία των χρονοσειρών, στην διερεύνηση των κλιματικών τάσεων (όπως η υδρολογική εμμονή – το φαινόμενο Hurst).
Variable	Μεταβλητό ή και Ακανόνιστο	Πρωτογενείς, σποραδικές μετρήσεις ή αποτέλεσμα κάποιων επεξεργασιών (βλ. εξαγωγή παροχών από μετρήσεις στάθμης – ενότητα 5.4)

Με την χρήση της ετήσιας χρονοσειράς μπορούν να αναπαρασταθούν δεδομένα που έχουν προκύψει από εποχική συνάθροιση (π.χ. Βροχοπτώσεις Οκτωβρίου – Μαρτίου). Το μηνιαίο και το ετήσιο χρονικό βήμα έχουν την ιδιαιτερότητα για τα χρονικά διαστήματα μεταξύ των τιμών να μην είναι σταθερά. Έτσι σε μία χρονοσειρά με ετήσιο χρονικό βήμα, τα χρονικά διαστήματα είναι 365 ημερών ή 366 ημερών αν πρόκειται για δίσεκτο έτος. Παρομοίως στις μηνιαίες χρονοσειρές τα χρονικά διαστήματα είναι 28, 29, 30 ή 31 ημερών ανάλογα με τον μήνα και το έτος (στα δίσεκτα έτη). Συνεπώς οι επεξεργασίες που παράγουν αυτού του τύπου χρονοσειρές (συνάθροιση), λαμβάνουν υπόψη το θέμα του μεταβλητού χρονικού διαστήματος.

### 2.2.1 Χρήση υδρολογικού έτους – εποχιακά συναθροίσματα

Το υδρολογικό έτος είναι το χρονικό διάστημα μέσα στο οποίο πραγματοποιείται ένας πλήρης κύκλος περιοδικών υδρολογικών διακυμάνσεων (Κουτσογιάννης, 1997, σ. 87). Η εφαρμογή του είναι βασική έτσι ώστε οι ετήσιες τιμές να έχουν μεταξύ τους την μεγαλύτερη δυνατή στατιστική ανεξαρτησία. Στον Ελλαδικό χώρο θεωρείται πως το υδρολογικό έτος ξεκινάει την πρώτη Οκτώβρη και λήγει την 30<sup>η</sup> Σεπτέμβρη του επόμενου έτους. Ο δε συμβολισμός π.χ. 1987-88 αναφέρεται στο υδρολογικό έτος

με αρχή την 1/10/1987 και λήξη την 30/9/1988. Η αποθήκευση της τιμής (timestamp) γίνεται στην πρώτη ημέρα του υδρολογικού έτους (π.χ. 1/10/1987 00:00, βλ. Πίνακα 2.4).

Δίνεται επιπλέον δυνατότητα μέσω της λειτουργίας της συνάθροισης (βλ. ενότητα 4.3), παραγωγής ετησίων τιμών από συνάθροιση μηνιαίων σε συγκεκριμένη εποχή (π.χ. Απρίλιος – Σεπτέμβριος) ή και σε ένα πλήρες έτος με αρχή οποιοδήποτε μήνα (π.χ. Απρίλιος – Μάρτιος), ώστε να καταστεί δυνατή η εξέταση υδρολογικών ετών που υιοθετούνται από διαφορετικές χώρες του πλανήτη. Σε αυτήν την περίπτωση η αποθήκευση μπορεί να γίνει είτε σε χρονοσειρά κανονικού – ημερολογιακού έτους είτε σε χρονοσειρά του τυπικού Ελληνικού υδρολογικού έτους – κάτι το οποίο όμως πρέπει να αποφεύγεται για να μην δημιουργείται σύγχυση.

## 2.2.2 Αυστηρό χρονικό βήμα (Timestep strict) – σταθερή χρονική ολίσθηση

Προκειμένου να διεξαχθούν μία σειρά από επεξεργασίες οι χρονοσειρές πρέπει να αποκτήσουν σταθερό χρονικό βήμα. Η ειδική επεξεργασία για την μετατροπή του ακανόνιστου βήματος σε σταθερό – αυστηρό, με την κατάλληλη άρση των χρονικών μετατοπίσεων, περιγράφεται στην ενότητα 4.2. Οι χρονοσειρές μηνιαίου και ετήσιου χρονικού βήματος έχουν εξορισμού αυστηρό χρονικό βήμα, για τα υπόλοιπα χρονικά βήματα όμως, υπάρχει ειδική επισήμανση η οποία αποθηκεύεται μαζί με τα δεδομένα της χρονοσειράς είτε στην βάση δεδομένων είτε στα αρχεία.

Οι χρονοσειρές δεκάλεπτου, ωριαίου και ημερήσιου – αυστηρού χρονικού βήματος, επιτρέπεται να έχουν κάποια σταθερή χρονική ολίσθηση (time offset), έτσι ώστε στο δεκάλεπτο βήμα να μπορούν να έχουν μορφή 00:05, 00:15, 00:25, ..., (offset 5 λεπτών) στο ωριαίο 00:15, 01:15, 02:15, ..., (offset 15 λεπτών) και στο ημερήσιο 21/2/2003T08:00, 22/2/2003T08:00, ... (offset 480 λεπτών – 8 ωρών). Για τις μηνιαίες και ετήσιες χρονοσειρές το *timestamp* θα είναι πάντα αυτό που ορίζεται στον Πίνακα 2.4, επιπλέον όμως θα αποθηκεύεται σε χωριστή μεταβλητή το *time (date) offset* χρησιμοποιώντας τον αριθμό ημερών της μετατόπισης, κυρίως για να γνωρίζουμε την χρονική μετατόπιση της αρχικής χρονοσειράς που παρήγαγε την μηνιαία ή την αντίστοιχη ετήσια.

## 2.3 Τυποποίηση μεταβλητών

Πίνακας 2.3: Τύποι μεταβλητής των χρονοσειρών

Αγγλική ονομασία	Ελληνική μετάφραση	Χρήση
Unknown	Άγνωστη	Ο τύπος της μεταβλητής δεν έχει καθοριστεί
Instantaneous	Στιγμιαία	Οι τιμές αφορούν στιγμιαίες μετρήσεις, π.χ. την στάθμη ενός ταμειυτήρα, την ταχύτητα του ανέμου, κ.α.
Average	Μέσων τιμών	Οι τιμές αφορούν μέσους όρους χρονικών διαστημάτων, π.χ. η μέση θερμοκρασία 24h από ωριαίες μετρήσεις θερμοκρασίας
Cumulative	Αθροιστική	Οι τιμές είναι το άθροισμα των επιμέρους τιμών χρονικών διαστημάτων, π.χ. συναθροισμένη βροχόπτωση μήνα που προέρχεται από (ημερήσιες) μετρήσεις βροχόμετρων
Maximum	Μεγίστων	Μέγιστες τιμές χρονικών διαστημάτων, π.χ. η μέγιστη ωριαία τιμή της θερμοκρασίας για κάποιο εικοσιτετράωρο
Minimum	Ελαχίστων	Ελάχιστες τιμές χρονικών διαστημάτων, π.χ. η ελάχιστη ωριαία τιμή της θερμοκρασίας για κάποιο εικοσιτετράωρο

Στον Πίνακα 2.3 παρουσιάζονται οι τύποι της μεταβλητής για τις χρονοσειρές επεξεργασμένων δεδομένων που μπορούν να παραχθούν με τις λειτουργίες της συνάθροισης (βλ. Ενότητα 4.3).

Επιπλέον, οι αθροιστικές μεταβλητές μπορούν να αφορούν πρωτογενή δεδομένα με το αντίστοιχο φυσικό νόημα (όπως η βροχόπτωση, η διάρκεια ηλιοφάνειας, κ.α.).

Στον Πίνακα 2.4 παρουσιάζονται οι χρονικές στιγμές αποθήκευσης (timestamps) για τα δεδομένα των χρονοσειρών. Η μεθοδολογία αποθήκευσης τροποποιείται ανάλογα με τον τύπο μεταβλητής, αν πρόκειται δηλαδή για στιγμιαία μεταβλητή ή μεταβλητή που αναφέρεται σε χρονικό διάστημα.

Πίνακας 2.4: Χρονικές στιγμές (timestamps) – χρονικά διαστήματα που αναφέρονται οι μεταβλητές

Τύπος μεταβλητής	Χρονικό βήμα	Χρονικό διάστημα	Παράδειγμα
Στιγμιαία (αναφέρεται σε συγκεκριμένη χρονική στιγμή)	Όλα τα χρονικά βήματα	Η ακριβής χρονική στιγμή μέτρησης	Η μέτρηση στάθμης την 2005-09-06T00:10 είναι η στάθμη στις 00:10, την 6 <sup>η</sup> Σεπτεμβρίου του 2005
Όλες όσες αναφέρονται σε χρονικά διαστήματα (μέσων τιμών, αθροιστικές, μεγίστων και ελαχίστων)	Δεκάλεπτο, ωριαίο και ημερήσιο	Το χρονικό διάστημα ισούται με την υποδιαίρεση του χρόνου (χρονικό βήμα), αποτυπώνεται όμως στο τέλος του χρονικού διαστήματος	Η ημερήσια βροχόπτωση που μετρήθηκε μεταξύ 08:01 της 17 <sup>ης</sup> Νοεμβρίου του 1973 και 08:00 της 18 <sup>ης</sup> , θα αποτυπώνεται ως: 1973-11-18T08:00
	Μηνιαίο	Ο συγκεκριμένος μήνας, η τιμή αποτυπώνεται στην αρχή του μήνα	Η μηνιαία βροχόπτωση του Οκτωβρίου του 1981 απεικονίζεται ως 1981/10 ή και 1981-10-01T00:00
	Ετήσιο	Το συγκεκριμένο έτος. Η τιμή αποτυπώνεται στην πρώτη Ιανουαρίου	Η ετήσια βροχόπτωση του 1968 απεικονίζεται ως 1968 ή και 1968-01-01T00:00
	Ετήσιο (Τυπικό Ελληνικό Υδρολογικό έτος)	Το συγκεκριμένο υδρολογικό έτος. Η τιμή αποτυπώνεται στην πρώτη Οκτωβρίου	Η βροχόπτωση του υδρολογικού έτους 1989-90 απεικονίζεται ως 1989-90 ή και 1989-10-01T00:00

Η μεθοδολογία που υιοθετήθηκε για την χρονική στιγμή αποθήκευσης στις μεταβλητές που αναφέρονται σε χρονικό διάστημα μπορεί να προκαλέσει σύγχυση στον χρήστη, ειδικά στις χρονοσειρές ημερησίου χρονικού βήματος, εντούτοις αναπαριστά καλύτερα την φυσική διεργασία της συνάθροισης (χρονικής ολοκλήρωσης).

Η αναπαράσταση της χρονικής στιγμής γίνεται ανεστραμμένα, δηλαδή ξεκινάμε από το έτος και καταλήγουμε τελικά στα πρώτα λεπτά της ώρας. Έτσι για παράδειγμα ή δευτέρα πρωινή της έκτης Σεπτεμβρίου του 2005 κωδικοποιείται ως: 2005-09-06T02:00. Τα σύμβολα διαχωρισμού (-, T, :) μπορούν να τροποποιούνται ή και να παραλείπονται. Η ανεστραμμένη απεικόνιση βοηθά στην ορθότερη χρονολογική ταξινόμηση.

### 2.3.1 Κενές (null) εγγραφές – ελλείψεις

Οι κενές εγγραφές περιγράφουν δεδομένα για τα οποία αδυνατούμε να εκφράσουμε την αριθμητική τους τιμή λόγω ανύπαρκτης μέτρησης, χονδροειδούς σφάλματος ή ατελούς επεξεργασίας. Οι εγγραφές αυτές φέρουν κανονικά κάποιο *timestamp* αλλά και ειδική επισήμανση NULL (κενό), ενώ δεν γίνεται αποθήκευση αριθμητικής τιμής. Κατά την οπτικοποίηση των εγγραφών, γίνεται απεικόνιση της χρονικής στιγμής ενώ στην θέση της τιμής δεν απεικονίζεται ένα κενό κελί.

## 2.4 Σημαιοθέτηση

Πίνακας 2.5: Σημαίες των δεδομένων των χρονοσειρών.

Σημαία (αγγλικά)	Σύντομη μετάφραση	Χρήση σημαίας
RANGE	Εκτός ορίων	Η τιμή είναι εκτός ορίων
SPATIAL	Ασυνεπής χωρικά	Η τιμή είναι ασυνεπής στο χώρο
TEMPORAL	Ασυνεπής χρονικά	Η τιμή είναι χρονικά ασυνεπής
INTERNAL	Ασυνεπής εσωτερικά	Η τιμή είναι εσωτερικά ασυνεπής
ESTIMATED	Εκτιμημένη	Η τιμή είναι κατ' εκτίμηση
SUSPECT	Ύποπτη	Η τιμή είναι ύποπτη
SNOW	Χιόνι	Χιόνι κατά τη στιγμή της μέτρησης
ICE	Πάγος	Πάγος κατά τη στιγμή της μέτρησης
FROST	Παγετός	Παγετός κατά τη στιγμή της μέτρησης
DIVE	Πλημμυρισμένος	Σταθμός πλημμυρισμένος κατά τη στιγμή της μέτρησης
SPILL	Υπερχείλιση	Σταθμός σε κατάσταση υπερχείλισης κατά τη στιγμή της μέτρησης
PUMP	Άντληση	Τιμή υπό άντληση (για γεωτρήσεις)
LOGOVERRUN, LOGNOISY, LOGOUTSIDE, LOGRANGE	-	Σημαίες του logger της Delta-T
HOMOGEN	Από ομογενοποίηση	Η τιμή προέρχεται από ομογενοποίηση
INFILLING	Από συμπλήρωση	Η τιμή προέρχεται από συμπλήρωση
PENMAN	Penman	Η τιμή έχει υπολογιστεί με τη μέθοδο Penman
DATEINSERT	Εισηγμένη ημερομηνία	Η ημερομηνία δεν υπήρχε στην πρωτογενή χρονοσειρά και δημιουργήθηκε κατά τη μετατροπή σε σταθερό χρονικό βήμα
INCONSISTENT	Ασυνεπής	Η τιμή είναι ασυνεπής
AUTO	Αυτόματα	Η τιμή έχει παραχθεί αυτόματα
MISSING	Ελλείψεις	Η τιμή προέρχεται από χρονοσειρά μικρότερου χρονικού βήματος, και μερικές τιμές έλλειπαν, οπότε η εξαγωγή της τιμής (με συνάθροιση, μέσο όρο, κλπ.) έχει γίνει από τις υπόλοιπες

Οι τιμές των χρονοσειρών ή και οι κενές εγγραφές μπορούν να φέρουν κάποια ιδιαίτερα χαρακτηριστικά τα οποία απεικονίζονται με δυαδικές (on – off) σημαίες. Οι σημαίες που χρησιμοποιούνται από το σύστημα του «Υδρογνώμων» απεικονίζονται στον Πίνακα 2.5. Κάθε εγγραφή μπορεί να είναι σημαιοθετημένη με μία ή και περισσότερες σημαίες. Οι σημαίες αφορούν

είτε πρωτογενή δεδομένα (π.χ. DIVE, SPILL, SNOW, κ.α.) είτε επεξεργασμένα (π.χ. TEMPORAL, RANGE, MISSING, DATEINSERT, κ.α.).

Για την περίπτωση των πρωτογενών δεδομένων η σημαιοθέτηση μπορεί να γίνεται αυτόματα μέσω του λογισμικού μεταφοράς των δεδομένων από τον καταγραφέα (logger) στην βάση δεδομένων.

## 2.5 Χρήση γεωγραφικών οντοτήτων

Σύμφωνα με τον σχεδιασμό της βάσης δεδομένων, κάθε χρονοσειρά συσχετίζεται υποχρεωτικά με κάποια οντότητα στον χώρο. Οι γεωγραφικές οντότητες είναι αντικείμενα που έχουν υδρολογικό ενδιαφέρον και μπορούν να περιλάβουν χρονοσειρές για την περιγραφή των ιδιοτήτων τους ή των μεγεθών που μεταβάλλονται στον χρόνο. Τυπικά παραδείγματα αυτής της λογικής είναι οι βροχοπτώσεις που συσχετίζονται με κάποιον μετρητικό σταθμό και οι μετρήσεις στάθμης με κάποιο υδρομετρικό σταθμό ή κάποιον ταμιευτήρα. Μία χρονοσειρά δεν μπορεί να ανήκει σε παραπάνω από μία γεωγραφική οντότητα, το δε σύστημα της βάσης δεδομένων φροντίζει για την συνέπεια των δεδομένων έτσι ώστε να αποτρέπεται η δημιουργία και η ύπαρξη «ορφανών» χρονοσειρών. Μερικές κατηγορίες γεωγραφικών οντοτήτων είναι οι παρακάτω:

- Σημειακές οντότητες όπως οι μετρητικοί σταθμοί, τα φράγματα και τα διυλιστήρια
- Γραμμικές οντότητες είναι οι υδροφορείς (υδραγωγεία, υδατορεύματα ή ποτάμια)
- Επιφανειακές οντότητες όπως οι ταμιευτήρες, οι λίμνες και οι λεκάνες απορροής.

Στο περιβάλλον του «Υδρογνώμων» παρέχονται πεδία εισόδου και ρυθμίσεων για τις ειδικές ιδιότητες των οντοτήτων που δεν μπορούν (ή δεν έχουν νόημα) να παρασταθούν με την μορφή χρονοσειράς καθώς και γραφικά εργαλεία για την εισαγωγή και επισκόπηση κάθε είδους. καμπυλών (τύπου στάθμης – παροχής ή στάθμης – επιφάνειας, όγκου, κ.α.).

Στην ειδική κατηγορία των μετρητικών σταθμών συσχετίζεται ένας αριθμός μετρητικών οργάνων και καταγραφέων που μοιράζονται κοινές χρονοσειρές πρωτογενών δεδομένων μεταξύ τους (μία χρονοσειρά που μετρείται από κάποιο όργανο, καταγράφεται από τον καταγραφέα – ο οποίος μπορεί να διαθέτει περισσότερα του ενός όργανα – ο οποίος ανήκει σε κάποιον μετρητικό σταθμό – ο οποίος μπορεί να έχει περισσότερους καταγραφείς και όργανα).

### 2.5.1 Απεικόνιση γεωγραφικών οντοτήτων – γεωγραφικές επεξεργασίες

Το μεγαλύτερο μέρος των γεωγραφικών επεξεργασιών καθώς και της απεικόνισης των οντοτήτων καλύπτεται από την διασύνδεση του συστήματος επεξεργασίας χρονοσειρών «Υδρογνώμων» με το σύστημα γεωγραφικής πληροφορίας (GIS). Η διασύνδεση αυτή γίνεται με την κατάλληλη αντιστοίχιση των οντοτήτων του πραγματικού κόσμου (αυτών δηλαδή που απεικονίζονται από τον «Υδρογνώμων») και των οντοτήτων του συστήματος γεωγραφικής πληροφορίας.

Ο «Υδρογνώμων» εντούτοις, παρέχει επιπλέον εργαλεία για γεωγραφικές επεξεργασίες και απεικόνιση, τα οποία λειτουργούν συμπληρωματικά ή και αυτόνομα σε σχέση με το σύστημα γεωγραφικής πληροφορίας. Οι δυνατότητες αυτές που παρέχονται αυτόνομα από τον «Υδρογνώμων» είναι οι παρακάτω:

- Απεικόνιση των σημειακών και επιφανειακών οντοτήτων σε επίπεδο χάρτη χρησιμοποιώντας τις καρτεσιανές συντεταγμένες τους
- Απεικόνιση γραμμικών οντοτήτων
- Προσεγγιστική εισαγωγή συντεταγμένων χρησιμοποιώντας τον χάρτη
- Επιλογή οντοτήτων από τον χάρτη



- Επιλεκτική απεικόνιση σύμφωνα με διάφορα κριτήρια (εμφάνιση συγκεκριμένου τύπου οντότητας, συγκεκριμένου ονόματος, κλπ). Άμεση μετάβαση στον χάρτη επιλεγμένης οντότητας
- Μετατροπή συντεταγμένων μεταξύ γεωδαιτικών (ελλειψοειδών) και καρτεσιανών συστημάτων.

Ειδικά για το τελευταίο, ακολουθείται η διαδικασία μεταφοράς συντεταγμένων από  $(\varphi, \lambda)$  σε  $(X, Y)$  και αντίστροφα (γεωδαιτικές  $\leftrightarrow$  καρτεσιανές) σύμφωνα με συστήματα που χρησιμοποιούνται στον Ελληνικό χώρο (ΕΓΣΑ 87, WGS84) και τυπικές διαδικασίες μετατροπής και μεταφοράς των συστημάτων. Δεν θα γίνει περαιτέρω αναφορά στο ζήτημα των γεωδαιτικών μετατροπών, μπορεί κάποιος να ανατρέξει στην σχετική βιβλιογραφία (Βέης, κ.α., 1992, ΟΚΧΕ, 1989).

Η απεικόνιση των σημείων των οντοτήτων στον χάρτη γίνεται με χρήση του αφινικού μετασχηματισμού ο οποίος μετασχηματίζει τις πραγματικές συντεταγμένες  $(X, Y)$  σε συντεταγμένες οθόνης  $(x, y)$ . Ο αφινικός μετασχηματισμός περιλαμβάνει: Μεταφορά αρχής συστήματος συντεταγμένων, κλίμακα σε δύο άξονες, περιστροφή και στρέβλωση. Ο τύπος μετατροπής των συντεταγμένων είναι:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & X & Y & XY \\ 1 & X & Y & XY \end{bmatrix} \cdot \begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \end{bmatrix}^T \quad 2.1$$

Οι συντελεστές  $a_i, b_i$  υπολογίζονται επιλύοντας δύο ανεξάρτητα γραμμικά συστήματα τεσσάρων εξισώσεων τα οποία σχηματίζονται εύκολα εφαρμόζοντας τέσσερα γνωστά σημεία καλιμπραρίσματος στην εξίσωση (2.1). Το σύστημα των εξισώσεων είτε για τα  $x$  είτε για τα  $y$  ( $x|y$ ) γράφεται ως:

$$\begin{bmatrix} 1 & X_1 & Y_1 & X_1Y_1 \\ 1 & X_2 & Y_2 & X_2Y_2 \\ 1 & X_3 & Y_3 & X_3Y_3 \\ 1 & X_4 & Y_4 & X_4Y_4 \end{bmatrix} \cdot \begin{bmatrix} a_1 | b_1 \\ a_2 | b_2 \\ a_3 | b_3 \\ a_4 | b_4 \end{bmatrix} = \begin{bmatrix} x_1 | y_1 \\ x_2 | y_2 \\ x_3 | y_3 \\ x_4 | y_4 \end{bmatrix} \quad 2.2$$

Το αντίστροφο πρόβλημα, της μεταφοράς δηλαδή των συντεταγμένων οθόνης  $(x, y)$  δεν επιλύεται με μία απλή γραμμική αλλά χρησιμοποιώντας την (2.1) ως μη-γραμμική εξίσωση δύο αγνώστων ( $X$  και  $Y$ ). Το σύστημα εξισώσεων διατυπώνεται ως:

$$\left. \begin{matrix} f_1(X, Y) = 0 \\ f_2(X, Y) = 0 \end{matrix} \right\} \quad \text{ή} \quad \left. \begin{matrix} f_1(X, Y) = a_1 + a_2X + a_3Y + a_4XY - x = 0 \\ f_2(X, Y) = b_1 + b_2X + b_3Y + b_4XY - y = 0 \end{matrix} \right\} \quad 2.3$$

Το μη γραμμικό σύστημα λύνεται με διαδοχικές προσεγγίσεις θεωρώντας μία αρχική λύση  $[X_0, Y_0]$  (όπως  $[0, 0]$ ) υπολογίζουμε μία μετατόπιση  $[\delta X, \delta Y]$  της αρχικής λύσης. Στην συνέχεια αυτή η διαδικασία επαναλαμβάνεται μέχρι η νόρμα της μετατόπισης να γίνει π.χ. μικρότερη από ένα χιλιοστόμετρο (1 mm). Η μετατόπιση του κάθε βήματος  $[\delta X_i, \delta Y_i]$  υπολογίζεται μέσω του παρακάτω συστήματος  $2 \times 2$  εξισώσεων:

$$\begin{aligned} \frac{\partial f_1}{\partial X} \delta X_{i+1} + \frac{\partial f_1}{\partial Y} \delta Y_{i+1} &= -f_1(X_i, Y_i) \\ \frac{\partial f_2}{\partial X} \delta X_{i+1} + \frac{\partial f_2}{\partial Y} \delta Y_{i+1} &= -f_2(X_i, Y_i) \end{aligned} \quad \Leftrightarrow \quad 2.4$$

$$\begin{aligned} (a_2 + a_4Y_i)\delta X_{i+1} + (a_3 + a_4X_i)\delta Y_{i+1} &= -f_1(X_i, Y_i) \\ (b_2 + b_4Y_i)\delta X_{i+1} + (b_3 + b_4X_i)\delta Y_{i+1} &= -f_2(X_i, Y_i) \end{aligned} \quad 2.5$$

με αρχική τιμή όπως αναφέρθηκε παραπάνω:

$$[X_0, Y_0] = [0, 0] \quad 2.6$$

και οι τιμές των  $X, Y$  για το  $i+1$  βήμα υπολογίζονται ως:

$$\begin{bmatrix} X_{i+1} \\ Y_{i+1} \end{bmatrix} = \begin{bmatrix} X_i \\ Y_i \end{bmatrix} + \begin{bmatrix} \delta X_{i+1} \\ \delta Y_{i+1} \end{bmatrix} \quad 2.7$$

## 3 Έλεγχοι και αποκατάσταση της συνέπειας των δεδομένων

---

### 3.1 Εισαγωγή

Προτού χρησιμοποιηθούν περαιτέρω τα υδρολογικά δεδομένα – χρονοσειρές, πρέπει να γίνουν οι κατάλληλοι έλεγχοι για την συνέπεια των δεδομένων και στην συνέχεια οι κατάλληλες αναγωγές – διορθώσεις αν διαπιστωθούν προβληματικές μετρήσεις. Στον «Υδρογνώμων» υλοποιούνται κάποιοι βασικοί έλεγχοι που χωρίζονται σε δύο κατηγορίες:

- Έλεγχοι στα πρωτογενή δεδομένα (έλεγχοι ακραίων τιμών και χρονικής συνέπειας)
- Έλεγχος και αποκατάσταση ομοιογένειας επεξεργασμένων χρονοσειρών.

### 3.2 Έλεγχοι ακραίων τιμών και χρονικής συνέπειας

Οι έλεγχοι αυτοί αφορούν τα πρωτογενή δεδομένα όπου είναι δυνατόν να υπάρχουν χονδροειδή ή και συστηματικά σφάλματα κατά την διαδικασία των μετρήσεων. Τέτοια σφάλματα μπορούν να προέρχονται από βλάβες στα όργανα, μη αξιόπιστη λειτουργία καταγραφικής διάταξης, πτώσεις τάσεις, χονδροειδή και συστηματικά σφάλματα ψηφιοποίησης, κ.α.

Τα σφάλματα πρέπει να εντοπιστούν και στην συνέχεια αφού λάβει γνώση ο μελετητής για την ύπαρξή τους, να προβεί σε ενέργειες για την αποκατάστασή τους ή και να αφαιρέσει πλήρως τις προβληματικές τιμές από το δείγμα. Ο εντοπισμός τους γίνεται σύμφωνα με τις παραδοχές για την διακύμανση των τιμών του δείγματος εντός κάποιων φυσικών ορίων (έλεγχος ακραίων τιμών) και για την χρονική μεταβολή των μεγεθών να μην ξεπερνάνε κάποιο κατώφλι (έλεγχος χρονικής συνέπειας).

#### 3.2.1 Έλεγχος ακραίων τιμών

Ο έλεγχος ακραίων τιμών συνοψίζεται στην απλή ανισότητα:

$$l \leq x(t_i) \leq u \quad 3.1$$

Αν δεν ικανοποιείται η ανισότητα 3.1 για κάποια μέτρηση  $x(t_i)$ , τότε αίρεται η σημαία «εκτός ορίων» (“RANGE”) για αυτήν την μέτρηση. Τα όρια  $l$  και  $u$  τίθενται από τον χρήστη και οι τιμές τους εξαρτώνται από τον τύπο της μεταβλητής, έχοντας ως αφετηρία οδηγίες όπως:

- Για αθροιστικές – μη φραγμένες μεταβλητές, όπως το ύψος βροχόπτωσης, το κάτω όριο θα είναι το μηδέν, ενώ είναι δύσκολο να διατυπωθεί μία τιμή για το άνω όριο
- Στις αθροιστικές – φραγμένες μεταβλητές, όπως η διάρκεια ηλιοφάνειας, το κάτω όριο θα είναι το μηδέν ενώ το άνω όριο θα είναι η μέγιστη τιμή διάρκειας ηλιοφάνειας (π.χ. αν η μέτρηση γίνεται κάθε δέκα λεπτά, τότε θα είναι δέκα λεπτά)
- Για συνεχείς μεταβλητές, όπως η θερμοκρασία, θα τίθενται άνω και κάτω όρια σύμφωνα με τα κλιματικά χαρακτηριστικά της περιοχής
- Στις αδιάστατες μεταβλητές, όπως στη σχετική υγρασία, θα χρησιμοποιούνται οι τιμές 0% και 100%
- Σε μεταβλητές που είναι απόλυτα φραγμένες όπως οι σταθμημετρήσεις μπορούν να χρησιμοποιηθούν οι τιμές του άνω και κάτω φράγματος.

## Αυτόματη λειτουργία

Δίνεται η δυνατότητα ώστε το άνω ή το κάτω όριο ή και τα δύο ταυτοχρόνως να τίθενται αυτόματα από το λογισμικό εφαρμόζοντας την κανονική κατανομή. Ο χρήστης μπορεί να θέσει κάποιο επίπεδο εμπιστοσύνης (με προρυθμισμένη τιμή την  $\gamma=99\%$ ). Τότε οι τιμές των κάτω και των άνω ορίων να δίνονται από τις σχέσεις:

$$l_N = N^{-1}[(1-\gamma)/2, \mu, \sigma] \quad 3.2$$

$$u_N = N^{-1}[(1+\gamma)/2, \mu, \sigma] \quad 3.3$$

όπου  $\gamma$  το επίπεδο εμπιστοσύνης,  $N^{-1}$  το ποσοστημόριο της κανονικής κατανομής (βλ. παράγραφο 9.4.1) για μέση τιμή  $\mu$  και τυπική απόκλιση  $\sigma$ . Οι τιμές των  $\mu$  και  $\sigma$  υπολογίζονται από όλο το μήκος της χρονοσειράς σύμφωνα με την μεθοδολογία της παραγράφου 9.2.2. Η μεθοδολογία της αυτόματης λειτουργίας δεν εξυπηρετεί τόσο τον εντοπισμό ασυνεπών τιμών σε πρωτογενή δεδομένα όσο την επισήμανση ακραίων στατιστικά τιμών.

## Οπτική επισήμανση τιμών χρονοσειράς

Οι τιμές μίας χρονοσειράς που έχουν σημειοθετηθεί με επεξεργασία όπως ο εντοπισμός των ακραίων τιμών μπορούν να απεικονιστούν είτε με προβολή των σημαιών σε παρακείμενη στήλη στις τιμές είτε με κατάλληλο «φιλτράρισμα». Επιπλέον δίνεται ειδική δυνατότητα οπτικής επισήμανσης ορισμένων τιμών χρησιμοποιώντας κάποιο έντονο χρώμα ώστε ο χρήστης να έχει καλύτερη εποπτεία. Η οπτική επισήμανση μπορεί να ενεργοποιηθεί για μία από τις παρακάτω συνθήκες:

- Ύπαρξη σημειοθετημένων τιμών
- Ύπαρξη ελλείψεων
- Μέγιστη ή ελάχιστη τιμή
- Ακραίες τιμές οι οποίες δεν ικανοποιούν την ανισότητα  $\mu-3\sigma \leq x \leq \mu+3\sigma$ , όπου  $\mu$  η μέση τιμή και  $\sigma$  η τυπική απόκλιση των τιμών της χρονοσειράς. Με αυτόν τον τρόπο επισημαίνονται ως ακραίες οι τιμές που είναι εκτός αυτών των ορίων για επίπεδο εμπιστοσύνης  $\gamma \approx 99.5\%$ .

### 3.2.2 Έλεγχος χρονικής συνέπειας

Σύμφωνα με τον έλεγχο χρονικής συνέπειας θα σημειοθετείται ως «ύποπτη» (με την σημαία “TEMPORAL”) κάποια τιμή  $x_i$  της χρονοσειράς για την οποία ισχύει:

$$|x_i - x_{i-1}| > \varepsilon \quad 3.4$$

όπου  $\varepsilon$  είναι το κατώφλι της μέγιστης επιτρεπόμενης χρονικής μεταβολής. Η επιλογή της κατάλληλης τιμής για το  $\varepsilon$  είναι ένα πρόβλημα χωρίς μονοσήμαντη λύση και εξαρτάται από παράγοντες όπως ο τύπος της μεταβλητής, το χρονικό βήμα καθώς και η φυσική διακύμανση της μεταβλητής. Συνεπώς θα πρέπει να επιλέγεται με ιδιαίτερη προσοχή, οι δε τιμές που θα προκύψουν ως «ύποπτες» θα πρέπει να εξετάζονται ενδελεχώς από τον χρήστη προτού προβεί σε κάποια ενέργεια αποκατάστασης.

## 3.3 Έλεγχος ομοιογένειας – ομογενοποίηση δεδομένων με την μέθοδο της διπλής αθροιστικής καμπύλης

Ο έλεγχος ομοιογένειας (ή έλεγχος συνέπειας κατά τους Κουτσογιάννη και Ξανθόπουλο, 1999, σ. 107) αποσκοπεί στον εντοπισμό τεχνητών αλλαγών στις συνθήκες μέτρησης, οι οποίες επηρεάζουν το αποτέλεσμα της μέτρησης. Σε περίπτωση που εντοπιστούν τέτοιες αλλαγές, η αποκατάσταση της ομοιογένειας (γνωστή και ως ανόρθωση) αποσκοπεί στην τροποποίηση των μετρήσεων σε τρόπο ώστε να αρθούν τα αποτελέσματα των αλλαγών των συνθηκών. Τέτοιες αλλαγές συσχετίζονται συνήθως με μεταβολές στον τρόπο μέτρησης, στο φυσικό περιβάλλον ή ακόμα και στην τοποθεσία του σταθμού.

Η μεθοδολογία της διπλής αθροιστικής καμπύλης (double mass curve) που παρουσιάζεται και υλοποιείται στα πλαίσια του «Υδρογνώμων» αποτελεί έναν έλεγχο «χωρικής συνέπειας» μεταξύ μετρήσεων βροχόπτωσης γειτονικών σταθμών. Η μέθοδος είναι ημιεμπειρική και συνδυάζει την γραφική απεικόνιση βροχομετρικών δεδομένων με κάποιες στατιστικές υποθέσεις και την μέθοδο των ελαχίστων τετραγώνων. Η εφαρμογή της πρέπει να γίνεται σε σταθμούς με ισχυρή συσχέτιση λόγω κοντινής γεωγραφικής τους θέσης, η δε ανόρθωση είναι καλό να τεκμηριώνεται με στοιχεία από το ημερολόγιο του σταθμού. Η εφαρμογή της μεθόδου γίνεται σε ετήσια δεδομένα βροχόπτωσης χωρίς να αποκλείεται και εφαρμογή σε μηνιαίες χρονοσειρές. Οι συντελεστές που προκύπτουν από την ανόρθωση μπορούν να χρησιμοποιηθούν στην συνέχεια για αναγωγή όλων των χρονοσειρών μικρότερου χρονικού βήματος.

### 3.3.1 Γραφική κατάρτιση της διπλής αθροιστικής καμπύλης

Έστω  $x_0(t_i)$  και  $x_1(t_i)$  οι χρονοσειρές βροχόπτωσης που πρόκειται να ελεγχθούν ως προς την ομοιογένεια τους. Έστω  $x_0$  η χρονοσειρά του σταθμού βάσης, ώστε η ανόρθωση να γίνεται στην χρονοσειρά  $x_1$ . Θεωρούμε πως  $t_i$  είναι η κοινή περίοδο μετρήσεων ( $i=1 \dots n$ ) των χρονοσειρών η οποία καλό είναι να μην περιέχει κενά (τέτοιος έλεγχος δεν γίνεται και είναι στην ευθύνη του χρήστη η αποδοχή των αποτελεσμάτων της μεθοδολογίας στην περίπτωση ύπαρξης κενών). Η διπλή αθροιστική καμπύλη θα συντάσσεται από τις πλέον πρόσφατες κοινές μετρήσεις προς τις παλαιότερες ώστε η ανόρθωση να γίνεται στα παλαιότερα δεδομένα, εντούτοις δίνεται στον χρήστη η δυνατότητα αντίστροφης σύνταξης της καμπύλης. Έτσι οι μεταβλητές του γραφήματος της διπλής αθροιστικής θα είναι:

$$y_0(t_i) = \sum_{j=n}^{n-i+1} x_0(t_j) \quad 3.5$$

$$y_1(t_i) = \sum_{j=n}^{n-i+1} x_1(t_j) \quad 3.6$$

Έστω πως  $y_1'$ , οι ανορθωμένες τιμές της αθροιστικής μεταβλητής, μπορούμε να μεταβούμε εύκολα σε ανορθωμένες τιμές βροχόπτωσης αφαιρώντας δύο διαδοχικές τιμές:

$$x_1(t_i)' = y_1(t_i)' - y_1(t_{i+1})' \quad 3.7$$

όπου το  $i$  θα λαμβάνει τις τιμές μεταξύ  $1 \dots n-1$  και θα ισχύει επιπλέον η σχέση:

$$x_1(t_n)' = y_1(t_n)' \quad 3.8$$

Αν μεταξύ δύο διαδοχικών  $t_i$  υπάρχουν τιμές που έχουν τεθεί εκτός λόγω αλμάτων στην διπλή αθροιστική καμπύλη, θα αναζητείται η πρώτη συνεπής τιμή. Η ανόρθωση μπορεί να γίνει και με την απλή διαδικασία πολλαπλασιασμού των τιμών της χρονοσειράς βροχόπτωσης με τον συντελεστή ανόρθωσης  $\lambda$  που θα παρουσιαστεί στην συνέχεια:

$$x_1(t_i)' = \lambda x_1(t_i) \quad 3.9$$

Αυτή η διαδικασία μπορεί να εφαρμοστεί και σε όλες τις χρονοσειρές μικρότερου χρονικού βήματος, αρκεί να είναι γνωστό το σημείο «θλάσης» της διπλής αθροιστικής καμπύλης.

### 3.3.2 Αποκατάσταση της ομοιογένειας - ανόρθωση

Ο χρήστης μπορεί να υποδείξει πάνω στην γραφική απεικόνιση της διπλής αθροιστικής καμπύλης το σημείο θλάσης, το οποίο θα αντιστοιχεί σε κάποιο  $t_i$ . Στην συνέχεια υπολογίζεται ο συντελεστής ανόρθωσης  $\lambda$  ως εξής:

$$\lambda = \frac{b}{b'} \quad 3.10$$

όπου  $b$  η κλίση της ευθείας που αντιστοιχεί στις συνεπείς μετρήσεις και  $b'$  η κλίση της ευθείας που αντιστοιχεί σε μη συνεπείς μετρήσεις. Αν οι τιμές  $y_1$  της διπλής αθροιστικής πολλαπλασιαστούν με το  $\lambda$  τότε η καμπύλη θα ανορθωθεί. Λόγω της γραμμικότητας κατά την αποσυνάθροιση της καμπύλης σε τιμές βροχόπτωσης, είναι απολύτως ισοδύναμο να πολλαπλασιαστούν κατευθείαν με το  $\lambda$  όλες οι ασυνεπείς τιμές της χρονοσειράς  $x_1$ , ή και των χρονοσειρών μικρότερου χρονικού βήματος από τις οποίες προήλθαν οι ετήσιες (ή μηνιαίες). Η ανόρθωση μπορεί να επαναληφθεί όσες φορές χρειάζεται ώστε να αποκατασταθούν διάφορες ασυνέπειες οι οποίες απεικονίζονται με διαφορετικές θλάσεις. Δίνεται μάλιστα η δυνατότητα αντιστροφής της σειράς συνάθροισης ώστε να υπάρχει η δυνατότητα ανόρθωσης σε περιόδους που είναι κοντά στα άκρα της διπλής αθροιστικής καμπύλης.

Η προσαρμογή ευθειών στην διπλή αθροιστική καμπύλη θα γίνεται με την μέθοδο των ελαχίστων τετραγώνων. Το μέτρο της προσαρμογής των ευθειών στα σημεία της διπλής αθροιστικής καμπύλης θα είναι ο συντελεστής προσδιορισμού:

$$\delta = 1 - \frac{\sum_{i=1}^n W_i^2}{\sum_{i=1}^n y_{1i}^2 - n\bar{y}_1^2} \quad 3.11$$

όπου  $W_i$  είναι οι αποστάσεις των σημείων της διπλής αθροιστικής από κάποια ευθεία με κλίση  $b$  και σταθερό όρο  $a$ :

$$W_i = y_{1i} - (a + b y_{0i}) \quad 3.12$$

Αν περιλαμβάνεται θλάση θα πρέπει να χρησιμοποιούνται τα αντίστοιχα  $a$  και  $b$  προ και μετά της θλάσης. Προκειμένου να σχεδιαστεί η θλάση, το πρώτο τμήμα θα υπολογίζεται με την σχέση της ομογενούς ευθείας (Κουτσογιάννης, 1997, σ. 201), ώστε να είναι εξασφαλισμένη η τομή της ευθείας με την αρχή των αξόνων. Η σχέση που δίνει την κλίση της ευθείας είναι (σταθερός όρος  $a=0$ ):

$$b = \frac{\sum_{i=1}^{n_a} y_{0i} y_{1i}}{\sum_{i=1}^{n_a} y_{0i}^2} \quad 3.13$$

όπου  $i=n_a$  αντιπροσωπεύει το σημείο της θλάσης. Η ευθεία πέρα από το σημείο της θλάσης έχει τον περιορισμό να ξεκινάει από το τέλος της ευθείας του πρώτου τμήματος το οποίο έχει συντεταγμένες:

$$(y_0, y_1) = (y_{0n_a}, b y_{0n_a}) \quad 3.14$$

Συνεπώς η εξίσωση της ευθείας πέρα από το σημείο της θλάσης θα είναι:

$$y_1 = (b-b') y_{0n_a} + b' y_0 \quad 3.15$$

όπου  $b'$  είναι η ζητούμενη κλίση η οποία θα υπολογιστεί ελαχιστοποιώντας το σφάλμα του αθροίσματος των τετραγώνων:

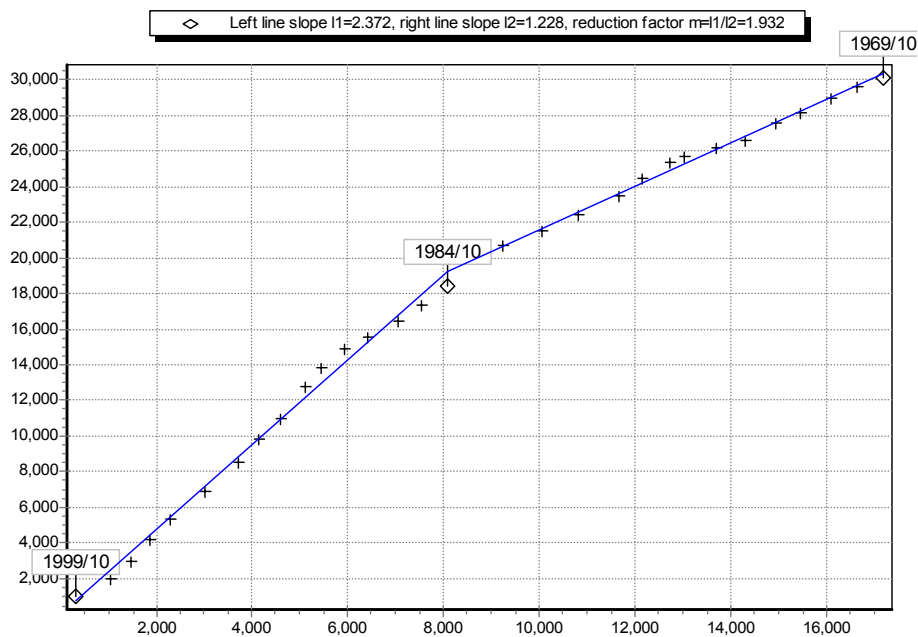
$$q = \sum_{i=n_a}^n W_i^2 = \sum_{i=n_a}^n [y_{1i} - (b-b') y_{0n_a} - b' y_{0i}]^2 \quad 3.16$$

$$\frac{\partial q}{\partial b'} = 0 \Rightarrow b' = \frac{n_b b y_{0n_a}^2 + \sum_{i=n_a}^n y_{0i} y_{1i} - y_{0n_a} \left( \sum_{i=n_a}^n y_{1i} + b \sum_{i=n_a}^n y_{0i} \right)}{n_b y_{0n_a}^2 + \sum_{i=n_a}^n y_{0i}^2 - 2 y_{0n_a} \sum_{i=n_a}^n y_{0i}} \quad 3.17$$

όπου:

$$n_b = n - n_a + 1 \quad 3.18$$

Η θλάση μπορεί να εντοπιστεί εύκολα εξετάζοντας όλα τα σημεία από  $i=3 \dots n-3$  ως πιθανά σημεία θλάσης, λαμβάνοντας τελικά το  $i(=n_a)$  που μεγιστοποιεί τον συντελεστή προσδιορισμού. Ο εξαντλητικός αυτός αλγόριθμος υλοποιείται ως «αυτόματη λειτουργία» ανεύρεσης του σημείου θλάσης.



Σχήμα 3.1: Ανόρθωση του δείγματος βροχόπτωσης του σταθμού «Άνω Μπράλος» χρησιμοποιώντας ως σταθμό βάσης τον σταθμό «Αταλάντη».

Στο Σχήμα 3.1 απεικονίζεται μία ενέργεια αποκατάστασης της ομοιογένειας. Ως μέτρο προσαρμογής χρησιμοποιήθηκε ο συντελεστής προσδιορισμού ο οποίος ήταν  $\delta=0.95$  πριν την αποκατάσταση και  $\delta=0.998$  μετά την αποκατάσταση. Η τιμή του συντελεστή αναγωγής βρέθηκε  $\lambda=1.932$ . Μία τόσο μεγάλη τιμή θα πρέπει να μας βάζει σε σκέψεις κατά πόσο έχει νόημα μία τέτοια αναγωγή και θα πρέπει να εξετάζεται ακόμα και ο αποκλεισμός της χρονοσειράς αν δεν μπορεί να αιτιολογηθεί επαρκώς αυτή η ασυνέπεια ανατρέχοντας π.χ. και στα ημερολόγια του σταθμού.

### 3.3.3 Εντοπισμός εξωκείμενων ετησίων τιμών

Ο χρήστης μπορεί να εντοπίσει εξωκείμενες τιμές με οπτικό έλεγχο στην διπλή αθροιστική καμπύλη. Οι εξωκείμενες τιμές απεικονίζονται ως άλματα της διπλής αθροιστικής καμπύλης. Ο χρήστης μπορεί να βοηθηθεί στον χαρακτηρισμό τιμών βροχόπτωσης ως εξωκείμενων με τον παρακάτω στατιστικό έλεγχο (Barnett, V. and Lewis, 1978, σ.390), ο οποίος προσφέρεται ως «αυτόματος εντοπισμός» τέτοιων σημείων στην διπλή αθροιστική καμπύλη.

Καταρχήν ταξινομείται το δείγμα  $x_1(t_i)$  κατά αύξουσα σειρά τέτοια ώστε:

$$x_1(t_i) > x_1(t_j) \quad \forall i > j \quad 3.19$$

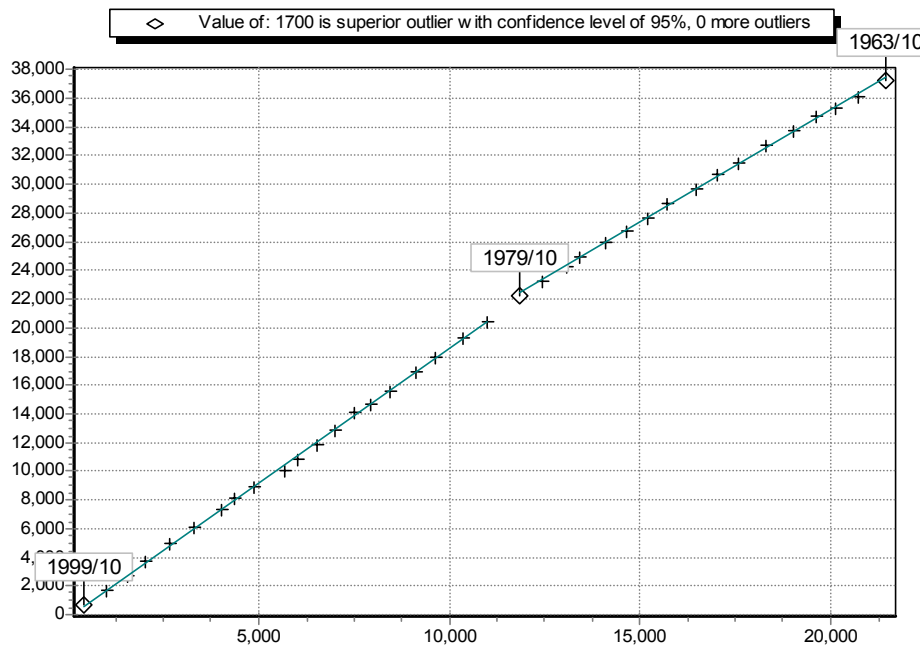
Ξεκινώντας από τις 2 μεγαλύτερες τιμές  $x_1(t_n)$  και  $x_1(t_{n-1})$  υπολογίζεται το μέγεθος  $D_{\text{sup}}$ , καθώς και για το 10% των μεγαλύτερων τιμών του δείγματος:

$$D_{\text{sup}} = \frac{x_1(t_i) - x_1(t_{i-1})}{x_1(t_n) - x_1(t_1)} \quad 3.20$$

Εξωκείμενες θα θεωρούνται οι τιμές  $x_1(t_i)$  για τις οποίες ισχύει:

$$D_{\text{sup}} > 0.182 + \frac{2.3}{n} \quad 3.21$$

Στα πλαίσια του γραφικού περιβάλλοντος επικοινωνίας, προτάσσεται η μεγαλύτερη εξωκείμενη τιμή ώστε ο χρήστης να αποφασίσει αν θα την δεχθεί ή όχι. Στην συνέχεια προβάλλεται ο αριθμός των υπόλοιπων πιθανών εξωκείμενων τιμών (που δεν ικανοποιούν τον παραπάνω έλεγχο).



Σχήμα 3.2: Εντοπισμός εξωκείμενης τιμής μέσω της διπλής αθροιστικής καμπύλης

Στο Σχήμα 3.2 παρουσιάζεται ο εντοπισμός μέσω του στατιστικού ελέγχου κάποιας εξωκείμενης τιμής. Βλέπουμε πως το άλμα της καμπύλης δεν είναι τέτοιο ώστε να γίνεται αντιληπτή μέσω της οπτικής εξέτασης. Οι ευθείες που παρουσιάζονται ενδεικτικά για καλύτερη οπτικοποίηση του άλματος υπολογίζονται μέσω απλής γραμμικής παλινδρόμησης:

$$b = \frac{n \sum_{i=0}^n y_{0i} y_{1i} - \sum_{i=0}^n y_{0i} \sum_{i=0}^n y_{1i}}{n \sum_{i=0}^n y_{0i}^2 - \left( \sum_{i=0}^n y_{0i} \right)^2} \quad 3.22$$

$$a = \bar{y}_1 - b \bar{y}_0 \quad 3.23$$



## 4 Χειρισμοί χρονοσειρών

### 4.1 Εισαγωγή

Για να καταστεί δυνατή η αξιοποίηση των πρωτογενών δεδομένων είναι απαραίτητο να γίνουν οι απαραίτητοι έλεγχοι συνέπειας (βλ. Κεφάλαιο 3) καθώς και κάποιες αρχικές επεξεργασίες που θα εξεταστούν στο παρόν κεφάλαιο. Οι «χειρισμοί χρονοσειρών» όπως καλούμε αυτές τις απλές επεξεργασίες, αφορούν την άρση των χρονικών ολισθήσεων των πρωτογενών δεδομένων με την επιβολή «σταθερού χρονικού βήματος», η εξαγωγή χρονοσειρών μεγαλύτερου χρονικού καθώς και οι γραμμικές πράξεις μεταξύ χρονοσειρών.

Οι επεξεργασίες αυτές μπορούν να εφαρμοστούν ακόμα και σε επεξεργασμένα δεδομένα (π.χ. κανονικοποίηση χρονικού βήματος σε χρονοσειρά από εξαγωγή παροχής εφαρμόζοντας καμπύλες στάθμης – παροχής – βλ. Κεφάλαιο 5) ή ακόμα και σε συνθετικά δεδομένα.

### 4.2 Κανονικοποίηση χρονικού βήματος (άρση χρονικών ολισθήσεων)

Η επεξεργασία της κανονικοποίησης του χρονικού βήματος ή άρσης των χρονικών ολισθήσεων είναι ένα απαραίτητο στάδιο προκειμένου να ακολουθήσουν πιο σύνθετες επεξεργασίες και αναλύσεις. Η μεθοδολογία που παρουσιάζεται προτείνεται από τον Χριστοφίδη (1998) και εφαρμόζεται σε χρονοσειρές με μικρή χρονική ολίσθηση (έως 25% του χρονικού βήματος). Επιπλέον για την εφαρμογή της, πρέπει να προκαθοριστεί στην χρονοσειρά ο τύπος του χρονικού βήματος στον οποίο θα γίνει η κανονικοποίηση.

Για την πραγματοποίηση της μετατροπής, ο χρήστης θα προσδιορίζει τη χρονοσειρά  $X$  την οποία επιθυμεί να μετατρέψει. Η χρονοσειρά  $X$  πρέπει να είναι μεταβλητού χρονικού βήματος  $T$ , όπου  $T = 10$  λεπτά, 1 ώρα, ή 24 ώρες. Οι μηνιαίες και ετήσιες χρονοσειρές έχουν εξορισμού σταθερό χρονικό βήμα οπότε και δεν χρειάζεται να πραγματοποιηθεί αυτή η επεξεργασία. Το σύστημα αγνοεί τυχόν αρχικές και τελικές κενές τιμές στη χρονοσειρά, οπότε η χρονοσειρά  $X$  θα αποτελείται από σημεία  $(\tau_i, x_i)$ ,  $i=1, \dots, m$ , όπου  $x_1 \neq \text{NULL}$  και  $x_m \neq \text{NULL}$ .

Η εξαγόμενη χρονοσειρά  $Y$  είναι σταθερού χρονικού βήματος  $T$  και αποτελείται από σημεία  $(t_j, y_j)$ ,  $i = 1, \dots, n$ , όπου, για κάθε  $j$ ,  $t_j - t_{j-1} = T$ . Το  $t_1$  ορίζεται ως η μεγαλύτερη δυνατή χρονική στιγμή που είναι μικρότερη ή ίση της  $\tau_1$ . Το  $t_n$  ορίζεται ως η μικρότερη δυνατή χρονική στιγμή που είναι μεγαλύτερη ή ίση της  $\tau_m$ . Τίθεται επίσης  $n = (t_n - t_1) / T$ . Είναι δυνατή δε η εισαγωγή σταθερής χρονικής μετατόπισης στις τιμές  $t_i$  ως προς τις ακέραιες υποδιαιρέσεις του χρονικού βήματος (π.χ. 480 λεπτά ώστε να έχουμε ημερήσια χρονοσειρά με τιμές την 08:00 κάθε ημέρας).

Στις στιγμιαίες μεταβλητές, για κάθε  $j$  τέτοιο ώστε να υπάρχει  $i$  τέτοιο ώστε  $t_j = \tau_i$ , τίθεται  $y_j = x_i$ . Επίσης, όλες οι σημαίες του  $x_i$  μεταφέρονται στο  $y_j$ . Στις μεταβλητές διαστήματος γίνεται το ίδιο με τον επιπλέον περιορισμό  $t_{j-1} = \tau_{i-1}$ , εκτός αν  $j = i = 1$ , οπότε δεν υπάρχει επιπλέον περιορισμός.

Στις στιγμιαίες μεταβλητές, για κάθε  $j$  τέτοιο ώστε  $\tau_i < t_j < \tau_{i+1}$  και  $x_i \neq \text{NULL}$  και  $x_{i+1} \neq \text{NULL}$  και  $\min(\tau_{i+1} - t_j, t_j - \tau_i) < 0.25T$  και  $\max(\tau_{i+1} - t_j, t_j - \tau_i) < 1.25T$ , το  $y_j$  υπολογίζεται από τον τύπο

$$y_j = x_i \frac{\tau_{i+1} - t_j}{\tau_{i+1} - \tau_i} + x_{i+1} \frac{t_j - \tau_i}{\tau_{i+1} - \tau_i} \quad 4.1$$

Στις αθροιστικές μεταβλητές, για κάθε  $j$  τέτοιο ώστε  $\tau_i \leq t_j \leq \tau_{i+k}$ , όπου  $k \geq 1$ , και  $\tau_{i+1} \neq \text{NULL}$ ,  $\tau_{i+2} \neq \text{NULL}$ , ...,  $\tau_{i+k} \neq \text{NULL}$ , και  $\min(\tau_{i+k} - t_j, t_j - \tau_{i+k-1}) < 0.25T$  και  $\max(\tau_{i+k} - t_j, t_j - \tau_{i+k-1}) < 1.25T$ , το  $y_j$  υπολογίζεται από την σχέση:

$$y_j = x_{i+1} \frac{\tau_{i+1} - t_{j-1}}{\tau_{i+1} - \tau_i} + x_{i+2} + \dots + x_{i+k-1} + x_{i+k} \frac{t_j - \tau_{i+k-1}}{\tau_{i+k} - \tau_{i+k-1}} \quad 4.2$$

Σε αμφοτέρως τις περιπτώσεις, μεταφέρονται στην τιμή  $y_j$  όλες οι σημαίες της πλησιέστερης τιμής της χρονοσειράς  $X$  (εκείνης, δηλαδή, που απέχει λιγότερο από  $0.25T$ ).

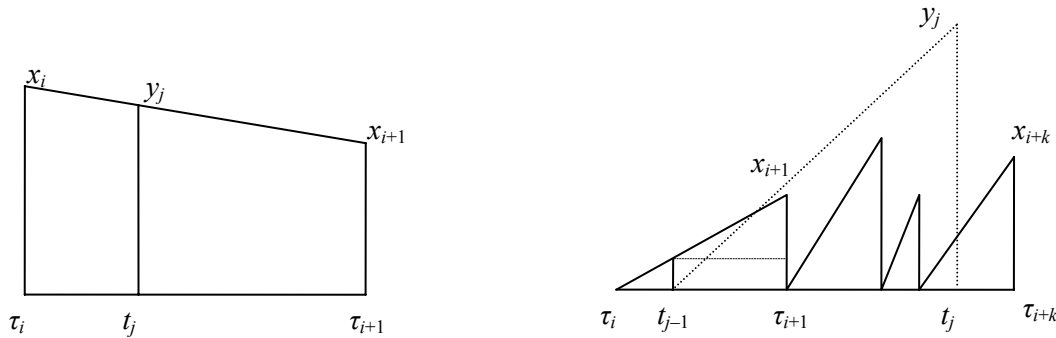
Αν  $0 < \tau_1 - t_1 < 0.25T$  και  $\tau_2 - \tau_1 < 1.25T$ , τότε, για στιγμιαίες μεταβλητές, το  $y_1$  υπολογίζεται από τον ίδιο τύπο όπως προηγουμένως, θέτοντας  $i = 1$  (δηλαδή με εξωτερική παρεμβολή στις δύο ακραίες τιμές). Για αθροιστικές μεταβλητές, αν  $0 < \tau_1 - t_1 < 0.25T$ , τότε το  $y_1$  υπολογίζεται από τον τύπο

$$y_1 = x_1 \left( 1 - \frac{\tau_1 - t_1}{T} \right) \quad 4.3$$

Αν  $0 < t_n - \tau_m < 0.25T$  και  $\tau_m - \tau_{m-1} < 1.25T$ , τότε, για στιγμιαίες μεταβλητές, το  $y_n$  υπολογίζεται από τον ίδιο τύπο όπως προηγουμένως, θέτοντας  $i = m-1$  (δηλαδή με εξωτερική παρεμβολή στις δύο ακραίες τιμές). Αντίστοιχος υπολογισμός για τις αθροιστικές μεταβλητές δεν γίνεται.

Για όσα  $j$  δεν έχει τεθεί τιμή σύμφωνα με τα παραπάνω, τίθεται  $y_j = \text{NULL}$ .

Για να μετατρέψουμε μια χρονοσειρά  $X$  μεταβλητού βήματος σε χρονοσειρά  $Y$  σταθερού βήματος, θεωρούμε ότι οι χρονοσειρές είναι σύνολα ζευγών τιμών  $(\tau_i, x_i)$  και  $(t_j, y_j)$ . Το  $t_{j+1} - t_j = T$  είναι σταθερό για κάθε  $j$  (το βήμα της χρονοσειράς). Το βήμα μιας χρονοσειράς σταθερού βήματος ισούται με το βήμα της χρονοσειράς μεταβλητού βήματος από την οποία προήλθε.



Σχήμα 4.1: Τεχνική άρσης χρονικών μετατοπίσεων

Για τις γραμμικές μεταβλητές (Σχήμα 4.1 αριστερά), αν  $t_j$  είναι χρονική στιγμή μεταξύ των  $\tau_i$  και  $\tau_{i+1}$ , η τιμή  $y_j$  υπολογίζεται ως εξής:

$$y_j = x_i \frac{\tau_{i+1} - t_j}{\tau_{i+1} - \tau_i} + x_{i+1} \frac{t_j - \tau_i}{\tau_{i+1} - \tau_i} \quad 4.4$$

Για τις αθροιστικές μεταβλητές (Σχήμα 4.1 δεξιά), ο υπολογισμός γίνεται ως εξής:

$$y_j = x_{i+1} \frac{\tau_{i+1} - t_{j-1}}{\tau_{i+1} - \tau_i} + x_{i+2} + \dots + x_{i+k-1} + x_{i+k} \frac{t_j - \tau_{i+k-1}}{\tau_{i+k} - \tau_{i+k-1}} \quad 4.5$$

Στις περισσότερες περιπτώσεις, το  $t_j$  συμπίπτει με κάποιο  $\tau_i$  και το  $t_{j-1}$  με το  $\tau_{i-1}$ , οπότε το  $y_j$  τίθεται ίσο με το αντίστοιχο  $x_i$ .

Παρόλο που οι παραπάνω κανόνες φαίνονται απλοί, υπάρχουν λεπτομέρειες που χρειάζονται μεγάλη προσοχή. Οι χρονοσειρές μεταβλητού βήματος έχουν ελλείπουσες τιμές, που σημειώνονται με  $x_i = \text{NULL}$ . Αν κάποια μέρα ο παρατηρητής δεν έχει σημειώσει τιμή στο έντυπο, ή αν η τιμή δεν μπορεί

να διαβαστεί ή είναι λανθασμένη, το πρόγραμμα πληκτρολόγησης θα σημειώσει στην αντίστοιχη θέση  $x_i = \text{NULL}$ . Αν μια ταινία ή ένα τμήμα της δεν μπορεί να διαβαστεί, το πρόγραμμα ψηφιοποίησης θα σημειώσει με  $\text{NULL}$  όλες τις ελαττωματικές τιμές. Όταν ο χρόνος  $t_j$  συμπίπτει με κάποιο χρόνο  $t_i$  τέτοιον ώστε  $x_i = \text{NULL}$ , τότε τίθεται  $y_j = \text{NULL}$ . Όταν όμως οι χρόνοι δεν συμπίπτουν, δεν μπορεί να γίνει αυτό, ούτε μπορούν να εφαρμοστούν οι παραπάνω τύποι. Ακόμα, οι παραπάνω τύποι δεν εφαρμόζονται πάντα· όταν η χρονική απόσταση μεταξύ των εγγραφών  $i$  και  $i+1$  είναι μεγάλη, δεν πρέπει να γίνει παρεμβολή μεταξύ τους, αλλά να τεθούν ελλείπουσες τιμές. Δεν μπορούμε να είμαστε σίγουροι ότι πάντα η χρονοσειρά που λαμβάνεται από το πρόγραμμα πληκτρολόγησης ή ψηφιοποίησης είναι συνεπής στις ελλείπουσες τιμές· μπορεί ο χρήστης, φτάνοντας σε κατεστραμμένη ταινία, να μην δώσει τις κατάλληλες εντολές στο πρόγραμμα, ώστε να σημειωθούν οι αντίστοιχες χρονικές θέσεις με την τιμή  $\text{NULL}$ , οπότε σ' εκείνο το σημείο θα υπάρχει απλώς μεγάλη χρονική απόσταση μεταξύ δύο διαδοχικών εγγραφών. Τέλος, η σημειοθέτηση μιας τιμής  $i$  της χρονοσειράς μεταβλητού βήματος δεν μπορεί εύκολα να μεταφερθεί στη χρονοσειρά σταθερού βήματος αν δεν συμπίπτει χρονικά με κάποια τιμή της.

Για όλα τα σημεία στα οποία υπάρχει και η παραμικρή αμφιβολία, πρέπει να αποφασίζει ο χρήστης, αλλά η εφαρμογή πρέπει να προτείνει τις πιθανότερες λύσεις.

Για την περίπτωση των διανυσματικών μεταβλητών  $a_i$ , θα παράγονται δύο νέες χρονοσειρές  $x_i, y_i$  ως εξής:

$$x_i = \cos(a_i), y_i = \sin(a_i) \quad 4.6$$

στις οποίες θα εφαρμόζεται η κανονικοποίηση του χρονικού βήματος με την μεθοδολογία των στιγμιαίων μεταβλητών. Τελικά θα λαμβάνεται η χρονοσειρά:

$$a_i = \tan^{-1}(y_i/x_i) \{ + 180^\circ \} \{ + 360^\circ \} \quad 4.7$$

όπου ο όρος των  $180^\circ$  θα προστίθεται όταν  $x_i < 0$  και ο όρος των  $360^\circ$  αν  $a_i < 0$ .

Τέλος παρέχεται στον χρήστη η τεχνική της απευθείας μετατόπισης της χρονικής τιμής (change time) κατά την οποία δεν θα γίνεται κάποια γραμμική παρεμβολή αλλά μετατόπιση των εγγραφών της ακανόνιστης χρονοσειράς ώστε να συμπίπτουν με την ακέραια υποδιαίρεση του χρονικού βήματος (συν τυχόν σταθερές χρονικές μετατοπίσεις).

#### 4.2.1 Εναλλακτική μεθοδολογία

Στον «Υδρογνώμων» παρέχεται μία επιπλέον εναλλακτική μεθοδολογία για την κανονικοποίηση του χρονικού βήματος. Η μεθοδολογία αυτή δίνει λύση και σε εξαιρετικές περιπτώσεις χρονικής μετατόπισης (η οποία μπορεί να φτάνει και στο 50% του χρονικού βήματος), πρέπει όμως να εφαρμόζεται κατόπιν κριτικής σκέψης και προσεκτικής εξέτασης των δεδομένων από τον χρήστη. Για λόγους διάκρισης από την παραπάνω μεθοδολογία, στο περιβάλλον του «Υδρογνώμων» θα ονομάζεται «μετατροπή από ακανόνιστο βήμα σε αυστηρό» (“irregular to strict”).

Σύμφωνα με αυτήν την μεθοδολογία η χρονοσειρά που εξετάζεται μπορεί να είναι δεκάλεπτου, ωριαίου, ημερήσιου ή ακανόνιστου χρονικού βήματος. Στην συνέχεια ο χρήστης καθορίζει το χρονικό βήμα της τελικής χρονοσειράς· έτσι είναι θεωρητικά δυνατόν π.χ. να παραχθεί μία χρονοσειρά ημερησίου βήματος από ακανόνιστη χρονοσειρά δεκάλεπτου βήματος, κάτι το οποίο υποκαθιστά και την εφαρμογή του αλγορίθμου της συνάθροισης. Τέτοιες εφαρμογές όμως θα πρέπει να εξετάζονται ενδελεχώς, ειδικά όταν έχουμε να κάνουμε με χρονοσειρές αθροιστικής μεταβλητής.

Στην συνέχεια καταρτίζεται ένα γράφημα με τις τιμές της ακανόνιστης χρονοσειράς συναρτήσει του χρόνου. Αν η μεταβλητή είναι αθροιστική τότε εφαρμόζεται αθροιστικό διάγραμμα ( $t_i, \Sigma x_i$ ). Οι τελικές, κανονικοποιημένες τιμές, λαμβάνονται με γραμμική παρεμβολή στα σημεία  $t_j$  της νέας

χρονοσειράς, εφόσον υπάρχει τέτοιο  $t_i$  που να διαφέρει από το  $t_j$  λιγότερο από το 50% του χρονικού βήματος. Αν δεν υπάρχει τιμή να ικανοποιεί αυτήν την συνθήκη, τότε η νέα τιμή τίθεται NULL (κενή). Ο χρήστης μπορεί να καθορίσει κάποια χρονική ολίσθηση ώστε οι νέες τιμές να έχουν μία σταθερή μετατόπιση ως προς την ακέραια υποδιαίρεση του χρονικού διαστήματος. Για παράδειγμα μπορεί να τεθεί μία σταθερή μετατόπιση 480 λεπτών στην παραγόμενη ημερήσια χρονοσειρά, ώστε οι τιμές να έχουν ως χρονικές στιγμές μέτρησης την 08:00 κάθε ημέρας.

### 4.3 Συνάθροιση – εξαγωγή χρονοσειρών μεγαλύτερου χρονικού βήματος (χρονική ολοκλήρωση)

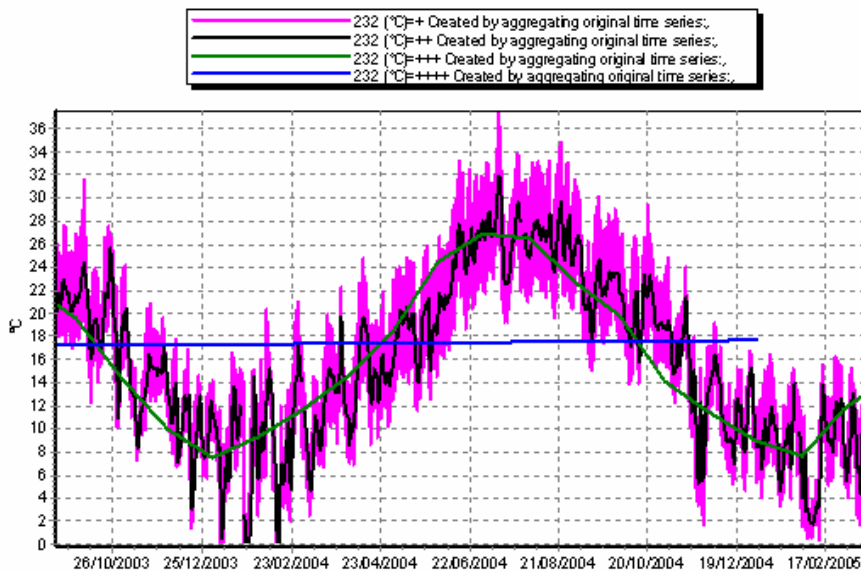
Μία από τις πλέον συνηθισμένες επεξεργασίες των χρονοσειρών είναι η συνάθροιση τους σε χρονοσειρές μεγαλύτερου χρονικού βήματος. Έτσι κατά την ενέργεια της συνάθροισης παράγεται η χρονοσειρά με το αμέσως μεγαλύτερο χρονικό βήμα: δεκάλεπτο → ωριαίο → ημερήσιο → μηνιαίο → ετήσιο. Η υπολογιστική διαδικασία είναι πολύ απλή και συνοψίζεται στον Πίνακα 4.1. Το  $d$  αντιπροσωπεύει το χρονικό διάστημα για το οποίο συναθροίζονται οι τιμές της αρχικής χρονοσειράς. Η επιλογή της χρονικής στιγμής που θα αντιπροσωπεύει το χρονικό διάστημα  $d$  στην νέα χρονοσειρά σύμφωνα με τον Πίνακα 2.4 για όλους τους τύπους συνάθροισης εκτός από την «πρώτη τιμή διαστήματος». Στην τελευταία αυτή περίπτωση παράγεται μία στιγμιαία μεταβλητή και η χρονική στιγμή αναφέρεται στην στιγμιαία τιμή μέτρησης του φαινομένου.

Πίνακας 4.1: Μεθοδολογίες παραγωγής χρονοσειράς με μεγαλύτερο χρονικό βήμα.

Τύπος	Μεθοδολογία
Μέση τιμή	$\sum_d x_i / count_d(x_i)$
Μέση τιμή διανυσματικού μεγέθους	Όπως παραπάνω, η τιμή της γωνίας μετατρέπεται σε δύο επιμέρους χρονοσειρές $x_i = \cos(a_i)$ , $y_i = \sin(a_i)$ όπου και υπολογίζονται οι μέσες τιμές τους. Τελικά λαμβάνεται το τόξο εφαπτομένης των μέσων τιμών των $x_i, y_i$ : $a_i = \tan^{-1}(y_i/x_i) \{+ 180^\circ\} \{+ 360^\circ\}$ .
Άθροισμα (συνάθροιση)	$\sum_d x_i$
Μέγιστη τιμή	$\max_d \{x_i\}$
Ελάχιστη τιμή	$\min_d \{x_i\}$
Πρώτη τιμή διαστήματος	$x_i  _{t_i=d_0}$

Προτού πραγματοποιηθεί η συνάθροιση πρέπει να οριστεί ο μέγιστος επιτρεπόμενος αριθμός ελλείψεων προκειμένου να εξαχθούν τιμές. Αν ο αριθμός των ελλείψεων υπερβεί αυτόν τον αριθμό, τότε το αποτέλεσμα θα είναι μία κενή τιμή (NULL) με την κατάλληλη σημαιοθέτηση (η προρυθμισμένη σημαία για αυτόν τον σκοπό είναι η “MISSING” ωστόσο ο χρήστης μπορεί να καθορίσει και κάποια άλλη σημαία).

Στην ειδική περίπτωση που η συνάθροιση γίνεται από το μηνιαίο χρονικό βήμα προς το ετήσιο, υπάρχει η επιλογή ώστε η συνάθροιση να γίνεται επί του υδρολογικού ή του ημερολογιακού έτους. Επιπλέον δίδεται η δυνατότητα να καταρτιστεί εποχιακή συνάθροιση χρησιμοποιώντας ένα μέρος των μηνών του έτους (π.χ. εποχές – μήνες καλοκαιριού – άνοιξης κλπ ή εξαμηνιαίες τιμές).



Σχήμα 4.2: Συνάθροιση ωριαίας χρονοσειράς μέσω τιμών θερμοκρασίας σε ημερήσια, μηνιαία και ετήσια χρονοσειρά μέσω τιμών. Σταθμός Πολυτεχνειούπολης Ζωγράφου.

#### 4.4 Γραμμικές πράξεις χρονοσειρών

Συχνά χρειάζεται να εξαχθεί χρονοσειρά που να είναι γραμμική σχέση άλλων χρονοσειρών. Ένα παράδειγμα είναι η απλή σχέση βροχής-απορροής, όπου η απορροή  $Q$  δίνεται από τον τύπο

$$Q = C i A \quad 4.8$$

όπου  $C$  ο συντελεστής απορροής,  $i$  η ένταση της βροχόπτωσης και  $A$  η έκταση της λεκάνης απορροής. Τα  $C$  και  $A$  είναι σταθερές. Το  $i$  μπορεί να είναι χρονοσειρά, οπότε και το  $Q$  προκύπτει χρονοσειρά. Άλλα τυπικά παραδείγματα είναι η μετατροπή των μονάδων θερμοκρασίας, η υψομετρική αναγωγή της βροχόπτωσης, η επιφανειακή ολοκλήρωση της βροχόπτωσης, κ.α.

Γενικά, αν  $x_1, x_2, \dots, x_n$  είναι χρονοσειρές, τότε η χρονοσειρά  $y$  που προκύπτει ως γραμμική πράξη αυτών είναι της μορφής

$$y = a_0 + \sum_{i=1}^n a_i x_i \quad 4.9$$

όπου  $a_i$  οι συντελεστές των χρονοσειρών και  $a_0$  ο σταθερός όρος. Οι χρονοσειρές  $x_i$  θα ονομάζονται ανεξάρτητες χρονοσειρές ενώ η χρονοσειρά  $y$  εξαρτημένη χρονοσειρά. Θα πρέπει να ορίζεται τουλάχιστον μία ανεξάρτητη χρονοσειρά, αν δε ορίζονται παραπάνω χρονοσειρές όλες θα πρέπει να έχουν το ίδιο χρονικό βήμα. Η εξαρτημένη χρονοσειρά που προκύπτει θα έχει τιμές για την κοινή περίοδο των ανεξάρτητων χρονοσειρών.

Οι γραμμικές πράξεις υλοποιούνται σε επίπεδο υπορουτίνας για χρήση από επεξεργασίες όπως η συμπλήρωση των ελλείψεων με χρήση πολλαπλής γραμμικής παλινδρόμησης (βλ. παράγραφο 7.2.2) ή την κατάρτιση ισοζυγίων (βλ. ενότητα 8.2). Ωστόσο παρέχεται και ως αυτόνομη λειτουργία όπου ο χρήστης μπορεί να εκτελέσει γραμμικές πράξεις μεταξύ των χρονοσειρών που απεικονίζονται στο «πλέγμα χρονοσειρών».

## 4.5 Απεικόνιση – πλέγματα χρονοσειρών

### 4.5.1 Προβολή ημερήσιων τιμών σε πίνακα για ένα έτος

Η μέθοδος συνίσταται στην προβολή ή εκτύπωση των τιμών σε πίνακα με την εξής μορφή:

Πίνακας 4.2: Απεικόνιση ημερησίων τιμών χρονοσειράς για συγκεκριμένο έτος

	Ιαν	Φεβ	Μαρ	Απρ	Μαϊ	Ιουν	Ιουλ	Αυγ	Σεπ	Οκτ	Νοε	Δεκ
1												
2												
3												
...												
29												
30												
31												
Πλήθος τιμών												
Ελλείπουσες τιμές												
Άθροισμα												
Μέση τιμή												
Τυπική απόκλιση												
Συντελ. διασποράς												
Μέγιστη τιμή												
Ελάχιστη τιμή												
Άνω όριο												
Πολύ ψηλές τιμές												
Κάτω όριο												
Πολύ χαμηλές τιμές												

Ο πίνακας μπορεί να ξεκινά από Ιανουάριο και να καταλήγει στο Δεκέμβριο, όπως ο εικονιζόμενος, ή να χρησιμοποιεί το υδρολογικό έτος, κατ' επιλογή του χρήστη. Η προκαθορισμένη επιλογή λαμβάνεται, ανάλογα με το είδος της χρονοσειράς, από τη βάση δεδομένων. Το άνω όριο ισούται με τη μέση τιμή συν ένα συντελεστή επί την τυπική απόκλιση (βλ. σ. 14, «Οπτική επισήμανση τιμών χρονοσειράς»). Το κάτω όριο ισούται με τη μέση τιμή πλην ένα συντελεστή επί την τυπική απόκλιση. Οι «πολύ ψηλές τιμές» είναι το πλήθος των τιμών που υπερβαίνουν το άνω όριο, και οι «πολύ χαμηλές τιμές» είναι το πλήθος των τιμών που βρίσκονται κάτω από το κάτω όριο. Στο πρώτο μέρος του πίνακα, όπου παρατίθενται οι τιμές, όσες τιμές βρίσκονται εκτός ορίων μπορούν να επισημανθούν οπτικά με κάποιο χρώμα (Highlight). Κάτω από τον πίνακα παρατίθενται και στατιστικά στοιχεία για όλο το έτος, και συγκεκριμένα το πλήθος των τιμών και των ελλειπουσών τιμών, το άθροισμα, η μέση, μέγιστη και ελάχιστη τιμή και η τυπική απόκλιση, και το συνολικό πλήθος των πολύ ψηλών και πολύ χαμηλών τιμών.

Ο χρήστης μπορεί να μεταφέρει τον πίνακα σε άλλες εφαρμογές (π.χ. λογιστικά φύλλα- spreadsheets). Οι ελλείπουσες τιμές είτε θα συμβολίζονται με κενό. Επιπλέον είναι δυνατό και το αντίστροφο (μεταφορά πίνακα από άλλη εφαρμογή στον «Υδρογνώμων»).

### 4.5.2 Προβολή των μηνιαίων τιμών σε πίνακα

Η μέθοδος συνίσταται στην προβολή ή εκτύπωση των μηνιαίων δεδομένων σε πίνακα με την εξής μορφή:

Πίνακας 4.3: Απεικόνιση μηνιαίων τιμών χρονοσειράς

	Οκτ	Νοε	Δεκ	Ιαν	Φεβ	Μαρ	Απρ	Μαϊ	Ιουν	Ιουλ	Αυγ	Σεπ	Έτο ς
1985-86													
1986-87													
...													
1995-96													
Πλήθος τιμών													
Ελλείπουσες τιμές													
Μέση τιμή													
Τυπική απόκλιση													
Συντελ. διασποράς													
Μέγιστη τιμή													
Ελάχιστη τιμή													
Άνω όριο													
Πολύ ψηλές τιμές													
Κάτω όριο													
Πολύ χαμηλές τιμές													

Ο πίνακας μπορεί να ξεκινά από Ιανουάριο και να καταλήγει στο Δεκέμβριο, ή να χρησιμοποιεί το υδρολογικό έτος, όπως ο εικονιζόμενος, κατ' επιλογή του χρήστη. Η προκαθορισμένη επιλογή λαμβάνεται, ανάλογα με το είδος της χρονοσειράς, από τη βάση δεδομένων. Η δεξιά στήλη, «Έτος», δίνει το ετήσιο άθροισμα ή την ετήσια μέση τιμή, ανάλογα με το αν η μεταβλητή είναι αθροιστική ή στιγμιαία. Το άνω όριο ισούται με τη μέση τιμή συν ένα συντελεστή επί την τυπική απόκλιση (βλ. σ. 14, «Οπτική επισήμανση τιμών χρονοσειράς»). Το κάτω όριο ισούται με τη μέση τιμή πλην ένα συντελεστή επί την τυπική απόκλιση. Οι «πολύ ψηλές τιμές» είναι το πλήθος των τιμών που υπερβαίνουν το άνω όριο, και οι «πολύ χαμηλές τιμές» είναι το πλήθος των τιμών που βρίσκονται κάτω από το κάτω όριο. Στο πρώτο μέρος του πίνακα, όπου παρατίθενται οι τιμές, όσες τιμές βρίσκονται εκτός ορίων μπορούν να επισημανθούν οπτικά με κάποιο χρώμα (Highlight). Κάτω από τον πίνακα παρατίθενται και στατιστικά υπερετήσια στοιχεία, και συγκεκριμένα το πλήθος των μηνιαίων τιμών και των ελλειπουσών μηνιαίων τιμών, η μέση, μέγιστη και ελάχιστη μηνιαία τιμή, και το συνολικό πλήθος των πολύ ψηλών και πολύ χαμηλών τιμών.

Για τη μεταφορά του πίνακα σε άλλες εφαρμογές ισχύει ότι και για τον πίνακα ημερήσιων τιμών.

#### 4.5.3 Προβολή δύο ή περισσότερων χρονοσειρών σε συγκριτικό πίνακα (πλέγμα χρονοσειρών–timeseries grid)

Η μέθοδος συνίσταται στην προβολή δύο έως είκοσι χρονοσειρών (ή υποσυνόλων τους) σε συγκριτικό πίνακα.. Ο πίνακας έχει την εξής μορφή:

Ημερομηνία	Χρονοσειρά 1	Χρονοσειρά 2	...

Οι χρονοσειρές μπορούν να έχουν διαφορετική διακριτικότητα (χρονικό βήμα). Το ποιες χρονοσειρές θα αντιπαραβληθούν επαφίεται στην κρίση του χρήστη. Ο χρήστης μπορεί να λάβει την απόφασή του χρησιμοποιώντας βοηθητικές επεξεργασίες (π.χ συσχέτιση). Μπορεί επίσης να συγκρίνει δυο χρονοσειρές που προκύπτουν η μια από την άλλη (π.χ. ομογενοποιημένη με αρχική). Τέλος είναι

δυνατή η προβολή, δίπλα στις εγγραφές των χρονοσειρών, οι σημαίες καθώς και η οπτική επισήμανση με χρώμα (Highlight) των εγγραφών που είναι σημειωμένες.

#### **4.5.4 Δημιουργία διαγραμμάτων**

Με τη βοήθεια διαγραμμάτων μπορούν να εντοπιστούν πολλά λάθη, η ύπαρξη ακραίων τιμών και ο άμεσος εντοπισμός μεγάλων περιόδων με ελλείψεις τιμών. Τα διαγράμματα μπορούν να παράγονται είτε μεμονωμένα για κάθε χρονοσειρά είτε πολλαπλά διαγράμματα που θα αντιπαραβάλλουν τις τιμές διαφορετικών χρονοσειρών. Δίνεται η επιλογή στον χρήστη ώστε το σύστημα να επιτρέπει ή όχι πολλαπλό διάγραμμα χρονοσειρών με διαφορετικές μονάδες μέτρησης. Τέτοιο παράδειγμα γραφήματος απεικονίζεται στο Σχήμα 4.2.



## 5 Καμπύλες στάθμης – παροχής, παρεμβολές

### 5.1 Εισαγωγή

Το θέμα της εξαγωγής τιμών κάποιου μεγέθους εφαρμόζοντας μαθηματικά μοντέλα παρεμβολής σε κάποιο μετρημένο μέγεθος (χρονοσειρά), καλύπτει μία μεγάλη κατηγορία έμμεσων υπολογισμών που έχουν ενδιαφέρον για την υδρολογία. Οι περισσότεροι από αυτούς τους υπολογισμούς αφορά την αξιοποίηση της μέτρησης της στάθμης του νερού ενός υδροφορέα ή υδάτινης μάζας εν γένει. Τέτοιες είναι και οι επεξεργασίες που εξετάζονται στο παρόν κεφάλαιο, συγκεκριμένα:

- Η εξαγωγή του όγκου και της επιφάνειας ενός ταμιευτήρα
- Η εκτίμηση των υπόγειων διαφυγών ενός ταμιευτήρα
- Η εξαγωγή της παροχής ενός υδατορεύματος μέσω σχέσεων στάθμης – παροχής.

Μία υποπερίπτωση της εξαγωγής της παροχής, αρκετά απλή, είναι η εκτίμηση της παροχής που διέρχεται από υπερχειλιστή, η οποία δεν θα τύχει ξεχωριστής ανάλυσης σε αυτό το κεφάλαιο ωστόσο υλοποιείται στο περιβάλλον του «Υδρογνώμων». Ιδιαίτερο βάρος δίνεται στις καμπύλες στάθμης – παροχής του σύνθετου προβλήματος. Θα παρουσιαστεί μία μεθοδολογία για την καλύτερη περιγραφή του φυσικού προβλήματος, μέθοδοι κατάρτισης των καμπύλων, υδραυλικές μέθοδοι για την εκτίμηση των καμπύλων επέκτασης καθώς και οι απαραίτητες διορθώσεις για να αντιμετωπιστούν τα μετρητικά σφάλματα.

### 5.2 Καμπύλες στάθμης – όγκου επιφάνειας ταμιευτήρων

Η εκτίμηση του αποθέματος καθώς και του εμβαδού της ελεύθερης επιφάνειας των ταμιευτήρων μπορεί να πραγματοποιηθεί με την εφαρμογή των καμπύλων στάθμης – όγκου επιφάνειας. Τέτοιες καμπύλες μπορούν να προκύψουν εύκολα από το τοπογραφικό ανάγλυφο της λεκάνης κατάκλισης χρησιμοποιώντας κάποιο Σύστημα Γεωγραφικής Πληροφορίας (GIS). Οι καμπύλες αυτές καταρτίζονται ως τεθλασμένες γραμμές από σημεία  $(h_i, y_i)$ , όπου  $h_i$  η απόλυτη στάθμη της ελεύθερης επιφάνειας και  $y_i$  το μέγεθος που μας ενδιαφέρει (απόθεμα ή επιφάνεια). Το μέγεθος  $y$  μπορεί να προκύψει μεταξύ δύο διαδοχικών σημείων  $(h_i, y_i)$  και  $(h_{i+1}, y_{i+1})$  με λογαριθμική παρεμβολή.

Έστω λοιπόν  $h$  μία στάθμη μεταξύ δύο διαδοχικών σταθμών  $h_i$  και  $h_{i+1}$  της καμπύλης στάθμης – όγκου επιφάνειας, τότε το ζητούμενο μέγεθος υπολογίζεται ως:

$$y = e^{\ln y_i + \frac{\ln y_{i+1} - \ln y_i}{\ln h_{i+1} - \ln h_i} (\ln h - \ln h_i)} \quad 5.1$$

Οι τιμές της στάθμης  $h$  προέρχονται από την αντίστοιχη χρονοσειρά η οποία προκύπτει είτε από κάποιο καταγραφικό όργανο, είτε από κάποιο σταθμήμετρο, είτε είναι μία χρονοσειρά σύνθεσης που έχει προκύψει συνδυάζοντας π.χ. σταθμήμετρο και σταθμηγράφο (βλ. ενότητα 5.4).

### 5.3 Εκτίμηση υπόγειων διαφυγών ταμιευτήρων

Προκειμένου να καταρτιστεί το υδατικό ισοζύγιο ταμιευτήρων είναι απαραίτητη η εκτίμηση του μεγέθους των υπόγειων διαφυγών νερού από τον πυθμένα. Κατά καιρούς έχουν προταθεί διάφορα

μοντέλα για την εκτίμηση των διαφυγών. Τα πιο απλά από αυτά προτείνουν μία γραμμική σχέση του τύπου (Ευστρατιάδης, κ.α., 2000):

$$Q_L = a (H - H_0) + b \quad 5.2$$

όπου  $Q_L$  παροχή,  $a$  ο συντελεστής των υπόγειων διαφυγών,  $H_0$  κάποια στάθμη αναφοράς,  $H$  η απόλυτη στάθμη της ελεύθερης υδατικής επιφάνειας και  $b$  κάποιος σταθερός όρος. Το μοντέλο αυτό μπορεί να ρυθμιστεί είτε με εμπειρικές μεθόδους, είτε με απλά μαθηματικά μοντέλα λαμβάνοντας υπόψη την διαπερατότητα των γεωλογικών σχηματισμών, ή, με επίλυση υδατικού ισοζυγίου εφόσον είναι γνωστή η απορροή στον ταμειυτήρα.

Στον «Υδρογνώμων» εφαρμόζεται ένα μαθηματικό μοντέλο με εξίσωση τρίτου βαθμού, λαμβάνοντας υπόψη την εποχική διαφοροποίηση του φαινομένου. Η μαθηματική έκφραση αυτού του μοντέλου συνοψίζεται στην εξίσωση:

$$Q_L = a_i H^3 + b_i H^2 + c_i H + d_i \quad 5.3$$

όπου  $a_i, b_i, c_i$  οι συντελεστές του πολυωνύμου,  $d_i$  ο σταθερός όρος και  $i=1 \dots 12$  αντιπροσωπεύοντας τους μήνες του έτους. Οι τιμές της στάθμης  $H$  προέρχονται από την αντίστοιχη χρονοσειρά η οποία προκύπτει είτε από κάποιο καταγραφικό όργανο, είτε από κάποιο σταθμήμετρο, είτε είναι μία χρονοσειρά σύνθεσης που έχει προκύψει συνδυάζοντας π.χ. σταθμήμετρο και σταθμηγράφο (βλ. ενότητα 5.4).

## 5.4 Εξαγωγή παροχής από μετρήσεις στάθμης και καμπύλες στάθμης - παροχής

Για την εξαγωγή της παροχής με την χρήση καμπυλών στάθμης παροχής ακολουθείται η μεθοδολογία που περιγράφεται από τον Χριστοφίδη (1998) καθώς και από τους Κουτσογιάννη και Ξανθόπουλο (1999, σ. 328, 330 και 339). Η υλοποίηση του αλγορίθμου υπολογισμού των παροχών στον «Υδρογνώμων» ακολουθεί αυτές τις περιγραφές: παρακάτω παρατίθεται σχεδόν αυτούσια αυτή του Χριστοφίδη (1998) με κατάλληλες προσαρμογές και επισημάνσεις για το σύστημα του «Υδρογνώμων».

Το πρόβλημα της εξαγωγής των τιμών της παροχής από μετρήσεις στάθμης σε κάποιον φυσικό υδροφορέα όπου δεν είναι δυνατή η εγκατάσταση κάποιας διατομής ελέγχου (π.χ. διώρυγα μέτρησης της παροχής) λύνεται με την διατύπωση της κατάλληλης σχέσης στάθμης παροχής της μορφής  $Q = f(h)$ , όπου  $Q$  η παροχή και  $h$  η μέτρηση στάθμης. Η διατύπωση αυτής της σχέσης απαιτεί συνήθως την εγκατάσταση κάποιου υδρομετρικού σταθμού στον οποίο θα γίνονται συστηματικές μετρήσεις στάθμης και παροχής π.χ. με χρήση μιλίσκων. Η σχέση  $Q = f(h)$  είναι γνωστή ως σχέση (ή καμπύλη) *στάθμης - παροχής*. Μία απλή μορφή για την καμπύλη στάθμης παροχής είναι η ακόλουθη (Κουτσογιάννης & Ξανθόπουλος, 1999):

$$Q = C (h - a)^N \quad 5.4$$

όπου  $a$  είναι η στάθμη η οποία αντιστοιχεί στην μηδενική παροχή και  $C, N$  παράμετροι που προσδιορίζονται με προσαρμογή της καμπύλης στις υδρομετρήσεις. Η γενικότερη σχέση για τις καμπύλες στάθμης παροχής, αποτελείται από ένα πλήθος  $n$  σημείων της καμπύλης  $(h_i, q_i), h_{i+1} > h_i \forall i$ , όπου οι παροχές εξάγονται για οποιοδήποτε  $h$  με χρήση της διπλής λογαριθμικής παρεμβολής:

$$Q = f(h) = e^{q_i + \frac{\ln q_{i+1} - \ln q_i}{\ln h_{i+1} - \ln h_i} (\ln h - \ln h_i)} \quad 5.5$$

όπου τα σημεία  $h_i$  και  $h_{i+1}$  επιλέγονται έτσι ώστε:  $h_i \leq h < h_{i+1}$ . Αν δεν υπάρχουν τέτοια σημεία ( $h > h_n$ ) τότε επιχειρείται επέκταση της καμπύλης (extrapolation) χρησιμοποιώντας την Εξίσωση (5.5) και τα

σημεία  $h_{n-1}$ ,  $h_n$ . Αν δε, προβλέπεται ειδική καμπύλη επέκτασης (βλ. στην συνέχεια) γίνεται παρεμβολή μεταξύ του τελευταίου σημείου της καμπύλης στάθμης παροχής και του πρώτου σημείου της καμπύλης επέκτασης χρησιμοποιώντας την Εξίσωση (5.5).

Στις τιμές της στάθμης ( $h$  καθώς και στα σημεία  $h_i$ ) προστίθεται κάποιος σταθερός όρος (offset)  $h_0$ . Αυτός ο όρος δεν εισάγει διαφοροποίηση στα αποτελέσματα, συνήθως επιλέγεται κάποιος αριθμός τέτοιος ώστε τα  $h_i+h_0$  να είναι θετικοί αριθμοί μεγαλύτεροι του μηδενός. Αυτό γίνεται για να αποφευχθούν αρνητικοί αριθμοί προς λογαρίθμηση στην Εξίσωση (5.5). Συχνά είναι χρήσιμη η αντίστροφη συνάρτηση παρεμβολής:

$$h = f^{-1}(Q) = e^{\frac{h_i + \frac{\ln h_{i+1} - \ln h_i}{\ln q_{i+1} - \ln q_i} (\ln Q - \ln q_i)}{\ln q_{i+1} - \ln q_i}} - h_0 \quad 5.6$$

Λόγω μεταβολών στα χαρακτηριστικά του υδατορεύματος με τον χρόνο οι καμπύλες στάθμης παροχής μεταβάλλονται επίσης με τον χρόνο. Αυτό μπορεί να διαπιστωθεί εύκολα όταν απεικονιστούν τα σημεία των υδρομετρήσεων ( $h_i, q_i$ ) σε κάποιο διπλό λογαριθμικό διάγραμμα. Επειδή η μαθηματική διατύπωση αυτής της μεταβολής είναι δύσκολη, επιλέγεται ο χωρισμός του δείγματος των υδρομετρήσεων σε  $m$  περιόδους:  $(h_i^{(j)}, q_i^{(j)})$ ,  $j=1 \dots m$ . Έτσι ο υπολογισμός της παροχής  $Q$  από την στάθμη  $h$  γίνεται αφού επιλεγεί η κατάλληλη καμπύλη που αντιστοιχεί στην περίοδο μέτρησης της  $h$ . Ο διαχωρισμός των υδρομετρήσεων (συνεπώς και των καμπυλών στάθμης – παροχής) σε περιόδους μπορεί να γίνει με κάποιον ημιγραφικό τρόπο, όπως δηλαδή στον «Υδρογνώμων» (βλ. παράγραφο 5.5.1) είτε με πλήρως αυτοματοποιημένη διαδικασία (βλ. Tsakalias and Koutsoyiannis, 1999).

Επειδή συνήθως οι μετρήσεις στάθμης – παροχής δεν καλύπτουν όλη την δυνατή γκάμα, ειδικά σε υψηλές τιμές που παρουσιάζονται στα πλημμυρικά επεισόδια, είμαστε υποχρεωμένοι να διατυπώσουμε ειδικές καμπύλες επέκτασης οι οποίες θα έχουν ισχύ πέρα από τις μετρήσεις στάθμης – παροχής που διαθέτουμε. Οι καμπύλες επέκτασης λαμβάνονται συνήθως σταθερές για όλη την περίοδο των μετρήσεων, καταρτίζονται δε χρησιμοποιώντας εξισώσεις της υδραυλικής (ομοιόμορφης ροής σε ανοικτούς αγωγούς). Για να καταρτίσουμε τις καμπύλες επέκτασης χρησιμοποιούμε μεταξύ άλλων τα τοπογραφικά δεδομένα του σταθμού (διατομή υδατορεύματος) καθώς και τις μεγαλύτερες από τις διαθέσιμες υδρομετρήσεις (βλ. Κουτσογιάννης και Ξανθόπουλος, 1999, σ. 333).

Με δεδομένες τις καμπύλες στάθμης-παροχής, φαίνεται λογικό ότι η εξαγωγή μιας χρονοσειράς παροχής από χρονοσειρά στάθμης θα ήταν απλή διαδικασία παρεμβολής στις καμπύλες. Στην πραγματικότητα όμως πρέπει να ληφθούν υπόψη οι αποκλίσεις των μετρήσεων παροχής από τις καμπύλες στάθμης-παροχής. Το γεγονός ότι οι μετρήσεις παροχής δεν πέφτουν ακριβώς πάνω στις καμπύλες δεν οφείλεται σε ανακρίβεια των μετρήσεων ή σε άλλα αμελητέα γεγονότα, αλλά σε φυσικά αίτια. Η Shaw (1994, σ. 118-119) εξηγεί ότι για την ίδια ακριβώς διατομή μπορεί να έχουμε σε δυο στιγμές ίδια στάθμη αλλά διαφορετική παροχή, γιατί η παροχή στη φάση ανόδου της στάθμης είναι διαφορετική απ' ό,τι στη φάση καθόδου. Η διαφορά αυτή έχει να κάνει με τις ιδιότητες του νερού ως ρευστού μέσου. Ακόμα, οι Linsley et al. (1949 σ. 204-205) εξηγούν επίσης και ότι μπορεί να υπάρχει βαθμιαία μεταβολή των χαρακτηριστικών της κοίτης (π.χ. ανάπτυξη φυτών, ανάπτυξη πάγου, βαθμιαία εναπόθεση φερτών).

Για τέτοιες μικρές μεταβολές δεν είναι κατάλληλη η δημιουργία διαφορετικής περιόδου, και, επομένως, διαφορετικής καμπύλης στάθμης-παροχής, διαδικασία κατάλληλη κυρίως για περιπτώσεις απότομης και μεγάλης αλλαγής, μετά την οποία η κατάσταση παραμένει γενικά σταθερή. Στις μικρές, προσωρινές μεταβολές, είναι καλύτερα να διατηρείται μία καμπύλη στάθμης – παροχής, και να λαμβάνονται υπόψη οι αποκλίσεις των μετρήσεων παροχής από την καμπύλη. Αυτό γίνεται με μέθοδο που είναι γνωστή ως *διόρθωση στάθμης* (stage correction). Μερικοί ερευνητές, όπως η Shaw (1994), την ονομάζουν και *διόρθωση Stout*. Λεπτομέρειες γι' αυτή τη διόρθωση δεν υπάρχουν πολλές στη

βιβλιογραφία, γι' αυτό η ανάπτυξη της μεθόδου που ακολουθεί έχει προκύψει από διερεύνηση του Α. Χριστοφίδη (Χριστοφίδης, 1998).

### 5.4.1 Η διόρθωση της στάθμης

Όπως περιγράφεται στο προηγούμενο παράγραφο, η καμπύλη στάθμης-παροχής είναι συνάρτηση  $f$  που συνδέει τη στάθμη  $h$  με την παροχή  $Q$ :

$$Q = f(h) \quad 5.7$$

Όμως, επειδή θα λάβουμε υπόψη μας τις αποκλίσεις των μετρήσεων παροχής από την καμπύλη στάθμης-παροχής, θα κάνουμε διόρθωση και η παροχή που τελικά θα υπολογίσουμε θα είναι ελαφρώς διαφορετική:

$$Q' = Q + \Delta Q \quad 5.8$$

Στην πράξη, η παροχή διορθώνεται έμμεσα, διορθώνοντας τη στάθμη:

$$Q' = f(h + \Delta h) \quad 5.9$$

Υπολογίζεται δηλαδή με τη μέθοδο που θα αναφέρουμε ο διορθωτικός όρος  $\Delta h$ , που προστίθεται στη στάθμη πριν γίνει η παρεμβολή στην καμπύλη στάθμης-παροχής. Γι' αυτό η μέθοδος ονομάζεται *διόρθωση στάθμης* (stage correction), αλλά ο όρος είναι παραπλανητικός: στην πραγματικότητα δεν θέλουμε να διορθώσουμε τη στάθμη, αλλά την παροχή. Η στάθμη που μετρήθηκε ήταν  $h$ , και όχι  $h + \Delta h$ . Ο όρος  $\Delta h$  είναι τεχνητή διόρθωση την οποία εισάγουμε προκειμένου να διορθώσουμε την ανακρίβεια της καμπύλης στάθμης-παροχής.

Η διόρθωση στάθμης  $\Delta h$  είναι συνάρτηση του χρόνου. Στις χρονικές στιγμές των μετρήσεων παροχής είναι εύκολο να υπολογιστεί. Αν  $hm_i$  και  $qm_i$  είναι η μετρημένη στάθμη και η μετρημένη παροχή της μέτρησης  $i$ , τότε τη χρονική στιγμή της μέτρησης η παροχή που υπολογίζεται χωρίς διόρθωση είναι

$$Q = f(hm_i) \quad 5.10$$

Επειδή οι μετρήσεις παροχής δεν πέφτουν ακριβώς πάνω στην καμπύλη στάθμης-παροχής, εν γένει  $Q \neq qm_i$ . Ο διορθωτικός όρος  $\Delta h$  σ' αυτή την περίπτωση υπολογίζεται έτσι ώστε η καμπύλη στάθμης-παροχής αντί  $Q$  να δώσει αποτέλεσμα  $qm_i$ :

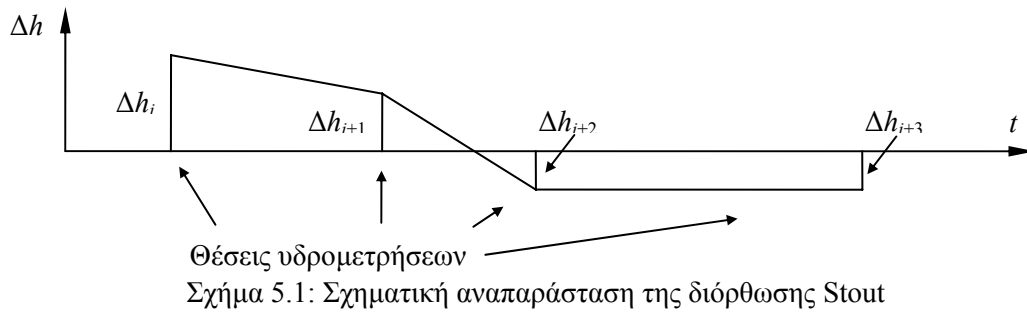
$$\begin{aligned} qm_i &= f(hm_i + \Delta h) \\ \Leftrightarrow hm_i + \Delta h &= f^{-1}(qm_i) \\ \Leftrightarrow \Delta h &= f^{-1}(qm_i) - hm_i \end{aligned} \quad 5.11$$

Με άλλα λόγια, η διορθωμένη στάθμη  $hm_i + \Delta h$  βρίσκεται αν παρεμβληθεί στην καμπύλη στάθμης-παροχής η παροχή  $qm_i$ .

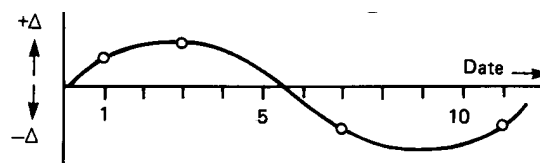
Για να υπολογίσουμε τη διόρθωση στάθμης  $\Delta h$  στις υπόλοιπες χρονικές στιγμές, εκτός δηλαδή των μετρήσεων παροχής, ο απλούστερος τρόπος είναι να θεωρήσουμε ότι η διόρθωση  $\Delta h$  μεταξύ δύο διαδοχικών μετρήσεων παροχής μεταβάλλεται γραμμικά, όπως στο Σχήμα 5.1.

Συνεπώς η μεταβολή της διόρθωσης συναρτήσει του χρόνου μεταξύ δύο διαδοχικών μετρήσεων μπορεί να διατυπωθεί με την εξίσωση:

$$\Delta h(t) = \Delta h(t_i) \frac{t_{i+1} - t}{t_{i+1} - t_i} + \Delta h(t_{i+1}) \frac{t - t_i}{t_{i+1} - t_i} \quad 5.12$$

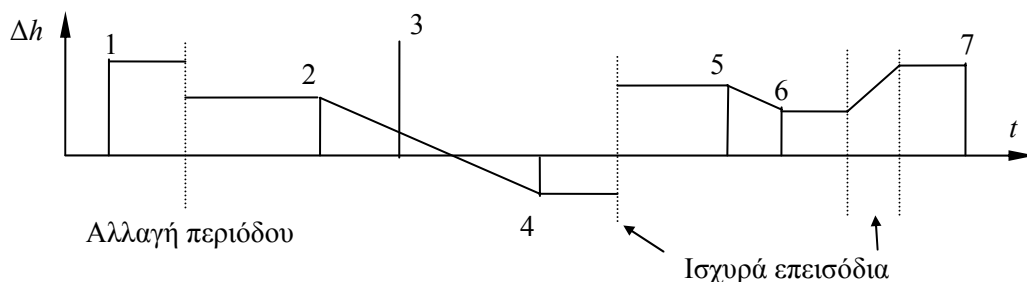


Η Shaw (1994 σ. 119) παραθέτει το Σχήμα 5.2, όπου φαίνεται πως πρέπει να προσαρμοστεί καμπύλη στις αποκλίσεις από τις μετρήσεις παροχής. Αυτό όμως μπορεί να γίνει μόνο σε ειδικές περιπτώσεις, όπου οι μετρήσεις είναι τακτικές και η διατομή του ποταμού έχει τέτοια χαρακτηριστικά ώστε να συνηγορεί υπέρ αυτής της λύσης. Στην Ελλάδα, οι κοίτες των ποταμών σπάνια παρουσιάζουν τέτοια ομαλή συμπεριφορά και κατάλληλη είναι μόνο η απλή και σίγουρη λύση της γραμμικής μεταβολής.



Σχήμα 5.2: Διόρθωση Stout σύμφωνα με την Shaw (1994)

Υπάρχουν περιπτώσεις (Linsley et al. 1949, σ. 205) στις οποίες δεν πρέπει να θεωρηθεί γραμμική η μεταβολή. Τέτοιες περιπτώσεις εικονίζονται στο Σχήμα 5.3. Υπάρχουν 7 μετρήσεις παροχής. Μεταξύ των μετρήσεων 5 και 6 έχει θεωρηθεί απλή γραμμική μεταβολή. Μεταξύ των μετρήσεων 1 και 2 αλλάζει η περίοδος, και, επομένως, η καμπύλη στάθμης-παροχής. Η μέτρηση 1 είναι η τελευταία μιας περιόδου και η 2 είναι η πρώτη της επόμενης περιόδου. Εφόσον έχει αλλάξει η καμπύλη στάθμης-παροχής, τα  $\Delta h$ , που είναι οι αποκλίσεις από τις καμπύλες, είναι ανόμοια και μη συγκρίσιμα και δεν γίνεται επομένως να θεωρήσουμε γραμμική μεταβολή μεταξύ τους. Η μόνη παραδοχή που μπορούμε να κάνουμε είναι ότι το  $\Delta h$  παραμένει σταθερό από την τελευταία μέτρηση ως το τέλος της περιόδου και από την αρχή της περιόδου ως την πρώτη μέτρηση.



Σχήμα 5.3: Διάφορες περιπτώσεις μεταβολής της διόρθωσης στάθμης

Η μέτρηση 3 έχει θεωρηθεί από το χρήστη ότι είναι λανθασμένη, και γι' αυτό έχει αγνοηθεί εντελώς, οπότε έχει θεωρηθεί γραμμική μεταβολή μεταξύ των μετρήσεων 2 και 4.

Μεταξύ των μετρήσεων 4 και 5 έχει συμβεί ένα ισχυρό επεισόδιο (π.χ. μια καταιγίδα), που ο χρήστης θεωρεί ότι είναι αυτό που άλλαξε τη συμπεριφορά της κοίτης. Γι' αυτό θεωρείται ότι η μεταβολή μεταξύ των 4 και 5 δεν ήταν γραμμική, αλλά συντελέστηκε απότομα κατά το ισχυρό επεισόδιο. Αντίστοιχη περίπτωση έχουμε μεταξύ των μετρήσεων 6 και 7, όπου το ισχυρό επεισόδιο είχε μεγάλη διάρκεια και θεωρείται ότι η μεταβολή του  $\Delta h$  έγινε γραμμικά κατά τη διάρκεια του επεισοδίου.

Από αυτές τις περιπτώσεις, μόνο η αλλαγή περιόδου φροντίζεται αυτόματα από την εφαρμογή· οι υπόλοιπες αποφασίζονται από το χρήστη. Το πότε έχει γίνει ισχυρό επεισόδιο, το ποια ήταν η διάρκεια του επεισοδίου, το αν το ισχυρό επεισόδιο είναι αυτό που άλλαξε τη συμπεριφορά, και το ποιες μετρήσεις είναι λανθασμένες επαφίενται όλα στην κρίση του χρήστη. Πάντως για μικρά  $\Delta h$  ο τρόπος μεταβολής δεν έχει μεγάλη σημασία. Πρέπει όμως να λαμβάνονται υπόψη και να εξετάζονται όλες οι μετρήσεις με μεγάλο  $\Delta h$ . Συγκεκριμένα, η εφαρμογή πρέπει να θέτει υπόψη του χρήστη όλες τις μετρήσεις που δεν είχαν ληφθεί υπόψη στην κατάρτιση της καμπύλης, καθώς και όλες τις μετρήσεις των οποίων το  $\Delta h$  υπερβαίνει ένα κατώφλι ορισμένο από το χρήστη.

Για την υλοποίηση αυτών των απαιτήσεων, το  $\Delta h$  πρέπει να θεωρείται χωριστή χρονοσειρά, με βήμα ίδιο με αυτό της χρονοσειράς στάθμης. Το σύστημα διαθέτει τα απαραίτητα οπτικά εργαλεία για να βοηθά το χρήστη να τη διαχειριστεί, καθώς και εύκολη πρόσβαση στις αντίστοιχες χρονοσειρές στάθμης και βροχοπτώσεων, ώστε ο χρήστης να διευκολύνεται στον εντοπισμό των ισχυρών επεισοδίων. Ο χρήστης μπορεί να προσδιορίσει τα ισχυρά επεισόδια και την έναρξη και λήξη τους, καθώς και τις μετρήσεις παροχής που δεν πρέπει να ληφθούν υπόψη, οπότε η εφαρμογή υπολογίζει τη διόρθωση στάθμης σύμφωνα με τους παραπάνω κανόνες. Ακόμα, ο χρήστης μπορεί να προσδιορίσει περιόδους στις οποίες δεν επιθυμεί να γίνει διόρθωση στάθμης, δηλαδή  $\Delta h = 0$ . Τέλος, ο χρήστης δύναται να επέμβει στη χρονοσειρά  $\Delta h$  με οποιονδήποτε τρόπο και να ορίσει τις τιμές που θεωρεί σωστές για οποιονδήποτε λόγο (δυνατότητα που καλύπτει τις προηγούμενες απαιτήσεις). Ο χρήστης στην συνέχεια, μπορεί να αποθηκεύσει τη χρονοσειρά  $\Delta h$  για περαιτέρω εξέταση.

Όταν η διόρθωση stout γίνεται σε στάθμη που ξεφεύγει από τα όρια της καμπύλης (είναι στην περιοχή επέκτασης), τότε μειώνεται με νόμο γραμμικής παρεμβολής έως και το μηδέν για στάθμη που βρίσκεται στην έναρξη της καμπύλης επέκτασης. Αυτή η μείωση δεν γίνεται όταν δεν έχει οριστεί καμπύλη επέκτασης οπότε η επέκταση γίνεται με extrapolation μίας κανονικής καμπύλης στάθμης παροχής.

#### **5.4.2 Διόρθωση στάθμης για τα σφάλματα των σταθμηγράφων**

Σε όλα τα παραπάνω έχουμε κάνει την παραδοχή ότι οι μετρήσεις ήταν σωστές, ή, τουλάχιστον, συνεπείς μεταξύ τους, και ότι οι αποκλίσεις των μετρήσεων παροχής από τις καμπύλες στάθμης – παροχής οφείλονται σε φυσικά αίτια. Στην πράξη υπάρχουν όμως επιπλέον προβλήματα οφειλόμενα στις ανακρίβειες των μετρήσεων, που στην Ελλάδα είναι ιδιαίτερα συνηθισμένες. Συγκεκριμένα, συμβαίνει συχνά να υπάρχουν ταυτόχρονες μετρήσεις από το σταθμηγράφο και από το σταθμήμετρο, και να είναι διαφορετικές. Ακόμα, συμβαίνει να έχει γίνει σταθμημέτρηση ταυτόχρονα με τη μέτρηση παροχής, και η στάθμη που έχει καταγράψει το συνεργείο μέτρησης παροχής να είναι διαφορετική από αυτήν που έχει καταγράψει ο παρατηρητής του σταθμημετρικού σταθμού. Η μέθοδος διόρθωσης στάθμης με τον τρόπο που περιγράφεται παρακάτω για να αντιμετωπίσει και αυτά τα προβλήματα προτείνεται από τους Κουτσογιάννη και Ξανθόπουλο (1999, σ. 330). Αυτός ο τρόπος μπορεί πολλές φορές να βελτιώσει τα στοιχεία και να δώσει καλύτερα αποτελέσματα. Σ' αυτό το εδάφιο περιγράφεται η χρήση της μεθόδου για διόρθωση των σφαλμάτων του σταθμηγράφου, και στο επόμενο η χρήση της μεθόδου για αντιμετώπιση των διαφορών των τακτικών σταθμημετρήσεων από τις μετρήσεις παροχής. Αυτές οι διορθώσεις πρέπει να γίνουν πριν από την ήδη περιγραφείσα.

Για να υπολογιστούν οι παροχές πρέπει να υπάρχει μία χρονοσειρά στάθμης. Αυτή η χρονοσειρά είναι οι μετρήσεις του σταθμημέτρου που καταγράφηκε κατά την διεξαγωγή των μετρήσεων παροχής και που χρησιμοποιήθηκε για την εξαγωγή των καμπυλών στάθμης-παροχής. Αν στη θέση αυτού του σταθμημέτρου μετρά και σταθμηγράφος, τότε μπορούν να χρησιμοποιηθούν και οι μετρήσεις του σταθμηγράφου. Γενικά δεν είναι σίγουρο ότι θα υπάρχουν σε όλες τις περιόδους μετρήσεις από αμφοτέρωτα τα όργανα. Σε περιόδους που ο σταθμηγράφος έχει βλάβη, θα υπάρχουν μόνο οι μετρήσεις από το σταθμήμετρο. Σε περιόδους που π.χ. ο παρατηρητής ασθενεί, ή σε περιπτώσεις απώλειας

εντύπων, θα υπάρχουν μόνο οι μετρήσεις από το σταθμηγράφο. Εν γένει τα δεδομένα μας είναι μια ημερήσια χρονοσειρά μετρήσεων  $sm$  προερχόμενη από το σταθμήμετρο και μια ωριαία ή δίωρη χρονοσειρά  $sg$  προερχόμενη από το σταθμηγράφο.

Αν δεν υπήρχαν σφάλματα στις μετρήσεις, τότε η ένωση  $ss$  αυτών των δύο χρονοσειρών θα μας έδινε μία χρονοσειρά από στάθμες, η οποία θα ήταν ακανόνιστου βήματος – άλλοτε ημερήσιου και άλλοτε ωριαίου ή δίωρου. Υπάρχουν χρονικές στιγμές στις οποίες υπάρχει μέτρηση και από το σταθμήμετρο και από το σταθμηγράφο. Κανονικά θα έπρεπε αυτές οι δύο μετρήσεις να συμπίπτουν, αλλά εν γένει διαφέρουν, είτε λόγω κακής ρύθμισης του σταθμηγράφου, είτε λόγω κακής τοποθέτησης του χαρτιού στο σταθμηγράφο, είτε λόγω του σφάλματος της ψηφιοποίησης, είτε λόγω άλλων σφαλμάτων. Εξαιτίας αυτής της διαφοράς, η απλή ένωση των δύο χρονοσειρών – σταθμημέτρου και σταθμηγράφου – είναι αδύνατη και απαιτούνται διορθώσεις.

Στις περιπτώσεις που υπάρχει διαφορά, μπορεί να θεωρηθεί σωστή η ένδειξη του σταθμημέτρου ή του σταθμηγράφου, ή να θεωρηθούν και οι δύο λάθος και να εξαχθεί π.χ. ο μέσος όρος τους. Η παραδοχή που γίνεται είναι ότι είναι σωστή η μέτρηση του σταθμημέτρου, για τους εξής λόγους:

- Ο σταθμηγράφος είναι όργανο ευαίσθητο, ενώ το σταθμήμετρο είναι αξιόπιστο και η ανάγνωσή του είναι απλή διαδικασία
- Στόχος μας είναι να υπολογίσουμε όσο το δυνατόν πιο σωστή παροχή. Επομένως δεν μας ενδιαφέρει αυτή καθαυτή η σωστή στάθμη, που είναι άλλωστε συμβατικό μέγεθος, ανάλογα με το πού έχει οριστεί το μηδέν, αλλά εκείνη η στάθμη που ακολουθεί την ίδια σύμβαση με τη στάθμη από την οποία έχει προκύψει η καμπύλη στάθμης-παροχής. Η καμπύλη έχει προκύψει από τις μετρήσεις παροχής, και η στάθμη που αναφέρεται στις μετρήσεις παροχής έχει διαβαστεί από το σταθμήμετρο, και όχι από το σταθμηγράφο.

Θεωρούμε δύο διαδοχικές μετρήσεις σταθμημέτρου  $sm_i$  και  $sm_{i+1}$ , μεταξύ των οποίων παρεμβάλλονται μετρήσεις σταθμηγράφου. Ταυτόχρονα με τις μετρήσεις  $sm_i$  και  $sm_{i+1}$  υπάρχουν και μετρήσεις σταθμηγράφου  $sg_a$  και  $sg_b$ . Στην πραγματικότητα όταν λέμε ταυτόχρονα υπάρχει ανοχή: αν η μέτρηση του σταθμημέτρου και του σταθμηγράφου έχουν ώρα 08:00 και 08:05, τότε προφανώς θεωρούνται ταυτόχρονες. Η ανοχή τίθεται περίπου μία ώρα. Όπως αναφέρθηκε, εν γένει οι  $sg_a$  και  $sg_b$  διαφέρουν από τις  $sm_i$  και  $sm_{i+1}$  κατά  $\Delta sg_a$  και  $\Delta sg_b$ :

$$\Delta sg_a = sm_i - sg_a \quad 5.13$$

$$\Delta sg_b = sm_{i+1} - sg_b \quad 5.14$$

Αν τώρα, θεωρήσουμε ότι το  $\Delta sg$  μεταβάλλεται γραμμικά μεταξύ των δύο μετρήσεων, μπορούμε να διορθώσουμε και όλες τις ενδιάμεσες στάθμες της χρονοσειράς  $sg$ , χρησιμοποιώντας π.χ. την Σχέση (5.12) όπου στην θέση των  $\Delta h$  χρησιμοποιούνται οι διαφορές  $\Delta sg_a$  και  $\Delta sg_b$ .

Η διόρθωση στάθμης της χρονοσειράς  $sg$  δίνει τη διορθωμένη χρονοσειρά  $sg'$ . Ακολούθως, η ένωση της χρονοσειράς  $sm$  με τη διορθωμένη χρονοσειρά  $sg'$  μας δίνει τη ζητούμενη ενιαία χρονοσειρά  $ss$ , που έχει το βήμα της χρονοσειράς  $sg$  (π.χ. ωριαίο). Στις περιόδους που υπάρχουν μόνο μετρήσεις σταθμημέτρου και όχι σταθμηγράφου, μεταξύ δυο μετρήσεων του σταθμημέτρου θα παρεμβάλλονται ελλείπουσες τιμές.

Παρόλο που η μέθοδος αυτή είναι χρήσιμη για τη διόρθωση μικρών λαθών, χρειάζεται ιδιαίτερη προσοχή στην εφαρμογή της. Αν η χρονοσειρά  $sm$  έχει προκύψει από σταθμήμετρο που μετρά σε άλλη θέση από αυτήν του σταθμηγράφου (γιατί εκείνο έχει καταγραφεί στις μετρήσεις παροχής), τότε τα στοιχεία του σταθμηγράφου δεν πρέπει να χρησιμοποιηθούν, τουλάχιστον όχι με αυτή τη μέθοδο. Υπάρχει κίνδυνος κάποιος όχι καλά διαβασμένος χρήστης να χρησιμοποιεί αυτή τη μέθοδο, αφού του την προσφέρει η εφαρμογή, νομίζοντας ότι δίνει σωστά αποτελέσματα σε όλες τις περιπτώσεις.

Πρέπει λοιπόν η μέθοδος να χρησιμοποιείται αυστηρά για τα λάθη περιορισμένης έκτασης για τα οποία προβλέπεται, ενώ κάθε διαφορά σταθμημέτρου με σταθμηγράφο πρέπει να εξεταστεί χωριστά από το χρήστη πριν αποφασιστεί τι ακριβώς θα γίνει.

### 5.4.3 Διόρθωση στάθμης για τις διαφορές κατά τις μετρήσεις παροχής

Το αποτέλεσμα της διόρθωσης δεδομένων σταθμηγράφου, που μόλις περιγράφηκε, είναι η χρονοσειρά  $ss$ , η οποία είναι υπερσύνολο της  $sm$ . Το νέο πρόβλημα που αντιμετωπίζουμε τώρα είναι ότι οι στάθμες της χρονοσειράς κατά τις μετρήσεις παροχής  $hm$  δεν ταυτίζονται πάντα με τις ταυτόχρονες τιμές της  $sm$ . Αυτό οφείλεται στους εξής λόγους:

- Η μέτρηση παροχής διαρκεί περίπου δύο ώρες, στη διάρκεια των οποίων η στάθμη μπορεί να μεταβάλλεται. Γι' αυτό καταγράφεται η στάθμη στην αρχή και στο τέλος της μέτρησης παροχής και εξάγεται μέσος όρος. Αυτή η διαδικασία μπορεί να δίνει μικροδιαφορές από την όποια τακτική σταθμημέτρηση του σταθμού θεωρούμε, με ορισμένη χρονική ανοχή, ότι έχει γίνει ταυτόχρονα
- Μπορεί οι πραγματικές ώρες λήψης των ενδείξεων του σταθμημέτρου να μην είναι οι αναγραφόμενες στα έντυπα. Μπορεί να έχει γίνει λανθασμένη καταγραφή ώρας, ή μπορεί ο παρατηρητής του σταθμημετρικού σταθμού να πηγαίνει σε ακανόνιστες ώρες και κάθε φορά να σημειώνει στο έντυπο ότι πήγε στις 08:00. Έτσι, μετρήσεις παροχής και μετρήσεις στάθμης που νομίζουμε ότι είναι ταυτόχρονες μπορεί στην πραγματικότητα να μην είναι ταυτόχρονες
- Υπάρχουν τα σφάλματα που οφείλονται στην καταγραφή λανθασμένου σταθμημέτρου. Αν και τα συνεργεία μέτρησης παροχής έχουν υποχρέωση να καταγράφουν τις ενδείξεις όλων των σταθμημέτρων, μερικές φορές καταγράφουν λιγότερα ή μόνο ένα. Μπορεί να μη γνωρίζουμε ποιο σταθμημέτρο είναι το καταγραφέν, ή μπορεί να είναι διαφορετικό από εκείνο που καταγράφεται από τον παρατηρητή στις τακτικές σταθμημετρήσεις (σε μερικούς σταθμούς με πολλά σταθμημέτρα δεν καταγράφονται όλα).

Στόχος μας είναι να υπολογίσουμε όσο το δυνατόν πιο σωστή παροχή. Επομένως δεν μας ενδιαφέρει αυτή καθαυτή η σωστή στάθμη, που είναι άλλωστε συμβατικό μέγεθος, ανάλογα με το πού έχει οριστεί το μηδέν, αλλά εκείνη η στάθμη που ακολουθεί την ίδια σύμβαση με τη στάθμη από την οποία έχει προκύψει η καμπύλη στάθμης-παροχής. Η καμπύλη έχει προκύψει από τις μετρήσεις παροχής. Επομένως, όταν υπάρχει διαφορά μεταξύ  $ss$  και μέτρησης παροχής, θεωρούμε ότι σωστή είναι η τιμή της μέτρησης παροχής. Η διόρθωση γίνεται με τον ίδιο τρόπο όπως και πριν: θεωρούμε δύο διαδοχικές μετρήσεις παροχής  $hm_i$  και  $hm_{i+1}$ , μεταξύ των οποίων παρεμβάλλονται μετρήσεις στάθμης της χρονοσειράς  $ss$ . Ταυτόχρονα με τις μετρήσεις  $hm_i$  και  $hm_{i+1}$  υπάρχουν και μετρήσεις στάθμης  $ss_a$  και  $ss_b$ . Όταν λέμε *ταυτόχρονα* υπάρχει ανοχή, που τίθεται περίπου μια ώρα.

Κατά τον ίδιο τρόπο, όπως προηγουμένως, έχουμε:

$$\Delta ss_a = hm_i - ss_a \quad 5.15$$

$$\Delta ss_b = hm_{i+1} - ss_b \quad 5.16$$

οπότε θεωρείται ότι το  $\Delta ss$  μεταβάλλεται γραμμικά μεταξύ των δύο μετρήσεων παροχής και εφαρμόζουμε κατά τα γνωστά την Σχέση (5.12) για να ανάγουμε τις ενδιάμεσες στάθμες χρησιμοποιώντας τα  $\Delta ss_a$  και  $\Delta ss_b$  στην θέση του  $\Delta h$ .

Με αυτή τη μέθοδο, οι στάθμες  $ss'$  που θα ληφθούν στα σημεία των μετρήσεων παροχής θα είναι σωστές (θα συμπίπτουν με τις στάθμες μετρήσεων παροχής). Δυστυχώς, δεν υπάρχει ένδειξη ότι στα σημεία μεταξύ των μετρήσεων παροχής η μέθοδος είναι ιδιαίτερα ωφέλιμη, καθόσον μάλιστα οι μετρήσεις παροχής μπορεί να απέχουν σημαντικό χρονικό διάστημα μεταξύ τους. Αν οι διαφορές  $\Delta ss_a$  και  $\Delta ss_b$  οφείλονται σε καταγραφή λανθασμένου σταθμημέτρου, ουσιαστικά δεν υπάρχει τίποτε που



να μπορεί να γίνει για να διορθωθούν με κάποια βεβαιότητα οι στάθμες των ενδιάμεσων σημείων. Όμως αυτή η μέθοδος έχει πιθανότητα να βελτιώσει, μερικές φορές, τις σύτως ή άλλως λανθασμένες τιμές των σημείων. Χρειάζεται όμως και εδώ, όπως και στην προηγούμενη περίπτωση, ιδιαίτερη προσοχή ώστε να μην γίνεται κατάχρηση της μεθόδου. Συγκεκριμένα, η μέθοδος είναι εντελώς ακατάλληλη για προσαρμογή μετρήσεων σταθμημέτρου διαφορετικού από εκείνου από το οποίο προέκυψαν οι μετρήσεις παροχής (και, επομένως, οι καμπύλες).

Για την υλοποίηση των δύο διορθώσεων που περιγράφηκαν σ' αυτό και στο προηγούμενο εδάφιο, τα  $\Delta sg$  και  $\Delta ss$  πρέπει να θεωρούνται χωριστές χρονοσειρές, με βήμα ίδιο με αυτό της χρονοσειράς στάθμης. Η εφαρμογή διαθέτει τα απαραίτητα οπτικά εργαλεία για να βοηθά το χρήστη να τις διαχειριστεί, καθώς και εύκολη πρόσβαση σε όλες τις σχετικές χρονοσειρές σταθμημέτρων, σταθμηγράφων και μετρήσεων παροχής, ώστε ο χρήστης να διευκολύνεται στον εντοπισμό των λαθών. Ο χρήστης μπορεί να προσδιορίσει μετρήσεις που δεν πρέπει να ληφθούν υπόψη. Τέλος, ο χρήστης δύναται να επέμβει στις χρονοσειρές  $\Delta sg$  και  $\Delta ss$  με οποιονδήποτε τρόπο και να ορίσει τις τιμές που θεωρεί σωστές για οποιονδήποτε λόγο. Ο χρήστης είναι δηλαδή αυτός που αποφασίζει σε λεπτομέρεια πώς θα γίνουν οι διορθώσεις. Η εφαρμογή εντοπίζει τα σημεία των διαφορών, και τις προβάλλει, με τα οπτικά βοηθήματα, περιμένοντας από το χρήστη εντολή για το πώς θα γίνουν οι διορθώσεις. Τελικά ο χρήστης έχει την δυνατότητα να αποθηκεύσει τις παραπάνω χρονοσειρές.

Οι δύο διορθώσεις που αναφέρθηκαν σ' αυτό και στο προηγούμενο εδάφιο είναι εμπειρικές και προέκυψαν από την ιδέα της μεθόδου διόρθωσης της στάθμης. Οι διορθώσεις που εισάγουν είναι γενικά επισφαλείς. Ο μελετητής που εξάγει παροχές μπορεί μεν να τις χρησιμοποιήσει για να διευκολυνθεί, αλλά η ύπαρξη ανωμαλιών που να απαιτούν αυτές τις δύο διορθώσεις φανερώνει κακή ποιότητα στα δεδομένα, η οποία πρέπει να ληφθεί σοβαρά υπόψη ώστε να εξαχθούν τα κατάλληλα συμπεράσματα για την αξιοπιστία της σειράς παροχών που προκύπτει.

#### 5.4.4 Η εξαγωγή των παροχών

Όπως είδαμε, συνολικά γίνονται τρεις διορθώσεις στη στάθμη: η διόρθωση  $\Delta sg$  των σφαλμάτων του σταθμηγράφου, η διόρθωση  $\Delta ss$  των αποκλίσεων σταθμημετρήσεων-μετρήσεων παροχής, και η διόρθωση  $\Delta h$  για τις αποκλίσεις των μετρήσεων παροχής από τις καμπύλες. Η διόρθωση  $\Delta sg$  πρέπει να γίνει στη χρονοσειρά του σταθμηγράφου  $sg$ . Η διόρθωση  $\Delta ss$  πρέπει να γίνει στην ενιαία σειρά στάθμης  $ss$  (η οποία ήδη περιλαμβάνει τη διόρθωση  $\Delta sg$ ). Τέλος, η διόρθωση  $\Delta h$  (stout) πρέπει να γίνει στη διορθωμένη ενιαία σειρά στάθμης  $ss'$  (η οποία ήδη περιλαμβάνει τις άλλες δύο διορθώσεις), οπότε οι παροχές υπολογίζονται από τη σχέση

$$Q = f(ss' + \Delta h) \quad 5.17$$

όπου  $f$  η εκάστοτε ισχύουσα καμπύλη στάθμης-παροχής. Για την ακρίβεια,

$$ss = sm \cup (sg + \Delta sg) \quad 5.18$$

$$Q = f(ss + \Delta ss + \Delta h) \quad 5.19$$

Αν τα δεδομένα δεν έχουν σφάλματα, τότε  $\Delta sg = \Delta ss = 0$ .

Ο χρήστης τελικά μπορεί να αποθηκεύσει την τελική χρονοσειρά παροχών. Πάντα όμως πρέπει να αποθηκεύονται και οι αντίστοιχες χρονοσειρές διορθώσεων  $\Delta sg$ ,  $\Delta ss$ ,  $\Delta h$  (αν δεν είναι ήδη αποθηκευμένες).

## 5.5 Κατάρτιση καμπυλών στάθμης – παροχής

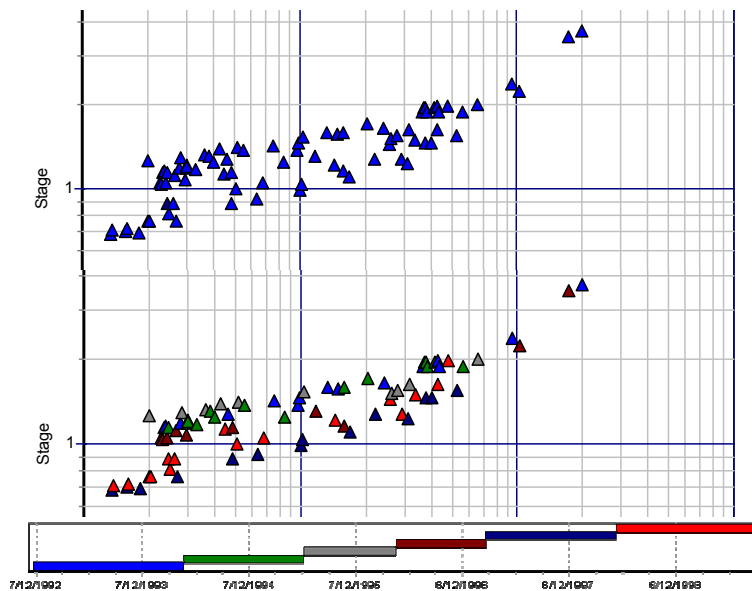
Το πρόβλημα της κατάρτισης καμπυλών στάθμης – παροχής είναι αρκετά σύνθετο, καθώς όπως είδαμε οι καμπύλες είναι χρονικά μεταβαλλόμενες, επιπλέον δεν διαθέτουμε υδρομετρήσεις για όλο το εύρος παροχών του υδατορεύματος. Για να απλοποιηθεί η διαδικασία γίνεται ο χωρισμός των εργασιών στα εξής στάδια:

- Χωρισμός των υδρομετρήσεων σε ομοιογενείς περιόδους
- Εντοπισμός των εξωκείμενων σημείων στάθμης – παροχής
- Χάραξη της καμπύλης στάθμης παροχής για κάθε περίοδο χωριστά
- Εκτίμηση της καμπύλης επέκτασης.

Τα παραπάνω στάδια μπορούν να αυτοματοποιηθούν σε μεγάλο βαθμό χρησιμοποιώντας τεχνικές όπως οι γενετικοί αλγόριθμοι. Οι Tsakalias και Koutsoyiannis (1999) είχαν κατασκευάσει ένα τέτοιο σύστημα το οποίο πρόκειται να αξιοποιηθεί σε μελλοντική έκδοση του συστήματος του «Υδρογνώμων». Προς το παρόν εφαρμόζεται μία ημιγραφική μέθοδος για τα δύο πρώτα στάδια (χωρισμός περιόδων – εντοπισμός των εξωκείμενων σημείων) η οποία έχει το πλεονέκτημα της φιλικότητας και του καλού ελέγχου από τον χρήστη.

### 5.5.1 Ορισμός χρονικών περιόδων – εντοπισμός εξωκείμενων σημείων

Αφού γίνεται μία καταρχήν εκτίμηση για τον αριθμό των διαφορετικών χρονικών περιόδων το σύστημα απεικονίζει τα σημεία των υδρομετρήσεων με διαφορετικό χρώμα για κάθε περίοδο. Ο χρήστης μεταβάλλει την έναρξη και το πέρας κάθε χρονικής περιόδου χρησιμοποιώντας απλά εργαλεία, όπως μία ολισθαίνουσα μπάρα. Ο χρήστης έχει την δυνατότητα όποτε το κρίνει σκόπιμο να εισάγει νέες καμπύλες ή και να διαγράψει. Η εμπειρία έχει δείξει πως πιθανές χρονικές στιγμές αλλαγής περιόδου είναι τα έντονα πλημμυρικά επεισόδια, τέτοια που μπορούν να αλλάξουν τα χαρακτηριστικά της κοίτης.



Σχήμα 5.4: Διαχωρισμός των υδρομετρήσεων σε ομοιογενείς περιόδους (με διαφορετικό χρωματισμό). Υδρομετρήσεις 1992-1998 από τον σταθμό «Πόρος Ρηγαίου» στον ποταμό Εύηνο.

Το σύστημα παρέχει διάφορες διευκολύνσεις στον χρήστη, όπως δυνατότητα εμφάνισης μόνο μίας περιόδου, επιλεκτικής εμφάνισης των σημείων ή των καμπυλών, απόκρυψη των εξωκείμενων σημείων κ.α. Ο χρήστης τελικά αποφασίζει τις θέσεις των αλλαγών των περιόδων ανάλογα με την τάση των σημείων να προσαρμόζονται σε απλά γεωμετρικά μοντέλα όπως μία ευθεία, ή μία θλάση,

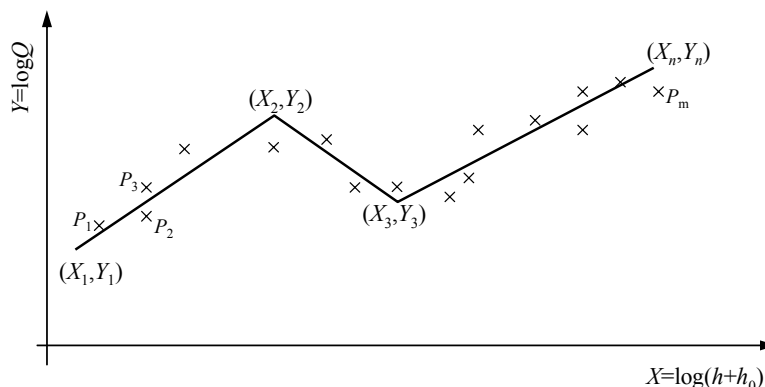
κλπ. Ο χρήστης μπορεί να δοκιμάσει διαφορετικές λύσεις και στην συνέχεια αφού προσαρμόσει και τις καμπύλες στάθμης παροχής, να τις αποθηκεύσει για να τις συγκρίνει μεταξύ τους σε κάποιο μεταγενέστερο στάδιο.

Τέλος ο χρήστης αποφασίζει ποια σημεία είναι εξωκείμενα και τα επισημαίνει με την κατάλληλη γραφική λειτουργία. Το κριτήριο είναι εμπειρικό και βασίζεται αποκλειστικά στις ενέργειες του χρήστη. Ο χρήστης μπορεί να δικαιολογήσει τον χαρακτηρισμό ενός σημείου ως εξωκείμενο βασισμένος σε μεθοδολογία με στατιστική τεκμηρίωση, παρατηρώντας δηλαδή την βελτίωση του συντελεστή προσδιορισμού αφαιρώντας τέτοια σημεία.

### 5.5.2 Προσαρμογή καμπυλών στις υδρομετρήσεις

Η προσαρμογή των καμπυλών στάθμης παροχής στα σημεία των υδρομετρήσεων γίνεται με την μέθοδο της γραμμικής παλινδρόμησης. Συγκεκριμένα γίνεται θεώρηση ενός αριθμού θλάσεων (από 0 – ευθύγραμμο τμήμα έως  $n-2$ ). Ο αριθμός των θλάσεων αποφασίζεται από τον χρήστη και μπορεί να είναι διαφορετικός για κάθε χρονική περίοδο. Στην συνέχεια μέσω ενός εξαντλητικού αλγορίθμου ελέγχονται όλες οι τιμές στάθμης των υδρομετρήσεων ως εν δυνάμει σημεία θλάσης, τελικά επιλέγεται το σετ των σημείων θλάσης που δίνει τον καλύτερο συντελεστή προσδιορισμού. Το ταξινομημένο σετ πρέπει να αποτελείται από αύξουσες τιμές της στάθμης και της παροχής.

Πριν ξεκινήσουμε τον προσδιορισμό των καμπυλών πρέπει να τεθεί κάποια τιμή για τον σταθερό όρο  $h_0$  (βλ. εισαγωγή της παρούσας ενότητας). Συνήθως επιλέγεται μία τιμή για όλες τις χρονικές περιόδους, τέτοια που αν προστεθεί στις μετρημένες τιμές της στάθμης  $h_i$ , αποτρέπει αρνητικές τιμές που θα προκαλούσαν σφάλματα κατά την λογαρίθμηση.



Σχήμα 5.5: Μία εν δυνάμει καμπύλη στάθμης παροχής. Τα  $X_i$  είναι δεδομένα και αναζητείται το βέλτιστο σετ  $Y_i$

Ο συντελεστής προσδιορισμού υπολογίζεται από την γενικευμένη Σχέση (7.9) και είναι το μέτρο για την ποιότητα της προσαρμογής κάθε καμπύλης χωριστά αλλά και όλων των καμπυλών για όλες τις χρονικές περιόδους. Έστω  $n-2$  ο αριθμός των θλάσεων που ορίζει  $n-1$  ευθύγραμμα τμήματα από  $n$  σημεία  $(X_i, Y_i)$ . Ως  $X_i$  ορίζονται οι τιμές της ανεξάρτητης μεταβλητής και είναι οι λογαριθμικοποιημένες τιμές της στάθμης (συν τον σταθερό όρο  $h_0$ ), αντίστοιχα  $Y_i$  οι τιμές της εξαρτημένης μεταβλητής και είναι οι λογαριθμικοποιημένες τιμές της παροχής. Τα σημεία  $X_i$  είναι δεδομένα και προέρχονται από ένα σετ τιμών στάθμης από τις υδρομετρήσεις. Ο σκοπός μας είναι να υπολογίσουμε το βέλτιστο σετ  $Y_i$  για τα δεδομένα  $X_i$  καθώς και έναν συντελεστή προσδιορισμού που θα χρησιμοποιηθεί για την τελική επιλογή. Αποδεκτά αποτελέσματα θα είναι αυτά που θα δίνουν αύξουσες τιμές για την παροχή.

Έστω  $P_j$  τα  $m$  σημεία των υδρομετρήσεων με συντεταγμένες  $(x_{Pj}, y_{Pj})$ , οι οποίες προέρχονται από την λογαριθμικοποίηση των τιμών της στάθμης και της παροχής κατά τα γνωστά. Εξετάζουμε κάθε

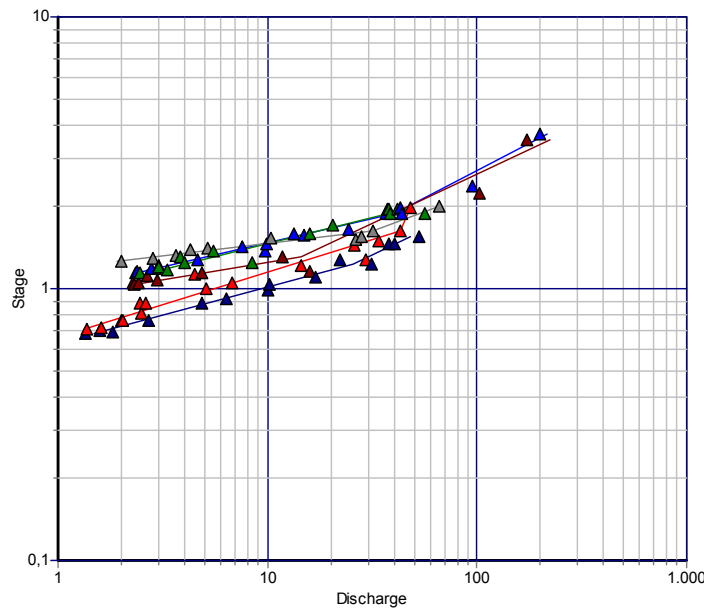
ευθύγραμμο τμήμα  $[i-1 - i]$  χωριστά χρησιμοποιώντας τα σημεία των υδρομετρήσεων  $P_j$  που αφορούν μόνο το συγκεκριμένο τμήμα ( $X_{i-1} \leq x_{Pj} < X_i$ ). Διατυπώνοντας το άθροισμα των τετραγώνων και στην συνέχεια παραγωγίζοντας το ως προς  $Y_{i-1}$  για να ελαχιστοποιηθεί το τετραγωνικό σφάλμα προκύπτει η ακόλουθη σχέση:

$$\sum_{X_{i-1} \leq x_{Pj} < X_i} \left( \frac{x_{Pj} - X_{i-1}}{X_i - X_{i-1}} - 1 \right)^2 Y_{i-1} - \sum_{X_{i-1} \leq x_{Pj} < X_i} \left[ \left( \frac{x_{Pj} - X_{i-1}}{X_i - X_{i-1}} \right) \left( \frac{x_{Pj} - X_{i-1}}{X_i - X_{i-1}} - 1 \right) \right] Y_i = \sum_{X_{i-1} \leq x_{Pj} < X_i} \left[ \left( 1 - \frac{x_{Pj} - X_{i-1}}{X_i - X_{i-1}} \right) y_{Pj} \right] \quad 5.20$$

Η παραπάνω σχέση για  $i=2 \dots n$  δίνει  $n-1$  εξισώσεις για  $n$  αγνώστους (τα  $Y_i$ ). Μπορούμε όμως να παράξουμε μία ακόμα εξίσωση παραγωγίζοντας ως προς  $Y_n$  το οποίο επιτυγχάνεται απλά αντιστρέφοντας τους συντελεστές των  $Y_{i-1}$  και  $Y_i$  της Σχέσης (5.20) για  $i=n$ . Τελικά η επίλυση του συστήματος  $n$  εξισώσεων δίνει τα άγνωστα  $Y_i$ . Τα σημεία ( $X_i, Y_i$ ) είναι μία εν δυνάμει λύση καμπύλης στάθμης παροχής. Τελικά επιλέγεται η λύση που θα δίνει μονότονα  $Y_i$  και τον βέλτιστο συντελεστή προσδιορισμού ο οποίος υπολογίζεται από την γενική Σχέση (7.9). Αφού υπολογιστεί το βέλτιστο σετ ( $X_i, Y_i$ ) μετά από ένα πλήθος δοκιμών, η καμπύλη στάθμης παροχής θα έχει τα εξής σημεία:

$$(h_i, Q_i) = (e^{X_i} - h_0, e^{Y_i}) \quad 5.21$$

Στην συνέχεια τα σημεία ( $h_i, Q_i$ ) αποτελούν μία καμπύλη στάθμης παροχής αφού εφαρμοστεί η λογαριθμική παρεμβολή σύμφωνα με την μεθοδολογία που περιγράφεται στην αρχή της παρούσας ενότητας.



Σχήμα 5.6: Χάραξη καμπυλών στάθμης παροχής. Καμπύλες με μία θλάση προσαρμοσμένες στον διαχωρισμό των περιόδων του Σχήματος 5.4

### 5.5.3 Εισαγωγή καμπυλών επέκτασης με εξισώσεις της υδραυλικής

Οι καμπύλες επέκτασης αφορούν τις σχέσεις που συνδέουν τις τιμές της στάθμης και παροχής πέρα από κάποιο μέγεθος, εκεί δηλαδή που δεν διαθέτουμε υδρομετρήσεις. Για την περίπτωση των καμπυλών επέκτασης γίνεται η παραδοχή πως μία καμπύλη επέκτασης αρκεί για να περιγράψει την σχέση στάθμης και παροχής για όλες τις περιόδους. Είναι δε παρατηρημένο πως οι καμπύλες στάθμης

– παροχής διαφορετικών περιόδων έχουν την τάση να συγκλίνουν για μεγάλες τιμές στάθμης και παροχής.

Η μεθοδολογία εξαγωγής των σχέσεων για τις καμπύλες επέκτασης δεν βασίζεται αποκλειστικά στις υδρομετρήσεις, καθώς δεν υπάρχουν διαθέσιμες αυτήν την περιοχή τιμών, συνήθως εφαρμόζεται κάποια μεθοδολογία που συνδυάζει εξισώσεις της υδραυλικής και τις υψηλότερες τιμές από τις υδρομετρήσεις. Η κατάρτιση μπορεί να γίνει και με αποκλειστική χρήση των σχέσεων της υδραυλικής εφόσον είναι γνωστά τα στοιχεία τραχύτητας και κλίσης του «αγωγού» που προσομοιάζει το υδατόρευμα. Επιπλέον ο χρήστης μπορεί να χρησιμοποιήσει αυτήν την μεθοδολογία κατάρτισης καμπύλων και για συνηθισμένες τιμές στάθμης – παροχής σε συνδυασμό με ή και χωρίς καθόλου υδρομετρήσεις.

### Οι σχέσεις των Manning και Chezy

Οι εξισώσεις που χρησιμοποιούνται για την κατάρτιση των καμπυλών επέκτασης είναι αυτές των Manning και Chezy (Κουτσογιάννης, 1999, Δημητρίου 1994). Πρόκειται για δύο από τις πλέον διαδεδομένες σχέσεις για τους υπολογισμούς των παροχών σε αγωγούς με ελεύθερη επιφάνεια. Η σχέση του Manning δίνει την ταχύτητα ροής σε αγωγό με ελεύθερη επιφάνεια ως:

$$V = \frac{1}{n} R^{2/3} \sqrt{i} \quad 5.22$$

όπου  $n$  ο αδιάστατος συντελεστής τραχύτητας,  $R$  η υδραυλική ακτίνα της υδάτινης διατομής και  $i$  η κλίση της ελεύθερης επιφάνειας. Για την ειδική περίπτωση της ομοιόμορφης ροής η κλίση  $i$  ταυτίζεται με την κλίση του πυθμένα  $J$ . Η δε παροχή κατά Manning προκύπτει πολλαπλασιάζοντας την ταχύτητα με την επιφάνεια της υδάτινης διατομής  $A$ :

$$Q = \frac{1}{n} AR^{2/3} \sqrt{i} \quad 5.23$$

Η σχέση του Chezy για την ταχύτητα ροής είναι αντίστοιχα:

$$V = C\sqrt{Ri} \quad 5.24$$

όπου  $C$  ο συντελεστής τραχύτητας του Chezy,  $R$  η υδραυλική ακτίνα της υδάτινης διατομής και  $i$  η κλίση της ελεύθερης επιφάνειας. Πολλαπλασιάζοντας με την επιφάνεια της διατομής  $A$  προκύπτει το μέγεθος της παροχής κατά Chezy:

$$Q = CA\sqrt{Ri} \quad 5.25$$

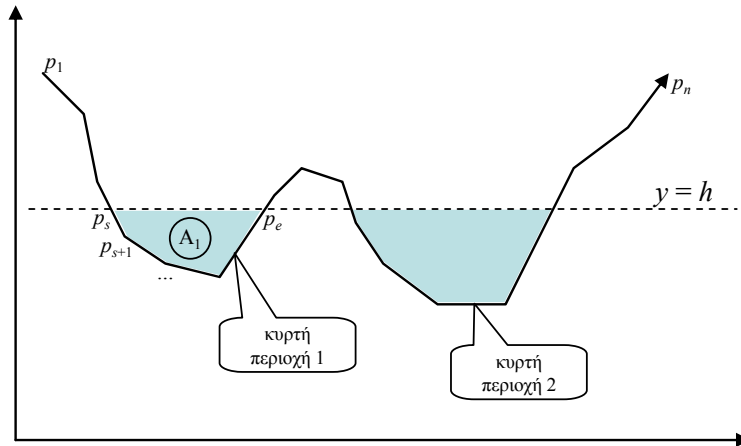
Η σχέση του Chezy δίνει «ηπιότερες» παροχές και ορισμένες φορές προτιμάται η χρήση της αντί της σχέσης του Manning. Τα στοιχεία της υγρής διατομής  $A$  και  $R$  είναι συνήθως γνωστά από τα στοιχεία της διατομής του υδατορεύματος και την τιμή της στάθμης. Όσον αφορά δε τους συντελεστές τραχύτητας ( $n$  για την σχέση του Manning και  $C$  για την σχέση του Chezy) καθώς και την υδραυλική κλίση  $i$  (ή  $J$  με θεώρηση ομοιόμορφης ροής), ο χρήστης έχει την δυνατότητα να τους θέσει ρητά ή να προσδιορίσει τα σχετικά γινόμενα ( $\sqrt{i}/n$  για την σχέση του Manning και  $C\sqrt{i}$  για την σχέση του Chezy) με χρήση των μεγαλύτερων υδρομετρήσεων και βέλτιστη προσαρμογή με την διαδικασία της γραμμικής παλινδρόμησης.

### Υπολογισμός των στοιχείων της υδάτινης διατομής

Προκειμένου να γίνουν οι υπολογισμοί παροχής με τις σχέσεις των Manning ή Chezy, πρέπει να υπολογιστούν τα στοιχεία της υδάτινης διατομής, συγκεκριμένα η υδραυλική ακτίνα  $R$  καθώς και η επιφάνεια της υδάτινης διατομής  $A$ . Η υδραυλική ακτίνα προκύπτει από την γνωστή σχέση:

$$R = \frac{A}{\Pi} \quad 5.26$$

όπου  $\Pi$  είναι η βρεχόμενη περίμετρος του υδατορεύματος. Αν η διατομή του υδατορεύματος αναπαρασταθεί με μία τεθλασμένη γραμμή με τα  $n$  σημεία  $p_i = (x_i, y_i)$ , τότε αρκεί να βρεθούν κυρτές περιοχές που ορίζονται από τις τομές με την ευθεία  $y = h$ , όπου  $h$  η στάθμη του νερού. Οι κυρτές περιοχές ορίζονται από τα σημεία τομής με την διατομή του υδατορεύματος  $(x_s, y_s)$  και  $(x_e, y_e)$  καθώς και τα ενδιάμεσα σημεία που περικλείονται  $i = k \dots l$  ( $l < k < l < n$ ).



Σχήμα 5.7: Η υγρή διατομή σε σχέση με την διατομή του υδατορεύματος και την στάθμη του νερού.

Ως επιφάνεια λαμβάνεται το άθροισμα των επιφανειών των επιμέρους κυρτών περιοχών, το αυτό και για την βρεχόμενη περίμετρο. Οι σχέσεις που δίνουν αυτά τα μεγέθη είναι:

$$\Pi = \sum_{i=s}^{e-1} \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2} \quad 5.27$$

και

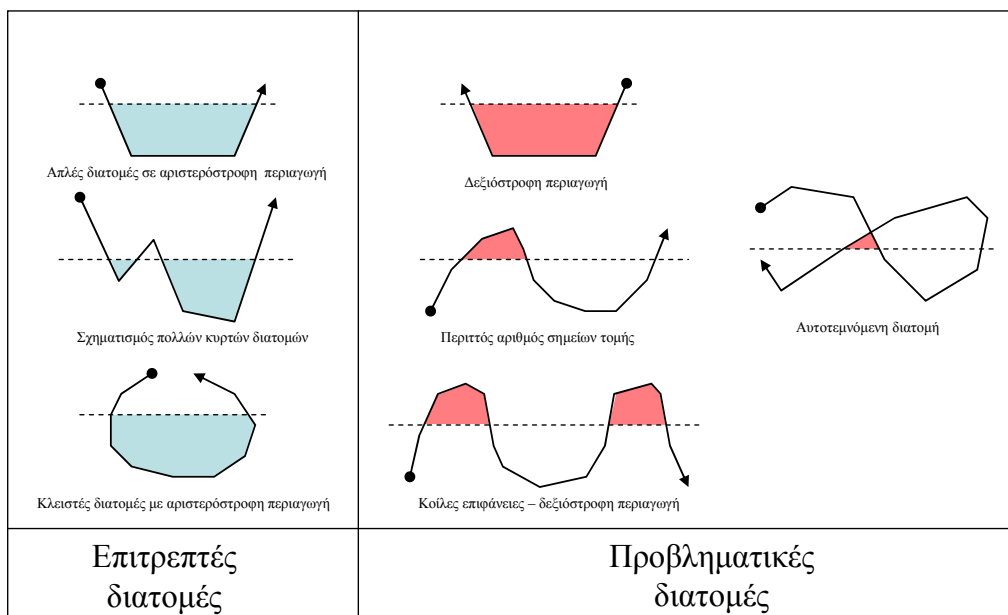
$$A = \frac{1}{2} \sum_{i=s}^e (x_i y_{i+1} - x_{i+1} y_i) \quad 5.28$$

όπου  $p_{e+1} \equiv p_s$ .

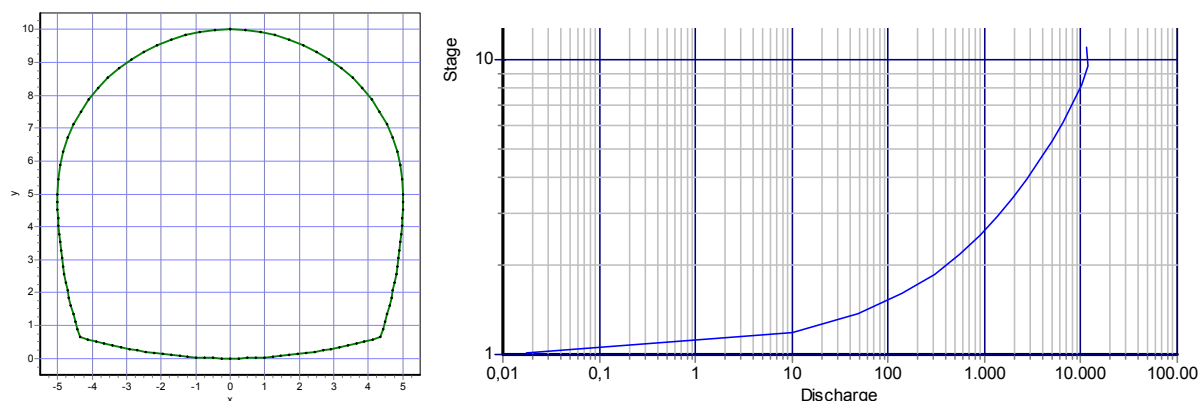
Οι τομές από τις οποίες προκύπτουν οι κυρτές επιφάνειες υπολογίζονται από την ευθεία  $y = h$  καθώς και από την εξίσωση ευθείας μεταξύ δύο σημείων  $i$  και  $i+1$ :

$$x = x_i + \frac{x_{i+1} - x_i}{y_{i+1} - y_i} (y - y_i) \quad 5.29$$

Ο αριθμός των τομών πρέπει να είναι άρτιος αριθμός ώστε να σχηματίζονται κλειστές επιφάνειες· σε αντίθετη περίπτωση επιστρέφεται στον χρήστη μήνυμα σφάλματος. Η τεθλασμένη γραμμή της διατομής πρέπει να ορίζεται αριστερόστροφα στο καρτεσιανό επίπεδο (π.χ. σε έναν ανοικτό αγωγό τα σημεία να δίνονται από τα αριστερά προς τα δεξιά), επιπλέον δεν πρέπει η γραμμή να τέμνεται με τον εαυτό της. Στο Σχήμα 5.8 παρουσιάζονται μερικά παραδείγματα διατομών που μπορούν να χρησιμοποιηθούν για την εξαγωγή των χαρακτηριστικών της υγρής διατομής καθώς και μερικών προβληματικών περιπτώσεων.



Σχήμα 5.8: Επιτρεπτές και προβληματικές διατομές για τον υπολογισμό των στοιχείων της υγρής διατομής



Σχήμα 5.9: Κατάρτιση καμπύλης στάθμης παροχής στο περιβάλλον του «Υδρογνώμων» για μία κλειστή διατομή. Χρήση σχέσης του Manning με  $n=0.0015$ ,  $i=0.012$

### Προσαρμογή των συντελεστών με χρήση της ομογενούς ευθείας

Οι συντελεστές για τις σχέσεις των Manning και Chezy μπορούν να τεθούν ρητά από τον χρήστη αν έχει στοιχεία π.χ. για την τραχύτητα της διατομής  $\eta$  για την υδραυλική κλίση  $i$  (για την οποία θα γίνεται η παραδοχή πως ακολουθεί την κλίση του πυθμένα). Εφόσον όμως οι σχέσεις χρησιμοποιούνται για την κατάρτιση καμπυλών επέκτασης, είναι δυνατόν να εκτιμηθούν με βέλτιστη προσαρμογή των σχέσεων στα δεδομένα των υδρομετρήσεων. Επιλέγουμε έναν ικανό αριθμό των υδρομετρήσεων από τις μεγαλύτερες τιμές της παροχής που έχουν παρουσιαστεί για όλες τις χρονικές περιόδους, αφού έχουμε ταξινομήσει τις υδρομετρήσεις κατά αύξουσα σειρά. Η προρυθμισμένη τιμή στο σύστημα του «Υδρογνώμων» για αυτόν τον αριθμό υδρομετρήσεων είναι  $\boxed{n=8}$ .

Έστω  $(h_i, Q_i)$  το δείγμα των υδρομετρήσεων που θα χρησιμοποιηθεί για την προσαρμογή των συντελεστών. Οι υδραυλικές σχέσεις έχουν την μορφή:

$$Q = G f(h) \quad 5.30$$

όπου  $f(h)$  μία σχέση που περιλαμβάνει όρους επιφανείας ( $A$ ) και υδραυλικής ακτίνας ( $R$ ), η οποία εξαρτάται τελικά από τα γεωμετρικά χαρακτηριστικά της διατομής καθώς και το μέγεθος της στάθμης  $h$  και  $G$  ένας συντελεστής ιδιοτήτων ο οποίος περιλαμβάνει τους όρους τραχύτητας και κλίσης. Για την σχέση του Manning είναι:

$$G = \frac{\sqrt{i}}{n} \text{ και } f(h) = AR^{2/3} \quad 5.31$$

Για την σχέση του Chezy αντίστοιχα:

$$G = C A \sqrt{i} \text{ και } f(h) = \sqrt{R} \quad 5.32$$

Η συνάρτηση  $f(h)$  μπορεί να υπολογιστεί από την μέτρηση της στάθμης καθώς και από τα στοιχεία της διατομής, σύμφωνα με αυτά που αναφέραμε στην προηγούμενη παράγραφο. Αν θεωρηθούν λοιπόν οι τιμές των  $f(h_i)$  ως τιμές μίας ανεξάρτητης μεταβλητής και οι μετρήσεις της παροχής  $Q_i$  ως τιμές της εξαρτημένης μεταβλητής, τότε είναι δυνατό να υπολογίσουμε τον συντελεστή ιδιοτήτων  $G$  με την μέθοδο της γραμμικής παλινδρόμησης – περίπτωση ομογενούς ευθείας (βλ. παράγραφο 7.2.3). Η τιμή του συντελεστή  $G$  δίδεται τελικά από την σχέση:

$$G = \frac{\sum_{i=1}^n f(h_i) Q_i}{\sum_{i=1}^n f(h_i)^2} \quad 5.33$$

Το μέτρο προσαρμογής του μοντέλου στις υδρομετρήσεις θα είναι ο συντελεστής προσδιορισμού ο οποίος υπολογίζεται από την γενικευμένη Σχέση (7.9). Στην συνέχεια ο χρήστης μπορεί αν το επιθυμεί να υπολογίσει κάποιο από τα  $i$ ,  $C$  ή  $n$  εφόσον είναι γνωστό κάποιο άλλο μέγεθος από αυτά (π.χ. να υπολογίσει το μέγεθος της τραχύτητας αν είναι γνωστή η κλίση).



## 6 Μοντέλα εκτίμησης της εξατμοδιαπνοής

---

### 6.1 Εισαγωγή

Η εκτίμηση της εξατμοδιαπνοής καθώς και της εξάτμισης υδάτινων όγκων έχει σημαντικό ρόλο στην κατάρτιση των υδατικών ισοζυγίων (βλ. Κεφάλαιο 8), στην εκτίμηση των υδατικών αναγκών καθώς και σε άλλες εφαρμογές υδρολογικού ενδιαφέροντος. Οι μεθοδολογίες που εξετάζονται αφορούν τον έμμεσο υπολογισμό με την χρήση υδρομετεωρολογικών μεταβλητών σε μηνιαίο χρονικό βήμα με κάποιες δυνατότητες εφαρμογής και σε ημερήσιο χρονικό βήμα. Οι μεθοδολογίες που υλοποιούνται στον «Υδρογνώμων» ταξινομημένες βάσει την απαίτηση σε αριθμό χρονοσειρών είναι οι εξής:

- Αναλυτικά μοντέλα (Penman για την εκτίμηση της εξάτμισης / Penman – Monteith για την εκτίμηση της εξατμοδιαπνοής) που απαιτούν τέσσερις υδρομετεωρολογικές μεταβλητές (μέση θερμοκρασία, σχετική υγρασία, ταχύτητα ανέμου, διάρκεια ηλιοφάνειας)
- Εμπειρικά μοντέλα για την εκτίμηση της εξατμοδιαπνοής (Thornthwaite, Hargreaves και Blaney-Criddle) που απαιτούν την ύπαρξη της χρονοσειρά θερμοκρασίας
- Ειδική μεθοδολογία για την επέκταση και συμπλήρωση χρονοσειρών εξάτμισης ή εξατμοδιαπνοής που απαιτεί την ύπαρξη της χρονοσειράς θερμοκρασίας καθώς και ένα ικανό δείγμα υπολογισμένων τιμών με την μέθοδο του Penman ή των Penman – Monteith.

### 6.2 Τα μοντέλα Penman / Penman - Monteith

Το μοντέλο του Penman για την εκτίμηση της εξάτμισης από υδάτινη επιφάνεια καθώς και η τροποποίηση του μοντέλου από τον Monteith (μοντέλο Penman – Monteith) για την εκτίμηση της δυναμικής εξατμοδιαπνοής βασίζονται σε σχέσεις της φυσικής και δίνουν την καλύτερη ακρίβεια σχετικά με οποιαδήποτε άλλη μεθοδολογία. Το μειονέκτημα αυτών των μεθόδων είναι η απαίτηση τους ως προς τον αριθμό των υδρομετεωρολογικών μεταβλητών. Συγκεκριμένα απαιτούνται χρονοσειρές για:

- Μέση θερμοκρασία  $T_a$ , °C
- Σχετική υγρασία  $U$ , %
- Σχετική ηλιοφάνεια  $n/N$ , %
- Ταχύτητα ανέμου  $u$ , m/s.

Οι χρονοσειρές αυτές έχουν μηνιαίο χρονικό βήμα, είναι δυνατή εντούτοις η χρήση χρονοσειρών ημερήσιου χρονικού βήματος για την εκτίμηση της ημερήσιας εξάτμισης ή εξατμοδιαπνοής, με μειωμένη όμως αξιοπιστία στα αποτελέσματα. Ως προς την σχετική ηλιοφάνεια όπως θα δούμε, ο χρήστης έχει εναλλακτικά την ευχέρεια εισαγωγής της διάρκειας ηλιοφάνειας σε λεπτά ή το μετρημένο μέγεθος της καθαρής ακτινοβολίας βραχέων κυμάτων  $S_{n,m}$ .

#### 6.2.1 Υπολογισμοί φυσικών μεταβλητών

Προκειμένου να υπολογιστούν τα μεγέθη της εξάτμισης ή της εξατμοδιαπνοής, απαιτούνται κάποια φυσικά μεγέθη τα οποία υπολογίζονται με την σειρά που παρουσιάζεται στην συνέχεια.

### Πίεση κορεσμού υδρατμών $e_*$ , hPa

Για την δεδομένη μέση θερμοκρασία του αέρα υπολογίζεται ως:

$$e_* = 6.11 e^{\frac{17.27T_a}{T_a+237.3}} \quad 6.1$$

### Κλίση καμπύλης πίεσης κορεσμού υδρατμών $\Delta$ , hPa/K

Για την δεδομένη μέση θερμοκρασία του αέρα και την ήδη υπολογισμένη πίεση κορεσμού υδρατμών  $e_*$  υπολογίζεται ως:

$$\Delta = \frac{4098 e_*}{(T + 237.2)^2} \quad 6.2$$

### Λανθάνουσα θερμότητα εξάτμισης $\lambda$ , kJ/kg

Ο υπολογισμός του  $\lambda$  απαιτεί την θερμοκρασία  $T_s$  στην επιφάνεια του νερού σε °C. Κάνουμε την παραδοχή πως ισούται με την θερμοκρασία του αέρα ( $T_s=T_a$ ). Η σχέση που δίνει το  $\lambda$  είναι:

$$\lambda = 2501 - 2.361 T_s \quad 6.3$$

### Ψυχομετρικός συντελεστής $\gamma$ , hPa/K

Υπολογίζεται από την σχέση:

$$\gamma = \frac{c_p p}{\varepsilon \lambda} \quad 6.4$$

όπου το  $\lambda$  υπολογίζεται ως άνω,  $c_p$  είναι ειδική θερμότητα αέρα για σταθερή πίεση και λαμβάνεται η τυπική τιμή  $c_p=1.013$  kJ/kg/°C,  $\varepsilon$  ο λόγος μοριακών βαρών νερού και ξηρού αέρα και λαμβάνεται η τιμή  $\varepsilon=0.622$  και  $p$  η ατμοσφαιρική πίεση όπου χρησιμοποιείται μία μέση τιμή συναρτήσει του υψόμετρου από την σχέση:

$$p = 1013.25 (1 - 2.256 \times 10^{-5} z)^{5.256} \quad 6.5$$

όπου,  $z$  το υψόμετρο σε m.

### Αριθμός ημέρας $J$

Για την περίπτωση υπολογισμού σε ημερήσιο χρονικό βήμα λαμβάνεται απλά ο αριθμός της ημέρας στο έτος από 1 για την πρώτη Ιανουαρίου έως 365 για την 31<sup>η</sup> Δεκεμβρίου. Για τους υπολογισμούς σε μηνιαίο χρονικό βήμα λαμβάνεται μία αντιπροσωπευτική τιμή για κάθε μήνα, η οποία δίνεται από την σχέση:

$$J = J_0 + (\mu + 1) \% 2 - 1 \quad 6.6$$

όπου  $\mu$  ο αριθμός των ημερών του μήνα,  $J_0$  ο αριθμός ημέρας της πρώτης ημέρας του μήνα και % το σύμβολο της ακεραίας διαίρεσης.

### Ηλιακή απόκλιση $\delta$ , rad

Δίνει το γεωγραφικό πλάτος στο οποίο πέφτουν κάθετα οι ακτίνες του ηλίου κατά τη μεσουράνησή του:

$$\delta = 0.4093 \cos \left( \frac{2\pi}{365} J - 2.98 \right) \quad 6.7$$

### Γωνία ώρας δύσης του ηλίου $\omega_s$ , rad

Χρησιμοποιώντας το υπολογισμένο  $\delta$  και το γεωγραφικό πλάτος της περιοχής  $\phi$ , η γωνία ώρας δύσης του ηλίου δίνεται από την σχέση:

$$\omega_s = \cos^{-1}(-\tan \phi \tan \delta) \quad 6.8$$

Η χρήση αυτού του μεγέθους περιορίζει τους υπολογισμούς σε γεωγραφικά πλάτη  $|\phi| < 66.5^\circ$ .

### Αστρονομική διάρκεια της ημέρας $N$ , h

Έχοντας υπολογίσει το  $\omega_s$ , η αστρονομική διάρκεια της ημέρας σε ώρες δίνεται από την σχέση:

$$N = \frac{24}{\pi} \omega_s \quad 6.9$$

### Εξωγήινη ηλιακή ακτινοβολία $S_0$ , kJ/m<sup>2</sup>/d

Έχοντας υπολογίσει τα  $\omega_s$ ,  $\delta$ , η εξωγήινη ηλιακή ακτινοβολία δίνεται από την σχέση:

$$S_0 = \frac{I_s t_d d_r}{\pi} (\omega_s \sin \phi \sin \delta + \sin \omega_s \cos \phi \cos \delta) \quad 6.10$$

όπου  $t_d$  η μέση διάρκεια ημέρας η οποία λαμβάνεται  $t_d=86400$  s,  $I_s$  ηλιακή σταθερά η οποία λαμβάνεται  $I_s=1.367$  kW/m<sup>2</sup> και  $d_r$  η εκκεντρότητα η οποία δίνεται από την σχέση:

$$d_r = 1 + 0.033 \cos\left(\frac{2\pi J}{365}\right) \quad 6.11$$

### Διάρκεια ηλιοφάνειας $n$ , h

Ο χρήστης έχει μία από τις εξής δυνατότητες ως προς την εισαγωγή αντιπροσωπευτικού μεγέθους για την ηλιοφάνεια:

- Εισαγωγή του ποσοστού ηλιοφάνειας  $n/N$
- Εισαγωγή της διάρκειας ηλιοφάνειας σε συνολικά λεπτά ηλιοφάνειας (min) (εφόσον η τιμή της διάρκειας ηλιοφάνειας αντιπροσωπεύει μήνα, θα πρέπει να διαιρεθεί με το πλήθος των ημερών του μήνα  $\mu$ ).
- Εισαγωγή της μετρημένης ακτινοβολίας  $S_{n,m}$ .

Εφόσον το ζητούμενο είναι ο υπολογισμός της διάρκειας ηλιοφάνειας σε ώρες (h) κάθε μία από τις τρεις περιπτώσεις αντιμετωπίζεται ως εξής:

- Πολλαπλασιασμός του ποσοστού με την αστρονομική διάρκεια της ημέρας  $N$
- Διαίρεση της διάρκειας σε λεπτά με το 60 ώστε να προκύψει σε ώρες
- Εφαρμογή της Σχέσης 6.12.

$$n = \frac{N}{b_s} \left( \frac{S_{n,m}}{S_0} - a_s \right) \quad 6.12$$

όπου οι δείκτες  $a_s$  και  $b_s$  εξετάζονται στην συνέχεια.

### Συντελεστής απορρόφησης της ατμόσφαιρας (Prescott) $f_s$

Ο συντελεστής δίνεται από την σχέση:

$$f_s = a_s + b_s n/N \quad 6.13$$

Για τα  $a_s$ ,  $b_s$  λαμβάνονται οι τυπικές τιμές  $a_s=0.25$ ,  $b_s=0.50$ . Για την διάρκεια ηλιοφάνειας, χρησιμοποιείται η υπολογισμένη τιμή της προηγούμενης παραγράφου.

### Καθαρή ακτινοβολία βραχέων κυμάτων $S_n$ , kJ/m<sup>2</sup>/d

Υπολογίζεται από την σχέση:

$$S_n = (1-a)f_s S_0 \quad 6.14$$

Εφόσον η διάρκεια ηλιοφάνειας υπολογίζεται μέσω της μετρημένης  $S_n$ , τότε η παραπάνω τιμή θα διαφέρει από την μετρημένη τιμή  $S_{n,m}$  κατά το γινόμενο με  $(1-a)$  καθώς η Σχέση 6.12 προκύπτει λύνοντας ως προς  $n$  την σχέση Prescott και αντικαθιστώντας το  $f_s$  με την τιμή που προκύπτει αν λύσουμε ως προς αυτήν την μεταβλητή την σχέση του  $S_n$  και θεωρώντας μηδενική τιμή albedo (εφόσον κάποια μετρητική διάταξη ακτινοβολίας έχει πρακτικά μηδενική ανακλαστικότητα).

Ο συντελεστής  $a$  λέγεται Λευκαύγεια (albedo) και είναι ο λόγος της ανακλώμενης προς την προσπίπτουσα ακτινοβολία. Θα εισάγεται από τον χρήστη, ωστόσο θα πρέπει να παρέχονται προρυθμισμένες οι εξής τυπικές τιμές. Για την εφαρμογή της μεθόδου του Penman όπου υπολογίζεται το ύψος της εξάτμισης θα λαμβάνεται ως τυπική τιμή  $a=0.08$ , για την εφαρμογή της μεθόδου των Penman-Monteith, θα λαμβάνεται ως τυπική τιμή  $a=0.25$  που είναι μία τυπική τιμή για καλλιέργειες.

### Συντελεστής επίδρασης της νέφωσης $f_L$

Δίνεται από την εξής σχέση:

$$f_L = a_L + b_L n/N \quad 6.15$$

όπου για τα  $a_L$ ,  $b_L$  θα λαμβάνονται οι τυπικές τιμές  $a_L=0.10$  και  $b_L=0.90$ .

### Καθαρή ικανότητα εκπομπής $\varepsilon_n$

Δίνεται από τον τύπο του Brunt:

$$\varepsilon_n = a_e - b_e \sqrt{e} \quad 6.16$$

όπου  $e$  η τάση των ατμών σε hPa και δίνεται από την σχέση:

$$e = U e^* \quad 6.17$$

όπου  $U$  το ποσοστό της σχετικής υγρασίας.

Οι αδιάστατοι παράμετροι  $a_e$  και  $b_e$  λαμβάνονται  $a_e=0.56$  και  $b_e=0.08$  για την μέθοδο του Penman και  $a_e=0.34$  και  $b_e=0.044$  για την μέθοδο Penman – Monteith.

### Καθαρή ακτινοβολία μακρών κυμάτων $L_n$ , kJ/m<sup>2</sup>/d

Δίνεται από την σχέση:

$$L_n = \varepsilon_n f_L \sigma (T_a + 273)^4 \quad 6.18$$

όπου  $\sigma$  η σταθερά Stefan-Boltzmann  $\sigma=4.9 \times 10^{-6}$  kJ/(m<sup>2</sup>K<sup>4</sup>d) και  $T_a$  η μέση θερμοκρασία του αέρα σε °C.

### Ολική καθαρή ενέργεια ακτινοβολίας στην επιφάνεια της γης $R_n$ , kJ/m<sup>2</sup>/d

Είναι η διαφορά της καθαρής ακτινοβολίας βραχέων κυμάτων και της καθαρής ακτινοβολίας μακρών κυμάτων:

$$R_n = S_n - L_n \quad 6.19$$

### 6.2.2 Το μοντέλο του Penman

Η εξάτμιση κατά το μοντέλο του Penman μπορεί να υπολογιστεί από τον συνδυασμό της εκτίμησης της εξάτμισης λόγω μεταφοράς και της εκτίμησης της εξάτμισης από το ενεργειακό ισοζύγιο. Τελικά το ύψος εξάτμισης ανά ημέρα (mm/d) δίνεται από την μία και μοναδική σχέση:

$$E' = \frac{\Delta}{\Delta + \gamma} \frac{R_n}{\lambda} + \frac{\gamma}{\Delta + \gamma} F(u) D \quad 6.20$$

Αν το χρονικό βήμα υπολογισμού είναι ο μήνας, τότε η παραπάνω σχέση πρέπει να πολλαπλασιαστεί με τον αριθμό των ημερών του μήνα για να δώσει την συνολική μηνιαία εξάτμιση. Τα μεγέθη της εξίσωσης του Penman υπολογίζονται σύμφωνα με τις προηγούμενες παραγράφους. Εισάγεται επιπλέον η συνάρτηση ανέμου  $F(u)$  καθώς και το έλλειμμα κορεσμού  $D$  που δίνονται από τις παρακάτω σχέσεις:

$$F(u) = 0.13 + 0.140 u \quad 6.21$$

όπου  $u$  η ταχύτητα του ανέμου σε m/s.

$$D = e^* (1 - U) \quad 6.22$$

όπου  $U$  η το ποσοστό της σχετικής υγρασίας και  $e^*$  η πίεση κορεσμού των υδρατμών.

### 6.2.3 Το μοντέλο των Penman – Monteith

Το μοντέλο των Penman – Monteith είναι μία τροποποίηση της πρωτότυπης σχέσης του Penman ώστε να είναι δυνατή η εκτίμηση του ύψους της δυνητικής εξατμοδιαπνοής. Η δυνητική εξατμοδιαπνοή (ύψος ανά ημέρα – mm/d) δίνεται από την σχέση:

$$E' = \frac{\Delta}{\Delta + \gamma'} \frac{R_n}{\lambda} + \frac{\gamma'}{\Delta + \gamma'} F(u) D \quad 6.23$$

Η συνάρτηση ανέμου  $F(u)$  δίνεται από την σχέση:

$$F(u) = \frac{90}{273 + T_a} u \quad 6.24$$

όπου  $u$  η ταχύτητα του ανέμου σε m/s και  $T_a$  η μέση θερμοκρασία σε °C. Ο ανηγμένος ψυχομετρικός συντελεστής  $\gamma'$  δίνεται από την σχέση:

$$\gamma' = \gamma (1 + 0.33 u) \quad 6.25$$

όπου  $\gamma$  η τιμή του ψυχομετρικού συντελεστή για τους υδρατμούς και  $u$  η ταχύτητα του ανέμου σε m/s.

### 6.2.4 Αριθμητική εφαρμογή – έλεγχος αποτελεσμάτων

Για να γίνει έλεγχος στην αξιοπιστία των αλγορίθμων για τους υπολογισμούς σύμφωνα με τα μοντέλα των Penman και Penman – Monteith, επιλύονται οι εφαρμογές των Κουτσογιάννη και Ξανθόπουλου (1999, σ. 210). Το γεωγραφικό πλάτος  $\varphi$  είναι  $38^\circ 45'$  ενώ το υψόμετρο  $H$  είναι 145 m.

Πίνακας 6.1: Συγκρίσεις μεταξύ των αποτελεσμάτων των μεθόδων Penman, Penman Monteith. Δεδομένα από Κουτσογιάννη και Ξανθόπουλο (1999, σ. 210).

Μήνας	$T_a$ °C	$U$ %	$u$ m/s	$n/N$	Penman (Υπολογ.)	Penman (Πρωτ.)	Penman – Monteith (Υπολογ.)	Penman – Monteith (Πρωτ.)
Οκτ	18.7	68	1.6	0.56	74.59	74.6	69.56	69.4
Νοε	13.3	69	1.6	0.70	36.58	36.6	41.58	41.4
Δεκ	8.5	72	2.0	0.44	25.79	26.3	30.96	31.1
Ιαν	10.4	74	3.7	0.33	43.30	43.7	44.74	44.7
Φεβ	9.8	70	3.4	0.34	51.34	51.9	48.6	48.6
Μαρ	7.1	67	3.3	0.45	73.18	73.2	61.89	61.6
Απρ	11.3	64	3.3	0.55	110.21	110.3	89.99	89.7
Μαι	17.6	64	2.4	0.54	147.35	147.0	121.89	121.3
Ιουν	22.3	58	2.0	0.72	185.41	185.0	156.3	155.6
Ιουλ	27.3	57	3.3	0.74	236.07	235.8	203.86	203.0
Αυγ	26.0	61	2.1	0.80	193.01	193.1	164.25	163.9
Σεπ	25.8	62	2.4	0.70	148.67	148.8	132.75	132.4
Συν.					<b>1325.50</b>	<b>1326.3</b>	<b>1166.37</b>	<b>1162</b>

Οι διαφορές μεταξύ των υπολογισμένων τιμών καθώς και του πρότυπου υπολογισμού από το βιβλίο των Κουτσογιάννη και Ξανθόπουλου (1999) διαφέρουν πολύ λίγο (τάξη μεγέθους 0.1%), διαφορές που οφείλονται μεταξύ άλλων και σε ελαφρά τροποποιημένες σχέσεις.

### 6.3 Τα μοντέλα Thornthwaite, Hargreaves και Blaney-Criddle

Τα μοντέλα που εξετάζονται σε αυτήν την ενότητα είναι φειδωλά ως προς τον αριθμό χρονοσειρών που απαιτούν, αρκεί η ύπαρξη χρονοσειράς θερμοκρασίας. Τα μοντέλα αυτά βασίζονται περισσότερο σε εμπειρικές θεωρήσεις παρά στην μοντελοποίηση των φυσικών μηχανισμών. Θα εξεταστούν οι μεθοδολογίες των Thornthwaite, Hargreaves και των Blaney-Criddle (Κουτσογιάννης και Ξανθόπουλος, 1999, σ. 225) οι οποίες και υλοποιούνται στον «Υδρογνώμων». Από αυτές, οι μέθοδοι των Thornthwaite και των Blaney-Criddle υπολογίζουν αποκλειστικά μηνιαία μεγέθη, ενώ αυτή του Hargreaves μπορεί να εφαρμοστεί και για την εκτίμηση της ημερήσιας εξατμοδιαπνοής εφόσον είναι διαθέσιμη ημερήσια χρονοσειρά μέσης θερμοκρασίας.

#### 6.3.1 Η μέθοδος Thornthwaite

Είναι η πλέον απλή ως προς την υπολογιστική διαδικασία μέθοδος η οποία όμως έχει τύχει ευρύτατης εφαρμογής αν και τείνει να ξεπεραστεί (χρονολογείται από το 1948). Η δυνητική εξατμοδιαπνοή κάποιου μήνα υπολογίζεται από την σχέση:

$$E_p = 16 \left( \frac{10 T_a}{I} \right)^a \frac{\mu N}{360} \quad 6.26$$

όπου  $E_p$  η δυνητική εξατμοδιαπνοή σε mm/μήνα,  $T_a$  η μέση θερμοκρασία του μήνα σε °C,  $\mu$  ο αριθμός των ημερών του μήνα,  $N$  η μέση αστρονομική διάρκεια της ημέρας (λαμβάνεται η τιμή στην μέση του μήνα),  $I$  εμπειρικός συντελεστής που ονομάζεται ετήσιος θερμικός δείκτης και  $a$  άλλος εμπειρικός συντελεστής, συνάρτηση του  $I$ . Οι εμπειρικοί συντελεστές υπολογίζονται από το δείγμα των μέσων μηνιαίων θερμοκρασιών του έτους ( $T_{aj}$ ,  $j=1..12$ ) ως εξής:

$$I = \sum_{i=1}^{12} i_j \quad 6.27$$

$$i_j = 0.09 T_{a_j}^{3/2} \quad 6.28$$

$$a = 0.016I + 0.5 \quad 6.29$$

Στην περίπτωση που λείπουν δεδομένα ώστε να είναι αδύνατος ο υπολογισμός κάποιων από τα  $i_j$ , γίνεται επισήμανση στον χρήστη και σταματάνε οι υπολογισμοί. Στον Πίνακα 6.2 γίνεται σύγκριση των υπολογισμένων τιμών εξατμοδιαπνοής για το παράδειγμα των Κουτσογιάννη και Ξανθόπουλου (1999, σ. 228). Οι διαφορές είναι μικρές και οφείλονται στην χρήση πίνακα για τον υπολογισμό των  $N$  του λυμένου παραδείγματος.

Πίνακας 6.2: Σύγκριση αποτελεσμάτων για την μέθοδο υπολογισμού της εξατμοδιαπνοής του Thornthwaite. Δεδομένα από Κουτσογιάννη και Ξανθόπουλο (1999, σ. 228).

Μήνας	$T_a$ (°C)	$E_p$ (mm) – πρότυπο	$E_p$ (mm) - υπολογισμένο
Οκτ	18.70	69.9	70.17
Νοε	13.30	33.7	33.69
Δεκ	8.50	15.1	15.07
Ιαν	10.40	22.1	21.97
Φεβ	9.80	19.7	19.62
Μαρ	7.10	13.9	13.81
Απρ	11.30	33.4	33.39
Μαι	17.60	80.9	80.90
Ιουν	22.30	123.0	123.26
Ιουλ	27.30	177.9	178.42
Αυγ	26.00	152.9	153.37
Σεπ	25.80	132.7	132.69
		<b>875.1</b>	<b>876.36</b>

### 6.3.2 Η μέθοδος Hargreaves

Η μέθοδος του Hargreaves χρησιμοποιείται για να εκτιμηθεί η εξατμοδιαπνοή της καλλιέργειας αναφοράς. Η μέθοδος απαιτεί την ύπαρξη χρονοσειράς θερμοκρασίες με μηνιαίο ή και ημερήσιο χρονικό βήμα, η δε τιμή της εξατμοδιαπνοής αναφοράς  $E_{rc}$  σε mm/d υπολογίζεται από την σχέση:

$$E_{rc} = \frac{0.0023 S_0}{\lambda} (T_a + 17.8)(T_{\max} - T_{\min})^{0.5} \quad 6.30$$

όπου  $S_0$  η εξωγήινη ακτινοβολία σε kJ/(m<sup>2</sup>d) η οποία υπολογίζεται στην συγκεκριμένη ημέρα προκειμένου για ημερήσιο υπολογισμό της εξατμοδιαπνοής ή στην μέση του μήνα για μηνιαίο σύμφωνα με την μεθοδολογία της ενότητας 6.2,  $\lambda$  είναι η λανθάνουσα θερμότητα εξατμίσεως σε kJ/kg και θα λαμβάνεται η τιμή  $\lambda=2460$ ,  $T_a$  η μέση θερμοκρασία και  $T_{\max}$ ,  $T_{\min}$ , η μέγιστη και ελάχιστη θερμοκρασία της χρονοσειράς μέσης θερμοκρασίας.

### 6.3.3 Η μέθοδος Blaney-Criddle

Η μέθοδος Blaney-Criddle έχει τύχει ευρύτατης εφαρμογής για την εκτίμηση των αρδευτικών αναγκών. Στον «Υδρογόνωμων» υλοποιείται η αρχική της μορφή που περιγράφεται από τους Κουτσογιάννη και Ξανθόπουλο (1999, σ. 225) που εκτιμά τη μηνιαία δυνητική εξατμοδιαπνοή μίας

καλλιέργειας χρησιμοποιώντας την παρακάτω σχέση η οποία έχει ως μόνη μετεωρολογική μεταβλητή την μέση μηνιαία θερμοκρασία  $T_a$ :

$$E_p = 0.254 k_c p (32 + 1.8 T_a) \quad 6.31$$

όπου  $E_p$  η δυνητική εξατμοδιαπνοή σε mm/μήνα,  $k_c$  ο συντελεστής καταναλωτικής χρήσης (φυτικός συντελεστής) και  $p$  το ποσοστό (%) των ωρών ημέρας του συγκεκριμένου μήνα σε σχέση με το σύνολο των ωρών ημέρας του έτους. Το ποσοστό αυτό υπολογίζεται από τη σχέση:

$$p = 100 \frac{\mu N}{365 \times 12} \quad 6.32$$

όπου  $N$  η μέση αστρονομική διάρκεια της ημέρας στο μέσο του μήνα σε ώρες και  $\mu$  ο αριθμός ημερών του συγκεκριμένου μήνα.

## 6.4 Επέκταση και συμπλήρωση χρονοσειρών εξατμοδιαπνοής

Η μεθοδολογία που παρουσιάζεται έχει προταθεί από τον Κουτσογιάννη, περιγράφεται από τους Κουτσογιάννη και Ξανθόπουλο (1997, σ. 222), έχει δε χρησιμοποιηθεί με επιτυχία στην επέκταση δειγμάτων εξάτμισης από λίμνες (Ευστρατιάδης κ.α. 2000) και αλλού. Υλοποιείται μέσω μίας σχετικά απλής αναλυτικής σχέσης που δίνει την εξάτμιση (ή την εξατμοδιαπνοή) συναρτήσει της θερμοκρασίας  $T_a$  και του μεγέθους της εξωγήινης ακτινοβολίας  $S_0$ , χρησιμοποιώντας τρεις συντελεστές  $a$ ,  $b$  και  $c$ . Οι τιμές αυτών των συντελεστών μπορούν να προκύψουν με διαδικασία βέλτιστης προσαρμογής χρησιμοποιώντας ένα δείγμα υπολογισμένων τιμών της εξάτμισης ή της εξατμοδιαπνοής με την μέθοδο Penman ή Penman-Monteith<sup>1</sup>. Η σχέση που προτείνεται είναι η εξής:

$$E = \frac{aS_0 - b}{1 - cT_a} \quad 6.33$$

όπου  $E$  το ύψος εξάτμισης ανά ημέρα (mm/d),  $T_a$  η μέση θερμοκρασία σε °C,  $a$ ,  $b$  και  $c$  οι παράμετροι προς προσδιορισμό και  $S_0$  η εξωγήινη ακτινοβολία σε kJ/m<sup>2</sup>/d. Η  $S_0$  υπολογίζεται από αναλυτικά στοιχεία και δεν προϋποθέτει μετρήσεις. Για ανάλυση με ημερήσιο χρονικό βήμα η  $S_0$  υπολογίζεται στην τρέχουσα ημέρα ενώ για μηνιαίο χρονικό βήμα στο μέσο του τρέχοντα μήνα (όπως γίνεται και στην μεθοδολογία για τον υπολογισμό με Penman). Τυπικές τιμές για τα  $a$ ,  $b$  και  $c$  που χρησιμοποιούνται και ως προτεινόμενες τιμές στον χρήστη είναι :  $a \approx 10^{-4}$ ,  $b \approx 0.5$  και  $c \approx 0.02$ . Σε καμιά όμως περίπτωση αυτές οι τιμές δεν πρέπει να εφαρμόζονται άκριτα στις εφαρμογές, σαφώς απαιτείται προσαρμογή στις τοπικές συνθήκες χρησιμοποιώντας κάποιο υπολογισμένο δείγμα.

### 6.4.1 Μεθοδολογία προσαρμογής

Εφαρμόζεται η Μέθοδος Ελαχίστων Τετραγώνων, συγκεκριμένα η μέθοδος μεταβολής των μεταβλητών (Μπαλοδήμου, 1991) που βασίζεται σε διαδοχικές προσεγγίσεις. Έστω  $a, b$  και  $c$  οι άγνωστες τιμές των παραμέτρων. Έστω  $a_0$ ,  $b_0$  και  $c_0$  μία αρχική χονδροειδής εκτίμηση, χρησιμοποιώντας π.χ. τις τυπικές τιμές  $a=10^{-4}$ ,  $b=0.5$  και  $c=0.02$ .

<sup>1</sup> Η λογική είναι πως διαθέτουμε πλήρη υδρομετεωρολογικά δεδομένα: χρονοσειρές θερμοκρασίας, σχετικής υγρασίας, ταχύτητας ανέμου και ηλιοφάνειας για μία μικρή μόνο περίοδο, όμως διαθέτουμε μεγάλο μήκος χρονοσειρά θερμοκρασίας (όπου η μέτρησή της είναι απλούστερη) την οποία την αξιοποιούμε αφού προσαρμόσουμε βέλτιστες τιμές στις παραμέτρους  $a, b$  και  $c$  χρησιμοποιώντας το μικρότερο σε μήκος πλήρες δείγμα. Δίνεται ακόμα η δυνατότητα προσαρμογής των παραμέτρων από το μηνιαίο δείγμα, κατόπιν εφαρμογή σε ημερήσιες χρονοσειρές.



Με την μεθοδολογία που θα ακολουθήσουμε, υπολογίζεται ένα διάνυσμα μεταβολής για τον  $i^{\text{th}}$  υπολογιστικό κύκλο:

$$\delta[x_i] = [\delta a_i, \delta b_i, \delta c_i] \quad 6.34$$

Οπότε η τιμή των παραμέτρων στον  $i^{\text{th}}$  υπολογιστικό κύκλο διαμορφώνεται ως:

$$[x_i] = [x_{i-1}] + \delta[x_i] \quad 6.35$$

Θεωρούμε πως έχουμε συγκλίνει στην τελική λύση όταν η ευκλείδεια νόρμα του  $\delta[x_i]$  πέσει κάτω από ένα κατώφλι  $\varepsilon$ :

$$|\delta[x_i]| \leq \varepsilon \quad 6.36$$

οπότε και σταματάει η διαδικασία των υπολογισμών. Το κατώφλι το θέτουμε  $\varepsilon=0.0001$ , επιπλέον δεν θα τερματίζονται οι υπολογιστικοί κύκλοι προτού το  $i>3$  αν δε υπερβούμε τους 10 υπολογιστικούς κύκλους η διαδικασία σταματάει και επιστρέφονται οι τιμές των παραμέτρων. Το  $\delta[x_i]$  υπολογίζεται ως λύση του παρακάτω γραμμικού συστήματος εξισώσεων:

$$[N] \delta[x] = [A]^T [\delta l] \quad 6.37$$

Όπου το  $[N]$  είναι μητρώο  $3 \times 3$ . Το μητρώο  $[A]$  καθώς και το διάνυσμα  $[\delta l]$  εξεταστούν στην συνέχεια. Το μητρώο  $[N]$  υπολογίζεται από την σχέση:

$$[N] = [A]^T [A] \quad 6.38$$

Έστω δείγμα υπολογισμένης εξάτμισης (ή εξατμοδιαπνοής) με την μέθοδο Penman (ή Penman – Monteith)  $E_j$ ,  $j=1 \dots n$  τις οποίες θα τις καλούμε καταχρηστικά αλλά για λόγους διάκρισης ως «μετρημένες τιμές». Έστω  $T_{aj}$ ,  $j=1 \dots n$  οι αντίστοιχες μετρήσεις της μέσης θερμοκρασίας του αέρα. Τότε το μητρώο  $[A]$  θα έχει διαστάσεις  $n$  γραμμές  $\times$  3 στήλες και ορίζεται ως:

$$[A] = \begin{bmatrix} \frac{\partial E(T_{a1})}{\partial a} & \frac{\partial E(T_{a1})}{\partial b} & \frac{\partial E(T_{a1})}{\partial c} \\ \vdots & \vdots & \vdots \\ \frac{\partial E(T_{an})}{\partial a} & \frac{\partial E(T_{an})}{\partial b} & \frac{\partial E(T_{an})}{\partial c} \end{bmatrix} \quad 6.39$$

Όπου  $E(T_{aj})$  η αναλυτική έκφραση της παραμετρικής εξίσωσης της εξάτμισης 6.33. Οι μερικές παράγωγοι ως προς τις παραμέτρους υπολογίζονται αναλυτικά ως:

$$\frac{\partial E(T_{aj})}{\partial a} = \frac{S_0}{1 - c_{i-1} T_{aj}}, \quad \frac{\partial E(T_{aj})}{\partial b} = \frac{-1}{1 - c_{i-1} T_{aj}}, \quad \frac{\partial E(T_{aj})}{\partial c} = \frac{(a_{i-1} S_0 - b_{i-1}) T_{aj}}{(1 - c_{i-1} T_{aj})^2} \quad 6.40$$

όπου  $a_{i-1}$ ,  $b_{i-1}$  και  $c_{i-1}$  οι επιτευχθείσες τιμές των παραμέτρων από τον προηγούμενο υπολογιστικό κύκλο. Το διάνυσμα  $[\delta l]$  διάστασης  $n$ , είναι η διαφορά της υπολογισμένης τιμής (χρησιμοποιώντας τις παραμέτρους του προηγούμενου υπολογιστικού κύκλου) και της μετρημένης τιμής της εξάτμισης και είναι ίσο με:

$$[\delta l] = \left[ E(T_{a1}) - \frac{E_1}{\mu}, \quad \dots, \quad E(T_{an}) - \frac{E_n}{\mu} \right] \quad 6.41$$

όπου  $\mu$  είναι ο αριθμός των ημερών του μήνα που αντιστοιχεί στην  $j$  μέτρηση εφόσον οι υπολογισμοί γίνονται σε μηνιαίο χρονικό βήμα. Αν οι υπολογισμοί γίνονται σε ημερήσιο βήμα τότε τίθεται  $\mu=1$ . Σύμφωνα με τα παραπάνω, ύστερα από λίγες προσεγγίσεις (συνήθως απαιτούνται 3-4 για να επιτευχθεί η επιθυμητή ακρίβεια) υπολογίζονται οι τελικές τιμές των  $a, b$  και  $c$ . Ο συντελεστής προσδιορισμού προκύπτει εύκολα σε κάθε βήμα από το διάνυσμα  $[\delta l]$ . Με αυτήν την παραδοχή ο

συντελεστής αναφέρεται στο  $i-1^{\text{th}}$  βήμα, ωστόσο δεν διαφοροποιείται ιδιαίτερα μετά τον τρίτο κύκλο, έτσι υιοθετούμε την τιμή που προκύπτει για το προτελευταίο βήμα. Το άθροισμα των τετραγώνων των σφαλμάτων είναι:

$$w = | [\delta I] [\delta I]^T | \quad 6.42$$

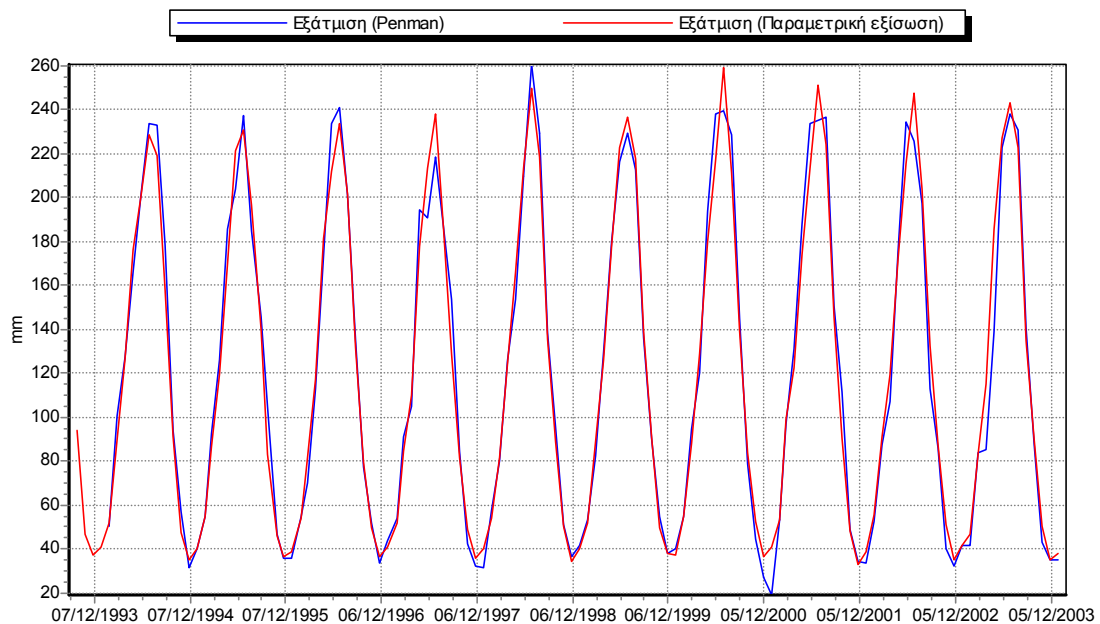
όπου  $||$  η ευκλείδεια νόρμα. Τελικά ο συντελεστής προσδιορισμού προκύπτει από την σχέση:

$$\delta_{i-1} = 1 - \frac{w}{(n-1)\sigma_{Ta}^2} \quad 6.43$$

όπου  $\sigma_{Ta}^2$  η διασπορά του δείγματος θερμοκρασιών για την κοινή περίοδο με τα μετρημένα  $E_j$ .

## 6.4.2 Εφαρμογή

Εξετάζεται η προσαρμογή του μοντέλου επέκτασης σε χρονοσειρά εξάτμισης που έχει υπολογιστεί με την μέθοδο του Penman. Χρησιμοποιούνται τα δεδομένα του μετεωρολογικού σταθμού του ΕΜΠ με ένα δείγμα εξάτμισης που εκτείνεται από τον Φεβρουάριο του 1994 έως τον Ιανουάριο του 2004. Η τοποθεσία του σταθμού βρίσκεται στο γεωγραφικό πλάτος  $\varphi=37^\circ 58' 24''$ . Η προσαρμογή του παραμετρικού μοντέλου απεικονίζεται στο Σχήμα 6.1.



Σχήμα 6.1: Προσαρμογή του παραμετρικού μοντέλου στο δείγμα εξάτμισης του σταθμού του Ε.Μ.Π.

Η προσαρμογή έδωσε ένα συντελεστή προσδιορισμού  $\delta = 0.977$  και τιμές παραμέτρων  $a = 0.958 \times 10^{-4}$ ,  $b = 0.519$ ,  $c = 0.0205$ , τιμές που όπως βλέπουμε δεν διαφέρουν πολύ από τις τυπικές τιμές.

## 7 Συσχέτιση – συμπλήρωση – επέκταση χρονοσειρών

### 7.1 Εισαγωγή

Η στατιστική συσχέτιση μεταξύ διαφορετικών χρονοσειρών μπορεί να μας δώσει χρήσιμα συμπεράσματα σχετικά με τον βαθμό της φυσικής συσχέτισης των δεδομένων (π.χ. λόγω γεωγραφικής θέσης). Επιπλέον, οι μεθοδολογίες της στατιστικής συσχέτισης χρησιμοποιούνται όταν θέλουμε να αντιμετωπίσουμε το πρόβλημα της συμπλήρωσης των ελλειπουσών τιμών των χρονοσειρών ή να επεκτείνουμε κάποιο δείγμα σε προγενέστερη ή μεταγενέστερη περίοδο από την περίοδο μέτρησης.

Οι μεθοδολογίες συσχέτισης (ή παλινδρόμησης) εξετάζονται χωριστά από την πραγματοποίηση της συμπλήρωσης ή της επέκτασης: εφόσον έχουν προσδιοριστεί οι κατάλληλοι συντελεστές με τις παραπάνω μεθοδολογίες στο στάδιο της συσχέτισης, στην συνέχεια χρησιμοποιούνται αυτοί οι ίδιοι συντελεστές για την συμπλήρωση ή την επέκταση. Δύο ακόμα θέματα που εξετάζονται είναι η εποχικότητα των μεταβλητών καθώς και η διατήρηση των στατιστικών χαρακτηριστικών του (προ της συμπλήρωσης – επέκτασης) δείγματος με την χρήση των κατάλληλων τεχνικών.

### 7.2 Μέθοδοι προσδιορισμού των παραμέτρων της παλινδρόμησης

Παρουσιάζονται οι μεθοδολογίες προσδιορισμού των παραμέτρων της παλινδρόμησης που είναι υλοποιημένες στο σύστημα του «Υδρογνώμων». Οι μεθοδολογίες αυτές καλύπτουν τόσο την ανάγκη εύρεσης του βαθμού συσχέτισης των χρονοσειρών όσο και τις ανάγκες για την επέκταση και την συμπλήρωση των χρονοσειρών.

#### 7.2.1 Μέση τιμή

Η μέση τιμή μπορεί να χρησιμοποιηθεί για την συμπλήρωση ελλειπουσών τιμών όταν δεν μπορεί να βρεθεί κάποια συσχέτιση που θα βοηθούσε σε αυτόν τον σκοπό. Η εφαρμογή της μέσης τιμής έχει ως αποτέλεσμα την υποεκτίμηση της τυπικής απόκλισης του συμπληρωμένου δείγματος· για να αντιμετωπιστεί αυτό το πρόβλημα μπορεί να εφαρμοστεί η μεθοδολογία εισαγωγής τυχαίου όρου (βλ. παράγραφο 7.3.1). Η εφαρμογή της μέσης τιμής μπορεί να αναπαρασταθεί σε όρους παλινδρόμησης με μηδενικό συντελεστή  $b$  και τιμή σταθερού όρου  $a$  ίση με την μέση τιμή:

$$\begin{aligned} a &= \bar{x} \\ b &= 0 \end{aligned} \quad 7.1$$

Η μεθοδολογία αυτή μπορεί να εφαρμοστεί κυρίως σε χρονοσειρές με μηνιαίο χρονικό βήμα λαμβάνοντας υπόψη την εποχικότητα, υπολογίζοντας δηλαδή δώδεκα διαφορετικές τιμές για τους σταθερούς όρους  $a$ . (βλ. παράγραφο 7.2.6). Οι τιμές της μέσης τιμής δίνονται στο Κεφάλαιο 9 στην ενότητα 9.2, είναι δε η τετριμμένη τιμή του μέσου όρου (άθροισμα τιμών μεταβλητής δια το πλήθος των τιμών).

#### 7.2.2 Απλή – πολλαπλή γραμμική παλινδρόμηση

Η απλή γραμμική παλινδρόμηση είναι η γνωστή μεθοδολογία βέλτιστης προσαρμογής μίας ευθείας στο καρτεσιανό επίπεδο όπου διατάσσονται διάφορα σημεία. Η λύση επιτυγχάνεται εφαρμόζοντας την

αρχή των ελάχιστων τετραγώνων, δηλαδή δίνεται μία λύση η οποία ελαχιστοποιεί το άθροισμα των τετραγώνων των αποστάσεων των σημείων από την ευθεία. Σύμφωνα με αυτήν την λύση, προσδιορίζεται ένας σταθερός όρος  $a$  και μία κλίση  $b$ , οι όροι δηλαδή της εξίσωσης:

$$y = a + b x \quad 7.2$$

όπου  $x$  θα είναι η ανεξάρτητη μεταβλητή και  $y$  η εξαρτημένη. Στην περίπτωση που χρησιμοποιείται το μοντέλο συσχέτισης για την συμπλήρωση, η μεταβλητή  $x$  θα προέρχεται από την χρονοσειρά βάσης ή αναφοράς και η μεταβλητή  $y$  θα αναφέρεται στην χρονοσειρά που υπάρχουν ελλείψεις. Οι συντελεστές την απλής γραμμικής συσχέτισης δίνονται από τις σχέσεις:

$$b = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - \left( \sum_{i=1}^n x_i \right)^2} \quad 7.3$$

και:

$$a = \bar{y} - b\bar{x} \quad 7.4$$

όπου  $n$  είναι το πλήθος των κοινών (ταυτόχρονων) μετρήσεων μεταξύ της εξαρτημένης μεταβλητής  $y$  και της ανεξάρτητης  $x$  αν μιλάμε για τιμές χρονοσειράς.

Στην περίπτωση ύπαρξης πολλών ανεξάρτητων μεταβλητών και μίας εξαρτημένης, θα αναζητείται μία βέλτιστη λύση για τους συντελεστές  $a_0 \dots a_m$  της εξίσωσης:

$$Y = a_0 + a_1 X_1 + \dots + a_m X_m \quad 7.5$$

όπου  $Y$  η εξαρτημένη μεταβλητή και  $X_1 \dots X_m$ , οι  $m$  ανεξάρτητες μεταβλητές. Η λύση επιτυγχάνεται μέσω της πολλαπλής γραμμικής παλινδρόμησης. Για να πραγματοποιηθεί η παλινδρόμηση, πρέπει το σύνολο των χρονοσειρών που θα χρησιμοποιηθούν, ανεξάρτητες και εξαρτημένη, να έχουν κοινή περίοδο μετρήσεων, δηλαδή σύνολο χρονικών στιγμών για τις οποίες υπάρχουν τιμές σε όλες τις χρονοσειρές. Στο εξής, όποτε λέμε *χρονοσειρά* θα εννοούμε το υποσύνολο της χρονοσειράς για την κοινή περίοδο. Επομένως, όποτε αναφερόμαστε στη διασπορά ή τη μέση τιμή μιας χρονοσειράς, εννοούμε τη διασπορά ή τη μέση τιμή των τιμών της κοινής περιόδου μόνο.

Οι σχέσεις που δίνουν τους συντελεστές  $a_i$  δίνονται από τον Χριστοφίδη (1994, σ. 16). Θεωρούμε τους εξής συμβολισμούς:

$\bar{x}$  : η μέση τιμή της χρονοσειράς  $x$

$s^2(x)$ : η διασπορά της χρονοσειράς  $x$

$s(xy)$ : η συνδιασπορά των χρονοσειρών  $x$  και  $y$

Τότε:

$$a_0 = \bar{y} - \sum_{i=1}^m a_i \bar{x}^{(i)} \quad 7.6$$

$$a_i = \sum_{j=1}^m d^{(ij)} s(x^{(j)} y) \quad i = 1, 2, \dots, m$$

όπου  $d^{(ij)}$  τα στοιχεία του μητρώου:

$$\begin{bmatrix} s^2(x^{(1)}) & s(x^{(1)}x^{(2)}) & \dots & s(x^{(1)}x^{(m)}) \\ s(x^{(2)}x^{(1)}) & s^2(x^{(2)}) & \dots & s(x^{(2)}x^{(m)}) \\ \vdots & & & \\ s(x^{(m)}x^{(1)}) & s(x^{(m)}x^{(2)}) & \dots & s^2(x^{(m)}) \end{bmatrix}^{-1} \quad 7.7$$

Η αμερόληπτη τιμή της συνδιασποράς μεταξύ δύο μεταβλητών  $x$  και  $y$  δίνεται από την σχέση:

$$s_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n-1} \quad 7.8$$

Οι τιμές της διασποράς δίνονται στο Κεφάλαιο 9 στην ενότητα 9.2.

Στον «Υδρογνώμων» υλοποιείται η πολλαπλή γραμμική παλινδρόμηση σύμφωνα με τα παραπάνω. Δεν υπάρχει χωριστή υλοποίηση για την απλή γραμμική παλινδρόμηση καθώς αποτελεί μία υποπερίπτωση. Ο συντελεστής προσδιορισμού σε κάθε περίπτωση υπολογίζεται από την γενικευμένη σχέση:

$$\delta = 1 - \frac{\sum_{i=1}^n W_i^2}{\sum_{i=1}^n y_i^2 - n\bar{y}^2} \quad 7.9$$

όπου  $W_i$  είναι το σφάλμα μεταξύ της εξαρτημένης μεταβλητής και της τιμής από το μοντέλο παλινδρόμησης,  $y_i$  η εξαρτημένη μεταβλητή,  $n$  το πλήθος των τιμών της εξαρτημένης μεταβλητής και  $\bar{y}$  η μέση τιμή της εξαρτημένης μεταβλητής. Η τετραγωνική ρίζα του συντελεστή προσδιορισμού είναι ο συντελεστής συσχέτισης  $r$ .

### 7.2.3 Ομογενής ευθεία

Σε πολλές περιπτώσεις είναι επιθυμητό να τίθεται ο περιορισμός  $a_0 = 0$ . Σ' αυτή την περίπτωση οι υπόλοιποι συντελεστές δίνονται από τον ίδιο τύπο όπως προηγουμένως:

$$a_i = \sum_{j=1}^m d^{(ij)} s(x^{(j)}y) \quad \text{για } i = 1, \dots, m \quad 7.10$$

με τη διαφορά ότι το  $s(x^{(j)}y)$ , και σ' αυτό τον τύπο και στο μητρώο (7.7), που δίνει τα  $d^{(ij)}$ , δεν είναι η συνδιασπορά αλλά το άθροισμα των γινομένων των τιμών των χρονοσειρών (Χριστοφίδης, 1994, σ. 18).

Στην περίπτωση μίας ανεξάρτητης μεταβλητής η κλίση της ευθείας που περνάει από την αρχή των αξόνων δίνεται από την σχέση (Κουτσογιάννης, 1997):

$$b = \frac{\sum_{i=1}^{n_a} x_i y_i}{\sum_{i=1}^{n_a} x_i^2} \quad 7.11$$

Ο συντελεστής προσδιορισμού θα υπολογίζεται από την γενικευμένη Σχέση (7.9).

#### 7.2.4 Οργανική συσχέτιση

Η οργανική συσχέτιση εφαρμόζεται κυρίως στην επέκταση δειγμάτων όπου είναι επιθυμητή η διατήρηση των στατιστικών χαρακτηριστικών του αρχικού (προ της συμπλήρωσης) δείγματος. Εγκαταλείποντας την απαίτηση ελαχιστοποίησης του μέσου τετραγωνικού σφάλματος, εκτιμώνται οι παράμετροι της συσχέτισης διατηρώντας την αρχική μέση τιμή και την διασπορά. Αυτός ο τρόπος εκτίμησης των παραμέτρων είναι γνωστός ως οργανική συσχέτιση ή και ως επέκταση διατήρησης διασποράς (Maintenance Of Variance Extension – MOVE.1). (Κουτσογιάννης, 1997, Χριστοφίδης, 1994).

Η εφαρμογή της οργανικής συσχέτισης γίνεται για μία ανεξάρτητη μεταβλητή  $x$ . Θα πρέπει να προσδιοριστούν οι συντελεστές της εξίσωσης:

$$y = a + b x \quad 7.12$$

$a$  και  $b$  έτσι ώστε να διατηρηθούν τα στατιστικά χαρακτηριστικά του αρχικού δείγματος. Οι συντελεστές δίνονται από τις σχέσεις (Κουτσογιάννης, 1997, σ. 202):

$$b = \text{sgn}(r_{xy}) \frac{s_y}{s_x} \quad 7.13$$

$$a = \bar{y} - b\bar{x} \quad 7.14$$

όπου  $s_x$ ,  $s_y$  οι δειγματικές τιμές της τυπικής απόκλισης των μεταβλητών  $x$ ,  $y$  και  $r_{xy}$  ο δειγματικός συντελεστής συσχέτισης:

$$r_{xy} = \frac{s_{xy}}{s_x s_y} \quad 7.15$$

όπου  $s_{xy}$  η δειγματική συνδιασπορά των μεταβλητών  $x$  και  $y$  από την οποία χρησιμοποιείται το πρόσημό της. Ο συντελεστής προσδιορισμού θα υπολογίζεται την Σχέση (7.21) (στο ίδιο αποτέλεσμα καταλήγουμε φυσικά, αν χρησιμοποιήσουμε την γενικευμένη Σχέση (7.9)).

#### 7.2.5 Αυτοσυσχέτιση

Οι υδρολογικές χρονοσειρές συχνά παρουσιάζουν αυτοσυσχέτιση· υπάρχει, δηλαδή, σχέση μεταξύ κάθε τιμής και προηγούμενων της (Χριστοφίδης 1994, Χριστοφίδης, 1998), δηλαδή:

$$y_i = a_0 + a_1 y_{i-1} + a_2 y_{i-2} + \dots + a_k y_{i-k} + \varepsilon \quad 7.16$$

όπου  $\varepsilon$  τυχαίος όρος που ακολουθεί κάποια κατανομή. Αυτή η σχέση, όπου κάθε τιμή σχετίζεται με τις  $k$  προηγούμενες της, λέγεται «ομοίωμα  $AR(k)$ ». Το  $k$  λέγεται «τάξη» του ομοιώματος. Ο προσδιορισμός των συντελεστών  $a_i$  γίνεται ακριβώς όπως στη γραμμική παλινδρόμηση, θεωρώντας ως ανεξάρτητες χρονοσειρές τις μετατοπίσεις της  $y$  κατά 1, 2, ...,  $k$  θέσεις (συσχέτιση της χρονοσειράς με τον εαυτό της  $k$  φορές, με χρονική υστέρηση κατά 1, 2, ...  $k$ ). Σε περίπτωση που υπάρχει εποχικότητα, οπότε γίνονται 12 ανεξάρτητες συσχετίσεις (π.χ., σε μηνιαία χρονοσειρά, οι μήνες Ιανουαρίου με τους μήνες Δεκεμβρίου, οι μήνες Φεβρουαρίου με τους μήνες Ιανουαρίου κλπ.), το μοντέλο λέγεται  $PAR(k)$ .

Ενίοτε γίνεται συνδυασμός αυτοσυσχέτισης με ετεροσυσχέτιση, ως εξής:

$$y_i = a_0 + a_1 y_{i-1} + a_2 y_{i-2} + \dots + a_k y_{i-k} + a_{k+1} x_i^{(1)} + a_{k+2} x_i^{(2)} + \dots + a_{k+m} x_i^{(m)} \quad 7.17$$

Εδώ οι ανεξάρτητες μεταβλητές είναι συνολικά  $k+m$ . Είναι οι  $k$  μετατοπίσεις της χρονοσειράς  $y$  συν τις υπόλοιπες  $m$  χρονοσειρές.

Στην υλοποίηση του «Υδρογνώμων» εισάγονται οι  $m$  χρονοσειρές ανεξάρτητων μεταβλητών της ετεροσυσχέτισης, επιπλέον εισάγονται  $k$  χρονοσειρές ως ανεξάρτητες μεταβλητές της αυτοσυσχέτισης, όπου είναι η χρονοσειρά της εξαρτημένης μεταβλητής μετατοπισμένη κατά  $1, 2, \dots, k$  χρονικές μονάδες. Παράδειγμα χρήσης της αυτοσυσχέτισης για συμπλήρωση ελλείψεων χρονοσειρών είναι οι μηνιαίες τιμές της παροχής ποταμού (βλ. Ευστρατιάδης, κ.α., 2000) όπου έτσι και αλλιώς οι τιμές έχουν ισχυρή στατιστική αυτοσυσχέτιση.

### 7.2.6 Εποχικότητα

Αρκετές από τις υδρομετεωρολογικές μεταβλητές παρουσιάζουν το φαινόμενο της εποχικότητας, σύμφωνα με το οποίο παρουσιάζεται κάποια περιοδικότητα στις τιμές που συσχετίζεται με την εναλλαγή των εποχών του έτους. Το φαινόμενο της εποχικότητας προσομοιώνεται με την θεώρηση δώδεκα διαφορετικών ομάδων που αντιστοιχούν στους δώδεκα μήνες του έτους. Έτσι για κάθε ομάδα σχηματίζεται ένα ανεξάρτητο δείγμα τιμών που προέρχονται από τον αντίστοιχο μήνα και προσδιορίζονται στην συνέχεια τιμές παραμέτρων παλινδρόμησης. Συνεπώς όλες οι μεθοδολογίες που παρουσιάστηκαν, θα εκτελούνται δώδεκα φορές για τα δώδεκα δείγματα που έχουν σχηματιστεί από τον διαχωρισμό των τιμών σύμφωνα με τους μήνες.

Τα αποτελέσματα της παλινδρόμησης θα παρουσιάζονται σε ένα μητρώο με δώδεκα γραμμές που θα αντιστοιχούν στους μήνες του έτους, όταν δε διενεργείται συμπλήρωση θα χρησιμοποιούνται οι κατάλληλες τιμές παραμέτρων που θα αντιστοιχούν στον μήνα της ελλείπουσας τιμής.

Για την συσχέτιση χρονοσειρών ετησίου χρονικού βήματος ο χρήστης αποτρέπεται από το σύστημα για διενέργεια εποχικής συσχέτισης. Για την περίπτωση όλων των χρονοσειρών με μικρότερο χρονικό βήμα, θα υπολογίζονται μηνιαίοι συντελεστές σύμφωνα με τις ημερομηνίες των τιμών των χρονοσειρών. Υπολογισμός πλέον των μηνιαίων συντελεστών (π.χ. ημερησίων) δεν έχει νόημα καθώς πέρα από την εποχικότητα εισάγεται η ισχυρά διαλείπουσα συμπεριφορά κάποιων φαινομένων, η ισχυρή αυτοσυσχέτιση καθώς περιοδικά φαινόμενα μικρότερης περιόδου, με κυριότερο την εναλλαγή ημέρας και νύχτας.

### 7.3 Συμπλήρωση - επέκταση

Συχνά οι χρονοσειρές περιέχουν κενά, π.χ. γιατί έπαθε βλάβη ένα όργανο ή καταγράφηκε λάθος μια τιμή. Αυτά τα κενά δυσχεραίνουν ορισμένες επεξεργασίες, οπότε είναι επιθυμητή η συμπλήρωσή τους. Η συμπλήρωση, ακόμα κι αν είναι πρόχειρη εκτίμηση με μεγάλα περιθώρια σφάλματος, μπορεί να είναι πολύ χρήσιμη, κυρίως όταν θέλουμε να εξαγάγουμε χρονοσειρές μικρότερης διακριτότητας. Για παράδειγμα, αν σε μια μηνιαία χρονοσειρά βροχοπτώσεων λείπει ο μήνας Ιανουάριος (π.χ. γιατί χάθηκε το έντυπο), τότε δεν μπορούμε να εξαγάγουμε την αντίστοιχη ετήσια τιμή, γιατί έχουμε πληροφορία μόνο για τους 11 μήνες του έτους. Είναι προτιμότερο τότε να συμπληρώσουμε κατ' εκτίμηση τον Ιανουάριο και να κάνουμε την συνάθροιση, ώστε να μη χαθεί η πληροφορία των υπόλοιπων 11 μηνών. Ακόμα κι αν η εκτίμηση έχει μεγάλο σφάλμα, η επιρροή του στην ετήσια τιμή θα είναι μικρή. Συμπλήρωση κάνουμε λοιπόν όταν υπάρχουν τέτοια σποραδικά κενά, π.χ. μέχρι τέσσερις (4) μήνες το χρόνο ή, σε ημερήσιες χρονοσειρές, π.χ. μέχρι δέκα (10) μέρες το μήνα.

Συχνά, όταν υπάρχουν εκτενέστερες χρονικά χρονοσειρές από γειτονικούς μετρητικούς σταθμούς, είναι δυνατό, με γραμμική παλινδρόμηση, να συμπληρωθούν μεγάλα χρονικά κενά ή να εκτιμηθούν τιμές πέραν των άκρων της υπάρχουσας χρονοσειράς. Αυτή η διαδικασία λέγεται *επέκταση* της χρονοσειράς. Η συμπλήρωση και η επέκταση είναι ουσιαστικά το ίδιο πράγμα — εκτίμηση τιμών που δεν υπάρχουν — αλλά έχουν διαφορετικούς στόχους και γίνονται με διαφορετικές μεθόδους. Η

συμπλήρωση γίνεται σε λίγες τιμές, με στόχο τη διευκόλυνση εξαγωγής χρονοσειρών μικρότερης διακριτότητας, και η ακρίβεια της εκτίμησης δεν έχει μεγάλη σημασία.

Η επέκταση γίνεται γιατί η επεκταμένη χρονοσειρά μπορεί, σε σχέση με την αρχική, να οδηγεί σε καλύτερη εκτίμηση των στατιστικών παραμέτρων ή να έχει χαρακτηριστικά πιο κοντά στα χαρακτηριστικά της διεργασίας που εκπροσωπεί. Απαιτεί επομένως κατάλληλες μεθόδους, αλλιώς τα στατιστικά χαρακτηριστικά θα αλλοιωθούν. Έτσι θα εφαρμόζεται η οργανική συσχέτιση η οποία εγγενώς διατηρεί τα στατιστικά χαρακτηριστικά του δείγματος ή κάποια άλλη μεθοδολογία σε συνδυασμό οπωσδήποτε με την εισαγωγή τυχαίου όρου (βλ. παράγραφο 7.3.1).

Η θεωρία συμπλήρωσης και επέκτασης περιγράφεται από τον Κουτσογιάννη (1997, σ. 227-236).

Σε κάθε περίπτωση, για να γίνει συμπλήρωση ή επέκταση, πρέπει να υπάρχουν μία ή περισσότερες χρονοσειρές από γειτονικούς σταθμούς, τις οποίες θα ονομάζουμε «ανεξάρτητες χρονοσειρές». Η χρονοσειρά που θα συμπληρωθεί ή θα επεκταθεί θα λέγεται «εξαρτημένη». Για κάθε χρονική θέση που πρόκειται να συμπληρωθεί ή να επεκταθεί, πρέπει να υπάρχουν τιμές σε όλες τις ανεξάρτητες χρονοσειρές. Το σύνολο αυτών των χρονικών θέσεων το ονομάζουμε «περίοδο συμπλήρωσης» ή περίοδο 2. Το σύνολο των χρονικών θέσεων για τις οποίες υπάρχουν τιμές σε όλες τις χρονοσειρές, ανεξάρτητες και εξαρτημένη, το ονομάζουμε «κοινή περίοδο» ή περίοδο 1. Το πλήθος των τιμών κοινής περιόδου και περιόδου συμπλήρωσης το συμβολίζουμε αντίστοιχα με  $N_1$  και  $N_2$ .

Για την διενέργεια της συμπλήρωσης ή της επέκτασης μπορεί να χρησιμοποιηθεί οποιαδήποτε από τις μεθόδους που παρουσιάστηκαν στην προηγούμενη ενότητα (απλή ή πολλαπλή γραμμική παλινδρόμηση, οργανική συσχέτιση, αυτοσυσχέτιση, κλπ.) με ή χωρίς να λαμβάνουμε υπόψη την εποχικότητα. Οι τιμές που έχουν προκύψει από συμπλήρωση ή από επέκταση θα πρέπει να σημαιοθετούνται κατάλληλα με χρήση σημαίας «από συμπλήρωση» (INFILLING) ώστε να είναι εύκολος ο εντοπισμός τους για περαιτέρω ελέγχους.

Αν θέσουμε ένα επίπεδο σημαντικότητας 5%, για να έχει νόημα η παλινδρόμηση (Κουτσογιάννης 1997, σ. 233) ο συντελεστής συσχέτισης πρέπει να έχει τιμή τουλάχιστον:

$$|r_{xy}| \geq \frac{2}{\sqrt{N_1}} \quad 7.18$$

Η παραπάνω ανισότητα σε όρους συντελεστή προσδιορισμού γράφεται ως:

$$\delta \geq \frac{4}{N_1} \quad 7.19$$

Οι παραπάνω σχέσεις, φυσικά, έχουν νόημα για τιμές του  $N_1 > 4$ . Για την περίπτωση της οργανικής συσχέτισης εισάγεται επιπλέον ένα άνω φράγμα:

$$|r_{xy, organic}| \geq \max \left\{ \frac{2}{\sqrt{N_1}}, 0.5 \right\} \quad 7.20$$

Εφόσον ο συντελεστής προσδιορισμού της οργανικής συσχέτισης προκύπτει από την σχέση:

$$\delta = 2|r_{xy}| - 1 \quad 7.21$$

τότε ο παραπάνω περιορισμός μπορεί να διατυπωθεί ως:

$$\delta \geq \frac{4}{\sqrt{N_1}} - 1 \quad 7.22$$



Να σημειωθεί πως η Σχέση (7.21) είναι αυτή που παρουσιάζεται στα αποτελέσματα της οργανικής συσχέτισης στο σύστημα του «Υδρογνώμων».

### 7.3.1 Εισαγωγή τυχαίου όρου

Κατά την συμπλήρωση των χρονοσειρών με χρήση της απλής ή της πολλαπλής γραμμικής παλινδρόμησης προκύπτει το πρόβλημα της υποεκτίμησης της διασποράς του συμπληρωμένου δείγματος. Ένας τρόπος αντιμετώπισης αυτού του προβλήματος είναι η εισαγωγή ενός τυχαίου όρου  $\varepsilon$  στις συμπληρωμένες τιμές ώστε να διατηρήσουμε την διασπορά στην αρχική προ της συμπλήρωσης τιμής (Χριστοφίδης, 1994). Χρησιμοποιείται ο τυχαίος όρος που προτείνεται από τους Matalas και Jacobs (1964):

$$\varepsilon = z\sigma_y a \sqrt{1 - R^2} \quad 7.23$$

όπου  $z$  οι τιμές που προκύπτουν από μία γεννήτρια τυχαίων αριθμών της τυποποιημένης κανονικής κατανομής (βλ. παράγραφο 9.7.1).  $R$  είναι ο συντελεστής (πολλαπλής) γραμμικής συσχέτισης. Ο συντελεστής  $a$  δίνεται από την σχέση:

$$a = \sqrt{\frac{N_2(N_1 - 2m - 2)(N_1 - 1)}{(N_2 - 1)(N_1 - m - 2)(N_1 - m - 1)}} \quad 7.24$$

όπου  $m$  ο αριθμός των χρονοσειρών αναφοράς (ανεξάρτητων μεταβλητών),  $N_1$  ο αριθμός των κοινών μετρήσεων μεταξύ των εξαρτημένων και ανεξάρτητων μεταβλητών και  $N_2$  ο αριθμός των ελλείψεων. Συνεπώς,  $N_1 + N_2$  θα είναι ο αριθμός των κοινών μετρήσεων μόνο μεταξύ των ανεξάρτητων μεταβλητών (χρονοσειρές αναφοράς). Η εισαγωγή τυχαίου όρου βλέπουμε πως δεν έχει νόημα για μικρό αριθμό ελλείψεων· σε τέτοια περίπτωση τα  $z$  που θα προκύψουν μπορεί να μην προσεγγίζουν καλά την κανονική κατανομή. Η αρχικοποίηση της γεννήτριας θα πρέπει να γίνεται σε συγκεκριμένη τιμή ώστε κάθε φορά να αναπαράγονται τα ίδια αποτελέσματα. Ωστόσο ο χρήστης έχει την δυνατότητα να αρχικοποιεί την γεννήτρια με κάποια τυχαία τιμή.

Η τεχνική της εισαγωγής του τυχαίου όρου μπορεί να χρησιμοποιηθεί εναλλακτικά της οργανικής συσχέτισης, μάλιστα διατηρώντας καλύτερα τον αρχικό συντελεστή συσχέτισης. Η μεθοδολογία έχει το μειονέκτημα της μη επαναληψιμότητας των αποτελεσμάτων για διαφορετικές γεννήτριες τυχαίων αριθμών ή διαφορετικών συνθηκών αρχικοποίησης.

Μία υπολογιστική προσέγγιση της παραπάνω διαδικασίας είναι αυτή της εύρεσης των στατιστικών χαρακτηριστικών του μετρημένου σφάλματος  $\varepsilon$  (βλ. Κουτσογιάννης, 1997, σ. 238). Αυτή η προσέγγιση ακολουθείται στον «Υδρογνώμων». Για την γενική περίπτωση της πολλαπλής γραμμικής παλινδρόμησης τα μετρημένα σφάλματα είναι:

$$\varepsilon_i = y_i - a_0 - \sum_{j=1}^m a_j x_{ij} \quad 7.25$$

όπου  $y$  η εξαρτημένη μεταβλητή και  $x_j$  οι ανεξάρτητες μεταβλητές. Αν βρούμε τα στατιστικά χαρακτηριστικά των  $\varepsilon_i$ , ήτοι την τυπική απόκλιση  $\sigma_w$  και την μέση τιμή  $\mu_w$ , στην συνέχεια μπορούμε να εκφράσουμε τους τυχαίους όρους σφάλματος χρησιμοποιώντας την γεννήτρια τυχαίων αριθμών της κανονικής κατανομής:

$$\varepsilon = N(\mu_w, \sigma_w) \quad 7.26$$

Η τιμή  $\mu_w$  (οφείλει να) είναι πολύ μικρή και μπορεί να ληφθεί πρακτικά ίση με το μηδέν.

### 7.3.2 Αποκοπή στο μηδέν

Για φυσικές μεταβλητές όπου οι αρνητικές τιμές δεν έχουν νόημα, όπως το ύψος βροχόπτωσης, υπάρχει η δυνατότητα αποκοπής στο μηδέν των πιθανών αρνητικών τιμών που θα παρουσιαστούν μετά την συμπλήρωση. Τέτοιες περιπτώσεις μπορούν να παρουσιαστούν όταν ο σταθερός όρος της γραμμικής παλινδρόμησης είναι μικρότερος του μηδέν. Η αποκοπή γίνεται με εφαρμογή της σχέσης:

$$y_i < 0 \Rightarrow y_i = 0 \quad 7.27$$

Η πρακτική αυτή έχει αρκετά μικρή επίπτωση στην τυπική απόκλιση του συμπληρωμένου δείγματος και σχετικά μικρή στην μέση τιμή.

## 8 Υδατικά ισοζύγια

### 8.1 Εισαγωγή

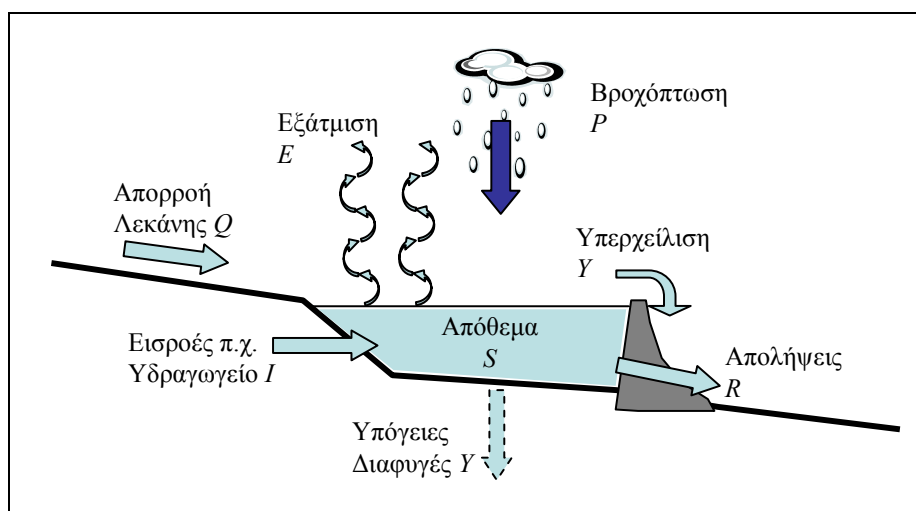
Τα υδατικά ισοζύγια χρησιμοποιούνται για να προσομοιάσουν τις υδρολογικές διεργασίες μίας λεκάνης απορροής με απώτερο στόχο την ορθολογική διαχείριση των υδατικών πόρων. Εφαρμόζοντας τα υδατικά ισοζύγια είναι δυνατή η έμμεση εκτίμηση υδρολογικών μεταβλητών που δεν μπορούν να μετρηθούν άμεσα όπως η απορροή. Στα πλαίσια του «Υδρογνώμων» παρέχονται τα εξής υπολογιστικά εργαλεία:

- Ισοζύγια υδάτινων φορέων με εφαρμογή σε ταμιευτήρες και υδραγωγεία. Οι υπολογισμοί βασίζονται σε απλές, γραμμικές σχέσεις
- Αδρομερής προσομοίωση των διεργασιών λεκάνης απορροής, μοντέλο βροχής – απορροής με την ονομασία «Ζυγός». Οι υπολογισμοί είναι σύνθετοι, ειδικοί αλγόριθμοι βελτιστοποίησης εφαρμόζονται για την ρύθμιση του μοντέλου.

### 8.2 Ισοζύγια υδάτινων φορέων

#### 8.2.1 Ισοζύγια ταμιευτήρων

Η κατάρτιση ισοζυγίων ταμιευτήρων είναι μία πολύ απλή αλλά χρήσιμη μεθοδολογία για την εκτίμηση υδρολογικών μεταβλητών που δεν μπορούν να μετρηθούν άμεσα. Η απλότητα της μεθοδολογίας έγκειται στις απολύτως γραμμικές σχέσεις των μεταβλητών που συμμετέχουν στο ισοζύγιο σε συνδυασμό με τη δυνατότητα ακριβούς εκτίμησης του αποθέματος και του εμβαδού της ελεύθερης επιφάνειας του ταμιευτήρα. Έτσι η μεταβολή του αποθέματος σε δεδομένο χρόνο μπορεί να εξισωθεί με την διαφορά εισροών και εκροών, η δε άγνωστη ποσότητα θα είναι το έλλειμμα του ισοζυγίου. Στο Σχήμα 8.1 σχηματοποιείται η γενική μορφή του ισοζυγίου ενός ταμιευτήρα.



Σχήμα 8.1: Σχηματική αναπαράσταση της γενικής μορφής ισοζυγίου ταμιευτήρα

Η μεθοδολογία περιγράφεται από τους Κουτσογιάννη και Ξανθόπουλο (1999, σ. 246) καθώς και από τους Ευστρατιάδη κ.α. (2000, σ. 21). Η γενική σχέση του ισοζυγίου ταμιευτήρα μπορεί να διατυπωθεί με το διαφορικό του αποθέματος  $S$ :

$$\frac{dS(t)}{dt} = I(t) - O(t) \quad 8.1$$

όπου  $I(t)$ ,  $O(t)$ , οι γενικές σχέσεις για τις εισροές και τις εκροές του ταμιευτήρα σε μονάδες παροχής. Η Σχέση 8.1 μπορεί να διατυπωθεί με όρους διαφορών ως:

$$\Delta S = P + Q + I - E - R - Y - G \quad 8.2$$

Η μεταβολή του αποθέματος  $\Delta S$  αφορά ένα συγκεκριμένο χρονικό διάστημα, στην περίπτωση του «Υδρογνώμων» είναι ο μήνας. Έτσι χρησιμοποιώντας την χρονοσειρά αποθέματος  $S(t)$ , υπολογίζεται η μεταβολή του αποθέματος σε ένα μήνα αφαιρώντας την τιμή της αρχής του μήνα από την τιμή στην αρχή του επόμενου μήνα. Η χρονοσειρά του αποθέματος προκύπτει από την χρονοσειρά της στάθμης και εφαρμόζοντας στην συνέχεια μεθόδους παρεμβολής (βλ. ενότητα 5.2). Στην συνέχεια εντάσσονται οι παρακάτω χρονοσειρές με τις μηνιαίες τιμές όγκου:

- Βροχόπτωση  $P$ . Χρησιμοποιείται η χρονοσειρά του ύψους βροχόπτωσης, ανάγεται στη μέση απόλυτη στάθμη του ταμιευτήρα και έπειτα πολλαπλασιάζεται με την μέγεθος της ελεύθερης επιφάνειας για να αναχθεί σε όγκο
- Ο όγκος απορροής  $Q$  είναι συνήθως το άγνωστο μέγεθος του ισοζυγίου, εντούτοις μπορεί να είναι και γνωστό π.χ. μέσω υδρομετρίας. Όταν ο όγκος απορροής είναι άγνωστη ποσότητα, τότε προκύπτει ως το έλλειμμα του ισοζυγίου. Στην συνέχεια είναι δυνατή η διαίρεση του με την επιφάνεια της λεκάνης απορροής (αφαιρώντας την τρέχουσα επιφάνεια του ταμιευτήρα) ώστε να προκύψει το ύψος απορροής
- Ο όγκος λοιπών γενικών εισροών  $I$  περιλαμβάνει εισερχόμενες παροχές από υδραγωγεία
- Η εξάτμιση  $E$  από την ελεύθερη επιφάνεια του ταμιευτήρα υπολογίζεται από την χρονοσειρά ύψους εξάτμισης πολλαπλασιασμένη με το μέγεθος της ελεύθερης επιφάνειας για να αναχθεί σε όγκο. Ο υπολογισμός της χρονοσειρά εξάτμισης μπορεί να γίνει χρησιμοποιώντας τυπικές μεθόδους όπως αυτή του Penman σύμφωνα με το Κεφάλαιο 6
- Οι απολήψεις  $R$  παρέχονται είναι ο όγκος των συνολικών απολήψεων για υδρευτικούς και αρδευτικούς σκοπούς καθώς και ενδεχόμενη οικολογική παροχή
- Η υπερχειλίση από τα φράγματα  $Y$ , εφόσον υπάρχουν εκτιμήσεις και χρονοσειρές για τους όγκους που διοχετεύτηκαν μέσω των υπερχειλιστών. Σε διαφορετική περίπτωση που είναι άγνωστη ποσότητα, μπορεί να εκτιμηθεί μέσω της μέγιστης αποθηκευτικότητας του ταμιευτήρα  $K$  (όγκος υπερχειλίσης) ως:  $Y = \max(S + P + Q + I - E - R - G - K, 0)$
- Οι υπόγειες διαφυγές  $G$  είναι 0 εφόσον ο ταμιευτήρας βρίσκεται σε στεγανή γεωλογικό σχηματισμό. Εφόσον ο πυθμένας έχει διαφυγές, αυτές πρέπει να εκτιμηθούν χρησιμοποιώντας μία απλή αναλυτική σχέση και την χρονοσειρά στάθμης (βλ. ενότητα 5.3). Είναι πιθανόν δε, οι υπόγειες διαφυγές να αποτελούν το άγνωστο μέγεθος του ισοζυγίου.

Τα μεγέθη που είναι εκφρασμένα σε ύψη όπως η βροχόπτωση και η εξάτμιση πρέπει να πολλαπλασιαστούν με το μέγεθος της ελεύθερης επιφάνειας του ταμιευτήρα. Η χρονοσειρά της επιφάνειας μπορεί να προκύψει από την χρονοσειρά στάθμης εφαρμόζοντας την μεθοδολογία παρεμβολής που περιγράφεται στην ενότητα 5.2. Ο χρήστης μπορεί να επιλέξει αν θα χρησιμοποιείται η τιμή της επιφάνειας στην αρχή του μήνα, ή μία μέση αντιπροσωπευτική τιμή που προκύπτει από την αρχή του τρέχοντος και του επόμενου μήνα.

## 8.2.2 Ισοζύγια υδραγωγείων

Τα ισοζύγια υδραγωγείων χρησιμοποιούνται για τον έλεγχο ύπαρξης διαρροών. Η μαθηματική τους διατύπωσή είναι απλούστερη από αυτήν του ισοζυγίου ταμιευτήρων καθώς δεν υπάρχει ο όρος της αποθήκευσης, συνοψίζεται δε στην παρακάτω σχέση:

$$I(t) = O(t) + L(t) \quad 8.3$$

όπου  $I(t)$  και  $O(t)$  η συνολική παροχή των εισροών και των εκροών αντίστοιχα. Ο όρος  $L(t)$  αντιστοιχεί στις υπολογισμένες διαρροές. Τα  $I(t)$ ,  $O(t)$  μετρούνται με την κατάλληλη μεθοδολογία, όπως με χρήση μονίμως εγκατεστημένων παροχόμετρων ή με σταθμημετρία και εφαρμογή σχέσεων στάθμης – παροχής. Το ισοζύγιο επιλύεται σειριακά από την πρώτη εισροή έως την τελευταία εκροή ώστε να υπολογίζονται ενδιάμεσες παροχές σε σημεία ελέγχου. Αν μάλιστα στα σημεία ελέγχου υπάρχει δυνατότητα μέτρησης της παροχής, τότε είναι δυνατή η εκτίμηση των διαρροών μεταξύ των σημείων ελέγχου.

Η υπολογιστική διαδικασία είναι απλή και στηρίζεται στον αλγόριθμο των γραμμικών πράξεων μεταξύ χρονοσειρών που περιγράφεται στην ενότητα 4.4. Οι χρονοσειρές εισόδου μπορούν να έχουν ημερήσιο, μηνιαίο ή ετήσιο βήμα, συνεπώς η χρονοσειρά των υπολογισμένων διαρροών θα έχει το μεγαλύτερο από τα χρονικά βήματα που χρησιμοποιούνται στο ισοζύγιο. Έτσι οι χρονοσειρές εισόδου συναθροίζονται στο τελικό χρονικό βήμα χρησιμοποιώντας τον αλγόριθμο συνάθροισης που περιγράφεται στην ενότητα 4.3.

## 8.3 Προσομοίωση διεργασιών λεκάνης – το μοντέλο «Ζυγός»

### 8.3.1 Περιγραφή του μοντέλου

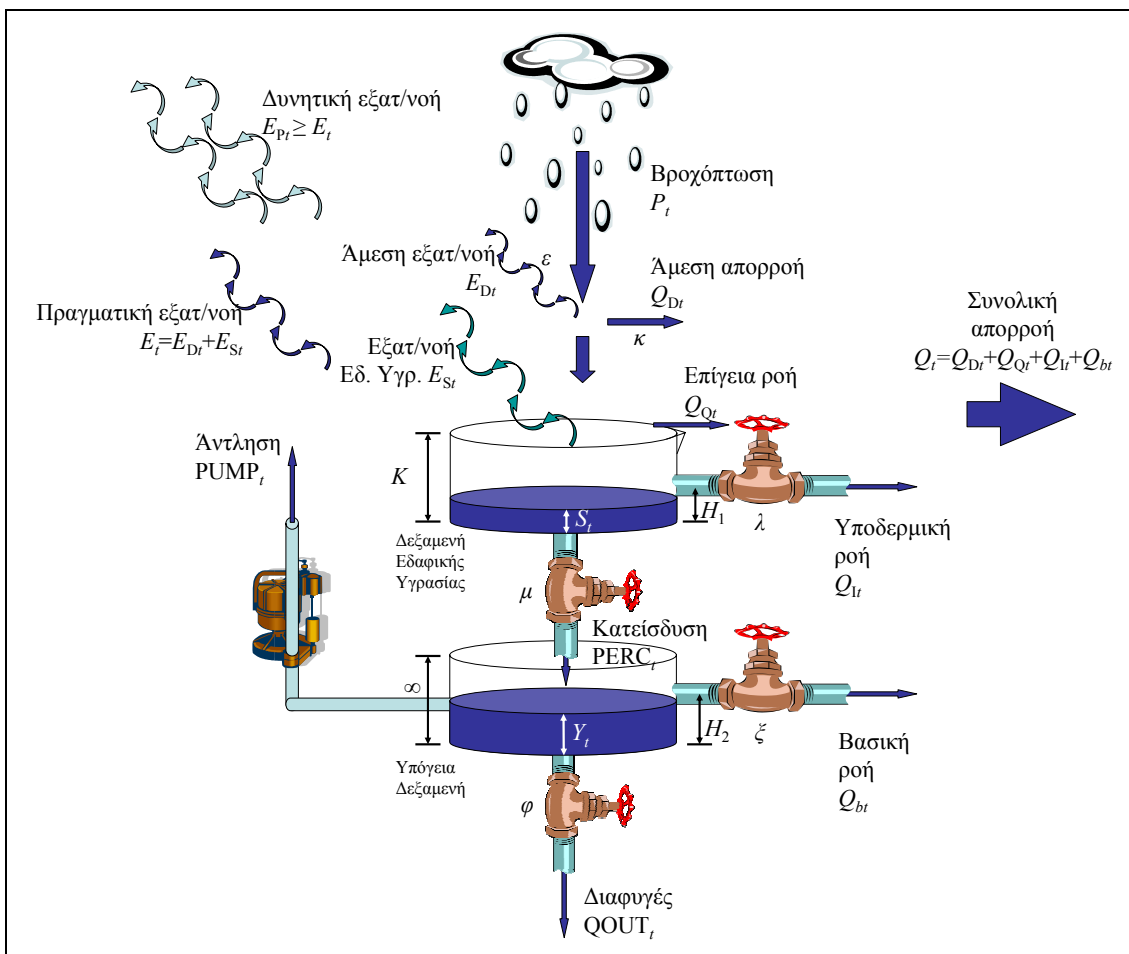
#### Γενικά

Η προσομοίωση των επιφανειακών διεργασιών της λεκάνης, η οποία αποσκοπεί στον επιμερισμό της βροχόπτωσης σε απορροή, απώλειες εξάτμισης και διαπνοής, και κατείδυση (βαθιά διήθηση), υλοποιείται μέσω ενός μοντέλου ισοζυγίου της εδαφικής υγρασίας, που αποτελεί μια τροποποιημένη εκδοχή του γνωστού μοντέλου Thornthwaite (Dingman, 1994, pp. 299-301· Κουτσογιάννης και Ξαnthόπουλος, 1999, σ. 236-243). Επιπλέον, η προσομοίωση των υπόγειων υδρολογικών διεργασιών σε υποτιθέμενο υπόγειο υδροφορέα που υπόκειται της λεκάνης, που αποσκοπεί στον επιμερισμό της κατείδυσης σε βασική απορροή και υπόγειες διαφυγές εκτός λεκάνης (π.χ. στη θάλασσα), υλοποιείται μέσω ενός μοντέλου ισοζυγίου γραμμικού ταμιευτήρα. Ο ταμιευτήρας, που έχει άπειρη χωρητικότητα, αναπαριστά τον υπόγειο υδροφορέα, και αποτελεί μια προσέγγιση απολύτως όμοια με εκείνη του τροποποιημένου μοντέλου Thornthwaite που χρησιμοποιείται για τις επιφανειακές διεργασίες. Με τον τρόπο αυτό, το πολύπλοκο υδρολογικό σύστημα μιας λεκάνης απορροής αναπαρίσταται, εννοιολογικά, μέσω των δύο επάλληλων δεξαμενών, ήτοι της δεξαμενής εδαφικής υγρασίας και της δεξαμενής υπόγειου νερού.

Στο Σχήμα 8.2 απεικονίζεται η δομή και ο τρόπος λειτουργίας του μοντέλου. Οι κρουνοί αναπαριστούν τις αδιάστατες παραμέτρους, όπου μία μηδενική τιμή αντιπροσωπεύει έναν κλειστό κρουνο που δεν επιτρέπει την ροή ενώ μία τιμή κοντά στην μονάδα αντιπροσωπεύει την δυνατότητα πλήρους στράγγισης. Έτσι θέτοντας κάποιες αδιάστατες παραμέτρους σταθερά μηδενικές (κλείσιμο των κρουनों), έχουμε την δυνατότητα να απλοποιήσουμε το μοντέλο μειώνοντας τους βαθμούς ελευθερίας. Η αναπαράσταση της δυνητικής εξατμοδιαπνοής  $E_p$ , δεν έχει έννοια απευθείας συμμετοχής της στο ισοζύγιο της λεκάνης αλλά ως μίας χρονοσειράς εισόδου και είναι ένα φυσικό όριο για την πραγματική εξατμοδιαπνοή (η οποία συμμετέχει άμεσα στο ισοζύγιο).

## Λειτουργία του μοντέλου

Το μοντέλο αποτελεί μια απλοποιημένη εκδοχή ενός ημικαταναμημένου σχήματος προσομοίωσης, το οποίο έχει αναπτυχθεί στο Ε.Μ.Π. (Ευστρατιάδης κ.ά., 2004, 2005; Rozos et al., 2004). Εντάσσεται στην κατηγορία των συγκεντρωτικών εννοιολογικών μοντέλων υδατικού ισοζυγίου, που έχουν τύχει ευρείας εφαρμογής διεθνώς, καθώς χρησιμοποιώντας απλές υποθέσεις σχετικά με τη λειτουργία των φυσικών μηχανισμών και μικρό πλήθος παραμέτρων, επιτυγχάνουν καλή προσαρμογή ακόμη και στην περίπτωση εξαιρετικά πολύπλοκων συστημάτων. Τα εν λόγω μοντέλα λειτουργούν, ως επί το πλείστον, σε μηνιαίο χρονικό βήμα (Makhlouf and Michel, 1994; Xu and Singh, 1998). Τονίζεται ότι οι παράμετροι των εννοιολογικών μοντέλων, αν και δεν έχουν απόλυτη φυσική ερμηνεία, μπορούν να θεωρηθούν αντιπροσωπευτικές κάποιων μέσων χαρακτηριστικών της λεκάνης, καθώς εξαρτώνται από παράγοντες όπως η εδαφολογία, η γεωλογία, η φυτοκάλυψη, κλπ..



Σχήμα 8.2: Σχηματική παράσταση αδρομερούς μοντέλου βροχόπτωσης – απορροής «Ζυγός»

Δεδομένα εισόδου του συγκεκριμένου μοντέλου είναι οι χρονοσειρές βροχόπτωσης,  $P_t$ , δυναμικής εξατμοδιαπνοής,  $E_{Pt}$ , καθώς και απόληψης υπόγειου νερού  $PUMP_t$ . Όλες οι χρονοσειρές θεωρούνται ολοκληρωμένες σε κλίμακα λεκάνης και δίνονται στη μορφή ισοδύναμου ύψους νερού. Η υπόγεια απόληψη θεωρείται ίση με την αντλούμενη ποσότητα από τις γεωτρήσεις στο σύνολο του υδροφορέα. Ο «Υδρογνώμων» ελέγχει τις χρονοσειρές εισόδου ώστε να έχουν το ίδιο χρονικό βήμα, χωρίς χρονικές ολισθήσεις (βλ. ενότητα 4.2 για την άρση των χρονικών ολισθήσεων). Επιπλέον οι χρονοσειρές εισόδου πρέπει να είναι «συμπαγείς» στο χρονικό διάστημα για το οποίο γίνονται οι υπολογισμοί χωρίς κενά και ελλείψεις. Η ύπαρξη τυχόν ελλείψεων επισημαίνεται στον χρήστη ο

οποίος καλείται να ορίσει μία «συμπαγή» περίοδο. Αυτή η απαίτηση δεν ισχύει για την μετρημένη απορροή που χρησιμοποιείται στην βαθμονόμηση του μοντέλου (βλ. παράγραφο 8.3.3).

Στην πλήρη μορφή του, το μοντέλο εκτιμά:

- την απορροή στην έξοδο της λεκάνης,  $Q_t$ ·
- την πραγματική εξατμοδιαπνοή,  $E_t$ ·
- την κατείδυση νερού προς τον υπόγειο υδροφορέα, PERC<sub>t</sub>·
- τις διαφυγές εκτός λεκάνης, QOUT<sub>t</sub>·
- την αποθήκευση εδαφικής υγρασίας,  $S_t$ ·
- την αποθήκευση υπόγειου νερού,  $Y_t$

Τα παραπάνω μεγέθη εξάγονται από τον «Υδρογνώμων» σε μορφή χρονοσειρών ιδίου χρονικού βήματος με τις χρονοσειρές εισόδου. Ειδικότερα, η απορροή περιλαμβάνει τέσσερις συνιστώσες, ήτοι:

- την άμεση απορροή,  $Q_{D_t}$ , που οφείλεται στην ύπαρξη αδιαπέρατων σχηματισμών, μέσω των οποίων ποσοστό της βροχής μετατρέπεται άμεσα σε απορροή·
- την επίγεια ροή,  $Q_{Q_t}$ , που είναι μια ταχεία απόκριση που οφείλεται στον κορεσμό του εδάφους·
- την υποδερμική ροή,  $Q_{I_t}$ , που είναι μια αργή απόκριση που οφείλεται στην πλευρική (οριζόντια) κίνηση του νερού που εισχωρεί στο έδαφος·
- τη βασική ροή,  $Q_{B_t}$ , ήτοι την απόκριση των κατώτερων εδαφικών στρωμάτων (υδροφορέας), μέσω πηγών.

Η πραγματική εξατμοδιαπνοή, που είναι εξ ορισμού μικρότερη ή ίση με τη δυναμική και θεωρείται υδρολογική απώλεια, περιλαμβάνει δύο συνιστώσες:

- την άμεση εξατμοδιαπνοή,  $E_{D_t}$ , που αναφέρεται στον ταχύ μετασχηματισμό μέρους της βροχόπτωσης σε απώλειες λόγω εξάτμισης, και
- την εδαφική εξατμοδιαπνοή,  $E_{S_t}$ , που αναφέρεται στην ικανοποίηση του ελλείμματος της δυναμικής εξατμοδιαπνοής μέσω της διαθέσιμης εδαφικής υγρασίας.

Η κατείδυση ή βαθιά διήθηση, PERC<sub>t</sub>, αναφέρεται στην κατακόρυφη μεταφορά νερού από τα ανώτερα (ακόρεστη ζώνη) προς τα κατώτερα (κορεσμένη ζώνη) στρώματα του εδάφους.

Οι διαφυγές εκτός λεκάνης αναφέρονται στο σύνολο των λοιπών υδρολογικών απωλειών, που οφείλονται στην πλευρική τροφοδοσία γειτονικών υδροφορέων ή την εκροή νερού προς τη θάλασσα.

Η αποθήκευση εδαφικής υγρασίας και η αποθήκευση υπόγειου νερού είναι μεταβλητές κατάστασης του μοντέλου, που προϋποθέτουν γνώση των αρχικών συνθηκών  $S_0$  και  $Y_0$ , αντίστοιχα. Επισημαίνεται ότι εφόσον η προσομοίωση ξεκινά στην αρχή του υδρολογικού έτους, η αρχική τιμή της εδαφικής υγρασίας μπορεί να θεωρηθεί ίση με μηδέν (όχι όμως η  $Y_0$ ).

### Παράμετροι μοντέλου

Στην πλήρη μορφή του, το μοντέλο έχει εννέα μεταβλητές ελέγχου (παραμέτρους), ενώ ως μεταβλητές ελέγχου μπορούν να θεωρηθούν και οι δύο προαναφερθείσες αρχικές συνθήκες, δεδομένου ότι δεν υπάρχει τρόπος να μετρηθούν, οπότε αναγκαστικά πρέπει να εκτιμηθούν. Το σύνολο των παραμέτρων του μοντέλου φαίνεται στον Πίνακα 8.1. Συγκεκριμένα:

Το ποσοστό της βροχόπτωσης που διατίθεται για ικανοποίηση της άμεσης εξατμοδιαπνοής,  $\varepsilon$ , είναι ένα αδιάστατο μέγεθος, και χρησιμοποιείται ώστε να εξομαλύνει τη διαφορετική χρονική κατανομή της βροχόπτωσης, που είναι μια διαλείπουσα διεργασία, σε σχέση με τη δυναμική εξατμοδιαπνοή, που είναι μια συνεχής διεργασία. Αναλυτικότερα, με την εφαρμογή του μοντέλου σε μηνιαία κλίμακα, η υπόθεση ότι το σύνολο της βροχόπτωσης μπορεί να μετασχηματιστεί σε άμεση εξάτμιση είναι μη

ρεαλιστική. Συνήθως, η μηνιαία βροχόπτωση πραγματοποιείται στα πλαίσια περιορισμένου αριθμού επεισοδίων βροχής, τα οποία λαμβάνουν χώρα σε μικρές χρονικές κλίμακες (ωριαία, ημερήσια). Είναι προφανές ότι σε τέτοιες κλίμακες, η βροχόπτωση υπερβαίνει την αντίστοιχη ζήτηση για εξάτμιση, κάτι που μπορεί να μην ισχύει εφόσον τα μεγέθη συναθροιστούν σε μηνιαία κλίμακα.

Πίνακας 8.1: Παράμετροι του αδρομερούς μοντέλου βροχόπτωσης – απορροής «Ζυγός»

Παράμετρος	Περιγραφή	Διαστάσεις
$\varepsilon$	Ποσοστό βροχόπτωσης που διατίθεται για ικανοποίηση της άμεσης εξάτμισηδιαπνοής	Αδιάστατο
$\kappa$	Ποσοστό της περίσσειας βροχόπτωσης που εμφανίζεται ως άμεση απορροή	Αδιάστατο
$K$	Χωρητικότητα Δεξαμενής Εδαφικής Υγρασίας	[L]
$\lambda$	Ρυθμός εκφόρτισης Δεξαμενής Εδαφικής Υγρασίας για παραγωγή υποδερμικής ροής	Αδιάστατο
$H_1$	Κατώφλι αποθέματος Δεξαμενής Εδαφικής Υγρασίας για παραγωγή υποδερμικής ροής	[L]
$\mu$	Ρυθμός εκφόρτισης Δεξαμενής Εδαφικής Υγρασίας για παραγωγή κατείσδυσης	Αδιάστατο
$\xi$	Ρυθμός εκφόρτισης Δεξαμενής Υπόγειου Νερού για παραγωγή βασικής ροής	Αδιάστατο
$H_2$	Κατώφλι αποθέματος Δεξαμενής Υπόγειου Νερού για παραγωγή βασικής ροής	[L]
$\varphi$	Ρυθμός εκφόρτισης Δεξαμενής Υπόγειου Νερού για παραγωγή υπόγειας διαφυγής	Αδιάστατο
$S_0$	Αρχικό απόθεμα Δεξαμενής Εδαφικής Υγρασίας	[L]
$Y_0$	Αρχικό απόθεμα Δεξαμενής Υπόγειου Νερού	[L]

Το ποσοστό της περίσσειας βροχόπτωσης που εμφανίζεται ως άμεση απορροή,  $\kappa$ , είναι ένα αδιάστατο μέγεθος, που εξαρτάται από την έκταση των αδιαπέρατων εδαφικών σχηματισμών της λεκάνης. Τυπικές τιμές του συντελεστή  $\kappa$  είναι μεταξύ 0.10 και 0.20.

Η χωρητικότητα της δεξαμενής εδαφικής υγρασίας, έχει διαστάσεις ισοδύναμου ύψους νερού, και εξαρτάται από τον τύπο του εδάφους (κυρίως από το πορώδες) και τη φυτοκάλυψη. Τυπικά, η χωρητικότητα αυτή, κυμαίνεται ανάμεσα στα 100-200 mm (Brutsaert, 1982, σ. 243).

Ο ρυθμός εκφόρτισης της δεξαμενής εδαφικής υγρασίας για παραγωγή υποδερμικής ροής,  $\lambda$ , είναι αδιάστατο μέγεθος, και εξαρτάται από τα χαρακτηριστικά του εδάφους.

Το κατώφλι του αποθέματος της δεξαμενής εδαφικής υγρασίας για παραγωγή υποδερμικής ροής,  $H_1$  ( $H_1 < K$ ), έχει διαστάσεις ισοδύναμου ύψους νερού, και εξαρτάται από τα χαρακτηριστικά του εδάφους.

Ο ρυθμός εκφόρτισης της δεξαμενής εδαφικής υγρασίας για παραγωγή κατείσδυσης,  $\mu$ , είναι αδιάστατο μέγεθος, και εξαρτάται από τα χαρακτηριστικά του εδάφους και του υδροφορέα.

Ο ρυθμός εκφόρτισης της δεξαμενής υπόγειου νερού για παραγωγή βασικής ροής,  $\xi$ , είναι αδιάστατο μέγεθος, και εξαρτάται από τα υδρογεωλογικά χαρακτηριστικά (π.χ., βαθμός καρστικοποίησης) του υδροφορέα.



Το κατώφλι αποθέματος της δεξαμενής υπόγειου νερού για παραγωγή βασικής ροής,  $H_2$ , έχει διαστάσεις ισοδύναμου ύψους νερού, και εξαρτάται από τα υδρογεωλογικά χαρακτηριστικά του υδροφορέα και την ύπαρξη πηγών.

Τέλος, ο ρυθμός εκφόρτισης της δεξαμενής υπόγειου νερού για παραγωγή υπόγειας διαφυγής,  $\varphi$ , είναι αδιάστατο μέγεθος, που εξαρτάται από το βαθμό επικοινωνίας του υδροφορέα με γειτονικούς υδροφορείς ή τη θάλασσα.

### 8.3.2 Περιγραφή υπολογιστικής διαδικασίας

#### Διεργασίες επιφανειακής υδρολογίας

Η υπολογιστική διαδικασία εφαρμογής του μοντέλου έχει ως εξής: Έστω ότι στην αρχή της περιόδου  $t$  είναι γνωστές οι είσοδοι του μοντέλου, ήτοι η βροχόπτωση,  $P_t$ , και η δυνητική εξατμοδιαπνοή,  $E_{Pt}$ . Αρχικά εκτιμάται η άμεση εξατμοδιαπνοή από τη σχέση:

$$E_{Dt} = \min (\varepsilon P_t, E_{Pt}) \quad 8.4$$

όπου  $\varepsilon$  το ποσοστό της βροχόπτωσης που μπορεί να μετατραπεί σε άμεση εξατμοδιαπνοή.

Η περίσσεια της βροχόπτωσης πέραν της άμεσης εξατμοδιαπνοής μετατρέπεται κατά ένα ποσοστό  $\kappa$  σε άμεση απορροή  $Q_{Dt}$  ως εξής:

$$Q_{Dt} = \kappa (P_t - E_{Dt}) \quad 8.5$$

Κατά συνέπεια, η διαθέσιμη εδαφική υγρασία,  $S_{\text{init}}$ , προκύπτει ως άθροισμα της αποθηκευμένης εδαφικής υγρασίας,  $S_{t-1}$ , και της περίσσειας βροχόπτωσης, ήτοι:

$$S_{\text{init}} = S_{t-1} + P_t - E_{Dt} - Q_{Dt} \quad 8.6$$

Εφόσον η διαθέσιμη εδαφική υγρασία ξεπερνά τη χωρητικότητα της δεξαμενής εδαφικής υγρασίας  $K$ , η περίσσεια απορρέει ως επίγεια ροή  $Q_{Qt}$

$$Q_{Qt} = \max (0, S_{\text{init}} - K) \quad 8.7$$

Η διαθέσιμη εδαφική υγρασία μειώνεται ως εξής:

$$S_{\text{init}} = S_{\text{init}} - Q_{Qt} \quad 8.8$$

Στη συνέχεια, θεωρείται η εξατμοδιαπνοή από την εδαφική υγρασία σύμφωνα με την κλασική προσέγγιση του μοντέλου Thornthwaite. Το έλλειμμα ικανοποίησης της ζήτησης για εξατμοδιαπνοή  $D$  είναι:

$$D = E_{Pt} - E_{Dt} \quad 8.9$$

ενώ η εδαφική εξατμοδιαπνοή, που εξαρτάται από το ποσοστό πλήρωσης της δεξαμενής εδαφικής υγρασίας και το έλλειμμα  $D$ , υπολογίζεται μέσω της σχέσης:

$$E_{St} = S_{\text{init}} (1 - e^{-D/K}) \quad 8.10$$

Αυτό μειώνει περαιτέρω την εδαφική υγρασία, ήτοι:

$$S_{\text{init}} = S_{\text{init}} - E_{St} \quad 8.11$$

Ακολουθεί η πραγματοποίηση της υποδερμικής ροής:

$$Q_{I_t} = \max \{0, \lambda (S_{init} - H_1)\} \quad 8.12$$

όπου  $H_1$  είναι το κατώφλι άνω του οποίου πραγματοποιείται η διεργασία και  $\lambda$  παράμετρος. Αυτό μειώνει περαιτέρω την εδαφική υγρασία, ήτοι:

$$S_{init} = S_{init} - Q_{I_t} \quad 8.13$$

Ως τελευταία διεργασία για τη δεξαμενή εδαφικής υγρασίας είναι η κατείσδυση στον υπόγειο υδροφορέα PERC<sub>t</sub>:

$$PERC_t = \max (0, \mu S_{init}) \quad 8.14$$

όπου  $\mu$  παράμετρος. Αυτό μειώνει την εδαφική υγρασία ως εξής:

$$S_{init} = S_{init} - PERC_t \quad 8.15$$

### Διεργασίες υπόγειας υδρολογίας

Η δεξαμενή του υπόγειου νερού με αρχικό απόθεμα  $Y_{t-1}$  δέχεται την κατείσδυση PERC<sub>t</sub> από τη δεξαμενή εδαφικής υγρασίας καθώς και υφίσταται απόληψη νερού ίση με την αντλούμενη ποσότητα από τις θεωρούμενες γεωτρήσεις στο σύνολο του υδροφορέα PUMP<sub>t</sub>. Το διαθέσιμο για περαιτέρω διεργασίες απόθεμα στη δεξαμενή  $Y_{init}$  είναι

$$Y_{init} = Y_{t-1} + PERC_t - PUMP_t \quad 8.16$$

Εφόσον το διαθέσιμο απόθεμα υπόγειου νερού υπερβαίνει ένα κατώφλι  $H_2$  παράγεται βασική ροή  $Q_{B_t}$  σύμφωνα με την σχέση:

$$Q_{B_t} = \max \{0, \zeta (Y_{init} - H_2)\} \quad 8.17$$

όπου  $H_2$  είναι το κατώφλι άνω του οποίου πραγματοποιείται η διεργασία και  $\zeta$  παράμετρος. Αυτό μειώνει περαιτέρω το διαθέσιμο απόθεμα υπόγειου νερού ως εξής:

$$Y_{init} = Y_{init} - Q_{B_t} \quad 8.18$$

Ως τελευταία διεργασία για τη δεξαμενή εδαφικής υγρασίας είναι η διαφυγή εκτός λεκάνης QOUT<sub>t</sub>:

$$QOUT_t = \varphi Y_{init} \quad 8.19$$

όπου  $\varphi$  παράμετρος. Το διαθέσιμο απόθεμα υπόγειου νερού μειώνεται ως εξής:

$$Y_{init} = Y_{init} - QOUT_t \quad 8.20$$

Σε κάθε χρονικό βήμα, όλες οι παραπάνω διεργασίες θεωρούνται σε μια διαδοχική σειρά  $n$  ίσων επί μέρους υποβημάτων το κάθε ένα από τα οποία ισούται με το κλάσμα  $(1 / n)$  του αρχικού βήματος του μοντέλου (π.χ. ένας μήνας). Αυτό εισάγει την αναγκαιότητα επιμερισμού όλων των δεδομένων εισόδου στα υποβήματα. Ο επιμερισμός που υλοποιήθηκε είναι ο απλούστερος δυνατός και, συγκεκριμένα, η απλή διαίρεση δια  $n$ . Στον «Υδρογνώμων» υιοθετείται μία τιμή του  $n = 10$ .

### 8.3.3 Βαθμονόμηση του μοντέλου

#### Γενικά

Οι παράμετροι των εννοιολογικών υδρολογικών μοντέλων δεν είναι μετρήσιμα μεγέθη αλλά εκτιμώνται έμμεσα, μέσω μιας συστηματικής διαδικασίας αξιολόγησης εναλλακτικών συνδυασμών

που καλείται βαθμονόμηση (calibration). Η βαθμονόμηση ενός μοντέλου γίνεται συγκρίνοντας τις προσομοιωμένες αποκρίσεις της λεκάνης με τις παρατηρημένες, και επιλέγοντας τον συνδυασμό εκείνο που επιτυγχάνει την καλύτερη προσαρμογή. Προφανώς, αυτό προϋποθέτει ένα επαρκές χρονικό διάστημα για το οποίο διατίθενται συστηματικές μετρήσεις (χρονοσειρές) απόκρισης της λεκάνης, έτσι ώστε να είναι δυνατή η σύγκριση των ιστορικών με τα προσομοιωμένα δεδομένα.

Η βαθμονόμηση ενός υδρολογικού μοντέλου μπορεί να διατυπωθεί ως ένα πρόβλημα μη γραμμικής βελτιστοποίησης της μορφής:

$$\begin{aligned} & \text{minimise } f(\mathbf{x}) \\ & \mathbf{x}^{\min} \leq \mathbf{x} \leq \mathbf{x}^{\max} \end{aligned} \quad 8.21$$

όπου  $f$  ένα βαθμωτό μέτρο σφάλματος που συνιστά την αντικειμενική συνάρτηση του προβλήματος βελτιστοποίησης,  $\mathbf{x} = (x_1, \dots, x_n)$  το διάνυσμα των παραμέτρων (μεταβλητών ελέγχου),  $n$  η διάσταση του προβλήματος (πλήθος παραμέτρων), και  $\mathbf{x}^{\min}$ ,  $\mathbf{x}^{\max}$  τα όρια διακύμανσης των παραμέτρων.

Ως μέτρο σφάλματος χρησιμοποιείται θεωρείται ο συντελεστής προσδιορισμού (coefficient of determination), που είναι το τυπικό μέτρο καλής προσαρμογής υδρολογικών μοντέλων (Nash and Sutcliffe, 1970). Ο εν λόγω συντελεστής ορίζεται ως:

$$R = 1 - \frac{\sum_{i=1}^n (x_i - y_i)^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad 8.22$$

όπου  $x_i$  είναι η παρατηρημένη χρονοσειρά απορροής,  $\bar{x}$  η μέση τιμή της,  $y_i$  η προσομοιωμένη χρονοσειρά απορροής και  $n$  ο χρονικός ορίζοντας ελέγχου. Ο συντελεστής προσδιορισμού λαμβάνει τιμές από  $-\infty$  έως 1. Αν  $R = 1$  υπάρχει απόλυτη ταύτιση της παρατηρημένης με την προσομοιωμένη χρονοσειρά. Προφανώς, κάτι τέτοιο δεν είναι ποτέ εφικτό για ένα μαθηματικό μοντέλο που αναπαριστά φυσικές διεργασίες και, ως εκ τούτου, υπόκειται τόσο σε δομικά σφάλματα όσο και σε σφάλματα μετρήσεων. Εφόσον  $R < 0$  η προσαρμογή του μοντέλου θεωρείται πολύ κακή, καθώς η προσομοιωμένη χρονοσειρά είναι χειρότερη εκτιμήτρια σε σχέση με την μέση παρατηρημένη τιμή  $\bar{x}$ .

### Αλγόριθμος βελτιστοποίησης

Ως μέθοδο αυτόματης βαθμονόμησης ο «Υδρογώνων» υποστηρίζει τον εξελικτικό αλγόριθμο ανόπτησης-απλόκου (evolutionary annealing-simplex method), που είναι μια ισχυρή μέθοδος μη γραμμικής βελτιστοποίησης, η οποία έχει αναπτυχθεί από ερευνητική ομάδα του ΕΜΠ (Ευστρατιάδης, 2001· Efstratiadis and Koutsoyiannis, 2002· Rozos et al., 2004). Πρόκειται για μια ευρετική τεχνική ολικής βελτιστοποίησης, που βασίζεται σε τρεις γνωστές τεχνικές βελτιστοποίησης, και συγκεκριμένα:

- μια στρατηγική εξελικτικής αναζήτησης (Duan et al., 1992), μέσω της οποίας παράγεται ένας αρχικός πληθυσμός τυχαίων εφικτών σημείων, η «μέση» καταλληλότητα του οποίου βελτιώνεται κατά τη διερεύνηση του εφικτού χώρου·
- ένα σύνολο κανόνων εξέλιξης του παραπάνω πληθυσμού, που χρησιμοποιούν ως βάση ένα γενικευμένο σχήμα κατερχόμενου απλόκου (Nelder and Mead, 1965)·
- μια τεχνική προσομοιωμένης ανόπτησης (Kirkpatrick et al., 1983), βάσει της οποίας ρυθμίζεται ο βαθμός τυχαιότητας στην αξιολόγηση των εναλλακτικών λύσεων κατά την διαδικασία αναζήτησης.

Η παραπάνω διαδικασία συνδυάζει την ευρωστία της προσομοιωμένης ανόπτωσης σε δυσχερείς χώρους αναζήτησης, με την αποτελεσματικότητα των μεθόδων κλίσης σε απλούς (κυρτούς) χώρους, παρέχοντας έτσι τη δυνατότητα εύρεσης πρόσφορων λύσεων με σχετικά μικρό αριθμό δοκιμών.

Η περιγραφή του αλγορίθμου ξεφεύγει από τα πλαίσια της παρούσης τεκμηρίωσης, ο αναγνώστης μπορεί να ανατρέξει στις δημοσιεύσεις των Ευστρατιάδη (2001), Efstratiadis and Koutsogiannis (2002) και Rozos et al (2004). Στον Πίνακα 8.2 μία παρουσίαση των παραμέτρων ελέγχου που υπαισέρχονται στον αλγόριθμο καθώς και των τιμών τους που υιοθετήθηκαν στο σύστημα του «Υδρογνώμων».

Πίνακας 8.2: Οι παράμετροι ελέγχου του εξελικτικού αλγορίθμου ανόπτωσης-απλόκου.

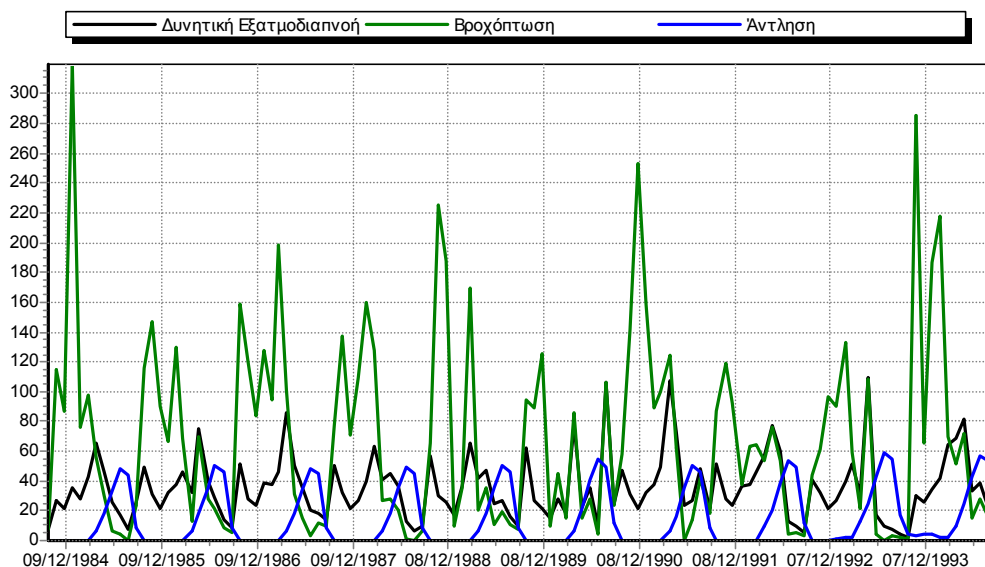
a/a	Παράμετρος	Επεξήγηση	Τιμές
1	$n$	Διάσταση του προβλήματος (αριθμός παραμέτρων του μοντέλου προς βελτιστοποίηση)	1 με 11
2	$m$	Πληθυσμός ( $m > n+1$ )	$4 \times n$
3	$x^{\min}$	Διάνυσμα κάτω ορίου παραμέτρων	
4	$x^{\max}$	Διάνυσμα άνω ορίου παραμέτρων	
5	$\sigma$	Κριτήριο σύγκλισης για πρόωρο υπολογισμό της συνάρτησης	0.005
6	$i_{\max}$	Μέγιστος αριθμός δοκιμών (υπολογισμού της αντικειμενικής συνάρτησης)	$10000 \times n$
7	$r$	Λόγος μείωσης της θερμοκρασίας, όταν έχει επιτευχθεί τοπικό ελάχιστο	0.99
8	$b$	Παράμετρος δρομολόγησης της ανόπτωσης	2
9	$p_{\text{mut}}$	Πιθανότητα αποδοχής νέας γενιάς από μετάλλαξη	0.10
10	maxclimbs	Μέγιστος αριθμός ανηφορικών βημάτων	5
11	INIFLG	Αρχικοποίηση σε συγκεκριμένο $x$	ΕΚΤΟΣ
12	RESTART	Επανάληψη της διαδικασίας αναζήτησης	ΕΚΤΟΣ

### Εκτίμηση παραμέτρων χωρίς μετρήσεις

Εφόσον δεν υπάρχουν συστηματικές μετρήσεις απορροής, η εκτίμηση των παραμέτρων ενός υδρολογικού μοντέλου είναι εξαιρετικά δύσκολη, καθώς είναι, ουσιαστικά, μη ελεγχόμενη. Στην περίπτωση αυτή, απαιτείται εξαιρετική εμπειρία από το χρήστη ώστε να προκύψουν ρεαλιστικές αποκρίσεις, που να είναι συμβατές με τις εκτιμήσεις των διαφόρων υδρολογικών μεγεθών (απορροή, εξάτμιση, κατείσδυση, διαφυγές) τουλάχιστο σε μέση ετήσια κλίμακα. Στην κατεύθυνση αυτή συμβάλλει αφενός ο περιορισμός των μεταβλητών ελέγχου του προβλήματος στο ελάχιστο δυνατό (με απλοποίηση της δομής του μοντέλου), καθώς και η εκμετάλλευση των διαθέσιμων πληροφοριών σχετικά με τα χαρακτηριστικά της λεκάνης που επηρεάζουν τις υδρολογικές διεργασίες (π.χ., γεωλογία, υδρογεωλογία, χρήσεις γης, κλπ.).

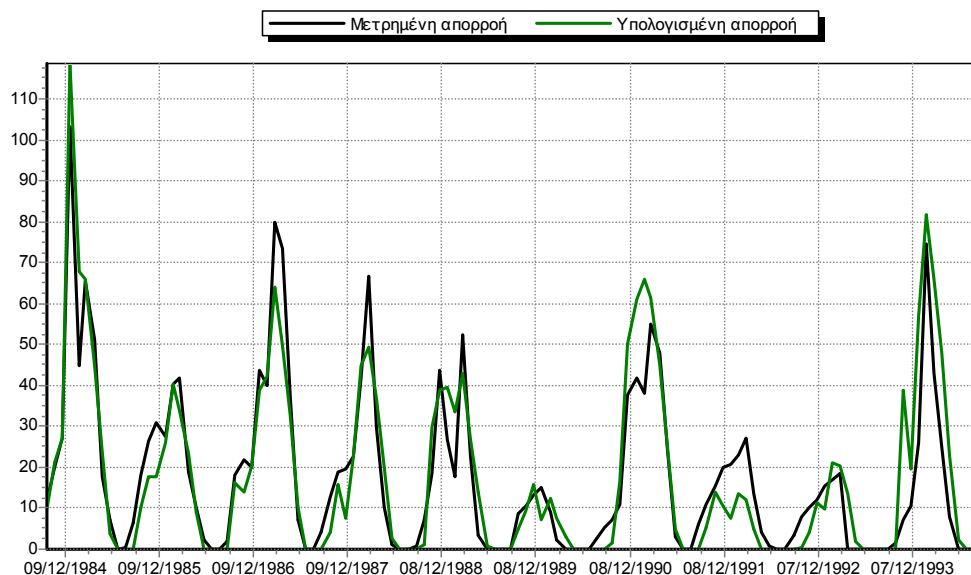
### Εφαρμογή

Στο Σχήμα 8.3 απεικονίζεται η γραφική παράσταση των μηνιαίων χρονοσειρών βροχόπτωσης, δυναμικής εξατμοδιαπνοής και αντλήσεων στην λεκάνη του Βοιωτικού Κηφισού. Οι χρονοσειρές εκτείνονται στα υδρολογικά έτη από 1984-85 έως 1993-94 και οι διαστάσεις τους είναι σε χιλιοστά. Θα επιχειρηθεί να ρυθμιστεί το μοντέλο «Ζυγός» χρησιμοποιώντας τις παραπάνω χρονοσειρές εισόδου καθώς και μία χρονοσειρά μέτρησης της απορροής. Η επιτυχής ρύθμιση δείχνει και την σωστή λειτουργία του αλγορίθμου που υλοποιεί το μοντέλο.



Σχήμα 8.3: Χρονοσειρές δυνητικής εξατμοδιαπνοής, βροχόπτωσης και αντλήσεων στην λεκάνη του Βοιωτικού Κηφισού

Η βελτιστοποίηση έγινε και στις 11 παραμέτρους του μοντέλου, έπειτα από 11000 δοκιμές επιτεύχθηκε μία τιμή της αντικειμενικής συνάρτησης ίση με 0.832. Η ρύθμιση που παρουσιάζεται είναι ενδεικτική για να δείξει τον τρόπο λειτουργίας του συστήματος. Στο Σχήμα 8.4 παρουσιάζεται η σύγκριση της μετρημένης και της υπολογισμένης απορροής. Στον Πίνακα 8.3 παρουσιάζονται τα όρια των παραμέτρων που τέθηκαν στην διαδικασία της βελτιστοποίησης καθώς και τα τελικά αποτελέσματα.



Σχήμα 8.4: Σύγκριση μεταξύ μετρημένης και υπολογισμένης απορροής στην λεκάνη του Βοιωτικού Κηφισού.

Πίνακας 8.3: Ενδεικτική ρύθμιση μοντέλου «Ζυγός», όρια παραμέτρων και τελικές τιμές. Λεκάνη Βοιωτικού Κηφισού.

Παράμετρος	$x_{\min}$	$x_{\max}$	$x_{\text{βέλτιστο}}$
$\varepsilon$	0.80	1.00	0.990
$\kappa$	0.05	0.40	0.152
$K$ (mm)	5	200	183.78
$\lambda$	0.00	1.00	0.784
$H_1$ (αδιαστατοποιημένο $H_1/K$ )	0.00	1.00	0.773
$\mu$	0.00	1.00	0.098
$\xi$	0.00	1.00	0.045
$H_2$ (mm)	5	300	97.37
$\varphi$	0.00	1.00	0.01
$S_0$ (αδιαστατοποιημένο $S_0/K$ )	0.00	1.00	0.016
$Y_0$ (mm)	5	300	128.36

## 9 Στατιστική επεξεργασία χρονοσειρών

---

### 9.1 Εισαγωγή

Η στατιστική επεξεργασία των χρονοσειρών έχει ιδιαίτερο θεωρητικό και πρακτικό ενδιαφέρον καθώς επιχειρεί να δώσει μία απάντηση στο ερώτημα: «ποιές είναι οι τιμές ενός φυσικού μεγέθους που μπορεί να παρατηρηθούν σε βάθος χρόνου;». Όταν έχουμε στην διάθεσή μας κάποιο φυσικό μοντέλο με μικρό αριθμό παραμέτρων (έτσι ώστε να έχουμε έλεγχο), είναι δυνατόν ορισμένες φορές να δοθεί η απάντηση μέσω των φυσικών ιδιοτήτων του συστήματος. Συχνά όμως τα φυσικά συστήματα είναι πολύπλοκα με ασαφείς μηχανισμούς και η απάντηση μπορεί να δοθεί μόνο με την στατιστική ανάλυση, εξετάζοντας την συμπεριφορά του συστήματος κατά το παρελθόν (μέσω των διαθέσιμων μετρήσεων).

Το πρόβλημα της στατιστικής ανάλυσης μπορεί να διατυπωθεί στα εξής διακριτά στάδια:

- Εύρεση των τυπικών στατιστικών χαρακτηριστικών του δείγματος, δηλαδή του μέτρου της θέσης (μέσω της μέσης τιμής), της διακύμανσης (μέσω της τυπικής απόκλισης – διακύμανσης ή μεταβλητότητας) καθώς και της μορφής (ασυμμετρία και κύρτωση) της τοποθέτησης των τιμών
- Χρησιμοποιώντας τα τυπικά στατιστικά χαρακτηριστικά προσαρμόζονται θεωρητικά μοντέλα στατιστικών κατανομών (με κατάλληλη προσαρμογή των παραμέτρων των κατανομών). Με αυτά τα μοντέλα μας δίνεται η δυνατότητα στατιστικής πρόγνωσης πέρα από τα όρια των μετρήσεων καθώς και η πιθανοτική προσέγγιση στα όρια διακύμανσης των φυσικών μεγεθών (μέσω πιθανοτήτων υπέρβασης ή μη-υπέρβασης των τιμών)
- Έλεγχος της καταλληλότητας των θεωρητικών κατανομών για την περιγραφή των ιδιοτήτων του πληθυσμού. Εξέταση προσαρμογής της θεωρητικής κατανομής σε εμπειρικές κατανομές, εκτέλεση στατιστικών δοκιμών
- Χρήση των θεωρητικών κατανομών για διαδικασίες στοχαστικής προσομοίωσης (με χρήση γεννητριών τυχαίων αριθμών των συγκεκριμένων κατανομών). Εφαρμογή στην εκτίμηση των διαστημάτων εμπιστοσύνης των συναρτήσεων κατανομής ή στην γένεση συνθετικών χρονοσειρών.

Όσον αφορά την χρήση συγκεκριμένων τύπων κατανομών μπορούμε να ακολουθούμε τις εξής κατευθυντήριες γραμμές και στην συνέχεια να γίνονται έλεγχοι καταλληλότητας:

- Κανονική κατανομή μπορούμε να εφαρμόσουμε σε μεγέθη που προέρχονται από συνάθροιση (ή μέσες τιμές) πάνω σε ένα χρονικό διάστημα, για παράδειγμα ετήσιες βροχοπτώσεις. Η καταλληλότητα σε αυτήν την περίπτωση τεκμηριώνεται μαθηματικά με το κεντρικό οριακό θεώρημα καθώς το άθροισμα των επιμέρους τιμών που ακολουθούν διαφορετικές κατανομές τείνει να περιγραφεί από την κανονική κατανομή. Αφού η κανονική κατανομή επεκτείνεται μέχρι το  $-\infty$  θα πρέπει να εξετάζεται το ποσοστό αρνητικών τιμών ώστε η πιθανότητα υπέρβασης της μηδενικής τιμής να είναι μεγάλη π.χ. 98% (Κουτσογιάννης, 1997 σ. 123). Ειδικά για την κανονική κατανομή ένα κριτήριο για να ικανοποιείται η απαίτηση της κυριαρχίας των θετικών τιμών είναι ο συντελεστής μεταβλητότητας ( $C_{vx}=\sigma_x/\mu_x$ ) να είναι  $C_{vx}<0.25$ . Αν όμως  $C_{vx}>0.5$  θα πρέπει να αποκλείεται η κανονική κατανομή

- Κατανομή γάμα (ή Pearson III ή και λογαριθμοκανονική 2-3 παραμέτρων) εφαρμόζουμε σε μεγέθη που παρουσιάζουν (θετική) ασυμμετρία. Τέτοια είναι π.χ. τα δείγματα που έχουν εποχικότητα (π.χ. χρονοσειρά βροχοπτώσεων συγκεκριμένου μήνα ετήσιου χρονικού βήματος)
- Εκθετική κατανομή για την περιγραφή υδρολογικών μεταβλητών σε μικρή χρονική κλίμακα
- Κατανομές ακραίων τιμών (AT) ή Log-Pearson III για την περιγραφή ακραίων τιμών σε ένα χρονικό διάστημα (όπως χρονοσειρές ετησίων μεγίστων βροχόπτωσης ή παροχής)
- AT-3 ελαχίστων (Weibull) για την περιγραφή παροχών ξηρασίας
- Pareto για την περιγραφή μεταβλητών που ξεπερνούν ένα δεδομένο κατώφλι.

Πάντα πρέπει να έχουμε υπόψη τι φυσικό μέγεθος περιγράφει η μεταβλητή και αυτό θα είναι βασικό κριτήριο για να επιλέξουμε τον τύπο της κατανομής. Δεν πρέπει σε καμία περίπτωση να θεωρείται ως βασικό κριτήριο η προσαρμογή της εμπειρικής κατανομής στην θεωρητική (πόσο δηλαδή τα σημεία των μετρήσεων «ταιριάζουν» στην καμπύλη της κατανομής), η προσαρμογή θα χρησιμοποιείται για έλεγχο καταλληλότητας. Άλλωστε η αστοχία στην προσαρμογή της κατανομής θα παρουσιαστεί στα σημεία της επέκτασης (στατιστική πρόγνωση), στην περιοχή των τιμών δηλαδή που δεν υπάρχουν διαθέσιμες μετρήσεις.

Το υποσύστημα στατιστικής επεξεργασίας χρονοσειρών του «Υδρογνώμων» (το οποίο φέρει μάλιστα την ονομασία «Πυθία») ικανοποιεί όλες τις παραπάνω απαιτήσεις, επιπλέον σε επίπεδο λογισμικού συστήματος παρέχει και το κατάλληλο περιβάλλον (user interface) για την καλύτερη εποπτεία αυτών των επεξεργασιών.

## 9.2 Εκτίμηση στατιστικών χαρακτηριστικών δείγματος

Οι τιμές μίας χρονοσειράς αποτελούν δείγμα ενός πληθυσμού. Έτσι αφού σχηματίσουμε ένα δείγμα  $x_i$  από τις  $n$  τιμές της χρονοσειράς (δεν λαμβάνονται υπόψη οι ελλείπουσες τιμές) μπορούμε να προχωρήσουμε στον προσδιορισμό των στατιστικών χαρακτηριστικών του δείγματος με την μέθοδο των ροπών ή την μέθοδο των L-ροπών, ώστε να είναι δυνατή τελικά η προσαρμογή των θεωρητικών κατανομών στο δείγμα.

Το δείγμα μπορεί να αποτελείται από τις τιμές χρονοσειρές σε τυχαία σειρά καθώς δεν θα εξετάζονται στα πλαίσια της στατιστικής ανάλυσης φαινόμενα όπως η αυτοσυσχέτιση, η εμμονή, η στασιμότητα, κλπ. Συγκεκριμένα από τις ιδιότητες της στοχαστικής ανέλιξης (που τέτοια αποτελεί η χρονοσειρά) θα εξετάζονται μόνο τα στατιστικά χαρακτηριστικά και οι συναρτήσεις κατανομής.

Εφόσον η χρονοσειρά παρουσιάζει κάποια περιοδικότητα (π.χ. εποχικότητα), θα πρέπει να καταρτίζονται ιδιαίτερα δείγματα ώστε να απαλείφεται αυτή. Έτσι σε μία μηνιαία χρονοσειρά βροχοπτώσεων, θα καταρτίζονται 12 διαφορετικά δείγματα, ένα για κάθε μήνα. Θα πρέπει να παραμένει φυσικά η δυνατότητα εξέτασης του πλήρους δείγματος της χρονοσειράς καθώς δεν παρουσιάζουν όλα τα μεγέθη εποχικότητα. Όσον αφορά την ετήσια κλίμακα θα χρησιμοποιείται το υδρολογικό έτος σε υδρολογικά μεγέθη όπως η βροχόπτωση ώστε οι τιμές της μεταβλητής να έχουν στατιστική ανεξαρτησία και την ελάχιστη αυτοσυσχέτιση (βλ. Κουτσογιάννης 1997, σ. 87).

Σχετικά με το χρονικό βήμα των χρονοσειρών από όπου θα παράγονται τα δείγματα θα παρέχεται το μηνιαίο και το ετήσιο χρονικό βήμα. Τα μεγέθη που έχουν ετήσιο βήμα θα έχουν την δυνατότητα να αναφέρονται είτε σε ημερολογιακό είτε σε υδρολογικό έτος.

### 9.2.1 Αντιμετώπιση των αρνητικών τιμών του δείγματος

Οι τυπικές υδρολογικές μεταβλητές έχουν θετικό πεδίο τιμών. Αυτό ισχύει ειδικότερα για τις μεταβλητές που συναθροίζονται στον χρόνο όπως η βροχόπτωση, η ένταση βροχόπτωσης, η απορροή,



η εξατμοδιαπνοή, η διάρκεια ηλιοφάνειας, το ποσοστό της σχετικής υγρασίας, η ηλιακή ακτινοβολία, κ.α. Είναι δυνατό από σφάλματα που εισάγονται στους υπολογισμούς (π.χ. αρνητικός σταθερός όρος κατά την συμπλήρωση με γραμμική παλινδρόμηση) να εξάγονται αρνητικές τιμές αυτών των μεγεθών. Σε αυτές τις περιπτώσεις μπορεί να εφαρμοστεί η αποκοπή των τιμών στο μηδέν καθώς οι αρνητικές τιμές δεν έχουν κάποιο φυσικό νόημα, ωστόσο η αποκοπή δεν πρέπει να γίνεται άκριτα καθώς μία μεγάλη αρνητική τιμή μπορεί να αντιπροσωπεύει κάποιο χονδροειδές σφάλμα και ενδεχομένως θα πρέπει να εξαιρείται από το δείγμα.

Έτσι ο «Υδρογνώμων» παρέχει τις εξής δυνατότητες ως προς τις αρνητικές τιμές του δείγματος:

- Διατήρηση των τιμών ως έχουν
- Αποκοπή στο μηδέν:  $x_i < 0 \Rightarrow x_i = 0$ .

Η τελευταία επιλογή κρίνεται καλύτερη από την εναλλακτική του αποκλεισμού τιμών, καθώς αλλοιώνει λιγότερο τα στατιστικά χαρακτηριστικά του δείγματος.

Ωστόσο θα πρέπει να εξετάζεται κατά περίπτωση από τον χρήστη ο αποκλεισμός των μεγάλων αρνητικών τιμών του δείγματος ως ένα πρόβλημα ακραίων τιμών (βλ. Κεφάλαιο 3).

Όσον αφορά μεταβλητές όπως η θερμοκρασία (ειδικά οι χρονοσειρές ελάχιστης θερμοκρασίας) και η διεύθυνση του ανέμου όπου το πεδίων τιμών μπορεί να είναι συνεχές και να προβλέπονται αρνητικές τιμές, θα πρέπει να εξετάζεται η εισαγωγή σταθερού όρου ώστε οι τιμές των χρονοσειρών να μετασχηματίζονται σε θετικές. Για παράδειγμα η θερμοκρασία μπορεί να εκφραστεί σε °K με την προσθήκη σταθερού όρου 273 στους °C ώστε οι τιμές να είναι πάντα θετικές.

## 9.2.2 Οι δειγματικές ροπές ως στατιστικοί δείκτες

Οι δειγματικές ροπές χρησιμοποιούνται για την εκτίμηση των παραμέτρων των κατανομών με την μέθοδο των ροπών η οποία είναι η πλέον συνηθισμένη αντιμετώπιση (θα εξεταστεί παρακάτω και η εφαρμογή της μεθόδου των L-ροπών). Στον Πίνακα 9.1 δίνονται οι μεροληπτικές εκτιμήτριες των δειγματικών ροπών. Στην συνέχεια θα δοθούν συντελεστές για την διόρθωση της μεροληψίας.

Πίνακας 9.1: Απλές (μεροληπτικές) εκτιμήτριες των δειγματικών ροπών μέχρι τρίτης τάξης

Στατιστικός δείκτης	Εκτιμήτρια
Μέση τιμή	$\mu_x = \frac{\sum x_i}{n}$
Τυπική απόκλιση	$\sigma_x = \sqrt{\frac{\sum (x_i - \mu_x)^2}{n}}$
Τρίτη κεντρική ροπή	$\mu_x^{(3)} = \frac{1}{n} \sum x_i^3 - 3\mu_x \sigma_x^2 - \mu_x^3$
Συντελεστής ασυμμετρίας	$C_{sx} = \frac{\mu_x^{(3)}}{\sigma_x^3}$

Για την περίπτωση της εκτίμησης παραμέτρων κατανομών με την έμμεση μέθοδο των ροπών (περίπτωση κατανομής Log-Pearson III) θα χρησιμοποιείται η μετασχηματισμένη μεταβλητή:  $y_i = \ln x_i$  στην θέση της  $x_i$ . Σε αυτήν την περίπτωση δεν θα λαμβάνονται υπόψη, αναγκαστικά, οι μη-θετικές τιμές της  $x_i$  λόγω της ύπαρξης του λογαρίθμου.

## Συντελεστές διόρθωσης μεροληψίας

Οι αμερόληπτες εκτιμήσεις των δειγματικών ροπών μπορούν να προκύψουν από τις αντίστοιχες μεροληπτικές πολλαπλασιάζοντας τις με τους συντελεστές διόρθωσης μεροληψίας του Πίνακα 9.2. Οι αμερόληπτες εκτιμήτριες είναι αυτές που προτιμώνται για την περαιτέρω εκτίμηση των παραμέτρων των κατανομών.

Πίνακας 9.2: Συντελεστές διόρθωσης μεροληψίας

Στατιστικός δείκτης	Συντελεστής διόρθωσης μεροληψίας
Μέση τιμή	-
Τυπική απόκλιση	$\sqrt{\frac{n}{n-1}}$
Τρίτη κεντρική ροπή	$\frac{n^2}{(n-1)(n-2)}$
Συντελεστής ασυμμετρίας	$\frac{\sqrt{n(n-1)}}{n-2}$

Όσον αφορά τον συντελεστή διόρθωσης για την ασυμμετρία είναι η προσέγγιση που προκύπτει αν αντικατασταθούν οι μεροληπτικές εκτιμήτριες των ροπών 2<sup>ης</sup> και 3<sup>ης</sup> τάξης από τις αντίστοιχες αμερόληπτες. Αυτή είναι μία προσέγγιση η οποία φυσικά δεν οδηγεί σε αυστηρώς αμερόληπτη εκτιμήτρια του συντελεστή ασυμμετρίας (Κουτσογιάννης 1997, σ. 59) ωστόσο καλύπτει τις απαιτήσεις του συστήματος.

### 9.2.3 Η χρήση των 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 κ.λ.π.). 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).

Από την εφαρμογή της μεθόδου των L-ροπών για την περιγραφή των στατιστικών χαρακτηριστικών ενός δείγματος καθώς και την εκτίμηση των παραμέτρων κατανομών προκύπτουν μία σειρά από πλεονεκτήματα όπως:

- Καλύτερη προσαρμογή των κατανομών ακροτάτων (π.χ. AT-1, AT-2, AT-3, ΓΑΤ) στο δείγμα
- Μικρή ευαισθησία σε εξωκείμενες τιμές (outliers values)
- Απλή υπολογιστική διαδικασία στην εξαγωγή παραμέτρων κατανομών.

Από τα παραπάνω φαίνεται πως πλεονεκτεί ιδιαίτερα η χρήση τους στις κατανομές AT, διότι μεταξύ άλλων η εφαρμογή των κλασσικών ροπών σε αυτές τις κατανομές απαιτεί προσεγγιστικές ή αριθμητικές λύσεις.

Οι L-ροπές θα συμβολίζονται με το ελληνικό γράμμα  $\lambda$  και έναν δείκτη 1, 2, 3, ... Η  $\lambda_1$  σχετίζεται με την παράμετρο θέσης και ταυτίζεται με την κλασσική μέση τιμή (ροπή πρώτης τάξης). Η  $\lambda_2$  σχετίζεται με την μεταβλητότητα, η  $\lambda_3$  με την ασυμμετρία και η  $\lambda_4$  με την κύρτωση. Εφόσον οι L-ροπές είναι γραμμικοί συνδυασμοί των τιμών του δείγματος, η διάσταση τους είναι η ίδια με αυτή του φυσικού μεγέθους του οποίου αντιπροσωπεύουν τα στατιστικά χαρακτηριστικά του. Ωστόσο μπορούμε να ορίσουμε τους εξής αδιάστατους συντελεστές:

- Συντελεστής μεταβλητότητας:  $\tau_2 = \lambda_2/\lambda_1$

- Συντελεστής ασυμμετρίας:  $\tau_3 = \lambda_3/\lambda_2$
- Συντελεστής μεταβλητότητας:  $\tau_4 = \lambda_4/\lambda_2$

### Υπολογιστική διαδικασία

Οι L-ροπές υπολογίζονται εύκολα αν έχουν υπολογιστεί προηγουμένως οι πιθανοτικά σταθμισμένες ροπές – Π.Σ.Ρ. (Probability-Weighted Moment, PWM). Οι Π.Σ.Ρ. εισήχθησαν από τους Greenwood et al. (1979) και υπολογίζονται ως γραμμικοί συνδυασμοί των τιμών ενός ταξινομημένου δείγματος.

Αν είναι γνωστές οι ΠΣΡ  $\beta_0, \beta_1, \beta_2, \beta_3$ , τότε οι L-ροπές μπορούν να υπολογιστούν από τις σχέσεις:

$$\lambda_1 = \beta_0 \quad 9.1$$

$$\lambda_2 = 2\beta_1 - \beta_0 \quad 9.2$$

$$\lambda_3 = 6\beta_2 - 6\beta_1 + \beta_0 \quad 9.3$$

$$\lambda_4 = 20\beta_3 - 30\beta_2 + 12\beta_1 - \beta_0 \quad 9.4$$

Όπως βλέπουμε οι L-ροπές είναι γραμμικοί συνδυασμοί των Π.Σ.Ρ., συνεπώς διατηρούν την ιδιότητά τους να είναι και γραμμικοί συνδυασμοί των τιμών του δείγματος. Η εκτίμηση των Π.Σ.Ρ. μπορεί να γίνει είτε μεροληπτικά είτε αμερόληπτα. Η μεροληπτική εκτίμηση έχει μικρή εφαρμοσιμότητα ωστόσο πρέπει να παρέχεται από τον «Υδρογνώμων». Για την αμερόληπτη εκτίμηση της  $i$  τάξης Π.Σ.Ρ.  $\beta_i$ , απαιτούνται τουλάχιστον  $i+1$  τιμές στο δείγμα. Στον «Υδρογνώμων», μας ενδιαφέρουν οι υπολογισμοί των τριών πρώτων Π.Σ.Ρ. καθώς και των τριών πρώτων L-ροπών καθώς δεν απαιτείται η κύρτωση για την εκτίμηση των παραμέτρων των κατανομών του συστήματος.

Εφόσον το δείγμα είναι ταξινομημένο σε αύξουσα σειρά τιμών έτσι ώστε  $x_{(j)} \geq x_{(i)}, \forall j > i$  οι μεροληπτικές εκτιμήσεις είναι:

$$\beta_{0\mu} = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad 9.5$$

$$\beta_{1\mu} = \sum_{i=1}^n \left(1 - \frac{n-i-0.65}{n}\right) x_{n-i+1} \quad 9.6$$

$$\beta_{2\mu} = \sum_{i=1}^n \left(1 - \frac{n-i-0.65}{n}\right)^2 x_{n-i+1} \quad 9.7$$

Αντίστοιχα οι αμερόληπτες εκτιμήσεις των Π.Σ.Ρ. είναι:

$$\beta_0 = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad 9.8$$

$$\beta_1 = \sum_{i=2}^n \frac{i-1}{n-1} x_{n-i+1} \quad 9.9$$

$$\beta_2 = \sum_{i=3}^n \frac{(i-1)(i-2)}{(n-1)(n-2)} x_{n-i+1} \quad 9.10$$

### 9.3 Οι εμπειρικές συναρτήσεις κατανομής

Οι εμπειρικές συναρτήσεις κατανομής μας δίνουν την δυνατότητα να εκτιμήσουμε την συνάρτηση κατανομής (πιθανότητα μη-υπέρβασης) για κάθε τιμή του δείγματος χωρίς την χρήση κάποιου

θεωρητικού μοντέλου. Μέσω της εκτίμησής της έχουμε στην συνέχεια την δυνατότητα ελέγχου της προσαρμογής κάποιας θεωρητικής συνάρτησης κατανομής στο δείγμα.

Ο «Υδρογνώμων» παρέχει την δυνατότητα εκτίμησης των εμπειρικών κατανομών κατά Weibull, Blom, Cunnane και Gringorten (Κουτσογιάννης, 1997). Η εμπειρική κατανομή που είναι καταρχήν ενεργοποιημένη είναι αυτή του Weibull, η οποία έχει την απλούστερη σχέση και είναι γνωστή στην βιβλιογραφία ως *αμερόληπτη ως προς την πιθανότητα υπέρβασης θέση σχεδίασης*.

Αν  $X_1, X_2, X_3, \dots, X_n$  οι τιμές του δείγματος, ορίζουμε μία ακολουθία που προκύπτει αν ταξινομηθεί το δείγμα σε αύξουσα αριθμητική σειρά τέτοια ώστε:  $X_{(1)}, X_{(2)}, X_{(3)}, \dots, X_{(n)}$  και  $X_{(1)} \leq X_{(2)} \leq X_{(3)} \leq \dots \leq X_{(n)}$ .

Τότε η πιθανότητα μη υπέρβασης (συνάρτηση κατανομής) για κάθε τιμή  $X_{(i)}$  μπορεί να εκφραστεί με μία συνάρτηση  $f$  της μορφής:

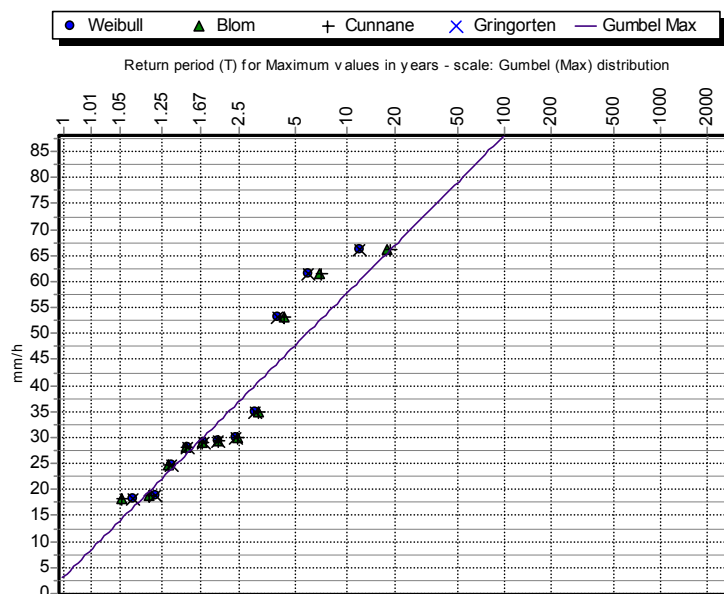
$$F(X_{(i)}) = q_i = f(i, n) \quad 9.11$$

όπου  $i$  είναι η σειρά της τιμής  $X_{(i)}$  στο διατεταγμένο δείγμα και  $n$  το μέγεθος του δείγματος ( $i=1 \dots n$ ). Οι σχέσεις των εμπειρικών συναρτήσεων κατανομής παρουσιάζονται στον Πίνακα 9.3.

Πίνακας 9.3: Οι εμπειρικές συναρτήσεις κατανομής (πιθανότητες μη-υπέρβασης ταξινομημένου δείγματος σε αύξουσα σειρά) που χρησιμοποιούνται στον «Υδρογνώμων»

Όνομα κατανομής	Σχέση $q_i =$
Weibull	$\frac{i}{n+1}$
Blom	$1 - \frac{(n-i+1) - 0.375}{n+0.25}$
Cunnane	$1 - \frac{(n-i+1) - 0.4}{n+0.2}$
Gringorten	$1 - \frac{(n-i+1) - 0.44}{n+0.12}$

(Κουτσογιάννης 1997, σ. 117)



Σχήμα 9.1: Εμπειρικές κατανομές Weibull, Blom, Cunnane, Gringorten. Ετήσια μέγιστα έντασης οριαίων βροχοπτώσεων. Σταθμός «Πολυτεχνειούπολης» 1993-4 – 2003-4

## 9.4 Οι θεωρητικές συναρτήσεις κατανομής

Η χρήση των θεωρητικών συναρτήσεων κατανομής έχει μία πληθώρα εφαρμογών όπως η στατιστική πρόγνωση (βλ. ενότητα 9.4.2) και η στοχαστική προσομοίωση. Στον χρήστη παρέχονται μία σειρά από κατανομές που ταξινομούνται σε δύο κύριες οικογένειες (οικογένεια κατανομών κανονικής κατανομής και κατανομής γάμμα) όπως παρουσιάζονται στον Πίνακα 9.4. Για τους παραπάνω σκοπούς πρέπει να παρέχονται από τον «Υδρογνώμων» οι συναρτήσεις κατανομής (τιμή τυχαίας μεταβλητής ως συνάρτηση πιθανότητας μη-υπέρβασης) καθώς και οι αντίστροφες συναρτήσεις κατανομής (τιμή πιθανότητας ως συνάρτηση της τυχαίας μεταβλητής).

Στην ενότητα αυτή παρουσιάζονται οι υπολογισμοί για τις παραπάνω συναρτήσεις ακριβώς όπως υλοποιούνται στον «Υδρογνώμων». Στην συνέχεια παρουσιάζεται η μεθοδολογία υπολογισμού των παραμέτρων των κατανομών. Οι παράμετροι των κατανομών εκτιμώνται από τα στατιστικά στοιχεία του δείγματος είτε όπως έχουν προκύψει από την μέθοδο ροπών είτε από την μέθοδο των L-ροπών.

Ειδικά για την κατανομή Log-Pearson III όπως θα δούμε, η εκτίμηση των παραμέτρων γίνεται μέσω της έμμεσης μεθόδου των ροπών, εφαρμόζοντας δηλαδή την μέθοδο των ροπών σε ένα μετασχηματισμένο (με λογαρίθμους) δείγμα.

Όταν υπεισέρχονται παράμετροι σχήματος, κλίμακας και θέσης, χρησιμοποιούνται οι συμβολισμοί  $\kappa$ ,  $\lambda$  και  $\psi$ . Εν γένει το πρόσημο του  $\kappa$  συσχετίζεται με το πρόσημο της ασυμμετρίας της κατανομής, το δε  $\lambda$  είναι θετικό. Επιπλέον όταν μπαίνει ως παράμετρος η τυπική απόκλιση  $\sigma$ , αυτή είναι θετική εξορισμού.

Στον Πίνακα 9.5 παρουσιάζονται τα συνεχή πεδία ορισμού της τυχαίας μεταβλητής ανά κατανομή. Κάποιες από τις κατανομές έχουν ως πεδίο ορισμού όλο το  $\mathbb{R}$ , αρκετές κατανομές ωστόσο έχουν ως κάτω φράγμα της τυχαίας μεταβλητής το 0 ή κάποια άλλη τιμή που εξαρτάται από τις παραμέτρους της κατανομής. Οι δε κατανομές ΓΑΤ και Pareto μπορούν να παρουσιάσουν και άνω φράγμα ανάλογα τις τιμές της παραμέτρου σχήματος  $\kappa$ .

Πίνακας 9.4: Οι συναρτήσεις κατανομής που περιλαμβάνονται στον «Υδρογνώμων» μαζί με τον αριθμό των παραμέτρων

α/α	Όνομασία Κατανομής	Όνομασία Κατανομής (Αγγλική)	Παράμετροι	Ανεξάρτητων Παραμέτρων Αριθμ.	Μέθοδος Ροπών	Μέθοδος λ-Ροπών	Έμφαση μεθόδους Ροπών
<i>Οικογένεια κανονικής κατανομής</i>							
1	Κανονική ή Gauss	Normal or Gauss	$\mu, \sigma$	2	✓	✓	
2	Λογαριθμο-κανονική	Log-Normal	$\mu_y, \sigma_y$	2	✓		
3	Λογαριθμο-κανονική 3 παραμέτρων	Log-Normal 3p (Galton)	$\mu_y, \sigma_y, c$	3	✓		
<i>Οικογένεια κατανομής Γ</i>							
4	Εκθετική 2 παραμέτρων	Exponential	$\lambda, \psi$	2	✓	✓	
5	Γάμα 2 παραμέτρων	Gamma 2p	$\kappa, \lambda$	2	✓		
6	Γάμα 3 παραμέτρων	Gamma 3p (Pearson III)	$\kappa, \lambda, \psi$	3	✓		
7	Λογαριθμική Γάμμα 3 παραμέτρων	Log-Pearson III	$\kappa, \lambda, \psi$	3			✓
<i>Ασυμπτωτικές κατανομές ακροτάτων</i>							
8	Μεγίστων, Ακραίων τιμών τύπου I (AT-1) ή Gumbel	Maximum, Extreme values type I (EV-1) or Gumbel	$\lambda, \psi$	2	✓	✓	
9	Μεγίστων, AT-2	Maximum, EV-2	$\kappa, \lambda$	2	✓	✓	
10	Μεγίστων, Γενική Ακραίων τιμών (ΓΑΤ)	Maximum, General Extreme values (GEV)	$\kappa, \lambda, \psi$	3 ή 2 (*)	✓	✓	
11	Ελαχίστων, AT-1 ή Gumbel	Minimum, EV-1 or Gumbel	$\lambda, \psi$	2	✓	✓	
12	Ελαχίστων, AT-3 ή Weibull	Minimum, EV-2 or Weibull	$\kappa, \lambda$	2	✓	✓	
13	Ελαχίστων, Γενική Ακραίων τιμών (ΓΑΤ)	Minimum, General Extreme values (GEV)	$\kappa, \lambda, \psi$	3 ή 2 (*)	✓	✓	
<i>Pareto</i>							
14	Pareto 3 παραμέτρων	Pareto 3p	$\kappa, \lambda, \psi$	3	✓	✓	

(\*) Οι κατανομές τύπου ΓΑΤ έχουν 3 παραμέτρους. Είτε προσδιορίζονται και οι τρεις με την μέθοδο των ροπών ή την μέθοδο των L-ροπών (περίπτωση 3 ανεξάρτητων παραμέτρων) είτε δίνεται σταθερή τιμή στην παράμετρο σχήματος ( $\kappa$ ) και υπολογίζονται οι υπόλοιπες δύο παράμετροι κατά τα γνωστά (περίπτωση 2 ανεξάρτητων παραμέτρων).

Η δυνατότητα εφαρμογής κάποιας συνάρτησης κατανομής εξαρτάται από πολλούς διαφορετικούς παράγοντες (βλ. Κουτσογιάννης 1997, σ.120). Αναφέρουμε επιγραμματικά την πιθανοθεωρία, την

γενική υδρολογική εμπειρία καθώς και την εξέταση του συγκεκριμένου δείγματος (π.χ. ένα δείγμα με μικρή ασυμμετρία μπορεί να περιγραφεί από την κανονική κατανομή ή κάποια άλλη συμμετρική κατανομή). Επιπλέον μέσω της οπτικοποίησης των σημείων των εμπειρικών κατανομών είναι δυνατός ο εμπειρικός έλεγχος για την προσαρμογή της κατανομής στο δείγμα. Τέλος ο «Υδρογνώμων» παρέχει κάποια εργαλεία στατιστικών ελέγχων (Δοκιμή  $\chi^2$  και δοκιμή Kolmogorov-Smirnov) για τον έλεγχο υποθέσεων απόρριψης κάποιας κατανομής.

Πίνακας 9.5: Πεδίο ορισμού της τυχαίας μεταβλητής ανά κατανομή

a/a	Ονομασία Κατανομής	Παράμετροι	Κάτω όριο	Άνω όριο
1	Κανονική	$\mu, \sigma$	$-\infty$	$\infty$
2	Λογαριθμο-κανονική	$\mu_y, \sigma_y$	0	$\infty$
3	Λογαριθμο-κανονική 3 παραμέτρων	$\mu_y, \sigma_y, c$	$c$	$\infty$
4	Εκθετική 2 παραμέτρων	$\lambda, \psi$	$\psi$	$\infty$
5	Γάμα 2 παραμέτρων	$\kappa, \lambda$	0	$\infty$
6	Γάμα 3 παραμέτρων	$\kappa, \lambda, \psi$	$\psi$	$\infty$
7	Λογαριθμική Γάμμα 3 παραμέτρων	$\kappa, \lambda, \psi$	$e^\psi$	$\infty$
8	Μεγίστων ή Ελαχίστων, Ακραίων τιμών τύπου I (AT-1) ή Gumbel	$\lambda, \psi$	$-\infty$	$\infty$
9	Μεγίστων, AT-2	$\kappa, \lambda$	0	$\infty$
10	Μεγίστων ή Ελαχίστων, Γενική Ακραίων τιμών (ΓΑΤ)	$\kappa, \lambda, \psi$	$\lambda\left(\psi - \frac{1}{\kappa}\right)$ για $\kappa > 0$ , $-\infty$ για $\kappa \leq 0$	$\lambda\left(\psi - \frac{1}{\kappa}\right)$ για $\kappa < 0$ , $\infty$ για $\kappa \geq 0$
12	Ελαχίστων, AT-3 ή Weibull	$\kappa, \lambda$	0	$\infty$
14	Pareto 3 παραμέτρων	$\kappa, \lambda, \psi$	$\lambda\psi$	$\lambda(\psi + \kappa)$ για $\kappa > 0$ , $\infty$ για $\kappa \leq 0$

#### 9.4.1 Υπολογισμοί συναρτήσεων κατανομής

Οι υπολογισμοί των συναρτήσεων κατανομής (πιθανότητα μη-υπέρβασης  $F(x)$  συναρτήσεως της τυχαίας μεταβλητής  $x$ ) καθώς και των αντιστρόφων συναρτήσεων κατανομής (τυχαία μεταβλητή  $x$  ως συνάρτηση του ποσοστημορίου  $u=F(x)$ ,  $x(u)=F^{-1}(u)$ ) διακρίνονται σε:

- Υπολογισμούς με μαθηματικές σχέσεις που έχουν κλειστή μορφή και
- Υπολογισμούς με προσεγγίσεις της αριθμητικής ανάλυσης.

Οι υπολογισμοί των συναρτήσεων κατανομής της οικογένειας της κανονικής κατανομής καθώς και της κατανομής Γάμα βασίζονται σε αριθμητικές προσεγγίσεις: οι υπολογισμοί στις υπόλοιπες κατανομές (Ακραίων τιμών, εκθετική κατανομή και κατανομή Pareto) γίνονται με σχέσεις κλειστής μορφής. Η διάκριση αυτή παρουσιάζεται στον Πίνακα 9.6.

Στην συνέχεια, η πιθανότητα υπέρβασης υπολογίζεται από την συνάρτηση κατανομής με την χρήση της σχέσης:

$$F_1(x) = 1 - F(x) \quad 9.12$$

Εφόσον το δείγμα το οποίο εξετάζεται αποτελείται από τιμές με τακτικά χρονικά διαστήματα δειγματοληψίας, η πιθανότητα υπέρβασης μπορεί να εκφραστεί με την τιμή της συχνότητας παρουσίας της τυχαίας μεταβλητής  $x$  ως:

$$T = 1/F_1(x) \quad 9.13$$

Το μέγεθος  $T$  καλείται *περίοδος επαναφοράς*, αν η περίοδος δειγματοληψίας είναι το έτος τότε και η περίοδος επαναφοράς είναι εκφρασμένη σε έτη, αφορά δε την συχνότητα υπέρβασης του μεγέθους  $x$ . Η  $T$  παρουσιάζεται στον «Υδρογνώμων» για λόγους διάκρισης και ως  $T(\max)$  καθώς είναι δυνατό να εκφραστεί και η *συχνότητα μη-υπέρβασης ενός μεγέθους* ως:

$$T(\min) = 1/F(x) = [1 - 1/T(\max)]^{-1}$$

Πίνακας 9.6: Υπολογισμοί συναρτήσεων κατανομής με κλειστή ή προσεγγιστική λύση

<b>Κατανομές που υπολογίζονται με προσεγγιστική λύση</b>	
<b>Κατανομές</b>	<b>Βασισμένες στον</b>
Κανονική κατανομή, Λογαριθμοκανονική κατανομή, κατανομή Galton	Υπολογισμό της συνάρτησης σφάλματος (Error function) με αριθμητική μέθοδο
Κατανομή Γάμα 2 παραμέτρων, κατανομή Γάμα 3 παραμέτρων (Person III), Log-Pearson III	Υπολογισμό της κατανομής Γάμα με αριθμητική μέθοδο
<b>Κατανομές που υπολογίζονται με λύση κλειστής μορφής</b>	
Εκθετική κατανομή Κατανομές AT1, AT2 μεγίστων και AT1, AT3 ελαχίστων Κατανομές ΓΑΤ ελαχίστων και μεγίστων Κατανομή Pareto τριών παραμέτρων	

### Οι υπολογισμοί κανονικής κατανομής και της συνάρτησης σφάλματος (Error function)

Η συνάρτηση πυκνότητας πιθανότητας της κανονικής κατανομής με παραμέτρους  $\mu$ ,  $\sigma$  ορίζεται ως (Κουτσογιάννης, 1997):

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad 9.14$$

Η συνάρτηση κατανομής προκύπτει από την ολοκλήρωση της (9.14) ως:

$$F(x) = \int_{-\infty}^x f(t) dt = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt \quad 9.15$$

Το παραπάνω ολοκλήρωμα μπορεί να υπολογιστεί εύκολα αν είναι γνωστή η τιμή των συναρτήσεων σφάλματος  $Erf(x)$  και  $Erfc(x)$ :

$$Erf(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \quad 9.16$$

$$Erfc(x) = 1 - Erf(x) = \frac{2}{\sqrt{\pi}} \int_x^{+\infty} e^{-t^2} dt$$

Στην συνέχεια η κανονική κατανομή θα υπολογίζεται ως:



$$F(x) = \begin{cases} \frac{1}{2} + \frac{1}{2} \operatorname{Erf}\left(\frac{z}{\sqrt{2}}\right) & z \geq 0 \\ \frac{1}{2} \operatorname{Erfc}\left(\frac{-z}{\sqrt{2}}\right) & z < 0 \end{cases} \quad 9.17$$

όπου  $z$  η τυποποιημένη μεταβλητή:

$$z = \frac{x - \mu}{\sigma} \quad 9.18$$

Οι υπολογισμοί των  $\operatorname{Erf}(x)$ ,  $\operatorname{Erfc}(x)$  περιγράφονται από τον Miller (Miller, 1981, σελ. 288) και παρουσιάζονται στην συνέχεια.

### Αριθμητικός υπολογισμός $\operatorname{Erf}(x)$ για $0 < x \leq 1.5$

Ορίζεται κατώφλι ακρίβειας  $\varepsilon$ .

Για την περίπτωση του συστήματός μας θέτουμε ως μία αποδεκτή τιμή ακρίβειας  $\varepsilon = 10^{-10}$ .

Αν  $x < \varepsilon$  τότε  $\operatorname{Erf}(x) = 0$

Ορίζουμε τις αναδρομικές σχέσεις  $\alpha_i$ ,  $\gamma_i$ , ως εξής:

$$\alpha_0 = \gamma_0 = x \quad 9.19$$

$$\gamma_{i+1} = 2\gamma_i \frac{x^2}{1+2i} \quad 9.20$$

$$\alpha_{i+1} = \alpha_i + \gamma_{i+1} \quad 9.21$$

Και ως κριτήριο σύγκλισης τίθεται:

$$\gamma_i < \varepsilon \alpha_i \quad 9.22$$

Αν  $i = n_0$  το σημείο σύγκλισης, τελικά η συνάρτηση σφάλματος δίνεται από την σχέση:

$$\operatorname{Erf}(x) = 2\alpha_{n_0} \frac{e^{-x^2}}{\sqrt{\pi}} \quad (0 < x \leq 1.5) \quad 9.23$$

### Αριθμητικός υπολογισμός $\operatorname{Erf}(x)$ για $x > 1.5$

Ορίζεται ένα ακέραιο όριο  $l$ .

Για την περίπτωση του συστήματός μας τίθεται  $l = 12$ .

Ορίζεται η παρακάτω αναδρομική σχέση  $a_i$  ως εξής:

$$a_l = 1 + \frac{l+1}{2x^2} \quad 9.24$$

$$a_{i-1} = 1 + \frac{i}{2a_i x^2} \quad 9.25$$

Η παραπάνω επαναληπτική διαδικασία ξεκινάει από  $i = l$  μέχρι να υπολογιστεί το  $a_0$ . Τελικά η συνάρτηση σφάλματος δίνεται από την σχέση:

$$Erf(x) = 1 - Erfc(x) = 1 - \frac{e^{-x^2}}{a_0 x \sqrt{\pi}} \quad (x > 1.5) \quad 9.26$$

### Αριθμητικός υπολογισμός της $F^{-1}$ της κανονικής κατανομής

Για τον υπολογισμό της αντίστροφης συνάρτησης κατανομής (συνάρτηση ποσοστημορίου) της κανονικής κατανομής, μπορούμε να εφαρμόσουμε την μέθοδο Newton – Raphson:

$$x_{n+1} = x_n - \frac{G(x_n)}{G'(x_n)} \quad 9.27$$

όπου  $x_i$  μία αναδρομική ακολουθία που συγκλίνει στην τυχαία μεταβλητή που αντιστοιχεί στην πιθανότητα μη υπέρβασης  $u$ . Η συνάρτηση  $G$  ορίζεται μέσω της συνάρτησης κατανομής και της  $u$  ως:

$$G = F(x) - u \quad 9.28$$

Η  $F(x)$  υπολογίζεται αριθμητικά σύμφωνα με την μεθοδολογία που αναπτύχθηκε στις προηγούμενες παραγράφους. Η δε παράγωγος της  $G$ , η  $G'$ , συμπίπτει εξορισμού με την συνάρτηση πυκνότητας πιθανότητας της κανονικής κατανομής (9.14).

Ως αρχική τιμή για την  $x_i$  τίθεται η δειγματική μέση τιμή  $\mu$ :

$$x_0 = \mu \quad 9.29$$

Η ακολουθία (9.27) συγκλίνει μόλις ικανοποιηθεί ένα από τα παρακάτω κριτήρια:

$$|x_i - x_{i-1}| \leq \varepsilon \quad 9.30$$

$$i > i_{\max} \quad 9.31$$

Για τον «Υδρογνώμων» τίθενται οι τιμές  $\varepsilon = 10^{-8}$  και  $i_{\max} = 25$ .

Η λύση του συγκεκριμένου προβλήματος με την μέθοδο Newton-Raphson είναι μονοσήμαντη καθώς η  $F(x)$  είναι μονότονη (αύξουσα) στο πεδίο ορισμού της.

Οι υπόλοιπες συναρτήσεις κατανομής της οικογένειας της κανονικής κατανομής βασίζονται στους παραπάνω υπολογισμούς κατόπιν κατάλληλων μετασχηματισμών και εισαγωγής παραμέτρων. Θα παρουσιαστούν παρακάτω σε ξεχωριστές παραγράφους.

Πίνακας 9.7: Επαλήθευση αλγορίθμων μέσω των τιμών της (τυποποιημένης) κανονικής κατανομής

Τυποποιημένη μεταβλητή $z$	$F(z)$ (υπολογισμένη)	$F(z)$ (πρότυπο) (*)
0.2500	0.5987063	0.59871
0.5244	0.6999998	0.70000
1.0000	0.8413447	0.84134
1.6449	0.9500048	0.95000

(\*) Πίνακας Π1 (Κουτσογιάννης 1997, σ. 301).

### Ο υπολογισμός της κατανομής γάμα

Η συνάρτηση πυκνότητας πιθανότητας της κατανομής γάμα 2 παραμέτρων ( $\kappa, \lambda > 0$ ) δίνεται από την σχέση (Κουτσογιάννης, 1997):

$$f(x) = \frac{\lambda^\kappa}{\Gamma(\kappa)} x^{\kappa-1} e^{-\lambda x} \quad 9.32$$

Η συνάρτηση κατανομής  $F$  προκύπτει από ολοκλήρωση της (9.33) ως εξής:

$$F(x) = \int_{-\infty}^x f(t) dt = \int_{-\infty}^x \frac{\lambda^\kappa}{\Gamma(\kappa)} t^{\kappa-1} e^{-\lambda t} dt \quad 9.33$$

Οι υπολογισμοί των παραπάνω συναρτήσεων βασίζονται αφενός στον αριθμητικό υπολογισμό της συνάρτησης  $\Gamma(x)$ , αφετέρου σε επιμέρους αριθμητικούς υπολογισμούς του ολοκληρώματος της (9.33).

Ο αριθμητικός υπολογισμός της συνάρτησης  $\Gamma(x)$

Η συνάρτηση  $\Gamma$  ορίζεται από την σχέση (Κουτσογιάννης, 1997):

$$\Gamma(a) = \int_0^{\infty} y^{a-1} e^{-y} dy \quad 9.34$$

Το ολοκλήρωμα συγκλίνει για  $a > 0$ . Παρακάτω δίνονται μερικές χαρακτηριστικές ιδιότητες της συνάρτησης.

$$\begin{aligned} \Gamma(1) &= 1 & \Gamma(1/2) &= \sqrt{\pi} & \Gamma(a+1) &= a\Gamma(a) \\ \Gamma(n+1) &= n! & \Gamma(n + \frac{1}{2}) &= (n - \frac{1}{2})(n - \frac{3}{2}) \cdots \frac{3}{2} \sqrt{\pi} & n &= 1, 2, \dots \end{aligned} \quad 9.35$$

Οι παραπάνω ιδιότητες είναι ιδιαίτερα χρήσιμες για τον υπολογισμό καθώς:

- Μπορούμε να υπολογίσουμε την τιμή του παραγοντικού για τις περιπτώσεις ακέραιου ορίσματος
- Μέσω της ιδιότητας  $\Gamma(a+1) = a\Gamma(a)$ , μπορούμε να ανάγουμε κάθε υπολογισμό σε γινόμενα ακεραίων με την προσέγγιση της  $\Gamma(a)$  για  $a < l$  όπου  $l$  κάποιο πραγματικό όριο πέρα του οποίου έχουμε ικανοποιητική προσέγγιση της  $\Gamma(a)$ .

Θέτουμε ως όριο για τον «Υδρογόνων»  $l=10$ .

Έστω πως θέλουμε να υπολογίσουμε την  $\Gamma(a)$ ,  $a > 0$ . Ορίζουμε την ανηγμένη μεταβλητή  $y$ . Η συνάρτηση  $\Gamma$  για την αυτήν την μεταβλητή είναι:

$$\Gamma(y) = \sqrt{\frac{2\pi}{y}} e^{y \ln(y) + \frac{1 - \frac{1}{30y^2}}{12y} - y} \quad 9.36$$

όπου η ανηγμένη μεταβλητή ορίζεται ως:

$$y = a + m \quad 9.37$$

Η  $m$  είναι μία ακέραια μεταβλητή που ορίζεται ως:

$$m = \min(0, l - [a]) \quad 9.38$$

όπου  $[a]$  το ακέραιο μέρος του  $a$ . Για την περίπτωση που  $a \geq l+1$  τότε η  $a$  και η  $y$  ταυτίζονται και η  $\Gamma(a)$  μπορεί να υπολογιστεί κατευθείαν από την  $\Gamma(y)$ :

$$\Gamma(a) = \Gamma(y) \quad a \geq l+1 \quad 9.39$$

Για  $a < l+1$ :

$$\Gamma(a) = \Gamma(y)/p \quad a < l+1 \quad 9.40$$

όπου:

$$p = \prod_{i=0}^{l-1} a + i \quad 9.41$$

### Υπολογισμός της συνάρτησης πυκνότητας πιθανότητας της κατανομής Γ

Έχοντας ορίσει τον υπολογισμό της συνάρτησης Γ μπορούμε να προχωρήσουμε στον υπολογισμό της συνάρτησης πυκνότητας πιθανότητας. Για βελτίωση της απόδοσης του συστήματος, ο λόγος:

$$r(x, \kappa) = x^{\kappa-1} / \Gamma(\kappa) \quad 9.42$$

, υπολογίζεται απευθείας από αυτήν την σχέση μέχρι κάποια οριακή τιμή του  $\kappa = \kappa_{\max}$ .

Στον «Υδρογνώμων» θέτουμε  $\kappa_{\max} = 25$ .

Για τιμές  $\kappa > \kappa_{\max}$ , εφαρμόζουμε την παρακάτω αναδρομική σχέση εκμεταλλευόμενοι τις αναδρομικές ιδιότητες της συνάρτησης Γ καθώς και του εκθέτη.

$$r(x, \kappa) = \frac{x}{\kappa - 1} r(x, \kappa - 1) \quad \kappa > \kappa_{\max} \quad 9.43$$

Τελικά η συνάρτηση πυκνότητας πιθανότητας της κατανομής Γ υπολογίζεται μέσω της σχέσης:

$$f(x) = \lambda e^{-\lambda x} r(\lambda x, \kappa) \quad 9.44$$

### Υπολογισμός της συνάρτησης κατανομής Γ

Για την περίπτωση που το  $\kappa$  είναι μεγαλύτερο από κάποιο όριο  $\kappa_{\text{normal}}$ , εφαρμόζοντας το κεντρικό οριακό θεώρημα, μπορούμε να εκτιμήσουμε την της πιθανότητας  $F$  χρησιμοποιώντας την κανονική κατανομή  $N$  με παραμέτρους  $\mu = \kappa / \lambda$  και  $\sigma = \sqrt{\kappa / \lambda}$ . Αυτό μας δίνει πλεονέκτημα απόδοσης καθώς από κάποιο σημείο και έπειτα οι υπολογισμοί της Γ είναι χρονοβόροι λόγω των διαδοχικών γινομένων.

Για τον «Υδρογνώμων» μας θέτουμε ως όριο χρήσης της κανονικής κατανομής  $\kappa_{\text{normal}} = 150$ , τιμή που δίνει μία ισορροπία καλής ακρίβειας και απόδοσης.

Για  $\kappa < \kappa_{\text{normal}}$  θα εφαρμόζεται η σχέση:

$$F(x) = x f(x) \sum_{i=0}^{\infty} t(i) \quad 9.45$$

όπου  $f(x)$  η συνάρτηση πυκνότητας πιθανότητας η οποία υπολογίζεται αριθμητικά όπως προδιαγράφεται στην προηγούμενη παράγραφο και  $t(i)$  όροι κάποιας αριθμητικής απειροσειράς. Πρακτικά χρησιμοποιείται ένας πεπερασμένος αριθμός όρων μέχρι την επίτευξη κάποιας ακρίβειας όχι όμως περισσότεροι από  $n_{\max}$ .

Στον «Υδρογνώμων» τίθεται ως  $n_{\max} = 100$ .

Οι όροι  $t(i)$  ορίζονται από την σχέση:

$$t(i) = \frac{(\lambda x)^i}{\prod_{j=0}^i (\kappa + j)} \quad 9.46$$

Η οποία ισοδυναμεί με την αναδρομική σχέση η οποία τελικά χρησιμοποιείται στους υπολογισμούς:

$$t_{i+1} = t_i \frac{\lambda x}{\kappa + i}, \quad t_0 = \frac{1}{\lambda x}$$

Η απαιτούμενη ακρίβεια επιτυγχάνεται μόλις προκύψει κάποιος όρος  $t(i)$  μικρότερος από το κατώφλι  $\varepsilon$ :

$$|t(i)| < \varepsilon \quad 9.47$$

Στον «Υδρογνώμων» εφαρμόζεται  $\varepsilon = 10^{-10}$ .

Οι υπόλοιπες συναρτήσεις κατανομής της οικογένειας της κατανομής  $\Gamma$  βασίζονται στους παραπάνω υπολογισμούς κατόπιν κατάλληλων μετασχηματισμών και εισαγωγής παραμέτρων. Θα παρουσιαστούν παρακάτω σε ξεχωριστές παραγράφους.

#### Αριθμητικός υπολογισμός της $F^{-1}$ της συνάρτησης κατανομής $\Gamma$

Εφαρμόζεται η μέθοδος Newton-Raphson όπως και στην περίπτωση της κανονικής κατανομής (σ. 84). Εφαρμόζεται λοιπόν η σχέση (9.27) όπου  $G, G'$  είναι ορισμένες μέσω της συνάρτησης κατανομής  $\Gamma$  και της συνάρτησης πυκνότητας πιθανότητας ως άνω.

Στην περίπτωση που  $\kappa > \kappa_{\text{normal}}$  χρησιμοποιείται η αντίστροφη κανονική κατανομή  $N$  με παραμέτρους  $\mu = \kappa/\lambda$  και  $\sigma = \sqrt{\kappa/\lambda}$ . Όπως και στην περίπτωση της συνάρτησης κατανομής  $\Gamma$  το  $\kappa_{\text{normal}}$  τίθεται ίσο με 150.

Για τις περιπτώσεις που  $\kappa < \kappa_{\text{normal}}$  εφαρμόζεται η αριθμητική μέθοδος των Newton Raphson. Η ακολουθία θεωρείται πως συγκλίνει μόλις ικανοποιηθεί κάποιο από τα παρακάτω κριτήρια:

$$\left| \frac{x_{i+1} - x_i}{x_{i+1}} \right| < \varepsilon \quad 9.48$$

$$i > i_{\text{max}} \quad 9.49$$

Για τον «Υδρογνώμων» τίθενται οι τιμές  $\varepsilon = 10^{-10}$  και  $i_{\text{max}} = 100$ .

Για να είναι ταχύτερη η σύγκλιση εφαρμόζονται οι παρακάτω ανισοτικοί μετασχηματισμοί:

$$x_{i+1} < \frac{x_i}{2} \Rightarrow x_{i+1} = \frac{x_i}{2} \quad 9.50$$

$$x_{i+1} > \frac{5x_i}{2} \Rightarrow x_{i+1} = \frac{5x_i}{2} \quad 9.51$$

Πίνακας 9.8: Επαλήθευση αλγορίθμων μέσω των τιμών της κατανομής  $\Gamma$

$\mu, \kappa$	Ποσοστημόριο $k_u$ (υπολογ.)	Ποσοστημόριο $k_u$ (πρότυπο) (*)
$\mu=0.05, \kappa=25$	-1.52357	-1.52
$\mu=0.10, \kappa=4$	-1.12762	-1.13
$\mu=0.90, \kappa=0.69$	1.26150	1.26
$\mu=0.995, \kappa=0.391$	5.08608	5.09

(\*) Πίνακας Π4α/β (Κουτσογιάννης 1997, σ. 301).

#### Η λογαριθμοκανονική κατανομή δύο παραμέτρων και η κατανομή Galton (λογ/κανονική τριών παραμέτρων)

Ο υπολογισμός αυτών των συναρτήσεων κατανομής βασίζεται στον υπολογισμό της κανονικής κατανομής ο οποίος παρουσιάζεται στις παραπάνω παραγράφους.

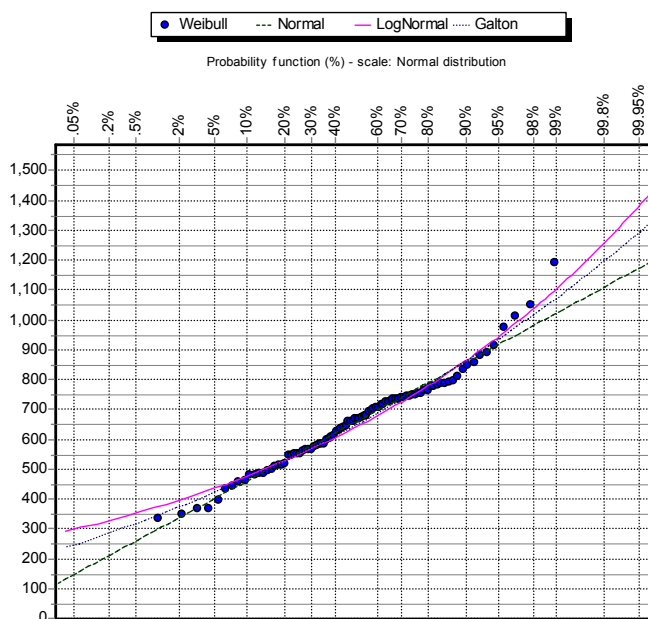
## Η λογαριθμοκανονική κατανομή δύο παραμέτρων

Αν  $\mu_y$  και  $\sigma_y$  οι παράμετροι της λογαριθμοκανονικής κατανομής, τότε η τυχαία μεταβλητή  $x$  ακολουθεί την λογαριθμοκανονική κατανομή αν η  $y=\ln(x)$  ακολουθεί την κανονική κατανομή  $N(x, \mu_y, \sigma_y)$ . Η συνάρτηση κατανομής λοιπόν θα δίνεται από την σχέση:

$$F(x) = \ln[N(x, \mu_y, \sigma_y)] \quad 9.52$$

όπου  $N$  η κανονική κατανομή. Η δε αντίστροφη συνάρτηση κατανομής θα δίνεται από την σχέση:

$$F^{-1} = e^{N^{-1}(u, \mu_y, \sigma_y)} \quad 9.53$$



Σχήμα 9.2: Κανονική, λογαριθμοκανονική και Galton κατανομή. Σημεία εμπειρικής πιθανότητας Weibull. Δείγμα: ετήσιες βροχοπτώσεις στον σταθμό «Αλιάρτος», 1907-8 – 2000-1

## Η λογαριθμοκανονική κατανομή τριών παραμέτρων (Galton)

Αν στην λογαριθμοκανονική κατανομή δύο παραμέτρων εισάγουμε μία παράμετρο θέσης  $c$  που μετασχηματίζει την τυχαία μεταβλητή σε  $x-c$ , τότε λαμβάνουμε την κατανομή Galton. Αν δηλαδή  $\mu_y$ ,  $\sigma_y$  και  $c$  οι παράμετροι της κατανομής Galton, τότε η τυχαία μεταβλητή  $x$  ακολουθεί την λογαριθμοκανονική κατανομή αν η  $y=\ln(x-c)$  ακολουθεί την κανονική κατανομή  $N(x, \mu_y, \sigma_y)$ . Η συνάρτηση κατανομής λοιπόν θα δίνεται από την σχέση:

$$F(x) = \ln[N(x-c, \mu_y, \sigma_y)] \quad 9.54$$

όπου  $N$  η κανονική κατανομή. Η δε αντίστροφη συνάρτηση κατανομής θα δίνεται από την σχέση:

$$F^{-1} = c + e^{N^{-1}(u, \mu_y, \sigma_y)} \quad 9.55$$

Η κατανομή Galton δεν ορίζεται για αρνητικό συντελεστή ασυμμετρίας του δείγματος.

## Η ομάδα των κατανομών γάμα

### Εκθετική κατανομή

Η εκθετική κατανομή είναι από τις απλούστερες κατανομές στην διατύπωση της μαθηματικής σχέσης. Η μορφή της κατανομής που είναι υλοποιημένη στον «Υδρογνώμων» έχει δύο παραμέτρους, μία παράμετρο κλίμακας  $\lambda$  και μία παράμετρο θέσης  $\psi$ .

Η συνάρτηση κατανομής δίνεται από την σχέση:

$$F(x) = 1 - e^{-\lambda(x-\psi)} \quad 9.56$$

Η δε αντίστροφη συνάρτηση κατανομής λύνοντας την παραπάνω σχέση ως προς  $x$ :

$$F^{-1}(u) = x(u) = \psi - \frac{\ln(1-u)}{\lambda} \quad 9.57$$

### Κατανομή γάμα τριών παραμέτρων (Pearson III)

Η κατανομή Pearson III προκύπτει εύκολα από την κατανομή γάμα αν προσθέσουμε μία παράμετρο θέσης  $\psi$ . Έτσι η τυχαία μεταβλητή  $x$  ακολουθεί την κατανομή Pearson III, αν η μεταβλητή  $y=(x-\psi)$  ακολουθεί την κατανομή γάμα. Αν λοιπόν  $G(x, \kappa, \lambda)$  η κατανομή γάμα με παραμέτρους  $\kappa$  και  $\lambda$ , η συνάρτηση κατανομής της Pearson III θα είναι:

$$F(x) = G(x-\psi, \kappa, \lambda) \quad 9.58$$

Η δε αντίστροφη συνάρτηση κατανομής της Pearson III:

$$x(u) = G^{-1}(u, \kappa, \lambda) + \psi \quad 9.59$$

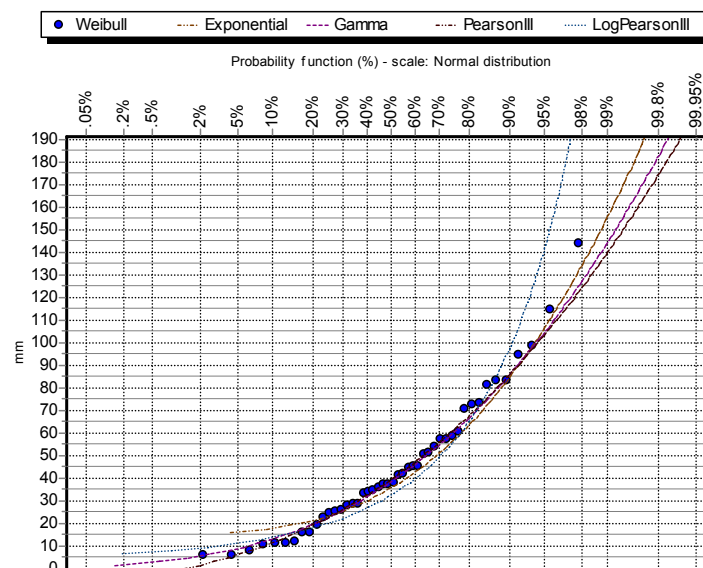
### Η κατανομή Log-Pearson III

Η κατανομή Log-Pearson III προκύπτει από την Pearson III με λογαριθμικό μετασχηματισμό. Συγκεκριμένα, η τυχαία μεταβλητή  $x$  ακολουθεί την κατανομή Log-Pearson III, αν η μεταβλητή  $y=\ln(x)$  ακολουθεί την κατανομή Pearson III. Αν λοιπόν  $P(x, \kappa, \lambda, \psi)$  η κατανομή Pearson III με παραμέτρους  $\kappa, \lambda$  και  $\psi$ , η συνάρτηση κατανομής της Log-Pearson III θα είναι:

$$F(x) = P(\ln(x), \kappa, \lambda, \psi) \quad 9.60$$

Η δε αντίστροφη συνάρτηση κατανομής της Log-Pearson III:

$$x(u) = e^{P^{-1}(u, \kappa, \lambda, \psi)} \quad 9.61$$



Σχήμα 9.3 Εκθετική, γάμα, Pearson III, Log-Pearson III κατανομή. Σημεία εμπειρικής πιθανότητας Weibull. Δείγμα: βροχοπτώσεις Απριλίου στον σταθμό «Αμφίσσα», 1955 – 2000

## Ασυμπτωτικές κατανομές ακροτάτων (μεγίστων)

Η κατανομή Ακραίων Τιμών τύπου 1 (AT-1 μεγίστων) – Gumbel μεγίστων

Η κατανομή Gumbel μεγίστων με παραμέτρους κλίμακας  $\lambda$  και θέσης  $\psi$  δίνεται από την σχέση:

$$F(x) = e^{-e^{-\frac{x}{\lambda} + \psi}} \quad 9.62$$

Η δε αντίστροφη συνάρτηση κατανομής δίνεται από την σχέση:

$$x(u) = \lambda \psi - \lambda \ln(-\ln u) \quad 9.63$$

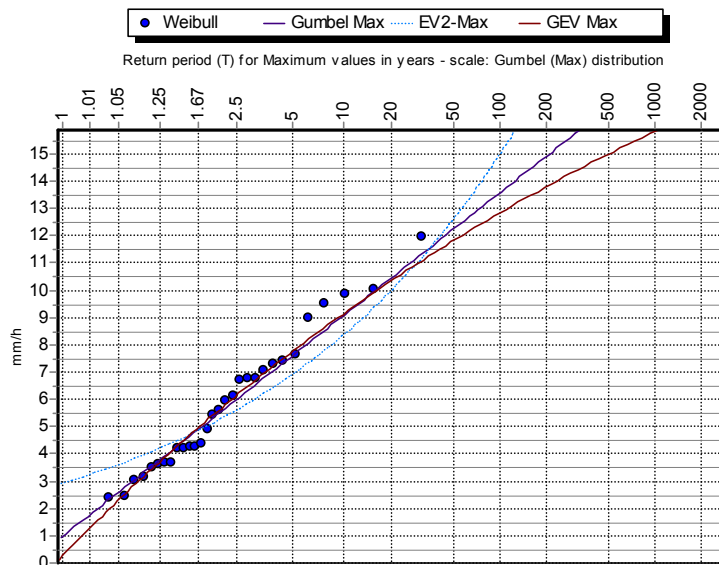
Η κατανομή Ακραίων Τιμών τύπου 2 (AT-2 μεγίστων)

Η κατανομή AT-2 μεγίστων με παραμέτρους σχήματος  $\kappa$  και κλίμακας  $\lambda$  δίνεται από την σχέση:

$$F(x) = e^{-\left(\frac{x}{\lambda}\right)^{-\frac{1}{\kappa}}} \quad 9.64$$

Η δε αντίστροφη συνάρτηση κατανομής δίνεται από την σχέση:

$$x(u) = \frac{\lambda}{\kappa} (-\ln u)^{-\kappa} \quad 9.65$$



Σχήμα 9.4 AT-1, AT-2, ΓΑΤ κατανομές μεγίστων. Σημεία εμπειρικής πιθανότητας Weibull. Δείγμα: μέγιστες ετήσιες εντάσεις βροχής εξάωρου στον σταθμό «Ελληνικό», 1957-58 – 1986-87 (Πηγή: Εθνική Μετεωρολογική Υπηρεσία).

Η κατανομή Γενική Ακραίων Τιμών (ΓΑΤ μεγίστων)

Η ΓΑΤ είναι μία γενίκευση των κατανομών ακραίων τιμών που συνδυάζει τις τρεις περιπτώσεις AT-1, AT-2 και AT-3 (Κουτσογιάννης, 1999). Οι παράμετροι της κατανομής είναι τρεις, μία παράμετρος σχήματος  $\kappa$ , μία παράμετρος κλίμακας  $\lambda$  και μία παράμετρος θέσης  $\psi$ . Όταν  $\kappa \rightarrow 0$  τότε λαμβάνουμε την AT-1, ενώ για  $\kappa > 0$  και  $\psi = 1/\kappa$  μπορούμε να λάβουμε την AT-2. Η σχέση της ΓΑΤ μεγίστων είναι η παρακάτω:

$$F(x) = e^{-\left[1 + \kappa \left(\frac{x}{\lambda} - \psi\right)\right]^{-1/\kappa}} \quad 9.66$$



Η παραπάνω σχέση ισχύει για τιμές του  $\kappa$  διάφορες του 0. Όταν το  $\kappa$  τείνει στο 0 και συγκεκριμένα στον «Υδρογνώμων» όταν  $|\kappa| < 0.001$  χρησιμοποιούμε την έκφραση της AT-1 μεγίστων. Η αντίστροφη συνάρτηση κατανομής δίνεται από την σχέση:

$$x(u) = \lambda\psi + \frac{\lambda}{\kappa} \left[ (-\ln u)^{-\kappa} - 1 \right] \quad 9.67$$

Αντίστοιχα, όταν το  $\kappa$  τείνει στο 0 θα χρησιμοποιείται η  $x(u)$  της AT-1 μεγίστων.

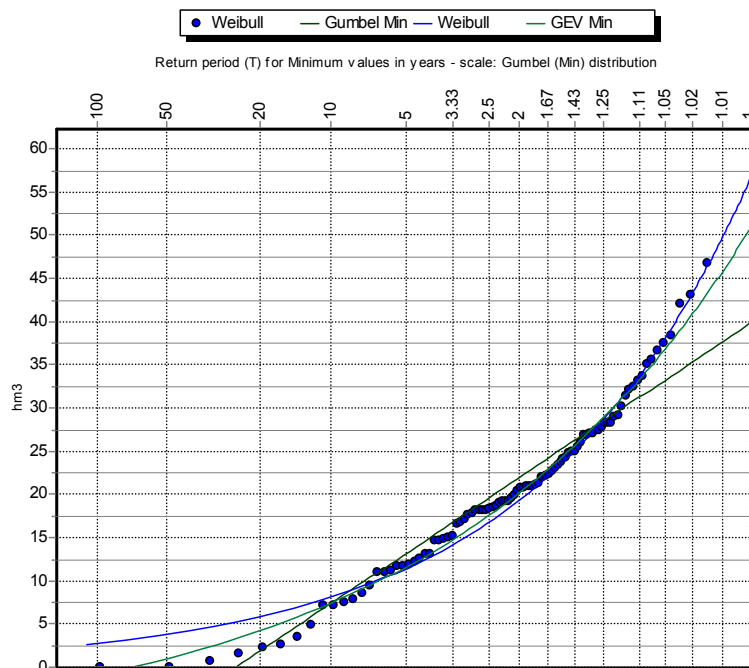
#### Ασυμπτωτικές κατανομές ακροτάτων (ελαχίστων)

Η κατανομή Ακραίων Τιμών τύπου 1 (AT-1 ελαχίστων) – Gumbel ελαχίστων  
 Η κατανομή Gumbel ελαχίστων με παραμέτρους κλίμακας  $\lambda$  και θέσης  $\psi$  δίνεται από την σχέση:

$$F(x) = 1 - e^{-e^{\frac{x}{\lambda} - \psi}} \quad 9.68$$

Η δε αντίστροφη συνάρτηση κατανομής δίνεται από την σχέση:

$$x(u) = \lambda\psi + \lambda \ln[-\ln(1-u)] \quad 9.69$$



Σχήμα 9.5: Ελάχιστες μηνιαίες παροχές περιόδου Οκτώβριος – Μάρτιος. AT-1, AT-3, ΓΑΤ κατανομές ελαχίστων. Μηνιαία παροχή Β. Κηφισού («Διώρυγα Καρδίτσας»), 1906-7 – 2003-4.

Η κατανομή Ακραίων Τιμών τύπου 3 (AT-3 ελαχίστων) ή κατανομή Weibull  
 Η κατανομή Weibull με παραμέτρους σχήματος  $\kappa$  και κλίμακας  $\lambda$  δίνεται από την σχέση:

$$F(x) = 1 - e^{-\left(\frac{\kappa x}{\lambda}\right)^{\kappa}} \quad 9.70$$

Η δε αντίστροφη συνάρτηση κατανομής δίνεται από την σχέση:

$$x(u) = \frac{\lambda}{\kappa} \left[ -\ln(1-u) \right]^{\kappa} \quad 9.71$$

### Η κατανομή Γενική Ακραίων Τιμών (ΓΑΤ ελαχίστων)

Η σχέση της ΓΑΤ ελαχίστων με παραμέτρους  $\kappa$ ,  $\lambda$ ,  $\psi$  (σχήμα, κλίμακα θέση) είναι η παρακάτω:

$$F(x) = 1 - e^{-\left[1 + \kappa \left(\frac{x}{\lambda} - \psi\right)\right]^{1/\kappa}} \quad 9.72$$

Η παραπάνω σχέση ισχύει για τιμές του  $\kappa$  διάφορες του 0. Όταν το  $\kappa$  τείνει στο 0 και συγκεκριμένα στον «Υδρογνώμων» όταν  $|\kappa| < 0.001$  χρησιμοποιούμε την έκφραση της AT-1 ελαχίστων. Η αντίστροφη συνάρτηση κατανομής δίνεται από την σχέση:

$$x(u) = \lambda\psi + \frac{\lambda}{\kappa} \left\{ \left[ -\ln(1-u) \right]^\kappa - 1 \right\} \quad 9.73$$

Αντίστοιχα, όταν το  $\kappa$  τείνει στο 0 θα χρησιμοποιείται η  $x(u)$  της AT-1 ελαχίστων.

#### Κατανομή Pareto τριών παραμέτρων

Η συνάρτηση κατανομής Pareto με παραμέτρους σχήματος  $\kappa$ , κλίμακας  $\lambda$  και θέσης  $\psi$ , δίνεται από την σχέση:

$$F(x) = 1 - \left[ 1 - \kappa \left( \frac{x}{\lambda} - \psi \right) \right]^{1/\kappa} \quad 9.74$$

Η παραπάνω σχέση δεν ορίζεται για  $\kappa=0$ , οπότε εφαρμόζοντας τον κανόνα του l'Hôpital λαμβάνουμε την παρακάτω σχέση την οποία χρησιμοποιούμε όταν  $|\kappa| < 0.001$ :

$$F(x) = 1 - e^{-\frac{\psi - x}{\lambda}} \quad 9.75$$

Η αντίστροφη συνάρτηση κατανομής δίνεται αντίστοιχα από τις σχέσεις:

$$x(u) = \frac{\lambda}{\kappa} \left[ 1 - (1-u)^\kappa + \psi\kappa \right] \quad 9.76$$

Όταν δε  $|\kappa| < 0.001$  θα χρησιμοποιείται η σχέση:

$$x(u) = \lambda[\psi - \ln(1-u)] \quad 9.77$$

#### 9.4.2 Στατιστική πρόγνωση

Σύμφωνα με τον Κουτσογιάννη (1997): «Η στατιστική πρόγνωση στην τεχνική υδρολογία αφορά στην εκτίμηση της τιμής ενός υδρολογικού μεγέθους για δεδομένη πιθανότητα υπέρβασης (ή περίοδο επαναφοράς). Η πρόγνωση αυτή είναι αρκετά εύκολη από υπολογιστική άποψη, εφόσον έχει ήδη υιοθετηθεί και προσαρμοστεί ένα συγκεκριμένο πιθανοτικό μοντέλο για το μέγεθος που ενδιαφέρει». Ορισμένες φορές μάλιστα, μας ενδιαφέρει και το αντίστροφο πρόβλημα: της εύρεσης, δηλαδή, της πιθανότητας υπέρβασης ή της περιόδου επαναφοράς για συγκεκριμένη τιμή ενός μεγέθους.

Η εκτίμηση των παραπάνω μεγεθών μπορεί να γίνει είτε με απευθείας εφαρμογή της αντιστρόφου συνάρτησης κατανομής (συνάρτησης ποσοστημορίου)  $F^{-1}$  (ή  $x(u)$ ) για την εκτίμηση του φυσικού μεγέθους συναρτήσει της πιθανότητας είτε της συνάρτησης κατανομής  $F$  για την εκτίμηση της πιθανότητας συναρτήσει του φυσικού μεγέθους.

Εφόσον δε, έχει σχεδιαστεί ένα διάγραμμα της κατανομής είναι δυνατή επιπλέον, η παρεμβολή του μεγέθους στα σημεία του διαγράμματος. Αυτή η μεθοδολογία εφαρμόζεται ειδικότερα στην εύρεση του διαστήματος εμπιστοσύνης σε τυχαία θέση όταν οι υπολογισμοί όπως θα δούμε στην ενότητα 9.7 είναι χρονοβόροι και γίνονται σε διακριτά σημεία. Στον Πίνακα 9.9 παρουσιάζεται ένα παράδειγμα στατιστικής πρόγνωσης σε συγκεκριμένη τιμή που υπολογίστηκε από τον «Υδρογνώμων».

Η  $F_1$  υπενθυμίζουμε πως είναι η πιθανότητα υπέρβασης ( $F_1 = 1 - F$ ,  $F$  η συνάρτηση κατανομής). Επιπλέον  $T(\max) = 1/F_1$  και  $T(\min) = 1/F$ .

Πίνακας 9.9: Παράδειγμα στατιστικής πρόγνωσης. Βροχοπτώσεις Οκτωβρίου από τον σταθμό «Αλιάρτος» 1907-8 – 2000-1.

Οκτώβριος Value= 50	F(%)	F1(%)	T(Max)(y)	T(Min)(y)
Normal	35.57	64.43	1.55	2.81
LogNormal	44.23	55.77	1.79	2.26
Galton	41.64	58.36	1.71	2.4
Exponential	46.73	53.27	1.88	2.14

### 9.4.3 Μεθοδολογία εκτίμησης των παραμέτρων των θεωρητικών κατανομών

Ήδη στον Πίνακα 9.4 έχουν παρουσιαστεί ονομαστικά οι μεθοδολογίες εκτίμησης των παραμέτρων των θεωρητικών κατανομών κάνοντας διάκριση σε υπολογισμούς με την μέθοδο των ροπών, της έμμεσης μεθόδου των ροπών καθώς και των L-ροπών. Προκειμένου να υπολογιστούν οι παράμετροι πρέπει να έχουμε στην διάθεση μας τα στατιστικά χαρακτηριστικά του δείγματος που παρουσιάστηκαν στην ενότητα 9.2 (μέση τιμή, τυπική απόκλιση και συντελεστή ασυμμετρίας  $\mu_x$ ,  $\sigma_x$  και  $C_s$  καθώς και τις L-ροπές  $\lambda_1$ ,  $\lambda_2$  και  $\lambda_3$  προκειμένου για εκτίμηση παραμέτρων κατανομής με την μέθοδο των L-ροπών). Επιπλέον για την κατανομή Log-Pearson III είναι απαραίτητα τα στατιστικά χαρακτηριστικά του λογαριθμικά μετασχηματισμένου δείγματος  $y = \ln(x)$ .

Μία ακόμη διάκριση που γίνεται στον υπολογισμό των παραμέτρων των κατανομών AT-3 ελαχίστων, AT-2 μεγίστων καθώς και ΓAT, με την μέθοδο των ροπών, είναι είτε η εφαρμογή της ακριβούς σχέσης υπολογισμού η οποία λύνεται μόνο με αριθμητική μέθοδο, είτε η εφαρμογή μίας προσέγγισης με σχέση κλειστής μορφής. Οι εκτιμήσεις των παραμέτρων με την μέθοδο των L-ροπών (όταν είναι αυτή δυνατή, βλ. Πίνακα 9.4) γίνονται πάντα με σχετικά απλές σχέσεις κλειστής μορφής.

Η χρήση των ακριβών σχέσεων μπορεί να χρησιμοποιείται για την εύρεση των παραμέτρων που θα χρησιμοποιηθούν στην συνέχεια για στατιστική πρόγνωση. Η χρήση προσεγγιστικών λύσεων πλεονεκτεί κατά την εφαρμογή της στοχαστικής προσομοίωσης. Συνεπώς για την εύρεση των διαστημάτων εμπιστοσύνης με τις μεθοδολογίες που αναλύονται στην ενότητα 9.7 θα χρησιμοποιούνται οι προσεγγιστικές λύσεις για τις παραμέτρους.

#### Η εκτίμηση των παραμέτρων των κατανομών με την μέθοδο των ροπών

##### Η κανονική κατανομή

Είναι η πλέον απλή περίπτωση, λαμβάνονται απλά οι δειγματικές τιμές  $\mu_x$  και  $\sigma_x$  ως παράμετροι θέσης  $\mu = \mu_x$  και κλίμακας  $\sigma = \sigma_x$ .

##### Η λογαριθμοκανονική κατανομή δύο παραμέτρων

Οι παράμετροι της λογαριθμοκανονικής κατανομής δύο παραμέτρων είναι οι  $\mu_y$  (παράμετρος θέσης) και  $\sigma_y$  (παράμετρος κλίμακας) και μπορούν να προκύψουν από τις δειγματικές τιμές  $\mu_x$  και  $\sigma_x$  με την μέθοδο των ροπών ως εξής:

$$\sigma_y = \sqrt{\ln\left(1 + \frac{\sigma_x^2}{\mu_x^2}\right)} \quad 9.78$$

$$\mu_y = \ln \mu_x - \frac{\sigma_y^2}{2} \quad 9.79$$

### Η λογαριθμοκανονική κατανομή τριών παραμέτρων (Galton)

Εισάγοντας την παράμετρο θέσης  $c$  η  $\mu_y$  γίνεται παράμετρος κλίμακας και η  $\sigma_y$  παράμετρος σχήματος. Για την εκτίμηση των παραμέτρων απαιτείται πλέον και η δειγματική τιμή του συντελεστή ασυμμετρίας  $C_{sx}$ . Η κατανομή Galton έχει εξορισμού θετική ασυμμετρία και δεν είναι δυνατή η διατύπωση της αν έχουμε αρνητική δειγματική τιμή ασυμμετρίας. (Σε αυτήν την περίπτωση στον «Υδρογνώμων» αναιρούνται οι υπολογισμοί παραμέτρων, επιστρέφεται δε αρνητικό  $\sigma_y$  ώστε να μην γίνει η επεξεργασία από το λογισμικό). Οι υπολογισμοί γίνονται με την εξής σειρά (υιοθετώντας και δύο ενδιάμεσες παραμέτρους  $\omega$  και  $\phi$ ).

$$\omega = \frac{-C_{sx} + \sqrt{C_{sx}^2 + 4}}{2}, \quad \phi = \frac{1 - \omega^{2/3}}{\omega^{1/3}} \quad 9.80$$

$$\sigma_y = \sqrt{\ln(1 + \phi^2)} \quad 9.81$$

$$\mu_y = \ln \frac{\sigma_x}{\phi} - \frac{\sigma_y^2}{2} \quad 9.82$$

$$c = \mu_x - \frac{\sigma_x}{\phi} \quad 9.83$$

### Η εκθετική κατανομή δύο παραμέτρων

$$\lambda = \frac{1}{\sigma_x} \quad 9.84$$

$$\psi = \mu_x - \sigma_x \quad 9.85$$

### Η κατανομή γάμα δύο παραμέτρων

Η κατανομή γάμα είναι θετικά ασύμμετρη καθώς και ορίζεται μόνο για θετικές τιμές της μεταβλητής. Έτσι επειδή υπάρχει η απαίτηση του συντελεστή κλίμακας να είναι θετικός, δεν μπορεί να οριστεί η κατανομή γάμα για αρνητικές τιμές της δειγματικής μέσης τιμής  $\mu_x$ . Οι παράμετροι σχήματος  $\kappa$  και κλίμακας  $\lambda$  δίνονται από τις παρακάτω σχέσεις:

$$\kappa = \frac{\mu_x^2}{\sigma_x^2} \quad 9.86$$

$$\lambda = \frac{\mu_x}{\sigma_x^2} \quad 9.87$$

### Η κατανομή γάμα τριών παραμέτρων (Pearson III)

Εισάγοντας την παράμετρο θέσης  $\psi$  στην κατανομή γάμα λαμβάνουμε την κατανομή Pearson III. Σε αυτήν την περίπτωση η παράμετρος σχήματος  $\kappa$  ελέγχεται μόνο από τον συντελεστή ασυμμετρίας  $C_{sx}$ . Οι παράμετροι της κατανομής υπολογίζονται με την μέθοδο των ροπών από τις παρακάτω σχέσεις:

$$\kappa = \frac{4}{C_{sx}} \quad 9.88$$

$$\lambda = \frac{\sqrt{\kappa}}{\sigma_x} \quad 9.89$$

$$\psi = \mu_x - \frac{\kappa}{\lambda} \quad 9.90$$

Η κατανομή AT-1 μεγίστων (Gumbel max)

$$\lambda = 0.78\sigma_x \quad 9.91$$

$$\psi = \frac{\mu_x}{\lambda} - 0.5772 \quad 9.92$$

Η κατανομή AT-1 ελαχίστων (Gumbel min)

$$\lambda = 0.78\sigma_x \quad 9.93$$

$$\psi = \frac{\mu_x}{\lambda} + 0.5772 \quad 9.94$$

Η κατανομή AT-2 μεγίστων

Η σχέση που συνδέει την παράμετρο σχήματος της AT-2 μεγίστων με τα στατιστικά χαρακτηριστικά του δείγματος είναι της μορφής:

$$\frac{\Gamma(1-2\kappa)}{\Gamma^2(1-\kappa)} = \frac{\sigma_x^2}{\mu_x^2} + 1 \quad 9.95$$

Η λύση της παραπάνω εξίσωσης ως προς  $\kappa$  γίνεται με αριθμητική μέθοδο. Προσεγγιστική λύση για το  $\kappa$  θα παρουσιαστεί σε επόμενη παράγραφο. Εφόσον έχουμε λύση για την  $\kappa$ , η παράμετρος κλίμακας  $\lambda$  μπορεί να υπολογιστεί από την σχέση:

$$\lambda = m \frac{\kappa}{\Gamma(1-\kappa)} \quad 9.96$$

Εξετάζοντας την σχέση (9.95) συμπεραίνουμε πως το  $\kappa$  πρέπει να κυμαίνεται μεταξύ 0 και 0.5 ώστε να μην έχουμε αρνητικό όρισμα της συνάρτησης  $\Gamma$ . Επιπλέον το αριστερό σκέλος της εξίσωσης είναι αύξων ως προς το  $\kappa$ . Η αριθμητική λύση της (9.95) θα επιτυγχάνεται με την μέθοδο της εσφαλμένης θέσης (Charpa and Canale, 1988) με τιμές εκκίνησης της  $\kappa$ :

$\kappa_{\min} = 0, \kappa_{\max} = 0.5$
--

Και συνεχή μετακίνηση των ορίων  $\kappa_{\min}$  και  $\kappa_{\max}$  (λαμβάνοντας ως νέα τιμή ενός από τα δύο όρια την μέση τιμή τους ανάλογα με το πρόσημο της παράστασης – εκμετάλλευση της μονοτονίας). Αν  $i$  ο κύκλος υπολογισμού, ο υπολογισμός θεωρείται επιτυχής μόλις ικανοποιηθεί κάποιο από τα παρακάτω κριτήρια:

$$\kappa_{\max} - \kappa_{\min} \leq \varepsilon \quad 9.97$$

$$i \geq i_{\max} \quad 9.98$$

όπου για τον «Υδρογώνων» υιοθετούμε  $i_{\max}=100$  και  $\varepsilon=0.0001$ .

### Η κατανομή AT-3 ελαχίστων (Weibull)

Η σχέση που συνδέει την παράμετρο σχήματος της AT-3 ελαχίστων με τα στατιστικά χαρακτηριστικά του δείγματος είναι της μορφής:

$$\frac{\Gamma(1+2\kappa)}{\Gamma^2(1+\kappa)} = \frac{\sigma_x^2}{\mu_x^2} + 1 \quad 9.99$$

Η λύση της παραπάνω εξίσωσης ως προς  $\kappa$  γίνεται με αριθμητική μέθοδο. Προσεγγιστική λύση για το  $\kappa$  θα παρουσιαστεί σε επόμενη παράγραφο. Εφόσον έχουμε λύση για την  $\kappa$ , η παράμετρος κλίμακας  $\lambda$  μπορεί να υπολογιστεί από την σχέση:

$$\lambda = m \frac{\kappa}{\Gamma(1+\kappa)} \quad 9.100$$

Δεν είναι δυνατόν να υιοθετηθεί η αριθμητική μέθοδος που εφαρμόστηκε στην AT-2 μεγίστων καθώς η παράμετρος  $\kappa$  δεν είναι πλέον φραγμένη από δύο όρια αλλά μόνο από ένα κάτω όριο  $\kappa = -0.5$  (το οποίο έχει νόημα στην κατανομή Weibull και θα λαμβάνεται ως κάτω όριο το 0). Έτσι ακολουθούμε την εξής στρατηγική: Χρησιμοποιώντας την ιδιότητα της συνάρτησης  $\Gamma$ :  $\Gamma(\alpha+1)=\alpha\Gamma(\alpha)$ , η σχέση (9.99) μετασχηματίζεται σε:

$$\frac{2\kappa\Gamma(2\kappa)}{\Gamma^2(1+\kappa)} = \frac{\sigma_x^2}{\mu_x^2} + 1 \quad 9.101$$

Και λύνοντας ως προς  $\kappa$  λαμβάνουμε την εξής πεπλεγμένη εξίσωση:

$$\kappa = \frac{\Gamma^2(1+\kappa)}{2\Gamma(2\kappa)} \left( \frac{\sigma_x^2}{\mu_x^2} + 1 \right) \quad 9.102$$

Η (9.102) μπορεί να λυθεί με διαδοχικές επαναλήψεις θέτοντας μία αρχική τιμή στο  $\kappa$  του δεξιού σκέλους και κάθε νέα τιμή που προκύπτει στο αριστερό σκέλος. Τελικά η (9.102) συγκλίνει προς την τελική τιμή του  $\kappa$ . Αν  $\kappa_i$  η τρέχουσα τιμή του  $\kappa$  στον  $i$  κύκλο υπολογισμού και  $\kappa_{i+1}$  η νέα τιμή που προκύπτει από το δεξί σκέλος της (9.102) η ταχύτητα σύγκλισης αυξάνεται σημαντικά αν χρησιμοποιηθεί ο παρακάτω μετασχηματισμός:

$$\kappa_{i+1} = \frac{1}{2}(3\kappa_i - \kappa_{i+1}) \quad 9.103$$

Κατ'αυτόν τον τρόπο «εξομαλύνονται» οι διακυμάνσεις της ακολουθίας και συγκλίνει γρηγορότερα. Στην συνέχεια δε γίνεται ο εξής έλεγχος και μετασχηματισμός ώστε να μην έχουμε απόκλιση σε περίπτωση αρνητικών τιμών του  $\kappa$ :

$$\kappa_{i+1} < 0 \Rightarrow \kappa_{i+1} = \frac{\kappa_i}{2}$$

Η σύγκλιση επιτυγχάνεται μόλις ικανοποιηθεί κάποιο από τα παρακάτω κριτήρια σύγκλισης:

$$\left| \frac{\kappa_{i+1} - \kappa_i}{\kappa_{i+1}\kappa_i} \right| \leq \varepsilon \quad 9.104$$

$$i \geq i_{\max} \quad 9.105$$

όπου για τον «Υδρογνώμων» υιοθετούμε  $i_{\max}=100$  και  $\varepsilon=0.001$ .

Ως αρχική τιμή εκκίνησης  $\kappa_0$  θα λαμβάνεται η προσεγγιστική τιμή του  $\kappa$  της σχέσης (9.124).

### Η κατανομή ΓΑΤ μεγίστων

Η σχέση που συνδέει την παράμετρο σχήματος της ΓΑΤ μεγίστων με τα στατιστικά χαρακτηριστικά του δείγματος (ουσιαστικά μόνο με τον συντελεστή ασυμμετρίας  $C_{sx}$ ) είναι της μορφής:

$$\operatorname{sgn}(\kappa) \frac{\Gamma(1-3\kappa) - 3\Gamma(1-2\kappa)\Gamma(1-\kappa) + 2\Gamma^3(1-\kappa)}{[\Gamma(1-2\kappa) - \Gamma^2(1-\kappa)]^{3/2}} = C_{sx} \quad 9.106$$

Σε αυτό το σημείο έχουμε τις εξής δυνατότητες:

- Λύνουμε αριθμητικά την σχέση (9.106) ως προς  $\kappa$ . Σε αυτήν την περίπτωση οι βαθμοί ελευθερίας του προβλήματος είναι 3 (3 ανεξάρτητες παράμετροι της κατανομής)
- Θεωρούμε μία τιμή του  $\kappa$  και την θέτουμε ρητά ως παράμετρο της κατανομής. Σε αυτήν την περίπτωση αγνοούμε την δειγματική τιμή του συντελεστή ασυμμετρίας. Ο Δ. Κουτσογιάννης προτείνει μία καταρχήν εκτίμηση του συντελεστή  $\kappa = 0.15$  η οποία φαίνεται να ταιριάζει καλά σε ακραία γεγονότα βροχόπτωσης. Σε αυτήν την περίπτωση οι βαθμοί ελευθερίας του προβλήματος είναι 2 (2 ανεξάρτητες παράμετροι της κατανομής). Ουσιαστικά σε αυτήν την περίπτωση (και εφόσον φυσικά  $\kappa > 0$ ), η κατανομή ΓΑΤ παίρνει την μορφή μίας AT-2 μεγίστων.

Εφόσον λοιπόν έχουμε κάποια τιμή για το  $\kappa$  με έναν από τους δύο παραπάνω τρόπους υπολογίζουμε τις τιμές των  $\lambda$  και  $\psi$  ως εξής:

$$\lambda = \frac{|\kappa| \sigma_x}{\sqrt{\Gamma(1-2\kappa) - \Gamma^2(1-\kappa)}} \quad 9.107$$

$$\psi = \frac{\mu_x}{\lambda} - \frac{\Gamma(1-\kappa) - 1}{\kappa} \quad 9.108$$

Οι παραπάνω σχέσεις γίνονται απροσδιόριστες για  $\kappa=0$ . Αν λοιπόν προκύψει  $\kappa=0$  (πρακτικά όταν  $|\kappa| \leq 0.001$ ), χρησιμοποιούνται οι αντίστοιχες σχέσεις για τα  $\lambda$ ,  $\psi$  για την κατανομή AT-1 μεγίστων (Gumbel max).

Η αριθμητική λύση της (9.106) γίνεται ακριβώς με την μεθοδολογία που εφαρμόζεται και για την AT-2 μεγίστων (βλ. σ. 95). Ως αρχικά όρια για την μεταβλητή  $\kappa$  θα λαμβάνονται  $\kappa_{\min} = -10$  (ένα αυθαίρετο όριο, πρακτικά δεν συναντάμε τόσο αρνητική τιμή του  $\kappa$ ) και  $\kappa_{\max} = 1/3$  (το οποίο προκύπτει από την απαίτηση να έχουμε όρισμα στην συνάρτηση  $\Gamma$  θετικό).

Αν  $i$  ο κύκλος υπολογισμού, ο υπολογισμός θεωρείται επιτυχής μόλις ικανοποιηθεί κάποιο από τα παρακάτω κριτήρια:

$$\kappa_{\max} - \kappa_{\min} \leq \varepsilon \quad 9.109$$

$$i \geq i_{\max} \quad 9.110$$

όπου για τον «Υδρογνώμων» υιοθετούμε  $i_{\max}=100$  και  $\varepsilon=0.0005$ .

### Η κατανομή ΓΑΤ ελαχίστων

Η σχέση που συνδέει την παράμετρο σχήματος της ΓΑΤ ελαχίστων με τα στατιστικά χαρακτηριστικά του δείγματος είναι της μορφής:

$$\operatorname{sgn}(\kappa) \frac{\Gamma(1+3\kappa) - 3\Gamma(1+2\kappa)\Gamma(1+\kappa) + 2\Gamma^3(1+\kappa)}{[\Gamma(1+2\kappa) - \Gamma^2(1+\kappa)]^{3/2}} = C_{sx} \quad 9.111$$

Για την εύρεση του  $\kappa$  ισχύουν οι ίδιες παρατηρήσεις με την ΓΑΤ μεγίστων, δηλαδή, είτε θα επιλύουμε αριθμητικά την άνω σχέση είτε θα θέτουμε κάποια τιμή  $\kappa$  ρητά.

Εφόσον λοιπόν έχουμε κάποια τιμή για το  $\kappa$  με έναν από τους δύο παραπάνω τρόπους υπολογίζουμε τις τιμές των  $\lambda$  και  $\psi$  ως εξής:

$$\lambda = \frac{|\kappa|\sigma_x}{\sqrt{\Gamma(1+2\kappa) - \Gamma^2(1+\kappa)}} \quad 9.112$$

$$\psi = \frac{\mu_x}{\lambda} + \frac{1 - \Gamma(1+\kappa)}{\kappa} \quad 9.113$$

Οι παραπάνω σχέσεις γίνονται απροσδιόριστες για  $\kappa=0$ . Αν λοιπόν προκύψει  $\kappa=0$  (πρακτικά όταν  $|\kappa| \leq 0.001$ ), χρησιμοποιούνται οι αντίστοιχες σχέσεις για τα  $\lambda$ ,  $\psi$  για την κατανομή AT-1 ελαχίστων (Gumbel min).

Η αριθμητική λύση της (9.111) γίνεται ακριβώς με την μεθοδολογία που εφαρμόζεται και για την AT-2 μεγίστων (βλ. σ. 95) καθώς και ΓΑΤ μεγίστων. Ως αρχικά όρια για την μεταβλητή  $\kappa$  θα λαμβάνονται  $\kappa_{\min} = -1/3$  (το οποίο προκύπτει από την απαίτηση να έχουμε όρισμα στην συνάρτηση  $\Gamma$  θετικό) και  $\kappa_{\max} = 10$  (ένα αυθαίρετο όριο).

Αν  $i$  ο κύκλος υπολογισμού, ο υπολογισμός θεωρείται επιτυχής μόλις ικανοποιηθεί κάποιο από τα παρακάτω κριτήρια:

$$\kappa_{\max} - \kappa_{\min} \leq \varepsilon \quad 9.114$$

$$i \geq i_{\max} \quad 9.115$$

όπου για τον «Υδρογνώμων» υιοθετούμε  $i_{\max}=100$  και  $\varepsilon=0.0001$ .

### Η κατανομή Pareto τριών παραμέτρων

Η παράμετρος σχήματος της κατανομής Pareto  $\kappa$  συνδέεται με τον συντελεστή ασυμμετρίας του δείγματος  $C_{sx}$  με την εξής σχέση:

$$\frac{2(1-\kappa)\sqrt{1+2\kappa}}{1+3\kappa} = C_{sx} \quad 9.116$$

Η παραπάνω εξίσωση λύνεται αριθμητικά με την μέθοδο της εσφαλμένης θέσης (ακριβώς όπως στην AT-2 μεγίστων και στις ΓΑΤ), με όρια  $\kappa_{\min}=-1/3$  και  $\kappa_{\max}=1$  για θετική ασυμμετρία ( $C_{sx}>0$ ) και  $\kappa_{\min}=1$  και  $\kappa_{\max}=100$  για αρνητική ασυμμετρία ( $C_{sx}<0$ ). Ως μέγιστο επιτρεπόμενο σφάλμα σύγκλισης υιοθετείται η τιμή  $\varepsilon=0.0005$  ενώ ο μέγιστος αριθμός υπολογιστικών κύκλων  $i_{\max}=100$ .

Στην συνέχεια, αφού υπολογιστεί η παράμετρος σχήματος  $\kappa$ , οι παράμετροι κλίμακας και θέσης  $\lambda$  και  $\psi$  υπολογίζονται από τις σχέσεις:

$$\lambda = \sigma_x (1+\kappa)\sqrt{1+2\kappa} \quad 9.117$$



$$\psi = \frac{\mu_x}{\lambda} - \frac{1}{1 + \kappa} \quad 9.118$$

### Η έμμεση μέθοδος των ροπών – η περίπτωση της Log-Pearson III

Στην περίπτωση της Log-Pearson III εφαρμόζεται η έμμεση μέθοδος των ροπών. Εναλλακτικά θα μπορούσαμε να επιλύσουμε το μη γραμμικό σύστημα εξισώσεων που προκύπτει από τις μαθηματικές εκφράσεις της μέσης τιμής, διασποράς και της ροπής τρίτης τάξεως (Κουτσογιάννης 1997, σ. 152). Σύμφωνα λοιπόν με την έμμεση μέθοδο των ροπών, θα υπολογίζονται τα στατιστικά χαρακτηριστικά του μετασχηματισμένου δείγματος  $y_i = \ln x_i$ . Προφανώς θα λαμβάνονται υπόψη μόνο οι τιμές  $x_i > 0$  και αυτή πρέπει να είναι απαίτηση για το δείγμα (η Log-Pearson III έχει ως κάτω όριο της μεταβλητής την τιμή  $e^{\psi} > 0$ ).

Αφού προσδιοριστούν τα στατιστικά μεγέθη του μετασχηματισμένου δείγματος  $y_i$  ήτοι μέση τιμή, τυπική απόκλιση και ασυμμετρία ( $\mu_y$ ,  $\sigma_y$  και  $C_{sy}$ ) εφαρμόζονται οι σχέσεις της Pearson III:

$$\kappa = \frac{4}{C_{sy}} \quad 9.119$$

$$\lambda = \frac{\sqrt{\kappa}}{\sigma_y} \quad 9.120$$

$$\psi = \mu_y - \frac{\kappa}{\lambda} \quad 9.121$$

### Προσεγγιστικές σχέσεις εκτίμησης των παραμέτρων για τις AT-2 μεγίστων, AT-3 ελαχίστων και ΓΑΤ με την μέθοδο των ροπών

Στις προηγούμενες παραγράφους παρουσιάστηκαν μεθοδολογίες για την εκτίμηση της παραμέτρου σχήματος  $\kappa$  των κατανομών AT-2 μεγίστων, AT3-ελαχίστων και ΓΑΤ με την μέθοδο των ροπών υιοθετώντας την ακριβή λύση και εφαρμόζοντας αριθμητικές μεθόδους. Οι μεθοδολογίες αυτές έχουν το μειονέκτημα της χαμηλής απόδοσης σε ταχύτητα υπολογισμών. Έτσι θα διατυπωθούν μερικές προσεγγίσεις που θα δίνουν την παράμετρο  $\kappa$  για αυτές τις κατανομές σε ένα υπολογιστικό βήμα και θα χρησιμοποιούνται στην στοχαστική προσομοίωση (π.χ. Monte-Carlo). Οι προσεγγίσεις αυτές είναι πρωτότυπες και διατυπώθηκαν από τον Δ. Κουτσογιάννη.

#### Η προσέγγιση του $\kappa$ για την AT-2 μεγίστων

$$\kappa = \frac{1}{2} - \frac{1}{e^{2.59[\ln(C_v^2+1)]^{0.475}} + 1} \quad 9.122$$

όπου  $C_v$  είναι ο συντελεστής μεταβλητότητας:

$$C_v = \frac{\sigma_x}{\mu_x} \quad 9.123$$

#### Η προσέγγιση του $\kappa$ για την AT-3 ελαχίστων (Weibull)

$$\kappa = 2.56 \left\{ e^{0.41[\ln(C_v^2+1)]^{0.58}} - 1 \right\} \quad 9.124$$

όπου  $C_v$  ο συντελεστής μεταβλητότητας όπως ορίζεται στην Σχέση (9.123).

Η προσέγγιση του  $\kappa$  για την ΓΑΤ μεγίστων

$$\kappa = \frac{1}{3} - \frac{1}{0.31 + 0.91C_{sx} + \sqrt{(0.91C_{sx})^2 + 1.8}} \quad 9.125$$

Η προσέγγιση του  $\kappa$  για την ΓΑΤ ελαχίστων

$$\kappa = \frac{1}{0.28 - 0.9C_{sx} + 0.998\sqrt{(0.9C_{sx})^2 + 1.93}} - \frac{1}{3} \quad 9.126$$

### Η εκτίμηση των παραμέτρων των κατανομών με την μέθοδο των L-ροπών

Όπως αναφέρθηκε, η εκτίμηση με την μέθοδο των L-ροπών γίνεται μέσω απλών σχετικά σχέσεων που έχουν κλειστή μορφή. Υπενθυμίζουμε πως η  $\lambda_1$  ταυτίζεται με την μέση τιμή, η  $\lambda_2$  σχετίζεται με την διασπορά του δείγματος ενώ η  $\lambda_3$  με την ασυμμετρία. Συνεπώς, η  $\lambda_1$  και η  $\lambda_2$  υπεισέρχονται πάντα (όπως συμβαίνει με τις  $\mu_x$ ,  $\sigma_x$  στην μέθοδο των ροπών) ενώ η  $\lambda_3$  συσχετίζεται με την παράμετρο σχήματος ( $\kappa$  συνήθως) όταν η κατανομή προβλέπει τρεις παραμέτρους. Οι εκτιμήσεις των παραμέτρων για τις κατανομές ακροτάτων περιγράφονται από τον Κουτσογιάννη (1999, σ. 27).

Αρχικά ορίζουμε τον συντελεστή μεταβλητότητας  $\tau_2$  καθώς και τον συντελεστή ασυμμετρίας  $\tau_3$  ως:

$$\tau_2 = \frac{\lambda_2}{\lambda_1} \quad 9.127$$

$$\tau_3 = \frac{\lambda_3}{\lambda_2} \quad 9.128$$

Η κανονική κατανομή

$$\mu = \lambda_1, \quad \sigma = \sqrt{\pi} \lambda_2 \quad 9.129$$

Η εκθετική κατανομή

$$\lambda = \frac{1}{2\lambda_2}, \quad \psi = \lambda_1 - 2\lambda_2 \quad 9.130$$

Η κατανομή AT-1 μεγίστων (Gumbel μεγίστων)

$$\lambda = \frac{\lambda_2}{\ln 2}, \quad \psi = \frac{\lambda_1}{\lambda} - 0.5772 \quad 9.131$$

Η κατανομή AT-2 μεγίστων

$$\kappa = \frac{\ln(1 + \tau_2)}{\ln 2} \quad 9.132$$

$$\lambda = \lambda_1 \frac{\kappa}{\Gamma(1 - \kappa)} \quad 9.133$$

Η κατανομή ΓΑΤ μεγίστων

$$\kappa = 7.8c - 1.43c^2, \text{ όπου } c = \frac{\ln 2}{\ln 3} - \frac{2}{3 + \tau_3} \quad 9.134$$

$$\lambda = \frac{\kappa \lambda_2}{\Gamma(1 - \kappa)(2^\kappa - 1)} \quad 9.135$$

$$\psi = \frac{\lambda_1}{\lambda} - \frac{\Gamma(1 - \kappa) - 1}{\kappa} \quad 9.136$$

Όπως και στην μέθοδο των ροπών, έτσι και στην μέθοδο των L-ροπών η παράμετρος  $\kappa$  μπορεί να υπολογιστεί από τα στατιστικά χαρακτηριστικά του δείγματος (L-ροπές στην συγκεκριμένη περίπτωση) ή να τεθεί ρητά σε κάποια τιμή π.χ.  $\kappa=0.15$ . Οι παράμετροι  $\lambda$  και  $\psi$  υπολογίζονται στην συνέχεια βάσει της τιμής του  $\kappa$  που έχει θεωρηθεί.

Η κατανομή AT-1 ελαχίστων (Gumbel ελαχίστων)

$$\lambda = \frac{\lambda_2}{\ln 2}, \quad \psi = \frac{\lambda_1}{\lambda} + 0.5772 \quad 9.137$$

Η κατανομή AT-3 ελαχίστων (Weibull)

$$\kappa = \frac{-\ln(1 - \tau_2)}{\ln 2} \quad 9.138$$

$$\lambda = \lambda_1 \frac{\kappa}{\Gamma(1 + \kappa)} \quad 9.139$$

Η κατανομή ΓΑΤ ελαχίστων

$$\kappa = 7.8c + 4.71c^2, \text{ όπου } c = \frac{2}{3 - \tau_3} - \frac{\ln 2}{\ln 3} \quad 9.140$$

$$\lambda = \frac{\kappa \lambda_2}{\Gamma(1 + \kappa)(1 - 2^{-\kappa})} \quad 9.141$$

$$\psi = \frac{\lambda_1}{\lambda} + \frac{1 - \Gamma(1 + \kappa)}{\kappa} \quad 9.142$$

Ισχύει τυπικά η παρατήρηση για την παράμετρο  $\kappa$  (θα υπολογίζεται από την παραπάνω σχέση ή θα τίθεται ρητά σε κάποια τιμή).

Η κατανομή Pareto τριών παραμέτρων

$$\kappa = \frac{1 - 3\tau_3}{1 + \tau_3} \quad 9.143$$

$$\lambda = \lambda_2(1 + \kappa)(2 + \kappa) \quad 9.144$$

$$\psi = \frac{\lambda_1}{\lambda} - \frac{1}{1 + \kappa} \quad 9.145$$

## 9.5 Μετασχηματισμοί - απεικόνιση

Τα σημεία των μετρήσεων απεικονίζονται μέσω των εμπειρικών συναρτήσεων κατανομής ως ζεύγος τιμών (πιθανότητας μη-υπέρβασης, τιμή μεγέθους). Αντίστοιχα απεικονίζονται και οι θεωρητικές κατανομές, ως ένα σημειοσύνολο δηλαδή τιμών  $(F(x), x)$ . Η απεικόνιση των σημείων αυτών μέσω κάποιου μετασχηματισμού ξεφεύγει από τα όρια του περιβάλλοντος επικοινωνίας (user interface) και είναι ένα θέμα της στατιστικής ανάλυσης. Ο σκοπός του μετασχηματισμού είναι η καλύτερη απεικόνιση του δείγματος και των συναρτήσεων κατανομής, όσο το δυνατόν ευθειοποιημένα ώστε να εξετάζεται καλύτερα η προσαρμογή του δείγματος στις συγκεκριμένες συναρτήσεις κατανομής. Οι μετασχηματισμοί αυτοί λέγονται και «χαρτιά κατανομών» λόγω της ύπαρξης τέτοιων χαρτιών που ο μηχανικός είχε την δυνατότητα να σχεδιάσει με μολύβι τα σημεία του δείγματος.

Ο «Υδρογνώμων» παρέχει τα εξής «χαρτιά κατανομών»:

- Απλή γραμμική κλίμακα (χωρίς μετασχηματισμό)
- Κανονικής κατανομής
- Λογαριθμοκανονικής κατανομής
- Κατανομής Gumbel (AT-1) μεγίστων
- Κατανομής Gumbel (AT-1) ελαχίστων
- Κατανομής Weibull (AT-3) ελαχίστων
- Κατανομής ΓΑΤ μεγίστων (θέτοντας την παράμετρο σχήματος  $\kappa$ )
- Κατανομής ΓΑΤ ελαχίστων (θέτοντας την παράμετρο σχήματος  $\kappa$ ).

Το χαρτί της λογαριθμοκανονικής κατανομής έχει τον ίδιο μετασχηματισμό με αυτό της κανονικής κατανομής με την διαφορά πως οι τιμές  $x$  απεικονίζονται με λογαριθμικό μετασχηματισμό (αυτή η δυνατότητα παρέχεται και από το περιβάλλον επικοινωνίας). Το ίδιο συμβαίνει και με το χαρτί Weibull το οποίο έχει τον ίδιο μετασχηματισμό με το χαρτί της Gumbel ελαχίστων.

Όσον αφορά τα χαρτιά της κατανομής ΓΑΤ υπάρχει δυνατότητα εισαγωγής της παραμέτρου σχήματος  $\kappa$  από τον χρήστη, ωστόσο παρέχεται αρχικοποιημένη τιμή  $\kappa=0.15$ . Στον Πίνακα 9.10 παρουσιάζονται οι ορθοί (από την πιθανότητα υπέρβασης  $F_{1x}$  στην γραμμικοποιημένη μεταβλητή  $z$  που θα χρησιμοποιείται στην απεικόνιση) καθώς και οι αντίστροφοι μετασχηματισμοί (από  $z$  στην συνάρτηση κατανομής  $F$ ).

Πίνακας 9.10: Μετασχηματισμοί πιθανοτήτων για την απεικόνιση σε «χαρτιά πιθανοτήτων»

«Χαρτί» κατανομής	Ορθός μετασχηματισμός $F_{1x} \rightarrow z$	Αντίστροφος μετασχηματισμός $z \rightarrow F$
Γραμμικό	$z = 1 - F_{1x}$	$F = z$
Κανονική (και λογαριθμοκανονική)	$z = N_{1x}^{-1}(F_{1x}, \mu, \sigma), \mu=0, \sigma=1$	$F = N(z, \mu, \sigma), \mu=0, \sigma=1$
AT-1 μεγίστων	$z = -\ln[-\ln(1 - F_{1x})]$	$F = e^{-e^{-z}}$
AT-1 ελαχίστων (και AT-3 ελαχίστων)	$z = \ln[-\ln(-F_{1x})]$	$F = 1 - e^{-e^z}$
ΓΑΤ μεγίστων	$z = \frac{[-\ln(1 - F_{1x})]^{-\kappa} - 1}{\kappa}$	$F = e^{-(1+z\kappa)^{-1/\kappa}}$
ΓΑΤ ελαχίστων	$z = \frac{[-\ln(F_{1x})]^\kappa - 1}{\kappa}$	$F = 1 - e^{-(1+z\kappa)^{1/\kappa}}$

## 9.6 Στατιστικοί έλεγχοι

Ο έλεγχος καταλληλότητας με την στατιστική θεωρία ελέγχου υποθέσεων είναι ένα εργαλείο το οποίο μας επιτρέπει να απορρίψουμε ή όχι μία θεωρητική συνάρτηση κατανομής για την περιγραφή του δείγματος. Συγκεκριμένα ελέγχουμε την μηδενική υπόθεση  $H_0$ . Αν η  $H_0$  απορριφθεί ισχύει μία εναλλακτική υπόθεση  $H_1$ :

- $H_0$ : Η θεωρητική συνάρτηση κατανομής που έχει υποτεθεί ταυτίζεται με την πραγματική συνάρτηση κατανομής για κάθε σημείο  $j$  του δείγματος  $x_j$
- $H_1$ : Η θεωρητική συνάρτηση κατανομής που έχει υποτεθεί διαφέρει από την πραγματική συνάρτηση κατανομής για κάποια σημεία  $j$  του δείγματος  $x_j$ .

Αν η  $H_0$  δεν απορρίπτεται θα λέμε πως «η συνάρτηση κατανομής δεν απορρίπτεται για επίπεδο σημαντικότητας  $\alpha$ ». Σε αντίθετη περίπτωση θα λέμε «η συνάρτηση κατανομής απορρίπτεται για επίπεδο σημαντικότητας  $\alpha$ ». Ο «Υδρογνώμων» διατυπώνει αυτές τις φράσεις απλά ως «Απορρίπτεται» ή «Δεν απορρίπτεται». Το επίπεδο σημαντικότητας ορίζεται ως ποσοστό τάξης μεγέθους π.χ. 1% ή 5%.

### 9.6.1 Η δοκιμή $\chi^2$

Η δοκιμή  $\chi^2$  είναι από τους πλέον διαδεδομένους στατιστικούς ελέγχους. Εξετάζεται η παραλλαγή των «σοπίθανων κλάσεων» (Κουτσογιάννης, 1997, σ. 126). Τα υπολογιστικά βήματα είναι τα παρακάτω:

1. Επιλέγονται  $k$  κλάσεις. Για το  $k$  ισχύουν οι εξής περιορισμοί:
  - Πρέπει  $k \geq r+2$ , όπου  $r$  ο αριθμός των παραμέτρων της κατανομής που εκτιμώνται από το δείγμα (π.χ.  $r=2$  για την κανονική κατανομή,  $r=3$  για την Pearson III)
  - Το θεωρητικό δυναμικό των κλάσεων να είναι μεγαλύτερο από 5, το οποίο μεταφράζεται στην σχέση  $k \leq n/5$  ( $n$  το μέγεθος του δείγματος).

Ο «Υδρογνώμων» πραγματοποιεί τους παραπάνω ελέγχους για το  $k$  πριν την διεξαγωγή του ελέγχου. Επιπλέον δίνει μία προτεινόμενη τιμή του  $k$  ως:

$$k_{\text{προτεινόμενο}} = \max[ r+2, (n\%5+r+2)\%2 ] \quad 9.146$$

Μία μέση τιμή δηλαδή των δύο περιορισμών με ελάχιστο τον πρώτο. Το σύμβολο % αναφέρεται στην ακέραια διαίρεση. Εφόσον το  $k$  τίθεται μαζικά για ελέγχους σε πολλές συναρτήσεις κατανομής μπορούμε να δεχτούμε την τιμή  $r=2$  για το  $k_{\text{προτεινόμενο}}$ . Τελικά ο χρήστης αποφασίζει αν θα δεχθεί το  $k_{\text{προτεινόμενο}}$  ή θα τροποποιήσει αυτήν την τιμή.

2. Χωρίζουμε το διάστημα πιθανότητας  $[0, 1]$  σε  $k$  ίσα υποδιαστήματα με όρια  $u_j=j/k$  ( $j=0, \dots, k$ ).
3. Υπολογίζουμε τα όρια  $x_j$  των κλάσεων μέσω της αντιστρόφου συνάρτησης κατανομής:  $x_j=F^{-1}(u_j)$ .
4. Υπολογίζουμε το πραγματικό δυναμικό  $n_j$  της κλάσης καταμετρώντας τις τιμές του δείγματος που βρίσκονται στα όρια αυτής της κλάσης  $n_j=\text{count}(x_{j-1} \leq x_i < x_j)$ , όπου  $x_i$  οι τιμές του δείγματος. Η διαδικασία αυτή απλοποιείται αν προηγουμένως το δείγμα έχει ταξινομηθεί.
5. Υπολογίζεται η στατιστική συνάρτηση  $q$  του Pearson:

$$q = \frac{k}{n} \sum_{j=1}^k n_j^2 - n \quad 9.147$$

6. Για ένα αποδεκτό επίπεδο σημαντικότητας  $\alpha$  (π.χ. 1%, 5% ή 10%) υπολογίζουμε την κρίσιμη τιμή της στατιστικής συνάρτησης ελέγχου  $q_c=q_{1-\alpha}$ . Για το σκοπό αυτό χρησιμοποιούμε την κατανομή  $\chi^2$  για  $k-r-1$  βαθμούς ελευθερίας όπου  $r$  είναι ο αριθμός των παραμέτρων της κατανομής που εκτιμώνται από το δείγμα.
7. Απορρίπτουμε τη μηδενική υπόθεση αν  $q > q_c$ .

Η υλοποίηση της συνάρτησης κατανομής  $\chi^2$  καθώς και της αντίστροφης συνάρτησης κατανομής βασίζεται στην κατανομή γάμα (καθώς και στην αντίστοιχη αντίστροφη συνάρτηση κατανομής), όπου

ως παράμετρος σχήματος  $\kappa$  τίθεται  $\kappa=m/2$  και ως παράμετρο κλίμακας  $\lambda$  τίθεται  $\lambda=1/2$ , όπου  $m$  οι βαθμοί ελευθερίας της κατανομής  $\chi^2$ .

Ο «Υδρογνώμων» εκτός από τον έλεγχο  $q>q_c$  για τις τρεις τυπικές τιμές του  $a$  (1%, 5% και 10%) θα επιστρέφει την παράμετρο Pearson καθώς και το επιτευχθέν ποσοστό σημαντικότητας  $a_{\text{επιτευχθέν}}=1-\chi^2(m=k-r-1, q)$  για την υπολογισμένη παράμετρο Pearson  $q$ , ώστε να έχουμε μία καλύτερη αίσθηση για την απόρριψη ή μη της στατιστικής υπόθεσης.

Πίνακας 9.11: Ενδεικτικά αποτελέσματα από την εφαρμογή του ελέγχου  $\chi^2$  στο δείγμα βροχοπτώσεων Οκτωβρίου του σταθμού «Αλιάρτος». Μέγεθος δείγματος  $n=93$ , κλάσεων  $k=11$ .

Έλεγχος $\chi^2$ για Οκτώβριος	$a=1\%$	$a=5\%$	$a=10\%$	Επιτευχθέν $a$	Παράμ. Pearson $q$
Κανονική	ΑΠΟΡΡΙΨΗ	ΑΠΟΡΡΙΨΗ	ΑΠΟΡΡΙΨΗ	0.0%	40.06
Κανονική (L-Ροπές)	ΑΠΟΡΡΙΨΗ	ΑΠΟΡΡΙΨΗ	ΑΠΟΡΡΙΨΗ	0.0%	39.59
Λογαριθμοκανονική	ΑΠΟΡΡΙΨΗ	ΑΠΟΡΡΙΨΗ	ΑΠΟΡΡΙΨΗ	0.4%	22.80
Galton	ΔΕΝ ΑΠΟΡ.	ΔΕΝ ΑΠΟΡ.	ΔΕΝ ΑΠΟΡ.	13.0%	11.20
Εκθετική	ΔΕΝ ΑΠΟΡ.	ΔΕΝ ΑΠΟΡ.	ΔΕΝ ΑΠΟΡ.	21.7%	10.73
Εκθετική (L-Ροπές)	ΔΕΝ ΑΠΟΡ.	ΔΕΝ ΑΠΟΡ.	ΔΕΝ ΑΠΟΡ.	26.4%	10.02
Γάμμα	ΔΕΝ ΑΠΟΡ.	ΔΕΝ ΑΠΟΡ.	ΑΠΟΡΡΙΨΗ	9.4%	13.57
Pearson III	ΔΕΝ ΑΠΟΡ.	ΑΠΟΡΡΙΨΗ	ΑΠΟΡΡΙΨΗ	4.6%	14.28
Log Pearson III	ΑΠΟΡΡΙΨΗ	ΑΠΟΡΡΙΨΗ	ΑΠΟΡΡΙΨΗ	0.0%	33.20
AT1-Max (Gumbel)	ΔΕΝ ΑΠΟΡ.	ΔΕΝ ΑΠΟΡ.	ΔΕΝ ΑΠΟΡ.	21.7%	10.73

### 9.6.2 Ο έλεγχος Kolmogorov-Smirnov

Κατά τον έλεγχο Kolmogorov-Smirnov (Τσακίρης, 1995), ελέγχεται η μέγιστη διαφορά μεταξύ της θεωρητικής και της εμπειρικής συνάρτησης κατανομής ως εξής: για κάθε τιμή  $x_i$  από το δείγμα υπολογίζεται η ποσότητα:

$$D_i = |F_0(x_i) - F(x_i)| \quad 9.148$$

όπου  $F_0$  η εμπειρική και  $F$  η θεωρητική συνάρτηση κατανομής. Ως εμπειρική συνάρτηση κατανομής θα χρησιμοποιούμε την εμπειρική κατανομή του Weibull (ενότητα 9.3). Τελικά υπολογίζεται η παράμετρος  $D_{\max}$  ως η μέγιστη τιμή  $D_i$ :

$$D_{\max} = \max(D_i) \quad 9.149$$

Η τιμή  $D_{\max}$  συγκρίνεται με μία κρίσιμη τιμή  $D_{cr}$  σύμφωνα και με το επιθυμητό επίπεδο σημαντικότητας  $a$ . Η υπόθεση  $H_0$  απορρίπτεται αν:

$$D_{\max} \geq D_{cr} \quad 9.150$$

Υπολογιστικά αυτό θα αντιμετωπίζεται αντίστροφα, δηλαδή για δεδομένο  $D_{\max}$  του δείγματος και της θεωρητικής συνάρτησης κατανομής που έχει προσαρμοστεί, θα υπολογίζεται το επίπεδο σημαντικότητας Kolmogorov-Smirnov  $a_D$ . Αν το υπολογισμένο επίπεδο σημαντικότητας  $a_D$  είναι μικρότερο από το επιθυμητό  $a$  τότε η απορρίπτουμε την υπόθεση  $H_0$ .

Το επίπεδο σημαντικότητας υπολογίζεται από την σχέση (η διερεύνηση για τον υπολογισμό του  $a_D$  από τον Δ. Κουτσογιάννη):

$$a_D = 1 - L(z) \quad 9.151$$

όπου:

$$L(z) = \frac{\sqrt{2\pi}}{z} \sum_{k=1}^{\infty} e^{-\frac{\sqrt{(2k-1)\pi}}{z}} \quad \text{αν } 0.3 \leq z \leq 8.5 \quad 9.152$$

$$L(z)=0 \text{ αν } z < 0.3 \text{ και } L(z)=1 \text{ αν } z > 8.5$$

Για τους φθίνοντες όρους της απειροσειράς που δίνουν το  $L(z)$  σταματάμε τον υπολογισμό περαιτέρω όρων όταν γίνουν μικρότεροι από ένα επίπεδο ακρίβειας  $\varepsilon$ . Στον «Υδρογνώμων» θέτουμε αυτό το επίπεδο ακρίβειας στην τιμή  $\varepsilon=10^{-12}$ . Η παράμετρος  $z$  μπορεί να υπολογιστεί από την σχέση:

$$z = D_{\max} \sqrt{n} + c \quad 9.153$$

όπου  $n$  το μέγεθος του δείγματος. Ο διορθωτικός όρος  $c$  εισάγεται όταν το δείγμα είναι μικρότερο από 50:

$$c = \begin{cases} 0, & n \geq 50 \\ \frac{50-n}{500}, & n < 50 \end{cases} \quad 9.154$$

Πίνακας 9.12: Ενδεικτικά αποτελέσματα από την εφαρμογή του ελέγχου Kolmogorov - Smirnov στο δείγμα βροχοπτώσεων Οκτωβρίου του σταθμού «Αλιάρτος». Μέγεθος δείγματος  $n=93$ .

<b>Kolmogorov-Smirnov test για: Οκτώβριος</b>	<b><math>\alpha=1\%</math></b>	<b><math>\alpha=5\%</math></b>	<b><math>\alpha=10\%</math></b>	<b>Επιτευχθέν <math>\alpha</math></b>	<b><math>D_{\max}</math></b>
Κανονική	ΔΕΝ ΑΠΟΡ.	ΑΠΟΡΡΙΨΗ	ΑΠΟΡΡΙΨΗ	1.3%	0.16
Κανονική (L-Ροπές)	ΔΕΝ ΑΠΟΡ.	ΑΠΟΡΡΙΨΗ	ΑΠΟΡΡΙΨΗ	1.4%	0.16
Λογαριθμοκανονική	ΔΕΝ ΑΠΟΡ.	ΔΕΝ ΑΠΟΡ.	ΔΕΝ ΑΠΟΡ.	50.3%	0.09
Galton	ΔΕΝ ΑΠΟΡ.	ΔΕΝ ΑΠΟΡ.	ΔΕΝ ΑΠΟΡ.	58.9%	0.08
Εκθετική	ΔΕΝ ΑΠΟΡ.	ΔΕΝ ΑΠΟΡ.	ΔΕΝ ΑΠΟΡ.	83.9%	0.06
Εκθετική (L-Ροπές)	ΔΕΝ ΑΠΟΡ.	ΔΕΝ ΑΠΟΡ.	ΔΕΝ ΑΠΟΡ.	64.9%	0.08
Γάμμα	ΔΕΝ ΑΠΟΡ.	ΔΕΝ ΑΠΟΡ.	ΔΕΝ ΑΠΟΡ.	92.5%	0.06
Pearson III	ΔΕΝ ΑΠΟΡ.	ΔΕΝ ΑΠΟΡ.	ΔΕΝ ΑΠΟΡ.	71.9%	0.07
Log Pearson III	ΑΠΟΡΡΙΨΗ	ΑΠΟΡΡΙΨΗ	ΑΠΟΡΡΙΨΗ	0.1%	0.21
AT1-Max (Gumbel)	ΔΕΝ ΑΠΟΡ.	ΔΕΝ ΑΠΟΡ.	ΔΕΝ ΑΠΟΡ.	38.9%	0.09

## 9.7 Εκτίμηση των διαστημάτων εμπιστοσύνης μέσω στατιστικής προσομοίωσης (Monte – Carlo).

Στην ενότητα αυτή περιγράφονται κάποιες μεθοδολογίες για την εκτίμηση των διαστημάτων εμπιστοσύνης. Οι κλασικές μεθοδολογίες της στατιστικής βασίζονται σε κάποιες σχέσεις κλειστής μορφής οι οποίες δεν μπορούν να διατυπωθούν για κάθε συνάρτηση κατανομής, επιπλέον μπορούν να ισχύσουν μόνο όταν έχουμε δείγμα ικανού μεγέθους ώστε να μπορούμε να δεχθούμε πως η διακύμανση του δείγματος προσεγγίζει την διακύμανση του πληθυσμού.

Οι μέθοδοι της στατιστικής προσομοίωσης μπορούν να δώσουν μία λύση σε αυτό το πρόβλημα αρκετά αξιόπιστη και με εφαρμογή σε κάθε περίπτωση κατανομής – παραμέτρων (π.χ. χρήση L-ροπών). Οι μέθοδοι αυτοί βασίζονται στην πραγματοποίηση κάποιων «πειραμάτων» ώστε να προσπαθήσουμε να αναπαραστήσουμε το φυσικό πρόβλημα (μέθοδοι Monte-Carlo). Τα πειράματα θα είναι «τεχνητά» δείγματα που θα παράγονται από γεννήτριες τυχαίων αριθμών που θα περιγραφούν στην συνέχεια.

### 9.7.1 Γεννήτριες τυχαίων αριθμών

Μία γεννήτρια ψευδοτυχαίων  $R(a,b)$  αριθμών της ομοιόμορφης κατανομής είναι μία ακολουθία αριθμών  $x_i$  στο διάστημα  $[a,b]$  όπου η συνάρτηση πυκνότητας πιθανότητας  $f_u$  είναι μία σταθερή ευθεία:

$$f_u(x) = \frac{1}{b-a} \quad 9.155$$

Τέτοιες ακολουθίες παρέχοντα από τα περισσότερα υπολογιστικά συστήματα. Στον «Υδρογνώμων» χρησιμοποιείται αυτή της *Borland Delphi*. Η ακολουθία είναι συγκεκριμένη για κάθε αριθμό αρχικοποίησης. Στις προσομοιώσεις χρησιμοποιούμε κάποια σταθερή τιμή αρχικοποίησης ώστε να έχουμε κάποια «επαναληψιμότητα» στα αποτελέσματα. Αν ο όρος της αρχικοποίησης επιλεγεί τυχαία (π.χ. από την μέτρηση της ώρας του υπολογιστή κατά την ώρα εκτέλεσης του προγράμματος) τότε η ακολουθία θα είναι μία ακολουθία πραγματικά τυχαίων αριθμών (η εφαρμογή της θα είναι περιορισμένη στον «Υδρογνώμων» και μόνο όταν το ζητάει ο χρήστης παρακάμπτοντας την σταθερά αρχικοποίησης).

Έστω η γεννήτρια ψευδοτυχαίων αριθμών στο διάστημα  $[0,1]$   $R(0,1)$  ή  $R_{0-1}$ . Χρησιμοποιώντας την συνάρτηση ποσοστημορίου  $x(u)=F^{-1}(u)$  κάποιας κατανομής  $F$ , τότε η:

$$F^{-1}(R_{0-1}) \quad 9.156$$

Θα είναι μία γεννήτρια ψευδοτυχαίων αριθμών της κατανομής  $F$ . Έτσι οι τυχαίοι αριθμοί που παράγονται θα διατηρούν τις ιδιότητες της κατανομής  $F$  με τις παραμέτρους της.

Έτσι για όλες τις κατανομές που υπολογίζονται με κλειστές σχέσεις, δηλαδή η εκθετική κατανομή, όλες οι ΑΤ καθώς και η ΓΑΤ και η κατανομή Pareto, ως γεννήτριες ψευδοτυχαίων αριθμών θα χρησιμοποιείται η σχέση (9.156). Για τις συναρτήσεις κατανομής όπου χρησιμοποιούνται μη-κλειστές σχέσεις (συναρτήσεις κατανομής οικογένειας κανονικής κατανομής και κατανομής γάμα) δεν θα χρησιμοποιείται η (9.156)· κάτι τέτοιο θα είχε πολύ κακή απόδοση, ειδικά όταν απαιτείται ένας αριθμός τυχαίων τιμών της τάξεως μεγέθους του  $10^5$  με  $10^6$  για τις εφαρμογές που θα εξετάσουμε. Έτσι θα διατυπωθούν δύο ειδικές γεννήτριες – μία για την κανονική κατανομή και μία για την κατανομή γάμα. Στην συνέχεια θα ακολουθείται η τακτική που αναλύθηκε στην παράγραφο 9.4.1, για την εύρεση τυχαίων τιμών κατανομών των οικογενειών τους (π.χ. από κανονική κατανομή → λογαριθμοκανονική, Galton ή από κατανομή γάμα → Pearson III ή Log-Pearson III).

#### Γεννήτρια ψευδοτυχαίων αριθμών της κανονικής κατανομής

Εφαρμόζεται η μεθοδολογία που εισήχθη από τους Box & Muller (1958) (βλ. Kottegoda, N.T., “Stochastic Water Resources Technology”. Έστω  $u_i, u_{i+1}$  δύο διαδοχικοί όροι της ακολουθίας ψευδοτυχαίων αριθμών της ομοιόμορφης κατανομής  $R_{0-1}$ . Θα δείξουμε πως υπολογίζουμε δύο διαδοχικούς όρους της ακολουθίας ψευδοτυχαίων αριθμών της κανονικής τιμής  $n_i, n_{i+1}$  με παραμέτρους  $\mu$  και  $\sigma$ .

Αρχικά υπολογίζουμε τις ποσότητες  $z$  και  $f$ :

$$z = \sqrt{-2 \ln u_i} \quad 9.157$$

$$f = 2\pi u_{i+1} \quad 9.158$$

Οι διαδοχικοί όροι της ακολουθίας ψευδοτυχαίων αριθμών της κανονικής κατανομής θα είναι:

$$n_i = \mu + \sigma z \cos f \quad 9.159$$

$$n_{i+1} = \mu + \sigma z \sin f \quad 9.160$$



## Γεννήτρια ψευδοτυχαίων αριθμών της κατανομής γάμα

Η γένεση ψευδοτυχαίων αριθμών που ακολουθούν την κατανομή γάμα είναι ένα πολύπλοκο πρόβλημα που απαιτεί αρκετή βελτιστοποίηση ώστε να έχουμε καλή ισορροπία μεταξύ απόδοσης και ακρίβειας. Στην συνέχεια παρουσιάζεται η αντιμετώπιση που προέκυψε από διερεύνηση που έχει κάνει ο Δ. Κουτσογιάννης. Η κατανομή γάμα έχει παραμέτρους  $\kappa$  και  $\lambda$  (παραμέτρος σχήματος και παράμετρος κλίμακας).

Αρχικά ορίζουμε τους αριθμούς  $\varepsilon$ ,  $N_L$  και  $I_L$  όπου αντιστοιχούν στο κατώφλι ακρίβειας, στο όριο χρήσης της κανονικής κατανομής και στο όριο χρήσης του ακεραίου τύπου. Στον «Υδρογνώμων» χρησιμοποιούμε τις τιμές:  $\varepsilon=10^{-5}$ ,  $N_L=100$  και  $I_L=25$ . Ο υπολογισμός λοιπόν ενός ψευδοτυχαίου αριθμού  $g_i$  που θα ακολουθεί την κατανομή γάμα θα γίνεται με την εξής διαδοχή:

- Αν  $\kappa > N_L$  τότε η κατανομή γάμα προσεγγίζεται από την κανονική κατανομή και επιστρέφεται μία τιμή από την ακολουθία ψευδοτυχαίων τιμών της κανονικής κατανομής με παραμέτρους  $\mu = \kappa/\lambda$  και  $\sigma = \sqrt{\kappa/\lambda}$ .
- Αν όμως  $\kappa \leq N_L$  τότε εξετάζεται αν το  $\kappa$  είναι ακέραιος με κάποια προσέγγιση  $\varepsilon$ ,  $|\lceil \kappa \rceil - \kappa| < \varepsilon$ . Αν είναι ακέραιος τότε αν ισχύει και η συνθήκη  $\kappa < I_L$  επιστρέφεται τιμή από την γεννήτρια της κατανομής γάμα για ακέραια  $\kappa$ .
- Αν δεν ισχύει η προηγούμενη συνθήκη, επιστρέφεται η τιμή της γεννήτριας Cheng & Feast εφόσον το  $\kappa > 1$ .
- Στις υπόλοιπες περιπτώσεις (δηλαδή αν  $0 < \kappa < 1$ ) επιστρέφεται η τιμή της γεννήτριας Ahrens & Dieter.

### Η γεννήτρια Ahrens & Dieter, 1974

Η γεννήτρια Ahrens-Dieter έχει εφαρμογή για  $0 < \kappa < 1$  (βλ. Ripley, 1987, p. 88).

Έστω δύο διαδοχικές τιμές της ακολουθίας ψευδοτυχαίων αριθμών της ομοιόμορφης κατανομής  $u_i, u_{i+1}$ .

Εάν ο  $u_0 > \frac{e}{e + \kappa}$  (όπου  $e$  η βάση των φυσικών λογαρίθμων) τότε υπολογίζουμε την ποσότητα  $u_i$  ως εξής:

$$x_i = -\ln \frac{(\kappa + e)(1 - u_i)}{\kappa e}, \quad u_i = \{x_i\}^{\kappa-1} \quad 9.161$$

Εάν  $u_0 \leq \frac{e}{e + \kappa}$  τότε υπολογίζουμε την ποσότητα  $u_i$  ως εξής:

$$x_i = \left[ \frac{(\kappa + e)u_i}{e} \right]^{1/\kappa}, \quad u_i = e^{-x_i} \quad 9.162$$

Ελέγχουμε αν  $u_{i+1} > u_i$ . Αν δεν ισχύει αυτή η συνθήκη επαναλαμβάνουμε τους παραπάνω υπολογισμούς μέχρι να ισχύσει. Τελικά η τιμή του ψευδοτυχαίου αριθμού που ακολουθεί την κατανομή γάμα είναι:

$$g_i = \frac{x_i}{\lambda} \quad 9.163$$

### Η γεννήτρια Cheng & Feast, 1979

Η γεννήτρια Cheng & Feast είναι κατάλληλη για τιμές  $\kappa > 1$  (βλ. Ripley, 1987, p. 90).

Ορίζουμε τις παρακάτω σταθερές:

$$c_1 = \kappa - 1, \quad c_2 = \frac{\kappa - 1}{6\kappa}, \quad c_3 = \frac{2}{c_1}, \quad c_4 = c_3 + 2, \quad c_5 = \frac{1}{\sqrt{\kappa}} \quad 9.164$$

Στην συνέχεια ακολουθείται ο παρακάτω αλγόριθμος:

ΕΠΑΝΕΛΛΑΒΕ

ΕΠΑΝΕΛΛΑΒΕ

Έστω δύο αριθμοί της ακολουθίας ψευδοτυχαίων αριθμών της ομοιόμορφης κατανομής  $u_i$  και  $u_{i+1}$ .

ΕΑΝ  $\kappa > 2.5$  ΤΟΤΕ  $u_i = u_{i+1} + c_5(1 - 1.86u_i)$

ΜΕΧΡΙ ΝΑ ΙΣΧΥΣΕΙ  $0 < u_i < 1$

Θέτουμε:  $w = \frac{c_2 u_{i+1}}{u_i}$

ΜΕΧΡΙ ΝΑ ΙΣΧΥΣΕΙ  $(c_3 u_i + w + \frac{1}{w} \leq c_4) \vee (c_3 \ln u_i - \ln w + w \leq 1)$

Τελικά η τιμή του ψευδοτυχαίου αριθμού που ακολουθεί την κατανομή γάμα είναι:

$$g_i = \frac{c_1 w}{\lambda} \quad 9.165$$

Η γεννήτρια της κατανομής γάμα για ακέραια  $\kappa$

Ο παρακάτω αλγόριθμος μπορεί να χρησιμοποιείται όταν το  $\kappa$  είναι ακέραιος αριθμός. Όπως είδαμε παραπάνω ο αλγόριθμος αυτός χάνει την απόδοσή του για μεγάλες τιμές του  $\kappa$ . Ένα λογικό όριο που θέτουμε είναι το  $\kappa=25$ . Έστω η ακολουθία  $\kappa$  ψευδοτυχαίων αριθμών της ομοιόμορφης κατανομής  $u_j$ . Ο ψευδοτυχαίος αριθμός της κατανομής γάμα μπορεί να υπολογιστεί από την σχέση:

$$g_i = \frac{-\ln\left(\prod_{j=1}^{\kappa} u_j\right)}{\lambda} \quad 9.166$$

## 9.7.2 Εύρεση διαστημάτων εμπιστοσύνης του δείγματος

Εφόσον είναι γνωστά τα στατιστικά χαρακτηριστικά του δείγματος, αν θεωρήσουμε πως οι δειγματικές τιμές τους αντιπροσωπεύουν τον πληθυσμό τότε θα είναι εύκολος ο προσδιορισμός των ορίων εμπιστοσύνης του ποσοστημορίου  $x(u)$  με την μέθοδο Monte-Carlo για οποιαδήποτε κατανομή  $F(x)$ .

Η έκφραση των ορίων εμπιστοσύνης για την συνάρτηση ποσοστημορίου της κατανομής  $F(x)$  γίνεται μέσω δύο συναρτήσεων, του κάτω ορίου  $x_L(u)$  και του άνω ορίου  $x_U(u)$ , έτσι ώστε για συγκεκριμένη πιθανότητα μη υπέρβασης  $u$ , να ισχύει:

$$P(x_L(u) \leq x \leq x_U(u)) = \gamma \quad 9.167$$

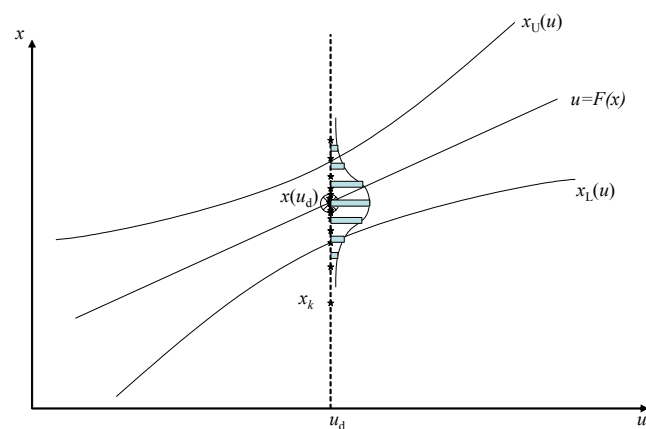
όπου  $\gamma$  το επίπεδο επιθυμητό εμπιστοσύνης. Βλέπουμε πως τα όρια εμπιστοσύνης είναι δύο συνεχείς συναρτήσεις που μπορούν να απεικονιστούν στο χαρτί κατανομής ως δύο καμπύλες που περικλείουν την καμπύλη της συνάρτησης κατανομής. Η πρακτική που θα ακολουθούμε θα είναι ο υπολογισμός του διαστήματος εμπιστοσύνης σε διακριτά  $u$  (σε 5 ή 10 ή 20 σημεία ανάλογα με την απαίτηση του χρήστη) και στην συνέχεια η χάραξη τεθλασμένων γραμμών που θα προσεγγίζουν τα όρια εμπιστοσύνης.

Στην συνέχεια θα παρουσιαστεί η μεθοδολογία που ακολουθείται για την εύρεση των ορίων εμπιστοσύνης σε μία τιμή της πιθανότητας μη-υπέρβασης  $u$ .

### Ορισμός του πειράματος

Πείραμα ονομάζουμε την διαδικασία που θα μας δώσει μία τιμή  $x_k$  που θα αντιστοιχεί στο επιθυμητό  $u_d$  και θα προκύψει από διαδικασίες γένεσης τυχαίων αριθμών. Οι τιμές  $x_k$  διαφόρων πειραμάτων διαφέρουν και έχουν διακύμανση και άλλα στατιστικά χαρακτηριστικά που εξαρτώνται από την κατανομή που έχει υιοθετηθεί (βλ. Σχήμα 9.6). Για να προκύψει η τιμή  $x_k$  ενός μοναδιαίου πειράματος ακολουθούμε τα εξής βήματα:

- Χρησιμοποιώντας την γεννήτρια τυχαίων αριθμών  $R_F$  της κατανομής  $F$ , παράγουμε ένα δείγμα τυχαίων αριθμών με  $n$  στοιχεία, όπου  $n$  ο αριθμός των στοιχείων του πραγματικού δείγματος
- Από το «τεχνητό δείγμα»  $y_i$  που παρήχθη, προσδιορίζουμε τα στατιστικά του χαρακτηριστικά σύμφωνα με αυτά που έχουν διατυπωθεί στην ενότητα 9.2, δηλαδή ροπές ή L-ροπές ανάλογα την μεθοδολογία που υπολογίζονται οι παράμετροι της  $F$
- Χρησιμοποιώντας ανάλογα την μέθοδο των ροπών ή των L-ροπών, υπολογίζονται οι παράμετροι της κατανομής (εφαρμόζοντας προσεγγιστικές σχέσεις αν υπάρχει αυτή η δυνατότητα βλ. παράγραφο 9.4.3)
- Από τις προσδιορισθέντες παραμέτρους του προηγούμενου βήματος και χρησιμοποιώντας την συνάρτηση ποσοστημορίου  $x(u) = F^{-1}(u)$ , υπολογίζουμε τελικά την τιμή του πειράματος:  $x_k = F^{-1}(u_d)$ .



Σχήμα 9.6: Όρια εμπιστοσύνης  $x_U$ ,  $x_L$ , πειράματα  $x_k$  και κατανομή τους περί του σημείου  $x(u_d)$ .

### Εύρεση δειγματικών ορίων

Έστω ένας ικανός αριθμός πειραμάτων  $m$  και  $x_{k1}, x_{k2}, \dots, x_{km}$  οι τιμές που προκύπτουν από τα πειράματα. Οι τιμές αυτές διατάσσονται όπως βλέπουμε και στο Σχήμα 9.6 γύρω από την κεντρική τιμή  $x(u_d)$ . Η δε κατανομή τους είναι άγνωστη αν και εξαρτάται από τα στατιστικά χαρακτηριστικά του πραγματικού δείγματος και από την κατανομή  $F$  που έχουμε υιοθετήσει. Το δείγμα αυτό με τα  $m$  πειράματα θα το λέμε δείγμα Monte-Carlo και από αυτό θα προκύπτουν άμεσα οι δειγματικές τιμές των ορίων. Η τάξη μεγέθους του  $m$  είναι  $10^4$ - $10^5$  ανάλογα με τον επιθυμητό βαθμό ακρίβειας.

Για να υπολογίσουμε τα  $x_U(u_d)$  και  $x_L(u_d)$  αρκεί να ταξινομήσουμε το δείγμα Monte-Carlo κατά αύξουσα σειρά και να λάβουμε ως  $x_U(u_d)$  και  $x_L(u_d)$  τα  $m(1-\gamma)/2$  και  $m(1+\gamma)/2$  στοιχεία του ταξινομημένου δείγματος Monte-Carlo αντίστοιχα:

$$x_L(u_d) = x_{k, m \frac{1-\gamma}{2}} \quad 9.168$$

$$x_U(u_d) = x_{k, m \frac{1+\gamma}{2}} \quad 9.169$$

Επειδή όμως τα  $m(1-\gamma)/2$  και  $m(1+\gamma)/2$  δεν είναι εν γένει ακέραιοι αριθμοί, κρατάμε το ακέραιο μέρος τους και παρεμβάλουμε το δεκαδικό μέρος χρησιμοποιώντας και το επόμενο στοιχείο. Οπότε οι παραπάνω σχέσεις μετασχηματίζονται ως:

$$x_L(u_d) = x_{k, \left[ m \frac{1-\gamma}{2} \right]} + \left( x_{k, \left[ m \frac{1-\gamma}{2} \right] + 1} - x_{k, \left[ m \frac{1-\gamma}{2} \right]} \right) \left( m \frac{1-\gamma}{2} - \left[ m \frac{1-\gamma}{2} \right] \right) \quad 9.170$$

$$x_U(u_d) = x_{k, \left[ m \frac{1+\gamma}{2} \right]} + \left( x_{k, \left[ m \frac{1+\gamma}{2} \right] + 1} - x_{k, \left[ m \frac{1+\gamma}{2} \right]} \right) \left( m \frac{1+\gamma}{2} - \left[ m \frac{1+\gamma}{2} \right] \right) \quad 9.171$$

όπου οι αγκύλες  $[\ ]$  συμβολίζουν το ακέραιο μέρος.

Μπορούμε να κάνουμε μία σύγκριση των τιμών που προκύπτουν με προσομοίωση ακολουθώντας την παραπάνω διαδικασία και κάποιας θεωρητικής λύσης. Χρησιμοποιήθηκε η θεωρητική λύση για τα όρια εμπιστοσύνης της μέσης τιμής ενός δείγματος (η οποία μέση τιμή, ακολουθεί την κανονική κατανομή εφόσον το  $n$  είναι αρκετά μεγάλο). Τα αποτελέσματα της σύγκρισης απεικονίζονται στον Πίνακα 9.13. Αν το επίπεδο εμπιστοσύνης είναι  $\gamma$  και  $z_{(1+\gamma)/2}$  η συνάρτηση ποσοστημορίου της τυποποιημένης κανονικής κατανομής στην τιμή  $u=(1+\gamma)/2$ , τότε τα όρια εμπιστοσύνης της μέσης τιμής είναι:

$$P\left( \mu_x - \frac{z_{(1+\gamma)/2} \sigma_x}{\sqrt{n}} < \bar{X} < \mu_x + \frac{z_{(1+\gamma)/2} \sigma_x}{\sqrt{n}} \right) = \gamma \quad 9.172$$

Πίνακας 9.13: Σύγκριση ορίων εμπιστοσύνης που προκύπτουν με την μέθοδο της προσομοίωσης και της θεωρητικής τιμής για την μέση τιμή (κανονική κατανομή,  $u_d=50\%$ , δείγμα  $x_i=i, i=1..19, n=19$ ).

$\gamma$	Προσομοίωση $m=30 \times 10^3$		Προσομοίωση $m=60 \times 10^3$		Προσομοίωση $m=120 \times 10^3$		Θεωρητική τιμή	
	$x_L$	$x_U$	$x_L$	$x_U$	$x_L$	$x_U$	$x_L$	$x_U$
90%	7.866	12.137	7.870	12.129	7.880	12.123	<b>7.876</b>	<b>12.124</b>
95%	7.453	12.522	7.466	12.524	7.47	12.517	<b>7.470</b>	<b>12.530</b>
99%	6.679	13.298	6.667	13.312	6.652	13.321	<b>6.675</b>	<b>13.325</b>

Συγκρίνοντας τις θεωρητικές και τις υπολογισμένες τιμές του Πίνακα 9.13 παρατηρούμε πως το σφάλμα του διαστήματος  $\delta$  ( $\delta=x_U-x_L$ ), δεν ξεπερνά το 0.1% στις περισσότερες περιπτώσεις. Η δε βελτίωση της ακρίβειας από  $m=60 \times 10^3$  προς  $m=120 \times 10^3$  δεν είναι θεαματική ειδικά για  $\gamma=99\%$ . Επιπλέον όπως είναι αναμενόμενο για μικρότερες τιμές του  $\gamma$  έχουμε καλύτερη ακρίβεια καθώς υπάρχει μεγαλύτερη συγκέντρωση του δείγματος Monte-Carlo για να γίνει η κατάλληλη παρεμβολή.

### 9.7.3 Εύρεση διαστημάτων εμπιστοσύνης του πληθυσμού

Στις περισσότερες πρακτικές περιπτώσεις η διασπορά (τα στατιστικά χαρακτηριστικά εν γένει) του πληθυσμού είναι άγνωστη και το μόνο που γνωρίζουμε είναι τα στατιστικά χαρακτηριστικά ενός πεπερασμένου δείγματος. Όσο το δείγμα είναι μικρότερο μάλιστα, τόσο η διαφορά μεταξύ δειγματικών τιμών της διασποράς και του πληθυσμού είναι μεγαλύτερη.

Έτσι τα όρια εμπιστοσύνης που αναφέρονται στον πληθυσμό διαφέρουν από τα όρια εμπιστοσύνης του δείγματος που υπολογίσαμε στην προηγούμενη παράγραφο, τόσο περισσότερο όσο μικραίνει το μέγεθος του δείγματος. Αποδεικνύεται πως αν γνωρίζουμε τα δειγματικά όρια  $x_U$  και  $x_L$ , την τιμή της συνάρτησης ποσοστημορίου  $x(u_d)$  καθώς και την κλίση της συνάρτησης ποσοστημορίου ως προς  $x_U$  και  $x_L$ , τότε μπορούμε να υπολογίσουμε τα όρια εμπιστοσύνης του πληθυσμού  $X_L$  και  $X_U$  ως:

$$X_L = x(u_d) + [x(u_d) - x_U(u_d)] \frac{dx(u)}{dx_U(u)} \quad 9.173$$

$$X_U = x(u_d) + [x(u_d) - x_L(u_d)] \frac{dx(u)}{dx_L(u)} \quad 9.174$$

Όλα τα μεγέθη των παραπάνω εξισώσεων είναι γνωστά και προσδιορίζονται με την μεθοδολογία της προηγούμενης ενότητας, οι κλίσεις όμως είναι άγνωστες. Κατά καιρούς διάφορες προσεγγίσεις έχουν προταθεί, ο «Υδρογνώμων» χρησιμοποιεί μία προσέγγιση που έχει προταθεί από τους Koutsoyiannis και Efstradiadis (2004) και βασίζεται στην διαδοχική παραγωγή:

Σύμφωνα με αυτές τις προσεγγίσεις, οι κλίσεις θα δίνονται από τις σχέσεις:

$$\frac{dx(u)}{dx_L(u)} = \frac{\sum_{i=1}^r \frac{\partial x(u)}{\partial p_i} \left( \frac{\partial x(u)}{\partial p_i} + \frac{\partial x_U(u)}{\partial p_i} \right) \sigma_{p_i}^2}{\sum_{i=1}^r \frac{\partial x_L(u)}{\partial p_i} \left( \frac{\partial x(u)}{\partial p_i} + \frac{\partial x_U(u)}{\partial p_i} \right) \sigma_{p_i}^2} \quad 9.175$$

$$\frac{dx(u)}{dx_U(u)} = \frac{\sum_{i=1}^r \frac{\partial x(u)}{\partial p_i} \left( \frac{\partial x(u)}{\partial p_i} + \frac{\partial x_L(u)}{\partial p_i} \right) \sigma_{p_i}^2}{\sum_{i=1}^r \frac{\partial x_U(u)}{\partial p_i} \left( \frac{\partial x(u)}{\partial p_i} + \frac{\partial x_L(u)}{\partial p_i} \right) \sigma_{p_i}^2} \quad 9.176$$

όπου  $p_i$  οι 1 έως  $r$  παράμετροι που χαρακτηρίζουν τα στατιστικά χαρακτηριστικά του πειραματικού δείγματος (π.χ. μέση τιμή και τυπική απόκλιση ή και ασυμμετρία ή L-ροπές). Ο αριθμός  $r$  των παραμέτρων θα είναι συνήθως 2 ή 3 ανάλογα με τον αριθμό ανεξάρτητων παραμέτρων της κατανομής που έχει υιοθετηθεί.  $\sigma_{p_i}^2$  θα είναι η διασπορά της παραμέτρου  $p_i$  η οποία προκύπτει από τις  $m$  τιμές των παραμέτρων που προκύπτουν από τα αντίστοιχα πειράματα. Οι μερικοί παράγωγοι ως προς τις παραμέτρους  $p_i$  προκύπτουν αριθμητικά με την εξής μεθοδολογία:

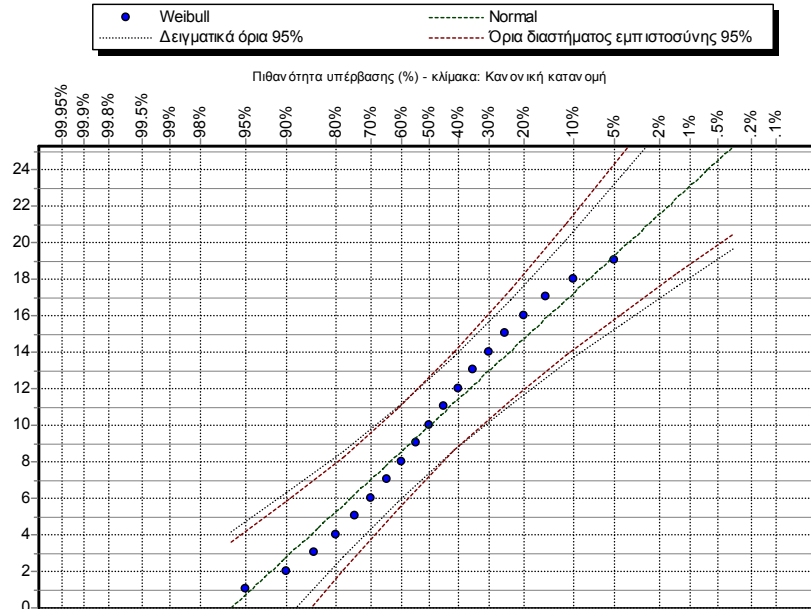
Αρχικά προσδιορίζονται οι δειγματικές τιμές κατά τα γνωστά. Σε αυτήν την διαδικασία προσδιορίζουμε και τις διασπορές των παραμέτρων  $\sigma_{p_i}^2$  από τα  $m$  πειράματα. Η διαδικασία επαναλαμβάνεται  $r$  φορές (επανυπολογίζονται δηλαδή δειγματικά όρια) για μικρές μεταβολές των στατιστικών παραμέτρων του αρχικού (πραγματικού) δείγματος:

$$(1+\delta)p_i \quad 9.177$$

όπου για το  $\delta$  υιοθετούμε την τιμή  $\boxed{\delta=0.15}$ . Στην συνέχεια η μερική παράγωγος κάποιας μεταβλητής (έστω της  $x_L(u)$ ) ως προς την παράμετρο  $p_i$  θα είναι:

$$\frac{\partial x_L(u)}{\partial p_i} = \frac{x_L(u) - x_L(u)|_{(1+\delta)p_i}}{\delta \cdot p_i} \quad 9.178$$

Με παρόμοιο τρόπο υπολογίζονται και οι υπόλοιπες μερικές παράγωγοι δηλαδή των  $x_L$  και  $x(u)$ , τέλος υπολογίζονται οι ζητούμενες κλίσεις και τα όρια εμπιστοσύνης του πληθυσμού  $X_L$  και  $X_U$ .



Σχήμα 9.7: Διαστήματα εμπιστοσύνης 95% κανονικής κατανομής του δείγματος και του πληθυσμού, για το δείγμα:  $x_i=i, i=1..19, n=19$

Στον Πίνακα 9.14 παρουσιάζεται μία σύγκριση των ορίων εμπιστοσύνης που προκύπτουν με την παραπάνω μεθοδολογία με μία θεωρητική λύση για τα όρια εμπιστοσύνης της μέσης τιμής όταν είναι άγνωστη η διασπορά του πληθυσμού. Στην περίπτωση αυτή τα όρια εμπιστοσύνης είναι:

$$P\left(\bar{X} - \frac{t_{(1+\gamma)/2, n-1} \sigma_x}{\sqrt{n}} < \mu_x < \bar{X} + \frac{t_{(1+\gamma)/2, n-1} \sigma_x}{\sqrt{n}}\right) = \gamma \quad 9.179$$

όπου  $t_{(1+\gamma)/2, n-1}$  είναι το ποσοστημόριο της κατανομής  $t$  ή Student για  $n-1$  βαθμούς ελευθερίας.

Πίνακας 9.14: Σύγκριση ορίων εμπιστοσύνης του πληθυσμού που προκύπτουν με την μέθοδο της προσομοίωσης και της θεωρητικής τιμής για την μέση τιμή (κανονική κατανομή,  $u_d=50\%$ , δείγμα  $x_i=i, i=1..19, n=19$ ).

$\gamma$	Προσομοίωση $m=30 \times 10^3$		Προσομοίωση $m=60 \times 10^3$		Προσομοίωση $m=120 \times 10^3$		Θεωρητική τιμή	
	$x_L$	$x_U$	$x_L$	$x_U$	$x_L$	$x_U$	$x_L$	$x_U$
90%	7.732	12.177	7.768	12.151	7.771	12.183	<b>7.767</b>	<b>12.233</b>
95%	7.321	12.611	7.297	12.632	7.342	12.649	<b>7.289</b>	<b>12.711</b>
99%	6.341	13.745	6.376	13.728	6.336	13.617	<b>6.282</b>	<b>13.718</b>

Ισχύουν τα ίδια συμπεράσματα με την σύγκριση των δειγματικών τιμών των ορίων εμπιστοσύνης. Τα σφάλματα των διαστημάτων είναι καλύτερα από 1% εν γένει και με έναν αριθμό πειραμάτων  $m=60 \times 10^3$  (ο οποίος ουσιαστικά πολλαπλασιάζεται επί 3 ή 4 για να προκύψουν οι μερικές παράγωγοι) έχουμε ικανοποιητικά αποτελέσματα ως προς την ακρίβεια.

Στο Σχήμα 9.7 παρουσιάζεται μία αναπαράσταση των ορίων εμπιστοσύνης (95%) του δείγματος και του πληθυσμού, για το δείγμα που κάνουμε τις συγκρίσεις. Το δείγμα είναι σχετικά μεγάλο ( $n=19$ ) οπότε οι διαφορές μεταξύ των ορίων δεν είναι μεγάλες.





## 10 Ανάλυση εξαιρετικών υδρολογικών γεγονότων – κατάρτιση όμβριων καμπυλών

---

### 10.1 Εισαγωγή

Ο διεθνής όρος για τις «όμβριες καμπύλες» είναι «καμπύλες έντασης – διάρκειας – περιόδου επαναφοράς βροχόπτωσης (Intensity – Duration – Frequency curves / IDF curves) και είναι ένα εργαλείο της στατιστικής υδρολογίας το οποίο μπορεί να δώσει τα απαραίτητα μεγέθη για τον σχεδιασμό αντιπλημμυρικών έργων. Για περισσότερες λεπτομέρειες πάνω στην χρήση των όμβριων καμπυλών μπορεί κάποιος να ανατρέξει σε βιβλία Υδρολογίας όπως αυτό των Κουτσογιάννη και Ξανθόπουλου (1999). Οι όμβριες καμπύλες είναι αναλυτικές ή γραφικές εκφράσεις της μέγιστης έντασης βροχόπτωσης  $i$  συναρτήσει της διάρκειας του επεισοδίου βροχόπτωσης  $d$  και της περιόδου επαναφοράς  $T$ .

Η κατάρτιση των όμβριων καμπυλών από τις μετρήσεις των βροχοπτώσεων (από τους βροχογράφους κυρίως) μπορεί να χωριστεί σε δύο στάδια που αποτελούν ουσιαστικά δύο υποσυστήματα του «Υδρογνώμων». Στο πρώτο στάδιο γίνεται κατάρτιση χρονοσειρών ετησίων μεγίστων μέσων εντάσεων  $i_{jt}$  για μία γκάμα διαρκειών συνάθροισης όπως  $d_j = \{5\text{min}, 10\text{min}, 30\text{min}, 1\text{h}, 2\text{h}, 6\text{h}, 12\text{h}, 24\text{h}, 48\text{h}\}$  από τα δεδομένα της χρονοσειράς μέτρησης της βροχόπτωσης. Στο δεύτερο στάδιο χρησιμοποιούνται τα δείγματα  $i_{jt}$  που παρήχθησαν για να καταρτιστούν οι τελικές όμβριες καμπύλες. Το υποσύστημα κατάρτισης των όμβριων καμπυλών, ειδικά, φέρει την ονομασία «Όμβρος».

Στα πλαίσια του «Υδρογνώμων» προτείνεται η εφαρμογή συνεπούς μεθοδολογίας η οποία δίνει ενιαία μαθηματική έκφραση για τις όμβριες καμπύλες  $i(d, T)$ , επιπλέον προτείνεται η χρήση στατιστικής κατανομής με χαρακτηριστικά AT-2 και μέθοδο L-ροπών για την εξαγωγή των παραμέτρων (σε αντίθεση με την κλασική θεώρηση που εφαρμόζει AT-1 – Gumbel και απλή μέθοδο ροπών). Ωστόσο παρέχεται και η συμβατική μεθοδολογία για την εκτίμηση καμπύλης σε συγκεκριμένη περίοδο επαναφοράς.

### 10.2 Εκτίμηση μηνιαίων – ετήσιων μεγίστων μέσων εντάσεων

Η παραγωγή των σειρών μηνιαίων ή ετησίων μεγίστων μέσων εντάσεων (εφεξής θα παραλείπεται η λέξη «μέσων»), είναι μία απλή υπολογιστική διαδικασία συνάθροισης και εύρεσης κάποιας μέγιστης τιμής. Η παραγωγή ξεκινάει από κάποια χρονοσειρά μέτρησης της βροχόπτωσης η οποία έχει μικρό χρονικό βήμα (π.χ. πεντάλεπτο, δεκάλεπτο, ωριαίο και κάποιες φορές και ημερήσιο). Στον «Υδρογνώμων» δεν περιλαμβάνεται το πεντάλεπτο χρονικό βήμα οπότε η παραγωγή θα γίνεται από τις δεκάλεπτες, ωριαίες ή ημερήσιες χρονοσειρές βροχόπτωσης.

Η παραγωγή των σειρών μεγίστων μηνιαίου χρονικού βήματος ειδικά, εξυπηρετεί περισσότερο την κατάρτιση πινάκων εποχικών στατιστικών και λιγότερο την κατάρτιση όμβριων καμπυλών. Για την κατάρτιση των όμβριων καμπυλών απαιτούνται σειρές ετησίων μεγίστων (αναφερόμενοι μάλιστα στο υδρολογικό έτος). Ωστόσο αν έχουμε στην διάθεσή μηνιαίες χρονοσειρές μεγίστων, μπορούμε εύκολα να μεταβούμε σε ετήσια χρονοσειρά μεγίστων με τις διαδικασίες του «Υδρογνώμων» για την παραγωγή χρονοσειρών μεγαλύτερου χρονικού βήματος – συνάθροιση (βλ. Ενότητα 4.3). Αν πρόκειται οι χρονοσειρές μεγίστων να χρησιμοποιηθούν αποκλειστικά για την κατάρτιση ομβρίων

καμπυλών, τότε προτείνεται η απευθείας παραγωγή χρονοσειράς μεγίστων ετησίου βήματος για να αποφύγουμε μεταξύ άλλων, προβλήματα τιμών περιθωρίου.

Η υπολογιστική διαδικασία είναι απλή, έστω πως  $d$  είναι η διακριτότητα της χρονοσειράς μέτρησης (ουσιαστικά το χρονικό βήμα: δεκάλεπτο, ωριαίο ή ημερήσιο). Τότε μπορούμε να παράγουμε μία χρονοσειρά μεγίστων που θα αναφέρεται σε διάρκεια βροχόπτωσης  $d$ , όταν το  $d$  είναι ακέραιο πολλαπλάσιο του  $\delta$  ( $n \times \delta$ ). Αυτό γίνεται δημιουργώντας ένα «παράθυρο» διαστάσεων  $d = n \times \delta$  το οποίο μετακινείται στην χρονοσειρά μετρήσεων και συναθροίζει τις τιμές της βροχόπτωσης ώστε να προκύψει το συνολικό ύψος  $h$  που αντιστοιχεί σε αυτό το «παράθυρο». Στην συνέχεια επιλέγουμε την μέγιστη τιμή  $h$  για κάθε μήνα ή έτος, ανάλογα με το χρονικό βήμα της σειράς που θέλουμε να παράγουμε. Η αφετηρία του παραθύρου μπορεί και μετακινείται έως την τελευταία τιμή της βροχόπτωσης του μήνα ή του έτους που εξετάζεται, άρα μπορεί να συναθροίσει τιμές και από το επόμενο χρονικό βήμα. Οπότε η μέγιστη τιμή ίσως να περιλαμβάνει και τιμές που περιέχονται σε δύο αλληπάλληλα βήματα.

Η σειρά  $h$ , μεγίστων υψών βροχής για συγκεκριμένη διάρκεια  $d$ , μετατρέπεται εύκολα σε χρονοσειρά μεγίστων εντάσεων βροχής  $i$ , αρκεί να διαιρέσουμε το  $h$  με το  $d$ . Ωστόσο παρέχεται η δυνατότητα διατήρησης των μεγίστων υψών ως τελικό αποτέλεσμα.

Η διαδικασία όπως φάνηκε είναι ιδιαίτερα απλή αρκεί να δοθεί προσοχή σε δύο ζητήματα: Πρέπει να γίνεται επισήμανση στον χρήστη μέσω κατάλληλης σημαιοθέτησης των τιμών (βλ. ενότητα 2.4) στην περίπτωση ύπαρξης ελλειπουσών τιμών ή τιμών περιθωρίου. Έτσι θα σημαιοθετείται μία εγγραφή με την σημαία «έλλειψης» (MISSING) όταν στην μέγιστη τιμή που προέκυψε περιλαμβάνονται ελλείψεις. Υπάρχει ωστόσο η επιλεκτική δυνατότητα απόρριψης των «παραθύρων» που περιλαμβάνουν ελλείψεις. Επιπλέον, θα σημαιοθετείται με την σημαία «τιμών περιθωρίου» όταν η μέγιστη τιμή προκύπτει από κάποιο παράθυρο που συνορεύει με περίοδο ελλείψεων. Η σημαιοθέτηση αυτή επιστρά την προσοχή στον μελετητή, καθώς υπάρχει περίπτωση τα μέγιστα που προέκυψαν, αν συνορεύουν με ελλείψεις, να ήταν μεγαλύτερα, οπότε πρέπει να αποφασίσει για απόρριψη ή μη του έτους. Επιπλέον αυτή είναι μία απαραίτητη πληροφορία αν θέλουμε να μεταβούμε από σειρά μηνιαίων μεγίστων σε σειρά ετησίων μεγίστων. Τέλος παρέχεται προαιρετικά στήλη (χρονοσειρά) με το ποσοστό των ελλείψεων σε μονάδες τοις εκατό (%).

## 10.3 Έλεγχοι - αναγωγές

### 10.3.1 Έλεγχος συνέπειας χρονοσειρών

Οι χρονοσειρές που θα χρησιμοποιηθούν για την κατάρτιση των όμβριων καμπυλών θα πρέπει να ελέγχονται αν είναι συνεπείς. Συγκεκριμένα:

- Όταν αυξάνεται η διάρκεια  $d$ , οι τιμές των μεγίστων υψών βροχόπτωσης για κάθε έτος χωριστά είναι αύξουσες:  $h(d_1) > h(d_2)$  για  $d_1 > d_2$
- Αντίθετα, όσον αφορά τις εντάσεις βροχόπτωσης, είναι φθίνουσες για αντίστοιχη αύξηση της διάρκειας:  $i(d_1) < i(d_2)$  για  $d_1 > d_2$ .

Στους παραπάνω ελέγχους εισάγεται ένα περιθώριο σφάλματος  $\epsilon = 0.02$ , το οποίο χρησιμοποιείται αυτούσιο στον έλεγχο των εντάσεων, ενώ προσαυξάνεται πολλαπλασιαζόμενο με την διάρκεια για τον έλεγχο των υψών. Ο έλεγχος μπορεί να καταλήγει σε αυτόματη διόρθωση των ασυνεπών τιμών εξισώνοντας δύο διαδοχικές τιμές εντάσεων ή υψών (έτσι ώστε να μην παραβιάζεται καμία από τις δύο συνθήκες).

Ο έλεγχος μπορεί να απενεργοποιηθεί, χρήσιμη δυνατότητα όταν εξετάζονται συνθετικές χρονοσειρές όταν δεν λαμβάνεται ειδική μέριμνα για την συνέπεια των τιμών (π.χ. δείγμα που έχει προέλθει από την ενοποίηση πολλών σταθμών).

Ένας ακόμα έλεγχος είναι ο ελάχιστος αριθμός χρονοσειρών ο οποίος τίθεται σε 2 (δύο) όταν θέλουμε να προσδιορίσουμε τα στοιχεία του παρανομαστή  $b(d)$ . Όταν όμως θέτουμε ρητά τα στοιχεία του παρονομαστή ( $\eta$  και  $\theta$ ), τότε ο ελάχιστος αριθμός περιορίζεται σε μία χρονοσειρά. Αυτή η δυνατότητα είναι χρήσιμη καθώς μπορούμε για παράδειγμα να βγάλουμε τα στοιχεία  $\eta$  και  $\theta$  από ένα δείγμα βροχογράφου, στην συνέχεια όμως να καταρτίσουμε τις τελικές όμβριες από μία χρονοσειρά βροχόμετρου χρησιμοποιώντας τα  $\eta$  και  $\theta$  που προέκυψαν από τον βροχογράφο.

### 10.3.2 Αναγωγές χρονικής ευκρίνειας – σειρών υπεράνω κατωφλίου

#### Επίδραση χρονικής ευκρίνειας

Έχει παρατηρηθεί πως όταν οι χρονοσειρές μέγιστης έντασης που προκύπτουν από την συνάθροιση διακριτών τιμών υποεκτιμούν την πραγματική μέγιστη ένταση για την δεδομένη διάρκεια συνάθροισης (Κουτσογιάννης, 1997). Αυτή η υποεκτίμηση είναι τόσο μεγαλύτερη όσο η διάρκεια βροχόπτωσης πλησιάζει στην χρονική ευκρίνεια. Αν η χρονική ευκρίνεια είναι  $\delta$ , μπορούμε να εκφράσουμε τον λόγο  $d/\delta$  και στην συνέχεια να εξάγουμε συντελεστές άρσης του σφάλματος διακριτοποίησης. Οι συγκεκριμένοι συντελεστές έχουν προταθεί από τους Linsley et al. (1975, σ. 375), παίρνουν δε την μέγιστη τιμή 1.13 όταν η διάρκεια είναι ίση με την ευκρίνεια.

Πίνακας 10.1: Τιμές του συντελεστή άρσης του σφάλματος διακριτοποίησης

Λόγος διάρκειας προς ευκρίνεια ( $d/\delta$ )	Συντελεστής άρσης του σφάλματος διακριτοποίησης
1	1.13
2	1.04
3-4	1.03
5-8	1.02
9-24	1.01

Μπορούμε να κάνουμε τις εξής παρατηρήσεις στην χρήση των παραπάνω συντελεστών διόρθωσης:

- Όταν οι χρονοσειρές των μεγίστων βροχοπτώσεων έχουν προκύψει από συνάθροιση χρονοσειρών μικρής χρονικής ευκρίνειας (π.χ. από βροχογράφους με διακριτότητα 5 ή 10 λεπτά), η εφαρμογή των συντελεστών δίνει όμβριες καμπύλες με καλύτερη εκτίμηση (βάρος) στις μικρές διάρκειες
- Όταν οι όμβριες καταρτίζονται από δείγμα βροχόμετρου (π.χ. ημερήσιου), η χρήση του συντελεστή 1.13 (ή και του 1.04 για τις χρονοσειρές διάρκειας  $d=48h$ ) είναι απαραίτητη καθώς η μη-χρήση υποεκτιμά την ένταση σε όλη την γκάμα χρονικών διαρκειών
- Όταν χρησιμοποιούνται ταυτοχρόνως δείγματα από βροχόμετρα και βροχογράφους θα πρέπει να εφαρμόζονται οι κατάλληλοι συντελεστές με διαφορετικά κριτήρια σε κάθε δείγμα ώστε να αίρεται η ανομοιογένεια (τα βροχόμετρα δηλαδή πρέπει να συμμετέχουν με μεγαλύτερο βάρος).

Ο χρήστης του συστήματος μπορεί να καθορίσει ανά χρονοσειρά την χρονική ευκρίνεια (ειδικά όταν χρησιμοποιούνται ταυτοχρόνως χρονοσειρές από βροχόμετρα και βροχογράφους) ή αν δεν είναι γνωστή να μην την εισάγει. Στην συνέχεια το υποσύστημα κατάρτισης των όμβριων καμπυλών μπορεί να λάβει ή όχι υπόψη την επίδραση (προσοχή: η προρυθμισμένη λειτουργία είναι να μην λαμβάνεται υπόψη).

### Σχέση σειρών ετήσιων μεγίστων και σειρών υπεράνω κατωφλίου

Για μικρές περιόδους επαναφοράς ( $T < 10$ ), πιο αντιπροσωπευτικές όμβριες καμπύλες είναι αυτές που προκύπτουν από χρονοσειρές υπεράνω κατωφλίου (Κουτσογιάννης, 1997, σ. 287). Οι χρονοσειρές που έχουμε στην διάθεσή μας όμως συνήθως, είναι χρονοσειρές ετήσιων μεγίστων. Είναι δυνατή δε η αναγωγή των αποτελεσμάτων που προέρχονται από επεξεργασία σειρών ετήσιων μεγίστων με τέτοιο τρόπο ώστε να αντιστοιχούν προσεγγιστικά σε αυτά που προέρχονται από επεξεργασία υπεράνω κατωφλίου. Αν υποθέσουμε πως  $T$  είναι η περίοδος επαναφοράς που αντιστοιχεί στην χρονοσειρά ετήσιων μεγίστων και  $T'$  η περίοδος επαναφοράς που αντιστοιχεί στην χρονοσειρά μεγίστων υπεράνω κατωφλίου, τότε η παρακάτω σχέση συνδέει τα  $T$  και  $T'$ :

$$T = \frac{1}{1 - e^{-\frac{1}{T'}}} \Leftrightarrow T' = \frac{1}{-\ln\left(1 - \frac{1}{T}\right)} \quad 10.1$$

Η παραπάνω σχέση μπορεί να προσεγγιστεί με ακρίβεια δεύτερου δεκαδικού ψηφία από την απλούστερη:

$$T = T' + 0.5 \quad 10.2$$

Η αναγωγή αυτή (είτε με την ακριβή σχέση είτε με την παραπάνω προσέγγιση) έχει νόημα για  $2 < T < 10$ . Είναι δυνατόν δε να ανάγουμε τις εξισώσεις των ομβριών καμπυλών αντικαθιστώντας το  $T$  με την έκφραση που περιέχει το  $T'$ . Η αναγωγή αυτή των εξισώσεων δίνει απλούστερες εκφράσεις, ειδικά στην περίπτωση των ΑΤ κατανομών (βλ. ενότητα 10.4)

## 10.4 Στατιστικές κατανομές

Οι στατιστικές κατανομές που χρησιμοποιούνται από το σύστημα εξαγωγής όμβριων καμπυλών είναι ένα υποσύνολο των κατανομών που περιγράφονται στην ενότητα 9.4 του Κεφαλαίου 9. Έτσι θα χρησιμοποιούνται κατανομές της ομάδας κατανομών γάμα, οι κατανομές τύπου ΑΤ μεγίστων καθώς και η κατανομή Pareto. Οι παράμετροι των κατανομών θα εξάγονται είτε με την μέθοδο των ροπών είτε με την μέθοδο των L-ροπών στις κατανομές όταν αυτό είναι εφικτό (βλ. ενότητα 9.2.3). Η μέθοδος των L-ροπών θα προτιμάται, εφόσον έχουμε να κάνουμε με κατανομές ακραίων υδρολογικών γεγονότων (βλ. ενότητα 9.2.3).

Στον Πίνακα 10.2 απεικονίζονται όλοι οι τύποι κατανομών που χρησιμοποιούνται για την κατάρτιση των όμβριων καμπυλών, γίνεται επισήμανση για τις μεθοδολογίες υπολογισμού παραμέτρων και απεικονίζεται η μαθηματική σχέση της συνάρτησης ποσοστημορίου συναρτήσει της περιόδου επαναφοράς  $T$ ,  $x(T)$ . Οι παράμετροι των κατανομών αυτών μπορούν να προσδιορίζονται χωριστά για κάθε δείγμα συγκεκριμένης διάρκειας  $d$  (σύμφωνα με την συμβατική μέθοδο) είτε από κάποιο ενοποιημένο δείγμα (σύμφωνα με συνεπείς μεθόδους εξαγωγής όμβριων καμπυλών). Στην τελευταία στήλη του Πίνακα 10.2 απεικονίζεται η ενιαία σχέση όμβριας καμπύλης που θα προκύψει αν τελικά εφαρμοστεί η μέθοδος ενοποίησης διαρκειών. Σε αυτήν την περίπτωση η γενικευμένη σχέση είναι:

$$i = x(T)/b(d) \quad 10.3$$

όπου η  $x(T)$  είναι η συνάρτηση ποσοστημορίου για την κατανομή που τελικά εφαρμόζεται και με παραμέτρους που αναφέρονται στο ενοποιημένο δείγμα, ενώ ο παρονομαστής  $b(d)$  είναι μία σχέση που περιέχει έναν σταθερό όρο  $\theta$  και έναν εκθέτη  $\eta$  και έχει την εξίσωση:

$$b(d) = (d + \theta)^\eta \quad 10.4$$

Πίνακας 10.2: Οι κατανομές που χρησιμοποιούνται για την κατάρτιση των όμβριων καμπυλών, συναρτήσει της περιόδου επαναφοράς  $T$ . Γενικευμένες εξισώσεις όμβριων καμπυλών που προκύπτουν από την διαίρεση με κάποιον όρο  $b(d)$  ( $d$ : διάρκεια βροχοπτώσεως).

$n/a$	Κατανομή	Μέθοδος ροπών	Μέθοδος L-ροπών	Συνάρτηση ποσοστημορίου συναρτήσει της περιόδου επαναφοράς $T$ : $x(T)$	Ενιαία εξίσωση όμβριας καμπύλης $i$ (ένταση) συναρτήσει της περιόδου επαναφοράς $T$ και της διάρκειας $d$
1	Εκθετική	✓	✓	$\psi + \frac{\ln(T)}{\lambda}$	$\frac{\lambda\psi + \ln(T)}{\lambda(d + \theta)^\eta}$
2	Γάμα	✓		$G^{-1}\left(1 - \frac{1}{T}, \kappa, \lambda\right)$	$\frac{G^{-1}\left(1 - \frac{1}{T}, \kappa, \lambda\right)}{(d + \theta)^\eta}$
3	Log-Pearson III	✓		$e^{G^{-1}\left(1 - \frac{1}{T}, \kappa, \lambda\right) + \psi}$	$\frac{e^{G^{-1}\left(1 - \frac{1}{T}, \kappa, \lambda\right) + \psi}}{(d + \theta)^\eta}$
4	ΑΤ-1 μεγίστων	✓	✓	$\lambda\psi - \lambda \ln\left(-\ln\left(1 - \frac{1}{T}\right)\right)$	$\frac{\lambda\psi - \lambda \ln\left(-\ln\left(1 - \frac{1}{T}\right)\right)}{(d + \theta)^\eta}$
5	ΑΤ-2 μεγίστων	✓	✓	$\frac{\lambda}{\kappa} \left(-\ln\left(1 - \frac{1}{T}\right)\right)^{-\kappa}$	$\frac{\lambda \left(-\ln\left(1 - \frac{1}{T}\right)\right)^{-\kappa}}{\kappa(d + \theta)^\eta}$
6	ΓΑΤ μεγίστων	✓	✓	$\lambda\psi + \frac{\lambda}{\kappa} \left[ \left(-\ln\left(1 - \frac{1}{T}\right)\right)^{-\kappa} - 1 \right]$	$\frac{\lambda\psi + \frac{\lambda}{\kappa} \left[ \left(-\ln\left(1 - \frac{1}{T}\right)\right)^{-\kappa} - 1 \right]}{(d + \theta)^\eta}$
7	ΓΑΤ μεγίστων με προ- καθορισμένο $\kappa$	✓	✓		
8	Pareto	✓	✓	$\frac{\lambda}{\kappa} \left[ 1 - \left(\frac{1}{T}\right)^\kappa + \psi\kappa \right]$	$\frac{\lambda \left[ 1 - \left(\frac{1}{T}\right)^\kappa + \psi\kappa \right]}{\kappa(d + \theta)^\eta}$

Η κλασική κατανομή που χρησιμοποιείται για την κατάρτιση των όμβριων καμπυλών είναι η ΑΤ-1 μεγίστων (Gumbel max), ωστόσο σύμφωνα με νεότερες έρευνες η κατανομή των μεγίστων των βροχοπτώσεων προσεγγίζεται καλύτερα από μία τύπου ΑΤ-2 μεγίστων, ως εκ τούτου προτείνεται στον χρήστη (ως προρυθμισμένη κατανομή) η ΓΑΤ (GEV) μεγίστων με καθορισμένη παράμετρο σχήματος  $\kappa$ ,  $\kappa=0.15$  (οπότε η ΓΑΤ μεταπίπτει σε ΑΤ-2 αφού  $\kappa>0$ ). Ο χρήστης στην συνέχεια έχει την ευχέρεια να εξετάσει και τους άλλους τύπους κατανομών, ελέγχοντας παράλληλα την προσαρμογή της θεωρητικής κατανομής στα σημεία της εμπειρικής κατανομής (βλ. Κεφάλαιο 9).

Για τον έλεγχο της προσαρμογής χρησιμοποιείται κάποιο «χαρτί κατανομής» (βλ. ενότητα 9.5) που μετασχηματίζει τον άξονα των πιθανοτήτων (δηλαδή των περιόδων επαναφοράς). Εφόσον η ΓΑΤ

μεγίστων είναι η προτεινόμενη κατανομή, έτσι και το προρυθμισμένο χαρτί κατανομής θα είναι αυτό της ΓΑΤ μεγίστων (με  $\kappa=0.15$ ) ώστε να γραμμικοποιείται η απεικόνιση της θεωρητικής κατανομής.

Ως εμπειρική κατανομή για την απεικόνιση των σημείων των μετρήσεων θα χρησιμοποιείται αυτή του Weibull (βλ. ενότητα 9.3).

Στον Πίνακα 10.3 παρουσιάζονται οι μετασχηματισμένες συναρτήσεις ποσοστημορίου για τις κατανομές ΑΤ, εφόσον χρησιμοποιείται η περίοδος επαναφοράς  $T'$  που αντιστοιχεί σε χρονοσειρά μεγίστων υπεράνω κατωφλίου. Ο μετασχηματισμός γίνεται εισάγοντας την έκφραση της  $T$  συναρτήσει της  $T'$ , Σχέση (10.1). Τελικά οι σχέσεις που προκύπτουν έχουν απλούστερη μορφή και μπορούν να χρησιμοποιούνται εναλλακτικά για μικρές περιόδους επαναφοράς (βλ. παράγραφο 10.3.2). Στο υπολογιστικό σύστημα αρκεί να κάνουμε τον απλό μετασχηματισμό  $T = T' + 0.5$  όταν θέλουμε να εξάγουμε μία όμβρια καμπύλη για συγκεκριμένη περίοδο επαναφοράς.

Πίνακας 10.3: Κατανομές ΑΤ που χρησιμοποιούνται στις όμβριες καμπύλες, για (ανηγμένη) περίοδο επαναφοράς  $T'$  που θα αντιστοιχούσε σε χρονοσειρά μεγίστων υπεράνω κατωφλίου  $T'$ .

$n/n$	Κατανομή	Συνάρτηση ποσοστημορίου συναρτήσει της περιόδου επαναφοράς $T'$ : $x(T')$	Ενιαία εξίσωση όμβριας καμπύλης $i$ (ένταση) συναρτήσει της περιόδου επαναφοράς $T'$ και της διάρκειας $d$
1	ΑΤ-1 μεγίστων	$\lambda(\psi + \ln T')$	$\frac{\lambda(\psi + \ln T')}{(d + \theta)^\eta}$
2	ΑΤ-2 μεγίστων	$\frac{\lambda}{\kappa} T'^{\kappa}$	$\frac{\lambda T'^{\kappa}}{\kappa(d + \theta)^\eta}$
3	ΓΑΤ μεγίστων	$\lambda\psi + \frac{\lambda}{\kappa} (T'^{\kappa} - 1)$	$\frac{\lambda\psi + \frac{\lambda}{\kappa} (T'^{\kappa} - 1)}{(d + \theta)^\eta}$
4	ΓΑΤ μεγίστων με προκαθ. $\kappa$		

## 10.5 Κατάρτιση όμβριων καμπυλών με την συμβατική μέθοδο

Η κατάρτιση ομβρίων καμπυλών βασίζεται πρωτίστως σε συνεπείς μεθόδους εξαγωγής όπως είναι η μέθοδος ενοποίησης διαρκειών που θα εξεταστεί στην ενότητα 10.6. Για λόγους πληρότητας ο «Υδρογνώμων» μπορεί να εκτιμήσει όμβριες καμπύλες για συγκεκριμένη περίοδο επαναφοράς (αυτό που καλούμε συμβατική μέθοδο). Πρέπει να έχουμε υπόψη όμως πως η μέθοδος αυτή έχει τα παρακάτω μειονεκτήματα:

- Δεν μπορεί να εξαχθεί ενιαία μαθηματική σχέση για τις όμβριες καμπύλες που να ισχύει για κάθε  $d$  και κάθε  $T$
- Οι καμπύλες που προκύπτουν δεν είναι συνεπείς, έτσι είναι δυνατόν για παράδειγμα να τέμνονται σε κάποια διάρκεια, πράγμα φυσικά απαράδεκτο.

Τα παραπάνω μειονεκτήματα αίρονται με τις συνεπείς μεθόδους εξαγωγής.

Θεωρούμε λοιπόν πως κάποια όμβρια καμπύλη για συγκεκριμένη περίοδο επαναφοράς περιγράφεται από την εξής σχέση:

$$i = \frac{\omega}{d^n} \quad 10.5$$

Όπως βλέπουμε χρησιμοποιούμε την πλέον απλή μορφή για τον παρονομαστή  $b(d)=d^n$ , ώστε να μπορούμε να προσδιορίσουμε τα  $a$  και  $\eta$  με απλή γραμμική παλινδρόμηση. Η μεθοδολογία αναλύεται στα εξής βήματα:

- Θέτουμε κάποια περίοδο επαναφοράς  $T$  για την οποία θα καταρτιστεί η καμπύλη
- Επιλέγεται χρήση στατιστικής κατανομής
- Έστω, διαθέτουμε  $k$  χρονοσειρές ετησίων μεγίστων και  $k$  τιμές διάρκειας  $d_i$  ( $i=1\dots k$ )
- Κάθε χρονοσειρά συγκεκριμένης διάρκειας  $d_i$  θεωρείται ως ένα αυτόνομο δείγμα. Έτσι για κάθε χρονοσειρά προσδιορίζονται τα στατιστικά χαρακτηριστικά του δείγματος και προσαρμόζονται οι παράμετροι της κατανομής σύμφωνα με τις μεθοδολογίες του Κεφαλαίου 9
- Διαθέτοντας τις παραμέτρους της κατανομής για κάθε χρονοσειρά, κάνουμε στατιστική πρόγνωση για την περίοδο επαναφοράς  $T$  που έχουμε θέσει. Για αυτόν τον σκοπό μπορούμε να χρησιμοποιήσουμε τις σχέσεις του ποσοστημορίου  $x(T)$  του Πίνακα 10.2
- Τελικά έχουμε  $k$  σημεία  $[d_i, x_i(T)]$ . Αν αυτά τα σημεία τα βάλουμε σε ένα διπλό λογαριθμικό χαρτί παρατηρούμε πως διατάσσονται πάνω σε μία ευθεία με πολύ καλή προσέγγιση. Έτσι εφαρμόζοντας απλή γραμμική παλινδρόμηση (βλ. Κεφάλαιο 7 καθώς και ενότητα 3.3) στα σημεία  $x_i = \ln d_i, y_i = \ln x_i(T)$ , μπορούμε να προσαρμόσουμε κάποια ευθεία του τύπου:

$$y = ax + b \quad 10.6$$

Τελικά οι παράμετροι της όμβριας καμπύλης  $i = \frac{\omega}{d^n}$  για το  $T$  που έχουμε θέσει, υπολογίζονται από τις σχέσεις:

$$\omega = e^a \text{ και } \eta = -b \quad 10.7$$

Στον χρήστη παρέχονται τελικά οι τιμές των  $\omega$  και  $\eta$  καθώς και ο συντελεστής προσδιορισμού της προσαρμογής.

## 10.6 Κατάρτιση όμβριων καμπυλών με την μέθοδο ενοποίησης διαρκειών

Η μεθοδολογία αυτή προτείνεται από τον Κουτσογιάννη, περιγράφεται στην Στατιστική υδρολογία (Κουτσογιάννης, 1997, σ. 270). Επιτρέπει δε, την εξαγωγή ενιαίας μαθηματικής σχέσης για τις όμβριες καμπύλες συναρτήσει της διάρκειας  $d$  και της περιόδου επαναφοράς  $T$ . Ως ενιαίες σχέσεις μπορούν να χρησιμοποιηθούν αυτές που εμφανίζονται στην τελευταία στήλη του Πίνακα 10.2. Η μεθοδολογία απαιτεί μία σειρά αρκετών υπολογισμών και για αυτό πλεονεκτεί ο προγραμματισμός της σε λογισμικό σύστημα.

Ο προσδιορισμός των στοιχείων του παρονομαστή  $b(d)=(d+\theta)^n$  ( $d, \theta$  π.χ. σε ώρες) θα γίνεται με διαδικασία εξαντλητικών δοκιμών. Συγκεκριμένα, θα γίνεται υπόθεση τιμών ελέγχου  $\eta$  και  $\theta$  και στην συνέχεια θα υπολογίζεται ένας στατιστικός δείκτης  $h$ . Μετά από πολλές δοκιμές για διαφορετικά  $\eta$  και  $\theta$ , θα υιοθετείται εκείνο το ζεύγος που ελαχιστοποιεί αυτόν τον στατιστικό δείκτη  $h$ . Έτσι η κατάρτιση των όμβριων καμπυλών θα χωρίζεται αδρομερώς στα εξής δύο στάδια:

- Προσδιορισμός των στοιχείων του παρονομαστή  $b(d)$   $\eta$  και  $\theta$
- Προσδιορισμός των παραμέτρων της κατανομής (π.χ.  $\kappa, \lambda$  και  $\psi$ ) για το ενοποιημένο δείγμα.

### 10.6.1 Ενοποίηση διαρκειών – εκτίμηση των $\eta$ και $\theta$

Υποθέτουμε πως είναι γνωστές οι παράμετροι  $\eta$  και  $\theta$ . Θα δείξουμε πως θα υπολογίζεται μία στατιστική παράμετρος  $h$  συναρτήσει των δειγμάτων των βροχοπτώσεων και των παραμέτρων του

παρονομαστή. Οι βέλτιστες τιμές  $\eta$  και  $\theta$  προκύπτουν με την ελαχιστοποίηση αυτής της στατιστικής παραμέτρου  $h$ .

Έστω οι  $k$  χρονοσειρές ετησίων μεγίστων που αντιστοιχούν σε διάρκειες  $d_j, j=1..k$ . Αν  $n_j$  το μέγεθος κάθε δείγματος (χρονοσειράς) τότε ενοποιώντας όλα τα δείγματα, παράγουμε ένα δείγμα μεγέθους:

$$m = \sum_{j=1}^k n_j \quad 10.8$$

Για να γίνει η ενοποίηση, κάθε χρονοσειρά  $i_{jl}$  πολλαπλασιάζεται με  $b(d_j)=(d_j+\theta)^n$ :  $y_{jl}=i_{jl}b(d_j)$ , στην συνέχεια δε λαμβάνεται το σύνολο των τιμών  $y_{jl}$  ως ενοποιημένο δείγμα. Το ενοποιημένο δείγμα στην συνέχεια ταξινομείται σε φθίνουσα σειρά. Ωστόσο για λόγους καλύτερης προσαρμογής της συνάρτησης  $b(d)$  στην περιοχή των υψηλότερων εντάσεων, είναι σκόπιμο να μην χρησιμοποιείται σε αυτό το στάδιο υπολογισμού το σύνολο των δεδομένων αλλά ένα μέρος αυτών. Έτσι πριν την ταξινόμηση του συνόλου του δείγματος γίνεται μία ταξινόμηση ανά χρονοσειρά και λαμβάνεται ένα ποσοστό (π.χ. 1/3) από τις μέγιστες τιμές των χρονοσειρών για να σχηματιστεί το ενιαίο δείγμα (τελικά δε, το  $m$  είναι μικρότερο και εξαρτάται από αυτό το ποσοστό).

Το προτεινόμενο ποσοστό από τον «Υδρογνώμων» είναι 1/3, ωστόσο παρέχονται εναλλακτικές τιμές στον χρήστη 10%, 20%, 1/3, 50%, 2/3, 100%. Η επιρροή του ποσοστού στα αποτελέσματα δεν είναι μεγάλη, ως εκ τούτου μπορούμε να χρησιμοποιούμε την προρυθμισμένη τιμή. Κάποια αναγωγή αυτού του ποσοστού γίνεται για τα μικρά δείγματα: Αν το επιθυμητό ποσοστό είναι  $\rho$  και το μέγιστο πλήθος των τιμών από τις χρονοσειρές είναι  $n_{\max} = \max(n_j)$ , τότε ως τελικό ποσοστό  $q$  λαμβάνεται:

$$q = \begin{cases} \rho, & \rho n_{\max} > 10 \\ 10/n_{\max}, & n_{\max} > 10 \wedge \rho n_{\max} \leq 10 \\ 1, & n_{\max} \leq 10 \wedge \rho n_{\max} \leq 10 \end{cases} \quad 10.9$$

Αφού ληφθούν οι  $q n_j$  μεγαλύτερες τιμές από κάθε χρονοσειρά και σχηματιστεί το ενιαίο δείγμα που τελικά θα έχει μέγεθος:

$$m = \sum_{j=1}^k q n_j \quad 10.10$$

Αφού γίνει τελικά η ταξινόμηση κατά φθίνουσα σειρά του ενοποιημένου δείγματος, αποδίδονται βαθμοί (ranks)  $r_{jl}$  για κάθε τιμή του δείγματος, δηλαδή η αύξουσα αριθμητική σειρά (π.χ. 1 για την μεγαλύτερη τιμή, 2 για την αμέσως επόμενη κλπ). Στην περίπτωση ταυτόσημων τιμών  $y_{jl}$  χρησιμοποιείται η μέση τιμή των αντίστοιχων βαθμών. Τελικά επανερχόμαστε σε κάθε δείγμα χωριστά και υπολογίζεται η μέση τιμή των βαθμών για κάθε χρονοσειρά:

$$\bar{r}_j = \frac{1}{n_j} \sum_{l=1}^k r_{jl} \quad 10.11$$

όπου στην θέση του  $n_j$  χρησιμοποιούμε ποσοστό του  $q n_j$ . Αν όλα τα επιμέρους δείγματα είχαν την ίδια κατανομή, τότε κάθε  $\bar{r}_j$  θα βρισκόταν πολύ κοντά στην τιμή  $(m+1)/2$ . Για την επίτευξη αυτού του στόχου και τελικά για τον υπολογισμό των βέλτιστων τιμών  $\eta$  και  $\theta$  χρησιμοποιείται η στατιστική παράμετρος ελέγχου Kruskal-Wallis, η οποία εξετάζει τους μέσους βαθμούς από όλα τα επιμέρους δείγματα:

$$h = \frac{12}{m(m+1)} \sum_{j=1}^k n_j \left( \bar{r}_j - \frac{m+1}{2} \right)^2 \quad 10.12$$



Η ελαχιστοποίηση του  $h$  επιτυγχάνεται έπειτα από κάποια εξαντλητικά τεστ. Στον «Υδρογνώμων» χρησιμοποιείται ένα μητρώο  $31 \times 31$  όπου οι τιμές των  $\eta$  και  $\theta$  κυμαίνονται στο διάστημα (0,1). Αφού επιτευχθεί ο βέλτιστος συνδυασμός  $\eta$ ,  $\theta$ , σχηματίζεται ένα νέο μητρώο  $31 \times 31$  με κέντρο αυτόν το συνδυασμό και διαστάσεις  $1/32 \times 1/32$  ώστε να υπολογιστεί μία ακριβέστερη λύση που θα δώσει ακρίβεια τρίτου δεκαδικού ψηφίου στα  $\eta$  και  $\theta$ . Έτσι απαιτούνται συνολικά 1922 δοκιμές για την εξαγωγή των τελικών τιμών.

Στην συνέχεια γίνεται σύγκριση μεταξύ των επιτευχθέντων τιμών  $\eta$  και  $\theta$  που εξάγονται στο παράδειγμα του βιβλίου «Στατιστική Υδρολογία» (Κουτσογιάννης, 1997, σ. 280). Οι χρονοσειρές που χρησιμοποιούνται για την επίλυση παρουσιάζονται στον Πίνακα 10.4. Οι ενδιάμεσες πράξεις δεν παρουσιάζονται, γίνεται μόνο μία σύγκριση των τελικών αποτελεσμάτων.

Πίνακας 10.4: Καταγραμμένες ετήσιες μέγιστες εντάσεις βροχής στο σταθμό Ελληνικό (Πηγή Εθνική Μετεωρολογική Υπηρεσία).

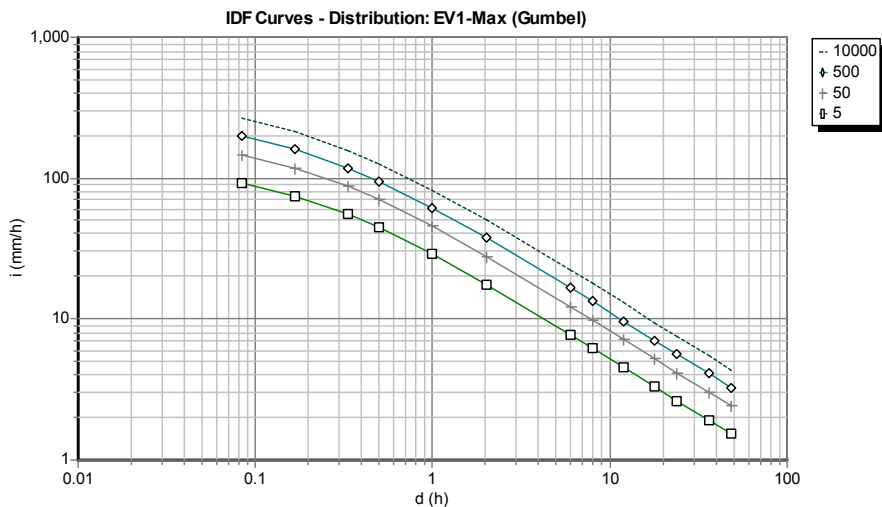
Υδρ. έτος	Μέγιστη ένταση βροχής $i$ (mm/h) για διάρκεια $d=$								
	5min	10min	30min	1h	2h	6h	12h	24h	
1957-58	81.60	66.00	53.20	35.00	26.90	8.97	6.27		
1958-59	58.80	48.00	33.00	21.50	11.20	6.75	3.98		
1959-60	39.60	34.80	20.00	11.60	6.85	2.40	1.82		
1960-61	54.00	34.80	18.40	11.00	6.65	3.62	2.28		
1961-62	120.00	85.80	41.80	24.80	19.30	7.32	3.73		
1962-63	67.20	60.00	23.60	13.80	7.20	3.03	1.94		
1963-64	78.00	48.00	27.80	14.30	8.50	3.52	2.72		
1964-65	96.00	63.00	28.00	15.50	10.65	4.28	2.17		
1965-66	38.40	36.00	23.00	12.00	6.55	2.45	1.69		
1966-67	74.40	63.60	28.40	15.10	7.55	4.88	2.46		
1967-68	36.00	24.60	16.60	10.20	6.65	3.65	2.75	1.58	
1968-69	126.00	69.00	43.20	26.80	15.15	5.93	2.97	1.48	
1969-70	82.80	64.20	41.60	24.50	12.45	5.45	2.75	1.76	
1970-71	42.00	42.00	25.20	17.70	8.95	3.70	3.09	1.55	
1971-72	117.60	85.20	65.20	35.90	19.75	10.02	5.01	2.92	
1972-73	68.40	49.80	39.40	33.50	17.75	6.78	5.27	2.68	
1973-74	60.00	42.00	28.60	15.20	9.85	4.20	3.47	2.00	
1974-75	48.00	48.00	30.60	15.90	8.30	4.27	2.60	1.30	
1975-76	120.00	120.00	74.00	40.90	21.50	7.38	4.54	2.27	
1976-77	115.20	87.60	41.40	23.20	14.90	6.12	3.30	1.65	
1977-78	56.40	46.20	38.60	32.70	20.15	6.73	3.37	1.68	
1978-79	78.00	66.60	47.60	30.00	19.55	11.93	6.12	3.37	
1979-80	67.20	40.80	17.20	13.30	8.60	4.22	2.81	1.62	
1980-81	58.80	56.40	30.40	19.40	11.10	5.58	3.27	1.95	
1981-82	67.20	64.80	40.60	24.70	13.05	4.35	2.28	1.14	
1982-83	141.60	79.80	49.60	36.20	22.90	7.63	4.52	2.29	
1983-84	102.00	69.00	50.40	29.00	17.70	7.03	3.63	1.82	
1984-85	40.80	31.80	16.40	12.90	12.15	9.87	6.00	3.40	
1985-86	74.40	66.00	29.20	15.60	9.40	3.13	1.57	0.83	
1986-87			32.20	29.10	18.55	9.50	7.24	3.85	

Χρησιμοποιούμε το 1/3 των μέγιστων τιμών από κάθε χρονοσειρά. Τελικά σύμφωνα με το παράδειγμα του βιβλίου, οι τιμές  $\eta$  και  $\theta$  είναι:  $\eta=0.796$  και  $\theta=0.189$ . Κάνοντας τον υπολογισμό με τον «Υδρογνώμων» προκύπτουν τιμές  $\eta=0.792$  και  $\theta=0.186$ . Η διαφορά είναι πολύ μικρή (τάξη μεγέθους 1 %) και δικαιολογείται από την προσεγγιστική διαδικασία που ακολουθείται.

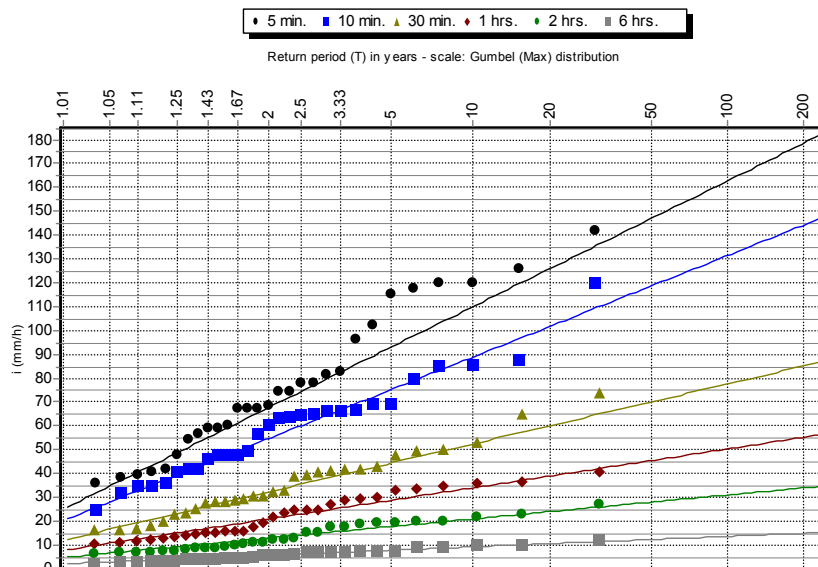
Στον «Υδρογνώμων» παρέχεται η δυνατότητα ώστε κάθε χρονοσειρά χωριστά να καθοριστεί αν θα χρησιμοποιείται στον προσδιορισμό των  $\eta$  και  $\theta$ , ή αν θα χρησιμοποιείται για τον προσδιορισμό των παραμέτρων της κατανομής ή και για τους δύο σκοπούς. Μπορούμε για παράδειγμα να εισάγουμε χρονοσειρές από βροχογράφο για να προκύψουν τα  $\eta$  και  $\theta$  ενώ οι παράμετροι της κατανομής να προκύπτουν είτε από τις χρονοσειρές που αναφέρονται σε μεγάλες χρονικές διάρκειες (π.χ. 12, 24, 48 h) ή και να χρησιμοποιείται μία χρονοσειρά που έχει προκύψει π.χ. από βροχόμετρο. Τέλος υπάρχει δυνατότητα να τεθούν ρητά τιμές  $\eta$  και  $\theta$  από τον χρήστη, π.χ. κάποιες τιμές που χαρακτηρίζουν κάποια περιοχή και έχουν προκύψει από την ενοποίηση δειγμάτων διαφορετικών σταθμών (βλ. Κουτσογιάννης 1997, σ. 284).

## 10.6.2 Προσδιορισμός των παραμέτρων της κατανομής

Το τελικό βήμα για την κατάρτιση των όμβριων καμπυλών είναι ο προσδιορισμός των παραμέτρων της κατανομής που υιοθετείται. Έχοντας πια τις τιμές για τα  $\eta$  και  $\theta$  (οι οποίες είτε έχουν υπολογιστεί είτε έχουν τεθεί ρητά από τον χρήστη), χρησιμοποιούμε το πλήρες ενοποιημένο δείγμα που προκύπτει πολλαπλασιάζοντας κάθε  $i_{ji}$  με  $b(d_j)=(d_j+\theta)^{\eta}$ :  $y_{ji}=i_{ji}b(d_j)$ . Τελικά υπολογίζονται τα στατιστικά χαρακτηριστικά για το ενοποιημένο δείγμα  $y_{ji}$  και τελικά οι παράμετροι της συνάρτησης κατανομής που υιοθετείται. Σύμφωνα με το παράδειγμα (Κουτσογιάννης, 1997, σ. 282), εφαρμόζεται η συνάρτηση κατανομής AT-1 μεγίστων (Gumbel) και προκύπτουν τιμές για τα  $\lambda$  και  $\psi$ :  $\lambda=7.962$  και  $\psi=2.652$ . Οι αντίστοιχες τιμές που προκύπτουν με εφαρμογή του «Υδρογνώμων» είναι:  $\lambda=7.95$  και  $\psi=2.64$ . Οι τιμές αυτές είναι πολύ κοντά και οι διαφορές μπορεί να προέρχονται και από σφάλματα στρωγγυλοποιήσεων. Εφαρμογή της ΓΑΤ μεγίστων με  $\kappa=0.15$  και μέθοδο L-ροπών (προτεινόμενη – προρυθμισμένη κατανομή του συστήματος) θα έδινε αντίστοιχα:  $\lambda=7.04$  και  $\psi=2.88$ .



Σχήμα 10.1: Όμβριες καμπύλες για περιόδους επαναφοράς  $T=5, 10, 50, 500, 5000$  στο Ελληνικό, υπολογισμένες από την συνεπή ενιαία εξίσωση και συνάρτηση κατανομής AT-1 μεγίστων



Σχήμα 10.2: Συναρτήσεις κατανομής Gumbel (max) και Weibull (σημεία εμπειρικής κατανομής) για τον έλεγχο προσαρμογής της κατανομής

### 10.6.3 Τα όρια εμπιστοσύνης των όμβριων καμπυλών

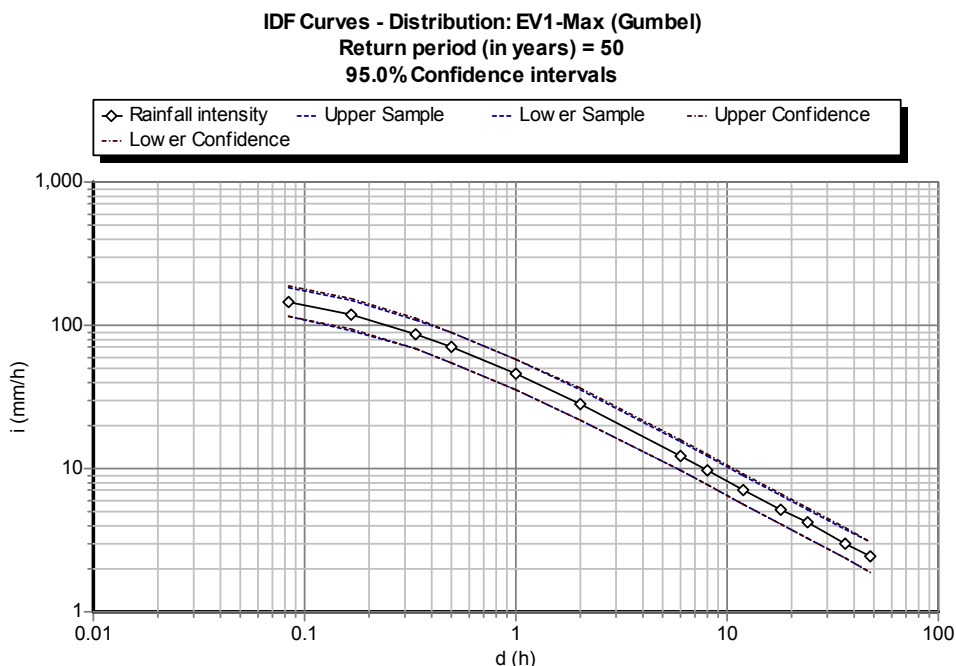
Αφού έχουν προσαρμοστεί οι παράμετροι της κατανομής, είναι εύκολο να εκτιμηθούν τα όρια εμπιστοσύνης κάποιας συγκεκριμένης όμβριας καμπύλης που αντιστοιχεί σε κάποια περίοδο επαναφοράς  $T$  (ή πιθανότητα μη-υπέρβασης  $u=(1-1/T)$ ). Εφαρμόζοντας την μεθοδολογία της ενότητας 9.7 (Monte-Carlo) εκτιμούμε τα όρια εμπιστοσύνης στην συγκεκριμένη θέση  $u=(1-1/T)$  τόσο για το δείγμα όσο και για τον πληθυσμό. Αφού βρεθούν τα όρια εμπιστοσύνης της  $x(T)$  (σε συγκεκριμένη θέση)  $x_L$  και  $x_U$ , τελικά τα όρια εμπιστοσύνης για την όμβρια καμπύλη  $i=x(T)/b(d)$  θα είναι απλά:

$$i_L = \frac{x_L}{(d + \theta)^n}, \quad i_U = \frac{x_U}{(d + \theta)^n} \quad 10.13$$

Ουσιαστικά έχουμε δύο «όμβριες καμπύλες εμπιστοσύνης» που δίνουν τα όρια εμπιστοσύνης της καμπύλης  $i$  για συγκεκριμένο  $T$ . Ωστόσο η δυσκολία του προβλήματος έγκειται στις παραδοχές ως προς το μέγεθος του δείγματος (Κουτσογιάννης, 1987, σ. 289). Αν λάβουμε ως μέγεθος δείγματος αυτό του ενοποιημένου δείγματος  $m$ , θα ήταν μία παραδοχή που θα έδινε πολύ μικρό διάστημα εμπιστοσύνης. Επιπλέον, αυτή η παραδοχή δεν είναι μαθηματικά ορθή καθώς το δείγμα προκύπτει από επιμέρους δείγματα με ισχυρή στατιστική εξάρτηση. Στον «Υδρογνώμων» ως μέγεθος του δείγματος  $n_m$  για την προσομοίωση λαμβάνεται η μέση τιμή από τα πλήρη δείγματα των χρονοσειρών:

$$n_m = \frac{\sum_{j=1}^k n_j}{k} \quad 10.14$$

Παραδοχή που δίνει σχετικά μεγάλα όρια εμπιστοσύνης. Στο Σχήμα παρουσιάζεται μία εφαρμογή βασισμένη στο παράδειγμα που αναπτύχθηκε, για περίοδο επαναφοράς  $T=50$ . Παρατηρούμε πως η διαφορά μεταξύ της δειγματικής τιμής των ορίων και των ορίων εμπιστοσύνης του πληθυσμού είναι μικρή, φαινόμενο που οφείλεται μεταξύ άλλων στην χρήση λογαριθμικών αξόνων που αμβλύνουν τις διαφορές.



Σχήμα 10.3: Διαστήματα εμπιστοσύνης 95% της όμβριας καμπύλης για  $T=50$ . Χρήση κατανομής AT-1 μεγίστων (Gumbel) στο δείγμα του Ελληνικού.



## Αναφορές - βιβλιογραφία

---

- Βέης Γ., Μπιλλήρης Χ., Παπαζήση Κ., Ανώτερη Γεωδαισία, Εθνικό Μετσόβιο Πολυτεχνείο, 1992
- Δημητρίου Ι. Δ., Εφαρμοσμένη Υδραυλική, τεύχος Α' (Θεωρία), Εκδόσεις Γρηγ. Φούντας, 2<sup>η</sup> Έκδοση, 1994
- Δημητρίου Ι. Δ., Εφαρμοσμένη Υδραυλική, τεύχος Β' (Εφαρμογές), Εκδόσεις Γρηγ. Φούντας, 2<sup>η</sup> Έκδοση, 1994
- Ευστρατιάδης, Α., Ι. Ναλμπάντης, και Ν. Μαμάσης, Υδρομετεωρολογικά δεδομένα και επεξεργασίες, *Εκσυγχρονισμός της εποπτείας και διαχείρισης του συστήματος των υδατικών πόρων ύδρευσης της Αθήνας*, Τεύχος 8, 129 σελίδες, Τομέας Υδατικών Πόρων, Υδραυλικών και Θαλάσσιων Έργων - Εθνικό Μετσόβιο Πολυτεχνείο, Αθήνα, Δεκέμβριος 2000
- Ευστρατιάδης, Α., Ι. Ναλμπάντης, και Ε. Ρόζος, Μοντέλο προσομοίωσης του υδρολογικού κύκλου στη λεκάνη Βοιωτικού Κηφισού - Υλίκης, *Εκσυγχρονισμός της εποπτείας και διαχείρισης του συστήματος των υδατικών πόρων ύδρευσης της Αθήνας*, Τεύχος 21, 196 σελίδες, Τομέας Υδατικών Πόρων, Υδραυλικών και Θαλάσσιων Έργων - Εθνικό Μετσόβιο Πολυτεχνείο, Αθήνα, Ιανουάριος 2004
- Ευστρατιάδης, Α., *Διερεύνηση μεθόδων αναζήτησης ολικού βελτίστου σε προβλήματα υδατικών πόρων*, MSc thesis, 139 σελίδες, Τομέας Υδατικών Πόρων, Υδραυλικών και Θαλάσσιων Έργων - Εθνικό Μετσόβιο Πολυτεχνείο, Αθήνα, 2001
- Ευστρατιάδης, Α., κ.α., Θεωρητική τεκμηρίωση μοντέλου προσομοίωσης υδρολογικών-υδρογεωλογικών διεργασιών λεκάνης απορροής, *Ολοκληρωμένη Διαχείριση Υδατικών Συστημάτων σε Σύζευξη με Εξελιγμένο Υπολογιστικό Σύστημα*, Τομέας Υδατικών Πόρων, Υδραυλικών και Θαλάσσιων Έργων – Εθνικό Μετσόβιο Πολυτεχνείο, Αθήνα, 2005
- Χριστοφίδης Α., Καραβοκυρός Γ. και Στάυρακας Ι., Σχεδιασμός βάσης δεδομένων, *Ολοκληρωμένη Διαχείριση Υδατικών Συστημάτων σε Σύζευξη με Εξελιγμένο Υπολογιστικό Σύστημα*, Τομέας Υδατικών Πόρων, Υδραυλικών και Θαλάσσιων Έργων – Εθνικό Μετσόβιο Πολυτεχνείο, Αθήνα, 2005
- Κουτσογιάννης Δ. και Ξανθόπουλος Α., *Τεχνική Υδρολογία*, Εθνικό Μετσόβιο Πολυτεχνείο, 1999
- Κουτσογιάννης Δ., *Στατιστική Υδρολογία*, Εθνικό Μετσόβιο Πολυτεχνείο, 1997
- Κουτσογιάννης Δ., *Σχεδιασμός αστικών δικτύων αποχέτευσης*, Εθνικό Μετσόβιο Πολυτεχνείο, 1999
- Κουτσογιάννης Δ., *Σημειώσεις Προχωρημένης Υδρολογίας - Μέρος Α*, 52 σελίδες, Εθνικό Μετσόβιο Πολυτεχνείο, Αθήνα, 1999
- Μπαλοδήμου – Αγατζά Α.Μ., *Θεωρία σφαλμάτων και Μ.Ε.Τ.*, Ε.Μ.Π., 1991
- Ο.Κ.Χ.Ε, Γεωδαιτικά Συστήματα Αναφοράς – Το Ελληνικό Σύστημα «ΕΓΣΑ 87», 1989
- Παπακόστας, Ν., Σχεδιασμός και υλοποίηση κεντρικής βάσης δεδομένων, *Εκσυγχρονισμός της εποπτείας και διαχείρισης του συστήματος των υδατικών πόρων ύδρευσης της Αθήνας*, Τεύχος 19, 106 σελίδες, Τομέας Υδατικών Πόρων, Υδραυλικών και Θαλάσσιων Έργων - Εθνικό Μετσόβιο Πολυτεχνείο, Αθήνα, Ιανουάριος 2004

- Τσακίρης Γ., *Υδατικοί Πόροι: Ι. Τεχνική Υδρολογία*, Εκδόσεις Συμμετρία, Αθήνα, 1995
- Χριστοφίδης, Α., *Συμπλήρωση ελλειπόν υδρομετεωρολογικών χρονοσειρών σε κατανεμημένες σχεσιακές βάσεις δεδομένων*, ΕΜΠ, διπλωματική εργασία, Αθήνα, Φεβρουάριος 1994
- Χριστοφίδης Α., *Απαιτήσεις συστήματος*, Εσωτερική έκθεση στα πλαίσια της Εθνικής Τράπεζας Υδρολογικής και Μετεωρολογικής Πληροφορίας, ΕΜΠ, ΤΥΠΥΘΕ, 1998
- Χριστοφίδης, Α., και Κοζάνης Σ., *Υδρογνώμονας (έκδοση 1.0) - Λογισμικό διαχείρισης δεδομένων, Εκσυγχρονισμός της εποπτείας και διαχείρισης του συστήματος των υδατικών πόρων ύδρευσης της Αθήνας*, Τεύχος 22, 90 σελίδες, Τομέας Υδατικών Πόρων, Υδραυλικών και Θαλάσσιων Έργων - Εθνικό Μετσόβιο Πολυτεχνείο, 2004
- Barnett, V. and Lewis, T., *Outliers in statistical data*, Chichester: John Wiley and Sons, 1978
- Brutsaert, W., *Evaporation Into the Atmosphere*, Reidel, Dordrecht, 1982
- Chapra S.C., Canale R.P., *Numerical methodes for engineers*, McGraw-Hill, 1988
- Chow, V. T., *Open-channel Hydraulics*, McGraw-Hill, 1959
- Dingman, S. L., *Physical Hydrology*, Prentice-Hall, Englewood Cliffs, New Jersey, 1994
- Duan, Q., S. Sorooshian, and V. Gupta, Effective and efficient global optimization for conceptual rainfall-runoff models, *Water Resources Research*, 28(4), 1015-1031, 1992
- Efstratiadis, A., and D. Koutsoyiannis, An evolutionary annealing-simplex algorithm for global optimisation of water resource systems, *Proceedings of the Fifth International Conference on Hydroinformatics*, Cardiff, UK, 1423-1428, International Water Association, 2002
- Hosking, J. R. M., *L-moments: analysis and estimation of distributions using linear combinations of order statistics*, Journal of the Royal Statistical Society, Series B, 52, 105-124, 1990
- Kirkpatrick, S., C. D. Gelatt, and M. P. Vecchi, Optimization by simulated annealing, *Science*, 220, 671-680, 1983
- Koutsoyiannis, D., and A. Efstratiadis, Climate change certainty versus climate uncertainty and inferences in hydrological studies and water resources management (solicited), *1st General Assembly of the European Geosciences Union, Geophysical Research Abstracts, Vol. 6*, Nice, European Geosciences Union, 2004
- Kozanis, S., A. Christofides, N. Mamassis, A. Efstratiadis, and D. Koutsoyiannis, Hydrognomon - A hydrological data management and processing software tool, *2nd General Assembly of the European Geosciences Union, Geophysical Research Abstracts, Vol. 7*, Vienna, 04644, European Geosciences Union, 2005
- Linsley, R. K., M. A. Kohler, and J. L. Paulhus, *Applied Hydrology*, McGraw-Hill, 1949
- Linsley, R. K., M. A. Kohler, and J. L. Paulhus, *Hydrology for Engineers*, McGraw-Hill, Tokyo, 2<sup>nd</sup> edition, 1975
- Makhlouf, Z. and C. Michel, A two-parameter monthly water balance model for French watersheds, *Journal of Hydrology*, 162, 299-318, 1994
- Matalas, N. C., and B. Jacobs, "A correlation procedure for augmenting hydrologic data", U.S. Geological Survey Professional Paper, 434-E, pp. E1-E7, 1964.
- Miller, Alan R., *BASIC programs for Scientists and Engineers*, Sybex, 1981

- Nelder, J. A., and R. Mead, A simplex method for function minimization, *Computer Journal*, 7(4), 308-313, 1965
- Nash, J. E., and J. V. Sutcliffe, River flow forecasting through conceptual models, I, A discussion of principles, *Journal of Hydrology*, 10(3), 282-290, 1970
- Rozos, E., A. Efstratiadis, I. Nalbantis, and D. Koutsoyiannis, Calibration of a semi-distributed model for conjunctive simulation of surface and groundwater flows, *Hydrological Sciences Journal*, 49(5), 819-842, 2004
- Shaw, E. M., *Hydrology in Practice*, Third Edition, Chapman & Hall, 1994
- Tsakalias, G., and D. Koutsoyiannis, A comprehensive system for the exploration and analysis of hydrological data, *Water Resources Management*, 13, 269-302, 1999
- Xu, C.-Y., and V. P. Singh, A review of monthly water balance models for water resources investigations, *Water Resources Management*, 12, 31-50, 1998