ΥΠΟΥΡΓΕΙΟ ΠΕΡΙΒΑΛΛΟΝΤΟΣ, ΧΩΡΟΤΑΞΙΑΣ & ΔΗΜΟΣΙΩΝ ΕΡΓΩΝ
ΓΕΝΙΚΗ ΓΡΑΜΜΑΤΕΙΑ ΔΗΜΟΣΙΩΝ ΕΡΓΩΝ
Δ/ΝΣΗ ΕΡΓΩΝ ΥΔΡΕΥΣΗΣ & ΑΠΟΧΕΤΕΥΣΗΣ

ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΤΟΜΕΑΣ ΥΔΑΤΙΚΩΝ ΠΟΡΩΝ - ΥΔΡΑΥΛΙΚΩΝ
& ΘΑΛΑΣΣΙΩΝ ΕΡΓΩΝ

MINISTRY OF ENVIRONMENT, REGIONAL PLANNING & PUBLIC WORKS
GENERAL SECRETARIAT OF PUBLIC WORKS
SECRETARIAT OF WATER SUPPLY & SEWAGE

NATIONAL TECHNICAL UNIVERSITY OF ATHENS
DIVISION OF WATER RESOURCES - HYDRAULIC
& MARITIME ENGINEERING

ΕΡΕΥΝΗΤΙΚΟ ΕΡΓΟ

ΕΚΤΙΜΗΣΗ ΚΑΙ ΔΙΑΧΕΙΡΙΣΗ ΤΩΝ ΥΔΑΤΙΚΩΝ ΠΟΡΩΝ ΤΗΣ ΣΤΕΡΕΑΣ ΕΛΛΑΔΑΣ

ΦΑΣΗ Β

ΤΕΥΧΟΣ 12

ΜΟΝΤΕΛΟ ΣΤΟΧΑΣΤΙΚΗΣ ΠΡΟΣΟΜΟΙΩΣΗΣ ΥΔΡΟΛΟΓΙΚΩΝ ΧΡΟΝΟΣΕΙΡΩΝ ΜΕ ΑΠΛΗ ΤΕΧΝΙΚΗ ΕΠΙΜΕΡΙΣΜΟΥ - ΕΓΧΕΙΡΙΔΙΟ ΧΡΗΣΗΣ ΠΡΟΓΡΑΜΜΑΤΟΣ

RESEARCH PROJECT

EVALUATION AND MANAGEMENT OF THE WATER RESOURCES OF STEREA HELLAS

PHASE B

VOLUME 12

A MODEL FOR STOCHASTIC SIMULATION OF HYDROLOGICAL TIME SERIES USING A SIMPLE DISAGGREGATION TECHNIQUE - USER'S MANUAL

ΣΥΝΤΑΞΗ: Δ. ΚΟΥΤΣΟΓΙΑΝΝΗΣ, Α. ΜΑΝΕΤΑΣ
ΕΠΙΣΤΗΜΟΝΙΚΟΣ ΥΠΕΥΘΥΝΟΣ: Θ. ΞΑΝΘΟΠΟΥΛΟΣ
ΚΥΡΙΟΣ ΕΡΕΥΝΗΤΗΣ: Δ. ΚΟΥΤΣΟΓΙΑΝΝΗΣ

ΑΘΗΝΑ - ΣΕΠΤΕΜΒΡΙΟΣ 1995

BY: D. KOUTSOYIANNIS, A. MANETAS
SCIENTIFIC DIRECTOR: TH. XANTHOPOULOS
PRINCIPAL INVESTIGATOR: D. KOUTSOYIANNIS

ATHENS - SEPTEMBER 1995

Περιεχόμενα

1. Εισαγωγή

1.1 Αντικείμενο του τεύχους - Ιστορικό

Το τεύχος αυτό συνοδεύει τα προγράμματα ηλεκτρονικού υπολογιστή για την στοχαστική προσομοίωση υδρολογικών χρονοσειρών. Συγκεκριμένα παρέχει οδηγίες χρήσης και πλήρη θεωρητική και τεχνική τεκμηρίωση των προγραμμάτων. Τα προγράμματα αυτά καταρτίστηκαν στα πλαίσια του ερευνητικού έργου *Εκτίμηση και διαχείριση των υδατικών πόρων της Στερεάς Ελλάδας* (φάση Β). Σημειώνεται ότι η θεωρητική βάση των προγραμμάτων δεν καλύπτεται από τη βιβλιογραφία αλλά σε μεγάλο μέρος οφείλεται σε πρωτότυπη θεωρητική εργασία.

Το εν λόγω ερευνητικό έργο ανατέθηκε και χρηματοδοτήθηκε από τη Διεύθυνση Ύδρευσης και Αποχέτευσης του ΥΠΕΧΩΔΕ (απόφαση Δ6/21609/8-9-1993) σε ερευνητική ομάδα του Τομέα Υδατικών Πόρων, Υδραυλικών και Θαλάσσιων Έργων του ΕΜΠ με επιστημονικό υπεύθυνο τον καθηγητή Θ. Ξανθόπουλο και συντονιστή το Δ. Κουτσογιάννη. Η συγκεκριμένη εργασία την οποία καλύπτει το τεύχος αυτό προδιαγράφεται στο Παράρτημα της απόφασης ανάθεσης (άρθρο 2.2.3, εδάφιο α και 2.3.3 εδάφιο α: Συνέχιση της ανάπτυξης μοντέλου στοχαστικής προσομοίωσης χρονοσειρών επιφανειακής υδρολογίας).

Η μεθοδολογία που ακολουθείται στο πρόγραμμα αυτό βασίζεται κατ' αρχήν σε παρόμοια εργασία που έγινε στην Α' φάση η οποία είχε τίτλο *Προγράμματα παραγωγής συνθετικών σειρών με στοχαστικά μοντέλα.*

Η νέα εργασία αποτελεί εξέλιξη της προηγούμενης τόσο σε θεωρητικό όσο και σε προγραμματιστικό επίπεδο. Θεωρητικά είναι απλούστερη σε σύλληψη και ταυτόχρονα εξίσου αποτελεσματική. Προγραμματιστικά έχει γίνει μετάβαση από το περιβάλλον DOS στο περιβάλλον Windows και έτσι το πρόγραμμα είναι πολύ πιο φιλικό προς το χρήστη.

Όπως είχε αναφερθεί στο σχετικό τεύχος της Α φάσης του ερευνητικού έργου φιλοδοξία της ερευνητικής ομάδας είναι η διαρκής εξέλιξη του έργου τόσο ως προς το θεωρητικό υπόβαθρο όσο και ως προς τις προσφερόμενες δυνατότητες και την επικοινωνία με το χρήστη.

1.2 Διάρθρωση του τεύχους

Το τεύχος αποτελείται από το κύριο μέρος και δύο παραρτήματα. Το κύριο μέρος απευθύνεται στον βασικό χρήστη των προγραμμάτων και παρέχει οδηγίες χρήσης και συνοπτικά στοιχεία της μεθοδολογίας στην οποία στηρίζονται τα προγράμματα. Τα παραρτήματα απευθύνονται στον εξειδικευμένο χρήστη και παρέχουν αναλυτική θεωρητική και τεχνική τεκμηρίωση. Συγκεκριμένα, το Παράρτημα 1 καλύπτει τη μεθοδολογία και τους αλγορίθμους στα οποία στηρίζονται τα προγράμματα προσομοίωσης, ενώ το Παράρτημα 2 δίνει μερικές μαθηματικές αποδείξεις που τεκμηριώνουν τη μεθοδολογία του πρώτου παραρτήματος.

2. Συνοπτικό θεωρητικό υπόβαθρο

2.1 Εισαγωγή

Τα προγράμματα προσομοίωσης υδρολογικών χρονοσειρών δίνουν τη δυνατότητα γέννησης συνθετικών σειρών υδρολογικών μεταβλητών (απορροής, βροχής, εξάτμισης κ.ά.) σε πολλές θέσεις ταυτόχρονα. Οι συνθετικές αυτές σειρές μπορούν να χρησιμοποιηθούν σε προσομοιώσεις συστημάτων υδατικών πόρων για υποβοήθηση της διαχείρισης τους. Πιο συγκεκριμένα οι συνθετικές χρονοσειρές βοηθούν στον προγραμματισμό και σχεδιασμό των απαραίτητων υδραυλικών έργων, στον προγραμματισμό ή τη βελτιστοποίηση της λειτουργίας τους και στο λειτουργικό τους έλεγχο κάτω από εναλλακτικές πολιτικές διαχείρισης.

Προς το παρόν, οι περισσότερες μελέτες προγραμματισμού, σχεδιασμού ή λειτουργίας υδραυλικών έργων, και στη χώρα μας αλλά και διεθνώς, βασίζονται στις ιστορικές υδρολογικές χρονοσειρές και ακολουθούν ανάλογες καθιερωμένες μεθοδολογίες. Ωστόσο, έχει αποδειχτεί (πχ. *Vogel and Stedinger* [1988]) ότι η χρήση της στοχαστικής υδρολογίας οδηγεί σε ακριβέστερες εκτιμήσεις των απαιτούμενων μεγεθών (πχ. χωρητικότητες στο σχεδιασμό ταμιευτήρων) από αυτές στις οποίες οδηγεί η χρήση μεθόδων που στηρίζονται μόνο στα ιστορικά δεδομένα.

Η στοχαστική υδρολογία και η μέθοδος της προσομοίωσης προσφέρει τη δυνατότητα λεπτομερέστερης και ακριβέστερης μελέτης των συστημάτων υδατικών πόρων, με βάση συνθετικές χρονοσειρές οι οποίες αναπαράγουν τη στατιστική δομή και τις στατιστικές παραμέτρους των ιστορικών δεδομένων. Με βάση τις συνθετικές χρονοσειρές μπορούμε να καταρτίσουμε την πιθανοτική περιγραφή της συμπεριφοράς ενός συστήματος υδατικών πόρων και να αποκτήσουμε εικόνα των μεγεθών που ενδιαφέρουν για ακραία επίπεδα πιθανότητας (πχ. 1:100, 1:1 000 κτλ.) πράγμα που δεν μπορεί να γίνει μόνο με τα ιστορικά δείγματα που κατά κανόνα είναι διαθέσιμα για μικρή μόνο χρονική περίοδο.

Προϋπόθεση για τη γέννηση συνθετικών χρονοσειρών είναι η υιοθέτηση ενός πιθανοτικού/στοχαστικού μοντέλου που να περιγράφει την από κοινού συνάρτηση κατανομής των υδρολογικών μεταβλητών που ενδιαφέρουν. Ιδιαίτερα ενδιαφέρει η στοχαστική εξάρτηση των μεταβλητών ως προς το χώρο και το χρόνο. Η χωρική εξάρτηση αντιστοιχεί στην εμφανή συγγένεια της ταυτόχρονης υδρολογικής δίαιτας σε γειτονικές θέσεις ή λεκάνες. Αντίστοιχα, η χρονική εξάρτηση αντιστοιχεί στη διαπιστωμένη εμμονή των υδρολογικών (και γενικότερα των γεωφυσικών) μεγεθών.

Εφόσον υιοθετηθεί ένα συγκεκριμένο στοχαστικό μοντέλο για τις μεταβλητές που ενδιαφέρουν, το επόμενο βήμα είναι η εκτίμηση των παραμέτρων του. Κατά κανόνα ενδιαφέρουν οι στατιστικές παράμετροι που καλύπτονται με το γενικό όρο *στατιστικές ροπές* (μέσες τιμές, διασπορές, συνδιασπορές, τρίτες ροπές κτλ.). Η εκτίμηση των παραμέτρων αυτών γίνεται από τα ιστορικά δείγματα με καθιερωμένες μεθόδους της στατιστικής.

Θα πρέπει να τονιστεί ότι και η επιλογή ενός συγκεκριμένου στοχαστικού μοντέλου και η εκτίμηση των παραμέτρων του βασίζεται πάντα στο διαθέσιμο ιστορικό δείγμα, το οποίο αποτελεί τη μόνη πρωτογενή πηγή πληροφορίας. Η γέννηση συνθετικών χρονοσειρών (που κατά κανόνα έχει μήκος πολλαπλάσιο του μήκους του διαθέσιμου ιστορικού δείγματος) δεν προσθέτει ουσιαστική πληροφορία, ούτε επαυξάνει τη διάρκεια του συγκεκριμένου ιστορικού δείγματος. Για παράδειγμα, στο πρόβλημα της εκτίμησης της πλημμύρας εκατονταετίας σε συγκεκριμένη θέση ποταμού, για την οποία υπάρχει δείγμα πχ. 30 ετών, η χρησιμοποίηση συνθετικών χρονοσειρών δεν εξυπηρετεί σε τίποτε. Στην καλύτερη περίπτωση η εκτίμηση με συνθετικές χρονοσειρές θα είναι ίδια με την άμεση εκτίμηση, την οποία δίνει η συνάρτηση κατανομής που έχει υιοθετηθεί για τη συγκεκριμένη μεταβλητή. Κατά συνέπεια, η χρήση συνθετικών χρονοσειρών αποκτά νόημα όταν εξετάζονται αλληλοεξαρτώμενα μεγέθη που συνδυάζονται σε ένα αρκετά πολύπλοκο σύστημα, των οποίων η συνάρτηση κατανομής δεν μπορεί να είναι εξ αρχής γνωστή ή να προσδιοριστεί αναλυτικά. Κλασικό παράδειγμα είναι η περίπτωση συστήματος ταμιευτήρων (ή ακόμη και ενός μεμονωμένου ταμιευτήρα), όπου ενδιαφέρει η στατιστική κατανομή των απολήψεων, οι οποίες εξαρτώνται με ένα αρκετά πολύπλοκο τρόπο από τις εισροές, τις καταναλώσεις, τους κανόνες λειτουργίας κοκ.

Για πρώτη φορά συνθετικές σειρές χρησιμοποίησε στην υδρολογία ο Hazen το 1914 σε μελέτες αξιοπιστίας υδατικών πόρων [*Grygier and Stedinger*, 1990]. Όμως, η μέθοδος που χρησιμοποίησε ήταν εμπειρική, βασισμένη σε αναγωγή και ενοποίηση ιστορικών δειγμάτων, και δεν έχει ομοιότητες με τις σύγχρονες μεθόδους της στοχαστικής υδρολογίας. Με τη σύγχρονη της μορφή, η μέθοδος εφαρμόστηκε πολύ αργότερα, ξεκινώντας το 1954 με τον Barnes (γέννηση ασυσχέτιστων ετήσιων δεδομένων με κανονική κατανομή σε μία θέση) και συνεχίζοντας το 1962 με τους Maass κ.α. και Thomas & Fiering (γέννηση χρονικά συσχετισμένων δεδομένων με μη κανονικές κατανομές, το 1965 με τον Beard και το 1967 με τον Matalas (γέννηση παράλληλων χρονοσειρών σε διάφορες θέσεις) [*Grygier and Stedinger*, 1990]. Πολύ νωρίτερα, στη δεκαετία του 1940, είχε εισαχθεί από μαθηματικούς και φυσικούς (Ulam, von Neumann, Fermi, Metropolis) η μέθοδος Monte-Carlo για τη μελέτη φαινομένων πυρηνικής φυσικής [Metropolis, 1989, Eckhardt, 1989], η οποία απετέλεσε τη βάση και για την ανάπτυξη της υδρολογικής προσομοίωσης. Σημαντική για την ανάπτυξη και διάδοση της μεθόδου ήταν και η συμβολή του κλασικού βιβλίου των *Box and Jenkins* [1970], το οποίο πραγματεύεται την ανάλυση και σύνθεση των χρονοσειρών, την ταξινόμηση των στοχαστικών μοντέλων και τις εφαρμογές τους στην προσομοίωση και την πρόγνωση. Η έρευνα στο θέμα αυτό εξακολούθησε και τις επόμενες δεκαετίες και ακόμη και σήμερα συνεχίζεται.

2.2 Το γενικό σχήμα προσομοίωσης των προγραμμάτων

Από τα διάφορα σχήματα προσομοίωσης που έχουν μελετηθεί, θεωρήθηκε ως πλεονεκτικότερο και υιοθετήθηκε ως βάση για τα προγράμματα υδρολογικής προσομοίωσης, ένα σχήμα πολλών μεταβλητών (θέσεων) και δύο διαδοχικών επιπέδων ή φάσεων: Στο πρώτο επίπεδο (γνωστό ως υψηλότερο επίπεδο) γίνεται γέννηση των παράλληλων χρονοσειρών των διάφορων θέσεων σε μια αραιή χρονική κλίμακα. Η χρονική ισοδιάσταση αυτής της κλίμακας λέγεται *περίοδος*. Στο δεύτερο επίπεδο (γνωστό ως χαμηλότερο επίπεδο) γίνεται γέννηση των χρονοσειρών σε πυκνότερη χρονική κλίμακα και στη συνέχεια εφαρμόζεται μια διαδικασία διόρθωσης που εξασφαλίζει τη συμβατότητα των σειρών χαμηλότερου επιπέδου με αυτές του υψηλότερου επιπέδου (ικανοποίηση της αθροιστικής ιδιότητας). Η χρονική ισοδιάσταση αυτής της πυκνότερης κλίμακας λέγεται *υποπερίοδος*. Ως χρονική κλίμακα του πρώτου επιπέδου έχει επιλεγεί η ετήσια για διάφορους λόγους, ο κυριότερος από τους οποίους είναι ότι σε αυτή την κλίμακα εξαφανίζονται οι ετήσιες περιοδικότητες και έτσι οι χρονοσειρές εμφανίζουν στάσιμο (stationary) χαρακτήρα. Για το δεύτερο επίπεδο δεν υπάρχει καθορισμένη χρονική κλίμακα και, ανάλογα με το πρόβλημα που μας ενδιαφέρει, μπορούμε να επιλέξουμε κατά περίπτωση εποχική, μηνιαία, δεκαπενθήμερη ή άλλη κλίμακα. Θεωρητικά το σχήμα αυτό θα μπορούσε να επεκταθεί και με επόμενες φάσεις πύκνωσης σε ακόμη λεπτομερέστερες χρονικές κλίμακες, αλλά, ωστόσο, τεχνικά αυτό δεν υποστηρίζεται από την τρέχουσα έκδοση του προγράμματος.

Το παραπάνω σχήμα προσομοίωσης είναι σαφώς πλεονεκτικότερο από το πιο διαδεδομένο σχήμα που γεννά τις μεταβλητές σειριακά, τη μια μετά την άλλη, σε μια και μοναδική φάση που έχει μια μοναδική χρονική κλίμακα αναφοράς (ίδια με την πυκνότερη από τις δύο κλίμακες του παραπάνω σχήματος, δηλαδή τη χρονική κλίμακα του χαμηλότερου επιπέδου). Το βασικό πλεονέκτημα του σχήματος δύο επιπέδων είναι ότι παρέχει τη δυνατότητα διατήρησης των σημαντικών στατιστικών χαρακτηριστικών των χρονοσειρών σε πολλαπλή χρονική κλίμακα. Για παράδειγμα, στην γέννηση μηνιαίων χρονοσειρών, το σχήμα δύο επιπέδων επιτρέπει αφενός τη διατήρηση των στατιστικών χαρακτηριστικών των μηνιαίων απορροών (οι οποίες αποτελούν τις μεταβλητές χαμηλότερου επιπέδου) και αφετέρου τη διατήρηση των στατιστικών χαρακτηριστικών των ετήσιων απορροών (μεταβλητές υψηλότερου επιπέδου), αφού οι δεύτερες γεννώνται ανεξάρτητα και πριν από τις πρώτες με βάση διαφορετικό μοντέλο. Αντίθετα, το σειριακό σχήμα μπορεί να διατηρεί μόνο τα χαρακτηριστικά των μηνιαίων απορροών και να υπολογίζει τις ετήσιες απορροές ως αθροίσματα των μηνιαίων. Σε αυτή όμως την περίπτωση, λόγω συσσώρευσης σφαλμάτων και λόγω αναντιστοιχιών των μοντέλων με τη φυσική πραγματικότητα, δεν διατηρούνται από το σειριακό σχήμα επακριβώς, παρά μόνο σε πρώτη προσέγγιση, τα στατιστικά χαρακτηριστικά των ετήσιων χρονοσειρών. Βεβαίως, το σχήμα δύο επιπέδων που υιοθετήθηκε έχει και μειονεκτήματα, το κυριότερο από τα οποία είναι η πολυπλοκότητα του.

Αξίζει να σημειωθεί ότι το σχήμα που στηρίζεται σε πολλαπλά επίπεδα με χρήση μοντέλων επιμερισμού έχει υιοθετηθεί και σε δύο από τα πλέον δεδομένα διεθνώς πακέτα προγραμμάτων γέννησης συνθετικών σειρών, το LAST [*Lane and Frevert*, 1990] και το SPIGOT [*Grygier and Stedinger*, 1990]. Το δικό μας πρόγραμμα παρουσιάζει κάποιες ομοιότητες με αυτά, αλλά διαφέρει σημαντικά κυρίως στο νέο μοντέλο επιμερισμού που χρησιμοποιεί, το οποίο αναλύεται στο Παράρτημα 1.

2.3 Στατιστικές παράμετροι που διατηρούνται

Ανεξάρτητα από τη χρονική κλίμακα και το επίπεδο προσομοίωσης (υψηλότερο ή χαμηλότερο), το σύνολο των στατιστικών παραμέτρων των μεταβλητών που χρησιμοποιούνται από το πρόγραμμα και τελικά αναπαράγονται (διατηρούνται) στις συνθετικές χρονοσειρές αποτελείται από τις ακόλουθες ομάδες:

Παράμετροι των περιθώριων συναρτήσεων κατανομής κάθε μεταβλητής

(1) Μέσες τιμές των μεταβλητών.

(2) Διασπορές των μεταβλητών.

(3) Συντελεστές ασυμμετρίας των μεταβλητών (και, κατά συνέπεια, τρίτες ροπές).

Παράμετροι των από κοινού συναρτήσεων κατανομής των μεταβλητών

(4) Συντελεστές αυτοσυσχέτισης με μοναδιαίο χρονικό βήμα μεταξύ μεταβλητών της ίδιας θέσης.

(5) Συντελεστές ετεροσυσχέτισης με μηδενικό χρονικό βήμα μεταξύ μεταβλητών διαφορετικής θέσης.

Πρόκειται για το ελάχιστο σύνολο ουσιωδών στατιστικών παραμέτρων που κατά κανόνα ενδιαφέρουν [*Matalas and Wallis*, 1976, σσ. 60, 63]. Η επιλογή της ελαχιστοποίησης του αριθμού των παραμέτρων έγινε με σκοπό να είναι το πρόγραμμα κατά το δυνατόν εύχρηστο και γρήγορο, οι σχετικοί αλγόριθμοι κατά το δυνατόν απλούστεροι και η απαιτούμενη προεργασία εκτίμησης παραμέτρων σχετικά απλή και άμεση, χωρίς παράλληλα να χάνεται ουσιώδης και χρήσιμη στατιστική πληροφορία (βλ. και επόμενο υποκεφάλαιο). Με τον τρόπο αυτό επιτεύχθηκε η αποκαλούμενη φειδωλή χρήση παραμέτρων (parsimony of parameters) και μάλιστα σε βαθμό που ξεπερνάει άλλα γνωστά μοντέλα της βιβλιογραφίας (βλ. Παράρτημα 1). Βέβαια, είναι θεωρητικά δυνατό, στο βαθμό που κρίνονται ουσιώδεις, να εισαχθούν και να αναπαραχθούν και άλλες ομάδες παραμέτρων, πράγμα που υποστηρίζεται θεωρητικά από τη μεθοδολογία που αναπτύχθηκε αλλά προς το παρόν δεν υποστηρίζεται από το συγκεκριμένο πρόγραμμα.

Σημειώνεται ότι το παραπάνω σύνολο παραμέτρων αποτελεί την *είσοδο* στα συγκεκριμένα μοντέλα. Οι παράμετροι αυτές μπορούν να υπολογιστούν από το ίδιο το πρόγραμμα, αν είναι διαθέσιμα τα κατάλληλα ιστορικά δεδομένα, ή να εισαχθούν απ' ευθείας από το χρήστη, αν είναι εξ αρχής γνωστές οι τιμές τους.

Διευκρινίζεται ότι το παραπάνω σύνολο παραμέτρων αφορά κατά περίπτωση και στις μεταβλητές υψηλότερου επιπέδου και στις μεταβλητές χαμηλότερου επιπέδου. Συγκεκριμένα, για την ακολουθία των μεταβλητών υψηλότερου επιπέδου (ετήσιων μεταβλητών), η οποία θεωρείται στάσιμη, χρειάζεται ένα σύνολο τέτοιων παραμέτρων. Αντίστοιχα για την ακολουθία των μεταβλητών χαμηλότερου χρειάζονται τόσα σύνολα παραμέτρων όσες είναι και οι μεταβλητές μιας περιόδου (πχ. για μηνιαίες μεταβλητές χρειάζονται 12 σύνολα παραμέτρων).

2.4 Μοντέλο γέννησης ετήσιων μεταβλητών (μεταβλητών υψηλότερου επιπέδου)

Έχει αποδειχτεί ότι οι ετήσιες χρονοσειρές εμφανίζουν το φαινόμενο της εμμονής (persistence), δηλαδή την τάση ομαδοποίησης των ετών υψηλής υδροφορίας και αντίστοιχα των περιόδων χαμηλής υδροφορίας. Το φαινόμενο αυτό μπορεί εν μέρει να περιγραφεί και να μοντελοποιηθεί μαθηματικά με ένα μη μηδενικό συντελεστή αυτοσυσχέτισης των ετήσιων τιμών της υπόψη μεταβλητής. Ωστόσο, η πληρέστερη μαθηματική αναπαράσταση της μακροπρόθεσμης εμμονής απαιτεί την εισαγωγή της λεγόμενης παραμέτρου Hurst (από το όνομα του ερευνητή που την εισήγαγε και τη μελέτησε, το 1950) της οποίας ο ορισμός και ο τρόπος εκτίμησης είναι αρκετά πολύπλοκος και ξεφεύγει από τους στόχους αυτού του κειμένου. Μοντέλα προσομοίωσης που μπορούν να αναπαράγουν την παράμετρο αυτή αναπτύχθηκαν τη δεκαετία του 1970 από τους Mandelbrot, Wallis, O'Connell, Mejia κ.ά. Τα μοντέλα αυτά είναι αρκετά πολύπλοκα και ο ενδιαφερόμενος αναγνώστης παραπέμπεται σε εγχειρίδια στοχαστικής υδρολογίας, πχ. *Bras and Rodriguez-Iturbe* [1985].

Πολλοί ερευνητές έχουν αξιολογήσει συγκριτικά τα πολύπλοκα μοντέλα που αναπαριστούν την μακροπρόθεσμη εμμονή των υδρολογικών χρονοσειρών σε σχέση με απλούστερα μοντέλα που αναπαριστούν μόνο συντελεστές αυτοσυσχέτισης των χρονοσειρών. Το γενικό συμπέρασμα των ερευνών ήταν ότι τα απλούστερα μοντέλα δίνουν ικανοποιητικά αποτελέσματα σε πρακτικά προβλήματα υδρολογικής προσομοίωσης ενώ η χρήση των πολυπλοκότερων μοντέλων δεν κρίνεται γενικά απαραίτητη. Εξ άλλου, η διαφορά στα αποτελέσματα των δύο τύπων μοντέλων κρίνεται ως αμελητέα αν συγκριθεί με την επίδραση της αβεβαιότητας στην εκτίμηση των παραμέτρων και στους δύο τύπους μοντέλων. Για τους λόγους αυτούς στα κυριότερα διεθνώς διαδεδομένα επιχειρησιακά πακέτα προγραμμάτων έχουν υιοθετηθεί τα απλούστερα μοντέλα τύπου AR(0), AR(1) ή AR(2) (πχ. στο SPIGOT έχουν υιοθετηθεί τα AR(0) και AR(1) [*Grygier and Stedinger*, 1990] και στο LAST τα AR(1) και AR(2) [*Lane and Frevert*, 1990]).

Στο δικό μας προγράμματα έχει υιοθετηθεί το μοντέλο AR(1) που διατηρεί ακριβώς το σύνολο στατιστικών παραμέτρων που περιγράφηκε στο προηγούμενο υποκεφάλαιο. Είναι επίσης απλό να υλοποιηθεί και το μοντέλο τύπου AR(0): αρκεί να μηδενιστούν οι συντελεστές αυτοσυσχέτισης των μεταβλητών (ομάδα παραμέτρων 4 του προηγούμενου υποκεφαλαίου).

2.5 Μοντέλο γέννησης μεταβλητών χαμηλότερου επιπέδου

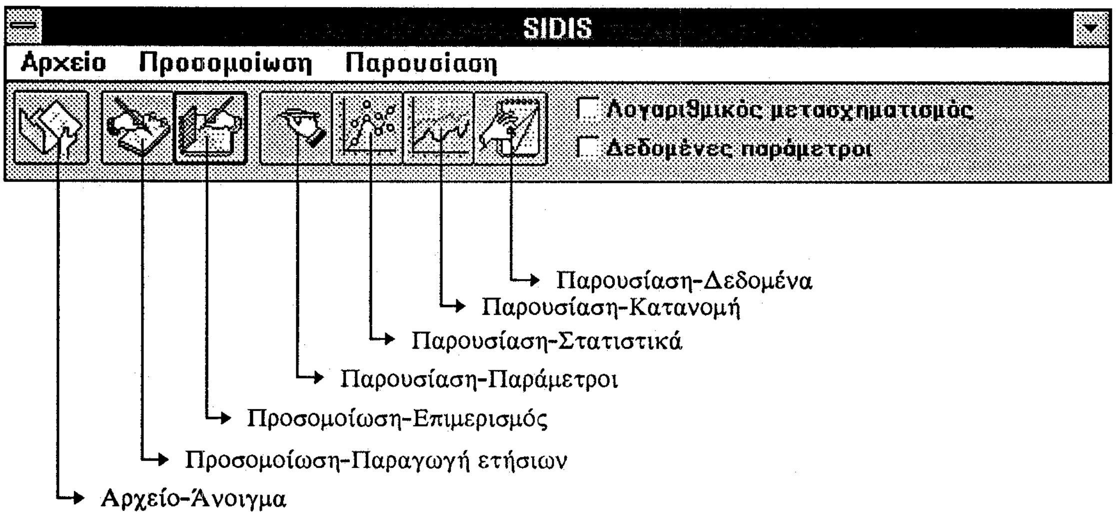
Όπως προαναφέρθηκε, η γέννηση των μεταβλητών χαμηλότερου επιπέδου, πχ. μηνιαίων, γίνεται σε δεύτερη φάση και σε τρόπο ώστε το ετήσιο άθροισμα των μεταβλητών να είναι ίσο με την γνωστή τιμή της ετήσιας μεταβλητής. Η τελευταία είναι γνωστή δεδομένου ότι έχει προηγηθεί η εφαρμογή του μοντέλου γέννησης των ετήσιων μεταβλητών. Κατά συνέπεια αυτό που χρειάζεται εδώ είναι ένα μοντέλο επιμερισμού, δηλαδή ένα μοντέλο που να επιμερίζει ένα άθροισμα στις συνιστώσες του. Ένα τέτοιο μοντέλο είναι το *απλό μοντέλο επιμερισμού*, το οποίο υιοθετήθηκε στα προγράμματα προσομοίωσης. Ας σημειωθεί ότι το μοντέλο αυτό δεν είναι μοντέλο της βιβλιογραφίας, αλλά αναπτύχθηκε εξ ολοκλήρου από την ερευνητική ομάδα στα πλαίσια της παρούσας φάσης του ερευνητικού έργου. Το μοντέλο αυτό διατηρεί, εκτός από την αθροιστική ιδιότητα σε σχέση με τη μεταβλητή υψηλότερου επιπέδου, και το σύνολο στατιστικών παραμέτρων που περιγράφηκε στο υποκεφάλαιο 2.3. Τονίζεται ότι στο απλό μοντέλο επιμερισμού δεν χρειάζεται εισαγωγή άλλων δευτερευουσών παραμέτρων, όπως συμβαίνει με άλλα μοντέλα επιμερισμού της βιβλιογραφίας. Εξ άλλου, με το ακολουθούμενο σχήμα προσομοίωσης δεν είναι απαραίτητη λεπτομερέστερη περιγραφή της στατιστικής δομής των μεταβλητών χαμηλότερου επιπέδου, που ενδεχομένως θα χρειάζονταν αν ακολουθούσαμε ένα αυστηρά σειριακό σχήμα προσομοίωσης, γιατί στην τελευταία περίπτωση θα ήταν πιθανή η συσσώρευση σφαλμάτων στη χρονική κλίμακα υψηλότερου επιπέδου (ετήσια).

Οι μεταβλητές χαμηλότερου επιπέδου (π.χ. μηνιαίες) αρχικά γεννιούνται από ένα μοντέλο PAR(1) ανεξάρτητα από τις ήδη γνωστές μεταβλητές υψηλότερου επιπέδου (π.χ. ετήσιες). Αυτό έχει ως αποτέλεσμα το άθροισμα των μεταβλητών χαμηλότερου επιπέδου για τη χρονική διάρκεια μιας περιόδου να μην ισούται με την αντίστοιχη μεταβλητή υψηλότερου επιπέδου. Την προκύπτουσα διαφορά καλούνται να μειώσουν δύο διαφορετικές τεχνικές: (α) η επανάληψη και (β) η διόρθωση. Έτσι, το πρόγραμμα, αντί να κάνει μια μόνο γέννηση τιμών των μεταβλητών χαμηλότερου επιπέδου μιας περιόδου, γεννά περισσότερες σειρές και τελικά κρατά αυτή που δίνει το μικρότερο σφάλμα σε σχέση με τις ήδη γνωστές τιμές των μεταβλητών υψηλότερου επιπέδου. Σε αυτή τη σειρά επεμβαίνει στη συνέχεια η διαδικασία διόρθωσης, η οποία μηδενίζει το παραπάνω σφάλμα, τροποποιώντας κατάλληλα τις τιμές των μεταβλητών χαμηλότερου επιπέδου. Η τροποποίηση γίνεται με τέτοιο τρόπο ώστε να μην εισάγεται μεροληψία στις στατιστικές παραμέτρους που ενδιαφέρουν. Ο χρήστης καθορίζει το μέγιστο αριθμό επαναλήψεων και το ανώτατο επιτρεπτό σφάλμα. Το πρόγραμμα σταματάει τις επαναλήψεις είτε όταν το σφάλμα γίνει μικρότερο από το επιτρεπτό είτε όταν φτάσει το μέγιστο αριθμό επαναλήψεων. Επίσης, ο χρήστης καθορίζει ποια διορθωτική διαδικασία θα ακολουθήσει το πρόγραμμα, επιλέγοντας μία από τις τρεις διαθέσιμες επιλογές.

Αναλυτική περιγραφή του δυναμικού μοντέλου επιμερισμού, ως προς τη θεωρητική του βάση και τους αλγορίθμους του, δίνεται στο Παράρτημα 1.

3. Οδηγίες χρήσης του προγράμματος

3.1 Κύριο παράθυρο εφαρμογής



Όλες οι βασικές επιλογές για τη λειτουργία του προγράμματος καθορίζονται απ' αυτό το παράθυρο. Μπορεί ο χρήστης να γεννήσει μεταβλητές υψηλότερου επιπέδου, να κάνει επιμερισμό, να διαχειριστεί τα ιστορικά δεδομένα και να σχεδιάσει διαγράμματα που περιγράφουν και συγκρίνουν τις ιστορικές με τις συνθετικές χρονοσειρές. Οι εντολές που εκτελούνται από τα μενού ή τα αντίστοιχα κουμπιά αυτού του παραθύρου είναι:

➢ `Αρχείο-Άνοιγμα` Με την εντολή αυτή εμφανίζεται ένα πλαίσιο διαλόγου απ' όπου ο χρήστης επιλέγει το αρχείο ιστορικών δεδομένων, από το οποίο θα υπολογιστούν οι στατιστικές παράμετροι για την εκτέλεση των προσομοιώσεων. Τα αρχεία χωρίζονται σε δύο κατηγορίες. Τα αρχεία μεταβλητών χαμηλότερου επιπέδου με κατάληξη `.pdt` και τα αρχεία μεταβλητών υψηλότερου επιπέδου με κατάληξη `.adt`.

➢ `Αρχείο-Έξοδος` Με την εντολή αυτή ο χρήστης εγκαταλείπει το πρόγραμμα.

➢ `Προσομοίωση-Παραγωγή ετήσιων` Με την εντολή αυτή γίνεται γέννηση μεταβλητών υψηλότερου επιπέδου. Για να λειτουργήσει απαιτείται πρώτα η εισαγωγή ενός αρχείου με ιστορικά δεδομένα υψηλότερου επιπέδου. Αν δεν έχει γίνει αυτή η ενέργεια το πρόγραμμα ειδοποιεί το χρήστη ότι πρέπει να την κάνει. Αν έχει γίνει η επιλογή `Δεδομένες παράμετροι` τότε δεν απαιτείται αρχείο ιστορικών δεδομένων. Πρέπει βέβαια ο χρήστης να εισάγει τις απαραίτητες παραμέτρους στην ειδική φόρμα (βλ. επεξήγηση πιο κάτω καθώς και στο υποκεφάλαιο 3.2).

➢ `Προσομοίωση-Επιμερισμός` Με την εντολή αυτή γίνεται επιμερισμός των μεταβλητών υψηλότερου επιπέδου σε μεταβλητές χαμηλότερου. Για να εκτελεστεί η εντολή αυτή απαιτούνται δύο αρχεία. Ένα με τα ιστορικά δεδομένα των μεταβλητών χαμηλότερου επιπέδου και ένα με τα συνθετικά δεδομένα των μεταβλητών υψηλότερου επιπέδου που θα επιμεριστούν. Όπως και στην προηγούμενη περίπτωση με την επιλογή `Δεδομένες παράμετροι` δεν απαιτείται η ύπαρξη αρχείου ιστορικών δεδομένων για τις μεταβλητές χαμηλότερου επιπέδου. Το αρχείο των μεταβλητών υψηλότερου επιπέδου απαιτείται πάντοτε.

➢ `Παρουσίαση-Παράμετροι` Οι απαραίτητες παράμετροι για να λειτουργήσουν τα μοντέλα γέννησης και επιμερισμού μπορούν ή να υπολογιστούν από το πρόγραμμα, εφόσον υπάρχει αρχείο ιστορικών δεδομένων, ή να εισαχθούν απ' ευθείας από το χρήστη αν δεν είναι διαθέσιμο το αρχείο με τα ιστορικά δεδομένα. Στην τελευταία περίπτωση, χρησιμοποιώντας αυτή την εντολή, εμφανίζεται ένα παράθυρο στο οποίο ο χρήστης μπορεί να εισαγάγει τις παραμέτρους των μοντέλων που χρησιμοποιεί το πρόγραμμα (βλ. υποκεφάλαιο 3.2).

➢ `Παρουσίαση-Στατιστικά` Με την εντολή αυτή εμφανίζεται ένα παράθυρο στο οποίο ο χρήστης μπορεί υπολογίσει τα στατιστικά χαρακτηριστικά των ιστορικών και συνθετικών χρονοσειρών και να κάνει γραφικές παραστάσεις.

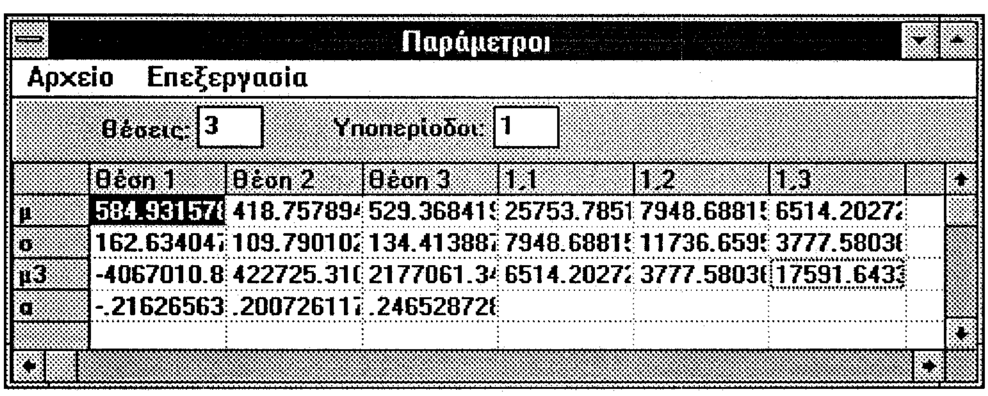
➢ `Παρουσίαση-Κατανομή` Με την εντολή αυτή εμφανίζεται ένα παράθυρο στο οποίο είναι δυνατή η σχεδίαση των εμπειρικών συναρτήσεων κατανομής των ιστορικών και συνθετικών χρονοσειρών σε "χαρτί" κανονικής κατανομής.

➢ `Παρουσίαση-Δεδομένα` Με την εντολή αυτή ανοίγει το παράθυρο διαχείρισης και εισαγωγής των ιστορικών δεδομένων του προγράμματος.

➢ Η επιλογή *Λογαριθμικός μετασχηματισμός* υποδηλώνει στο πρόγραμμα ότι πρέπει να λειτουργήσει με τους λογαρίθμους των μεταβλητών, παρά με τις πραγματικές τιμές τους.

➢ Η επιλογή *Δεδομένες παράμετροι* πρέπει να ενεργοποιηθεί στην περίπτωση που δεν υπάρχει αρχείο με ιστορικά δεδομένα.

3.2 Παράθυρο παραμέτρων



Όπως προαναφέρθηκε, το πρόγραμμα δίνει τη δυνατότητα στο χρήστη να εισαγάγει μόνος του τις στατιστικές παραμέτρους των μοντέλων. Η διαδικασία αυτή γίνεται στο παράθυρο *Παράμετροι*. Οι εντολές που εκτελούνται σ' αυτό το επίπεδο είναι:

➢ `Αρχείο-Νέο` Με την εντολή αυτή το πρόγραμμα γράφει στο φύλλο εργασίας κάποια στοιχεία που βοηθούν το χρήστη στην εισαγωγή των παραμέτρων του μοντέλου. Για να εκτελεστεί η εντολή πρέπει πρώτα να συμπληρωθούν τα πεδία *Θέσεις* και *Υποπερίοδοι* με τις επιθυμητές τιμές.

➢ Αρχείο-Αποθήκευση Με την εντολή αυτή αποθηκεύονται σε αρχείο οι παράμετροι που έχουν εισαχθεί στο φύλλο εργασίας ώστε να μπορούν να χρησιμοποιηθούν και σε άλλες προσομοιώσεις.

➢ Αρχείο-Εισαγωγή Με την εντολή αυτή εισάγεται ένα ήδη αποθηκευμένο αρχείο παραμέτρων.

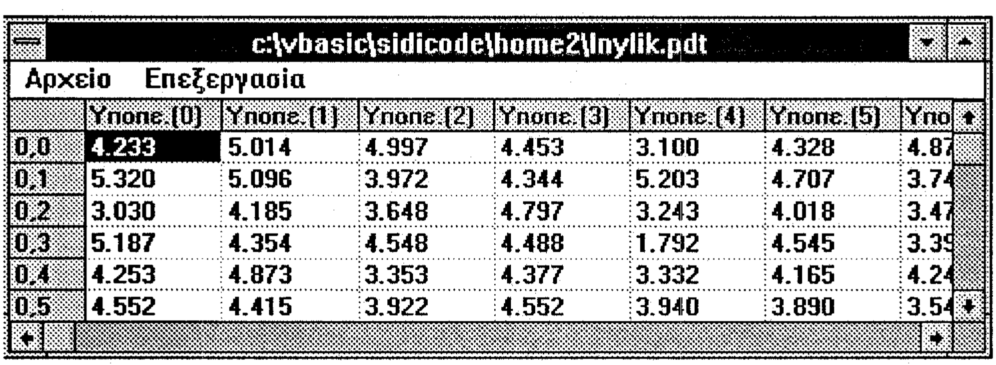
➢ Αρχείο-Έξοδος Με την εντολή αυτή κλείνει το παράθυρο παραμέτρων.

➢ Επεξεργασία-Αντιγραφή Με την εντολή αυτή αντιγράφονται στην προσωρινή μνήμη του υπολογιστή τα πεδία του φύλλου εργασίας που έχει επιλέξει ο χρήστης.

➢ Επεξεργασία-Κόλληση Με την εντολή αυτή τοποθετούνται στο φύλλο εργασίας τα περιεχόμενα της προσωρινής μνήμης.

Το παράθυρο *Παράμετροι* μπορεί να εμφανιστεί μετά από εντολή του χρήστη ή να το εμφανίσει το ίδιο το πρόγραμμα. Η περίπτωση που εμφανίζεται από το πρόγραμμα αφορά στη γέννηση μεταβλητών χωρίς να υπάρχει αρχείο ιστορικών δεδομένων. Συγκεκριμένα, αν ο χρήστης έχει διαλέξει στο κύριο παράθυρο (SIDIS) την επιλογή *Παράμετροι από χρήστη* και στη συνέχεια έχει εκτελέσει την εντολή *Προσομοίωση-Παραγωγή ετήσιων* τότε το πρόγραμμα, επειδή δεν υπάρχει αρχείο ιστορικών δεδομένων για να υπολογίσει αυτόματα τις απαραίτητες παραμέτρους, τις ζητάει από το χρήστη εμφανίζοντας το σχετικό παράθυρο. Μόλις τελειώσει η συμπλήρωση των παραμέτρων πρέπει να πατηθεί το δεξί πλήκτρο του ποντικιού για να ξεκινήσει η γέννηση δεδομένων. Η εμφάνιση του παραθύρου αυτού μετά από εντολή του χρήστη και αφού έχει προηγηθεί επιμερισμός ή γέννηση μεταβλητών, με δεδομένο αρχείο ιστορικών δεδομένων, του δίνει τη δυνατότητα να δει και να τροποποιήσει τις παραμέτρους που υπολόγισε το πρόγραμμα.

3.3 Παράθυρο διαχείρισης ιστορικών δεδομένων



	Υποπε.[0]	Υποπε.[1]	Υποπε.[2]	Υποπε.[3]	Υποπε.[4]	Υποπε.[5]	Υπο
0,0	4.233	5.014	4.997	4.453	3.100	4.328	4.87
0,1	5.320	5.096	3.972	4.344	5.203	4.707	3.74
0,2	3.030	4.185	3.648	4.797	3.243	4.018	3.47
0,3	5.187	4.354	4.548	4.488	1.792	4.545	3.39
0,4	4.253	4.873	3.353	4.377	3.332	4.165	4.24
0,5	4.552	4.415	3.922	4.552	3.940	3.890	3.54

Το παράθυρο αυτό προσφέρει τη δυνατότητα στο χρήστη να δει ένα αρχείο ιστορικών δεδομένων, να το τροποποιήσει, να φυλάξει τις αλλαγές, ή να δημιουργήσει ένα νέο αρχείο ιστορικών δεδομένων. Η παρουσίαση των δεδομένων γίνεται σε γραμμές και στήλες. Οι στήλες αντιπροσωπεύουν τις υποπεριόδους ενώ οι γραμμές τις περιόδους. Στο παράθυρο του σχήματος η γραμμή 0,0 περιέχει όλες τις υποπεριόδους της πρώτης περιόδου της πρώτης θέσης για την οποία είναι διαθέσιμα τα ιστορικά δεδομένα. Οι εντολές σε αυτό το επίπεδο είναι:

➢ Αρχείο-Νέο Με την εντολή αυτή μπορεί ο χρήστης να δημιουργήσει ένα νέο αρχείο ιστορικών δεδομένων. Όταν εκτελεστεί εμφανίζεται ένα πλαίσιο διαλόγου που ζητάει από το χρήστη το πλήθος των θέσεων και το πλήθος των περιόδων και υποπεριόδων. Μόλις δοθούν τα στοιχεία αυτά το πρόγραμμα δημιουργεί ένα φύλλο εργασίας με τις κατάλληλες θέσεις για την εισαγωγή των δεδομένων.

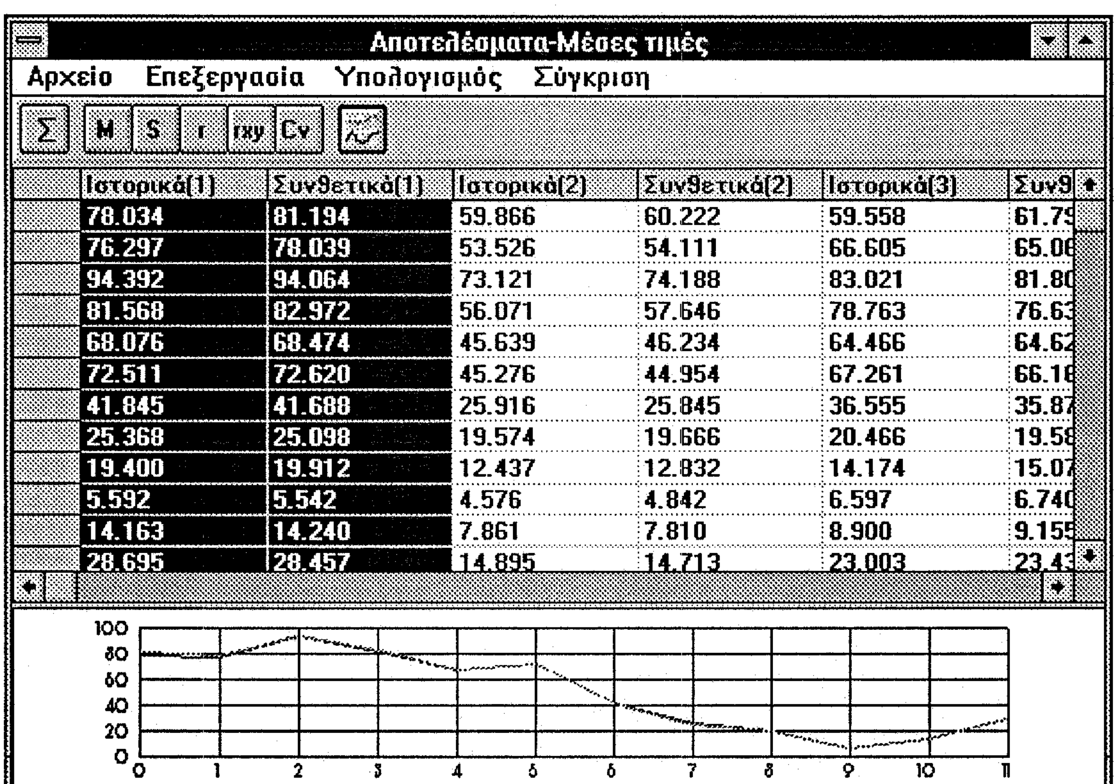
➢ Αρχείο-Άνοιγμα Με την εντολή αυτή το πρόγραμμα ανοίγει ένα ήδη υπάρχον αρχείο ιστορικών δεδομένων.

➢ Αρχείο-Αποθήκευση Με την εντολή αυτή αποθηκεύονται τα περιεχόμενα του φύλλου εργασίας σε ένα αρχείο του οποίου το όνομα ορίζει ο χρήστης.

➢ Αρχείο-Έξοδος Με την εντολή αυτή κλείνει το παράθυρο *Δεδομένα.*

Οι εντολές της επιλογής *Επεξεργασία* είναι όμοιες με αυτές που περιγράφηκαν στο υποκεφάλαιο 3.2.

3.4 Παράθυρο Στατιστικών



Ιστορικά(1)	Συνθετικά(1)	Ιστορικά(2)	Συνθετικά(2)	Ιστορικά(3)	Συνθ
78.034	81.194	59.866	60.222	59.558	61.7
76.297	78.039	53.526	54.111	66.605	65.0
94.392	94.064	73.121	74.188	83.021	81.8
81.568	82.972	56.071	57.646	78.763	76.6
68.076	68.474	45.639	46.234	64.466	64.6
72.511	72.620	45.276	44.954	67.261	66.1
41.845	41.688	25.916	25.845	36.555	35.8
25.368	25.098	19.574	19.666	20.466	19.5
19.400	19.912	12.437	12.832	14.174	15.0
5.592	5.542	4.576	4.842	6.597	6.74
14.163	14.240	7.861	7.810	8.900	9.15
28.695	28.457	14.895	14.713	23.003	23.4

Στο παράθυρο αυτό μπορεί ο χρήστης να κάνει συγκρίσεις των στατιστικών χαρακτηριστικών των ιστορικών και συνθετικών χρονοσειρών. Όλες οι εντολές των μενού μπορούν να εκτελεστούν και από τα κουμπιά που βρίσκονται πάνω από το φύλλο εργασίας. Το παράθυρο χωρίζεται σε δύο μέρη. Το πάνω, όπου βρίσκεται το φύλλο εργασίας παρουσιάζει τα στατιστικά μεγέθη που έχει ζητήσει ο χρήστης. Το κάτω παρουσιάζει τη γραφική παράσταση αυτών των στατιστικών μεγεθών για να διευκολυνθεί η σύγκριση. Οι εντολές που μπορούν να εκτελεστούν είναι:

➢ Αρχείο-Φύλαξη Με την εντολή αυτή μπορεί ο χρήστης να φυλάξει σε αρχείο τις συνθετικές χρονοσειρές χαμηλότερου επιπέδου που γέννησε το πρόγραμμα. Οι χρονοσειρές που προκύπτουν από τη διαδικασία του επιμερισμού δεν φυλάσσονται αυτόματα από το πρόγραμμα σε αντίθεση με τις χρονοσειρές υψηλότερου επιπέδου που φυλλάσσονται αυτόματα με όνομα ίδιο με αυτό του αρχείου ιστορικών δεδομένων και κάποια άλλη κατάληξη (*.agt). Αυτό γίνεται γιατί τα αρχεία σειρών χαμηλότερου επιπέδου είναι γενικώς μεγάλου μεγέθους και ο χρήστης είναι πιθανό να μη θέλει να δεσμεύσει αποθηκευτικό όγκο για αυτά. Άλλος ένας λόγος που δεν απαιτείται η φύλαξη είναι η δυνατότητα του προγράμματος να επαναλάβει επακριβώς μια διαδικασία επιμερισμού.

➢ Αρχείο-Άνοιγμα Με την εντολή αυτή μπορούν να εισαχθούν στο πρόγραμμα τα περιεχόμενα αρχείων ιστορικών δεδομένων, αρχείων με συνθετικές σειρές υψηλότερου επιπέδου και αρχείων με σειρές χαμηλότερου επιπέδου. Μόλις γίνει η εισαγωγή μπορεί ο χρήστης να δει επιλεγμένα στατιστικά μεγέθη των χρονοσειρών που περιέχουν τα παραπάνω αρχεία.

➢ Αρχείο-Έξοδος Με την εντολή αυτή κλείνει το παράθυρο *Στατιστικά.*

➢ Υπολογισμός-Απόκλιση Με την εντολή αυτή και αφού ο χρήστης έχει επιλέξει δύο στήλες του φύλλου εργασίας υπολογίζεται το σφάλμα μεταξύ των δεδομένων των στηλών, θεωρούμενης της πρώτης από τις δύο ως σωστής.

➢ Σύγκριση-Μέσες τιμές Με την εντολή αυτή υπολογίζονται οι μέσες τιμές των χρονοσειρών που έχουν ανοιχτεί.

➢ Σύγκριση-Τυπικές αποκλίσεις Με την εντολή αυτή υπολογίζονται οι τυπικές αποκλίσεις των χρονοσειρών που έχουν ανοιχτεί.

➢ Σύγκριση-Αυτοσυσχετίσεις Υπολογίζονται οι αυτοσυσχετίσεις με βήμα ένα μεταξύ των μεταβλητών της ίδιας θέσης

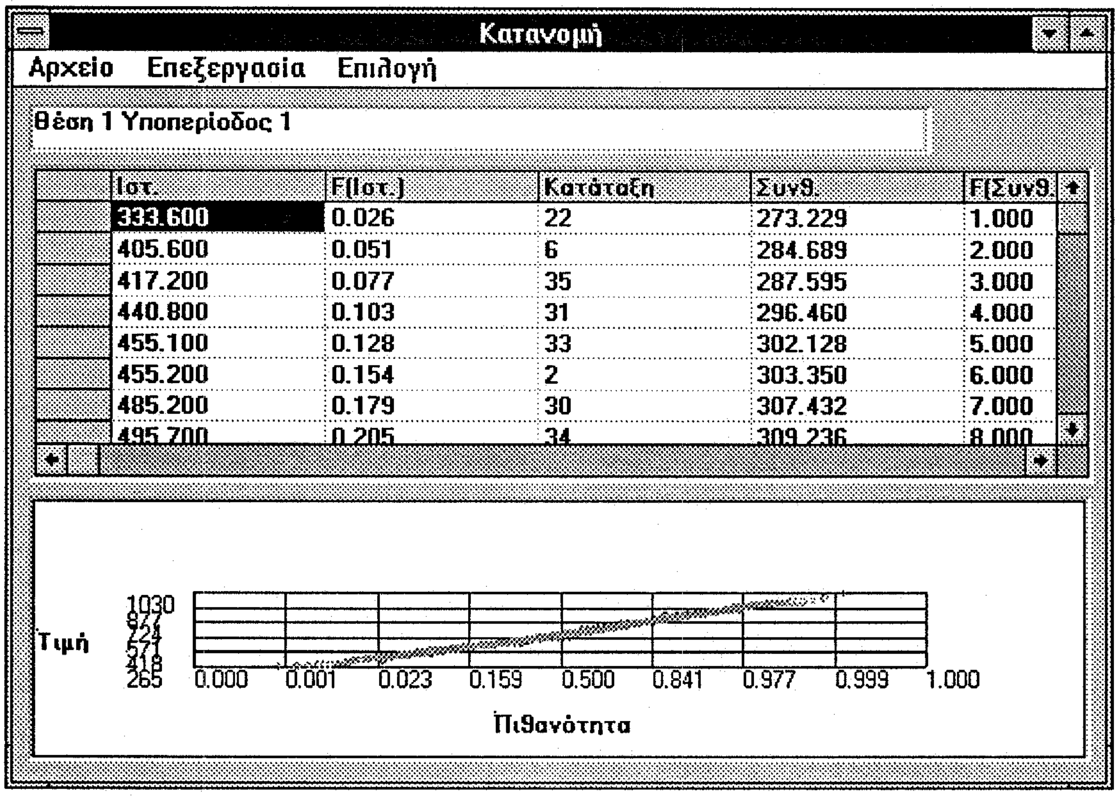
➢ Σύγκριση-Ετεροσυσχετίσεις Υπολογίζονται οι συντελεστές συσχέτισης με μηδενικό χρονικό βήμα μεταξύ μεταβλητών διαφορετικής θέσης.

➢ Σύγκριση-Ασυμμετρίες Η εντολή αυτή υπολογίζει τους συντελεστές ασυμμετρίας των χρονοσειρών που έχουν ανοιχτεί.

➢ Σύγκριση-Γράφημα Με την εντολή αυτή και αφού έχει γίνει επιλογή μίας ή περισσότερων στηλών, το πρόγραμμα εμφανίζει στην οθόνη την σχετική γραφική παράσταση.

Οι εντολές του επιλογής *Επεξεργασία* λειτουργούν με τον τρόπο που αναφέρθηκε στο υποκεφάλαιο 3.2. Άν πατηθεί το κουμπί του ποντικιού δύο φορές πάνω στο γραφικό πεδίο τότε το πεδίο καταλαμβάνει όλο το παράθυρο. Για να πάρει την αρχική του μορφή πρέπει να ξαναγίνει η ίδια διαδικασία.

3.5 Παράθυρο Κατανομής



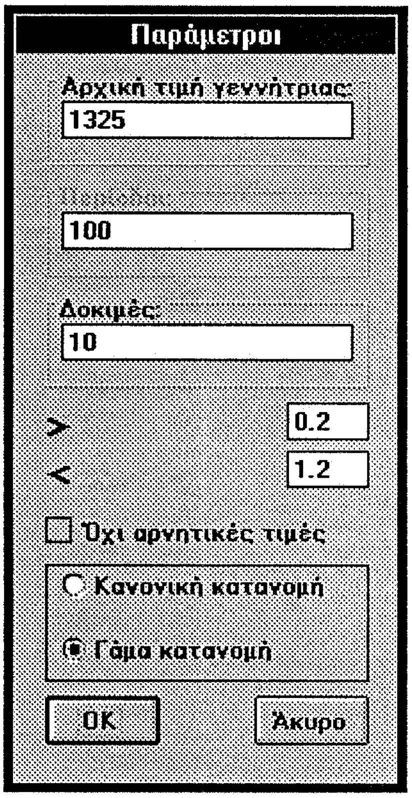
Στο παράθυρο αυτό ο χρήστης μπορεί να σχεδιάσει την εμπειρική συνάρτηση κατανομής της ιστορικής και της συνθετικής χρονοσειράς, για μία συγκεκριμένη υποπερίοδο και θέση, σε χαρτί κανονικής κατανομής. Ταυτόχρονα μπορεί να δει και τις τιμές που απαρτίζουν τις χρονοσειρές, με ή χωρίς κατάταξη σε αύξουσα σειρά, και την εμπειρική πιθανότητά τους. Οι λειτουργίες που εκτελούνται σε αυτό το επίπεδο είναι:

➢ `Αρχείο-Έξοδος` Με την εντολή αυτή κλείνει το παράθυρο *Κατανομή.*

➢ `Επιλογή-Θέση-Υποπερίοδος` Για να εμφανιστεί η γραφική παράσταση πρέπει πρώτα να επιλεγούν μια θέση και μια υποπερίοδος. Με την εντολή αυτή εμφανίζεται ένα πλαίσιο διαλόγου στο οποίο ο χρήστης συμπληρώνει τα δύο παραπάνω στοιχεία. Μόλις επιβεβαιώσει την επιλογή του το πρόγραμμα εμφανίζει τη σχετική γραφική παράσταση.

Οι εντολές του μενού *Επεξεργασία* είναι όμοιες με αυτές που αναφέρονται. στο υποκεφάλαιο 3.2.

3.6 Παράθυρο λειτουργικών δεδομένων



Το παράθυρο αυτό εμφανίζεται αυτόματα από το πρόγραμμα κάθε φορά που γίνεται γέννηση συνθετικών δεδομένων. Μέσω αυτού του παραθύρου ο χρήστης μπορεί να κάνει τις παρακάτω ρυθμίσεις:

- Να θέσει την αρχική τιμή της γεννήτριας τυχαίων αριθμών. (Χρησιμοποιείται μόνο για να υπάρχει η δυνατότητα επακριβούς επανάληψης μιας προσομοίωσης).
- Να ορίσει το πλήθος των παραγόμενων περιόδων.
- Να ορίσει το πλήθος των επιτρεπόμενων δοκιμών.
- Να καθορίσει το μέγιστο και ελάχιστο επιτρεπόμενο όριο απόκλισης.
- Να απαγορεύσει ή όχι την γέννηση αρνητικών μεταβλητών.
- Να επιλέξει την κατανομή της γεννήτριας τυχαίων αριθμών.

Στην περίπτωση του επιμερισμού το πεδίο που αναφέρεται στο πλήθος των παραγόμενων περιόδων είναι ανενεργό. Οι παραγόμενες περίοδοι μεταβλητών χαμηλότερου επιπέδου, μέσω της διαδικασίας του επιμερισμού, καθορίζονται από το πλήθος των περιόδων του αρχείου συνθετικών δεδομένων για τις μεταβλητές υψηλότερου επιπέδου. Επίσης όταν στο κύριο παράθυρο της εφαρμογής έχει επιλεγεί *Λογαριθμικός μετασχηματισμός* τότε η κατανομή Γάμα δεν μπορεί να επιλεγεί.

4. Παραδείγματα εφαρμογής

4.1 Παραγωγή ετήσιων χρονοσειρών με αρχείο δεδομένων

Στο παράδειγμα αυτό θα περιγραφεί η γέννηση 2000 χρόνων ετήσιων δεδομένων. Χρησιμοποιούμε το αρχείο ιστορικών δεδομένων `ryliki.adt`, το οποίο περιέχει 38 συνεχή χρόνια ετήσιων βροχοπτώσεων για την περιοχή της Υλίκης. Οι θέσεις-σταθμοί είναι τρεις: Αλίαρτος, Μουρίκι και Καλλιθέα. Ακολουθούμε την εξής διαδικασία:

1. Εισάγουμε το αρχείο στη μνήμη του προγράμματος. Για να γίνει αυτό εκτελούμε την εντολή `Αρχείο-Άνοιγμα` και διαλέγουμε το αρχείο `ryliki.adt`.
2. Δίνουμε τα λειτουργικά δεδομένα με τον ακόλουθο τρόπο:
   - Εκτελούμε την εντολή `Προσομοίωση-Παραγωγή ετήσιων`.
   - Στο παράθυρο που εμφανίζεται πληκτρολογούμε στο πεδίο *Περίοδοι* τον αριθμό 2000.
   - Οι δοκιμές και τα όρια δεν χρησιμοποιούνται γι' αυτό δεν εισάγουμε στα αντίστοιχα πεδία τίποτε.
   - Επιλέγουμε (με κλικ) *Όχι αρνητικές τιμές.*
   - Επιλέγουμε *Κανονική κατανομή.*
3. Πατάμε το κουμπί *ΟΚ* για να επιβεβαιώσουμε τις επιλογές μας.

Το πρόγραμμα ξεκινάει την γέννηση των συνθετικών χρονοσειρών. Η λήξη του σημαίνεται με την εμφάνιση του μηνύματος *Τέλος.*

Στο σημείο αυτό μπορούμε να δούμε τα αποτελέσματα της γέννησης. Τα βήματα που πρέπει να κάνουμε είναι:

1. Εκτελούμε την εντολή `Παρουσίαση-Στατιστικά`.
2. Στο παράθυρο που εμφανίζεται εκτελούμε την εντολή `Σύγκριση-Μέσες τιμές`. Αμέσως το πρόγραμμα εμφανίζει τις μέσες τιμές των ιστορικών και των συνθετικών χρονοσειρών. Με παρόμοιο τρόπο μπορούμε να συγκρίνουμε και τα υπόλοιπα στατιστικά μεγέθη.
3. Οι συνθετικές χρονοσειρές έχουν φυλαχτεί στο δίσκο σε αρχείο με το όνομα `ryliki.arg`. Αν επιθυμούμε να τις φυλάξουμε σε αρχείο με κάποιο άλλο όνομα πρέπει να εκτελέσουμε την εντολή `Αρχείο-Φύλαξη`. Στο πλαίσιο διαλόγου που εμφανίζεται, δίνουμε το όνομα που επιθυμούμε και πατούμε το κουμπί *ΟΚ.*

Εκτός από τα στατιστικά μεγέθη μπορούμε να συγκρίνουμε και τις εμπειρικές κατανομές των ιστορικών και συνθετικών χρονοσειρών. Για να γίνει αυτό πρέπει:

1. Να εκτελέσουμε την εντολή `Παρουσίαση-Κατανομή` του κύριου παραθύρου (SIDIS).

2. Να εκτελέσουμε, στο παράθυρο που εμφανίζεται, την εντολή Επιλογή-Θέση-Υποπερίοδος.

3. Να πληκτρολογήσουμε, στα πεδία *Θέση* και *Υποπερίοδος* του πλαισίου διαλόγου, τους αριθμούς που θέλουμε, λαμβάνοντας υπόψη ότι η προσομοίωση γίνεται σε τρεις θέσεις και μία υποπερίοδο (ετήσια προσομοίωση). Μόλις γίνει αυτό εμφανίζεται γραφική παράσταση της κατανομής σε χαρτί κανονικής κατανομής.

4.2 Επιμερισμός με τη χρήση αρχείου δεδομένων

Για να γίνει επιμερισμός πρέπει να εισαχθεί ένα αρχείο με μηνιαία δεδομένα. Το αρχείο που χρησιμοποιείται στο παράδειγμα έχει όνομα `ryliki.pdt`. Μόλις γίνει η εισαγωγή του αρχείου το πρόγραμμα ψάχνει το αρχείο συνθετικών ετήσιων δεδομένων με όνομα `ryliki.arg` για να το επιμερίσει. Το αρχείο αυτό υπάρχει από το προηγούμενο παράδειγμα. Αναλυτικά τα βήματα που πρέπει να ακολουθήσουμε είναι:

1. Εκτελούμε την εντολή Αρχείο-Άνοιγμα. Επιλέγουμε το αρχείο `ryliki.pdt`.

2. Εκτελούμε την εντολή Προσομοίωση-Επιμερισμός.

3. Στο παράθυρο παραμέτρων της προσομοίωσης που εμφανίζεται παρατηρούμε ότι το πεδίο *Περίοδοι* είναι ανενεργό. Αυτό γίνεται γιατί οι περίοδοι της διαδικασίας επιμερισμού καθορίζονται από το μήκος των συνθετικών χρονοσειρών του ετήσιου αρχείου. Στο παράδειγμα το αρχείο αυτό έχει χρονοσειρές μήκους 2000 και έτσι 2000 θα είναι και οι περίοδοι του επιμερισμού. Το πεδίο *Δοκιμές* τώρα παίζει ρόλο στην προσομοίωση, σε αντίθεση με το προηγούμενο παράδειγμα. Ο αριθμός που μπορούμε να εισαγάγουμε στο πεδίο αυτό δεν μπορεί να υπερβεί την τιμή 32767. Βεβαίως, όσο πιο μεγάλη είναι η τιμή που θα εισαχθεί στο πεδίο αυτό, τόσο μεγαλύτερος θα είναι και ο χρόνος που θα χρειαστεί για την προσομοίωση. Το πεδίο που αφορά στο άνω όριο επιτρεπτού σφάλματος (<) είναι εξίσου σημαντικό. Η ταχύτητα εκτέλεσης της προσομοίωσης εξαρτάται και απ' αυτό. Αν εισαγάγουμε ένα πολύ μικρό αριθμό, τότε το πρόγραμμα, στην προσπάθειά του να ικανοποιήσει τη συνθήκη του άνω επιτρεπτού ορίου, θα εξαντλεί όλες τις δοκιμές που επιτρέπονται, και έτσι θα ελαττωθεί η ταχύτητα εκτέλεσης της προσομοίωσης. Στο παράδειγμα οι δοκιμές είναι 100 και το άνω όριο σφάλματος 0.3 (το κάτω όριο 0). Η κατανομή που υιοθετείται για τις μεταβλητές είναι *Κανονική*.

4. Μόλις πατήσουμε το κουμπί *OK* ξεκινάει ο επιμερισμός. Με το ξεκίνημα εμφανίζεται ένα παράθυρο με δύο οριζόντιες "μπάρες". Η πάνω δείχνει τις περιόδους που έχουν ολοκληρωθεί ενώ η κάτω δείχνει το πλήθος των δοκιμών για κάθε περίοδο που ολοκληρώνεται.

Το πέρας του επιμερισμού σημαίνεται με την εμφάνιση του μηνύματος *Ο επιμερισμός τελείωσε*. Η σύγκριση των στατιστικών παραμέτρων και κατανομών των ιστορικών και συνθετικών σειρών γίνεται με τον τρόπο που περιγράφηκε πιο πάνω (υποκεφάλαιο 4.1).

4.3 Παραγωγή ετήσιων χρονοσειρών χωρίς αρχείο δεδομένων

Το αρχείο ιστορικών δεδομένων απαιτείται για να υπολογίσει το πρόγραμμα τις διάφορες παραμέτρους που χρειάζονται τα μοντέλα προσομοίωσης. Αν δεν υπάρχει αρχείο με ιστορικά δεδομένα τότε πρέπει ο χρήστης να τα εισαγάγει μόνος του. Στο παράδειγμα θα γεννηθούν 2000 χρόνια ετήσιων βροχών για τη λεκάνη της Υλίκης χωρίς αυτή τη φορά να χρησιμοποιηθεί το αρχείο με τα ιστορικά δεδομένα. Τα βήματα που απαιτούνται είναι:

1. Στο κύριο παράθυρο (SIDIS) σημειώνουμε (με κλικ) την επιλογή *Παράμετροι από χρήστη.*
2. Εκτελούμε την εντολή `Προσομοίωση-Παραγωγή ετήσιων.`
3. Στο πλαίσιο διαλόγου που εμφανίζεται, συμπληρώνουμε τα λειτουργικά δεδομένα (όπως στα προηγούμενα παραδείγματα).
4. Στο παράθυρο που ακολουθεί εισάγουμε στα πεδία *Θέσεις* και *Υποπερίοδοι* τους αριθμούς 3 και 1 αντίστοιχα. Εκτελούμε την εντολή `Αρχείο-Νέο`. Το πρόγραμμα εμφανίζει στο φύλλο εργασίας κάποια στοιχεία που μας καθοδηγούν στη συμπλήρωση των παραμέτρων. Μόλις τελειώσει η συμπλήρωση για να ξεκινήσει η γέννηση πρέπει να πατήσουμε το δεξί πλήκτρο του ποντικιού.

4.4 Επιμερισμός χωρίς αρχείο δεδομένων

Κατά τη διάρκεια του προηγούμενου παραδείγματος το πρόγραμμα φύλαξε τα συνθετικά δεδομένα σε ένα αρχείο για να μπορούν να επιμεριστούν στη συνέχεια. Το όνομα του αρχείου αυτού είναι `from_ar1` και κάθε φορά που γίνεται γέννηση ετήσιων σειρών χωρίς αρχείο δεδομένων τα αποτελέσματα γράφονται σε αυτό το αρχείο. Για να γίνει ο επιμερισμός του πρέπει να ακολουθήσουμε τα παρακάτω βήματα:

1. Στο κύριο παράθυρο (SIDIS) σημειώνουμε (με κλικ) την επιλογή *Παράμετροι από χρήστη.*
2. Εκτελούμε την εντολή `Προσομοίωση-Επιμερισμός.`
3. Στο πλαίσιο διαλόγου που εμφανίζεται συμπληρώνουμε τα λειτουργικά δεδομένα (όπως στα προηγούμενα παραδείγματα).
4. Στο παράθυρο που εμφανίζεται εισάγουμε στα πεδία *Θέσεις* και *Υποπερίοδοι* τους αριθμούς 3 και 12 αντίστοιχα. Εκτελούμε την εντολή `Αρχείο-Νέο`. Το πρόγραμμα εμφανίζει στο φύλλο εργασίας κάποια στοιχεία που μας καθοδηγούν στη συμπλήρωση των παραμέτρων. Μόλις τελειώσει η συμπλήρωση για να ξεκινήσει η γέννηση πρέπει να πατήσουμε το δεξί πλήκτρο του ποντικιού.

Παράρτημα Α: Περιγραφή του απλού μοντέλου επιμερισμού

1. Συμβολισμοί και μεθοδολογία

Συμβολίζουμε με $p = 1, 2, \ldots$, το δείκτη που αναφέρεται στο χρονικό βήμα των μεταβλητών υψηλότερου επιπέδου (δηλαδή τον αύξοντα αριθμό της περιόδου), με $s = 1, 2, \ldots, k$ τον αντίστοιχο δείκτη για τις μεταβλητές χαμηλότερου επιπέδου (δηλαδή τον αύξοντα αριθμό της υποπεριόδου μέσα στη δεδομένη περίοδο), όπου k είναι ο αριθμός των υποπεριόδων μιας περιόδου (πχ. 12 μήνες), και με l το δείκτη που ξεχωρίζει τις μεταβλητές διαφορετικής θέσης (n το πλήθος). Τότε:

${}_pX_s^l$ είναι η μεταβλητή χαμηλότερου επιπέδου που αντιστοιχεί στη θέση l την περίοδο p και την υποπερίοδο s.

${}_pZ^l$ είναι η μεταβλητή υψηλότερου επιπέδου που αντιστοιχεί στην περίοδο p και τη θέση l.

${}_p\mathbf{X}_s$ είναι το διάνυσμα που περιέχει τις μεταβλητές χαμηλότερου επιπέδου όλων των θέσεων για την υποπερίοδο s και την περίοδο p (μεγέθους n).

${}_p\mathbf{X}^l$ είναι το διάνυσμα που περιέχει τις μεταβλητές χαμηλότερου επιπέδου όλων των υποπεριόδων για την περίοδο p και τη θέση l (μεγέθους k).

${}_p\mathbf{Z}$ είναι το διάνυσμα των μεταβλητών υψηλότερου επιπέδου σε όλες τις θέσεις (μεγέθους n).

Τα διανύσματα ${}_p\mathbf{Z}$ και ${}_p\mathbf{X}_s$ ικανοποιούν την αθροιστική ιδιότητα, δηλαδή

$$\sum_{s=1}^{k} {}_p\mathbf{X}_s = {}_p\mathbf{Z} \tag{1}$$

Ο δείκτης p μπορεί να παραλειφθεί αν αναφερόμαστε σε κάποια συγκεκριμένη περίοδο. Στην περίπτωση αυτή τις μεταβλητές χαμηλότερου επιπέδου της προηγούμενης περιόδου τις συμβολίζουμε με $X_0^l, X_{-1}^l, \ldots$, ενώ αυτές της επόμενης με $X_{k+1}^l, X_{k+2}^l, \ldots$, όπου $X_0^l := {}_{p-1}X_k^l$, $X_{k+1}^l := {}_{p+1}X_1^l$, κοκ. Ο δείκτης θέσης l δεν ξεχωρίζει κατ' ανάγκη μεταβλητές διαφορετικής θέσης. Μπορεί να ξεχωρίσει και μεταβλητές της ίδιας θέσης που περιγράφουν διαφορετικά υδρολογικά μεγέθη (π.χ. βροχή και απορροή στην ίδια θέση). Επίσης μπορεί να παραλειφθεί αν αναφερόμαστε σε μονοδιάστατο πρόβλημα. Ο δείκτης υποπεριόδου s δεν μπορεί σε καμία περίπτωση να παραλειφθεί και έτσι με το συμβολισμό $\mathbf{X}$ (χωρίς δείκτες) θα αναφερόμαστε σ' ένα διάνυσμα μεγέθους k που περιέχει τις μεταβλητές χαμηλότερου επιπέδου για κάποια συγκεκριμένη περίοδο και θέση. Ο παραπάνω συμβολισμός ταιριάζει σε πολυδιάστατο χρονικό επιμερισμό αλλά μπορεί εύκολα να τροποποιηθεί και για χωρικό επιμερισμό.

Στην εργασία αυτή εξετάζουμε λεπτομερώς την περίπτωση όπου οι μεταβλητές χαμηλότερου επιπέδου περιγράφονται από ένα περιοδικό μοντέλο AR(1) (ή PAR(1)), οπότε ικανοποιούν τη σχέση

$$\mathbf{X}_s = \mathbf{a}_s\, \mathbf{X}_{s-1} + \mathbf{b}_s\, \mathbf{V}_s \tag{2}$$

όπου $\mathbf{a}_s$ είναι ένας $(n \times n)$ διαγώνιος πίνακας δηλαδή $\mathbf{a}_s = diag\left(a_s^1, \ldots, a_s^n\right)$, $\mathbf{b}_s = \left[b_s^{lj}\right]$ είναι ένας $(n \times n)$ πίνακας συντελεστών και $V_s = \left[V_s^1, \ldots, V_s^n\right]^T$ είναι ένα διάνυσμα ανεξάρτητων, στο χώρο και το χρόνο, τυχαίων μεταβλητών, όχι απαραίτητα κανονικής κατανομής. Η εξίσωση (2) δεν είναι περιοριστική για το μοντέλο. Τη θέση της θα μπορούσε να πάρει η εξίσωση ενός PAR ή PARMA μεγαλύτερου βαθμού. Στη συγκεκριμένη εργασία όμως, επειδή θέλαμε το μοντέλο να είναι απλό στη μαθηματική του διατύπωση και φειδωλό σε παραμέτρους χρησιμοποιήσαμε την (2). Στην παρούσα μορφή διατηρούνται τα παρακάτω στατιστικά μεγέθη:

1. μέσες τιμές των μεταβλητών χαμηλότερου επιπέδου δηλαδή k διανύσματα $\boldsymbol{\xi}_s = E[\mathbf{X}_s]$ μεγέθους n·
2. διασπορές και μηδενικού βήματος ετεροσυσχετίσεις για τις μεταβλητές χαμηλότερου επιπέδου δηλαδή k το πλήθος πίνακες $\boldsymbol{\sigma}_s = \mathrm{Cov}[\mathbf{X}_s, \mathbf{X}_s]$ $= E[(\mathbf{X}_s - \boldsymbol{\xi}_s)(\mathbf{X}_s - \boldsymbol{\xi}_s)^T]$ μεγέθους $(n \times n)$·
3. αυτοσυσχετίσεις μοναδιαίου βήματος των μεταβλητών χαμηλότερου επιπέδου δηλαδή k διανύσματα $\boldsymbol{\tau}_s = \left[\mathrm{Cov}[X_s^l, X_{s-1}^l]\right] = \left[E[(X_s^l - \xi_s^l)(X_{s-1}^l - \xi_{s-1}^l)]\right]$ μεγέθους n·
4. τρίτες ροπές των μεταβλητών χαμηλότερου επιπέδου δηλαδή k διανύσματα $\boldsymbol{\gamma}_s = \left[E[(X_s^l - \xi_s^l)^3]\right]$ μεγέθους n·

Χωρίς απώλεια της γενικότητας, στην παραπάνω περιγραφή παραλείψαμε το δείκτη της περιόδου γιατί παραδεχόμαστε στασιμότητα ανάμεσα στις διαδοχικές περιόδους. Λαμβάνοντας υπόψη τη συμμετρία του πίνακα σ_s ο ολικός αριθμός παραμέτρων δεύτερης τάξης (ομάδα 2 και 3) για όλες τις υποπεριόδους και θέσεις είναι $kn(n + 3)/2$ και μπορεί να μειωθεί σε $k(3n - 1)$ αν διατηρηθεί μόνο μία ετεροσυσχέτιση μηδενικού βήματος ανά μεταβλητή. Οι παράμετροι αυτές είναι σημαντικά λιγότερες από τις αντίστοιχες των κοινών επιμεριστικών μοντέλων [βλ. *Koutsoyiannis,* 1992].

Οι παράμετροι $\mathbf{a}_s$ και $\mathbf{b}_s$ τις εξίσωσης (2) σχετίζονται με τις ομάδες παραμέτρων 2 και 3 με τις παρακάτω σχέσεις:

$$a_s^l = \tau_s^l / \sigma_{s-1}^{ll} \tag{3}$$

$$\mathbf{b}_s \mathbf{b}_s^T = \boldsymbol{\sigma}_s - \mathbf{a}_s \boldsymbol{\sigma}_{s-1} \mathbf{a}_s \tag{4}$$

Οι εξισώσεις αυτές αποτελούν επέκταση των εξισώσεων του στάσιμου Μαρκοβιανού μοντέλου που δίνονται από τους *Matalas and Wallis* [1976, p. 63] για να ληφθεί υπόψη και η εποχιακή διακύμανση. Ο πίνακας $\mathbf{b}_s$ προκύπτει από το Γκραμιανό $\mathbf{b}_s\mathbf{b}_s^T$ είτε με τη μέθοδο της τριγωνικής διάσπασης είτε με τη μέθοδο των ιδιοτιμών.

Οι στατιστικές παράμετροι του τυχαίου διανύσματος V_s^l, που χρειάζονται για να ικανοποιηθεί η σχέση (2), δίνονται από τις παρακάτω εξισώσεις:

$$E[\mathbf{V}_s] = \mathbf{b}_s^{-1}\{E[\mathbf{X}_s] - \mathbf{a}_s E[\mathbf{X}_{s-1}]\} \tag{5}$$

$$\mathrm{Var}[V_s^l] = 1 \tag{6}$$

$$\mu_3[\mathbf{V}_s] = \left(\mathbf{b}_s^{(3)}\right)^{-1}\left(\boldsymbol{\gamma}_s - \mathbf{a}_s^{(3)}\boldsymbol{\gamma}_{s-1}\right) \tag{7}$$

όπου $\mu_3[\mathbf{V}_s]$ είναι το διάνυσμα των τρίτων κεντρικών ροπών της $\mathbf{V}_s$ και ο εκθέτης (3) στον πίνακα δείχνει ότι κάθε στοιχείο του πίνακα πρέπει να υψωθεί στον κύβο.

Μια εναλλακτική περίπτωση της σχέσης (2), που εξετάστηκε εξίσου στην εργασία αυτή, είναι ο μη γραμμικός μετασχηματισμός των μεταβλητών χαμηλότερου επιπέδου μέσω της σχέσης:

$$\mathbf{X}_s^* := \ln(\mathbf{X}_s - \mathbf{c}_s) \quad (8)$$

όπου $\mathbf{c}_s$ είναι ένα διάνυσμα παραμέτρων που υπολογίζεται με τέτοιο τρόπο ώστε οι μεταβλητές $\mathbf{X}_s^*$ να ακολουθούν την κανονική κατανομή. Βάση του παραπάνω μετασχηματισμού η σχέση (2) γίνεται:

$$\mathbf{X}_s^* = \mathbf{a}_s\,\mathbf{X}_{s-1}^* + \mathbf{b}_s\,\mathbf{V}_s \quad (9)$$

Τα στατιστικά μεγέθη που μας ενδιαφέρουν είναι όμοια με αυτά των ομάδων 1 έως 3 που αναφέραμε πιο πάνω. Βεβαίως, αφορούν στις μετασχηματισμένες μεταβλητές χαμηλότερου επιπέδου $\mathbf{X}_s^*$. Στην ομάδα 4 οι τρίτες ροπές τώρα αντικαθιστώνται από τις παραμέτρους $\mathbf{c}_s$. Οι εξισώσεις (3) μέχρι (6) ισχύουν επακριβώς αν αντικαταστήσουμε τον όρο $\mathbf{X}_s$ με $\mathbf{X}_s^*$. Η εξίσωση (7) αντικαθίσταται από την παραδοχή μηδενικής τρίτης ροπής για το διάνυσμα των τυχαίων μεταβλητών (κανονική κατανομή).

Τα μοντέλα που περιγράφονται με τις σχέσεις (2) και (8)-(9) είναι κατάλληλα για την παραγωγή μεταβλητών χαμηλότερου επιπέδου $\mathbf{X}_s$ αλλά δεν λαμβάνουν υπόψη τις ήδη γνωστές μεταβλητές υψηλότερου επιπέδου $\mathbf{Z}$, δηλαδή δεν είναι μοντέλα επιμερισμού. Αυτό που συνήθως κάνουν τα μοντέλα επιμερισμού είναι μια υπό συνθήκη παραγωγή μεταβλητών χαμηλότερου επιπέδου περιλαμβάνοντας στις εξισώσεις πληροφορία σχετική και με τις μεταβλητές υψηλότερου επιπέδου. Η μέθοδος που παρουσιάζουμε εδώ είναι αρκετά διαφορετική και κατά πολύ απλούστερη. Τα βήματά της είναι τα ακόλουθα:

1. Με τη χρήση της εξίσωσης (2) ή (8)-(9) παράγονται οι μεταβλητές χαμηλότερου επιπέδου, $\tilde{\mathbf{X}}_s$, μιας περιόδου χωρίς καμία αναφορά στη δεδομένη μεταβλητή υψηλότερου επιπέδου για την περίοδο αυτή.
2. Υπολογίζεται το άθροισμα $\tilde{\mathbf{Z}} = \sum_{s=1}^{k} \tilde{\mathbf{X}}_s$ και η απόστασή του από την πραγματική τιμή $\Delta Z = \|\mathbf{Z} - \tilde{\mathbf{Z}}\|$ (βλ. Κεφάλαιο 4).
3. Επαναλαμβάνονται τα βήματα 1 και 2 μέχρις ότου η απόσταση ΔZ να είναι μικρότερη από ένα προκαθορισμένο επιτρεπτό όριο. Μόλις ικανοποιηθεί η συνθήκη αυτή ακολουθεί το επόμενο βήμα. Το βήμα 3 είναι προαιρετικό. (βλ. Κεφάλαιο 4)
4. Στις μεταβλητές χαμηλότερου επιπέδου, $\tilde{\mathbf{X}}_s$, που μόλις έχουν γεννηθεί, παρεμβαίνει η διορθωτική διαδικασία και προκύπτουν οι νέες μεταβλητές $\mathbf{X}_s$ που ικανοποιούν τη σχέση $\sum_{s=1}^{k} \mathbf{X}_s = \mathbf{Z}$.

5. Επαναλαμβάνονται τα βήματα 1 έως 4 για όλες τις περιόδους με γνωστό **Z**.

Η διαδικασία διόρθωσης μπορεί να περιγραφεί ως μια συνάρτηση

$$\mathbf{X}_s = f\left(\tilde{\mathbf{X}}_s, \tilde{\mathbf{Z}}, \mathbf{Z}, \text{στατιστικές παράμετροι}\right) \tag{10}$$

Αν δοθούν τα κατάλληλα ορίσματα στην συνάρτηση αυτή, τότε προκύπτουν οι μεταβλητές χαμηλότερου επιπέδου που ικανοποιούν τη σχέση (1). Κανονικά η συνάρτηση μπορεί να εκφραστεί ως σύνολο n ισοδύναμων μονοδιάστατων εξισώσεων, μια για κάθε θέση. Συγκεκριμένες μορφές της συνάρτησης (10) που διατηρούν επακριβώς ορισμένα στατιστικά μεγέθη εξετάζονται παρακάτω.

Αυτή η επιμεριστική μέθοδος μπορεί να δώσει αποτελέσματα ισάξια με αυτά ενός PAR(1) όσον αφορά τη διατήρηση των στατιστικών μεγεθών που αναφέραμε παραπάνω. Όσο μικρότερο είναι το όριο του βήματος 3 τόσο μεγαλύτερη είναι η προσέγγιση της μεθόδου προς το μοντέλο PAR(1). Οστόσο τα πολύ μικρά όρια ελαττώνουν την ταχύτητα της μεθόδου. Είναι σημαντικό να τονίσουμε πάντως ότι σε πολλές εφαρμογές που εξετάστηκαν κατά τη δοκιμή του μοντέλου μία επανάληψη ήταν αρκετή για να δώσει ικανοποιητικά αποτελέσματα. Η ακρίβεια της διαδικασίας διόρθωσης είναι υπεύθυνη για το παραπάνω συμπέρασμα.

Η προτεινόμενη μέθοδος μπορεί να τροποποιηθεί πολύ εύκολα αν αλλάξουμε την εξίσωση του PAR(1) και στη θέση της τοποθετήσουμε την εξίσωση οποιουδήποτε άλλου γραμμικού μοντέλου. Η μετατροπή αυτή βέβαια θα επιφέρει αύξηση των παραμέτρων του μοντέλου.

2. Ακριβείς διορθωτικές διαδικασίες

Στην ενότητα αυτή παρουσιάζουμε τρεις διαδικασίες διόρθωσης μαζί με το θεωρητικό τους υπόβαθρο. Η πρώτη είναι μια περιορισμένης εμβέλειας διαδικασία γιατί μπορεί να εφαρμοστεί μόνο σε μεταβλητές χαμηλότερου επιπέδου που ακολουθούν κατανομή γάμα, με κάποιους περιορισμούς. Η δεύτερη είναι πιο γενική και μπορεί να εφαρμοστεί σε οποιασδήποτε κατανομής μεταβλητές διατηρώντας τις ροπές μέχρι και δεύτερης τάξης. Επίσης μπορεί να διατηρήσει πλήρως την από κοινού κατανομή για κανονικές μεταβλητές. Η τρίτη διαδικασία είναι μια παραλλαγή της δεύτερης που εφαρμόζεται μόνο στην περίπτωση θετικών μεταβλητών χαμηλότερου επιπέδου. Σε ειδικές περιπτώσεις ταυτίζεται και με την πρώτη. Η τρίτη διαδικασία δεν είναι ακριβής, εκτός από ορισμένες περιπτώσεις οι οποίες καλύπτονται και από τις δύο προηγούμενες, αλλά έχει το πλεονέκτημα ότι επιστρέφει μόνο θετικές μεταβλητές. Μια και όλες οι παραπάνω διαδικασίες εφαρμόζονται σε κάθε θέση ξεχωριστά, στο κεφάλαιο αυτό θα παραλείψουμε το δείκτη εκείνο που αναφέρεται στη θέση. Τονίζεται σ' αυτό το σημείο ότι όλες οι διαδικασίες διόρθωσης μπορούν να ενσωματωθούν σε οποιοδήποτε μοντέλο επιμερισμού που χρειάζεται διόρθωση των μεταβλητών που παράγει.

2.1 Διατήρηση περιθώριων κατανομών Γάμα (Αναλογική διορθωτική διαδικασία)

Σε μια πρόσφατη εργασία [*Koutsoyiannis*, 1994] είχε μελετηθεί η ακόλουθη πρόταση:

Πρόταση 1: Αν με $\tilde{X}_s$ συμβολίσουμε κάποιες ανεξάρτητες μεταβλητές κατανομής γάμα με παραμέτρους κ_s και λ ($s = 1, \dots, k$) και με Z μια μεταβλητή ανεξάρτητη των $\tilde{X}_s$ με κατανομή επίσης γάμα και παραμέτρους

$$\kappa := \sum_{s=1}^{k} \kappa_s \tag{11}$$

και λ τότε οι μεταβλητές

$$X_s := \frac{\tilde{X}_s}{\sum_{j=1}^{k} \tilde{X}_j} Z \qquad (s = 1, \dots k) \tag{12}$$

είναι ανεξάρτητες και έχουν κατανομή γάμα με παραμέτρους κ_s και λ.

Η απόδειξη της παραπάνω πρότασης δίνεται από τον *Koutsoyiannis* [1994· Appendix B]. Σημειώνεται ότι οι μεταβλητές X_s έχουν την ίδια κατανομή με τις αντίστοιχες $\widetilde{X}_s$ και αθροιζόμενες μας δίνουν τη μεταβλητή Z. Τονίζουμε ότι η παραπάνω πρόταση δεν μπορεί να επεκταθεί θεωρητικά για κατανομές γάμα με διαφορετικές παραμέτρους κλίμακας ή για κανονικές κατανομές ακόμη και στην περίπτωση που αυτές υπακούουν στους ίδιους περιορισμούς. Η πρόταση αυτή οδηγεί στην παρακάτω διορθωτική διαδικασία

$$X_s = \frac{Z}{\widetilde{Z}} \widetilde{X}_s \tag{13}$$

όπου $\widetilde{Z}$ είναι το άθροισμα όλων των $\widetilde{X}_s$. Αυτή η διαδικασία ονομάζεται *αναλογική διορθωτική διαδικασία*. Δίνει τον πιο φυσικό τρόπο μετατροπής των $\widetilde{X}_s$ σε X_s και συναντάται σε πολλά άλλα μοντέλα επιμερισμού [π.χ. *Grygier and Stedinger*, 1988; *Lane and Frevert*, 1990, σελ. V-22]. Παρ' όλα αυτά, σύμφωνα με την εμπειρία των συγγραφέων, δεν έχει μελετηθεί, θεωρητικά ή πρακτικά, σε προηγούμενες εργασίες το πεδίο και τα όρια εφαρμογής της.

Η πιο πάνω θεωρητική τεκμηρίωση της αναλογικής διαδικασίας φανερώνει και τους ισχυρούς περιορισμούς της: (α) οι μεταβλητές X_s πρέπει να ακολουθούν τη διπαραμετρική κατανομή γάμα· (β) πρέπει να έχουν κοινή παράμετρο κλίμακας, γεγονός που απαιτεί το πηλίκο $\mathrm{E}[X_s]/\mathrm{Var}[X_s]$ να είναι σταθερό για όλες τις υποπεριόδους· και (γ) όλες οι X_s πρέπει να είναι αμοιβαία ανεξάρτητες. Παρόλα αυτά, μετά από πρακτικές εφαρμογές προέκυψε ότι η διαδικασία διόρθωσης αυτή δίνει ικανοποιητικά αποτελέσματα ακόμη και σε περιπτώσεις που δεν ικανοποιούνται οι παραπάνω περιορισμοί, μολονότι δεν μπορεί να θεωρηθεί ακριβής. Ικανοποιητική προσέγγιση μπορεί να προκύψει κάτω από τους ακόλουθους ελαστικότερους όρους: (α) οι μεταβλητές X_S ακολουθούν κατά προσέγγιση κατανομή γάμα δύο παραμέτρων· (β) η ποσότητα $\mathrm{E}[X_s]/\mathrm{Var}[X_s]$ είναι σχετικά σταθερή ανάμεσα στις διαφορετικές υποπεριόδους· και (γ) οι μεταβλητές X_S έχουν σχετικά μικρό συντελεστή συσχέτισης $\mathrm{Corr}[X_s, X_j]$ (μια τιμή της τάξης 0.60-0.70 διατηρείται προσεγγιστικά από την εν λόγω διαδικασία χωρίς πρόβλημα). Οι παραπάνω όροι ικανοποιούνται, για παράδειγμα, σε χρονοσειρές βροχής μικρής κλίμακας (π.χ. ωριαίας) και γι' αυτό η διαδικασία έχει χρησιμοποιηθεί για τον επιμερισμό του ολικού ύψους επεισοδίων βροχής σε ωριαία ύψη [*Koutsoyiannis*, 1994].

2.2 Διατήρηση στατιστικών μεγεθών δευτέρας τάξης (Γραμμική διορθωτική διαδικασία)

Η δεύτερη διορθωτική διαδικασία βασίζεται στην παρακάτω πρόταση:

Πρόταση 2: Αν $\widetilde{X}_s$ ($s = 1, \ldots, k$) είναι οποιεσδήποτε τυχαίες μεταβλητές με μέσες τιμές $\xi_s = E[X_s]$ και πίνακα διασποράς-συνδιασποράς **σ** με στοιχεία $\sigma_{sj} = \mathrm{Cov}[X_s, X_j] = E[(X_s - \xi_s)(X_j - \xi_j)]$, Z είναι μια άλλη ανεξάρτητη της $\widetilde{X}_s$ μεταβλητή με μέση τιμή

$$\xi_Z := E[Z] = \sum_{s=1}^{k} \xi_s \tag{14}$$

και διασπορά

$$\sigma_{ZZ} := \mathrm{Var}[Z] = \sum_{s=1}^{k} \sum_{j=1}^{k} \sigma_{sj} \tag{15}$$

τότε οι μεταβλητές

$$X_s := \widetilde{X}_s + \lambda_s \left(Z - \sum_{j=1}^{k} \widetilde{X}_j \right) \qquad (s = 1, \ldots, k) \tag{16}$$

έχουν ίδιες μέσες τιμές και πίνακα διασπορών - συνδιασπορών με τις $\widetilde{X}_s$, εφόσον ο συντελεστής λ_s οριστεί κατάλληλα δηλαδή

$$\lambda_s := \sigma_{sZ} / \sigma_{ZZ} \tag{17}$$

όπου

$$\sigma_{sZ} := \sum_{j=1}^{k} \sigma_{sj} \tag{18}$$

Η απόδειξη της πρότασης βρίσκεται στο Παράρτημα 2. Με βάση τα παραπάνω ορίζεται η δεύτερη διορθωτική διαδικασία

$$X_s = \widetilde{X}_s + \lambda_s \left(Z - \widetilde{Z} \right) \tag{19}$$

όπου με $\widetilde{Z}$ εννοούμε το άθροισμα όλων των μεταβλητών $\widetilde{X}_s$. Η διαδικασία αυτή θα ονομάζεται *γραμμική διορθωτική διαδικασία*. Συναντάται και σε άλλα επιμεριστικά μοντέλα [π.χ. *Grygier and Stedinger*, 1988· *Lane and Frevert*, 1990, σελ. V-22], μόνο που οι συντελεστές λ_s αντί να ορίζονται με βάση τις συνδιασπορές με τις μεταβλητές υψηλότερου επιπέδου, όπως στη σχέση (17), θεωρούνται ανάλογοι των τυπικών αποκλίσεων των μεταβλητών χαμηλότερου επιπέδου δηλαδή

$$\lambda_s = \sqrt{\sigma_{ss}} \,/ \sum_{j=1}^{k} \sqrt{\sigma_{jj}} \tag{20}$$

Πρέπει να τονίσουμε στο σημείο αυτό ότι οι συντελεστές λ_s, όπως ορίζονται από τη σχέση (17), είναι οι μόνες που μπορούν να διατηρήσουν τον πίνακα διασπορών-συνδιασπορών $\boldsymbol{\sigma}$ (βλέπε και την απόδειξη του Παραρτήματος 2). Η διαφορά των σχέσεων (20) και (17) είναι περισσότερο εμφανής στην περίπτωση που οι μεταβλητές X_s είναι ανεξάρτητες πράγμα που μέσω της (17) οδηγεί στην

$$\lambda_s = \sigma_{ss} \,/ \sum_{j=1}^{k} \sigma_{jj} \tag{21}$$

Η σχέση αυτή φανερώνει ότι οι συντελεστές λ_s κανονικά είναι ανάλογοι των διασπορών και όχι των τυπικών αποκλίσεων.

Οι τελικές παράμετροι της αυτής της διορθωτικής διαδικασίας είναι οι συνδιασπορές σ_{sZ} των μεταβλητών χαμηλότερου επιπέδου με αυτές του υψηλότερου. Το άθροισμά τους για όλες τις υποπεριόδους s ισούται με τη διασπορά σ_{ZZ} και γι' αυτό οι συντελεστές λ_s αθροιζόμενοι δίνουν τη μονάδα και οι μεταβλητές X_s αθροιζόμενες δίνουν τη Z. Με τη χρήση της (18) μπορούμε να υπολογίσουμε αυτές τις συνδιασπορές με βάση τους όρους του πίνακα $\boldsymbol{\sigma}$. Βεβαίως δεν είναι όλοι οι όροι του πίνακα συνδιασπορών παράμετροι του μοντέλου. Για παράδειγμα στο PAR(1) μόνο οι συνδιασπορές διαδοχικών μεταβλητών χαμηλότερου επιπέδου είναι παράμετροι του μοντέλου. Κάθε συνδιασπορά σ_{sj} για $j > s + 1$ μπορεί να υπολογιστεί από τη σχέση

$$\sigma_{sj} = \frac{\sigma_{s,s+1}\,\sigma_{s+1,s+2}\cdots\sigma_{j-1,j}}{\sigma_{s+1,s+1}\cdots\sigma_{j-1,j-1}} \tag{22}$$

που είναι μια συνέπεια του μοντέλου PAR(1). Παρόμοιες, αλλά πολυπλοκότερες, σχέσεις μπορούν να προκύψουν και για άλλα σειριακά μοντέλα.

Η εν λόγω διαδικασία έχει γενική διατύπωση και μπορεί να χρησιμοποιηθεί ανεξαρτήτως της κατανομής που ακολουθούν οι μεταβλητές χαμηλότερου επιπέδου X_s. Διατηρεί και τις μέσες τιμές και τον πίνακα διασπορών-συνδιασπορών για τις μεταβλητές χαμηλότερου επιπέδου. Στην περίπτωση που η κατανομή των μεταβλητών χαμηλότερου επιπέδου είναι κανονική η διαδικασία διατηρεί την πολυμεταβλητή συνάρτηση κατανομής των μεταβλητών χαμηλότερου επιπέδου για προβλήματα που αναφέρονται σε μία θέση. Τονίζουμε ότι η συνθήκη (14) ικανοποιείται πάντα στα μοντέλα επιμερισμού, σε αντίθεση με την (15) που η ισχύς της εξαρτάται από το βαθμό σύγκλισης των συνδιασπορών σ_{sj} με τις πραγματικές.

2.3 Τροποποίηση για θετικές μεταβλητές χαμηλότερου επιπέδου (Διορθωτική διαδικασία δύναμης)

Όπως προκύπτει από τη σχέση (19) η δεύτερη διορθωτική διαδικασία είναι δυνατό να δώσει αρνητικές τιμές στις μεταβλητές χαμηλότερου επιπέδου, πράγμα που δεν είναι επιθυμητό γιατί οι περισσότερες υδρολογικές μεταβλητές είναι θετικές. Η πρώτη διαδικασία δίνει μόνο θετικές μεταβλητές αλλά είναι μειωμένης εφαρμογής λόγω των ισχυρών περιορισμών της. Οι δύο αυτές διαδικασίες, λόγω της διαφορετικής μαθηματικής τους διατύπωσης, δεν είναι συμβιβαστές μεταξύ τους. Η τρίτη διαδικασία που προτείνουμε είναι ένας συνδυασμός των άλλων δύο με αποτέλεσμα να παράγει κατά τη διόρθωση μόνο θετικές τιμές, να ταυτίζεται με την πρώτη αν οι σχετικοί περιορισμοί ικανοποιούνται και τέλος να ταυτίζεται με τη δεύτερη σε κάποια περιοχή και συγκεκριμένα στην γειτονιά των μέσων τιμών δηλαδή γύρω από το σημείο $\left(X_s, Z, \tilde{Z}\right) = \left(\xi_s, \xi_Z, \xi_Z\right)$.

Παρατηρούμε ότι οι διαδικασίες 1 και 2 μπορούν να γραφτούν με ένα κοινό τύπο

$$\frac{X_s}{\tilde{X}_s} = f\left(\frac{Z}{\tilde{X}_s}, \frac{\tilde{Z}}{\tilde{X}_s}\right) \tag{23}$$

όπου

$$f(u,w) = u / w \tag{24}$$

για την πρώτη διαδικασία και

$$f(u,w) = 1 + \lambda_s (u - w) \tag{25}$$

για την δεύτερη.

Κατ' αντιστοιχία για τη νέα διαδικασία θέτουμε

$$f(u,w) = u^{\mu} / w^{\nu} \tag{26}$$

όπου μ και ν είναι παράμετροι προς εκτίμηση. Η εξίσωση αυτή δίνει πάντα θετικές τιμές στις μεταβλητές χαμηλότερου επιπέδου και είναι γενικότερη από την (24). Η συμφωνία της με την (24) επιβάλλει $\mu = \nu = 1$, αν οι μεταβλητές χαμηλότερου επιπέδου είναι ανεξάρτητες και ταυτόχρονα ισχύει

$$\xi_s / \sigma_{ss} = \xi_Z / \sigma_{ZZ} \tag{27}$$

Για να εξετάσουμε τη συμβατότητα των (26) και (25) στην περιοχή γύρω από τις μέσες τιμές παίρνουμε τους τρεις πρώτους όρους του πολυώνυμου Taylor για τη σχέση (26) γύρω από το σημείο $(u, w) = (1/\eta_s, 1/\eta_s)$ όπου $\eta_s = \xi_s/\xi_Z$

$$\begin{aligned} f(u,w) &\approx 1/\eta_s^{\mu-\nu} + \left(\mu/\eta_s^{\mu-\nu-1}\right)\left(u - 1/\eta_s\right) - \left(\nu/\eta_s^{\mu-\nu-1}\right)\left(w - 1/\eta_s\right) \\ &\approx \left(1/\eta_s^{\mu-\nu}\right)\left(1 - \mu + \nu + \mu\eta_s u - \nu\eta_s w\right) \end{aligned} \tag{28}$$

Συγκρίνοντας τη (28) με την (25) καταλήγουμε στο συμπέρασμα

$$\mu = \nu = \lambda_s / \eta_s \tag{29}$$

Αν οι μεταβλητές χαμηλότερου επιπέδου είναι ανεξάρτητες και ικανοποιούν την (27) τότε η (29) γίνεται

$$\mu = \nu = \frac{\lambda_s}{\eta_s} = \frac{\sigma_{sZ}}{\sigma_{ZZ}} \frac{\xi_Z}{\xi_s} = \frac{\sigma_{ss}}{\sigma_{ZZ}} \frac{\xi_Z}{\xi_s} = 1 \tag{30}$$

όπως απαιτούσαμε αρχικά. Συνεπώς

$$f(u,w) = (u/w)^{\lambda_s/\eta_s} = (u/w)^{\frac{\sigma_{sZ}}{\sigma_{ZZ}}\frac{\xi_Z}{\xi_s}} \tag{31}$$

Καταλήγουμε έτσι στην ακόλουθη τελική σχέση, που θα ορίζει τη *διορθωτική διαδικασία δύναμης*:

$$X_s = \widetilde{X}_s\left(Z/\widetilde{Z}\right)^{\lambda_s/\eta_s} = \widetilde{X}_s\left(Z/\widetilde{Z}\right)^{\frac{\sigma_{sZ}}{\sigma_{ZZ}}\frac{\xi_Z}{\xi_s}} \tag{32}$$

Αυτή η διαδικασία δεν ικανοποιεί την αθροιστική ιδιότητα μεμιάς. Έτσι η εφαρμογή της πρέπει να είναι επαναληπτική μέχρις ότου το άθροισμα των μεταβλητών χαμηλότερου επιπέδου να εξισωθεί με τη δεδομένη μεταβλητή υψηλότερου επιπέδου.

Στην περίπτωση όπου οι X_s και Z έχουν κατώτερο όριο c_s και c_Z ($\neq 0$), αντίστοιχα, η (32) μπορεί να τροποποιηθεί στην

$$X_s - c_s = \left(\tilde{X}_s - c_s\right)\left(\frac{Z - c_Z}{\tilde{Z} - c_Z}\right)^{\lambda_s/\eta_s} = \left(\tilde{X}_s - c_s\right)\left(\frac{Z - c_Z}{\tilde{Z} - c_Z}\right)^{\frac{\sigma_{sZ}}{\sigma_{ZZ}}\frac{\xi_Z - c_Z}{\xi_s - c_s}} \tag{33}$$

όπου τώρα ο συντελεστής η_s προσδιορίζεται ως $\eta_s = \left(\xi_s - c_s\right)/\left(\xi_Z - c_Z\right)$. Άν λογαριθμίσουμε την (33) και συμβολίσουμε με $X_s^* = \ln\left(X_s - c_s\right)$, $Z^* = \ln\left(Z - c_Z\right)$, τότε

$$X_s^* = \tilde{X}_s^* + \lambda_s^*\left(Z^* - \tilde{Z}^*\right) \tag{34}$$

με

$$\lambda_s^* := \frac{\lambda_s}{\eta_s} = \frac{\sigma_{sZ}}{\sigma_{ZZ}}\frac{\xi_Z - c_Z}{\xi_s - c_s} \tag{35}$$

Η εξίσωση (34) είναι μια γραμμικοποιημένη μορφή της διόρθωσης δύναμης, που μοιάζει αρκετά με τη γραμμική διορθωτική διαδικασία. Πρέπει να τονίσουμε ότι ο συντελεστής λ_s^* εξαρτάται από τις ροπές (πρώτη και δεύτερη) των αρχικών μεταβλητών και όχι από τις ροπές των μετασχηματισμένων. Παρόλα αυτά αν οι X_s και Z είναι από κοινού λογαριθμοκανονικές μπορεί να δειχτεί με ευκολία ότι

$$\lambda_s^* = \frac{e^{\sigma_{sZ}^*} - 1}{e^{\sigma_{ZZ}^*} - 1} \tag{36}$$

δηλαδή ο συντελεστής λ_s^* εξαρτάται από τη δεύτερη ροπή των μετασχηματισμένων μεταβλητών ($\sigma_{sZ}^* = \mathrm{Cov}\left[X_s^*, Z^*\right]$, $\sigma_Z^{*2} = \mathrm{Var}\left[Z^*\right]$) με ένα τρόπο όμοιο (αλλά όχι ταυτόσημο) με αυτόν της εξίσωσης (17). Βέβαια αυτή είναι μια πολύ σπάνια περίπτωση. Σε μια πιο συνηθισμένη περίπτωση οι διαδοχικές μεταβλητές χαμηλότερου επιπέδου θεωρούνται από κοινού λογαριθμοκανονικές, περίπτωση κατά την οποία οι X_s και Z δεν μπορεί να είναι από κοινού λογαριθμοκανονικές. Στην περίπτωση αυτή ο λ_s^* πρέπει να εκτιμηθεί από την

(35). Για να προσδιορίσουμε τις ροπές των αρχικών μεταβλητών με δεδομένες τις ροπές των μετασχηματισμένων χρησιμοποιούμε τις ακόλουθες εξισώσεις

$$\xi_s - c_s = e^{\xi_s^* + \sigma_{ss}^*/2} \tag{37}$$

$$\sigma_{sj} = e^{\xi_s^* + \xi_j^* + \left(\sigma_{ss}^* + \sigma_{jj}^*\right)/2}\left(e^{\sigma_{sj}^*} - 1\right) \tag{38}$$

(όπου η (38) ισχύει και για $j = s$). Στη συνέχεια για τον καθορισμό του λ_s^*, υπολογίζουμε τα σ_{sZ}, σ_{ZZ} και ξ_Z από τις σχέσεις (18), (15) και (14) αντίστοιχα.

3. Πρακτικά θέματα

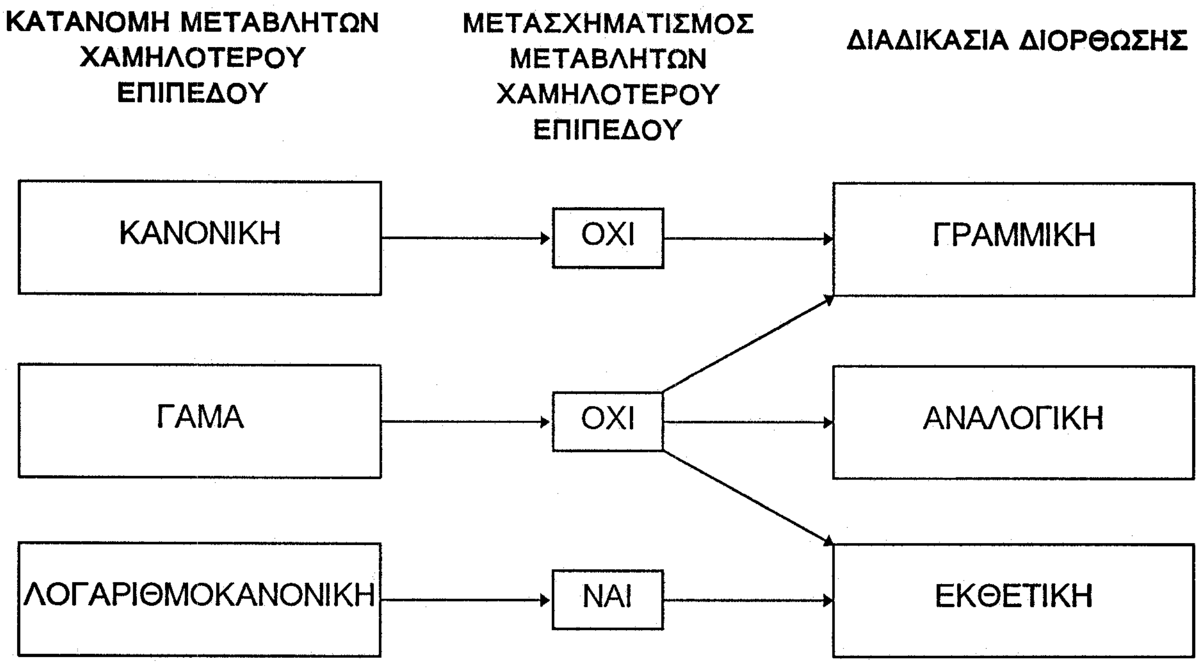
3.1 Επιλογή της διαδικασίας διόρθωσης

Η επιλογή της κατάλληλης διαδικασίας διόρθωσης εξαρτάται από τη συνάρτηση κατανομής των μεταβλητών χαμηλότερου επιπέδου (βλ. Σχήμα 1). Για μεταβλητές με κανονική κατανομή, η πιο κατάλληλη διαδικασία είναι η γραμμική, η οποία διατηρεί τη συνολική πολυμεταβλητή κανονική κατανομή. Αν η παρουσία αρνητικών μεταβλητών δεν είναι επιθυμητή τότε στην περίπτωση εμφάνισής τους είτε επαναλαμβάνεται η διαδικασία γέννησης είτε μηδενίζονται οι τυχόν αρνητικές τιμές. Στην τελευταία περίπτωση το λάθος που επιφέρει ο μηδενισμός εξαλείφεται αν ξαναεφαρμοστεί η διαδικασία διόρθωσης. Πιο συγκεκριμένα, αν επιλεγεί η γραμμική διαδικασία για αυτή τη διόρθωση τότε θα απαιτηθούν μερικές επαναλήψεις γιατί κάθε φορά που εφαρμόζεται δίνει κάποιο αρνητικό αριθμό. Οι επαναλήψεις σταματάνε όταν παύσει η εμφάνιση αρνητικών μεταβλητών ή οι αρνητικές τιμές είναι τόσο μικρές που μπορούν να αμεληθούν. Αντίθετα, αν για αυτή τη διόρθωση επιλεγεί η αναλογική διαδικασία ή η διαδικασία δύναμης, τότε δεν θα απαιτηθούν επαναλήψεις.

Αν οι μεταβλητές ακολουθούν διπαραμετρική ή τριπαραμετρική κατανομή γάμα μπορούν να εφαρμοστούν και οι τρεις διαδικασίες. Για τέτοιου είδους μεταβλητές δεν συνιστάται ο μετασχηματισμός Wilson-Hilferty γιατί δεν είναι ακριβής και εισάγει λάθη στις συναρτήσεις κατανομής. Αν ικανοποιούνται οι περιορισμοί της αναλογικής διαδικασίας (έστω και στην χαλαρωμένη μορφή τους) τότε η εφαρμογή της είναι ο απλούστερος και καλύτερος τρόπος διόρθωσης των μεταβλητών χαμηλότερου επιπέδου. Σε κάθε άλλη περίπτωση μπορούμε να χρησιμοποιήσουμε μια από τις άλλες δύο διαδικασίες. Η γραμμική είναι καλή μόνο στην περίπτωση όπου η πιθανότητα εμφάνισης αρνητικής μεταβλητής είναι μικρή δηλαδή οι συντελεστές μεταβλητότητας όλων των μεταβλητών είναι σχετικά μεγάλοι. Η εκθετική μπορεί να εφαρμοστεί σε οποιαδήποτε περίπτωση μια και δεν έχει περιορισμούς. Τα μειονεκτήματα της είναι ότι δεν είναι ακριβής με την αυστηρή σημασία του όρου και ότι απαιτεί επαναλήψεις για να ικανοποιήσει την αθροιστική ιδιότητα, πράγμα που την κάνει αργή.

Για λογαριθμοκανονικές μεταβλητές θα μπορούσαμε να χρησιμοποιήσουμε τις ίδιες διαδικασίες όπως στην κατανομή γάμα δηλαδή τη γραμμική ή την εκθετική. Παρ' όλα αυτά είναι πιο λογικό να γίνει η παραγωγή μετασχηματισμένων μεταβλητών και να χρησιμοποιηθεί για τη διόρθωση τους η γραμμικοποιημένη διαδικασία δύναμης της εξίσωσης (34).

Τέλος, για οποιοδήποτε άλλο τύπο κατανομής μπορούμε να προχωρήσουμε όπως στην πιο κοντινή από τις πιο πάνω περιπτώσεις, εμπιστευόμενοι την γέννηση τυχαίων μεταβλητών της συγκεκριμένης κατανομής στο μοντέλο PAR(1).



Σχήμα 1. Πεδίο εφαρμογής των διαδικασιών διόρθωσης για τις πιο συνηθισμένες περιθώριες συναρτήσεις κατανομής των μεταβλητών χαμηλότερου επιπέδου.

3.2 Κριτήρια σύγκλισης και επαναληπτική διαδικασία

Μέχρι τώρα δεν ερευνήσαμε την επίδραση της διορθωτικής διαδικασίας στην ασυμμετρία των μεταβλητών χαμηλότερου επιπέδου, ούτε την επιρροή της στους συντελεστές ετεροσυσχέτισης ανάμεσα στις μεταβλητές διαφορετικών θέσεων. Ο λόγος ήταν ότι τα παραπάνω στατιστικά μεγέθη δε διατηρούνται μετά από τη διορθωτική παρέμβαση εκτός από ορισμένες περιπτώσεις. Για παράδειγμα η αναλογική διαδικασία διατηρεί τους συντελεστές ασυμμετρίας, (όπως και τη συνολική συνάρτηση κατανομής) για τη διπαραμετρική κατανομή γάμα, αν ικανοποιούνται οι συγκεκριμένες περιοριστικές συνθήκες. Επίσης η γραμμική διαδικασία μπορεί να διατηρήσει μηδενικές ασυμμετρίες. Μπορεί να δειχτεί ότι οι συντελεστές ετεροσυσχέτισης διατηρούνται στις παρακάτω ακραίες περιπτώσεις: (α) οι μεταβλητές διαφορετικών θέσεων είναι ανεξάρτητες μεταξύ τους (μηδενικός συντελεστής ετεροσυσχέτισης), (β) οι μεταβλητές διαφορετικών θέσεων είναι πλήρως εξαρτημένες (μοναδιαίος συντελεστής ετεροσυσχέτισης). Στη γενική περίπτωση πάντως οι διορθωτικές διαδικασίες δίνουν μια προσεγγιστική τιμή των παραπάνω στατιστικών μεγεθών, συνήθως μικρότερη της πραγματικής, που μπορεί να μην είναι επαρκής. Παρ' όλα αυτά το σφάλμα μπορεί να μειωθεί με τη μέθοδο της επανάληψης. Αυτή προσπαθεί να βρει την ανεξάρτητη εκείνη γέννηση, μεταβλητών χαμηλότερου επιπέδου, που πλησιάζει περισσότερο στην πραγματική και ήδη γνωστή.

Στην εργασία αυτή υιοθετήσαμε τον απλούστερο τρόπο επαναληπτικής διαδικασίας όπως αυτός περιγράφεται στην ενότητα 2. Ο τρόπος αυτός αποδείχτηκε πολύ αποτελεσματικός όπως θα δειχτεί στην ενότητα 5, όπου δίνουμε κάποια στοιχεία για την απόδοσή του. Για κάθε περίοδο παράγονται ακαθόριστες στο πλήθος εναλλακτικές τιμές των μεταβλητών χαμηλότερου επιπέδου, μέσω του μοντέλου PAR(1), μέχρις ότου η απόσταση $\Delta Z := \left\| \mathbf{Z} - \tilde{\mathbf{Z}} \right\|$ να γίνει μικρότερη από κάποιο προκαθορισμένο όριο. Το όριο αυτό καθορίζεται από την επιθυμητή ακρίβεια. Όσο πιο μικρό τόσο μεγαλύτερη είναι η ακρίβεια και τόσο μεγαλύτερος ο αριθμός των επαναλήψεων. Είναι σκόπιμο να τεθεί ένα ανώτατο όριο επαναλήψεων γιατί κάποιες περίοδοι μπορεί να συγκλίνουν μετά από πάρα πολλές επαναλήψεις.

Η απόσταση ΔZ μπορεί να οριστεί ως εξής:

$$\Delta Z = \sum_{l=1}^{n} \left(Z^l - \tilde{Z}^l \right)^2 \tag{39}$$

Πάντως είναι προτιμότερο να ορίσουμε την απόσταση με ένα αδιάστατο τύπο ανεξάρτητο από το πλήθος των θέσεων του κάθε προβλήματος όπως:

$$\Delta Z = \frac{1}{n} \sum_{l=1}^{n} \frac{\left| Z^l - \tilde{Z}^l \right|}{\sqrt{\sigma_{ZZ}^l}} \tag{40}$$

Ο παραπάνω τύπος χρησιμοποιείται στο μοντέλο μας.

Άλλο ένα πρόβλημα, που συναντάται στα περισσότερα μοντέλα επιμερισμού, είναι η διατήρηση του συντελεστή συσχέτισης της πρώτης με την τελευταία υποπερίοδο κάποιας περιόδου. Το πρόβλημα αυτό πρώτοι εντοπίσανε οι *Stedinger and Vogel* [1984], ενώ οι *Lin* [1990] και *Koutsoyiannis* [1992] έκαναν προσπάθειες να το επιλύσουν. Στο παρόν επιμεριστικό μοντέλο δε χρησιμοποιήθηκαν στατιστικά μεγέθη που αφορούν τις συσχετίσεις μεταξύ μεταβλητών χαμηλότερου και υψηλότερου επιπέδου, τα οποία είναι υπεύθυνα για το συγκεκριμένο πρόβλημα [*Stedinger and Vogel*, 1984]. Παρόλα αυτά είναι και πάλι αναμενόμενη η εμφάνισή του. Η κατάσταση σε σχέση με το παρόν μοντέλο έχει ως εξής: όταν ξεκινάει η παραγωγή μεταβλητών για την περίοδο p είναι ήδη γνωστές όλες οι μεταβλητές χαμηλότερου επιπέδου της προηγούμενης περιόδου $p - 1$. Στη συνέχεια για την παραγωγή της πρώτης υποπεριόδου της περιόδου p, με τη χρήση PAR(1), χρησιμοποιείται η τελευταία υποπερίοδος k της προηγούμενης περιόδου $p - 1$. Η διαδικασία αυτή διατηρεί το συντελεστή συσχέτισης μεταξύ των συγκεκριμένων υποπεριόδων, αλλά όταν επεμβεί η διορθωτική διαδικασία τροποποιεί μόνο τη μεταβλητή της περιόδου p εισάγοντας με αυτό τον τρόπο μεροληψία στη συσχέτισή της με την τελευταία μεταβλητή της περιόδου $p - 1$.

Η εισαγόμενη αυτή μεροληψία μπορεί να επηρεάσει και τις γειτονικές, της πρώτης, υποπεριόδους σε μικρότερο βέβαια βαθμό. Η ελαχιστοποίηση της απόστασης ΔZ μέσω επαναλήψεων είναι ένας απλός τρόπος αντιμετώπισης και αυτού του προβλήματος παρόλο που θα μπορούσαν να εφευρεθούν και πιο πολύπλοκες διαδικασίες αντιμετώπισης. Στην ενότητα 5 γίνεται φανερό ότι το συγκεκριμένο πρόβλημα είναι το δυσκολότερο να αντιμετωπιστεί από όλα όσα αναφέραμε.

3.3 Προβλήματα σχετικά με το PAR(1)

Υπάρχουν δύο πολύ γνωστά προβλήματα που συσχετίζονται με το μοντέλο PAR(1) (όπως και με κάθε άλλο γραμμικό μοντέλο). Το πρώτο σχετίζεται με τη διάσπαση του πίνακα $\mathbf{d}_s = \mathbf{b}_s\mathbf{b}_s^T$, και το δεύτερο με τους μεγάλους συντελεστές ασυμμετρίας που απαιτεί το μοντέλο για το τυχαίο διάνυσμα $\mathbf{V}_s$. Και τα δύο αυτά προβλήματα αντιμετωπίστηκαν με την αποδυνάμωση των μη διαγώνιων στοιχείων του πίνακα $\boldsymbol{\sigma}_s$ με μια μέθοδο παρόμοια με αυτή των *Mejía and Millán* [1974], η οποία αναφέρεται και από τους *Bras and Rodriguez-Iturbre* [1985, σελ. 98]. Η μέθοδος αυτή είναι κατάλληλη για τη διάσπαση όπως και για το πρόβλημα των ασυμμετριών γιατί αυτές είναι ανάλογες με τον αντίστροφο πίνακα $\mathbf{b}_s^{(3)}$ (βλέπε εξίσωση (7)).

Προτού περιγραφεί η μέθοδος αυτή πρέπει να ποσοτικοποιηθούν οι απαιτήσεις του προβλήματος. Για την πρώτη περίπτωση, είναι γνωστό, ότι ο $\mathbf{d}_s$ πρέπει να είναι θετικά ορισμένος. Για τη δεύτερη πρέπει να τεθεί ένα ανώτατο όριο στις ασυμμετρίες που μπορεί να δώσει μια γεννήτρια τυχαίων αριθμών. Για το σκοπό αυτό θα ανακαλέσουμε ένα τύπο των [*Wallis et al*., 1974, *Kirby*, 1974] που καθορίζει τη μέγιστη ασυμμετρία ενός συνθετικού δείγματος ανάλογα με το πλήθος N των στοιχείων του:

$$C_{s_{\max}} = \frac{N-2}{\sqrt{N-1}} \approx \sqrt{N} \tag{41}$$

Έτσι, είναι σκόπιμο να τεθεί ένα ανώτατο όριο σύμφωνα με τον τύπο:

$$C_{s_{\text{acc}}} = \varepsilon\, C_{s_{\max}} \tag{42}$$

όπου $0 < \varepsilon < 1$ (π.χ. $\varepsilon = 0.5 \div 0.75$). Το μειονέκτημα αυτού του ορίου είναι η εξάρτηση του από το μέγεθος N του δείγματος. Παρ' όλα αυτά οποιαδήποτε άλλη διατύπωση θα ήταν αυθαίρετη.

Μπορούμε τώρα να παραθέσουμε τον αλγόριθμο της μεθόδου:

1. Γίνεται ο καθορισμός του άνω ορίου με βάση το μέγεθος N του δείγματος
2. Επιλέγεται ένας συντελεστής ϕ κοντά στη μονάδα, πάντα μικρότερος απ' αυτή (π.χ. 0.99)
3. Υπολογίζονται οι ιδιοτιμές του πίνακα $\mathbf{d}_s$. Αν βρεθεί έστω και μια αρνητική δηλαδή ο πίνακας $\mathbf{d}_s$ δεν είναι θετικά ορισμένος) ακολουθείται η διαδικασία του βήματος 4 αλλιώς ακολουθεί το βήμα 5.
4. Όλα τα εκτός διαγωνίου στοιχεία του πίνακα $\mathbf{d}_s$ πολλαπλασιάζονται με ϕ και ξαναεκτελείται το βήμα 3.
5. Υπολογίζεται ο πίνακας $\mathbf{b}_s$ με τη μέθοδο των ιδιοτιμών.
6. Υπολογίζονται οι ασυμμετρίες του τυχαίου διανύσματος με τη βοήθεια του τύπου (7). Αν κάποια απ' αυτές είναι μεγαλύτερη (απολύτως) από το επιτρεπόμενο όριο ακολουθεί το βήμα 4 αλλιώς ακολουθεί το βήμα 7.
7. Τέλος.

Η διάσπαση με τη μέθοδο των ιδιοτιμών προτιμήθηκε γιατί στις περισσότερες περιπτώσεις δίνει χαμηλότερους συντελεστές ασυμμετρίας από τη μέθοδο της διάσπασης σε τριγωνικό πίνακα. Τονίζουμε ότι η παραπάνω διαδικασία δε μεταβάλλει τη διασπορά ή την ασυμμετρία των μεταβλητών χαμηλότερου επιπέδου. Αν όμως εφαρμοστεί, έστω και μια φορά, η διαδικασία του βήματος 4 μειώνονται οι συντελεστές συσχέτισης μηδενικού βήματος μεταξύ μεταβλητών διαφορετικών θέσεων, πράγμα που είναι αναπόφευκτο. Στις εφαρμογές που πραγματοποιήθηκαν για να δοκιμαστεί το μοντέλο, φαινόμενα όπως αυτά που περιγράφουμε παραπάνω εμφανίστηκαν πολύ σπάνια, αλλά και όταν εμφανίστηκαν σε δύο με τέσσερις επαναλήψεις της μεθόδου επιλύθηκαν, επιφέροντας ένα σφάλμα της τάξης του 2-4% στους συντελεστές ετεροσυσχέτισης. Οι άλλες μέθοδοι που χρησιμοποιήθηκαν, όπως αυτή των *Grygier and Stedinger* [1990, σελ. 31-33] και *Koutsoyiannis* [1992], επέφεραν μεγαλύτερα σφάλματα στους πίνακες συνδιασποράς.

3.4 Γεννήτρια τυχαίων αριθμών

Το μοντέλο περιλαμβάνει διάφορες ακριβείς γεννήτριες τυχαίων αριθμών, αποφεύγοντας να χρησιμοποιήσει γεννήτριες προσεγγιστικού τύπου (όπως πχ. αυτή που βασίζεται στο μετασχηματισμό Wilson-Hilferty). Έτσι, για την κανονική (και τη λογαριθμοκανονική) κατανομή γίνεται χρήση του ακριβούς αλγόριθμου Πολικών Συντεταγμένων [*Papoulis*, 1990, p. 266]. Για την κατανομή γάμα χρησιμοποιήθηκε ο επίσης ακριβής αλγόριθμος του *Whittaker* [1973]. Στο Παράρτημα 2 περιγράφεται μια καινούρια ακριβής μέθοδος παραγωγής τυχαίων αριθμών που αναπτύξαμε, η οποία συμπεριφέρεται πολύ ικανοποιητικά στις περιπτώσεις μεγάλων ασυμμετριών. Πρέπει να τονιστεί ότι και οι δύο αλγόριθμοι κατανομής γάμα είναι ακριβείς μόνο για τις μεταβλητές που παράγουν απευθείας (για την περίπτωσή μας τις V_s). Αυτό σημαίνει ότι η απόλυτη ακρίβεια χάνεται όταν προστίθενται διάφορες μεταβλητές για να προκύψουν τα $\mathbf{X}_s$. Έτσι το μοντέλο καταφέρνει να διατηρήσει τις τρεις πρώτες ροπές και ταυτόχρονα να δώσει μια ικανοποιητική προσέγγιση της συνάρτησης κατανομής (βλ. σχήμα 5). Η ακριβής διατήρηση της κατανομής των $\mathbf{X}_s$ απαιτεί τη χρήση μιας άλλης κατηγορίας μοντέλων όπως τα μοντέλα αυτοπαλινδρόμησης γάμα (GAR) ή τα περιοδικά μοντέλα αυτοπαλινδρόμησης γάμα (PAGAR, PMGAR). Τέτοια μοντέλα έχουν διερευνηθεί θεωρητικά για στάσιμες [*Lawrance and Lewis*, 1981] και περιοδικές σειρές [*Fernandez and Salas*, 1986]. Έχουν εφαρμοστεί δε για την προσομοίωση παροχής υδατορευμάτων [π.χ. *Fernandez and Salas*, 1990]. Παρ' όλα αυτά η εφαρμογή τους εξαντλείται σε μονοδιάστατα προβλήματα και για το λόγο αυτό δεν ενσωματώθηκαν στο δικό μας μοντέλο, που είναι προσανατολισμένο σε πολυδιάστατες εφαρμογές.

4. Εφαρμογές

Στο κεφάλαιο αυτό παρουσιάζονται τρεις εφαρμογές του μοντέλου με σκοπό να δείξουν την ικανότητά του στη διατήρηση συγκεκριμένων στατιστικών χαρακτηριστικών. Η πρώτη εφαρμογή ασχολείται με την προσομοίωση μηνιαίων βροχοπτώσεων σε τρεις βροχομετρικούς σταθμούς που βρίσκονται στη λεκάνη της Υλίκης και συγκεκριμένα στις θέσεις Αλίαρτος, Μουρίκι και Καλλιθέα. Η μέση μηνιαία βροχόπτωση στους σταθμούς αυτούς κυμαίνεται από 420 mm. μέχρι 610 mm, η κατανομή των βροχών έχει μέγιστο το μήνα Δεκέμβρη με διακύμανση υψών από 70 έως 100 mm. και ελάχιστο τον Ιούλιο ή Αύγουστο με 5 mm. βροχής περίπου. Το υδρολογικό έτος για την περιοχή ξεκινάει την πρώτη Οκτωβρίου (έτσι ο μήνας 1 στο Σχήμα 2 είναι ο Οκτώβριος). Το καλοκαίρι οι περιθώριες κατανομές των βροχοπτώσεων είναι αρκετά ασύμμετρες (Σχήμα 2b). Ο συντελεστής συσχέτισης μεταξύ των θέσεων φτάνει το 0.80 τις φθινοπωρινές και ανοιξιάτικες περιόδους (Σχήμα 2d) ενώ οι αυτοσυσχετίσεις μοναδιαίου βήματος είναι χαμηλές, όπως εξ άλλου αναμενόταν για μηνιαίες βροχοπτώσεις.

Για την εφαρμογή αυτή υιοθετήθηκε η κατανομή γάμα και για τα μηνιαία αλλά και για τα ετήσια ύψη. Η παραδοχή αυτή επιβεβαιώθηκε και από τα ιστορικά δεδομένα. (βλ. Σχήμα 5). Έγινε παραγωγή 8000 περιόδων συνθετικών δεδομένων βροχόπτωσης ετήσιου βήματος και στη συνέχεια έγινε ο επιμερισμός τους σε μηνιαία. Οι παραδοχές της μεθόδου επιμερισμού συνοψίζονται στον Πίνακα 1. Για να εξερευνήσουμε τη συμπεριφορά του μοντέλου υπό διαφορετικές συνθήκες κάναμε μια εφαρμογή χωρίς επανάληψη (περίπτωση 1.1) και μια με επανάληψη (περίπτωση 1.2). Επιπροσθέτως, και για λόγους σύγκρισης, έγινε και μια εφαρμογή PAR(1), χωρίς επιμερισμό (περίπτωση 1.3). Τα στατιστικά μεγέθη των συνθετικών δειγμάτων και των ιστορικών δεδομένων φαίνονται στο Σχήμα 2. Παρατηρούμε ότι το μοντέλο, ακόμη και χωρίς επαναλήψεις, πλησιάζει τα στατιστικά μεγέθη που ενδιαφέρουν. Η απόδοσή του όσον αφορά τους συντελεστές συσχέτισης και τις ασυμμετρίες βελτιώνεται ακόμη περισσότερο με τη χρήση επαναλήψεων. Σε γενικές γραμμές πάντως τα αποτελέσματα που προκύπτουν είναι απολύτως συγκρίσιμα με αυτά του PAR(1) (χωρίς επιμερισμό), που είναι ένα καθιερωμένο μοντέλο για παρόμοιες προσομοιώσεις. Στο Σχήμα 5 γίνεται γραφική σύγκριση της εμπειρικής συνάρτησης κατανομής των συνθετικών δεδομένων, για μια θέση και μια υποπερίοδο, με την αντίστοιχη των ιστορικών και τη συνάρτηση κατανομής γάμα. Τα αποτελέσματα είναι ικανοποιητικά.

Η δεύτερη εφαρμογή προσομοιώνει τις μηνιαίες παροχές του Νείλου στην περιοχή του Aswan (μονοδιάστατη προσομοίωση). Σε μια παρόμοια εργασία ο Todini [1980] μελέτησε το πρόβλημα χρησιμοποιώντας 105 χρόνια ιστορικών δεδομένων. Στόχος του ήταν η διατήρηση των μηνιαίων ασυμμετριών μέσω ενός μοντέλου επιμερισμού. Από τους πίνακες και τα διαγράμματα αυτής της εργασίας αντλήσαμε τις στατιστικές παραμέτρους των ιστορικών δεδομένων που ήταν απαραίτητες για την προσομοίωση που κάναμε. Η μέση μηνιαία παροχή έχει μέγιστο τα 23000 m^3 το μήνα Σεπτέμβριο και ελάχιστο το 1600 m^3 το μήνα Μάιο. Το υδρολογικό έτος ξεκινάει τον Ιούλιο (έτσι ο μήνας 1 στα Σχήματα 3 και 4 είναι ο Ιούλιος). Τους μήνες χαμηλών παροχών οι συντελεστές ασυμμετρίας είναι πολύ μεγάλοι ξεπερνώντας την τιμή 2.0 για το μήνα Απρίλιο και Μάιο (Σχήμα 3b). Οι μηνιαίοι συντελεστές αυτοσυσχέτισης βήματος ένα είναι επίσης μεγάλοι φτάνοντας την τιμή 0.939 για τους μήνες Φεβρουάριο και Μάρτιο (Σχήμα 2c). Όπως και στην προηγούμενη εφαρμογή υιοθετήθηκε η κατανομή γάμα και η γραμμική διαδικασία διόρθωσης (βλ. Πίνακα 1). Παρήχθησαν 10000 περίοδοι συνθετικών ετήσιων δεδομένων και στη συνέχεια οι αντίστοιχες μηνιαίες σειρές με τρεις τρόπους: πρώτα με επιμερισμό χωρίς επανάληψη (περίπτωση 2.1), μετά με επιμερισμό και επανάληψη (περίπτωση 2.2) και τέλος χωρίς επιμερισμό (περίπτωση 2.3, για λόγους σύγκρισης). Η απόδοση του μοντέλου ήταν εξίσου ικανοποιητική με αυτή της πρώτης εφαρμογής. Η επανάληψη βελτίωσε την απόδοσή του όσον αφορά τη διατήρηση των ασυμμετριών (Σχήμα 3b). Επίσης ελάττωσε την απόσταση του πρώτου συντελεστή αυτοσυσχέτισης των συνθετικών από τα ιστορικά δεδομένα (Σχήμα 3c, μήνας 1). Παρόλα αυτά η απόσταση είναι ακόμα εμφανής (Σχήμα 3c) ενώ από την τρίτη υποπερίοδο και μετά οι δύο γραμμές ταυτίζονται. Τονίζεται ότι η συγκεκριμένη περίπτωση είναι η δυσμενέστερη εξαιτίας του συντελεστή αυτοσυσχέτισης με τιμή $\approx$0.5 ο οποίος διατηρείται με μεγάλη δυσκολία [*Koutsoyiannis*, 1992, Fig. 4]. Αν η τιμή του ήταν κοντά στο μηδέν ή τη μονάδα το σφάλμα θα ήταν σημαντικά μικρότερο. Παρατηρούμε πάντως ότι το σφάλμα στο Σχήμα 3c μπορεί να μειωθεί περισσότερο αν ελαττωθεί η μέγιστη επιτρεπόμενη απόσταση ΔZ.

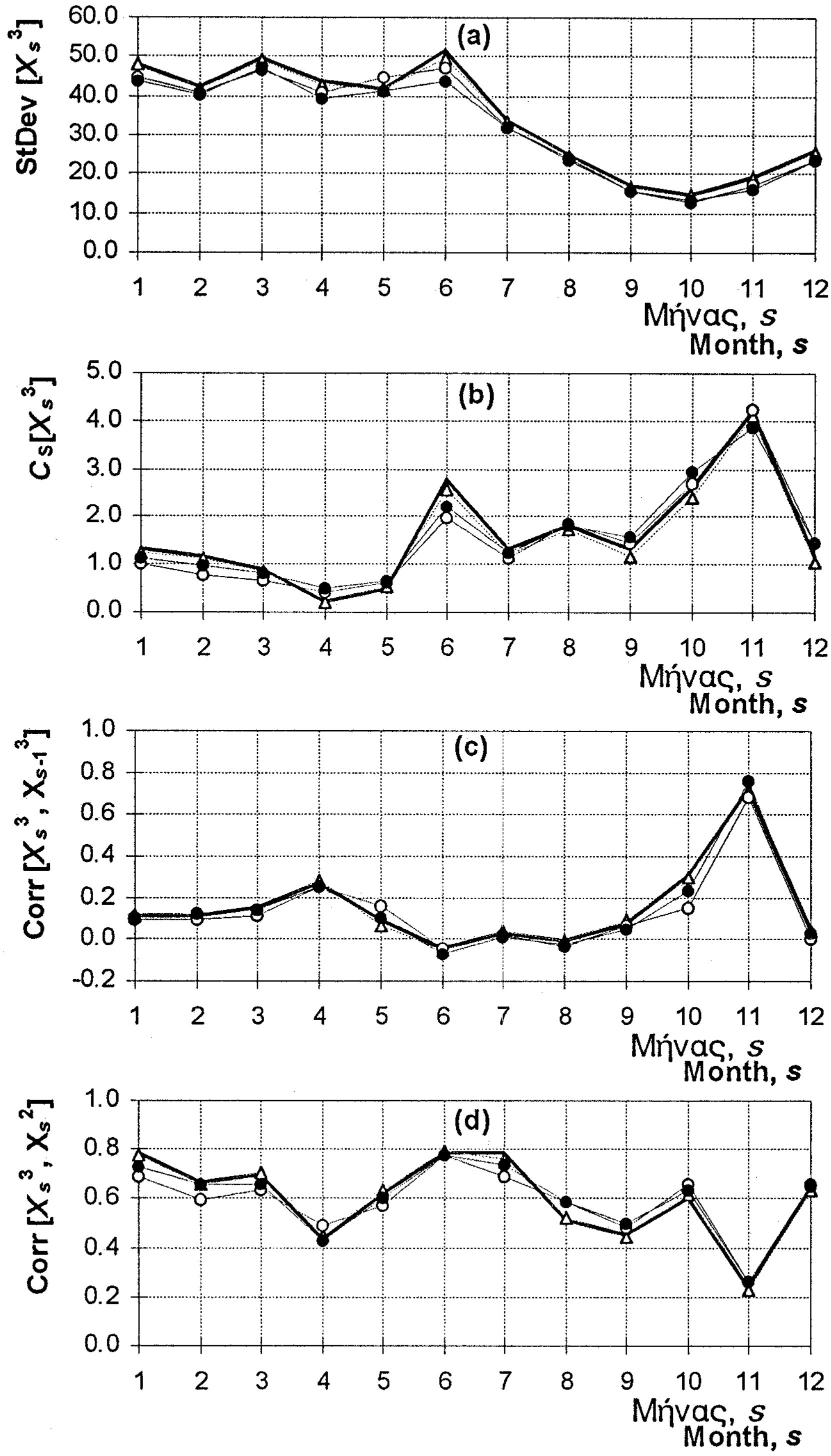
Για να εξερευνήσουμε την απόδοση του μοντέλου κάτω από λογαριθμικό μετασχηματισμό των μεταβλητών και με τη χρήση της διαδικασίας διόρθωσης δύναμης κάναμε μια ακόμη εφαρμογή (εφαρμογή 3 στον πίνακα 1) πάλι για το Νείλο. Η περίπτωση είναι παρόμοια με αυτή της εφαρμογής 2 μόνο που υιοθετείται η λογαριθμοκανονική κατανομή για τις μεταβλητές χαμηλότερου επιπέδου, γίνεται μετασχηματισμός των μεταβλητών με τη σχέση (8) και εφαρμόζεται η διαδικασία διόρθωσης δύναμης. Η εφαρμογή αυτή έγινε περισσότερο για τη δοκιμή του μοντέλου κάτω από ειδικές συνθήκες παρά για πραγματική προσομοίωση. Τα απαραίτητα στατιστικά μεγέθη για την πραγματοποίηση της εφαρμογής προέκυψαν με τη μέθοδο των ροπών γιατί δεν ήταν διαθέσιμα τα ιστορικά δεδομένα. Οι περίοδοι 2 και 3 είχαν αρνητικούς συντελεστές ασυμμετρίας, πολύ κοντά στο μηδέν, και εξισώθηκαν με μια μικρή θετική τιμή (0.01). Στο Σχήμα 4 παρουσιάζονται τα στατιστικά χαρακτηριστικά των λογαριθμισμένων χρονοσειρών (8000 περίοδοι) σε σύγκριση με τα ιστορικά. Είναι προφανές ότι το μοντέλο χωρίς τη βοήθεια της επανάληψης δε συμπεριφέρεται καλά, κάτι που δε συνέβαινε στις προηγούμενες εφαρμογές. Η εξήγηση έχει να κάνει με την διαδικασία διόρθωσης δύναμης, η οποία, όπως προαναφέρθηκε, είναι ακριβής μόνο κάτω από συγκεκριμένους περιορισμούς. Παρόλα αυτά η κατάσταση βελτιώθηκε δραστικά με τις επαναλήψεις φτάνοντας στην απόδοση τις προηγούμενες εφαρμογές.

Πίνακας 1. Χαρακτηριστικά στοιχεία των δοκιμαστικών εφαρμογών του μοντέλου

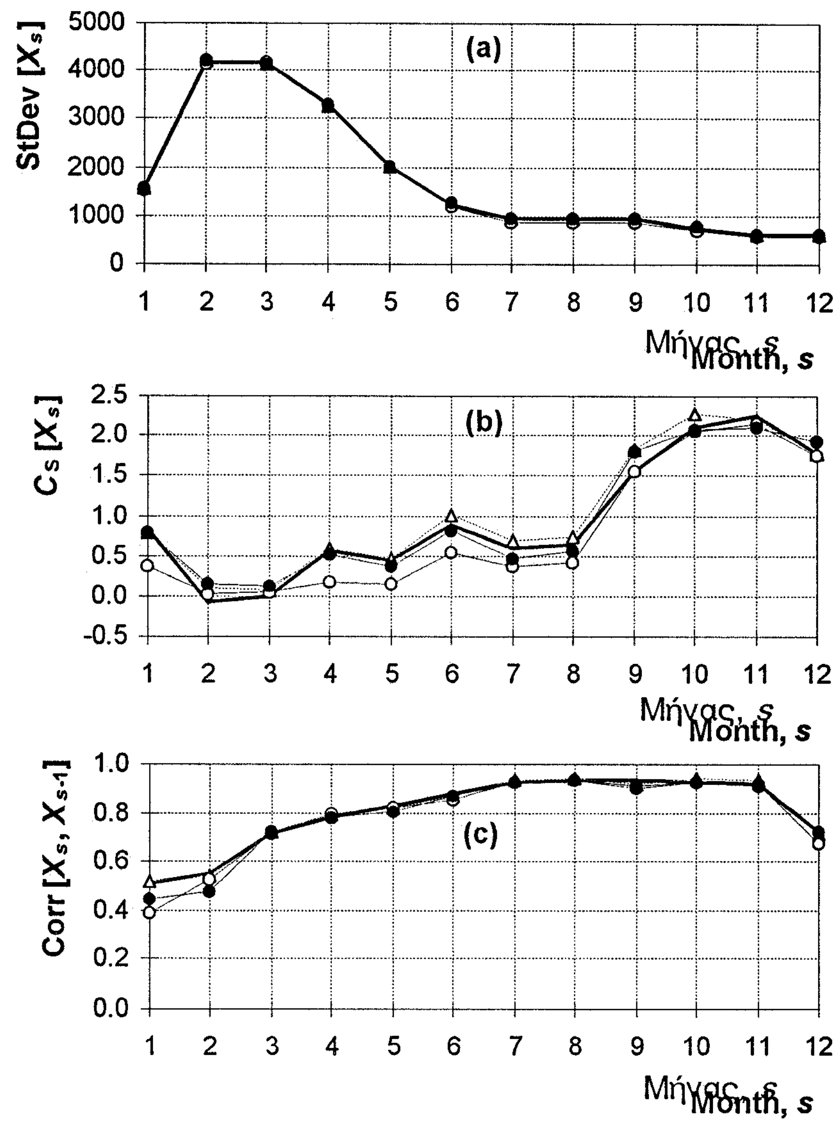
Περίπτωση	Περίοδοι (έτη)	Περιθώρια κατανομή	Μετασχηματισμός	Διορθωτική διαδικασία	Διορθωτική διαδικασία αρνητικών τιμών	Μέγιστη απόσταση ΔZ [§]	Μέγιστες επιτρεπόμενες επαναλ.
Εφαρμογή 1: Δεδομένα βροχής από τρεις σταθμούς στη λεκάνη Υλίκης. Δείγμα 38 ετών							
1.1	8 000	Γάμα	Όχι	Γραμμική	Γραμμική	–	1
1.2	8 000	Γάμα	Όχι	Γραμμική	Γραμμική	0.4	1000
1.3	8 000	Γάμα	Όχι	–	Δεν εφαρμόστηκε	–	–
Εφαρμογή 2: Δεδομένα απορροής στο φράγμα Aswan στο Νείλο. Ιστορικά δεδομένα Todini [1980].							
2.1	10 000	Γάμα	Όχι	Γραμμική	Γραμμική	–	1
2.2	10 000	Γάμα	Όχι	Γραμμική	Γραμμική	0.01	3000
2.3	10 000	Γάμα	Όχι	–	Δεν εφαρμόστηκε	–	–
Εφαρμογή 3: Δεδομένα απορροής στο φράγμα Aswan στο Νείλο, βλ. Εφαρμογή 2.							
3.1	8 000	Λογαριθμοκανονική	Λογαριθμικός	Δύναμης	Όχι απαραίτητη	–	1
3.2	8 000	Λογαριθμοκανονική	Λογαριθμικός	Δύναμης	Όχι απαραίτητη	0.0007	3000

[§] Στη συγκεκριμένη έκδοση του προγράμματος η διαφορά ΔZ υπολογίστηκε από τον τύπο $\Delta Z = \sum_{l=1}^{n} |Z^l - \tilde{Z}^l| \,/\, Z^l$ αντί του πιο κατάλληλου που δίνεται από τη σχέση (40). Η διαφορετική τάξη μεγέθους των ΔZ στις διάφορες εφαρμογές οφείλεται στις διαφορετικές τυπικές αποκλίσεις των μεταβλητών υψηλότερου επιπέδου.

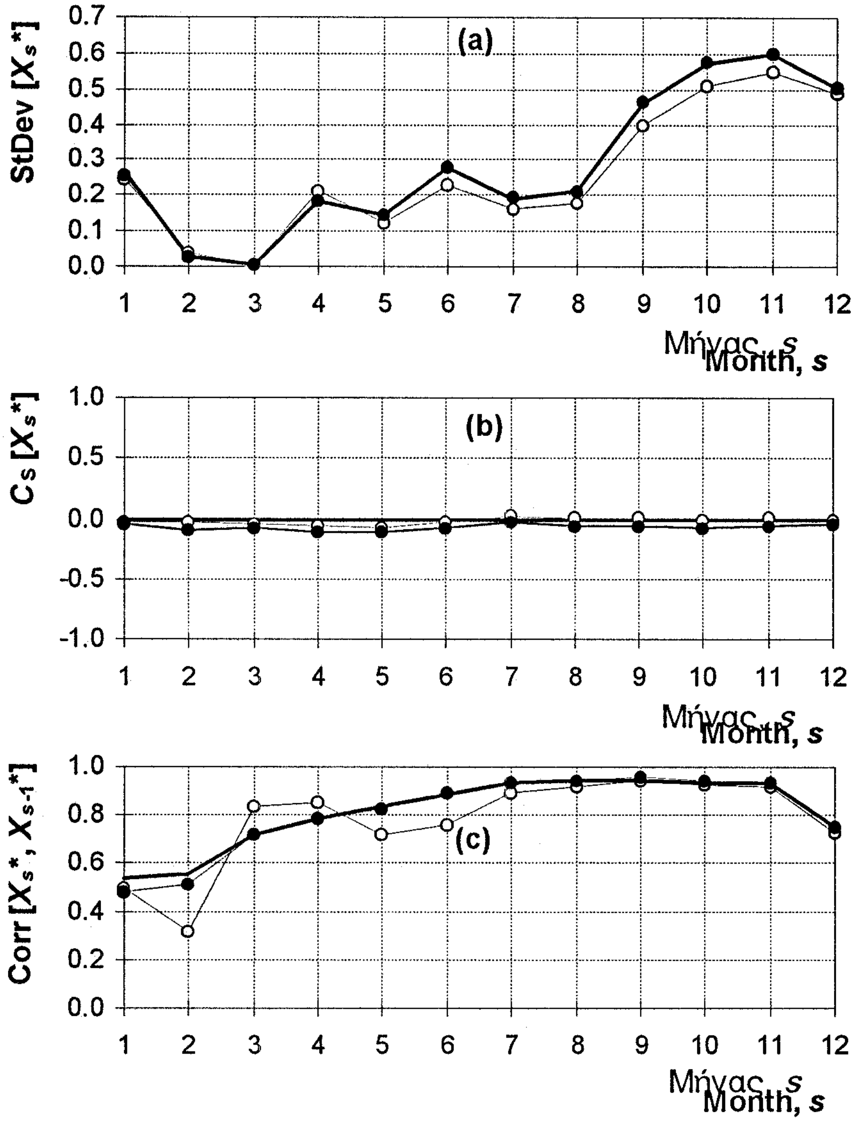
Τέλος παραπέμπουμε στην εργασία του *Koutsoyiannis* [1994] για μια ειδική εφαρμογή της αναλογικής διορθωτικής διαδικασίας. Αυτή η εφαρμογή ασχολείται με μικρής κλίμακας μονοδιάστατο επιμερισμό βροχοπτώσεων όπου οι μεταβλητές χαμηλότερου επιπέδου ακολουθούν την κατανομή γάμα σχήματος J.



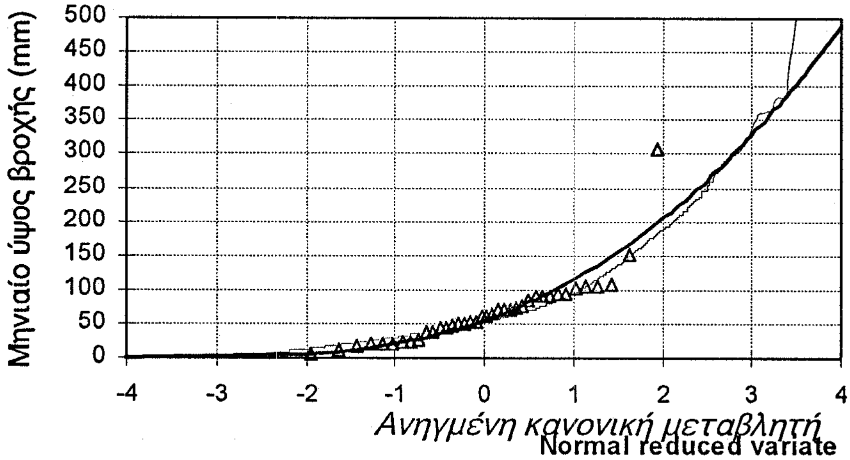
Σχήμα 2. Σύγκριση των στατιστικών χαρακτηριστικών ιστορικών και συνθετικών δεδομένων μηνιαίας βροχόπτωσης στη λεκάνη της Υλίκης (Εφαρμογή 1, 8000 περίοδοι συνθετικών δεδομένων): (a) τυπικές αποκλίσεις στη θέση 3, (b) συντελεστές ασυμμετρίας στη θέση 3, (c) αυτοσυσχετίσεις βήματος 1 στη θέση 3 και (d) ετεροσυσχετίσεις βήματος 0 μεταξύ των θέσεων 2 και 3. Σε όλα τα διαγράμματα η παχιά γραμμή αντιστοιχεί στα στατιστικά μεγέθη των ιστορικών δεδομένων, η λεπτή γραμμή με τους άσπρους κύκλους αντιστοιχεί στα στατιστικά μεγέθη επιμερισμού χωρίς επανάληψη (Περίπτωση 1.1), η λεπτή γραμμή με τους μαύρους κύκλους αντιστοιχεί στα συνθετικά δεδομένα επιμερισμού με επανάληψη (Περίπτωση 1.2) και τέλος η εστιγμένη γραμμή με τα τρίγωνα αντιστοιχεί σε συνθετικά δεδομένα χωρίς επιμερισμό (Περίπτωση 1.3).



Σχήμα 3. Σύγκριση των στατιστικών χαρακτηριστικών ιστορικών και συνθετικών δεδομένων μηνιαίας παροχής στο Νείλο (Εφαρμογή 2, 10000 περίοδοι συνθετικών δεδομένων): (a) τυπικές αποκλίσεις, (b) συντελεστές ασυμμετρίας, (c) αυτοσυσχετίσεις βήματος 1. Σε όλα τα διαγράμματα η παχιά μαύρη γραμμή αντιστοιχεί στα στατιστικά μεγέθη των ιστορικών δεδομένων, η λεπτή γραμμή με τους άσπρους κύκλους αντιστοιχεί στα στατιστικά μεγέθη που προέκυψαν από επιμερισμό χωρίς επανάληψη (Περίπτωση 2.1), η λεπτή γραμμή με τους μαύρους κύκλους αντιστοιχεί στα συνθετικά δεδομένα επιμερισμού με επανάληψη (Περίπτωση 2.2) και τέλος η εστιγμένη γραμμή με τα τρίγωνα αντιστοιχεί σε συνθετικά δεδομένα χωρίς επιμερισμό (Περίπτωση 2.3).



Σχήμα 4. Σύγκριση των στατιστικών χαρακτηριστικών ιστορικών και συνθετικών δεδομένων, με τη χρήση λογαριθμικού μετασχηματισμού, μηνιαίας παροχής στο Νείλο (Εφαρμογή 3, 8000 περίοδοι συνθετικών δεδομένων): (a) τυπικές αποκλίσεις λογαρίθμων, (b) συντελεστές ασυμμετρίας λογαρίθμων, (c) αυτοσυσχετίσεις βήματος 1 λογαρίθμων. Σε όλα τα διαγράμματα η παχιά μαύρη γραμμή αντιστοιχεί στα στατιστικά μεγέθη των ιστορικών δεδομένων, η λεπτή γραμμή με τους άσπρους κύκλους αντιστοιχεί στα στατιστικά μεγέθη που προέκυψαν από επιμερισμό χωρίς επανάληψη (Περίπτωση 3.1), η λεπτή γραμμή με τους μαύρους κύκλους αντιστοιχεί στα συνθετικά δεδομένα επιμερισμού με επανάληψη (Περίπτωση 3.2).



Σχήμα 5. Θεωρητική συνάρτηση κατανομής γάμα (παχιά γραμμή), εμπειρική κατανομή που προκύπτει από τα ιστορικά δεδομένα (τρίγωνα) και τα συνθετικά δεδομένα (λεπτή γραμμή) της Εφαρμογής 1 (Περίπτωση 1.2, θέση 3 και υποπερίοδος 6) .

5. Συμπεράσματα

Τα πολυδιάστατα μοντέλα επιμερισμού επιτρέπουν την προσομοίωση υδρολογικών χρονοσειρών εξασφαλίζοντας τη διατήρηση στατιστικών χαρακτηριστικών σε δύο τουλάχιστον χρονικές κλίμακες. Η μέθοδος επιμερισμού που παρουσιάζουμε στην εργασία αυτή βασίζεται σε τρεις απλές ιδέες που αποδείχτηκαν πολύ αποτελεσματικές. Πρώτον, ξεκινάει με τη χρήση ενός τυπικού μοντέλου PAR(1) το οποίο είναι το πιο φειδωλό σε παραμέτρους ανάμεσα στα γραμμικά στοχαστικά μοντέλα. Το μοντέλο αυτό παράγει μεταβλητές χαμηλότερου επιπέδου χωρίς να λαμβάνει υπόψη την υπάρχουσα πληροφορία για τις μεταβλητές υψηλότερου επιπέδου. Δεύτερον, χρησιμοποιεί μια ακριβή διορθωτική διαδικασία πάνω στις μεταβλητές χαμηλότερου επιπέδου για να ικανοποιήσει την αθροιστική ιδιότητα, δηλαδή για να "κλείσει" το άθροισμά των μεταβλητών χαμηλότερου επιπέδου στις ήδη γνωστές τιμές των μεταβλητών υψηλότερου επιπέδου. Για την παραπάνω διόρθωση αναπτύχθηκαν και μελετήθηκαν θεωρητικά και εμπειρικά τρεις διαφορετικές διαδικασίες. Τρίτον, γίνεται η χρήση επανάληψης, ώστε να ελαχιστοποιηθεί το σφάλμα στα στατιστικά μεγέθη των μεταβλητών χαμηλότερου επιπέδου που δε διατηρούνται από τη διορθωτική διαδικασία.

Το μοντέλο μπορεί να χειριστεί από κανονικές κατανομές μέχρι κατανομές σχήματος J. Στην παρούσα πολυμεταβλητή μορφή του μπορεί να φανεί χρήσιμο για πληθώρα υδρολογικών προβλημάτων όπως ο επιμερισμός ετήσιων βροχών σε μηνιαίες ή και εβδομαδιαίες, ή ακόμη ο επιμερισμός του ύψους βροχής ενός επεισοδίου σε ωριαία ή και μικρότερης διάρκειας ύψη. Για να δοκιμαστεί το μοντέλο έγιναν τέτοιες υδρολογικές εφαρμογές οι οποίες απέδειξαν την πολύ ικανοποιητική απόδοσή του.

Τα πλεονεκτήματα του μοντέλου είναι η απλότητά του, όσον αφορά τη μαθηματική διατύπωση, η φειδωλή χρήση παραμέτρων και η μαθηματική και υπολογιστική ευκολία του. Πιο σημαντικό όμως είναι το γεγονός ότι ο τρόπος ανάπτυξής του είναι τέτοιος που επιτρέπει την εύκολη μετατροπή του. Αποτελείται από ανεξάρτητα και αυτόνομα τμήματα, τα οποία ανάλογα με τις απαιτήσεις του χρήστη μπορούν να τροποποιηθούν ή να αντικατασταθούν. Για παράδειγμα η γέννηση των μεταβλητών χαμηλότερου επιπέδου μπορεί κάλλιστα να γίνει με κάποιο άλλο μοντέλο (π.χ. PARMA(p, q) αντί του PAR(1)). Επιπροσθέτως, τμήματα του μοντέλου, όπως η διαδικασία διόρθωσης, μπορούν να χρησιμοποιηθούν και σε άλλου τύπου επιμεριστικά σχήματα.

Η έρευνα που έγινε στην εργασία αυτή είχε ως πρόσθετο αποτέλεσμα ένα νέο αλγόριθμο για την αντιμετώπιση προβλημάτων σχετικών με το μοντέλο PAR(1) (όπως την αποφυγή μη θετικά ορισμένων πινάκων ή την αντιμετώπιση πολύ μεγάλων συντελεστών ασυμμετρίας) και ένα επίσης νέο αλγόριθμο παραγωγής τυχαίων μεταβλητών κατανομής γάμα.

Τέλος θα θέλαμε να σχολιάσουμε δυο θέματα που θα μπορούσαν να τύχουν κριτικής: την καταλληλότητα του PAR(1) για την προσομοίωση υδρολογικών προβλημάτων και την ταχύτητα του μοντέλου. Για το πρώτο θέμα η απάντηση είναι τριπλή. Πρώτον, το PAR(1) δεν αποτελεί κάποιο περιορισμό του επιμεριστικού μοντέλου που παρουσιάζουμε. Όπως αναφέραμε παραπάνω μπορεί ν' αντικατασταθεί από κάποιο άλλο αν παραστεί ανάγκη. Δεύτερον η φειδωλή χρήση παραμέτρων του μοντέλου δεν είναι μειονέκτημά του, αλλά το ισχυροποιεί ακόμη περισσότερο. Τρίτον, το PAR(1) μπορεί να μην είναι κατάλληλο αν εφαρμοστεί σειριακά αλλά στην περίπτωσή μας αποτελεί ένα μέρος της επιμεριστικής διαδικασίας που ευθύνεται μόνο για την ανεξάρτητη παραγωγή των μεταβλητών χαμηλότερου επιπέδου. Έτσι δεν οδηγεί σε συσσώρευση πιθανών λαθών λόγω πιθανής ατέλειας στην προσαρμογή του. Σχετικά με το δεύτερο θέμα την ταχύτητα του μοντέλου μπορεί να αναφερθούν τα εξής. Πρώτον, η μείωση της ταχύτητας που επιφέρει η επανάληψη αντισταθμίζεται εν μέρει με τη διαχείριση των κατά πολύ μειωμένων σε μέγεθος πινάκων. Δεύτερον, η μείωση του λάθους στα στατιστικά μεγέθη που δε διατηρούνται απευθείας είναι πιο σημαντική από την ταχύτητα προσομοίωσης. Τρίτον, πολλά από τα στατιστικά χαρακτηριστικά όπως οι μέσες τιμές, οι διασπορές και οι συντελεστές αυτοσυσχέτισης διατηρούνται και χωρίς επαναλήψεις ενώ και τα υπόλοιπα στατιστικά μεγέθη προσεγγίζονται ικανοποιητικά. Έτσι μπορούμε να τρέξουμε το μοντέλο χωρίς επαναλήψεις για να έχουμε μια αρχική ιδέα του τι κάνει και στην τελική φάση να χρησιμοποιήσουμε επαναλήψεις. Τέταρτο και τελευταίο, ο χρόνος που χρειάστηκε το μοντέλο για να κάνει τις προσομοιώσεις της Ενότητας 5 κυμάνθηκε από μερικά λεπτά μέχρι μια ώρα (σε PC/486) για την παραγωγή 10000 περιόδων μηνιαίων παροχών με επανάληψη, χρόνος που δεν είναι καθόλου σημαντικός.

Παράρτημα 2: Μαθηματικές αποδείξεις

1. Απόδειξη της πρότασης 2

Αν πάρουμε τις μέσες τιμές των μελών της εξίσωσης (16) τότε με τη χρήση της (14) προκύπτει

$$E[X_s] = E[\tilde{X}_s] \tag{A1}$$

Αποδεικνύεται έτσι το πρώτο μέρος της πρότασης που αφορά την ισότητα των μέσων τιμών. Με τη χρήση της (14) και της (A1) η (16) μπορεί να γραφεί

$$\left(X_s - E[X_s]\right) = \left(\tilde{X}_s - E[\tilde{X}_s]\right) + \lambda_s \left\{ \left(Z - E[Z]\right) - \sum_{j=1}^{k} \left(\tilde{X}_j - E[\tilde{X}_j]\right) \right\} \tag{A2}$$

Αν συμβολίσουμε με $X'_s = X_s - E[X_s]$, $\tilde{X}'_s = \tilde{X}_s - E[\tilde{X}_s]$, $Z' = Z - E[Z]$, και ορίσουμε τα διανύσματα $\mathbf{X}' = [X'_1, \cdots, X'_k]^T$, $\tilde{\mathbf{X}}' = [\tilde{X}'_1, \cdots, \tilde{X}'_k]^T$, $\boldsymbol{\lambda} = [\lambda_1, \cdots, \lambda_k]^T$ και $\boldsymbol{\Phi} = [1, \cdots, 1]$ (διάνυσμα-γραμμή με k στοιχεία) τότε η (A2) μπορεί να παρασταθεί, με συμβολισμό πινάκων, ως εξής

$$\mathbf{X}' = \tilde{\mathbf{X}}' - \boldsymbol{\lambda}\boldsymbol{\Phi}\tilde{\mathbf{X}}' + \boldsymbol{\lambda} Z' \tag{A3}$$

Στη συνέχεια αν αναστρέψουμε τα δυο μέλη της (A3) και πολλαπλασιάσουμε από δεξιά την (A3) προκύπτει

$$\mathbf{X}'\mathbf{X}'^T = \left(\tilde{\mathbf{X}}' - \boldsymbol{\lambda}\boldsymbol{\Phi}\tilde{\mathbf{X}}' + \boldsymbol{\lambda} Z'\right)\left(\tilde{\mathbf{X}}'^T - \tilde{\mathbf{X}}'^T \boldsymbol{\Phi}^T \boldsymbol{\lambda}^T + Z'\boldsymbol{\lambda}^T\right) \tag{A4}$$

που μετά από πράξεις γίνεται

$$\begin{aligned} \mathbf{X}'\mathbf{X}'^T &= \tilde{\mathbf{X}}'\tilde{\mathbf{X}}'^T - \tilde{\mathbf{X}}'\tilde{\mathbf{X}}'^T \boldsymbol{\Phi}^T \boldsymbol{\lambda}^T + \tilde{\mathbf{X}}' Z' \boldsymbol{\lambda}^T - \boldsymbol{\lambda}\boldsymbol{\Phi}\tilde{\mathbf{X}}'\tilde{\mathbf{X}}'^T + \boldsymbol{\lambda}\boldsymbol{\Phi}\tilde{\mathbf{X}}'\tilde{\mathbf{X}}'^T \boldsymbol{\Phi}^T \boldsymbol{\lambda}^T \\ &- \boldsymbol{\lambda}\boldsymbol{\Phi}\tilde{\mathbf{X}}' Z' \boldsymbol{\lambda}^T + \boldsymbol{\lambda} Z' \tilde{\mathbf{X}}'^T - \boldsymbol{\lambda} Z' \tilde{\mathbf{X}}'^T \boldsymbol{\Phi}^T \boldsymbol{\lambda}^T + \boldsymbol{\lambda} Z'^2 \boldsymbol{\lambda}^T \end{aligned} \tag{A5}$$

Παίρνοντας μέσες τιμές στην παραπάνω εξίσωση, λαμβάνοντας υπόψη ότι $E[\tilde{\mathbf{X}}'\tilde{\mathbf{X}}'^T] = \boldsymbol{\sigma}$, $E[\tilde{\mathbf{X}}' Z'] = \mathbf{0}$ (διάνυσμα με μηδενικές τιμές λόγω της ανεξαρτησίας των $\tilde{X}_s$ και Z) και $E[Z'^2] = \sigma_{ZZ}$ καταλήγουμε

$$E\left[\mathbf{X}'\mathbf{X}'^T\right]=\boldsymbol{\sigma}-\boldsymbol{\sigma}\boldsymbol{\Phi}^T\boldsymbol{\lambda}^T-\boldsymbol{\lambda}\boldsymbol{\Phi}\boldsymbol{\sigma}+\boldsymbol{\lambda}\boldsymbol{\Phi}\boldsymbol{\sigma}\boldsymbol{\Phi}^T\boldsymbol{\lambda}^T+\sigma_{ZZ}\boldsymbol{\lambda}\boldsymbol{\lambda}^T \tag{A6}$$

Ο πίνακας $\boldsymbol{\Phi\sigma}$ που εμφανίζεται στον τρίτο όρο της (Α6) είναι

$$\boldsymbol{\Phi\sigma}=\begin{bmatrix}1 & 1 & \cdots & 1\end{bmatrix}\begin{bmatrix}\sigma_{11} & \sigma_{12} & \cdots & \sigma_{1k}\\ \sigma_{12} & \sigma_{22} & \cdots & \sigma_{2k}\\ \vdots & \vdots & \ddots & \vdots\\ \sigma_{1k} & \sigma_{2k} & \cdots & \sigma_{kk}\end{bmatrix}=\begin{bmatrix}\sigma_{1Z} & \sigma_{2Z} & \cdots & \sigma_{kZ}\end{bmatrix}=\sigma_{ZZ}\boldsymbol{\lambda}^T \tag{A7}$$

Για την εξαγωγή της παραπάνω εξίσωσης χρησιμοποιήθηκαν οι εξισώσεις (18) και (17), που καθορίζουν τα σ_{sZ} και λ_s, αντίστοιχα. Επειδή ο πίνακας $\boldsymbol{\lambda\Phi\sigma}$ είναι συμμετρικός, ο δεύτερος όρος του δεξιού τμήματος της (Α6) ισούται με τον τρίτο. Επιπροσθέτως, ο πίνακας $\boldsymbol{\Phi\sigma\Phi}^T$ που εμφανίζεται στον τέταρτο όρο του δεξιού τμήματος της (Α6) είναι

$$\boldsymbol{\Phi\sigma\Phi}^T=\sigma_{ZZ}\begin{bmatrix}\lambda_1 & \lambda_2 & \cdots & \lambda_k\end{bmatrix}\begin{bmatrix}1\\1\\ \vdots\\1\end{bmatrix}=\sigma_{ZZ}\sum_{s=1}^{k}\lambda_s=\sigma_{ZZ} \tag{A8}$$

Συνεπώς η (Α6) γίνεται

$$E\left[\mathbf{X}'\mathbf{X}'^T\right]=\boldsymbol{\sigma}-\sigma_{ZZ}\boldsymbol{\lambda}\boldsymbol{\lambda}^T-\sigma_{ZZ}\boldsymbol{\lambda}\boldsymbol{\lambda}^T+\sigma_{ZZ}\boldsymbol{\lambda}\boldsymbol{\lambda}^T+\sigma_{ZZ}\boldsymbol{\lambda}\boldsymbol{\lambda}^T=\boldsymbol{\sigma} \tag{A9}$$

και έτσι αποδεικνύεται ότι ο πίνακας συνδιασπορών των μεταβλητών X_s ισούται με τον αντίστοιχο των μεταβλητών $\tilde{X}_s$, δηλαδή η διαδικασία διόρθωσης δεν μεταβάλλει τα στατιστικά μεγέθη δεύτερης τάξης των μεταβλητών χαμηλότερου επιπέδου.

Για να αποδείξουμε ότι οι συντελεστές $\boldsymbol{\lambda}$, όπως καθορίζονται από τη σχέση (17), είναι μονοσήμαντα ορισμένοι, δηλαδή κανένα άλλο διάνυσμα $\boldsymbol{\lambda}$ δεν μπορεί να ικανοποιήσει την αθροιστική ιδιότητα και ταυτόχρονα να διατηρήσει τις μέσες τιμές και τους πίνακες διασποράς-συνδιασποράς, υψώνουμε στο τετράγωνο τα μέλη της εξίσωσης (Α2) και παίρνουμε μέσες τιμές. Έτσι

$$\mathrm{Var}\left[X_s\right]=\mathrm{Var}\left[\tilde{X}_s\right]+2\lambda_s^2\ \mathrm{Var}\left[\tilde{Z}\right]-2\lambda_s\ \mathrm{Cov}\left[\tilde{X}_s,\tilde{Z}\right] \tag{A10}$$

Επειδή απαιτούμε $\mathrm{Var}\left[X_s\right]=\mathrm{Var}\left[\tilde{X}_s\right]$, η (Α10) γίνεται

$$\lambda_s^2\, \sigma_{ZZ} - \lambda_s\, \sigma_{sZ} = 0 \tag{A11}$$

Η παραπάνω εξίσωση έχει δύο ρίζες λ_s. Η λύση $\lambda_s = 0$ απορρίπτεται γιατί δεν ικανοποιεί την αθροιστική ιδιότητα. Η άλλη λύση $\lambda_s = \sigma_{sZ} / \sigma_{ZZ}$, είναι αυτή που καθορίζεται από την εξίσωση (17).

2. Μια καινούργια διαδικασία παραγωγής τυχαίων αριθμών κατανομής γάμα

Ας συμβολίσουμε με x μια τυχαία μεταβλητή που ακολουθεί κατανομή γάμα με συνάρτηση πυκνότητας πιθανότητας

$$f(x) = \frac{\lambda^{\kappa}}{\Gamma(\kappa)}(x-c)^{\kappa-1} e^{-\lambda(x-c)} \tag{B1}$$

και παραμέτρους $\kappa > 0$, $\lambda > 0$ και c. Το ζητούμενο είναι να καθορίσουμε μια διαδικασία γέννησης τυχαίων αριθμών για τη συγκεκριμένη κατανομή. Θα χρησιμοποιήσουμε την μέθοδο απόρριψης [*Papoulis*, 1990, pp. 261-263], που χρησιμοποιεί μια βοηθητική συνάρτηση πυκνότητας πιθανότητας $g(x)$. Χωρίς απώλεια της γενικότητας θεωρούμε $\kappa < 1$. (Αν $\kappa \geq 1$ μπορούμε να το χωρίσουμε στο δεκαδικό και το ακέραιο μέρος).

Η βοηθητική συνάρτηση ορίζεται ως εξής

$$g(x) = \begin{cases} \beta(x-c)^{\kappa-1}, & c \leq x \leq h \\ \beta\gamma e^{-\lambda(x-c)}, & x > h \end{cases} \tag{B2}$$

όπου β και γ είναι κάποιοι σταθεροί συντελεστές και h είναι ένας αριθμός μεγαλύτερος του c. Η απαίτηση για την συνέχεια της $g(x)$ οδηγεί στην εξίσωση

$$(h-c)^{\kappa-1} = \gamma e^{-\lambda(h-c)} \tag{B3}$$

απ' την οποία προκύπτει

$$\gamma = (h-c)^{\kappa-1} e^{\lambda(h-c)} \tag{B4}$$

Επειδή η $g()$ είναι συνάρτηση πυκνότητας πιθανότητας πρέπει

$$\int_c^{\infty} g(x)dx = 1 \tag{B5}$$

Από την τελευταία εξίσωση και μετά από πράξεις καταλήγουμε ότι

$$\beta = \frac{\kappa}{(h-c)^{\kappa-1}(h-c+\kappa/\lambda)} \tag{B6}$$

Συνεπώς η (B2) γίνεται

$$g(x) = \begin{cases} \dfrac{\kappa}{h-c+\kappa/\lambda} \left(\dfrac{x-c}{h-c}\right)^{\kappa-1}, & c \le x \le h \\ \dfrac{\kappa}{h-c+\kappa/\lambda}\, e^{-\lambda(x-h)}, & x > h \end{cases} \tag{B7}$$

και η αντίστοιχη συνάρτηση κατανομής είναι

$$G(x) = \begin{cases} \dfrac{h-c}{h-c+\kappa/\lambda} \left(\dfrac{x-c}{h-c}\right)^{\kappa}, & c \le x \le h \\ 1-\dfrac{\kappa/\lambda}{h-c+\kappa/\lambda}\, e^{-\lambda(x-h)}, & x > h \end{cases} \tag{B8}$$

Η $G(x)$ μπορεί να αντιστραφεί, οπότε προκύπτει

$$x(G) = \begin{cases} \left[\dfrac{h-c+\kappa/\lambda}{h-c} G\right]^{1/\kappa} (h-c)+c, & 0 \le G \le \dfrac{h-c}{h-c+\kappa/\lambda} \\ -\dfrac{1}{\lambda}\ln\left[(1-G)\dfrac{h-c+\kappa/\lambda}{\kappa/\lambda}\right]+h, & G > \dfrac{h-c}{h-c+\kappa/\lambda} \end{cases} \tag{B9}$$

και έτσι προκύπτει ο τρόπος παραγωγής των τυχαίων μεταβλητών μέσω της παραπάνω κατανομής.

Ας ορίσουμε τώρα τη συνάρτηση

$$r(x) = a\frac{f(x)}{g(x)} \tag{B10}$$

όπου a είναι μια σταθερή ποσότητα τέτοια ώστε $\max[r(x)] = 1$. Στην περίπτωσή μας

$$a = \frac{\kappa\Gamma(\kappa)}{\lambda^{\kappa}(h-c+\kappa/\lambda)(h-c)^{\kappa-1}} \tag{B11}$$

και συνεπώς

$$r(x)=\begin{cases} e^{-\lambda(x-c)}, & c\le x\le h \\ e^{-\lambda(h-c)}\left(\dfrac{x-c}{h-c}\right)^{\kappa-1}, & x>h \end{cases} \tag{B12}$$

Παρατηρούμε ότι η $r(x)$ είναι φθίνουσα με $r(0) = 1$ και $r(\infty) = 0$.

Ακολουθώντας την μέθοδο απόρριψης [*Papoulis*, 1990, pp. 261-263] μπορούμε να καθορίσουμε τα βήματα της διαδικασίας παραγωγής τυχαίων αριθμών:

1. Παράγουμε δύο ανεξάρτητες μεταβλητές ομοιόμορφης κατανομής u_1 και u_2 στο διάστημα [0, 1].
2. Με τη βοήθεια της (Β9) υπολογίζουμε τον τυχαίο αριθμό $x = x(u_1)$ με συνάρτηση κατανομής $G(x)$ (θέτουμε $G = u_1$ στην (Β9) και υπολογίζουμε το x).
3. Υπολογίζουμε το $r(x)$.
4. Αν ισχύει $u_2 \le r(x)$ δεχόμαστε την τιμή x ως τυχαίο αριθμό από την κατανομή γάμα με συνάρτηση πυκνότητας πιθανότητας (Β1). Αλλιώς απορρίπτουμε τη x και επαναλαμβάνουμε τη διαδικασία.

Για μαθηματική απλότητα εξισώνουμε το h με τη μέση τιμή του x, δηλαδή

$$h = E[x] = c + \kappa / \lambda \tag{B13}$$

κάτι που απλοποιεί όλες τις παραπάνω σχέσεις. Στην περίπτωση αυτή η διάμεση τιμή της $g(x)$ ισούται με h.

Αναφορές

Box, G. E. P., and Jenkins, G. M., *Time series analysis; Forecasting and control*, Holden Day, 1970.

Bras, R. L. and Rodriguez-Iturbe, I., *Random functions in hydrology*, Addison-Wesley, 1985.

Eckhardt, R., Stan Ulam, John von Neumann and the Monte Carlo method, in *From cardinals to chaos*, ed. N. G. Cooper, Cambridge University, NY., 1989.

Fernandez, B., and Salas, J. D., Gamma-autoregressive models for stream-flow simulation, *J. Hydraul. Eng.*, 116(11) 1403-1414, 1990.

Fernandez, B., and Salas, J. D., Periodic gamma autoregressive processes for operational hydrology, *Water Resour. Res.*, 22(10) 1385-1396, 1986

Grygier, J. C. and Stedinger, J. R., Condensed disaggregation procedures and conservation corrections for stochastic hydrology, *Water Resour. Res.*, 24(10), 1574-1584, 1988.

Grygier, J. C. and Stedinger, J. R., SPIGOT, A synthetic streamflow generation software package, Technical description, School of Civil and Environmental Engineering, Cornell University, Ithaca, NY., Version 2.5, 1990.

Hoshi, K. and Burges, S. J., Disaggregation of streamflow volumes, *Journal of the Hydraulics Division, Proceedings ASCE*, 105 (HY1), 27-41, 1979.

Kirby, W., Algebraic boundness of sample statistics, *Water Resour. Res.*, 10(2) 220-222, 1974.

Koutsoyiannis, D., A nonlinear disaggregation model with a reduced parameter set for simulation of hydrologic series, *Water Resour. Res.*, 28(12), 3175-3191, 1992.

Koutsoyiannis, D., A stochastic disaggregation method for design storm and flood synthesis, *Journal of Hydrology*, 156, 193-225, 1994.

Lane, W. L., Applied Stochastic Techniques, User's Manual, Bureau of Reclamation, Engineering and Research Center, Denver, Co., 1979.

Lane, W. L., Corrected parameter estimates for disaggregation schemes, in *Statistical Analysis of Rainfall and Runoff*, V. P. Singh, ed. Water Resources Publications, Littleton, Colo., 1982.

Lane, W. L. and Frevert, D. K., Applied Stochastic Techniques, User's Manual, Bureau of Reclamation, Engineering and Research Centre, Denver, Co., Personal Computer Version 1990.

Lawrance, A. J., and Lewis, P. A. W., A new autoregressive time series model in exponential variables [NEAR(1)], *Adv. Appl. Prob.*, 13(4), 826-845, 1981.

Lin, G.-F., Parameter estimation for seasonal to subseasonal disaggregation, *J. of Hydrol.* 120(1-4), 65-77, 1990.

Matalas, N. C. and Wallis, J. R., Generation of synthetic flow sequences, in *Systems Approach to Water Management*, edited by A. K. Biswas, McGraw-Hill, New York, 1976.

Mejía, J. M., and Millán, J., Una Metodología para tratar el problema de matrices inconsistentes en la generatión multivariada de series hidrológicas, *VI Congreso Latino Americano de Hydraúlica*, Bogotá, Colombia, 1974

Mejia, J. M. and Rousselle, J., Disaggregation models in Hydrology revisited, *Water Resour. Res.*, 12(2), 185-186, 1976.

Metropolis, N., The beginning of the Monte Carlo method, in *From cardinals to chaos*, ed. N. G. Cooper, Cambridge University, NY., 1989.

Oliveira, G. C., Kelman, J., Pereira, M. V. F., and Stedinger, J. R., Representation of spatial cross-correlation in a seasonal streamflow model, *Water Resour. Res.*, 24(5), 781-785, 1988.

Papoulis, A., *Probability and statistics*, Prentice-Hall, New Jersey, 1990.

Pereira, M. V. F., Oliveira, G. C., Costa, C. C. G., and Kelman, J., Stochastic streamflow models for hydroelectric systems, *Water Resour. Res.*, 20(3), 379-390, 1984.

Salas, J. D., Delleur, J. W., Yevjevich, V. and Lane, W. L., *Applied Modelling of Hydrologic Time Series*, Water Resources Publications, Littleton, Colo. 1980.

Stedinger, J. R. and Vogel, R. M., Disaggregation procedures for generating serially correlated flow vectors, *Water Resour. Res.*, 20(1) 47-56, 1984.

Stedinger, J. R., Pei, D. and Cohn, T. A., A condensed disaggregation model for incorporating parameter uncertainty into monthly reservoir simulations, *Water Resour. Res.*, 21(5) 665-675, 1985.

Tao, P.C. and Delleur, J. W., Multistation, multiyear synthesis of hydrologic time series by disaggregation, *Water Resour. Res.*, 12(6), 1303-1312, 1976.

Todini, E., The preservation of skewness in linear disaggregation schemes, *J. Hydrol.*, 47, 199-214, 1980.

Valencia, D. and Schaake, J. C., A disaggregation model for time series analysis and synthesis, Report no. 149, Ralph M. Parsons Laboratory for Water Resources and Hydrodynamics, Massachusetts Institute of Technology, Cambridge, Mass., 1972.

Valencia, D. and Schaake, J. C., Disaggregation processes in Stochastic Hydrology, *Water Resour. Res.*, 9(3) 211-219, 1973.

Vogel, R. M. and Stedinger, J. R., The value of stochastic sreamflow models in one-year reservoir design applications, *Water Resour. Res.*, 24(9), 1483-1490, 1988.

Whittaker, J., A note on the generation of gamma random variables with non-integral shape parameter, *Proceedings of the second international symposium on Hydrology*, Water Resources Publications, Fort Collins, Colorado, 591-594, 1973.

Wallis, J. R., Matalas, N. and Slack, J. R., Just a moment!, *Water Resour. Res.*, 10(2) 211-219, 1974.