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

ΤΕΥΧΟΣ 7
ΠΡΟΓΡΑΜΜΑΤΑ ΣΤΟΧΑΣΤΙΚΗΣ ΠΡΟΣΟΜΟΙΩΣΗΣ ΥΔΡΟΛΟΓΙΚΩΝ ΧΡΟΝΟΣΕΙΡΩΝ

RESEARCH PROJECT
EVALUATION AND MANAGEMENT OF THE WATER RESOURCES OF STEREA HELLAS

VOLUME 7
COMPUTER PROGRAMMES FOR STOCHASTIC SIMULATION OF HYDROLOGIC TIME SERIES

ΣΥΝΤΑΞΗ: Δ. ΚΟΥΤΣΟΓΙΑΝΝΗΣ
ΕΠΙΣΤΗΜΟΝΙΚΟΣ ΥΠΕΥΘΥΝΟΣ: Θ. ΞΑΝΘΟΠΟΥΛΟΣ
ΚΥΡΙΟΣ ΕΡΕΥΝΗΤΗΣ: Δ. ΚΟΥΤΣΟΓΙΑΝΝΗΣ

ΑΘΗΝΑ - ΟΚΤΩΒΡΙΟΣ 1992

BY: D. KOUTSOYIANNIS
SCIENTIFIC DIRECTOR: TH. XANTHOPOULOS
PRINCIPAL INVESTIGATOR: D. KOUTSOYIANNIS

ATHENS - OCTOBER 1992

ΠΡΟΓΡΑΜΜΑΤΑ ΣΤΟΧΑΣΤΙΚΗΣ ΠΡΟΣΟΜΟΙΩΣΗΣ ΥΔΡΟΛΟΓΙΚΩΝ ΧΡΟΝΟΣΕΙΡΩΝ

ΠΕΡΙΕΧΟΜΕΝΑ

1. ΕΙΣΑΓΩΓΗ

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 Α). 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 Δ6/20595/9-5-1991) σε ερευνητική ομάδα του Τομέα Υδατικών Πόρων, Υδραυλικών και Θαλάσσιων Έργων του ΕΜΠ με επιστημονικό υπεύθυνο τον καθηγητή Θ. Ξανθόπουλο. Η συγκεκριμένη εργασία την οποία καλύπτει το τεύχος αυτό προδιαγράφεται στο Παράρτημα της απόφασης ανάθεσης (άρθρο 2.2.4, εδάφιο γ: Προγράμματα παραγωγής συνθετικών σειρών με στοχαστικά μοντέλα).

Η μεθοδολογία που ακολουθείται στα προγράμματα αυτά χρησιμοποιήθηκε σε πρώτη μορφή σε προηγούμενο ερευνητικό έργο και συγκεκριμένα στο έργο *Διερεύνηση προσφερομένων δυνατοτήτων για την ενίσχυση της ύδρευσης μείζονος περιοχής Αθηνών*. Και αυτό το ερευνητικό έργο είχε ανατεθεί και χρηματοδοτηθεί από την ίδια διεύθυνση του ΥΠΕΧΩΔΕ στην ίδια ερευνητική ομάδα του ΕΜΠ. Κατά συνέπεια θα πρέπει να θεωρείται ότι τα προγράμματα έχουν ελεγχθεί αρκετά ως προς τη μεθοδολογία και τα αποτελέσματα τους. Ωστόσο, καμιά εργασία δεν μπορεί να θεωρηθεί ως ολοκληρωμένη, πολύ περισσότερο δε αν πρόκειται για πρόγραμμα ηλεκτρονικού υπολογιστή. Ας θεωρηθούν λοιπόν τα προγράμματα στην παρούσα τους μορφή ως μια πρώτη έκδοση που καλύπτει βασικές ανάγκες υδρολογικής προσομοίωσης. Φιλοδοξία μας είναι, εφόσον μας δοθεί η δυνατότητα, να βελτιώσουμε σημαντικά τα προγράμματα αυτά, και ως προς το θεωρητικό τους υπόβαθρο, και ως προς τις προσφερόμενες δυνατότητες, αλλά και ως προς τον κώδικα και τον τρόπο επικοινωνίας του χρήστη με τον υπολογιστή.

1.2 Διάρθρωση του τεύχους

Το τεύχος αποτελείται από ένα κύριο μέρος και τρία 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, τα δύο πρώτα παραρτήματα καλύπτουν τη μεθοδολογία και τους αλγορίθμους στα οποία στηρίζονται τα προγράμματα προσομοίωσης, ενώ το τρίτο παράρτημα δίνει πληροφορίες σχετικές με τη δομή των δυαδικών αρχείων που δημιουργούν τα προγράμματα προσομοίωσης

2. ΣΥΝΟΠΤΙΚΟ ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ

2.1 Εισαγωγή

Τα προγράμματα προσομοίωσης υδρολογικών χρονοσειρών δίνουν τη δυνατότητα γέννησης συνθετικών σειρών υδρολογικών μεταβλητών (απορροής, βροχής, εξάτμισης κ.ά.) σε πολλες 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, καί στη χώρα μας αλλά και διεθνώς, βασίζονται στις ιστορικές υδρολογικές χρονοσειρές και ακολουθούν ανάλογες καθιερωμένες μεθοδολογίες. Ωστόσο, έχει αποδειχτεί (πχ. *Vogel & 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 (με συνηθισμένο 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 & Stedinger*, 1990]. Όμως, η μέθοδος που χρησιμοποίησε ήταν εμπειρική, βασισμένη σε αναγωγή και ενοποίηση ιστορικών δειγματων, 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 κα. 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 & Stedinger*, 1990]. Πολυ 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 & 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. Στο δεύτερο επίπεδο (γνωστό ως χαμηλότερο επίπεδο) γίνεται λεπτομερέστερη γέννηση των χρονοσειρών σε πυκνότερη χρονική κλίμακα, χρησιμοποιώντας μια κατάλληλη τεχνική επιμερισμού. 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 ενδοετήσιες 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 των προγραμμάτων.

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. Αναλυτικότερα, τα συγκριτικά πλεονεκτήματα και μειονεκτήματα των δύο σχημάτων παρουσιάζονται στο Παράρτημα 2.

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 & Frevert*, 1990] και το SPIGOT [*Grygier & Stedinger*, 1990]. Τα δικά μας προγράμματα παρουσιάζουν κάποιες ομοιότητες με αυτά, αλλά διαφέρουν σημαντικά κυρίως στο νέο μοντέλο επιμερισμού που χρησιμοποιούν, το οποίο αναλύεται στο Παράρτημα 2.

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 & Wallis*, 1976, σ. 60, 63]: Η επιλογή της ελαχιστοποίησης του αριθμού των παραμέτρων έγινε με σκοπό να είναι τα προγράμματα κατά το δυνατόν εύχρηστα και γρήγορα, 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 2). Βέβαια, είναι θεωρητικά δυνατό, στο βαθμό που κρίνονται ουσιώδεις, να εισαχθούν και να αναπαραχθούν και άλλες ομάδες παραμέτρων, αλλά αυτό δεν υποστηρίζεται από τα συγκεκριμένα προγράμματα.

Σημειώνεται ότι το παραπάνω σύνολο παραμέτρων αποτελεί την "είσοδο" στα συγκεκριμένα προγράμματα. Κατά συνέπεια οι παράμετροι αυτές δεν υπολογίζονται από τα ίδια τα προγράμματα προσομοίωσης αλλά χρειάζεται να γίνει ξεχωριστός υπολογισμός τους από κατάλληλα ιστορικά δείγματα με βάση άλλα προγράμματα (π.χ. με το πρόγραμμα ΥΔΡΟΛΟΓΙΟ που εκπονήθηκε στα πλαίσια του ίδιου ερευνητικού έργου).

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 κά. Τα μοντέλα αυτά είναι αρκετά πολύπλοκα και ο ενδιαφερόμενος αναγνώστης παραπέμπεται σε εγχειρίδια στοχαστικής υδρολογίας, πχ. *Bras & 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) ήAR(2) (πχ. στο SPIGOT έχουν υιοθετηθεί τα AR(0) και AR(1) [*Grygier & Stedinger*, 1990]) και στο LAST τα AR(1) kai AR(2) [*Lane & Frevert*, 1990].

Στα δικά 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). Αναλυτική περιγραφή του μοντέλου δίνεται στο Παράρτημα 1.

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 *δυναμικό 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 του παρόντος και προηγούμενων ερευνητικών έργων. 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 δυναμικό 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).

Αναλυτική περιγραφή του δυναμικού μοντέλου επιμερισμού, ως προς τη θεωρητική του βάση και τους αλγορίθμους του, δίνεται στο Παράρτημα 2.

3. Η ΧΡΗΣΗ ΤΩΝ ΠΡΟΓΡΑΜΜΑΤΩΝ

3.1 Εκτελέσιμα προγράμματα

Όλοι οι απαιτούμενοι αλγόριθμοι για τη γέννηση και τον έλεγχο των χρονοσειρών έχουν ομαδοποιηθεί σε δύο εκτελέσιμα προγράμματα:

1. Το πρόγραμμα DDM που επιτρέπει τη γέννηση των συνθετικών χρονοσειρών και των δύο επιπέδων, και
2. Το πρόγραμμα DDMTEST που επιτρέπει τον έλεγχο των συνθετικών χρονοσειρών και συγκεκριμένα τον υπολογισμό των στατιστικών χαρακτηριστικών τους, των εμπειρικών συναρτήσεων κατανομής τους κτλ.

Τα προγράμματα τρέχουν σε οποιοδήποτε προσωπικό υπολογιστή με λειτουργικό σύστημα DOS, version 3.00 ή νεότερη. Δεν προϋποθέτουν ειδική σύνθεση του υπολογιστή αλλά για να λειτουργήσουν με ικανοποιητική ταχύτητα χρειάζονται μαθηματικό συνεπεξεργαστή και κατά το δυνατόν μεγαλύτερη συχνότητα CPU.

Οι περιορισμοί που έχουν τεθεί στα προγράμματα είναι οι εξής:

1. Μέχρι 10 διαφορετικές θέσεις (μεταβλητές)
2. Μέχρι 24 χρονικές υποδιαιρέσεις του έτους
3. Μέχρι 16 000 συνολικά δεδομένα για κάθε θέση και χρονική υποδιαίρεση

Ο τρόπος κλήσης των προγραμμάτων περιγράφεται στα κεφάλαια 3.4 και 3.5. Δεδομένου ότι τα προγράμματα χρειάζονται πολλά στοιχεία εισόδου που αν η εισαγωγή τους γίνονταν κάθε φορά από το πληκτρολόγιο θα ήταν χρονοβόρα, έχει υιοθετηθεί ο τρόπος εισαγωγής τους μέσω αρχείων τύπου κειμένου (text). Συγκεκριμένα για κάθε πρόβλημα χρειάζονται τρία αρχεία στοιχείων εισόδου, ένα γενικών πληροφοριών, ένα για τα στατιστικά χαρακτηριστικά των μεταβλητών υψηλότερου επιπέδου και ένα για τα στατιστικά χαρακτηριστικά των μεταβλητών χαμηλότερου επιπέδου. Αναλυτικότερα στοιχεία για τα αρχεία αυτά δίνονται στο υποκεφάλαιο 3.2 και στο κεφάλαιο 4.

Το πρόγραμμα DDM κατασκευάζει μέχρι δύο δυαδικά αρχεία συνθετικών δεδομένων. Το ένα από αυτά περιέχει τις χρονοσειρές υψηλότερου επιπέδου και το άλλο από κοινού τις χρονοσειρές υψηλότερου και χαμηλότερου επιπέδου. Αναλυτικότερα τα αρχεία αυτά περιγράφονται στο υποκεφάλαιο 3.3 και στο Παράρτημα 3. Επειδή τα εν λόγω αρχεία είναι δυαδικά δεν είναι άμεσα αναγνώσιμα από το χρήστη. Η ανάγνωση τους όμως μπορεί να γίνει με το πρόγραμμα DDMTEST.

Το πρόγραμμα DDMTEST δίνει όλα τα αποτελέσματα του στην οθόνη, αλλά είναι δυνατό να κατασκευάσει και αρχεία στα οποία θα γράψει τα αποτελέσματα του. Τα αρχεία αυτά είναι τύπου κειμένου (text) και επομένως άμεσα αναγνώσιμα από το χρήστη.

Όλα τα παραπάνω αρχεία, εκτός από το ενδεχόμενο αρχείο αποτελεσμάτων του DDMTEST έχουν στο όνομα τους ένα κοινό θέμα που είναι το όνομα του προβλήματος. Καθένα απ' αυτά ξεχωρίζει από την κατάληξη του, που είναι διαφορετική για κάθε τύπο αρχείου. Οι καταλήξεις έχουν τυποποιηθεί με τον τρόπο που εξηγείται στο επόμενο παράδειγμα: Έστω ότι στο πρόβλημα που μας απασχολεί δίνουμε το όνομα ABCD (μέχρι 8 το πολύ χαρακτήρες). Τα αρχεία που θα δημιουργηθούν είναι τα ακόλουθα

Αρχεία εισόδου (δημιουργούνται από το χρήστη)

1. ABCD.INF (αρχείο γενικών πληροφοριών, υποχρεωτικό)
2. ABCD.INH (αρχείο στατιστικών χαρακτηριστικών μεταβλητών υψηλότερου επιπέδου, υποχρεωτικό)
3. ABCD.INL (αρχείο στατιστικών χαρακτηριστικών μεταβλητών χαμηλότερου επιπέδου, υποχρεωτικό)

Αρχεία συνθετικών χρονοσειρών (δημιουργούνται από το πρόγραμμα DDM)

4. ABCD.OUH (αρχείο συνθετικών χρονοσειρών μεταβλητών χαμηλότερου επιπέδου)
5. ABCD.OUL (αρχείο συνθετικών χρονοσειρών μεταβλητών χαμηλότερου και υψηλότερου επιπέδου)

Παρακάτω για λόγους συντομίας ο κάθε τύπος αρχείου θα αναφέρεται με την κατάληξη του ονόματος του, όπως αυτή ορίζεται παραπάνω.

3.2 Αρχεία στοιχείων εισόδου

Όπως προαναφέρθηκε, ο χρήστης, προκειμένου να αξιοποιήσει τα προγράμματα, θα πρέπει να κατασκευάσει τρία αρχεία τύπου κειμένου (text) για το συγκεκριμένο πρόβλημα που καλείται να επιλύσει.

Το πρώτο αρχείο (INF) περιλαμβάνει τα γενικά δεδομένα του προβλήματος και συγκεκριμένα:

- τον αριθμό των (εναλλακτικών) χρονοσειρών που θα δημιουργήσει το πρόγραμμα DDM
- το μήκος της κάθε χρονοσειράς (αριθμός περιόδων)
- τον αριθμό των θέσεων
- τον αριθμό των υποπεριόδων κάθε περιόδου
- οδηγία για τα αρχεία που θα δημιουργήσει το πρόγραμμα DDM (εναλλακτικές δυνατότητες: δημιουργία ή όχι αρχείων τύπου OUH και OUL, χρήση ή όχι τυχόν υπάρχοντος (από προηγούμενη εκτέλεση του DDM) αρχείου τύπου OUH - βλ. και υποκεφ. 3.4 καθώς και κεφ. 4)
- συνθήκη για την αποδοχή ή όχι αρνητικών τιμών για τις μεταβλητές (κατά κανόνα οι αρνητικές τιμές δεν είναι αποδεκτές), και
- την αρχική τιμή της γεννήτριας τυχαίων αριθμών (random seed) που θα χρησιμοποιήσει το πρόγραμμα DDM (τυχών ακέραιος αριθμός)

Το δεύτερο αρχείο (INH) περιλαμβάνει τα στατιστικά χαρακτηριστικά των μεταβλητών χαμηλότερου επιπέδου (ετήσιων) και συγκεκριμένα:

- τον τύπο της στατιστικής κατανομής (στην τρέχουσα έκδοση μπορεί να είναι κανονική ή γάμα)
- τις μέσες τιμές των μεταβλητών
- τις διασπορές των μεταβλητών
- τις τρίτες ροπές των μεταβλητών
- τους συντελεστές αυτοσυσχέτισης
- τους συντελεστές ετεροσυσχέτισης
- τυχόν περιορισμούς (προαιρετικά)

Τέλος το τρίτο αρχείο (INL) περιέχει τους ίδιους τύπους στατιστικών χαρακτηριστικών όπως παραπάνω αλλά για τις μεταβλητές χαμηλότερου επιπέδου. Για κάθε τύπο δίνονται ξεχωριστά τα στατιστικά χαρακτηριστικά των μεταβλητών κάθε υποπεριόδου.

3.3 Αρχεία συνθετικών χρονοσειρών

Τα δύο αρχεία συνθετικών χρονοσειρών, όπως αναφέρθηκε και πιο πάνω, κατασκευάζονται από το πρόγραμμα DDM. Πρόκειται για δομημένα αρχεία που αποτελούνται από εγγραφές (files of records) δυαδικού (binary) τύπου και γι' αυτό δεν είναι άμεσα αναγνώσιμα από το χρήστη. Η επιλογή του δυαδικού τύπου έγινε για την ελαχιστοποίηση του χώρου αποθήκευσης τους και για τη μεγιστοποίηση της ταχύτητας ανάγνωσης και γραφής τους. Κάθε αρχείο περιέχει τις χρονοσειρές όλων των θέσεων.

Το αρχείο τύπου INH κάτω από ορισμένες συνθήκες (βλ. επόμενο υποκεφάλαιο) αποτελεί στοιχείο εισόδου για το πρόγραμμα DDM. Συγκεκριμένα είναι δυνατό με ένα πρώτο τρέξιμο του DDM να κατασκευστεί το αρχείο μεταβλητών υψηλότερου επιπέδου INH χωρίς να γίνει επιμερισμός του και σε δεύτερο τρέξιμο του ίδιου προγράμματος να διαβαστεί το αρχείο αυτό, να επιμεριστεί και να κατασκευαστεί το αρχείο τύπου OUL που περιέχει και τις μεταβλητές χαμηλότερου επιπέδου. Πάντως η κατασκευή του αρχείου τύπου INH δεν είναι γενικά απαραίτητη.

Το αρχείο τύπου OUL κατασκευάζεται από το DDM και διαβάζεται από το DDMTEST. Το τελευταίο πρόγραμμα δίνει τη δυνατότητα στο χρήστη να υπολογίσει στατιστικές παραμέτρους των χρονοσειρών αυτού του αρχείου, να δεί τις εγγραφές του και, ακόμη, να κάνει μικρές τροποποιήσεις σε αυτές.

Λεπτομέρειες για τη δομή των δύο τύπων αρχείων συνθετικών χρονοσειρών περιέχονται στο Παράρτημα 3.

3.4 Λειτουργία του προγράμματος DDM

Η κλήση του προγράμματος γίνεται, αφού πρώτα οδηγήσουμε τον έλεγχο του DOS σε εκείνο τον κατάλογο (directory) όπου υπάρχουν τά αρχεία εισόδου, με την ακόλουθη εντολή του DOS, στην οποία όσα στοιχεία είναι μέσα σε τετράγωνες αγγύλες ([ ]) είναι προαιρετικά:

`[διαδρομή]DDM [παράμετρος] όνομα[.INF] [> αρχείο_πληροφ]`

όπου

- `[διαδρομή]` είναι η πλήρης διαδρομή (path) του δέντρου αρχείων του δίσκου, στην οποία βρίσκεται το αρχείο `DDM.EXE`. Μπορεί να παραλειφθεί αν το αρχείο βρίσκεται στον τρέχοντα κατάλογο αρχείων (current directory).
- `[Παράμετρος]` είναι ένα προαιρετικό πεδίο που μπορεί να πάρει τις τιμές `/0`, `/1`, `/2`, `/3`, `/4` και καθορίζει την ποσότητα ενδιάμεσων πληροφοριών που θα εκτυπωθούν στην οθόνη κατά τη διάρκεια των υπολογισμών του προγράμματος. Το νόημα αυτών των τιμών είναι το εξής:

  `/0` : Χωρίς εκτύπωση ενδιάμεσων πληροφοριών

  `/1` (Προκαθορισμένη τιμή): Εκτύπωση λίγων ενδιάμεσων πληροφοριών

  `/2` .. `/4` : Εκτύπωση περισσότερων ενδιάμεσων πληροφοριών

  Το να μη δοθεί καμιά παράμετρος ισοδυναμεί με το να δοθεί η προκαθορισμένη παράμετρος `/1`.
- `όνομα[.INF]` είναι το όνομα του αρχείου γενικών πληροφοριών το οποίο ταυτίζεται με το όνομα του προβλήματος. Η κατάληξη `.INF` δεν είναι απαραίτητο να γραφεί.
- `[> αρχείο_πληροφ]` είναι ένα προαιρετικό πεδίο που γράφεται μόνο στην περίπτωση που θέλουμε να κάνουμε αναδιάταξη της τυπικής εξόδου (standard output), η οποία κανονικά είναι η οθόνη, σε ένα αρχείο και συγκεκριμένα στο αρχείο `αρχείο_πληροφ`. Το `αρχείο_πληροφ` είναι ένα οποιοδήποτε επιτρεπτό όνομα αρχείου του DOS. Τονίζεται ότι το σύμβολο > είναι υποχρεωτικό να μπεί στην περίπτωση που θέλουμε να κάνουμε αναδιάταξη. Στην περίπτωση που χρησιμοποιήσουμε τη δυνατότητα αναδιάταξης το πρακτικό αποτέλεσμα θα είναι να γραφούν όλες οι ενδιάμεσες πληροφορίες στο αρχείο `αρχείο_πληροφ` αντί να γραφούν στην οθόνη.

Αν στην εντολή κλήσης του προγράμματος γίνει κάποιο λάθος, δηλαδή δεν τηρηθούν οι πιο πάνω κανόνες, τότε το πρόγραμμα δεν θα εκτελεστεί κανονικά αλλά θα γράψει ένα μήνυμα σχετικό με την ορθή γραφή της εντολής κλήσης.

Ακολουθούν μερικά παραδείγματα κλήσης του προγράμματος με τις αντίστοιχες επεξηγήσεις:

- `DDM TEST`

  Εκτέλεση του DDM, το οποίο βρίσκεται στον τρέχοντα κατάλογο, με είσοδο πληροφοριών από τα αρχεία TEST.INF, TEST.INH και TEST.INL, τα οποία επίσης βρίσκονται στον τρέχοντα υποκατάλογο.

- `/SIMUL/DDM /0 TEST`

  Όπως στο παραπάνω παράδειγμα, αλλά το αρχείο DDM.EXE βρίσκεται στον κατάλογο /SIMUL και επιπλέον δεν ζητούνται ενδιάμεσες πληροφορίες στην οθόνη.

- `DDM /4 TEST > TEST.TXT`

  Όπως στο πρώτο παράδειγμα, αλλά με εκτύπωση όσο το δυνατόν περισσότερων πληροφοριών. Η εκτύπωση, αντί να γίνει στην οθόνη, θα γίνει στο αρχείο TEST.TXT, το οποίο θα δημιουργηθεί.

Οι πληροφορίες που τυπώνονται στην οθόνη ή στο αρχείο αναδιάταξης αφορούν σε (α) δεδομένα εισόδου που ξαναγράφονται όπως διαβάζονται από το κάθε αρχείο εισόδου, (β) αριθμητικά δεδομένα υπολογισμών που προηγούνται των διαδικασιών γέννησης των μεταβλητών (π.χ. υπολογισμοί παραμέτρων των διάφορων μοντέλων, στατιστικών ροπών κτλ.) και (γ) μηνύματα του προγράμματος.

Τα αρχεία συνθετικών χρονοσειρών που θα δημιουργηθούν από την εκτέλεση του προγράμματος ρυθμίζονται από την κατάλληλη οδηγία που υπάρχει στο αρχείο γενικών πληροφοριών (INF), η οποία παίρνει τις ακέραιες τιμές 0 έως 3. Το νόημα των τιμών αυτών προκύπτει από τον ακόλουθο πίνακα:

Τιμή οδηγίας:	0	1	2	3
Κατασκευή αρχείου INH	ΟΧΙ	ΟΧΙ	ΝΑΙ	ΝΑΙ
Χρησιμοποίηση αρχείου INH	ΟΧΙ	ΝΑΙ	ΝΑΙ	ΟΧΙ
Κατασκευή αρχείου INL	ΝΑΙ	ΝΑΙ	ΝΑΙ	ΟΧΙ

Ο χρήστης μπορεί να καθορίσει την τιμή της οδηγίας ανάλογα με το πρόβλημά του και με βάση τον παραπάνω πίνακα. Η συνηθέστερη περίπτωση αντιστοιχεί στην τιμή οδηγίας 0, δεδομένου ότι κατά κανόνα το αρχείο INH δεν είναι απαραίτητο. Αν χρειάζεται μόνο η χρονοσειρά των μεταβλητών υψηλότερου επιπέδου, τότε χρησιμοποιούμε την τιμή οδηγίας 3. Αν υπάρχει έτοιμο αρχείο INH από προηγούμενη εκτέλεση του προγράμματος και θέλουμε τη δημιουργία αρχείου INL, τότε χρησιμοποιούμε την τιμή οδηγίας 2, κοκ.

Το πρόγραμμα, αν ξεκινήσει σωστά, διαβάζει πρώτα τα τρία αρχεία εισόδου. Στη συνέχεια υπολογίζει τις βασικές παραμέτρους των δύο μοντέλων, πρώτα του μοντέλου ετήσιων τιμών και μετά του μοντέλου επιμερισμού. Αν περατωθεί σωστά αυτό το δεύτερο στάδιο τότε εισέρχεται στο τρίτο στάδιο, στο οποίο γίνεται η γέννηση των χρονοσειρών. Αν υπάρξει σφάλμα στην εκτέλεση του δεύτερου σταδίου, τότε η εκτέλεση του προγράμματος διακόπτεται και γράφεται ένα σχετικό μήνυμα στην οθόνη. Στην τελευταία περίπτωση για να λειτουργήσει σωστά το πρόγραμμα ο χρήστης θα πρέπει να κάνει επέμβαση στα αρχεία εισόδου και είτε να διορθώσει τυχόν λάθη που υπάρχουν, είτε να προσθέσει περιορισμούς στα αντίστοιχα τμήματα των αρχείων INH και INL.

3.5 Λειτουργία του προγράμματος DDMTEST

Η κλήση του προγράμματος γίνεται, αφού πρώτα οδηγήσουμε τον έλεγχο του DOS σε εκείνο τον κατάλογο (directory) όπου υπάρχουν τά αρχεία εισόδου και το αρχείο χρονοσειρών OUL, με την ακόλουθη εντολή του DOS, στην οποία όσα στοιχεία είναι μέσα σε τετράγωνες αγγύλες ([]) είναι προαιρετικά:

```
[διαδρομή] DDMTEST [παράμετρος] όνομα[.INF] [> αρχείο_πληροφ]
```

όπου

- `[διαδρομή]` είναι η πλήρης διαδρομή (path) του δέντρου αρχείων του δίσκου, στην οποία βρίσκεται το αρχείο `DDMTEST.EXE`. Μπορεί να παραλειφθεί αν το αρχείο βρίσκεται στον τρέχοντα κατάλογο αρχείων (current directory).
- `[Παράμετρος]` είναι ένα προαιρετικό πεδίο που, όπως και στο πρόγραμμα DDM, μπορεί να πάρει τις τιμές `/0`, `/1`, `/2`, `/3`, `/4` και καθορίζει την ποσότητα ενδιάμεσων πληροφοριών που θα εκτυπωθούν στην οθόνη κατά τη διάρκεια των υπολογισμών του προγράμματος. Το νόημα αυτών των τιμών είναι το εξής:

  `/0` : Χωρίς εκτύπωση ενδιάμεσων πληροφοριών

  `/1 .. /4` : Εκτύπωση ενδιάμεσων πληροφοριών

  Το να μη δοθεί καμιά παράμετρος ισοδυναμεί με το να δοθεί η προκαθορισμένη παράμετρος `/1`. Σημειώνεται ότι όλες οι παράμετροι από `/2` μέχρι `/4` είναι πλήρως ισοδύναμες με την παράμετρο `/1` στην περίπτωση του DDMTEST και έχουν προβλεφθεί ως έγγυρες παράμετροι μόνο για λόγους ομοιομορφίας και συμβατότητας με το DDM. Οι μόνες ενδιάμεσες πληροφορίες που εκτυπώνονται είναι τα δεδομένα εισόδου που ξαναγράφονται όπως διαβάζονται από το κάθε αρχείο εισόδου.
- `όνομα[.INF]` είναι το όνομα του αρχείου γενικών πληροφοριών το οποίο ταυτίζεται με το όνομα του προβλήματος. Η κατάληξη `.INF` δεν είναι απαραίτητο να γραφεί.
- `[> αρχείο_πληροφ]` είναι ένα προαιρετικό πεδίο που γράφεται μόνο στην περίπτωση που θέλουμε να κάνουμε αναδιάταξη της τυπικής εξόδου (standard output), η οποία κανονικά είναι η οθόνη, σε ένα αρχείο και συγκεκριμένα στο αρχείο `αρχείο_πληροφ`. Το `αρχείο_πληροφ` είναι ένα οποιοδήποτε επιτρεπτό όνομα αρχείου του DOS. Τονίζεται ότι το σύμβολο > είναι υποχρεωτικό να μπεί στην περίπτωση που θέλουμε να κάνουμε αναδιάταξη. Στην περίπτωση που χρησιμοποιήσουμε τη δυνατότητα αναδιάταξης το πρακτικό αποτέλεσμα θα είναι να γραφούν όλες οι ενδιάμεσες πληροφορίες στο αρχείο `αρχείο_πληροφ` αντί να γραφούν στην οθόνη.

Αν στην εντολή κλήσης του προγράμματος γίνει κάποιο λάθος, δηλαδή δεν τηρηθούν οι πιο πάνω κανόνες, τότε το πρόγραμμα δεν θα εκτελεστεί κανονικά αλλά θα γράψει ένα μήνυμα σχετικό με την ορθή γραφή της εντολής κλήσης.

Ακολουθούν μερικά παραδείγματα κλήσης του προγράμματος με τις αντίστοιχες επεξηγήσεις:

- `DDMTEST TEST`

  Εκτέλεση του DDMTEST, το οποίο βρίσκεται στον τρέχοντα κατάλογο, με είσοδο πληροφοριών από τα αρχεία TEST.INF, TEST.INH και TEST.INL, και είσοδο συνθετικών δεδομένων από το αρχείο TEST.OUL. Όλα αυτά τα αρχεία πρέπει να βρίσκονται στον τρέχοντα υποκατάλογο.
- `/SIMUL/DDMTEST /0 TEST`

  Όπως στο παραπάνω παράδειγμα αλλά το αρχείο DDMTEST.EXE βρίσκεται στον κατάλογο /SIMUL και επιπλέον δεν ζητούνται ενδιάμεσες πληροφορίες στην οθόνη.
- `DDMTEST TEST > TEST.RES`

  Όπως στο πρώτο παράδειγμα, αλλά η εκτύπωση, αντί να γίνει στην οθόνη, θα γίνει στο αρχείο TEST.RES, το οποίο θα δημιουργηθεί.

Όταν το πρόγραμμα ξεκινήσει θα διαβάσει κατ' αρχήν τα τρία αρχεία εισόδου (αν και δεν χρησιμοποιεί όλα τα δεδομένα που περιέχονται στα αρχεία αυτά) καθώς και την επικεφαλίδα του αρχείου OUL και θα γράψει ένα αρχικό μήνυμα όπως το ακόλουθο:

```
ΠΡΟΓΡΑΜΜΑ ΕΛΕΓΧΟΥ ΣΥΝΘΕΤΙΚΩΝ ΑΡΧΕΙΩΝ
Συνολικός αριθμός δεδομένων (εγγραφών):    100
Μήκος μιας χρονοσειράς                     100  100
Αριθμός χρονοσειρών                          1    1
Αριθμός θέσεων                               2    2
Αριθμός υποπεριόδων                         12   12
```

Το πρώτο αριθμητικό δεδομένο που γράφεται (ο συνολικός αριθμός των εγγραφών) προκύπτει από το μήκος του αρχείου OUL. Τα υπόλοιπα αριθμητικά δεδομένα γράφονται σε δύο στήλες: η πρώτη προκύπτει από τα αρχεία εισόδου και η δεύτερη από την επικεφαλίδα του αρχείου OUL. Εφόσον υπάρχει συμφωνία μεταξύ των δεδομένων των δύο στηλών καθώς και συμφωνία του συνολικού αριθμού δεδομένων με το μήκος μιας χρονοσειράς και τον αριθμό των χρονοσειρών (το γινόμενο των δύο τελευταίων θα πρέπει να είναι ίσο με το συνολικό αριθμό δεδομένων), τότε το πρόγραμμα θα συνεχίσει, αλλιώς θα τερματιστεί γράφοντας ένα σχετικό μήνυμα.

Από το σημείο αυτό και μετά το πρόγραμμα (εφόσον συνεχίσει) παρουσιάζει συνεχώς τον ακόλουθο κατάλογο επιλογών και τερματίζεται μόνο αν επιλεγεί η τελευταία επιλογή (αρ. 0).

ΕΡΓΑΣΙΑ : 1 = Εκτύπωση στατιστικών χαρακτηριστικών των μεταβλητών Υ.Ε.
2 = Εκτύπωση στατιστικών χαρακτηριστικών των μεταβλητών Χ.Ε.
3 = Εκτύπωση συσχετίσεων μεταξύ των μεταβλητών Χ.Ε. και Υ.Ε.
4 = Εκτύπωση πλήρους σειράς αυτοσυσχ/σεων μεταξύ μεταβλ. Χ.Ε.
5 = Δοκιμή Κολμογκόροφ - Σμίρνοφ
6 = Εκτύπωση συχνοτήτων
7 = Ανάγνωση εγγραφών / Επέμβαση στο αρχείο
0 = Τέλος προγράμματος
Σημείωση: Υ.Ε. = Υψηλότερο Επίπεδο (ετήσια κλίμακα)
Χ.Ε = Χαμηλότερο Επιπέδο (πχ. μηνιαία κλίμακα)

Οι τέσσερεις πρώτες επιλογές οδηγούν στην εκτύπωση διάφορων στατιστικών χαρακτηριστικών των χρονοσειρών χαμηλότερου ή υψηλότερου επιπέδου. Οι δύο επόμενες (5 και 6) οδηγούν είτε σε στατιστική σύγκριση με τη δοκιμή Κολμογκόροφ-Σμίρνοφ των θεωρητικών και εμπειρικών συχνοτήτων μιας χρονοσειράς υψηλότερου ή χαμηλότερου επιπέδου, είτε σε απλή εκτύπωση των εμπειρικών και θεωρητικών συχνοτήτων. Τέλος με την επιλογή 7 είναι δυνατή η ανάγνωση των διάφορων εγγραφών του αρχείου και η επέμβαση (μεταβολή ή διαγραφή) σε αυτές.

4. ΣΥΝΤΑΞΗ ΤΩΝ ΑΡΧΕΙΩΝ ΕΙΣΟΔΟΥ ΓΙΑ ΤΗ ΛΕΙΤΟΥΡΓΙΑ ΤΩΝ ΠΡΟΓΡΑΜΜΑΤΩΝ

4.1 Γενικά

Στο κεφάλαιο αυτό αντί γενικών οδηγιών για τη σύνταξη των αρχείων εισόδου προτιμήθηκε να δοθεί ένα παράδειγμα αρχείων που αντιστοιχεί στις ετήσιες και μηνιαίες απορροές του Μόρνου και της Υλίκης (δύο θέσεις - 12 υποπερίοδοι). Τα αρχεία αυτά περιέχονται και στη δισκέτα των προγραμμάτων. Αν και τα δεδομένα που υποχρεωτικά περιέχονται σε αυτά τα αρχεία είναι μόνο αριθμητικά, είναι επιτρεπτό (και συστήνεται) να μπαίνουν και γραμμές κειμένου με επεξηγηματικά σχόλια. Οι γραμμές αυτές θα πρέπει υποχρεωτικά να ξεκινούν με το σύμβολο #.

Τα αρχεία που δίνονται παρακάτω περιέχουν τέτοια σχόλια που δίνουν τις απαραίτητες επεξηγήσεις. Προσθέτουμε μόνο ότι οι περιορισμοί που μπαίνουν στις τελευταίες ενότητες των αρχείων INH και INL είναι προαιρετικοί. Αν χρειαστεί να μπουν τέτοιοι περιορισμοί τότε για κάθε υποπερίοδο (ή για την περίοδο) εισάγονται δύο αριθμοί. Ο πρώτος παίρνει τις τιμές 0 ή 1, όπου η τιμή 0 σημαίνει ότι το πρόγραμμα θα πρέπει να λάβει υπόψη όλους τους συντελεστές ετεροσυσχέτισης, ενώ η τιμή 1 σημαίνει ότι το πρόγραμμα θα λάβει υπόψη και θα διατηρήσει μόνο το μεγαλύτερο από τους συντελεστές ετεροσυσχέτισης. Ο δεύτερος αριθμός είναι ο συντελεστής R με τιμές στο διάστημα από 0 μέχρι 1 και αφορά στην αποσύνθεση του μητρώου παραμέτρων $\underline{b}^t\,(\underline{b}^t)^T$ (βλ. Παραρτήματα 2 και 3) Η τιμή 1 δεν θέτει κανένα περιορισμό, ενώ η τιμή 0 ουσιαστικά καθιστά τις μεταβλητές των διάφορων θέσεων ασυσχέτιστες. Οι ενδιάμεσες τιμές δίνουν τη δυνατότητα στο πρόγραμμα να περιορίσει τις τιμές των συντελεστών ετεροσυσχέτισης αν διαπιστωθεί πρόβλημα κατά την αποσύνθεση των μητρώων. Αναλυτικότερα στοιχεία για τους περιορισμούς και το νόημα των παραπάνω μεγεθών δίνονται στα Παραρτήματα 1 και 2. Γενικά συστήνεται στο χρήστη να ξεκινήσει την επίλυση του προβλήματος χωρίς να θέσει περιορισμούς και αν διαπιστώσει ότι υπάρχει πρόβλημα (το οποίο θα επισημανθεί κατάλληλα από το πρόγραμμα DDM) τότε να δοκιμάσει διάφορους περιορισμούς μέχρι να δοθεί ικανοποιητική λύση στο πρόβλημα.

4.2 Παράδειγμα αρχείου γενικών πληροφοριών (TEST.INF)

```
# ΓΕΝΙΚΑ ΔΕΔΟΜΕΝΑ
# Αριθμός χρονοσειρών
1
# Μήκος κάθε χρονοσειράς (αριθμός περιόδων)
100
# Αριθμός θέσεων
2
# Αριθμός υποπεριόδων κάθε περιόδου
12
# Ξεχωριστό αρχείο δεδομένων υψηλότερου επιπέδου (περιόδων)
# (0 = ΔΕΝ ΥΠΑΡΧΕΙ - ΝΑ ΜΗ ΔΗΜΙΟΥΡΓΗΘΕΙ,
#  1 = ΥΠΑΡΧΕΙ - ΝΑ ΧΡΗΣΙΜΟΠΟΙΗΘΕΙ,
#  2 = ΔΕΝ ΥΠΑΡΧΕΙ - ΝΑ ΔΗΜΙΟΥΡΓΗΘΕΙ ΚΑΙ ΝΑ ΓΙΝΕΙ ΚΑΙ ΕΠΙΜΕΡΙΣΜΟΣ,
#  3 = ΔΕΝ ΥΠΑΡΧΕΙ - ΝΑ ΔΗΜΙΟΥΡΓΗΘΕΙ ΚΑΙ ΝΑ ΜΗ ΓΙΝΕΙ ΕΠΙΜΕΡΙΣΜΟΣ)
2
# Συνθήκη αρνητικών τιμών
# 0 = ΕΠΙΤΡΕΠΟΝΤΑΙ ΑΡΝΗΤΙΚΕΣ ΤΙΜΕΣ, 1 = ΔΕΝ ΕΠΙΤΡΕΠΟΝΤΑΙ ΑΡΝΗΤΙΚΕΣ ΤΙΜΕΣ
1
# Αρχική τιμή γεννήτριας τυχαίων αριθμών (random seed)
1
# Τέλος (υποχρεωτικά ο αριθμός 99)
99
```

4.3 Παράδειγμα αρχείου στατιστικών δεδομένων μεταβλητών υψηλότερου επιπέδου (TEST.INH)

```
# ΣΤΑΤΙΣΤΙΚΑ ΔΕΔΟΜΕΝΑ ΠΕΡΙΟΔΩΝ (ΕΤΗΣΙΑ)
# Τύπος κατανομής (0 = Gauss, 1 = Γάμα)
1
# Μέσες Τιμές
# ΜΟΡΝΟΣ    ΥΛΙΚΗ
# απορροή   απορροή (σε mm)
 543.2   138.8
# Διασπορές
 25995.0 2840.8
# Μορφή με την οποία δίνονται οι τρίτες ροπές:
# (0 = δεν δίνονται, 1 = δίνονται τα μ3, 2 = δίνονται τα Cs,
#  3 = δίνονται τα Cs/Cv)
# Σημείωση: Cs = συντ. ασυμμετρίας, Cv = συντ. διασποράς
0
# Τρίτες Ροπές - Μόνο αν πρέπει να δοθούν

# Συντελεστές αυτοσυσχέτισης
 0.21 0.02
# Συντελεστές ετεροσυσχέτισης
# Εισάγονται τα στοιχεία μέχρι τη διαγώνιο
1.00
0.87 1.00
# Περιορισμοί (Προαιρετικοί, 0 1 = όχι περιορισμοί)
# 0 0.95
# Τέλος (υποχρεωτικά ο αριθμός 99)
99
```

4.4 Παράδειγμα αρχείου στατιστικών δεδομένων μεταβλητών χαμηλότερου επιπέδου (TEST.INL)

```
# ΜΗΝΙΑΙΑ ΔΕΔΟΜΕΝΑ
# Τύπος κατανομής (0 = Gauss, 1 = Γάμα)
1
# Μέσες Τιμές
# Μία γραμμή αντιστοιχεί στις διάφορες θέσεις της ίδιας υποπεριόδου
# ΜΟΡΝΟΣ    ΥΛΙΚΗ
# απορροή   απορροή
  20.2       8.2
  54.3      10.5
  88.3      18.8
  83.9      21.9
  79.0      21.7
  68.0      26.5
  57.6      17.0
  41.3       7.9
  22.7       1.5
  11.2       0.2
   8.7       0.5
   8.0       4.0
# Διασπορές
# Μία γραμμή αντιστοιχεί στις διάφορες θέσεις της ίδιας υποπεριόδου
  554.9   19.6
 1574.7   17.9
 2697.7  238.3
 3458.5  151.2
 2276.3   94.3
  713.4   84.2
  344.7   95.2
  385.4   40.6
  114.2    5.3
   43.2    0.2
   29.1    1.2
   26.0    8.8
# Μορφή με την οποία δίνονται οι τρίτες ροπές:
# (0 = δεν δίνονται, 1 = δίνονται τα μ3, 2 = δίνονται τα Cs,
#  3 = δίνονται τα Cs/Cv)
# Σημείωση: Cs = συντ. ασυμμετρίας, Cv = συντ. διασποράς
3
# Τρίτες Ροπές - Μόνο αν πρέπει να δοθούν
# Μία γραμμή αντιστοιχεί στις διάφορες θέσεις της ίδιας υποπεριόδου
2    2
2    2
2    2
2    2
2    2
2    2
2    1.3
2    1.3
2    2
2    1.3
2    1.3
2    2
```

```
# Συντελεστές αυτοσυσχέτισης
# Μία γραμμή αντιστοιχεί στις διάφορες θέσεις της ίδιας υποπεριόδου
 0.18 0.64
 0.30 0.69
 0.00 0.45
 0.23 0.59
 0.42 0.63
 0.37 0.42
 0.79 0.80
 0.83 0.82
 0.56 0.49
 0.00 0.56
 0.22 0.47
 0.77 0.54
# Συντελεστές ετεροσυσχέτισης
# Μία γραμμή αντιστοιχεί στους συντελεστές μιας θέσης με τις άλλες
# Εισάγονται τα στοιχεία μέχρι τη διαγώνιο
# υποπερίοδος 1
1.00
0.69 1.00

#2
1.00
0.32 1.00

#3
1.00
0.55 1.00

#4
1.00
0.81 1.00

#5
1.00
0.70 1.00

#6
1.00
0.77 1.00

#7
1.00
0.79 1.00

#8
1.00
0.79 1.00

#9
1.00
0.19 1.00

#10
 1.00
 0.00 1.00
```

```
#11
 1.00
 0.76 1.00

#12
 1.00
 0.76 1.00

# Περιορισμοί (Προαιρετικοί, 0 1 = όχι περιορισμός)
0 1
0 1
0 1
0 1
0 1
0 1
0 0.85
0 0.85
0 1
0 1
0 1
0 0.85
# Τέλος (υποχρεωτικά ο αριθμός 99)
99
```

ΑΝΑΦΟΡΕΣ

Box, G.E.P. & Jenkins, G.M. *Time Series Analysis; Forecasting and control*, Holden Day, 1970.

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

Eckhardt, R., Stan Ulam, John von Neumann and the Monte Carlo Method, in *From cardinals to chaos*, edited by N.G. Cooper, Cambrige University, NY, USA, 1989.

Grygier, J.C. & 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.

Matalas, N.C. and Wallis, J.R., Generation of synthetic flow sequences, in *Systems approach to water management*, A.K. Biswas editor, McGraw Hill, 1976.

Metropolis, N., The beginning of the Monte Carlo method, in *From cardinals to chaos*, edited by N.G. Cooper, Cambrige University, NY, USA, 1989.

Vogel, R.M. & Stedinger, J.R., The value of stochastic streamflow models in over-year reservoir design applications, *Water Resour. Res.* 24(9), 1483-90, 1988.

ΠΑΡΑΡΤΗΜΑ 1

ΓΕΝΙΚΗ ΠΕΡΙΓΡΑΦΗ ΤΟΥ ΣΕΙΡΙΑΚΟΥ ΜΑΡΚΟΒΙΑΝΟΥ ΜΟΝΤΕΛΟΥ ΠΟΛΛΩΝ ΜΕΤΑΒΛΗΤΩΝ

Στις περισσότερες πρακτικές εφαρμογές της υδρολογικής προσομοίωσης μας απασχολεί η ταυτόχρονη γέννηση πολλών χρονοσειρών στοχαστικά εξαρτημένων μεταξύ τους. Οι χρονοσειρές αυτές κατά κανόνα αναφέρονται στην ίδια υδρολογική μεταβλητή αλλά σε πολλές θέσεις στο χώρο, είναι όμως δυνατό να αναφέρονται σε διαφορετικές στοχαστικά εξαρτημένες υδρολογικές μεταβλητές σε μία ή περισσότερες θέσεις. Από μαθηματική άποψη το σύνολο των χρονοσειρών που ενδιαφέρει μπορεί να παρασταθεί από ένα διάνυσμα στοχαστικών ανελίξεων με n διαστάσεις:

$$\underline{X}^t = \left[X_1^t, X_2^t, \cdots, X_n^t\right]^T \tag{1}$$

όπου ο άνω δείκτης t συμβολίζει το χρόνο ο οποίος θεωρείται διακριτός. Κατά κανόνα οι μεταβλητές X_j^t, εκτός του ότι είναι στοχαστικά εξαρτημένες από θέση σε θέση, είναι και εξαρτημένες ως προς το χρόνο. Η απλούστερη εξάρτηση ως προς το χρόνο περιγράφεται από τη σχέση

$$\underline{X}^t = \underline{\underline{a}}^t \underline{X}^{t-1} + \underline{\underline{b}}^t \underline{V}^t \tag{2}$$

όπου $\underline{\underline{a}}^t$ και $\underline{\underline{b}}^t$ είναι μητρώα συντελεστών διαστάσεων ($n \times n$) και $\underline{V}^t = \left[V_j^t\right]$ είναι διάνυσμα n τυχαίων μεταβλητών, πλήρως ανεξάρτητων και ως προς το χρόνο και ως προς τη θέση. Υποθέτουμε ότι τα $\underline{\xi}^t = E\left[\underline{X}^t\right]$ και $\underline{\beta}^t = E\left[\underline{V}^t\right]$ δεν είναι απαραιτήτως μηδενικά. Θέτουμε (για λόγους μαθηματικής άνεσης) $\mathrm{Var}\left[V^t\right] = \underline{1}$ και συμβολίζουμε το διάνυσμα των τρίτων ροπών των μεταβλητών V με $\underline{\gamma}^t = \mu_3\left[\underline{V}^t\right]$. Η στασιμότητα της ακολουθίας των $\underline{X}^t$ δεν είναι απαραίτητη και γι' αυτό δεν υπονοείται στο συμβολισμό. Ωστόσο, αν αυτή υπάρχει, διευκολύνει τους υπολογισμούς. Το στοχαστικό μοντέλο που ορίζεται από την εξίσωση (2) είναι γνωστό ως μοντέλο αυτοπαλινδρόμησης τάξης 1 ή AR(1) (autoreggressive lag-one). Το μοντέλο απλοποιείται αν το μητρώο $\underline{\underline{a}}^t$ οριστεί ως διαγώνιο, ήτοι

$$\underline{\underline{a}}^t = \operatorname{diag}\left(a_1^t, a_2^t, \cdots, a_n^t\right) \tag{3}$$

Στην τελευταία περίπτωση το μοντέλο που προκύπτει είναι γνωστό και ως Μαρκοβιανό [*Matalas & Wallis*, 1976, σ. 63]. Εξ άλλου το μητρώο $\underline{\underline{b}}^t$ μπορεί σε κάθε περίπτωση να θεωρηθεί κάτω τριγωνικό. Έτσι η εξίσωση (2) μπορεί να γραφεί και με την ακόλουθη μορφή

$$X_j^t = a_j^t X_j^{t-1} + \sum_{r=1}^{j} b_{jr}^t V_r^t, \quad j = 1, \cdots, n \tag{4}$$

Το Μαρκοβιανό μοντέλο διατηρεί (δηλαδή αναπαράγει) τις ακόλουθες ομάδες στατιστικών παραμέτρων των μεταβλητών X_j^t [*Matalas & Wallis*, 1976, σ. 60, 63]:

(1) Μέσες τιμές των X_j^t (ξ_j^t).

(2) Διασπορές των X_j^t.

(3) Συντελεστές ασυμμετρίας των X_j^t.

(4) Συντελεστές αυτοσυσχέτισης (με βήμα ένα) μεταξύ των X_j^t και X_j^{t-1} (ίδια θέση).

(5) Συντελεστές ετεροσυσχέτισης (με μηδενικό βήμα) μεταξύ των X_j^t και X_r^t, για $j \neq r$ (ίδια περίοδος).

Πρόκειται για το ελάχιστο σύνολο ουσιωδών στατιστικών παραμέτρων που κατά κανόνα ενδιαφέρουν. Σε περίπτωση που ενδιαφέρουν και άλλες ομάδες παραμέτρων, π.χ. συντελεστές αυτοσυσχέτισης για μεγαλύτερα χρονικά βήματα, θα πρέπει να χρησιμοποιείται μοντέλο αυτοπαλινδρόμησης μεγαλύτερης τάξης ή μοντέλο τύπου αυτοπαλινδρόμησης - κινούμενων μέσων όρων (ARMA). Τα τελευταία είναι αρκετά πιο πολύπλοκα από το Μαρκοβιανό μοντέλο.

Οι ομάδες παραμέτρων $\underline{a}^t$, $\underline{b}^t$ και $\underline{\gamma}^t$ του μοντέλου υπολογίζονται από τις παραπάνω ομάδες στατιστικών παραμέτρων (2) έως (5) με τις ακόλουθες σχέσεις που αποτελούν γενικεύσεις των σχέσεων της στάσιμης περίπτωσης, όπως αυτές δίνονται από τους *Matalas & Wallis* [1976, σ. 63]

$$a_j^t = \frac{\mathrm{Cov}\left[X_j^t, X_j^{t-1}\right]}{\mathrm{Var}\left[X_j^{t-1}\right]} \tag{5}$$

$$\underline{b}^t \left(\underline{b}^t\right)^T = \underline{\sigma}^t - \underline{a}^t \underline{\sigma}^{t-1} \underline{a}^t \tag{6}$$

$$\gamma_j^t = \frac{\mu_3\left[X_j^t\right] - \left(a_j^t\right)^3 \mu_3\left[X_j^{t-1}\right] - \sum_{k=1}^{j-1} \left(b_{jk}^t\right)^3 \gamma_k^t}{\left(b_{jj}^t\right)^3} \tag{7}$$

όπου

$$\underline{\sigma}^t = \mathrm{Cov}\left[\underline{X}^t, \underline{X}^t\right] = E\left[\left(\underline{X}^t - \underline{\xi}^t\right)\left(\underline{X}^t - \underline{\xi}^t\right)^T\right] \tag{8}$$

Η εξίσωση (6) δεν δίνει άμεσα το μητρώο $\underline{b}^t$ αλλά μια δευτεροβάθμια μορφή του. Το μητρώο

$$\underline{c}^t = \underline{b}^t \left(\underline{b}^t\right)^T \tag{9}$$

είναι γνωστό ως Γκραμιανό (Gramian) μητρώο του $\underline{b}^t$. Αποδεικνύεται ότι, αν υπάρχει μια λύση της (9) ως προς $\underline{b}^t$, τότε υπάρχουν άπειρες λύσεις, δεδομένου ότι κάθε μητρώο της μορφής $\underline{b}^t \underline{d}$, όπου $\underline{d}$ τυχόν ορθογωνικό μητρώο (ήτοι $\underline{d}\,\underline{d}^T = \underline{I}$) αποτελεί λύση της (9). Από τις άπειρες αυτές λύσεις η ευνοϊκότερη από άποψη μαθηματικών χειρισμών είναι η λύση του κάτω τριγωνικού μητρώου. Η λύση αυτή δίνεται από τις εξισώσεις

$$b_{ij}^t = 0 \qquad (i < j) \tag{10α}$$

$$b_{ij}^t = \sqrt{c_{jj}^t - \sum_{k=1}^{j-1} \left(b_{jk}^t\right)^2} \qquad (i = j) \tag{10β}$$

$$b_{ij}^t = \frac{c_{ij}^t - \sum_{k=1}^{j-1} b_{jk}^t b_{ik}^t}{b_{jj}^t} \qquad (i > j) \tag{10γ}$$

Η διαδικασία υπολογισμού του μητρώου με επαναληπτική εφαρμογή των παραπάνω εξισώσεων είναι γνωστή ως αποσύνθεση του μητρώου $\underline{c}^t$. Είναι προφανές ότι η εξίσωση (10β) επιβάλλει περιορισμούς για την ύπαρξη λύσης της εξίσωσης (6). Σύμφωνα με τη μαθηματική ορολογία, η ύπαρξη λύσης προϋποθέτει να είναι θετικά ορισμένο το μητρώο $\underline{c}^t$. Εξ άλλου, το μητρώο $\underline{c}^t$ αποτελεί μητρώο συνδιασπορών, ήτοι

$$\underline{c}^t = \mathrm{Cov}\left[\left(\underline{X}^t - \underline{a}^t \underline{X}^{t-1}\right), \left(\underline{X}^t - \underline{a}^t \underline{X}^{t-1}\right)\right] \tag{11}$$

πράγμα που επιβάλλει και ένα δεύτερο περιορισμό στα στοιχεία του και συγκεκριμένα

$$-1 \le \frac{c_{ij}^t}{\sqrt{c_{ii}^t c_{jj}^t}} \le 1 \tag{12}$$

Είναι δυνατό οι δύο παραπάνω περιορισμοί (και ιδίως ο πρώτος) να μην ικανοποιούνται για συγκεκριμένα υδρολογικά δεδομένα, τα οποία χρησιμοποιούνται για τον υπολογισμό του $\underline{c}^t$ βάσει των (5) και (6). Επί πλέον, η χρήση της (7) μπορεί να οδηγήσει σε υπερβολικά μεγάλους συντελεστές ασυμμετρίας, δεδομένου ότι ο παρονομαστής στην (7) μπορεί να πάρει πολύ μικρές τιμές. Τα προβλήματα αυτά απαντούν κυρίως σε περιπτώσεις όπου τα υδρολογικά δεδομένα των διάφορων σταθμών αναφέρονται σε διαφορετικές περιόδους μέτρησης. Πρακτικές λύσεις για την αντιμετώπιση τέτοιων προβλημάτων μπορεί να είναι η μείωση των συντελεστών αυτοσυσχέτισης ή ετεροσυσχέτισης και η μείωση των συντελεστών ασυμμετρίας των ιστορικών υδρολογικών δεδομένων.

Σε μια παρόμοια θέση που αφορούσε δείγματα με διαφορετικά μήκη σε διαφορετικές θέσεις, οι *Grygier & Stedinger* [1990, σ. 31-33] συζητούν πρακτικές λύσεις σε ανάλογα προβλήματα, προτείνοντας τροποποιήσεις στα στοιχεία του προς αποσύνθεση Γκραμιανού μητρώου, όταν η διαδικασία της αποσύνθεσης αποτυγχάνει. Εδώ χρησιμοποιείται μια απλούστερη τεχνική. Πιο συγκεκριμένα, η διαδικασία αποσύνθεσης του $\underline{c}^t$ μπορεί να αποτύχει στο σημείο που πρόκειται να υπολογιστεί το διαγώνιο στοιχείο b_{jj}^t με βάση την εξίσωση (10β). Αν το δεξιό μέλος της εξίσωσης προκύψει αρνητικό, τότε δεν υπάρχει πραγματική λύση για το b_{jj}^t. Για να το αποφύγουμε αυτό μπορούμε να θέσουμε ένα κατώτερο όριο σε κάθε διαγώνιο στοιχείο b_{jj}^t έτσι ώστε

$$\left(b_{jj}^{t\ \min}\right)^2 = \left(1 - R^2\right) c_{jj}^t \tag{13}$$

όπου R είναι σταθερά ($0 \leq R < 1$). Αν το b_{jj}^t τεθεί ίσο με αυτό το όριο, τότε τα εκτός διαγωνίου στοιχεία της σειράς (b_{jk}^t, $k = 1, \ldots, j-1$) πρέπει να διορθωθούν με πολλαπλασιασμό επί ένα ενιαίο συντελεστή λ, υπολογισμένο έτσι ώστε να επανακτηθεί η ισχύς της (10β). Εύκολα μπορεί να δειχτεί ότι

$$\lambda = R \sqrt{c_{jj}^t \Big/ \sum_{k=1}^{j-1} \left(b_{jk}^t\right)^2} \tag{14}$$

Με αυτό τον τρόπο εξασφαλίζεται η διατήρηση της Var[X_j^t] και παράλληλα τίθεται ένα ελάχιστο όριο για την Var [V_j^t]. 'Ενα έμμεσο όφελος από αυτή τη διορθωτική επέμβαση είναι και το γεγονός ότι μειώνεται η ασυμμετρία της V_j^t και έτσι αποφεύγονται οι παράδοξα μεγάλες τιμές της. Ωστόσο, μειονέκτημα της λύσης αυτής είναι η μείωση των πραγματικών συντελεστών ετεροσυσχέτισης της X_j^t με μεταβλητές άλλων θέσεων. Μια διαφορετική λύση στο ίδιο πρόβλημα μπορεί να είναι η τροποποίηση του μητρώου $\underline{b}^t$ σε τρόπο ώστε να διατηρείται μόνο ο μεγαλύτερος συντελεστής ετεροσυσχέτισης, δηλαδή να υπάρχει μόνο ένα μη μηδενικό μη διαγώνιο στοιχείο σε κάθε σειρά του $\underline{b}^t$. Αυτό έχει συνέπεια την περαιτέρω μείωση του αριθμού των παραμέτρων του μοντέλου.

ΑΝΑΦΟΡΕΣ

Grygier, J.C. & 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.

Matalas, N.C. and Wallis, J.R., Generation of synthetic flow sequences, in *Systems approach to water management*, A.K. Biswas editor, McGraw Hill, 1976.

ΠΑΡΑΡΤΗΜΑ 2:

ΤΟ ΔΥΝΑΜΙΚΟ ΜΟΝΤΕΛΟ ΕΠΙΜΕΡΙΣΜΟΥ

1. ΕΙΣΑΓΩΓΗ

Τα μοντέλα επιμερισμού αποτελούν μια από τις πλέον διαδεδομένες τεχνικές γέννησης συνθετικών χρονοσειρών της στοχαστικής υδρολογίας [*Bras & Rodriguez-Iturbe*, 1984] και ενδεχομένως συνιστούν την καλύτερη διαθέσιμη μέθοδο για το σκοπό αυτό [*Lin*, 1990]. Ας σημειωθεί ότι δύο από τα πλέον διαδεδομένα διεθνώς σύγχρονα πακέτα προγραμμάτων γέννησης συνθετικών σειρών, το LAST [*Lane & Frevert*, 1990] και το SPIGOT [*Grygier & Stedinger*, 1990] έχουν βασιστεί εξ ολοκλήρου σε τεχνικές επιμερισμού.

Η κύρια διαφορά των μοντέλων επιμερισμού από τα σειριακά ή άμεσα μοντέλα βρίσκεται στο γεγονός ότι τα πρώτα, αντί να γεννούν τις συνθετικές τιμές στη σειρά, τη μια μετά την άλλη, επιμερίζουν γνωστές τιμές αθροιστικών μεταβλητών (ή μεταβλητών υψηλότερου επιπέδου, π.χ. ετήσιων απορροών) σε τιμές τμηματικών μεταβλητών (ή μεταβλητών χαμηλότερου επιπέδου, π.χ. εποχικών ή μηνιαίων απορροών). Έτσι έχουν το πλεονέκτημα να μετασχηματίζουν μια χρονοσειρά από μια υψηλότερη χρονική κλίμακα σε μια χαμηλότερη, παρέχοντας τη δυνατότητα διατήρησης των στατιστικών χαρακτηριστικών των χρονοσειρών σε πολλαπλή χρονική κλίμακα. Ωστόσο, τα μοντέλα επιμερισμού είναι πολύ πιο πολύπλοκα από τα σειριακά και έχουν και κάποια άλλα μειονεκτήματα που θα συζητηθούν παρακάτω.

Το γραμμικό μοντέλο επιμερισμού στην αρχική του διατύπωση αναπτύχθηκε από τους *Valencia & Schaake* [1972, 1973]. Το μοντέλο αυτό εξασφαλίζει τη διατήρηση των διασπορών και συνδιασπορών ανάμεσα στις ιστορικές και τις συνθετικές χρονοσειρές. Η διατήρηση αυτή αφορά μεταβλητές που περιέχονται στο εσωτερικό ενός χρονικού βήματος υψηλότερου επιπέδου, δηλαδή τις τμηματικές μεταβλητές που συνιστούν μια αθροιστική μεταβλητή. Ωστόσο, δεν διατηρεί τις συνδιασπορές των μεταβλητών χαμηλότερου επιπέδου που ανήκουν σε διαφορετικά (διαδοχικά) χρονικά βήματα υψηλότερου επιπέδου. Στην πραγματικότητα, αυτό το μοντέλο δεν υποθέτει καμιά σύνδεση των μεταβλητών χαμηλότερου επιπέδου που ανήκουν σε διαφορετικά χρονικά βήματα υψηλότερου επιπέδου, πράγμα αντίθετο στη φυσική πραγματικότητα.

Μετά την ανάπτυξη του πρώτου μοντέλου, προτάθηκαν διαφορετικές δομές μοντέλων και διαφορετικές τεχνικές εκτίμησης παραμέτρων με στόχο τη διατήρηση των συνδιασπορών μεταξύ μεταβλητών που ανήκουν σε διαδοχικά χρονικά βήματα υψηλότερου επιπέδου. Έτσι αναπτύχθηκαν τα μοντέλα των *Mejia & Rousselle* [1976], *Hoshi & Burges* [1979] και *Stedinger & Vogel* [1984]. Ωστόσο, όπως έδειξαν οι *Stedinger & Vogel* [1984], η ακριβής αναπαραγωγή των ιστορικών συντελεστών συσχέτισης ανάμεσα στις μεταβλητές χαμηλότερου επιπέδου είναι αδύνατη, εξαιτίας δομικών περιορισμών των μοντέλων επιμερισμού.

Στην περίπτωση που όλες οι μεταβλητές ακολουθούν πολυδιάστατη κανονική κατανομή, το βασικό γραμμικό σχήμα επιμερισμού των Valencia-Schaake εξασφαλίζει επίσης και τη διατήρηση των στατιστικών ροπών υψηλότερης τάξης, καθώς επίσης και των πλήρων συναρτήσεων κατανομής. Διαφορετικά, το μοντέλο στην πρώτη του διατύπωση δεν έχει τη δυνατότητα να διατηρεί τέτοιες ιδιότητες. Για το λόγο αυτό έχουν διατυπωθεί τροποποιήσεις του μοντέλου που επιτρέπουν τη διατήρηση μη κανονικών κατανομών. Οι τροποποιήσεις αυτές ανήκουν σε δύο κατηγορίες. Οι μέθοδοι της πρώτης κατηγορίας προσανατολίζονται στη διατήρηση μόνο της ασυμμετρίας των κατανομών των μεταβλητών χαμηλότερου επιπέδου [*Tao & Delleur*, 1976, *Todini*, 1980]. 'Ετσι, γεννούν ανεξάρτητες τυχαίες συνιστώσες που έχουν ασύμετρες κατανομές των οποίων οι συντελεστές ασυμμετρίας εκτιμώνται με βάση τις τρίτες ροπές των ιστορικών δεδομένων. Οι μέθοδοι της δεύτερης κατηγορίας χρησιμοποιούν μη γραμμικούς μετασχηματισμούς των δεδομένων, σε τρόπο ώστε τα μετασχηματισμένα δεδομένα να ακολουθούν κανονικές κατανομές [*Valencia & Schaake*, 1972, *Hoshi & Burges* 1979, *Stedinger & Vogel*, 1984]. Ωστόσο, όπως παρατηρεί ο *Todini* [1980], αυτό έχει συνέπεια να μην ισχύει πλέον η προσθετική ιδιότητα (δηλαδή το άθροισμα των μεταβλητών χαμηλότερου επιπέδου δεν είναι ίσο με τη μεταβλητή υψηλότερου επιπέδου). Για την αντιμετώπιση αυτού του προβλήματος των μεθόδων της δεύτερης κατηγορίας έχουν προταθεί διάφορα προσεγγιστικά διορθωτικά σχήματα [*Lane & Frevert* 1990, *Stedinger & Vogel*, 1984, *Grygier & Stedinger*, 1988, 1990].

'Ενα άλλο μειονέκτημα των μοντέλων επιμερισμού είναι ο υπερβολικά μεγάλος αριθμός παραμέτρων τους που οφείλεται στον μεγάλο αριθμό συνδιασπορών που επιχειρούν να διατηρήσουν. Για τη μείωση του αριθμού των παραμέτρων έχουν προταθεί δύο διαφορετικού τύπου μοντέλα: τα Σταδιακά Μοντέλα Επιμερισμού (Staged Disaggregation Models - SDMs) και τα Συμπυκνωμένα Μοντέλα Επιμερισμού (Condensed Disaggregation Models - CDMs). Τα πρώτα [*Lane*, 1979, 1982, *Salas κ.α.*, 1980, *Stedinger & Vogel*, 1984, *Grygier & Stedinger*, 1988, *Lane & Frevert*, 1990] χρησιμοποιούν μια τεχνική επιμερισμού που περιλαμβάνει διάφορα βήματα αποφεύγοντας την ταυτόχρονη γέννηση των χρονοσειρών για όλες τις υπό μελέτη θέσεις. Τα δεύτερα [*Lane*, 1979, 1982, *Pereira κ.α.*, 1984, *Oliveira κ.α.*, 1988, *Stedinger & Vogel*, 1984, *Stedinger κ.α.*, 1985, *Grygier & Stedinger*, 1988] μειώνουν τον αριθμό των παραμέτρων επιλέγοντας εκ των προτέρων τη διατήρηση ενός μόνο μέρους των συσχετίσεων μεταξύ των τμηματικών μεταβλητών.

Το μοντέλο που αναπτύσσεται εδώ, το οποίο αναφέρεται ως Δυναμικό Μοντέλο Επιμερισμού (Dynamic Disaggregation Model - DDM) [*Κουτσογιάννης*, 1988, *Koutsoyiannis & Xanthopoulos*, 1990, *Koutsoyiannis*, 1992] ακολουθεί γενικά μια διαφορετική λογική. Αποσκοπεί κυρίως στην μείωση του αριθμού των παραμέτρων που απαιτούνται για τον επιμερισμό και στην εξασφάλιση της δυνατότητας γέννησης μεταβλητών με ασύμμετρες κατανομές. Η διατύπωση του μοντέλου περιλαμβάνει μια γενική βήμα-προς-βήμα μεθοδολογία κατάλληλη για προβλήματα στοχαστικού επιμερισμού, η οποία επιτρέπει την ανάπτυξη διάφορων σχηματισμών του μοντέλου. Το μοντέλο περιλαμβάνει στη δομή του ένα αντίστοιχο σειριακό μοντέλο (π.χ. ένα μοντέλο τύπου PAR) και χρησιμοποιεί το ίδιο σύνολο παραμέτρων με το σειριακό μοντέλο. Με το ακολουθούμενο σχήμα επιμερισμού οι μεταβλητές χαμηλότερου επιπέδου γεννώνται στη σειρά η μία μετά την άλλη. 'Ετσι, ο επιμερισμός μιας μεταβλητής υψηλότερου επιπέδου στις συνιστώσες της διαχωρίζεται σε ισοδύναμα επί μέρους βήματα, που το καθένα αντιστοιχεί σε μια μεταβλητή χαμηλότερου επιπέδου. Σε κάθε βήμα εκτελούνται δύο διαφορετικές διαδικασίες. Πρώτα υπολογίζονται οι παράμετροι των εξισώσεων γέννησης, σε τρόπο ώστε να διατηρούνται οι απαιτούμενες περιθώριες και υπό συνθήκη ροπές. Στη δεύτερη διαδικασία οι παράμετροι αυτές χρησιμοποιούνται για την γέννηση μιας μεταβλητής χαμηλότερου επιπέδου και την ενημέρωση της ποσότητας που απομένει να επιμεριστεί στα επόμενα βήματα της μεθόδου. Οι δυο αυτές διαδικασίες αναφέρονται ως **διαδικασία προσδιορισμού ροπών** και **διαδικασία διχασμού**, αντίστοιχα.

Κάθε σχηματισμός του μοντέλου εξαρτάται από το ιδιαίτερο σχήμα της διαδικασίας διχασμού και της διαδικασίας προσδιορισμού ροπών που χρησιμοποιεί. Η μορφή της πρώτης εξαρτάται κυρίως από τον τύπο των περιθώριων κατανομών των μεταβλητών χαμηλότερου επιπέδου. Η μορφή της δεύτερης εξαρτάται από τον τύπο της στοχαστικής εξάρτησης μεταξύ των μεταβλητών.

Αν και οι πρώτοι σχηματισμοί του μοντέλου [*Koutsoyiannis & Xanthopoulos*, 1990], αφορούσαν σε μονοδιάστατα προβλήματα επιμερισμού που περιγράφονται από Μαρκοβιανές ακολουθίες με κατανομές κανονικές ή γάμα, ήδη το μοντέλο έχει γενικευτεί και χρησιμοποιηθεί σε πολυμεταβλητά προβλήματα. Ωστόσο η Μαρκοβιανή δομή της ακολουθίας των τμηματικών μεταβλητών έχει παραμείνει ως βασική υπόθεση και στο γενικευμένο πολυμεταβλητό μοντέλο, εξαιτίας της απλότητας της και του ελάχιστου αριθμού παραμέτρων που χρησιμοποιεί. Το μοντέλο, επίσης, προσεγγίζει ικανοποιητικά τις συσχετίσεις μεταβλητών που ανήκουν σε διαδοχικά χρονικά βήματα υψηλότερου επιπέδου. Επίσης έχει αναπτυχθεί και παρουσιάζεται ένας σχηματισμός του μοντέλου που διατηρεί τις συσχετίσεις μεταξύ μεταβλητών υψηλότερου και χαμηλότερου επιπέδου που ανήκουν σε διαδοχικά χρονικά βήματα υψηλότερου επιπέδου, ο οποίος αντιμετωπίζει την ασυνέπεια των μοντέλων επιμερισμού που κατέγραψαν οι *Stedinger & Vogel* [1984].

Το μοντέλο είναι κατάλληλο για επιμερισμό κάθε είδους υδρολογικής μεταβλητής όπως της απορροής, της βροχόπτωσης και της εξατμισοδιαπνοής.

2. ΒΑΣΙΚΑ ΣΤΟΙΧΕΙΑ ΤΟΥ ΔΥΝΑΜΙΚΟΥ ΜΟΝΤΕΛΟΥ ΕΠΙΜΕΡΙΣΜΟΥ

Ας θεωρήσουμε ένα συγκεκριμένο χρονικό βήμα υψηλότερου επιπέδου, ή περίοδο (π.χ. ένα έτος) και μια υποδιαίρεση της περιόδου σε k βήματα χαμηλότερου επιπέδου ή υποπεριόδους (π.χ. 12 μήνες) που καθεμιά συμβολίζεται με ένα χρονικό δείκτη $t = 1, ..., k$. Συμβολίζουμε με $\underline{Z} = (Z_1, Z_2, ..., Z_n)^{\mathrm{T}}$ τις μεταβλητές υψηλότερου επιπέδου (αθροιστικές) αυτής της περιόδου σε n θέσεις και με $\underline{X}^t = (X_1^t, X_2^t, ..., X_n^t)^{\mathrm{T}}$ τις μεταβλητές χαμηλότερου επιπέδου (τμηματικές) της υποπεριόδου t στις ίδιες n θέσεις. Οι μεταβλητές υψηλότερου και χαμηλότερου επιπέδου ικανοποιούν τη σχέση

$$\underline{X}^1 + \underline{X}^2 + ... + \underline{X}^t + ... + \underline{X}^k = \underline{Z} \tag{1}$$

Ορίζουμε επίσης το ακόλουθο τμηματικό άθροισμα των μεταβλητών χαμηλότερου επιπέδου, το οποίο θα αναφέρεται ως υπόλοιπο προς διάθεση:

$$\underline{S}^t = \underline{X}^t + \underline{X}^{t+1} + ... + \underline{X}^k \tag{2}$$

που εκφράζεται ισοδύναμα και από τη σχέση

$$\underline{S}^t = \underline{Z} - \underline{X}^1 - \underline{X}^2 - ... - \underline{X}^{t-1} \tag{3}$$

Με τη βοήθεια του παραπάνω συμβολισμού θα διατυπώσουμε τώρα ορισμένα από τα κύρια χαρακτηριστικά του μοντέλου, τα οποία θα μελετηθούν εκτενώς στις επόμενες παραγράφους.

α. Ο επιμερισμός της μεταβλητής υψηλότερου επιπέδου $\underline{Z}$, στις k συνιστώσες της $\underline{X}^1$, $\underline{X}^2$,.., $\underline{X}^k$ διαχωρίζεται σε $n(k-1)$ βήματα, καθένα από τα οποία αναφέρεται στην γέννηση μιας μόνο μεταβλητής χαμηλότερου επιπέδου σε μια θέση. (Για μια θέση χρειάζονται k–1 βήματα, δεδομένου ότι η τελευταία μεταβλητή χαμηλότερου επιπέδου μπορεί να υπολογιστεί από την (1)) .

β. Στο ξεκίνημα του βήματος t στη θέση j, έχουμε γνωστή στη διάθεσή μας την *προηγούμενη πληροφορία*, η οποία αποτελείται από (i) την πλήρη ακολουθία των μεταβλητών υψηλότερου επιπέδου και (ii) τις τιμές των μεταβλητών χαμηλότερου επιπέδου των προηγούμενων βημάτων. Κατά συνέπεια, από την (3), το υπόλοιπο προς διάθεση S_j^t είναι επίσης γνωστό.

γ. Με βάση την (2) και μια συγκεκριμένη υπόθεση σχετικά με τη στοχαστική εξάρτηση της ακολουθίας των $\underline{X}^t$, σε κάθε βήμα υπολογίζεται η από κοινού δεσμευμένη συνάρτηση κατανομής των (X_j^t, S_j^t), όπου η δέσμευση αναφέρεται στην προηγούμενη πληροφορία. Στην πραγματικότητα η εν λόγω συνάρτηση κατανομής μπορεί μόνο να προσεγγιστεί μέσω των στατιστικών ροπών της. Ἑτσι, σε κάθε βήμα έχουμε μια διαδικασία υπολογισμού ροπών. Σχετικά με τη στοχαστική εξάρτηση των $\underline{X}^t$, για λόγους μαθηματικής ευκολίας, είναι σκόπιμο να υποθέσουμε ότι η ακολουθία των $\underline{X}^t$ είναι ακολουθία αυτοπαλινδρόμησης (autoregressive sequence). Μια τέτοια υπόθεση επιτρέπει τον άνετο υπολογισμό των δεσμευμένων ροπών. Αυτή η συγκεκριμένη ακολουθία αυτοπαλινδρόμησης των $\underline{X}^t$ θα αναφέρεται ως το αντίστοιχο σειριακό μοντέλο του μοντέλου επιμερισμού. Θα πρέπει να τονιστεί ότι, δεδομένης της υπόθεσης για την εξάρτηση των $\underline{X}^t$, η εξάρτηση μεταξύ των $\underline{X}^t$ και $\underline{Z}^t$ (ή μεταξύ των $\underline{X}^t$ και $\underline{S}^t$) είναι πλήρως καθορισμένη από την (1) (ή, αντίστοιχα, την (2)).

δ. Αφού είναι δεδομένες οι περιθώριες και από κοινού ροπές των (X_j^t, S_j^t), δεσμευμένες από την προηγούμενη πληροφορία, η γέννηση της X_j^t μπορεί να θεωρηθεί ως ένα πλήρως απομονωμένο πρόβλημα με τρεις μόνο μεταβλητές, τις X_j^t, S_j^{t+1} και S_j^t, που ικανοποιούν την

$$X_j^t + S_j^{t+1} = S_j^t \tag{4}$$

Αυτή η απομονωμένη διαδικασία, που αναφέρεται ως διαδικασία διχασμού, αγνοεί όλες τις άλλες μεταβλητές του πολυμεταβλητού προβλήματος. Ουσιαστικά διαιρεί τη δεδομένη ποσότητα S_j^t σε δύο τμήματα X_j^t και S_j^{t+1}. Η σύνδεσή της με το πολυμεταβλητό μοντέλο βρίσκεται στο γεγονός ότι τροφοδοτείται με δεσμευμένες ροπές των (X_j^t, S_j^t) οι οποίες εμπεριέχουν τη στοχαστική δομή της πλήρους ακολουθίας των μεταβλητών. Η διαδικασία διχασμού μπορεί να παρει διάφορες μορφές, ανάλογα με την ιδιαίτερη περιθώρια κατανομή των μεταβλητών χαμηλότερου επιπέδου. Η διαδικασία βασίζεται σε ένα σχήμα γέννησης της μορφής

$$X_j^t = \mathrm{h}(S_j^t, W_j^t) \tag{5}$$

όπου η h(.,.) είναι γενικά μια μη γραμμική συνάρτηση που θα συζητηθεί παρακάτω και W_j^t είναι μια τυχαία μεταβλητή ανεξάρτητη από την S_j^t και όχι απαραίτητα κανονική.

ε. Μετά την εκτέλεση της διαδικασίας διχασμού ενημερώνεται η προηγούμενη πληροφορία, συμπεριλαμβάνοντας τώρα και τη γνωστή X_j^t, ενώ το νέο υπόλοιπο προς διάθεση S_j^{t+1} μεταφέρεται στο επόμενο βήμα που αφορά στην ίδια θέση.

Η κύρια καινοτομία της παραπάνω μεθόδου είναι η διαφορετική μεταχείριση που γίνεται στην εξάρτηση ανάμεσα στις μεταβλητές χαμηλότερου επιπέδου και στην εξάρτηση που επιβάλλει η αθροιστική ιδιότητα, όπως αυτή ορίζεται από την (1). Η πρώτη αντιμετωπίζεται μέσω της διαδικασίας υπολογισμού ροπών, ενώ η δεύτερη μέσω της διαδικασίας διχασμού. Αποτέλεσμα του διαφορετικού τρόπου μεταχείρισης είναι η αρθρωτή δομή του μοντέλου, η οποία είναι ευέλικτη και μπορεί να δώσει πολλούς διαφορετικούς σχηματισμούς του μοντέλου.

Το Δυναμικό Μοντέλο Επιμερισμού έχει εμφανώς πολλές διαφορές από άλλα μοντέλα επιμερισμού. Σε αντίθεση με τον εντελώς γραμμικό τύπο των άλλων μοντέλων [*Valencia & Schaake*, 1972, 1973, *Mejia & Rousselle*, 1976, *Tao & Delleur*, 1976, *Hoshi & Burges*, 1979, *Todini*, 1980 και *Stedinger & Vogel*, 1984], το DDM αποτελεί συνδυασμό ενός γραμμικού τμήματος (διαδικασία υπολογισμού ροπών) και ενός μη γραμμικού (διαδικασία διχασμού, βασισμένου στην (5)). Επί πλέον, το γραμμικό τμήμα του μοντέλου στην πραγματικότητα δεν μοιάζει με τα γραμμικά σχήματα που χρησιμοποιούνται από άλλα μοντέλα. ʼΟπως και το μοντέλο του *Todini* [1980] το DDM έχει τη δυνατότητα χειρισμού της ασυμμετρίας με άμεσο (explicit) τρόπο, χωρίς απώλεια της αθροιστικής ιδιότητας, αλλά τα δύο μοντέλα δεν έχουν άλλες ομοιότητες πέρα από αυτή. Το DDM μεταχειρίζεται μόνο ένα υποσύνολο των από κοινού ροπών δεύτερης τάξης μεταξύ των μεταβλητών χαμηλότερου επιπέδου, όπως κάνουν και τα συμπυκνωμένα μοντέλα επιμερισμού, αλλά, ωστόσο, υπάρχουν ουσιαστικές διαφορές ανάμεσα στα δύο μοντέλα, οι οποίες αναπτύσσονται πιο κάτω.

3. Η ΔΙΑΔΙΚΑΣΙΑ ΔΙΧΑΣΜΟΥ

Γενικά

Αφού το μοντέλο διχασμού είναι πλήρως απομονωμένο από το υπόλοιπο μοντέλο, μπορεί να μελετηθεί ξεχωριστά, πράγμα που γίνεται σε αυτό το κεφάλαιο. Για λόγους μαθηματικής άνεσης, ο συμβολισμός στο κεφάλαιο αυτό είναι ο απλούστερος δυνατός και αποφεύγεται η χρήση χρονικών και χωρικών δεικτών, οι οποίοι δεν είναι απαραίτητοι εδώ. Η διαδικασία διχασμού μπορεί να θεωρηθεί ως ένα απλό επιμεριστικό πρόβλημα με δύο μεταβλητές χαμηλότερου επιπέδου που αναφέρονται σε μια μόνο θέση. Θα πρέπει να τονιστεί ότι δεν είναι απαραίτητη ούτε πρόκειται να γίνει επέκταση της διαδικασίας διχασμού για την περίπτωση περισσότερων μεταβλητών και πολλών θέσεων. Αυτό που απαιτείται προκειμένου να έχουμε ένα πολυμεταβλητό μοντέλο πολλών θέσεων είναι η ανάπτυξη μιας ξεχωριστής διαδικασίας υπολογισμού ροπών και η σύνδεση των δύο διαδικασιών.

Ας θεωρήσουμε, λοιπόν, την περίπτωση δύο μεταβλητών χαμηλότερου επιπέδου X και Y με άθροισμα τη μεταβλητή υψηλότερου επιπέδου S, ήτοι

$$X + Y = S \tag{6}$$

Προκειμένου να γεννήσουμε την τιμή της X απαιτείται να καθορίσουμε την ακόλουθη έκφραση, όμοια με την (5)

$$X = \mathrm{h}(S,\ W) \tag{7}$$

όπου W είναι τυχαία μεταβλητή ανεξάρτητη της S και όχι αναγκαστικά κανονική. Αν δοθεί μια συγκεκριμένη συνάρτηση $\mathrm{h}(.,.)$ και μια συνάρτηση κατανομής $\mathrm{F}(w)$, η X μπορεί να παραχθεί άμεσα από την (7) και στη συνέχεια η Y μπορεί να υπολογιστεί από την (6). Ο καθορισμός των $\mathrm{h}(.,.)$ και $\mathrm{F}(w)$ πρέπει να γίνει σε τρόπο ώστε να διατηρείται η από κοινού συνάρτηση κατανομής των $(X,\ S)$. Ωστόσο, με την εξαίρεση ορισμένων ειδικών περιπτώσεων, η πλήρης διατήρηση της από κοινού κατανομής είναι αδύνατη. Στην πράξη όμως επαρκεί η διατήρηση των τριών πρώτων ροπών. Η συμπερίληψη και των τρίτων ροπών είναι απαραίτητη για να μπορεί το μοντέλο να χειρίζεται μη κανονικές κατανομές.

Δεδομένου ότι η S είναι ήδη γνωστή στο ξεκίνημα της διαδικασίας διχασμού, προϋποτίθεται ότι έχει εξασφαλιστεί η διατήρηση των περιθώριων ροπών της, δηλαδή των $\lambda_1' = \mathrm{E}[S]$ και $\lambda_i = \mathrm{E}\left[(S-\lambda_1')^i\right]$, $i = 2,3,\ldots$. Η πλήρης διατήρηση των πρώτων, δεύτερων και τρίτων ροπών των X, Y και S, σημαίνει ότι η διαδικασία διχασμού οφείλει να αναπαράγει τις ακόλουθες έξι ποσότητες:

Περιθώριες ροπές της X

$$\eta_1' = \mathrm{E}[X], \qquad \eta_2 = \mathrm{Var}[X] = \mathrm{E}\left[(X-\eta_1')^2\right] \qquad \eta_3 = \mu_3[X] = \mathrm{E}\left[(X-\eta_1')^3\right]$$

Από κοινού ροπές των (X, S)

$$\begin{aligned} \zeta_{11} &= \mathrm{Cov}[X, S] = \mathrm{E}\left[(X-\eta_1')(S-\lambda_1')\right] \\ \zeta_{12} &= \mu_{12}[X, S] = \mathrm{E}\left[(X-\eta_1')(S-\lambda_1')^2\right] \\ \zeta_{21} &= \mu_{21}[X, S] = \mathrm{E}\left[(X-\eta_1')^2(S-\lambda_1')\right] \end{aligned} \tag{8}$$

όπου το σύμβολο μ με ένα ή δύο δείκτες παριστάνει περιθώριες ή από κοινού ροπές και οι δείκτες συμβολίζουν την τάξη των ροπών. Εύκολα αποδεικνύεται ότι κάθε περιθώρια ή από κοινού ροπή μεταξύ των τριών μεταβλητών (X, Y, S) τάξης ≤ 3 διατηρείται, εφόσον διατηρούνται οι παραπάνω έξι ποσότητες και οι περιθώριες ροπές της S, ήτοι οι λ_1', λ_2, λ_3 (βλ. για παράδειγμα τις (21) έως (24) πιο κάτω). 'Ετσι, δεδομένου ότι οι παράμετροι λ διατηρούνται έξω από τη διαδικασία διχασμού και η Y υπολογίζεται από την (6), η διατήρηση των παραμέτρων η και ζ συνεπάγεται τη διατήρηση κάθε ροπής τάξης $\leq$ 3, περιθώριας ή από κοινού. Στην περίπτωση που ενδιαφέρει μόνο η διατήρηση των περιθώριων ροπών των X και Y απαιτείται να διατηρείται μόνο η διαφορά $\zeta_{12} - \zeta_{21}$, αντί των δύο επί μέρους τιμών.

Το γραμμικό σχήμα

Ας εξετάσουμε πρώτα τη γραμμική μορφή της συνάρτησης h(.,.) (το γραμμικό επιμεριστικό σχήμα), η οποία ορίζεται από την έκφραση

$$X = a\ S + b\ W \tag{9}$$

όπου a και b είναι παράμετροι που πρέπει να εκτιμηθούν. Υποθέτοντας, χωρίς απώλεια της γενικότητας, ότι $\mathrm{Var}[W] = 1$, οι άλλες δύο παράμετροι του μοντέλου είναι οι $\mathrm{E}[W]$ και $\mu_3[W]$. Εύκολα αποδεικνύεται ότι το παραπάνω γραμμικό σχήμα έχει τη δυνατότητα να διατηρεί τις πρώτες και δεύτερες ροπές αλλά όχι και τις τρίτες ροπές των X και Y, αν αυτές είναι μη μηδενικές. Πράγματι, τα a, b και $\mathrm{E}[W]$ μπορούν να προσδιοριστούν σε τρόπο ώστε να διατηρούνται οι πρώτες και δεύτερες ροπές ($a = \zeta_{11}/\lambda_2$, $b = \sqrt{\eta_2 - \zeta_{11}^2/\lambda_2}$,

$E[W] = (\eta_1' - a\lambda_1') / b)$. Ακολούθως, η $\mu_3[W]$ μπορεί να προσδιοριστεί σε τρόπο ώστε να διατηρείται η $\eta_3 = \mu_3[X]$. Αφού δεν υπάρχει άλλη παράμετρος προς προσδιορισμό, είναι προφανές ότι δεν υπάρχει τρόπος να διατηρηθούν τα ζ_{12} και ζ_{21}. Κατά συνέπεια, η $\mu_3[Y]$ δεν είναι δυνατό να διατηρηθεί από αυτό το σχήμα. Το ίδιο συμπέρασμα υπονοείται και από τον *Todini* [1980]. Στην πραγματικότητα, ο *Todini* [1980, σ. 203] χρησιμοποιεί ένα γραμμικό σχήμα με περισσότερες από δύο μεταβλητές, όπου η κατάσταση είναι διαφορετική, αν πρόκειται να διατηρηθούν μόνο οι περιθώριες τρίτες ροπές.

Η διατήρηση των πρώτων και δεύτερων ροπών είναι πιθανότατα το μόνο πρόβλημα που μπορεί να αντιμετωπιστεί με ακριβή τρόπο. Το γραμμικό μοντέλο είναι επαρκές για το πρόβλημα αυτό. Στην ειδική περίπτωση που η κατανομή όλων των μεταβλητών είναι κανονική, το γραμμικό μοντέλο μπορεί ακόμη να διατηρήσει και τις πλήρεις συναρτήσεις κατανομής όλων των μεταβλητών. Παρακάτω εξετάζονται ορισμένα μη γραμμικά σχήματα, τα οποία μπορούν να χειριστούν ροπές μεγαλύτερης τάξης στη γενική περίπτωση, αλλά δεν είναι ακριβή σχήματα, με την έννοια ότι δεν διατηρούν επακριβώς καμιά στατιστική παράμετρο, αλλά μόνο προσεγγιστικά.

Η χρήση μη γραμμικών μετασχηματισμών

Αποτελεί κοινή πρακτική στα στοχαστικά μοντέλα, ακόμη και στα μοντέλα επιμερισμού, η εφαρμογή μη γραμμικών μετασχηματισμών σε μεταβλητές με ασύμμετρες κατανομές, προκειμένου να επιτευχθούν συμμετρικές κατανομές που είναι ευκολότερες στο χειρισμό τους. Ας εξετάσουμε την πιο συνηθισμένη περίπτωση, το λογαριθμικό μετασχηματισμό, που δίνεται από σχέσεις της μορφής

$$X' = \ln(X - c_X) \tag{10}$$

$$Y' = \ln(Y - c_Y) \tag{11}$$

$$S' = \ln(S - c_S) \tag{12}$$

όπου c_X, c_Y και c_S είναι παράμετροι προς εκτίμηση. Προφανώς, λόγω της (6) και δεδομένου ότι η λογαριθμοκανονική κατανομή δεν είναι αναγεννήσιμη (δηλαδή το άθροισμα δύο μεταβλητών με λογαριθμοκανονική κατανομή δεν ακολουθεί λογαριθμοκανονική κατανομή), οι X', Y' και S' δεν μπορεί να ακολουθούν όλες κανονική κατανομή ταυτοχρόνως. Κατά συνέπεια ένας γραμμικός κανόνας γέννησης, όπως ο ακόλουθος:

$$X' = a_X S' + b_X W_X' \tag{13}$$

συνδυαζόμενος με την (6) δεν μπορεί να είναι ακριβής ως προς την πλήρη διατήρηση των κατανομών των τριών μεταβλητών. Επιπλέον, αν μειώσουμε τις απαιτήσεις στη διατήρηση μόνο των ροπών,. η κατάσταση θα είναι παρόμοια όπως στην περίπτωση του γραμμικού σχήματος. Πράγματι ο κανόνας της (13) έχει πάλι τέσσερις παραμέτρους προς προσδιορισμό, τις a_X, b_X, c_X και $\mathrm{E}[W'_x]$. Ας σημειωθεί ότι η c_S δεν είναι παράμετρος της διαδικασίας διχασμού, δεδομένου ότι αναφέρεται στην προϋποτιθέμενη κατανομή της S. Επίσης, η $\mu_3[W'_X]$ δεν είναι είναι παράμετρος· πρέπει αναγκαστικά να μηδενιστεί γιατί η W'_X θα πρέπει να υποτεθεί κανονική, διαφορετικά δεν υπάρχουν αναλυτικές σχέσεις μεταξύ των στατιστικών παραμέτρων των X και X'. 'Ετσι, το παραπάνω σχήμα, εξαιτίας της ανεπάρκειας του αριθμού των παραμέτρων του, δεν έχει τη δυνατότητα διατήρησης των από κοινού τρίτων ροπων ζ_{12} και ζ_{21}.

Μια άλλη μέθοδος, η οποία κυρίως χρησιμοποιείται στα μοντέλα επιμερισμού, αρχικά αγνοεί την (6) και γεννά ανεξάρτητα το X' από την (13) και το Y' από μια παρόμοια σχέση, ήτοι

$$Y' = a_Y S' + b_Y W'_Y \tag{14}$$

Παίρνοντας αντιλογαρίθμους στις (13) και (14) βρίσκουμε τις ακόλουθες πρώτες προσεγγίσεις των X και Y:

$$X^* = (S - c_S)^{a_X} W_X^{b_X} + c_X \tag{15}$$

$$Y^* = (S - c_S)^{a_Y} W_Y^{b_Y} + c_Y \tag{16}$$

όπου $W_X = \exp(W'_X)$ και $W_Y = \exp(W'_Y)$. Προφανώς το άθροισμα των X^* και Y^* δεν είναι ίσο με S. Για να ισχύσει η προσθετική ιδιότητα (6),θα πρέπει να διορθωθούν τα X^* και Y^*. Οι *Lane & Frevert* [1990, σ. V-22], *Stedinger & Vogel* [1984] και *Grygier & Stedinger* [1988, 1990] έχουν αναπτύξει διάφορες διορθωτικές διαδικασίες. Η απλούστερη από αυτές, η αναλογική, βασίζεται στις σχέσεις

$$X = \frac{S}{X^* + Y^*} X^*, \quad Y = \frac{S}{X^* + Y^*} Y^* \tag{17}$$

Συνδυάζοντας τις (15), (16) και (17) παίρνουμε τον ακόλουθη τελική λύση ως προς X

$$X = \frac{(S - c_S)^{a_X} W_X^{b_X} + c_X}{(S - c_S)^{a_X} W_X^{b_X} + c_X + (S - c_S)^{a_Y} W_Y^{b_Y} + c_Y} S \tag{18}$$

η οποία μπορεί να συνδυαστεί με την (6) για τη γέννηση του Y.

Η τελευταία σχέση έχει επαρκή αριθμό παραμέτρων (οκτώ) για την εξασφάλιση της διατήρησης των στατιστικών μεγεθών της (8). Ωστόσο, η πολυπλοκότητά της καθιστά αδύνατο τον αναλυτικό υπολογισμό οποιασδήποτε ροπής, ακόμη και της μέσης τιμής. Με αυτή την έννοια, το σχήμα δεν μπορεί να είναι ακριβές, αφού δεν μπορεί να αναπτυχθεί καμιά ακριβής διαδικασία εκτίμησης παραμέτρων.

Κάπως απλούστερη είναι η κατάσταση στην περίπτωση που χρησιμοποιείται μια άλλη διαδικασία διόρθωσης, βασισμένη στις τυπικές αποκλίσεις των μεταβλητών χαμηλότερου επιπέδου [*Lane & Frevert,* 1990, σ. V-22]

$$X = X^* + (S - X^* - Y^*)\ \delta_X, \quad Y = Y^* + (S - X^* - Y^*)\ (1 - \delta_X) \tag{19}$$

όπου $\delta_X = \sigma_X \ / \ (\sigma_X + \sigma_Y)$ και σ_X και σ_Y είναι οι τυπικές αποκλίσεις των X και Y. Συνδυάζοντας τις (15), (16) και (19) παίρνουμε τον ακόλουθο κανόνα γέννησης

$$X = (1-\delta_X) \left[(S-c_S)^{a_X} W_X^{b_X} + c_X\right] - \delta_X \left[(S-c_S)^{a_Y} W_Y^{b_Y} + c_Y\right] + \delta_X S \tag{20}$$

ο οποίος μπορεί να συνδυαστεί με την (6) για τη γέννηση της Y. Χρησιμοποιώντας την (20) μπορούμε να βρούμε αναλυτικές εκφράσεις για τις περιθώριες ροπές της X (η), οι οποίες όμως είναι πολύπλοκες, ενώ δεν είναι δυνατή η εξαγωγή ανάλογων σχέσεων για τις από κοινού ροπές των X και S (ζ). Κατά συνέπεια και πάλι δεν είναι δυνατή η εξαγωγή ακριβούς εκτιμήτριας των ροπών και κατά συνέπεια ούτε η (20) μπορεί να οδηγήσει σε ακριβές σχήμα επιμερισμού.

Πέρα από τα παραπάνω, οι (18) και (20) έχουν και μια άλλη πηγή ανακρίβειας που οφείλεται στο γεγονός ότι η λογαριθμοκανονική κατανομή δεν είναι αναγεννήσιμη. Ακόμη και αν υπήρχε ακριβής διαδικασία εκτήμησης παραμέτρων, αυτή θα στηρίζονταν στην υπόθεση ότι οι S, W_X και W_Y ακολουθούν λογαριθμοκανονικές κατανομές. Αλλά τότε οι X και Y δεν μπορούν να ακολουθούν λογαριθμοκανονική κατανομή. 'Ετσι, στην επόμενη εκτέλεση της διαδικασίας διχασμού, στην οποία η ενημερωμένη S είναι ίση με την προηγούμενη Y, η παραπάνω υπόθεση δεν είναι πλέον ακριβής.

Συμπερασματικά, η απαίτηση για ακρίβεια είναι συνυφασμένη με αξεπέραστες δυσκολίες. Προκειμένου να σχηματιστεί μια προσεγγιστική διαδικασία, συνήθως γίνεται η υπόθεση ότι οι στατιστικές παράμετροι της X είναι ίσες με αυτές της X^* και οι στατιστικές παράμετροι της Y είναι ίσες με αυτές της Y^*. Αυτή η προσέγγιση έχει δώσει καλά αποτελέσματα σε πολλά μοντέλα. Ωστόσο και πέρα από τη μη ακρίβεια της διαδικασίας, υπάρχει και ένα ακόμη σχετικό πρόβλημα: Ο αριθμός των παραμέτρων δεν είναι πλέον επαρκής για τη διατήρηση όλων των ροπών της (8). Πράγματι, τέσσερις παράμετροι και συγκεκριμένα οι a_X, b_X, c_X και $\mathrm{E}[W'_x]$, χρησιμοποιούνται τώρα για τη διατήρηση των τριών περιθώριων ροπών της X^* και της συνδιασποράς των X^* και S. Οι υπόλοιπες παράμετροι είναι οι a_Y, b_Y, c_Y και $\mathrm{E}[W'_Y]$. Δεδομένου ότι η γέννηση της Y^* είναι ανεξάρτητη από αυτή της X^* ($X^* + Y^* \neq S$), η πρώτη και οι δεύτερες ροπές της Y^* δεν διατηρούνται αυτόματα, όπως θα συνέβαινε στην περίπτωση που η Y προσδιορίζονταν από την (6). 'Ετσι, πρέπει να διατεθούν τρεις παράμετροι για τη διατήρηση αυτών των ροπών και συγκεκριμένα των $\mathrm{E}[Y^*]$, $\mathrm{Var}[Y^*]$ και $\mathrm{Cov}[Y^*, S]$. Απομένει λοιπόν μία παράμετρος, η οποία θα πρέπει να χρησιμοποιηθεί για τη διατήρηση της τρίτης ροπής της Y^*. Το αποτέλεσμα αυτού του τρόπου διάθεσης των παραμέτρων είναι ότι τελικά διατηρείται η διαφορά $\zeta_{12} - \zeta_{21}$ και όχι οι συγκεκριμένες τιμές των ζ_{12} και ζ_{21} (Σημειώνεται ότι η $\mu_3[Y]$ δίνεται από την παρακάτω σχέση (23), η λ_3 διατηρείται έξω από τη διαδικασία διχασμού - ή σε προηγούμενη εφαρμογή της - η η_3 προσεγγιστικά διατηρείται μέσω της αντίστοιχης ροπής της X^*, και έτσι η διατήρηση της $\mu_3[Y]$ είναι ισοδύναμη με τη διατήρηση της διαφοράς $\zeta_{12} - \zeta_{21}$). Προφανώς το πρόβλημα αυτό δεν είναι τόσο σοβαρό: οι από κοινού τρίτες ροπές των X και Y δεν διατηρούνται, αλλά οι περιθώριες ροπές που κυρίως ενδιαφέρουν διατηρούνται προσεγγιστικά. Τονίζεται ότι οι πιο πάνω δυσκολίες στην κατάρτιση ενός ακριβούς μοντέλου δεν αφορούν αποκλειστικά στο συγκεκριμένο μοντέλο που περιγράφεται εδώ, αλλά είναι κοινές σε όλα τα μοντέλα που χρησιμοποιούν λογαριθμικούς μετασχηματισμούς των μεταβλητών.

Η υλοποίηση μιας προσεγγιστικής διαδικασίας για την εκτίμηση των παραμέτρων (σύμφωνης με τους παραπάνω περιορισμούς), η οποία να προσαρμόζεται στις ειδικές απαιτήσεις της διαδικασίας διχασμού, είναι σχετικά απλή. Πρώτο, υπολογίζουμε τις ροπές της Y βάσει των ροπών της (8), χρησιμοποιώντας την (6), ήτοι

$$\mathrm{E}[Y] = \lambda'_1 - \eta'_1 \tag{21}$$

$$\mathrm{Var}[Y] = \lambda_2 + \eta_2 - 2\zeta_{11} \tag{22}$$

$$\mu_3[Y] = \lambda_3 - \eta_3 - 3\zeta_{12} + 3\zeta_{21} \tag{23}$$

$$\mathrm{Cov}[Y, S] = \lambda_2 - \zeta_{11} \tag{24}$$

Δεύτερο, υποθέτουμε ότι οι ροπές των X^* και Y^* είναι ίσες με τις αντίστοιχες των X και Y. Τρίτο, υπολογίζουμε τις ροπές των X' και Y' (και τα c_X και c_Y) βάσει των ροπών των X^* και Y^* με τη μέθοδο των ροπών [βλ. π.χ. *Charbeneau*, 1978, *Kottegoda*, 1980, σ. 137, or *Stedinger*, 1980] Τέταρτο, υπολογίζουμε τις παραμέτρους των γραμμικών σχέσεων (13) και (14), με τις καθιερωμένες μεθόδους.

Για τη γέννηση των τιμών των μεταβλητών ακολουθούμε την ακόλουθη πορεία: Πρώτο, γεννούμε τα W'_X και W'_Y από την κανονική κατανομή. Δεύτερο, υπολογίζουμε τα X' και Y' από τις (13) και (14) και παίρνουμε αντιλογαρίθμους, προκειμένου να υπολογίσουμε τα X^* και Y^* (επίσης προσθέτουμε τα c_X και c_Y). Τρίτο, χρησιμοποιούμε την (17) ή την (19) ή οποιαδήποτε άλλη διαδικασία διόρθωσης για να υπολογίσουμε τα X και Y.

Σημειώνεται ότι η παραπάνω διαδικασία διχασμού που βασίζεται σε λογαριθμικό μετασχηματισμό των μεταβλητών δεν έχει ακόμη χρησιμοποιηθεί σε πρακτικές εφαρμογές.

Με παρόμοιο τρόπο μπορούν να αναπτυχθούν και άλλες διαδικασίες διχασμού, βασισμένες σε άλλου τύπου μετασχηματισμούς, όπως π.χ. στο μετασχηματισμό Wilson-Hilferty. Τα κύρια προβλήματα αυτών των διαδικασιών θα είναι τα ίδια με αυτά που περιγράφηκαν για την περίπτωση του λογαριθμικού μετασχηματισμού.

Το δευτεροβάθμιο σχήμα

Τέλος θα εξετάσουμε ένα δευτεροβάθμιο σχήμα χωρίς μετασχηματισμούς των μεταβλητών, το οποίο ορίζεται από τις σχέσεις

$$X = \mathrm{g}(S) + \mathrm{f}(S)\ W \tag{25}$$

$$\mathrm{g}(S) = a_0 + a_1 S + a_2 S^2 \tag{26}$$

$$\mathrm{f}(S) = b_0 + b_1 S + b_2 S^2 \tag{27}$$

όπου a_i και b_i, $i = 0, 1, 2$, είναι παράμετροι προς εκτίμηση και W είναι τυχαία μεταβλητή, όχι απαραίτητα κανονική, ανεξάρτητη της S, με μηδενική μέση τιμή και μοναδιαία διασπορά. Το σχήμα αυτό, αν και δεν φαίνεται να έχει φυσικό νόημα, έχει εισαχθεί ως το πλέον άνετο από μαθηματική άποψη αναφορικά με το χειρισμό των ροπών, το οποίο παράλληλα έχει επαρκή αριθμό παραμέτρων για την κάλυψη των σχετικών περιορισμών. Ωστόσο, όπως εξηγείται παρακάτω, έχει ορισμένα δομικά προβλήματα, εξαιτίας των οποίων ούτε αυτό μπορεί να είναι ακριβές σχήμα, παρά μόνο προσεγγιστικό.

Στην παρακάτω ανάλυση για λόγους μαθηματικής ευκολίας και χωρίς απώλεια της γενικότητας, θα υποτεθεί ότι όλες οι τυχαίες μεταβλητές έχουν μηδενική μέση τιμή. Το σχήμα περιλαμβάνει επτά παραμέτρους, τους έξι συντελεστές a_i και b_i, συν την ασυμμετρία της W, $\theta_3 = \mathrm{E}[W^3]$. Συνεπώς, ο αριθμός των παραμέτρων υπερβαίνει τον αριθμό των περιορισμών της (8), παρέχοντας έτσι ένα βαθμό ελευθερίας.

Κάνοντας τις κατάλληλες πράξεις στην (25) και παίρνοντας στη συνέχεια τις αναμενόμενες τιμές, παίρνουμε εύκολα τις ακόλουθες εξισώσεις:

$$\mathrm{E}[X] = \mathrm{E}[\mathrm{g}(S)] \tag{28}$$

$$\mathrm{E}[X\,S] = \mathrm{E}[S\,\mathrm{g}(S)] \tag{29}$$

$$\mathrm{E}[X^2] = \mathrm{E}[\mathrm{g}^2(S)] + \mathrm{E}[\mathrm{f}^2(S)]\,\mathrm{E}[W^2] \tag{30}$$

$$\mathrm{E}[X\,S^2] = \mathrm{E}[S^2\mathrm{g}(S)] \tag{31}$$

$$\mathrm{E}[X^2\,S] = \mathrm{E}[S\,\mathrm{g}^2(S)] + \mathrm{E}[S\,\mathrm{f}^2(S)]\,\mathrm{E}[W^2] \tag{32}$$

$$\mathrm{E}[X^3] = \mathrm{E}[\mathrm{g}^3(S)] + 3\mathrm{E}[\mathrm{g}(S)\mathrm{f}^2(S)]\,\mathrm{E}[W^2] + \mathrm{E}[\mathrm{f}^3(S)]\,\mathrm{E}[W^3] \tag{33}$$

Είναι προφανές ότι οι εξισώσεις αυτές αποτελούν σχέσεις μεταξύ των ροπών των X και S και των άγνωστων παραμέτρων του μοντέλου. Συγκεκριμένα, οι εξισώσεις (28), (29) και (31) καθορίζουν πλήρως την $\mathrm{g}(S)$, δίνοντας απλές εκφράσεις για τον υπολογισμό των a_i. Κάπως πιο πολύπλοκη είναι η εξαγωγή της $\mathrm{f}(S)$, η οποία γίνεται από τις (30) και (32), θέτοντας $\mathrm{E}[W^2] = 1$. Τέλος, η θ_3 δίνεται από την (33). 'Ολες οι σχετικές πράξεις και οι τελικές εξισώσεις δίνονται στην Προσθήκη Α.

Υπάρχουν δύο οριακές περιπτώσεις στις οποίες το σχήμα είναι ακριβές με την έννοια ότι διατηρεί τις πλήρεις συναρτήσεις κατανομής και όχι μόνο τις ροπές. Η πρώτη περίπτωση αναφέρεται σε κανονικές μεταβλητές. Αν οι μεταβλητές X και Y (ή ισοδύναμα οι X και S) ακολουθούν διδιάστατη κανονική κατανομή, τότε το δευτεροβάθμιο σχήμα απλοποιείται σε γραμμικό και οι εξισώσεις (25) έως (27) μεταπίπτουν στην (9), όπως θεωρητικά αναμένεται (βλ. Προσθήκη Α). Προφανώς στην περίπτωση αυτή και η W είναι κανονική, ενώ η στατιστική διατήρηση επεκτείνεται στην πλήρη κατανομή των (X, Y). Η δεύτερη περίπτωση αναφέρεται σε μεταβλητές με κατανομή γάμα. Συγκεκριμένα, έστω ότι οι μεταβλητές X και Y είναι αποκλίσεις από τις μέσες τιμές δύο ανεξάρτητων μεταβλητών $\tilde{X}$

και $\tilde{Y}$, οι οποίες ακολουθούν κατανομές γάμα δύο παραμέτρων με κοινή παράμετρο κλίμακας. Σε αυτή την περίπτωση, αν $\tilde{S} = \tilde{X} + \tilde{Y}$ τότε η μεταβλητή

$$\tilde{W} = \tilde{X} \,/\, \tilde{S} \tag{34}$$

είναι ανεξάρτητη της $\tilde{S}$ και ακολουθεί κατανομή βήτα [*Johnson & Kotz,* 1972, σ. 234]. Κατά συνέπεια οι αποκλίσεις από τις μέσες τιμές συνδέονται με την

$$X = (\mathrm{E}[\tilde{X}] + \mathrm{E}[\tilde{W}]\ \mathrm{E}[\tilde{S}]) + \mathrm{E}[\tilde{W}]\ S + (\mathrm{E}[\tilde{S}] + S)\ W \tag{35}$$

όπου η τυχαία μεταβλητή $W = \tilde{W} - \mathrm{E}[\tilde{W}]$ είναι ανεξάρτητη της S. Η (35) είναι ειδική περίπτωση της (25) όπου οι συναρτήσεις $g(S)$ και $f(S)$ έχουν απλοποιηθεί σε γραμμικές συναρτήσεις ($a_2 = b_2 = 0$) ενώ εξασφαλίζεται η πλήρης διατήρηση των κατανομών των μεταβλητών.

Ας εξετάσουμε τώρα τα προβλήματα που συνδέονται με το δευτεροβάθμιο σχήμα στη γενική περίπτωση. Ένα πρώτο πρόβλημα ξεκινά από τις τριωνυμικές μορφές των $f(S)$ και $g(S)$, οι οποίες εμφανίζουν συγκεκριμένα ακρότατα (μέγιστα ή ελάχιστα). Εξαιτίας αυτών δεν είναι δυνατό να δώσουν λύσεις των εξισώσεων (28) έως (33) όλοι οι δυνατοί συνδυασμοί των λ, η και ζ. Οι σχετικοί περιορισμοί που αφορούν στην ύπαρξη των συναρτήσεων $f(S)$ και $g(S)$ δίνονται στην Προσθήκη Α (εξισώσεις (Α19) έως (Α21)). Οι περιορισμοί που ανακύπτουν από το πρόβλημα αυτό διερευνώνται περαιτέρω στο Κεφάλαιο 6.

Το δεύτερο πρόβλημα αφορά στην έμμεση παρουσία στις εξισώσεις (30), (32) και (33) ανώτερης τάξης ροπών της S και συγκεκριμένα των λ_4, λ_5 και λ_6. Το πρόβλημα αυτό επιβεβαιώνεται από τις σχέσεις της Προσθήκης Α που χρησιμοποιούνται για τον υπολογισμό των a και b. Αυτές οι ανώτερες ροπές θα πρέπει να είναι γνωστές προκειμένου να γίνει η εφαρμογή των παραπάνω σχέσεων και να διατηρηθούν οι χαμηλότερες ροπές των X και Y. Αν είναι γνωστή η πλήρης περιθώρια κατανομή της S, τότε και οι ανώτερες ροπές της είναι δυνατό να υπολογιστούν επακριβώς με κάποιο τρόπο. Ωστόσο, δεν είναι αυτή η περίπτωση που εξετάζουμε εδώ, δεδομένου ότι η διαδικασία διχασμού εκτελείται διαδοχικά σε όλα τα βήματα επιμερισμού. Επιπλέον, δεν υπάρχει τρόπος να υπολογιστούν επακριβώς σε κάθε βήμα οι ανώτερες ροπές του ενημερωμένου υπολοίπου προς διάθεση (του επόμενου βήματος, το οποίο στο τρέχον βήμα αντιπροσωπεύεται από την Y), δεδομένου ότι κάθε σχετικός υπολογισμός θα οδηγούσε σε εισαγωγή ακόμη μεγαλύτερης τάξης ροπών της S.

Για τις πρακτικές εφαρμογές, πάντως, μπορούμε να υποθέσουμε μια τυπική συνάρτηση κατανομής για την S και να υπολογίσουμε της ανώτερες ροπές αυτής της κατανομής αναλυτικά. Βεβαίως είναι προφανές ότι η αυστηρή ακρίβεια χάνεται με την εφαρμογή αυτής της μεθόδου. Για τις αριθμητικές εφαρμογές είναι χρήσιμη η εισαγωγή και χρήση των αθροιστών (cumulants), ήτοι

$$\lambda_4 = \kappa_4 + 3\,\lambda_2^2 \tag{36}$$

$$\lambda_5 = \kappa_5 + 10\,\lambda_3\,\lambda_2 \tag{37}$$

$$\lambda_6 = \kappa_6 + 15\,\kappa_4\,\lambda_2 + 10\,\lambda_3^2 + 15\,\lambda_2^3 \tag{38}$$

όπου κ_4, κ_5 και κ_6 είναι οι αθροιστές (cumulants) της *S*, τάξεως 4, 5 και 6 [*Kendall & Stuart*, 1962, σ.70]. Τονίζονται δύο ειδικές περιπτώσεις, χρήσιμες για τις εφαρμογές (βλ. *Kendall & Stuart* [1962, σ.70]). Η πρώτη είναι η περίπτωση της κανονικής κατανομής στην οποία

$$\kappa_4 = \kappa_5 = \kappa_6 = 0 \tag{39}$$

Η δεύτερη είναι η κατανομή γάμα στην οποία

$$\kappa_r = (r-1)\,\kappa_{r-1}\,\lambda_3\,/\,2\lambda_2\,,\ r > 3 \tag{40}$$

όπου $\kappa_3 = \lambda_3$. Η (39) ή η (40) μαζί με τις (36) έως (38) μπορούν επίσης να χρησιμοποιηθούν ως προσεγγίσεις και στην περίπτωση που η κατανομή της *S* δεν προσδιορίζεται επακριβώς, αλλά μπορεί να προσεγγιστεί από την κατανομή Gauss ή γάμα, αντίστοιχα.

Για να ανιχνευτεί η ευαισθησία του μοντέλου ως προς την υπόθεση συγκεκριμένου τύπου αθροιστών έγινε μια σχετική αριθμητική διερεύνηση. Συγκεκριμένα, υποτέθηκε μια μεταβλητή *S* με κατανομή γάμα, η οποία επιμερίστηκε σε δύο μεταβλητές χαμηλότερου επιπέδου. Έγιναν δύο διαφορετικές υποθέσεις σχετικές με τους αθροιστές: πρώτο, ότι τα κ_r δίνονται από την (40) (αληθής υπόθεση) και δεύτερο, ότι $\kappa_r = 0$ όπως συμβαίνει με την κανονική κατανομή (εσφαλμένη υπόθεση). Η δεύτερη υπόθεση δεν ισοδυναμεί με παραδοχή κανονικής κατανομής για την *S*: το λ_3 στις (36) έως (38) εξακολουθεί να είναι μη μηδενικό. Από τη διερεύνηση αυτή βρέθηκε ότι και στις δύο περιπτώσεις τα αποτελέσματα είναι παρόμοια (βλ. παράδειγμα στο Κεφάλαιο 6 και τον Πίνακα 2). Αυτό αποτελεί ένδειξη ότι το μοντέλο, αν και όχι θεωρητικά ακριβές, δεν επηρεάζεται πολύ από μια εσφαλμένη υπόθεση σχετικά με τους αθροιστές.

Στις εφαρμογές αυτής της μελέτης υιοθετήθηκε η χρήση των αθροιστών της κατανομής γάμα (εξ. 40) για τις περιπτώσεις που οι μεταβλητές μπορούσαν να προσεγγιστούν από την κατανομή γάμα. Περαιτέρω, για η κατανομή της W προσεγγίστηκε με την κατανομή βήτα (μετά από κατάλληλο γραμμικό μετασχηματισμό).

4. ΤΟ ΚΥΡΙΟ ΜΟΝΤΕΛΟ ΕΠΙΜΕΡΙΣΜΟΥ

Γενικά

'Ενα μοντέλο επιμερισμού μπορεί να θεωρηθεί ως μια τεχνική γέννησης μεταβλητών χαμηλότερου επιπέδου που ικανοποιούν την αθροιστική ιδιότητα (1). Οι μεταβλητές χαμηλότερου επιπέδου της τρέχουσας περιόδου $\underline{X}^1, ..., \underline{X}^k$ αποτελούν ένα υποσύνολο μιας στοχαστικής ακολουθίας η οποία επεκτείνεται αμφίδρομα στο χρόνο, ήτοι $(..., \underline{X}^{-1}, \underline{X}^0, \underline{X}^1, ..., \underline{X}^k, \underline{X}^{k+1}, ...)$. Επίσης, η μεταβλητή υψηλότερου επιπέδου της τρέχουσας περιόδου $\underline{Z}$ είναι ένας όρος μιας άλλης στοχαστικής ακολουθίας $(..., \underline{Z}^{-1}, \underline{Z}^0, \underline{Z}^1 \equiv \underline{Z}, \underline{Z}^2, ..., \underline{Z}^\tau, ...)$. Στις δύο παραπάνω ακολουθίες $\underline{X}^t$ και $\underline{Z}^\tau$ χρησιμοποιούνται διαφορετικοί χρονικοί δείκτες προκειμένου να αποφευχθούν οι διπλοί δείκτες· αν ο αριθμός των υποπεριόδων (ήτοι ο k) είναι ο ίδιος για όλες τις περιόδους, τότε $\underline{Z}^\tau = \underline{X}^{k(\tau-1)+1} + ... + \underline{X}^{k\tau}$. Ας σημειωθεί ακόμη η συμβολική απλοποίηση $\underline{Z} \equiv \underline{Z}^1$. Εξαιτίας της ανεξαρτησίας των διαδικασιών γέννησης των μεταβλητών υψηλότερου και χαμηλότερου επιπέδου, μπορεί να υποτεθεί ότι οι μεταβλητές υψηλότερου επιπέδου είναι όλες γνωστές πριν την έναρξη του επιμερισμού $(..., \underline{Z}^0 = \underline{z}^0, \underline{Z}^1 = \underline{z}^1, \underline{Z}^2 = \underline{z}^2, ...)$ Επίσης, υποτίθεται ότι η διαδικασία επιμερισμού έχει ολοκληρωθεί σε όλες τις προηγούμενες περιόδους και έτσι όλα τα προηγούμενα $\underline{X}^t$ έχουν γνωστές τιμές $(..., \underline{X}^{-2} = \underline{x}^{-2}, \underline{X}^{-1} = \underline{x}^{-1}, \underline{X}^0 = \underline{x}^0)$.

Ας θεωρήσουμε το επιμεριστικό βήμα t, στην τρέχουσα περίοδο και στη θέση j. Η γέννηση της μεταβλητής χαμηλότερου επιπέδου X_j^t χαρακτηρίζεται από την

$$X_j^t + S_j^{t+1} = S_j^t \tag{41}$$

όπου το υπόλοιπο προς διάθεση S_j^t έχει γνωστή τιμή που δίνεται από την (3), δεδομένου ότι τα προηγούμενα βήματα έχουν ολοκληρωθεί.

Δεδομένου ότι η (41) είναι όμοια με την (6), η γέννηση της X_j^t μπορεί να γίνει με τη διαδικασία διχασμού, όπως αυτή περιγράφηκε στο προηγούμενο κεφάλαιο. Κατά συνέπεια για το κύριο μοντέλο επιμερισμού η διαδικασία διχασμού μπορεί να θεωρηθεί ως μια διαδικασία τύπου μαύρου κουτιού (black-box). Τα εισαγόμενα (είσοδοι) προς αυτή τη διαδικασία από το κύριο μοντέλο είναι η τιμή της S_j^t και μια λίστα από περιθώριες και από κοινού στατιστικές ροπές των (X, S). Τα εξαγόμενα (έξοδοι) που παραλαμβάνει το κύριο μοντέλο είναι οι τιμές των X_j^t και S_j^{t+1}. Κάθε φορά που εκτελείται η διαδικασία διχασμού δεν έχει καμιά πληροφορία σχετικά με τις τιμές των προηγούμενων μεταβλητών χαμηλότερου επιπέδου ούτε σχετικά με την στοχαστική εξάρτηση με τις τιμές αυτές. Ωστόσο, υπάρχει δυνατότητα έμμεσης εισαγωγής στη διαδικασία διχασμού της εξάρτησης από τις προηγούμενες μεταβλητές, χωρίς να χρειαστεί να αλλάξει η δομή της. Αυτό μπορεί να γίνει με την εισαγωγή σε αυτή των δεσμευμένων ροπών των (X, S) δεδομένης της προηγούμενης πληροφορίας. Συγκεκριμένα οι δεσμευμένες αυτές ροπές, σε αντιστοιχία με τις ροπές της διαδικασίας διχασμού, είναι:

Δεσμευμένες περιθώριες ροπές της S_j^t

$$\lambda_1' = \mathrm{E}[S_j^t \mid \Omega_j^t], \qquad \lambda_2 = \mathrm{Var}[S_j^t \mid \Omega_j^t], \qquad \lambda_3 = \mu_3[S_j^t \mid \Omega_j^t]$$

Δεσμευμένες περιθώριες ροπές της X_j^t

$$\eta_1' = \mathrm{E}[X_j^t \mid \Omega_j^t], \qquad \eta_2 = \mathrm{Var}[X_j^t \mid \Omega_j^t], \qquad \eta_3 = \mu_3[X_j^t \mid \Omega_j^t] \tag{42}$$

Δεσμευμένες από κοινού ροπές των (X_j^t, S_j^t)

$$\zeta_{11} = \mathrm{Cov}[X, S_j^t \mid \Omega_j^t], \qquad \zeta_{12} = \mu_{12}[X, S_j^t \mid \Omega_j^t], \qquad \zeta_{21} = \mu_{21}[X, S_j^t \mid \Omega_j^t]$$

όπου Ω_j^t είναι συντομογραφία για την προηγούμενη πληροφορία. Ανάλογα με τη διαδοχή των χρονικών βημάτων και θέσεων σε κάθε περίοδο, μπορούν να υπάρξουν δύο διαφορετικές πορείες του μοντέλου επιμερισμού: η οριζόντια πορεία, στην οποία κάθε χρονικό βήμα ακολουθεί το προηγούμενο βήμα στην ίδια θέση και η κατακόρυφη πορεία, στην οποία κάθε χρονικό βήμα ακολουθείται από το ίδιο χρονικό βήμα στην επόμενη θέση. Στην πρώτη περίπτωση που έχει υιοθετηθεί εδώ, η Ω_j^t γενικά περιλαμβάνει τα ακόλουθα στοιχεία:

Την πλήρη ακολουθία των μεταβλητών Z	$(.., \underline{Z}^2 = \underline{z}^2, \underline{Z}^1 = \underline{z}^1, \underline{Z}^0 = \underline{z}^0, ..)$
Τις μεταβλητές X των προηγούμενων περιόδων σε όλες τις θέσεις	$(\underline{X}^0 = \underline{x}^0, \underline{X}^{-1} = \underline{x}^{-1}, ...)$

Τις μεταβλητές X της τρέχουσας περιόδου στην

- τρέχουσα περίοδο/προηγ. υποπεριόδους $(X_j^{t-1} = x_j^{t-1}, ..., X_j^1 = x_j^1)$
- προηγ. θέση/όλες τις υποπεριόδους $(X_{j-1}^k = x_{j-1}^k, ..., X_{j-1}^1 = x_{j-1}^1)$

$\vdots$ $\vdots$

- πρώτη θέση / όλες τις υποπεριόδους $(X_1^k = x_1^k, ..., X_1^1 = x_1^1)$

'Ομως, ο υπολογισμός των δεσμευμένων είναι πρακτικά αδύνατος αν πρέπει να συμπεριληφθεί στην Ω_j^t όλη η προηγούμενη πληροφορία, όπως αναλύεται παραπάνω. Γι' αυτό πρέπει να γίνουν γίνουν απλοποιήσεις, οι οποίες μπορούν να στηριχτούν σε μια κατάλληλη υπόθεση σχετικά με τη στοχαστική δομή της ακολουθίας των μεταβλητών χαμηλότερου επιπέδου. Μια τέτοια δομή, η οποία απλοποιεί τους υπολογισμούς, (ή μοντέλο) είναι η γραμμική ακολουθία αυτοπαλινδρόμησης. Αυτή η ακολουθία θα χρησιμοποιηθεί εδώ όχι για τη γέννηση των μεταβλητών χαμηλότερου επιπέδου (όπως συνήθως), αλλά για τον υπολογισμό των δεσμευμένων ροπών.

Το αντίστοιχο σειριακό μοντέλο

Η επιλογή του αντίστοιχου σειριακού μοντέλου, δηλαδή του συγκεκριμένου τύπου της ακολουθίας αυτοπαλινδρόμησης των μεταβλητών χαμηλότερου επιπέδου, εξαρτάται από τις στατιστικές παραμέτρους των τελευταίων που είναι επιθυμητό να διατηρηθούν. Εδώ, προκειμένου να κατασκευάσουμε το μοντέλο με τις λιγότερες δυνατές παραμέτρους, χρησιμοποιήσαμε ως αντίστοιχο σειριακό μοντέλο τη Μαρκοβιανή ακολουθία (ακολουθία αυτοπαλινδρόμησης α' τάξης), το σύνολο των στατιστικών παραμέτρων της οποίας αποτελείται από τις ακόλουθες ομάδες [*Matalas & Wallis*, 1976, σ. 60, 63]:

(α) Μέσες τιμές των X_j^t (ξ_j^t).

(β) Διασπορές των X_j^t.

(γ) Συντελεστές ασυμμετρίας των X_j^t.

(δ) Συντελεστές αυτοσυσχέτισης (με βήμα ένα) μεταξύ των X_j^t και X_j^{t-1} (ίδια θέση).

(ε) Συντελεστές ετεροσυσχέτισης (με μηδενικό βήμα) μεταξύ των X_j^t και X_r^t, για $j \neq r$ (ίδια υποπερίοδος).

Οι τρίτες από κοινού ροπές, που η διαδικασία διχασμού μπορεί επίσης να τις υποστηρίξει, δεν αποτελούν ανεξάρτητες παραμέτρους στην παραπάνω Μαρκοβιανή ακολουθία. 'Ομως, οι τιμές τους, όπου αυτές χρειάζονται, μπορούν να υπολογιστούν εύκολα ως εκφράσεις των παραπάνω ομάδων παραμέτρων (βλ. Κεφάλαιο 5 και Προσθήκη Β).

Η διατήρηση των ομάδων στατιστικών παραμέτρων (α) έως (ε) για όλες τις υποπεριόδους απαιτεί την ακόλουθη μόρφωση του Μαρκοβιανού μοντέλου της ακολουθίας $\underline{X}^t$, ως εποχικού μοντέλου AR(1) (το οποίο είναι γνωστό και ως PAR(1))

$$\underline{X}^t = \underline{a}^t \, \underline{X}^{t-1} + \underline{b}^t \, \underline{V}^t \tag{43}$$

ή ισοδύναμα

$$X_j^t = a_j^t \, X_j^{t-1} + \sum_{r=1}^{j} b_{jr}^t \, V_r^t, \; j = 1, \ldots, n \tag{44}$$

όπου $\underline{a}^t$ είναι μητρώο συντελεστών που μπορεί να θεωρηθεί διαγώνιο ($n \times n$), ήτοι $\underline{a}^t$ = diag $\left(a_1^t, a_2^t, \ldots, a_n^t\right)$, $\underline{b}^t = \left[b_{ij}^t\right]$ είναι μητρώο συντελεστών που μπορεί να θεωρηθεί κάτω τριγωνικό διαστάσεων ($n \times n$), και $\underline{V}^t = \left[V_j^t\right]$ είναι διάνυσμα n τυχαίων μεταβλητών, πλήρως ανεξάρτητων και ως προς το χρόνο και ως προς τη θέση. Υποθέτουμε ότι τα $\underline{\xi}^t$ = E[$\underline{X}^t$] και $\underline{\beta}^t$ = E[$\underline{V}^t$] δεν είναι απαραίτητα μηδενικά· θέτουμε (για λόγους μαθηματικής άνεσης) Var[$\underline{V}^t$] = $\underline{1}$ και συμβολίζουμε $\underline{\gamma}^t = \mu_3[\underline{V}^t]$. Η στασιμότητα της ακολουθίας δεν είναι απαραίτητη και γι' αυτό δεν υπονοείται στο συμβολισμό. Ωστόσο, η υπόθεση εποχικής στασιμότητας είναι χρήσιμη σε μερικές από τις παρακάτω αναλύσεις.

Τονίζεται ότι το μοντέλο δεν κάνει καμιά διάκριση για την πρώτη μεταβλητή χαμηλότερου επιπέδου, την X_j^1, για τη οποία εφαρμόζεται επίσης η (44) θεωρώντας ως X_j^0 την τελευταία μεταβλητή της προηγούμενης περιόδου.

Οι ομάδες παραμέτρων $\underline{a}^t$, $\underline{b}^t$ και $\underline{\gamma}^t$ υπολογίζονται από τις παραπάνω ομάδες στατιστικών παραμέτρων (β) έως (ε) με τις ακόλουθες σχέσεις που αποτελούν γενικεύσεις των σχέσεων της στάσιμης περίπτωσης, όπως αυτές δίνονται από τους *Matalas & Wallis* [1976, σ. 63]

$$a_j^t = \frac{\mathrm{Cov}\left[X_j^t, X_j^{t-1}\right]}{\mathrm{Var}\left[X_j^{t-1}\right]} \tag{45}$$

$$\underline{b}^t \left(\underline{b}^t\right)^{\mathrm{T}} = \underline{\sigma}^t - \underline{a}^t \, \underline{\sigma}^{t-1} \, \underline{a}^t \tag{46}$$

$$\gamma_j^t = \frac{\mu_3\left[X_j^t\right] - \left(a_j^t\right)^3 \mu_3\left[X_j^{t-1}\right] - \sum_{k=1}^{j-1} \left(b_{jk}^t\right)^3 \gamma_k^t}{\left(b_{jj}^t\right)^3} \tag{47}$$

όπου

$$\underline{\sigma}^t = \mathrm{Cov}\left[\underline{X}^t, \underline{X}^t\right] = \mathrm{E}\left[(\underline{X}^t - \underline{\xi}^t)\,(\underline{X}^t - \underline{\xi}^t)^{\mathrm{T}}\right] \tag{48}$$

και το κάτω τριγωνικό μητρώο $\underline{b}^t$ υπολογίζεται με αποσύνθεση του Γκραμιανού του $\underline{b}^t$ $(\underline{b}^t)^{\mathrm{T}}$.

Σύνδεση του σειριακού μοντέλου με το επιμεριστικό μοντέλο

Πρέπει να τονιστεί ότι το μοντέλο (43) με παραμέτρους που υπολογίζονται από τις (45) έως (47) αφορά μόνο στις μεταβλητές χαμηλότερου επιπέδου και δεν αποτελεί μοντέλο επιμερισμού. Ωστόσο, συνδυάζοντας την (43) με την (1) ή την (2) μπορούμε να υπολογίσουμε τις από κοινού ροπές μεταξύ των μεταβλητών χαμηλότερου και υψηλότερου επιπέδου, ή πιο γενικά, μεταξύ των μεταβλητών χαμηλότερου επιπέδου και των υπολοίπων προς διάθεση. Εννοείται ότι ο συνδυασμός αποσκοπεί αποκλειστικά στον υπολογισμό των ροπών της (42) και όχι στη γέννηση οποιασδήποτε μεταβλητής, για την οποία υπεύθυνη είναι η διαδικασία διχασμού. 'Ετσι, αυτό που απομένει είναι να αναπτυχθεί η διαδικασία προσδιορισμού ροπών, η οποία μπορεί να βασιστεί στο αντίστοιχο σειριακό μοντέλο. Η ανάπτυξη γίνεται σε δύο διαφορετικές εκδοχές στα Κεφάλαια 5 και 7.

Τέλος τονίζεται ότι το σύνολο των ανεξάρτητων παραμέτρων του μοντέλου επιμερισμού ταυτίζεται με το σύνολο παραμέτρων του αντίστοιχου σειριακού μοντέλου. Καμιά άλλη παράμετρος δεν χρειάζεται για τον προσδιορισμό των ροπών της (42) (βλ. Κεφ. 5 και 7) ενώ η διαδικασία διχασμού δεν τροφοδοτείται με καμιά εξωτερική παράμετρο.

5. Η ΔΙΑΔΙΚΑΣΙΑ ΠΡΟΣΔΙΟΡΙΣΜΟΥ ΡΟΠΩΝ - ΣΧΗΜΑΤΙΣΜΟΣ PAR(1)

Ας εξετάσουμε αναλυτικότερα τους όρους της προηγούμενης πληροφορίας οι οποίοι ορίστηκαν στο προηγούμενο κεφάλαιο. Κατ' αρχήν παρατηρούμε ότι οι μεταβλητές υψηλότερου επιπέδου των προηγούμενων περιόδων (ήτοι οι όροι $\underline{Z}^{-1} = \underline{z}^{-1}$, $\underline{Z}^{-2} = \underline{z}^{-2}$, ...) μπορούν να παραλειφθούν επειδή η σχετική πληροφορία περιέχεται (λόγω της (1)) στην ακολουθία των αντίστοιχων μεταβλητών χαμηλότερου επιπέδου. Με τη γενική πορεία που ακολουθεί το μοντέλο, ο όρος ($\underline{Z} = \underline{z}$) της τρέχουσας περιόδου μπορεί να αντικατασταθεί από το γνωστό υπόλοιπο προς διάθεση, του οποίου ο χειρισμός έχει αφεθεί στη διαδικασία διχασμού και αγνοείται στη διαδικασία προσδιορισμού ροπών. Αν παραλειφθεί από την Ω^t_j και η πληροφορία των επόμενων μεταβλητών υψηλότερου επιπέδου (ήτοι οι όροι $\underline{Z}^2 = \underline{z}^2$, $\underline{Z}^3 = \underline{z}^3$, ...), τότε το μοντέλο απλοποιείται σοβαρά. Ας σημειωθεί ότι παράλειψη των τελευταίων όρων γίνεται σε όλα τα μοντέλα επιμερισμού. Οι συνέπειες αυτής της παράλειψης θα συζητηθούν στο Κεφάλαιο 6.

'Ετσι, με σκοπό την απλοποίηση της διαδικασίας προσδιορισμού ροπών, στο κεφάλαιο αυτό θα αγνοηθούν όλοι οι όροι που αφορούν στην πληροφορία των μεταβλητών υψηλότερου επιπέδου. Παρακάτω, στο Κεφάλαιο 7, θα επιχειρηθεί να εισαχθεί και ο όρος ($\underline{Z}^2 = \underline{z}^2$). Στην Μαρκοβιανή περίπτωση που εξετάζεται η Ω^t_j απλοποιείται περαιτέρω, αφού (εξ ορισμού) την παρούσα κατάσταση την επηρεάζει μόνο η πληροφορία του πιο πρόσφατου χρονικού βήματος. Κατά συνέπεια η Ω^t_j τροποποιείται και γίνεται $(X^{t-1}_j = x^{t-1}_j) \cup (X^k_{j-1} = x^k_{j-1}, ..., X^t_{j-1} = x^t_{j-1}, X^{t-1}_{j-1} = x^{t-1}_{j-1}) \cup ... \cup (X^k_1 = x^k_1, ..., X^t_1 = x^t_1, X^{t-1}_1 = x^{t-1}_1)$.

Η ακόλουθη ανάλυση αποτελεί τη βάση της διαδικασίας προσδιορισμού των ροπών. 'Εστω $\underline{e}^t$ το $(n \times n)$ διαγώνιο μητρώο, του οποίου το (j,j) στοιχείο είναι $e^t_j = b^t_{jj}$. Μπορεί να δειχθεί εύκολα ότι το $(n \times n)$ μητρώο $\underline{d}^t$ που ορίζεται από την

$$\underline{d}^t = \underline{I} - \underline{e}^t \, (\underline{b}^t)^{-1} \tag{49}$$

όπου $\underline{I}$ το μοναδιαίο μητρώο, είναι κατώτερο τριγωνικό με μηδενικά στην κύρια διαγώνιο. Βάσει των $\underline{d}^t$ και $\underline{e}^t$ η (43) γίνεται

$$\underline{X}^t = \underline{a}^t \, \underline{X}^{t-1} + \underline{d}^t \, (\underline{X}^t - \underline{a}^t \, \underline{X}^{t-1}) + \underline{e}^t \, \underline{V}^t \tag{50}$$

και κατά συνέπεια

$$X_j^t = a_j^t X_j^{t-1} + \sum_{r=1}^{j-1} d_{jr}^t (X_r^t - a_r^t X_r^{t-1}) + e_j^t V_j^t \qquad (51)$$

Η τελευταία εξίσωση είναι μια τροποποίηση της (44), στην οποία εμφανίζεται μόνο ένας όρος V_j^t. (Είναι παρόμοια με τις εξισώσεις (10)-(12) των *Stedinger κ.α.* [1985]). Η (51) μπορεί να ξαναγραφεί με τη μορφή

$$X_j^t = T_j^t + U_j^t + W_j^t \qquad (52)$$

όπου

$$T_j^t = a_j^t X_j^{t-1} \qquad (53)$$

$$U_j^t = \sum_{r=1}^{j-1} d_{jr}^t (X_r^t - a_r^t X_r^{t-1}) \qquad (54)$$

$$W_j^t = e_j^t V_j^t \qquad (55)$$

Οι όροι T_j^t και U_j^t είναι εκφράσεις των μεταβλητών χαμηλότερου επιπέδου που περιέχονται στην Ω_j^t, ενώ ο όρος W_j^t είναι μια τυχαία μεταβλητή ανεξάρτητη των T_j^t και U_j^t. Ειδικότερα, ο όρος T_j^t σχετίζεται με την πληροφορία της τρέχουσας θέσης, ενώ ο όρος U_j^t σχετίζεται με τις προηγούμενες θέσεις. 'Ετσι, οι δεσμευμένες ροπές του X_j^t τάξης μεγαλύτερης από 1, δεδομένης της Ω_j^t, εκφράζονται βάσει των ροπών της W_j^t και μόνο. Επιπλέον, σε κάθε βήμα επιμερισμού, η δεσμευμένη μέση τιμή της X_j^t μπορεί εύκολα να προσδιοριστεί χρησιμοποιώντας τις παραπάνω εξισώσεις.

Με επαναλαπτική χρήση των (52) και (53) κάθε όρος X_j^r, $r = t,.., k$ μπορεί να εκφραστεί βάσει των X_j^{t-1}, U_j^r (όρων που περιέχονται στην Ω_j^t) και W_j^r. 'Ετσι, προσθέτοντας τις εξισώσεις για κάθε X_j^r παίρνουμε

$$S_j^t = a_j^t \pi_j^t X_j^{t-1} + \sum_{r=t}^{k} \pi_j^r U_j^r + \sum_{r=t}^{k} \pi_j^r W_j^r \tag{56}$$

όπου

$$\pi_j^r = 1 + \sum_{u=r+1}^{k} a_j^{r+1} a_j^{r+2} \cdots a_j^u \tag{57}$$

'Οπως και στην περίπτωση των δεσμευμένων ροπών της X_j^t που περιγράφηκε παραπάνω, οι δεσμευμένες ροπές της S_j^t, για δεδομένη Ω_j^t, μπορούν να υπολογιστούν χρησιμοποιώντας την (56). Ειδικότερα, οι ροπές τάξης > 1 είναι εκφράσεις των ροπών της W_j^r και μόνο. Επιπλέον, οι εξισώσεις (52) και (56) αν πολλαπλασιαστούν επιτρέπουν τον υπολογισμό των από κοινού δεσμευμένων ροπών των (X_j^t, S_j^t). Συστηματικός αλγόριθμος για τον υπολογισμό όλων των ροπών που απαιτούνται έχει εύκολα καταρτιστεί και δίνεται στην Προσθήκη Β.

6. ΑΝΑΣΚΟΠΗΣΗ ΤΟΥ ΜΟΝΤΕΛΟΥ

Γενικά

'Οπως περιγράφηκε παραπάνω, το μοντέλο αποτελείται από δύο μεμονωμένες διαδικασίες, τη διαδικασία διχασμού και τη διαδικασία προσδιορισμού ροπών. Μελετήθηκαν τρεις μορφές της πρώτης και μια μορφή της τελευταίας. Παρακάτω στο Κεφάλαιο 7 μελετάται και μια άλλη πιο σύνθετη μορφή της διαδικασίας προσδιορισμού ροπών.

Ανεξάρτητα από το σχηματισμό του μοντέλου, η συμπεριφορά του είναι παρόμοια με αυτή του σειριακού PAR(1) (Μαρκοβιανού) μοντέλου, εκτός από το γεγονός ότι οι μεταβλητές χαμηλότερου επιπέδου αφού αθροιστούν δίνουν τις εκ των προτέρων γνωστές τιμές των μεταβλητών υψηλότερου επιπέδου. Λόγω της τελευταίας αυτής ιδιότητάς του, το μοντέλο επιμερισμού υπερτερεί του αντίστοιχου σειριακού, επειδή δεν υπάρχει στατιστική "παραμόρφωση" στις μεταβλητές υψηλότερου επιπέδου, αφού αυτές παράγονται τελείως ανεξάρτητα, από άλλο μοντέλο το οποίο βασίζεται στη διατήρηση των δικών τους στατιστικών χαρακτηριστικών.

Πρέπει να σημειωθεί ότι, με την εξαίρεση κάποιων ειδικών περιπτώσεων που αναλύθηκαν στο Κεφάλαιο 3, η στατιστική διατήρηση που επιτυχάνεται με το μοντέλο επιμερισμού είναι μόνο προσεγγιστική. Η μη ακρίβεια του μοντέλου, με τη αυστηρή έννοια του όρου, οφείλεται σε δομικές ασυνέπειες της διαδικασίας διχασμού, οι οποίες συζητήθηκαν στο Κεφάλαιο 3 και συναντώνται και στα άλλα μοντέλα επιμερισμού.

Η βήμα-προς-βήμα δομή του μοντέλου επιτρέπει τη χρήση παράλληλων διαδικασιών, οι οποίες μπορούν να τροποποιούν κατάλληλα τις τιμές που γεννώνται, χωρίς απώλεια της αθροιστικής ιδιότητας. Π.χ., στην περίπτωση που οι μεταβλητές χαμηλότερου επιπέδου πρέπει να είναι θετικές, μια κατάλληλη παράλληλη διαδικασία μπορεί να χειρίζεται (να απορρίπτει ή να τροποποιεί) τις αρνητικές τιμές που τυχόν παράγονται. Σημειώνεται πάντως, ότι τέτοιες παράλληλες διαδικασίες εισάγουν μεροληψία στις προσομοιωμένες χρονοσειρές και μπορεί να επηρεάζουν το βαθμό ακρίβειας του μοντέλου.

Εξ άλλου, η βήμα-προς-βήμα πορεία του μοντέλου καθιστά εύκολο το χειρισμό της εξάρτησης ανάμεσα σε μεταβλητές χαμηλότερου επιπέδου που ανήκουν σε διαδοχικές μεταβλητές υψηλότερου επιπέδου. Στην πραγματικότητα, η μαθηματική διατύπωση στην περίπτωση αυτή δεν διαφέρει σε τίποτε από τη διατύπωση του μοντέλου για μεταβλητές χαμηλότερου επιπέδου που ανήκουν στην ίδια μεταβλητή υψηλότερου επιπέδου, όπως εξηγήθηκε στο Κεφάλαιο 4.

Σε πρώτη εξέταση, το δυναμικό μοντέλο επιμερισμού (DDM) εμφανίζει παρόμοια συμπεριφορά με τα συμπυκνωμένα μοντέλα επιμερισμού (CDM), αφού και τα δύο σχήματα αποφεύγουν την ταυτόχρονη παραγωγή όλων των μεταβλητών, ενώ χρησιμοποιούν υποσύνολα των παραμέτρων των μοντέλων τύπου Valencia-Schaake. Ωστόσο, υπάρχουν ουσιαστικές διαφορές ανάμεσα στις δύο προσεγγίσεις:

1. Η βήμα-προς-βήμα πορεία του DDM εφαρμόζεται όχι μόνο στα διάφορα χρονικά βήματα (περίπτωση των CDM) αλλά και στις διαφορετικές θέσεις.
2. Το DDM γεννά σε κάθε βήμα όχι μόνο την αντίστοιχη μεταβλητή χαμηλότερου επιπέδου αλλά, επίσης, και το υπόλοιπο προς διάθεση του επόμενου χρονικού βήματος. Τα CDM δεν χρησιμοποιούν την έννοια του υπολοίπου προς διάθεση.
3. Στα CDM κάθε μεταβλητή χαμηλότερου επιπέδου εκφράζεται ως γραμμική συνάρτηση της μεταβλητής υψηλότερου επιπέδου, ενώ το DDM χρησιμοποιεί στη θέση της τελευταίας το υπόλοιπο προς διάθεση.
4. Τα CDM χρησιμοποιούν μη γραμμικούς μετασχηματισμούς των πραγματικών υδρολογικών μεταβλητών, ενώ το DDM χρησιμοποιεί τις ίδιες τις πραγματικές μεταβλητές. Αν χρησιμοποιούνται μη γραμμικοί μετασχηματισμοί από το DDM, αυτοί τοποθετούνται μόνο μέσα στη διαδικασία διχασμού και δεν είναι ορατοί από το κύριο μοντέλο επιμερισμού.
5. Το DDM χειρίζεται τις τρίτες ροπές των μεταβλητών μέσω άμεσων αναλυτικών σχέσεων, ενώ τα CDM δεν κάνουν κανένα τέτοιο χειρισμό.
6. Τα CDM, όπως όλα τα μοντέλα τύπου Valencia-Schaake είναι πλήρως γραμμικά. Το DDM μπορεί να χρησιμοποιεί και ένα μη γραμμικό τμήμα (διαδικασία διχασμού).
7. Ένα από τα στοιχεία του συνόλου παραμέτρων των CDM είναι το μητρώο συνδιασπορών μεταξύ των μεταβλητών χαμηλότερου και υψηλότερου επιπέδου. Το DDM δεν χρησιμοποιεί αυτό το στοιχείο και έτσι περιορίζει το σύνολο των παραμέτρων του σε αυτό ενός τυπικού σειριακού μοντέλου (PAR). Το DDM για να προσδιορίσει συνδιασπορές μεταξύ μεταβλητών υψηλότερου και χαμηλότερου επιπέδου, ή γενικότερα, μεταξύ μεταβλητών χαμηλότερου επιπέδου και υπολοίπων προς διάθεση, αξιοποιεί τις ιδιότητες του αντίστοιχου σειριακού μοντέλου και υπολογίζει τις ποσότητες αυτές βάσει αναλυτικών σχέσεων.

Περιορισμοί του μοντέλου

Οι ουσιαστικοί περιορισμοί του μοντέλου σχετίζονται με τον υπολογισμό των παραμέτρων $\underline{b}^t$ και απαντούν επίσης σε όλα τα μοντέλα επιμερισμού και σειριακά. [*Matalas & Wallis*, 1976, σ. 69, *Bras & Rodriguez-Iturbe*, 1985, σ. 150, *Grygier & Stedinger*, 1990, σ. 31].

Το μητρώο $\underline{c}^t = \underline{b}^t \, (\underline{b}^t)^{\mathrm{T}}$ μπορεί να γραφεί και ως

$$\underline{c}^t = \mathrm{Cov}\left[(\underline{X}^t - \underline{a}^t \, \underline{X}^{t-1}),\ (\underline{X}^t - \underline{a}^t \, \underline{X}^{t-1})\right] \tag{58}$$

και κατά συνέπεια τα στοιχεία του πρέπει να ικανοποιούν τις ανισότητες

$$-1 \leq \frac{c^t_{jk}}{\sqrt{c^t_{jj}\, c^t_{kk}}} \leq 1 \quad \forall\ j, k \tag{59}$$

Επί πλέον, η ύπαρξη του $\underline{b}^t$ προϋποθέτει ότι το $\underline{c}^t$ είναι θετικά ορισμένο. Είναι δυνατό, όμως, να μην ικανοποιούνται οι δύο αυτοί δομικοί περιορισμοί για συγκεκριμένα υδρολογικά δεδομένα, τα οποία χρησιμοποιούνται για τον υπολογισμό του $\underline{c}^t$ βάσει των (45) και (46). Επί πλέον, είναι δυνατό η χρήση της (47) να οδηγήσει σε παράλογα μεγάλους συντελεστές ασυμμετρίας, π.χ. του μεγέθους που κατέγραψε ο *Todini* [1980] ($\gamma >$ 30). Στην πραγματικότητα, δεν υπάρχει όριο για το γ^t_j, δεδομένου ότι ο παρονομαστής $\left(b^t_{jj}\right)^3$ της (47) μπορεί να πάρει πολύ μικρές τιμές. Τα προβλήματα αυτά απαντούν κυρίως σε περιπτώσεις που τα ιστορικά δεδομένα των διάφορων σταθμών αναφέρονται σε διαφορετικές περιόδους μέτρησης.

Πρέπει να τονιστεί ότι αυτά τα προβλήματα συνδέονται με το αντίστοιχο σειριακό μοντέλο PAR [*Matalas & Wallis*, 1976, σ. 69] και όχι με το μοντέλο επιμερισμού αυτό καθεαυτό. Πρακτικές λύσεις για την αντιομετώπιση των προβλημάτων μπορεί να είναι η μείωση των συντελεστών αυτοσυχέτισης ή ετεροσυσχέτισης, ή και η μείωση των συντελεστών ασυμμετρίας των ιστορικών δεδομένων. Είναι προφανές ότι τέτοιες μειώσεις στατιστικών χαρακτηριστικών αποτελούν σημαντικό πρόβλημα. Ωστόσο, στην περίπτωση ενός μοντέλου επιμερισμού η συνολική επίδραση αυτών των τροποποιήσεων είναι μικρότερη, δεδομένου ότι δεν επηρεάζονται τα στατιστικά χαρακτηριστικά των μεταβλητών υψηλότερου επιπέδου.

Σε μια παρόμοια θέση που αφορούσε δείγματα με διαφορετικά μήκη σε διαφορετικές θέσεις, οι *Grygier & Stedinger* [1990, σ. 31-33] συζητούν πρακτικές λύσεις σε ανάλογα προβλήματα, προτείνοντας τροποποιήσεις στα στοιχεία του προς αποσύνθεση Γκραμιανού μητρώου, όταν η διαδικασία της αποσύνθεσης αποτυγχάνει. Πιο συγκεκριμένα, η διαδικασία αποσύνθεσης του $\underline{c}^t = \underline{b}^t \left(\underline{b}^t\right)^{\mathrm{T}}$ μπορει να αποτύχει στο σημείο που πρόκειται να υπολογιστεί το διαγώνιο στοιχείο b^t_{jj} με βάση την εξίσωση

$$\left(b^t_{jj}\right)^2 = c^t_{jj} - \sum_{k=1}^{j-1} \left(b^t_{jk}\right)^2 \tag{60}$$

Αν το δεξιό μέλος της εξίσωσης προκύψει αρνητικό, τότε δεν υπάρχει πραγματική λύση για το b^t_{jj}. Για να το αποφύγουμε αυτό μπορούμε να θέσουμε ένα κατώτερο όριο σε κάθε διαγώνιο στοιχείο b^t_{jj} έτσι ώστε

$$\left(b^{t\,\min}_{jj}\right)^2 = (1 - R^2)\, c^t_{jj} \tag{61}$$

όπου R είναι σταθερά ($0 \le R < 1$). Αν το b^t_{jj} τεθεί ίσο με αυτό το όριο, τότε τα εκτός διαγωνίου στοιχεία της σειράς (b^t_{jk}, $k = 1, ..., j-1$) πρέπει να διορθωθούν με πολλαπλασιασμό επί ένα ενιαίο συντελεστη λ, υπολογισμένοι έτσι ώστε να επανακτηθεί η ισχύς της (60). Εύκολα μπορεί να δειχτεί ότι $\lambda = R \sqrt{c^t_{jj} \,/ \sum_{k=1}^{j-1} (b^t_{jk})^2}$ Με αυτό τον τρόπο εξασφαλίζεται η διατήρηση της $\mathrm{Var}[X^t_j]$ και παράλληλα τίθεται ένα ελάχιστο όριο για την $\mathrm{Var}[W^t_j]$. 'Ενα έμμεσο όφελος από αυτή τη διορθωτική επέμβαση είναι και το γεγονός ότι μειώνεται η ασσυμετρία της W^t_j (ή της V^t_j) και έτσι αποφεύγονται οι παράδοξα μεγάλες τιμές της. Ωστόσο, μειονέκτημα της λύσης αυτής είναι η μείωση των πραγματικών συντελεστών ετεροσυσχέτισης της X^t_j με μεταβλητές χαμηλότερου επιπέδου άλλων θέσεων.

Μια διαφορετική λύση στο ίδιο πρόβλημα μπορεί να είναι η τροποποίηση του μητρώου $\underline{b}^t$ σε τρόπο ώστε να διατηρείται μόνο ο μεγαλύτερος συντελεστής ετεροσυσχέτισης, δηλαδή να υπάρχει μόνο ένα μη μηδενικό μη διαγώνιο στοιχείο σε κάθε σειρά του $\underline{b}^t$. Αυτό έχει συνέπεια την περαιτέρω μείωση του αριθμού των παραμέτρων του μοντέλου (βλ. επόμενο υποκεφάλαιο). 'Αλλες μέθοδοι εκτίμησης παραμέτρων για το σειριακό μοντέλο PAR(1) θα μπορούσε ενδεχομένως να είναι πιο αποτελεσματικές, αλλά δεν έχουν εξεταστεί στην παρούσα μελέτη.

Ορισμένοι δευτερεύοντες περιορισμοί του μοντέλου προκύπτουν από τις περιοριστικές σχέσεις του δευτεροβάθμιου σχήματος της διαδικασίας διχασμού, δηλαδή τις ανισότητες (A19) έως (A21), αν αυτό το σχήμα επιλεγεί στο συγκεκριμένο σχηματισμό του μοντέλου. 'Ενα ενδεικτικό διάγραμμα των ορίων εφαρμοσιμότητας του δευτεροβάθμιου σχήματος διχασμού δίνεται στο Σχήμα 1. Αυτό καταρτίστηκε μετά από αριθμητικές εφαρμογές του μοντέλου για διαφορετικούς συνδυασμούς παραμέτρων και μετά από επιλογή εκείνων των συνόλων παραμέτρων που οδηγούν σε οριακή ικανοποίηση των σχετικών περιορισμών. Από το διάγραμμα αυτό μπορούμε να συμπεράνουμε ότι, άν οι μεταβλητές χαμηλότερου επιπέδου είναι έντονα ασύμμετρες και ταυτόχρονα ισχυρά συσχετισμένες, τότε είναι ενδεχόμενο να ανακύψει πρόβλημα για το μοντέλο. Ωστόσο, όπως προέκυψε από διάφορες πραγματικές εφαρμογές, οι σχετικές περιοριστικές σχέσεις ικανοποιούνται αν οι συντελεστές ασυμμετρίας γ_j^t έχουν λογικές τιμές. Λύση για την περίπτωση πολύ μεγάλων συντελεστών γ_j^t μπορεί να αποτελεί η χρησιμοποίηση του λογαριθμοκανονικού σχήματος διχασμού αντί του δευτεροβάθμιου.

Παράμετροι του μοντέλου

'Οπως αναφέρθηκε στο Κεφάλαιο 4, ο παρών σχηματισμός του δυναμικού μοντέλου επιμερισμού χρησιμοποιεί το ελάχιστο δυνατό σύνολο παραμέτρων, ίδιο με αυτό του σειριακού Μαρκοβιανού μοντέλου. Αλλά είναι δυνατή (αν και όχι γενικά απαραίτητη) και περαιτέρω μείωση του αριθμού των παραμέτρων αν τα μητρώα $\underline{b}^t$ οριστούν έτσι ώστε να έχουν μόνο ένα μη μηδενικό μη διαγώνιο στοιχείο σε κάθε σειρά, σε τρόπο ώστε να διατηρείται μόνο ο μεγαλύτερος συντελεστής ετεροσυσχέτισης για κάθε θέση.

Ο Πίνακας 1 συνοψίζει τον αριθμό των παραμέτρων που χρειάζεται το μοντέλο, σε σύγκριση με άλλα γνωστά μοντέλα επιμερισμού. Είναι προφανές ότι το δυναμικό μοντέλο έχει τις λιγότερες παραμέτρους απ' όλα. Σημειώνεται ότι ο Πίνακας 1 αφορά μόνο στις στατιστικές παραμέτρους δεύτερης τάξης. Στους αριθμούς αυτούς πρέπει να προστεθούν και ο αριθμός των μέσων τιμών και των συντελεστών ασυμμετρίας, καθώς και ο αριθμός των παραμέτρων του μοντέλου που χρησιμοποιείται για τη γέννηση των μεταβλητών υψηλότερου επιπέδου.

Εμφανώς η μείωση του αριθμού των παραμέτρων συνεπάγεται κάποιο κόστος (παράλειψη όρων των πλήρων μητρώων συνδιασπορών), αλλά πάντως το κόστος αυτό δεν συναρτάται με την ικανοποίηση της αθροιστικής ιδιότητας ούτε με την διατήρηση συσχετίσεων ανάμεσα σε μεταβλητές χαμηλότερου επιπέδου που ανήκουν σε διαδοχικές περιόδους.

Διατήρηση περιθώριων κατανομών

Τα Σχήματα 2 (α και β) και 3 δίνουν παραδείγματα της συμπεριφοράς του μοντέλου σχετικά με τη διατήρηση χαρακτηριστικών της περιθώριας κατανομής των μεταβλητών χαμηλότερου επιπέδου. Προέρχονται από δύο εφαρμογές του μοντέλου [*Κουτσογιάννης κ.α.*, 1990], από τις οποίες η πρώτη αναφέρεται στην ταυτόχρονη προσομοίωση βροχής και απορροής των λεκανών Ευήνου, Μόρνου και Υλίκης/Β.Κηφισού, ενώ η δεύτερη αναφέρεται στην προσομοίωση της εξάτμισης από τους ταμιευτήρες των τριών λεκανών. Και στις δύο εφαρμογές συνθετικά ετήσια δεδομένα 5000 ετών επιμερίστηκαν με το μοντέλο σε μηνιαία συνθετικά δεδομένα. Στην πρώτη εφαρμογή υιοθετήθηκε κατανομή γάμα δύο παραμέτρων για τη μηνιαία βροχόπτωση και απορροή και χρησιμοποιήθηκε η δευτεροβάθμια διαδικασία διχασμού. Στη δεύτερη εφαρμογή υιοθετήθηκε κανονική κατανομή για τη μηνιαία εξάτμιση και χρησιμοποιήθηκε η γραμμική διαδικασία δίχασμού. Τα σχήματα δείχνουν καλή επίδοση του μοντέλου ως προς τη διατήρηση των χαρακτηριστικών των περιθώριων κατανομών και στις δύο εφαρμογές. 'Οπως φαίνεται στο Σχήμα 2 (α και β), παρόλο που το μοντέλο προχωρεί βήμα-βήμα στη γέννηση των τιμών των δώδεκα μηνών, ωστόσο δεν υπάρχει συνάθροιση σφαλμάτων στους συντελεστές ασυμμετρίας καθώς προχωρούμε από τον πρώτο στον τελευταίο μήνα.

Μια άλλου τύπου εφαρμογή μιας προηγούμενης έκδοσης του μοντέλου (σημειακής) που αναφέρεται στην γέννηση ωριαίων υψών βροχής, όπου η κατανομή τους ήταν γάμα ή Weibull σχήματος J, μπορεί να βρεθεί αλλού [*Κουτσογιάννης*, 1988, *Koutsoyiannis & Xanthopoulos*, 1990].

Στον Πίνακα 2 γίνεται σύγκριση του δυναμικού μοντέλου επιμερισμού με το μοντέλο Valencia-Schaake. Στο συγκεκριμένο παράδειγμα μια μεταβλητή υψηλότερου επιπέδου Z με κατανομή γάμα επιμερίζεται σε δύο συσχετισμένες μεταβλητές χαμηλότερου επιπέδου X^1 και X^2, ενώ υποτίθεται ότι δεν υπάρχει συσχέτιση μεταξύ διαδοχικών χρονικών βημάτων υψηλότερου επιπέδου. Το δυναμικό μοντέλο εφαρμόστηκε άμεσα στις πραγματικές (χωρίς μετασχηματισμούς) μεταβλητές με χρήση της δευτεροβάθμιας διαδικασίας διχασμού και μιας παράλληλης διαδικασίας που απέρριπτε τις αρνητικές τιμές (Σειρές 1 και 2). Πιο συγκεκριμένα, η Σειρά 1 γεννήθηκε με χρήση των πραγματικών αθροιστών (cumulants) της γνωστής κατανομής της Z. Για τη γέννηση της Σειράς 2 έγινε η εσφαλμένη υπόθεση ότι $\kappa_4 = \kappa_5 = \kappa_6 = 0$ προκειμένου να διερευνηθεί η επίδραση ενός τέτοιου σφάλματος στα στατιστικά χαρακτηριστικά της προσομοιωμένης σειράς. Τα αποτελέσματα δείχνουν ότι αυτή η επίδραση δεν είναι σημαντική.

Για την εφαρμογή του μοντέλου Valencia-Schaake στο παράδειγμα αυτό χρησιμοποιήθηκε ο μετασχηματισμός Wilson-Hilferty στις πραγματικές μεταβλητές, ακολουθούμενος από τον αντίστροφο μετασχηματισμό των συνθετικών δεδομένων. Επειδή υπάρχουν μόνο δύο μεταβλητές χαμηλότερου επιπέδου, το μοντέλο Valencia-Schaake είναι ισοδύναμο με μια διαδικασία διχασμού που χρησιμοποιεί το μετασχηματισμό Wilson-Hilferty σε τρόπο ανάλογο με το λογαριθμικό μετασχηματισμό που περιγράφηκε στο Κεφάλαιο 3. Στην περίπτωση του μοντέλου Valencia-Schaake (Σειρά 3) η αθροιστική ιδιότητα προφανώς δεν διατηρείται. Εφαρμόζοντας στα προσομοιωμένα δεδομένα, μετά τον αντίστροφο μετασχηματισμό τους, τη διορθωτική διαδικασία που περιγράφεται από την εξίσωση (19) η αθροιστική ιδιότητα επανακτήθηκε και επί πλέον διορθώθηκαν οι αρνητικές τιμές (Σειρά 4). Για λόγους σύγκρισης ο ευθύς μετασχηματισμός Wilson-Hilferty εφαρμόστηκε και στις Σειρές 1, 2 και 4 και οι σχετικές στατιστικές παράμετροι των μετασχηματισμένων σειρών φαίνονται επίσης στον Πίνακα 2.

Στον Πίνακα 2 μπορούμε να παρατηρήσουμε τη γενικώς καλή συμπεριφορά του δυναμικού μοντέλου επιμερισμού ως προς τη διατήρηση των στατιστικών παραμέτρων των πραγματικών μεταβλητών. Το μοντέλο Valencia-Schaake εμφανίζει μεγαλύτερες αποκλίσεις (κυρίως στη διασπορά και την τρίτη ροπή της X^2) αλλά αυτό δεν είναι αναγκαστικά σοβαρό μειονέκτημα του συγκεκριμένου μοντέλου, επειδή δεν όφειλε να αναπαράγει τις παραμέτρους αυτές. Σχετικά με τις στατιστικές παραμέτρους των μετασχηματισμένων μεταβλητών, το δυναμικό μοντέλο δεν εμφανίζει πλήρη προσαρμογή, αλλά ούτε όφειλε. Οι μετασχηματισμένες μεταβλητές της Σειράς 3 βρίσκονται σε συμφωνία με τις θεωρητικές τιμές των ροπών. Ωστόσο, η συμφωνία αυτή εξαφανίζεται μετά την εφαρμογή της διορθωτικής διαδικασίας, όπως φαίνεται από τα δεδομένα της Σειράς 4. Προφανώς τα μετασχηματισμένα και διορθωμένα δεδομένα δεν ακολουθούν πλέον κατανομή Gauss (δεν έχουν μηδενική τρίτη ροπή) και δεν έχουν μοναδιαία διασπορά ως όφειλαν. Η μεροληψία έχει εισαχθεί από τη διασπορά των συντελεστων διόρθωσης που εφαρμόζονται για την επανάκτηση της αθροιστικής ιδιότητας.

Διατήρηση των συσχετίσεων

Στο Σχήμα 2(γ), το οποίο προέρχεται από την εφαρμογή του μοντέλου που αναφέρθηκε πιο πάνω, δίνεται σαφής ένδειξη της διατήρησης των συντελεστών αυτοσυσχέτισης (για βήμα ένα) που επιτυγχανεται από το δυναμικό μοντέλο επιμερισμού. 'Ομοια είναι και η συμφωνία με τις θεωρητικές τιμές των ταυτόχρονων συντελεστών ετεροσυσχέτισης για διαφορετικές θέσεις. Από το ίδιο σχήμα προκύπτει ότι η καλή συμφωνία εμπειρικών και θεωρητικών τιμών επεκτείνεται επίσης και στη συσχέτιση της απορροής του πρώτου μήνα της τρέχουσας περιόδου με την απορροή του τελευταίου μήνα της προηγούμενης περιόδου. Σαφέστερη εικόνα για το ειδικότερο αυτό θέμα δίνει ο Πίνακας 3, ο οποίος αναφέρεται σε όλες τις σειρές απορροής και εξάτμισης των δύο εφαρμογών που αναφέρθηκαν παραπάνω (οι συντελεστές αυτοσυσχέτισης της βροχής δεν ήταν σημαντικά διάφοροι του μηδενός). Υπενθιμίζεται ότι για την επίτευξη της καλής αυτής επίδοσης του μοντέλου δεν χρειάζεται κανένας ειδικός χειρισμός, δεδομένου ότι το μοντέλο εφαρμόζεται χωρίς τροποποίηση και για την πρώτη μεταβλητή μιας περιόδου, X_j^1 χρησιμοποιώντας ως προηγούμενη πληροφορία την τελευταία μεταβλητή χαμηλότερου επιπέδου της προηγούμενης περιόδου ($X_j^0 = x_j^0$).

Από πρακτική άποψη η διατήρηση των των συντελεστών αυτοσυσχέτισης μεταξύ διαδοχικών μεταβλητών χαμηλότερου επιπέδου, που επιτυγχάνεται με το μοντέλο, είναι επαρκής. Ωστόσο, θεωρητικά δεν μπορούμε να μιλήσουμε για διατήρηση αλλά μόνο για προσέγγιση. Συστηματική διερεύνηση της συμπεριφοράς του μοντέλου ως προς το θέμα αυτό παρουσιάζεται στο Σχήμα 4, το οποίο συνοψίζει τα αποτελέσματα μιας σειράς προσομοιώσεων που αφορούν στον επιμερισμό μιας μεταβλητής υψηλότερου επιπέδου μιας θέσης σε δύο συνιστώσες. Ο συντελεστής συσχέτισης μεταξύ διαδοχικών μεταβλητών χαμηλότερου επιπέδου που ανήκουν σε διαδοχικές περιόδους κυμαίνεται από −1 και +1. Τα αποτελέσματα του σχήματος αυτού εξηγούνται παρακάτω.

Οι *Stedinger & Vogel* [1984] έδειξαν ότι τα μοντέλα επιμερισμού δεν είναι δυνατό να διατηρούν τις συσχετίσεις μεταξύ των μεταβλητών χαμηλότερου επιπέδου της προηγούμενης περιόδου με τις μεταβλητές υψηλότερου επιπέδου της τρέχουσας περιόδου, ή, ισοδύναμα και με το συμβολισμό που χρησιμοποιείται εδώ, τις ποσότητες Corr[$\underline{X}^t$, $\underline{Z}^2$], $t = 1,..,k$. Δεδομένου ότι οι ποσότητες αυτές χρησιμοποιούνται ως παράμετροι στο μοντέλο Mejia-Rousselle, το εν λόγω μοντέλο επηρεάζεται συνολικά από αυτή την ασυνέπεια. Το αυθεντικό μοντέλο Valencia-Schaake δεν χρησιμοποιεί τις ποσότητες αυτές αλλά ούτε αναπαράγει τις συσχετίσεις μεταξύ διαδοχικών μεταβλητών χαμηλότερου επιπέδου που ανήκουν σε διαδοχικές περιόδους. Το δυναμικό μοντέλο επιμερισμού, όπως και το μοντέλο Valencia-Schaake δεν χρησιμοποιεί αυτές τις ποσότητες ως παραμέτρους αλλά συμπεριφέρεται καλά ως προς τη διατήρηση των τελευταίων συσχετίσεων. 'Ομως, λεπτομερέστερη μελέτη δείχνει ότι η συμπεριφορά του είναι ίδια με αυτή του μοντέλου Mejia-Rousselle. Στα παραδείγματα του Σχήματος 4 τα δύο μοντέλα έδωσαν τις ίδιες ακριβώς καμπύλες οι οποίες δεν διακρίνονται μεταξύ τους (οι διακεκομμένες γραμμές αντιπροσωπεύουν ταυτόχρονα και τα δύο μοντέλα). Στο Σχήμα 4 μπορούμε να παρατηρήσουμε εμφανείς αποκλίσεις της Corr[X^t, Z^2] των προσομοιωμένων σειρών από τις θεωρητικές τιμές. Οι αποκλίσεις της Corr[X^1, X^2] είναι αμελητέες αλλά υπάρχει μια μικρή διαφορά στην Corr[X^0, X^1], η οποία παίρνει τη μεγαλύτερη τιμή της γύρω από τις θέσεις ±0.5.

Κατά συνέπεια, η αστοχία στην αναπαραγωγή των ποσοτήτων Corr[$\underline{X}^t$, $\underline{Z}^2$] επηρεάζει σε ένα βαθμό την πλήρη δομή αυτοσυσχέτισης του μοντέλου ανεξάρτητα από την παρουσία τους ή όχι στη διατύπωση του μοντέλου. Αλλά ποιός είναι ο λόγος μιας τέτοιας επίδρασης, αν αυτός δεν είναι η παρουσία ποσοτήτων όπως οι Corr[$\underline{X}^t$, $\underline{Z}^2$] στη διατύπωση του μοντέλου; Η απάντηση στο ερώτημα αυτό μπορεί να διατυπωθεί ως εξής: Τα μοντέλα επιμερισμού είναι τεχνικές που πυκνώνουν δειγματοσυναρτήσεις στοχαστικών ανελίξεων, οι οποίες είναι ήδη γνωστές σε μια αραιότερη κλίμακα (που καθορίζεται από το χρονικό βήμα υψηλότερου επιπέδου). Το αυθεντικό μοντέλο Valencia-Schaake δεν "βλέπει" έξω από κάθε χρονικό βήμα υψηλότερου επιπέδου και γι' αυτό εμφανίζει τη γνωστή αδυναμία στην αναπαραγωγή συσχετίσεων. Το μοντέλο Mejia-Rousselle καθώς και ο παραπάνω σχηματισμός του δυναμικού μοντέλου έχουν τρόπο να "δουν" προς το παρελθόν αλλά όχι και προς το μέλλον και έτσι εμφανίζουν παρόμοια συμπεριφορά. Αλλά στη συγκεκριμένη περίπτωση των μοντέλων επιμερισμού το μέλλον είναι γνωστό στη χρονική κλίμακα υψηλότερου επιπέδου και έτσι είναι δυνατή μια βελτιωμένη έκδοση του μοντέλου που να μπορεί να "δει" και προς τις δύο κατευθύνσεις. Αυτό επιχειρείται να γίνει στο επόμενο κεφάλαιο.

Πίνακας 1: Σύγκριση του αριθμού των στατιστικών παραμέτρων δεύτερης τάξης σε διάφορα μοντέλα επιμερισμού

Τύπος μοντέλου	Αριθμός στατ. παραμέτρων δεύτερης τάξης				
		Για k=12 και			
	Γενικά	$n=1$	$n=3$	$n=6$	$n=10$
"Πλήρη" μοντέλα επιμερισμού					
Valencia - Schaake	$kn(kn+2n+1)/2$	90	774	3060	8460
Mejia - Rousselle	$kn(3kn+2n+1)/2$	234	2070	8244	22860
Συμπυκνωμένα μοντέλα επιμερισμού					
Μοντέλο LAST (από τους *Grygier & Stedinger* [1988])	$kn(5n+1)/2 - n^2$	35	279	1080	2960
Μοντέλο *Stedinger-Pei-Cohn* [1985]	$kn(7n+1)/2 - 3n^2$	45	369	1440	3960
Σταδιακά μοντέλα επιμερισμού					
Μοντέλο *Stedinger and Vogel* [1984]	$kn(n+5)/2-n+90$	125	231	480	980
Δυναμικό μοντέλο επιμερισμού (βασισμένο στο σειριακό Μαρκοβιανό μοντ.)					
Πλήρες μέγεθος (όλες οι ετεροσυσχ.)	$kn(n+3)/2$	24	108	324	780
Μειωμένο μέγεθος (μία ετεροσυσχ.)	$k(3n-1)$	24	96	204	348

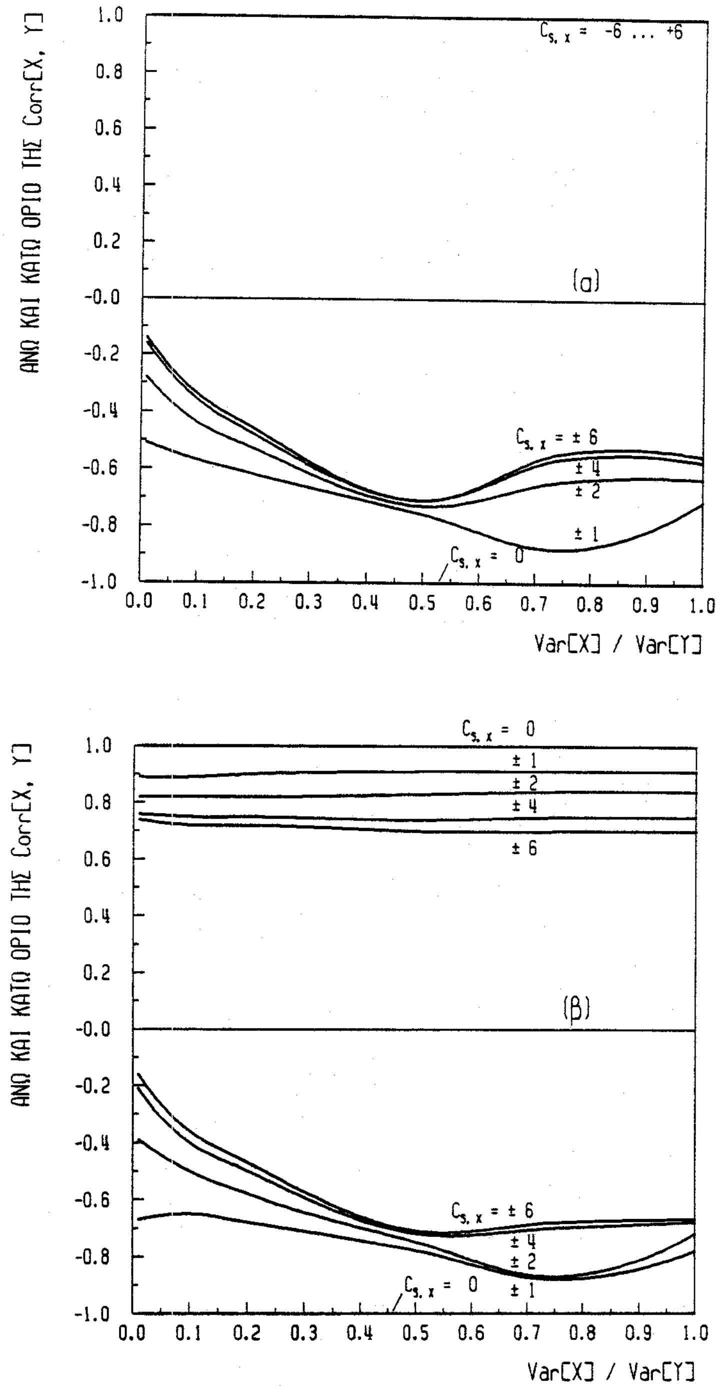
Πίνακας 2: Παράδειγμα σύγκρισης του δυναμικού μοντέλου επιμερισμού (DDM) με το μοντέλο Valencia-Schaake (VSM) για γάμα μεταβλητές. Το μοντέλο Valencia-Schaake χρησιμοποιεί το μετασχηματισμό Wilson-Hilferty και συνδυάζεται με μια διορθωτική διαδικασία για την επανάκτηση της αθροιστικής ιδιότητας. Τα μεγέθη των συνθετικών δειγμάτων είναι 16000.

Στατιστική παράμετρος	Θεωρητικές τιμές	Εμπειρικές τιμές συνθ. σειρών: DDM/δευτεροβάθμιος διχασμός		Εμπειρικές τιμές συνθ. σειρών: VSM/Wilson-Hilferty	
		Γάμα cumulants	Gauss cumulants	Αρχικά δεδομένα	Διορθωμένα δεδομένα
		Σειρά 1	Σειρά 2	Σειρά 3	Σειρά 4
Πραγματικές μεταβλητές					
$E[Z]$	3.000	3.067	3.069	3.031	3.034
$E[X^1]$	1.000	1.024	1.024	1.006	1.006
$E[X^2]$	2.000	2.043	2.045	2.019	2.028
$Var[Z]$	8.000	8.242	7.998	8.508	8.489
$Var[X^1]$	1.000	1.053	1.023	0.997	0.958
$Var[X^2]$	4.000	4.037	3.939	4.351	4.427
$\mu_3[Z]$	40.503	44.910	41.890	65.665	65.667
$\mu_3[X^1]$	2.000	2.311	2.326	2.061	1.807
$\mu_3[X^2]$	16.000	16.417	14.744	31.148	29.723
$Corr[X^1, X^2]$	0.750	0.765	0.756	0.753	0.754
$Corr[X^1, Z]$	0.884	0.893	0.888	0.853	0.880
$Corr[X^2, Z]$	0.972	0.973	0.972	0.978	0.975
Μετασχηματισμένες μεταβλητές (Wilson-Hilferty)					
$E[Z]$	0.000	0.053	0.057	0.010	0.023
$E[X^1]$	0.000	0.060	0.073	0.008	–0.030
$E[X^2]$	0.000	0.030	0.019	0.010	–0.051
$Var[Z]$	1.000	0.882	0.873	1.004	0.948
$Var[X^1]$	1.000	0.866	0.827	0.996	1.180
$Var[X^2]$	1.000	0.984	1.036	0.998	1.269
$\mu_3[Z]$	0.000	0.271	0.237	0.045	0.191
$\mu_3[X^1]$	0.000	0.299	0.287	0.019	–0.505
$\mu_3[X^2]$	0.000	0.005	–0.114	0.071	–0.486
$Corr[X^1, X^2]$	0.782*	0.797	0.777	0.783	0.781
$Corr[X^1, Z]$	0.872*	0.881	0.866	0.873	0.892
$Corr[X^2, Z]$	0.981*	0.982	0.980	0.981	0.956

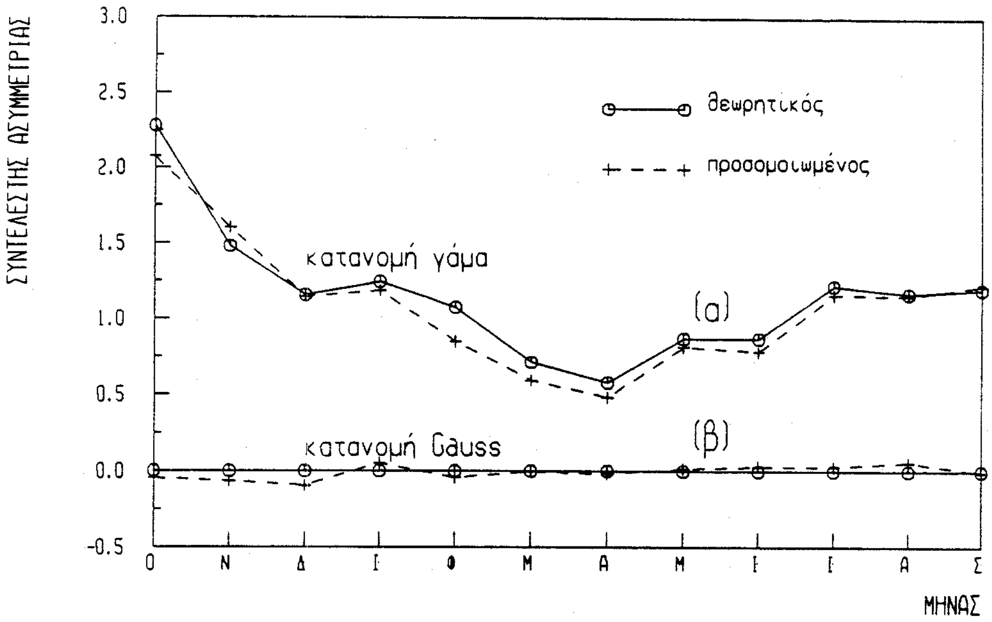
* εκτιμήσεις από προσομοίωση

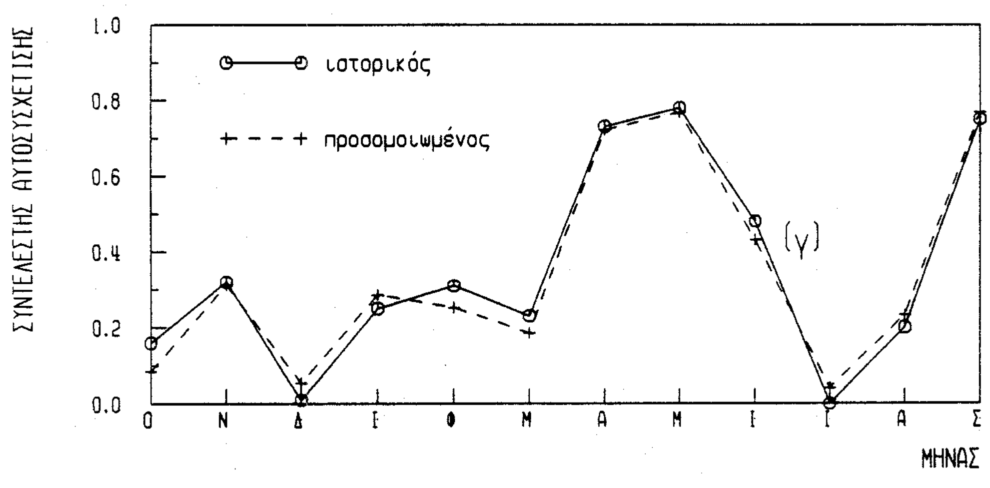
Πίνακας 3: Συντελεστές συσχέτισης των μεταβλητών χαμηλότερου επιπέδου μεταξύ της πρώτης υποπέριόδου της τρέχουσας περιόδου και της τελευταίας υποπεριόδου της προηγούμενης περιόδου, για τις εφαρμογές 1 και 2.

Εφαρ-μογή	Μεταβλητή - Θέση	Συνάρτηση κατανομής	Ιστορικά δεδομένα	Συνθετικά δεδομένα
1	Απορροή - Εύηνος	Γάμα	0.32	0.29
1	Απορροή - Μόρνος	Γάμα	0.16	0.09
1	Απορροή - Υλίκη	Γάμα	0.59	0.54
2	Εξάτμιση - Εύηνος	Gauss	0.75	0.76
2	Εξάτμιση - Μόρνος	Gauss	0.02	0.03
2	Εξάτμιση - Υλίκη	Gauss	0.07	0.07

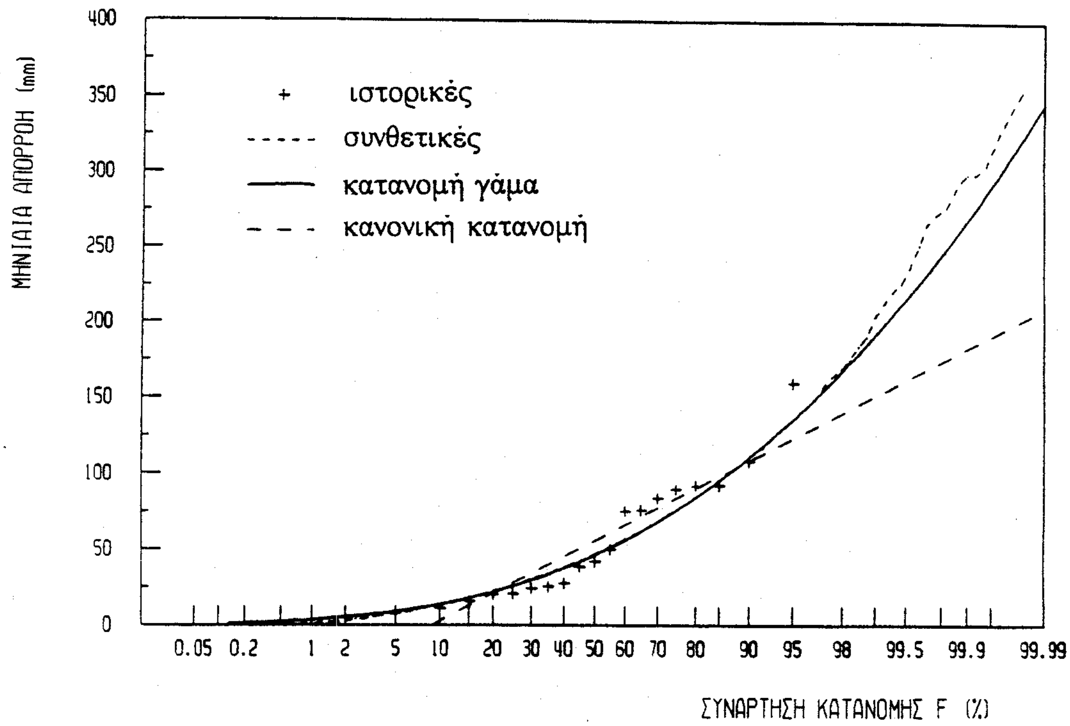


Σχήμα 1: Ενδεικτικό διάγραμμα για το πεδίο εφαρμοσιμότητας της δευτεροβάθμιας διαδικασίας διχασμού. Οι καμπύλες αντιπροσωπεύουν τα ανώτερα και κατώτερα όρια των συντελεστών συσχέτισης των μεταβλητών χαμηλότερου επιπέδου (Corr[X, Y]) που οδηγούν σε πραγματικές λύσεις των εξισώσεων. Τα όρια αυτά δίνονται συναρτήσει του λόγου διασπορών (Var[X] / Var[Y]) για δύο συνδυασμούς συντελεστών ασυμμετρίας των μεταβλητών χαμηλότερου επιπέδου: (α) $C_{S,X} = C_{S,Y}$ και (β) $C_{S,X} = 2C_{S,Y}$

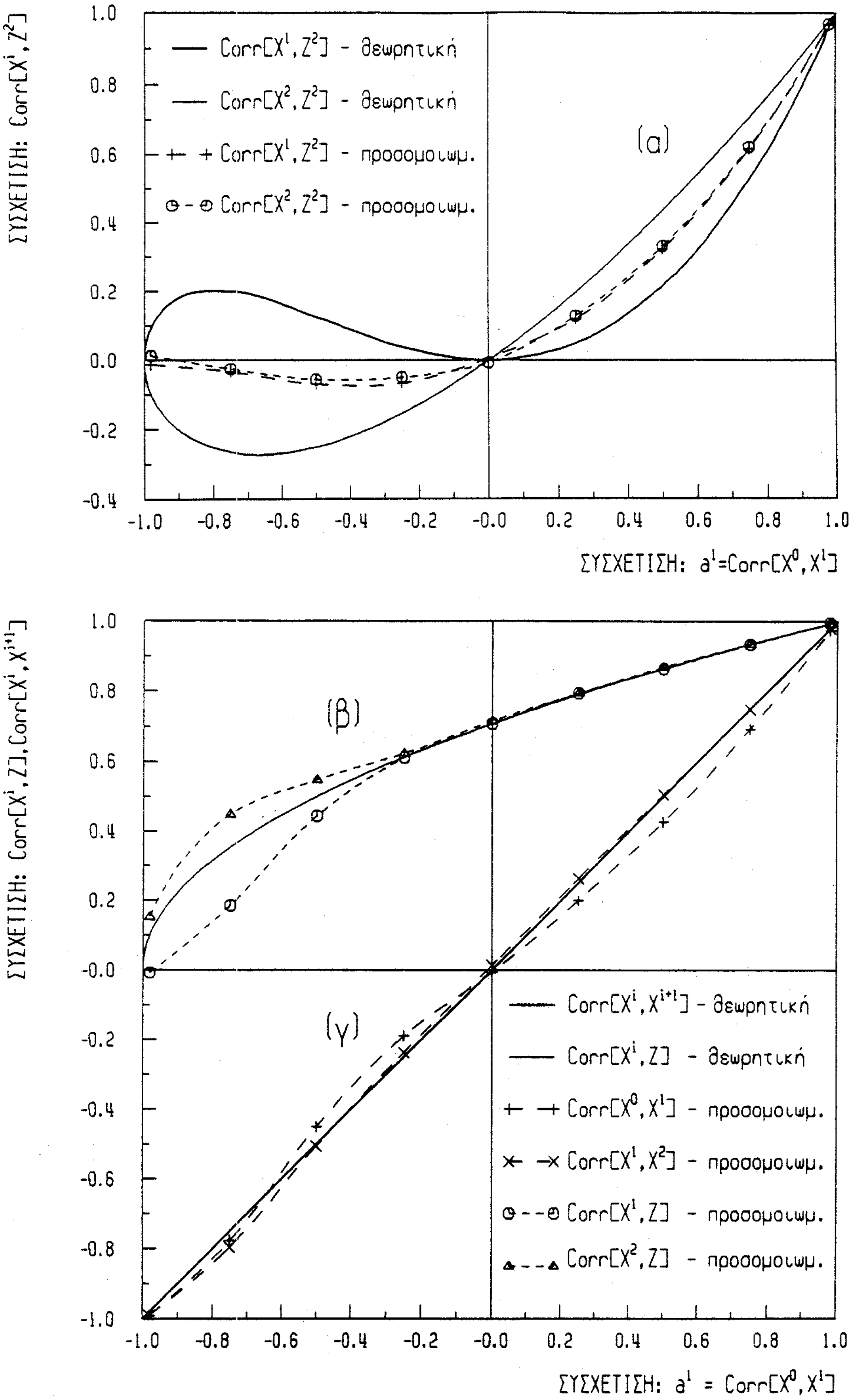




Σχήμα 2: Σύγκριση των μηνιαίων συντελεστών ασυμμετρίας και συσχέτισης συνθετικών δεδομένων (α) συντελεστής ασυμμετρίας, εφαρμογή 1, θέση 3 (απορροή Μόρνου), (β) συντελεστής ασυμμετρίας, εφαρμογή 2, θέση 2 (εξάτμιση ταμιευτήρα Μόρνου), (γ) συντελεστής αυτοσυσχέτισης για βήμα ένα των μεταβλητών χαμηλότερου επιπέδου, εφαρμογή 1, θέση 3 (απορροή Μόρνου).



Σχήμα 3: Σύγκριση της εμπειρικής και θεωρητικής κατανομής των μηνιαίων τιμών για την εφαρμογή 1, θέση 3, υποπερίοδο 2 (απορροή Νοεμβρίου στο Μόρνο). Η κατανομή γάμα δύο παραμέτρων υιοθετήθηκε για την προσομοίωση.



Σχήμα 4: Ενδεικτικό διάγραμμα της επίδοσης του σχηματισμού PAR(1) του μοντέλου σε σχέση με τη διατήρηση της συσχέτισης (α) των μεταβλητών χαμηλότερου επιπέδου της τρέχουσας περιόδου με τις μεταβλητές υψηλότερου επιπέδου της επόμενης περιόδου, (β) των μεταβλητών χαμηλότερου επιπέδου και υψηλότερου επιπέδου της τρέχουσας περιόδου και (γ) διαδοχικών μεταβλητών χαμηλότερου επιπέδου της ίδιας ή διαδοχικών περιόδων. Οι καμπύλες του μοντέλου Mejia-Rousselle συμπίπτουν με αυτές του DDM (διακεκομμένες γραμμές). Οι υποθέσεις για την κατάρτιση του διαγράμματος είναι (1) επιμερισμός μιας μεταβλητής υψηλότερου επιπέδου σε δύο συνιστώσες, X^1 και X^2, (2) $\mathrm{Var}[X^1] = \mathrm{Var}[X^2] = 1$ και (3) $\mathrm{Corr}[X^0, X^1] = \mathrm{Corr}[X^1, X^2]$. Το μέγεθος των συνθετικών δειγμάτων είναι 16000.

7. ΣΧΗΜΑΤΙΣΜΟΣ PARX(1)

Διατηρώντας τη Μαρκοβιανή δομή της ακολουθίας των μεταβλητών χαμηλότερου επιπέδου (εξισώσεις (43) έως (47)), θα επιχειρήσουμε να τροποποιήσουμε τη διαδικασία προσδιορισμού παραμέτρων σε τρόπο ώστε να παίρνει υπόψη τις γνωστές τιμές των μεταβλητών υψηλόερου επιπέδου, $\underline{Z}^2$. Αυτό μπορεί να γίνει με εφαρμογή της ακόλουθης εξίσωσης, στην οποία έχει εισαχθεί η $\underline{Z}^2$ ως "εξωγενής είσοδος":

$$\underline{X}^t = \underline{f}^t\, \underline{X}^{t-1} + \underline{g}^t\, \underline{Z}^2 + \underline{h}^t\, \underline{Q}^t \tag{62}$$

ή της ισοδύναμης

$$X^t_j = f^t_j\, X^{t-1}_j + g^t_j\, Z^2_j + \sum_{r=1}^{j} h^t_{jr}\, Q^t_r\,,\ j = 1, ..., n \tag{63}$$

όπου τα μητρώα $\underline{f}^t$ και $\underline{g}^t$ έχουν υποτεθεί $(n \times n)$ διαγώνια, ήτοι $\underline{f}^t$ = diag $(f^t_1, f^t_2, ..., f^t_n)$, $\underline{g}^t$ = diag $(g^t_1, g^t_2, ..., g^t_n)$, ενώ το $\underline{h}^t = [h^t_{ij}]$ είναι κατώτερο τριγωνικό μητρώο συντελεστών με διαστάσεις $(n \times n)$. Το $\underline{Q}^t = [Q^t_j]$ είναι διάνυσμα n τυχαίων μεταβλητών ανεξάρτητων πλήρως και ως προς το χρόνο και ως προς τη θέση και ανεξάρτητων από τις Z^2_j. Συμβολίζουμε επίσης $\underline{\zeta}^t = \mathrm{E}[\underline{Q}^t]$, $\underline{\eta}^t = \mu_3[\underline{Q}^t]$, και θέτουμε (για λόγους μαθηματικής ευκολίας) $\mathrm{Var}[\underline{Q}^t] = \underline{1}$.

Τα μητρώα παραμέτρων $\underline{f}^t$, $\underline{g}^t$ και $\underline{h}^t$ μπορούν να υπολογιστούν από το σύνολο παραμέτρων του αντίστοιχου σειριακού μοντέλου, χωρίς την εισαγωγή καμίας άλλης ανεξάρτητης παραμέτρου, βάσει των ακόλουθων εξισώσεων που εξάγονται στην Προσθήκη Γ.

$$\sigma^{t-1}_{jj}\, f^t_j + \alpha^{t-1}_j\, \sigma^{t-1}_{jj}\, g^t_j = a^t_j\, \sigma^{t-1}_{jj} \tag{64}$$

$$\alpha^{t-1}_j\, \sigma^{t-1}_{jj}\, f^t_j + \upsilon_{jj}\, g^t_j = \alpha^t_j\, \sigma^t_{jj} \tag{65}$$

$$\underline{h}^t (\underline{h}^t)^{\mathrm{T}} = \underline{\sigma}^t - \underline{f}^t\, \underline{\sigma}^{t-1}\, \underline{a}^t - \underline{g}^t\, \underline{\alpha}^t\, \underline{\sigma}^t \tag{66}$$

όπου

$$\underline{\alpha}^t = \mathrm{diag}\left(\alpha_1^t, \alpha_2^t, ..., \alpha_n^t\right) = \left[\sum_{l=1}^{k} \underline{a}^l ... \underline{a}^1\right] \underline{a}^k ... \underline{a}^{t+1} \tag{67}$$

$$\underline{\upsilon} = \mathrm{Cov}[\underline{Z}, \underline{Z}] \tag{68}$$

ενώ τα $\underline{\sigma}^t$ και $\underline{a}^t$ ορίζονται από τις (48) και (45) αντίστοιχα.

Συγκεκριμένα, οι εξισώσεις (64) και (65), επιλυόμενες ταυτοχρόνως, δίνουν τα f_j^t και g_j^t, ενώ το $\underline{h}^t$ εξάγεται από την (66) με αποσύνθεση.

Για τον πλήρη καθορισμό των παραμέτρων απομένουν τα διανύσματα $\underline{\zeta}^t$ και $\underline{\eta}^t$. Το πρώτο μπορεί εύκολα να υπολογιστεί παίρνοντας αναμενόμενες τιμές στην (62). Ο υπολογισμός του δεύτερου είναι πολύπλοκος, δεδομένου ότι υπεισέρχονται τρίτης τάξης από κοινού ροπές των $\underline{Z}^2$ και $\underline{X}^{t-1}$. μπορεί πάντως να γίνει με τρόπο ανάλογο αυτού που χρησιμοποιείται στην Προσθήκη Γ για την εξαγωγή παρόμοιων δεύτερης τάξης ροπών. Κατόπιν θα πρέπει να αξιοποιηθεί η ακόλουθη εξίσωση που προέκυψε από την (63):

$$\begin{aligned} \eta_j^t = (h_{jj}^t)^{-3} \Big\{ & \mu_3\left[X_j^t\right] - (f_j^t)^3 \mu_3\left[X_j^{t-1}\right] - (g_j^t)^3 \mu_3\left[Z_j^2\right] \\ & - 3(f_j^t)^2 g_j^t \mu_{21}\left[X_j^{t-1}, Z_j^2\right] - 3f_j^t (g_j^t)^2 \mu_{12}\left[X_j^{t-1}, Z_j^2\right] \\ & - \sum_{r=1}^{j-1} (h_{jr}^t)^3 \eta_r^t \Big\} \end{aligned} \tag{69}$$

Με τρόπο ανάλογο αυτού που που ακολουθήσαμε στο σχηματισμό PAR(1) (Κεφάλαιο 5), προχωρούμε στη διατύπωση των εξισώσεων που αποτελούν τη βάση για τον προσδιορισμό των ροπών σε κάθε βήμα. Εδώ το διαγώνιο μητρώο $\underline{e}^t$ ορίζεται με την $e_j^t = h_{jj}^t$ και το κατώτερο τριγωνικό μητρώο $\underline{d}^t$ (με μηδενικές τιμές στη διαγώνιο, όπως στην (49)) ορίζεται από την

$$\underline{d}^t = \underline{I} - \underline{e}^t \left(\underline{h}^t\right)^{-1} \tag{70}$$

Με τη χρήση των $\underline{d}^t$ και $\underline{e}^t$ η (62) γίνεται

$$\underline{X}^t = \underline{f}^t\,\underline{X}^{t-1} + \underline{g}^t\,\underline{Z}^2 + \underline{d}^t\,(\underline{X}^t - \underline{f}^t\,\underline{X}^{t-1} - \underline{g}^t\,\underline{Z}^2) + \underline{e}^t\,\underline{Q}^t \tag{71}$$

Αυτή η μορφή είναι όμοια με την (50) και πάλι καταλήγει στην

$$X_j^t = T_j^t + U_j^t + W_j^t \tag{52}$$

όπου οι ποσότητες T_j^t, U_j^t και W_j^t τώρα ορίζονται με διαφορετικό τρόπο, ήτοι

$$T_j^t = f_j^t\,X_j^{t-1} + g_j^t\,Z_j^2 \tag{72}$$

$$U_j^t = \sum_{r=1}^{j-1} d_{jr}^t\,(X_r^t - f_r^t\,X_r^{t-1} - g_r^t\,Z_r^2) \tag{73}$$

$$W_j^t = e_j^t\,Q_j^t \tag{74}$$

Αξιοποιώντας τις (52) και (72) έως (74), η S_j^t μπορεί να εκφραστεί με μια εξίσωση ανάλογη της (56):

$$S_j^t = f_j^t\,\pi_j^t\,X_j^{t-1} + \sum_{r=t}^{k} g_j^r\,\pi_j^r\,Z_j^2 + \sum_{r=t}^{k} \pi_j^r\,U_j^r + \sum_{r=t}^{k} \pi_j^r\,W_j^r \tag{75}$$

όπου

$$\pi_j^r = 1 + \sum_{u=r+1}^{k} f_j^{r+1}\,f_j^{r+2}\cdots f_j^u \tag{76}$$

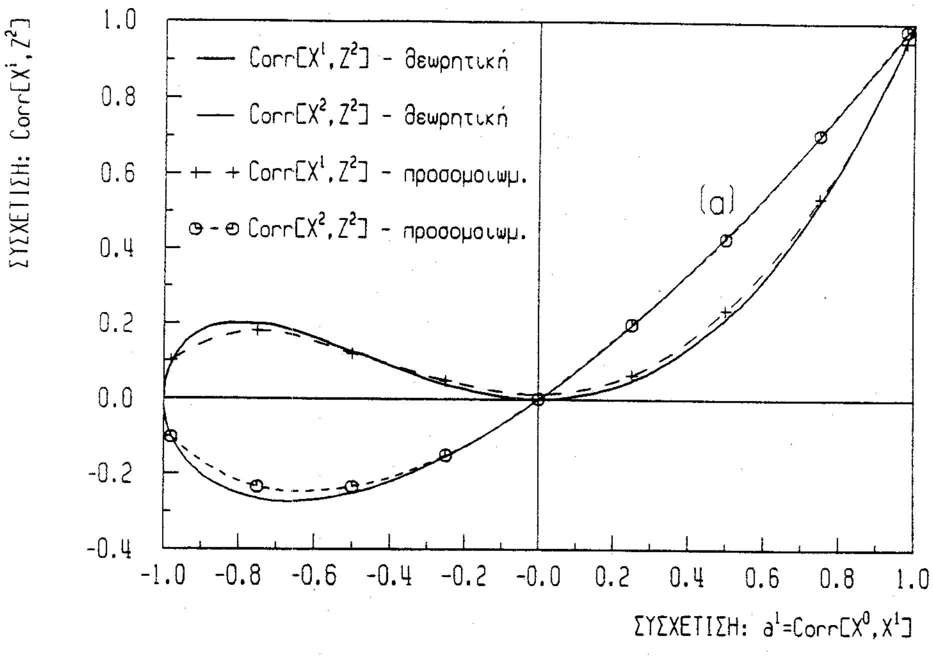
Οι εξισώσεις (52) και (75) μπορούν να εφαρμοστούν για τον υπολογισμό των δεσμευμένων περιθώριων και από κοινού ροπών των (X_j^t, S_j^t) και έτσι μπορεί να καταρτιστεί ένας συστηματικός αλγόριθμος ανάλογος με αυτόν της Προσθήκης Β.

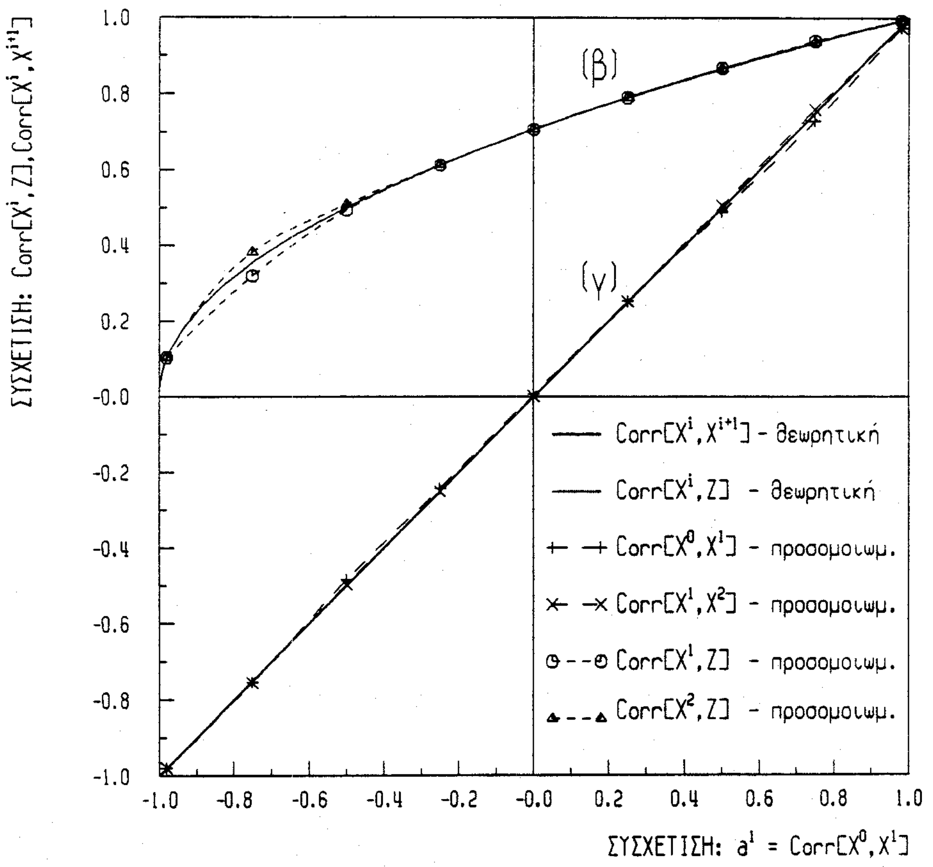
'Ενα δευτερεύον πρόβλημα που σχετίζεται με τον παραπάνω σχηματισμό προέρχεται από την ύπαρξη του όρου α_j^0 στην (64), όταν αυτή εφαρμόζεται για $t = 1$. Ο όρος αυτός αντιπροσωπεύει τη συνδιασπορά της μεταβλητής χαμηλότερου επιπέδου X_j^0, η οποία ανήκει στην προηγούμενη περίοδο, με τη μεταβλητή υψηλότερου επιπέδου Z_j^2 της επόμενης περιόδου. Δεδομένου ότι το μοντέλο διατηρεί συσχετίσεις μόνο μεταξύ διαδοχικών περιόδων, το α_j^0 δεν διατηρείται. Ωστόσο, η επίδραση αυτού του προβλήματος δεν είναι τόσο σημαντική. Πρακτικές λύσεις για την αντιμετώπιση του προβλήματος είναι (α) Να θεωρηθεί $\underline{\alpha}^0 = \underline{0}$ για τυπικές τιμές της συσχέτισης μεταξύ μεταβλητών υψηλότερου επιπέδου. (β) Για σημαντικά μεγάλες τιμές αυτής της συσχέτισης να υπολογιστεί το $\underline{\alpha}^0$ από την (67). (γ) Να χρησιμοποιηθεί ως προσέγγιση μια μέση τιμή $\underline{\alpha}'^0$ στη θέση του $\underline{\alpha}^0$, ήτοι

$$\underline{\alpha}'^0 = \frac{1}{k} \sum_{r=1-k}^{0} \underline{\alpha}^r \tag{77}$$

όπου το $\underline{\alpha}^r$ υπολογίζεται από την (67), με την παρατήρηση ότι $\underline{a}^{m-k} = \underline{a}^m$. Η προσέγγιση αυτή έδωσε καλά αποτελέσματα για οποιαδήποτε τιμή της συσχέτισης μεταξύ διαδοχικών μεταβλητών υψηλότερου επιπέδου.

Παρόλο που οι δύο σχηματισμοί PAR(1) και PARX(1) έχουν το ίδιο αντίστοιχο σειριακό μοντέλο, ο δεύτερος είναι σαφώς πολυπλοκότερος του πρώτου, ιδίως στο χειρισμό των τρίτων ροπών. Ωστόσο ο σχηματισμός PARX(1) συμπεριφέρεται πολύ καλά σε όλα τα είδη συσχέτισης μεταξύ μεταβλητων χαμηλότερου και υψηλότερου επιπέδου. Το συμπέρασμα αυτό προκύπτει από το Σχήμα 5 (παρόμοιο με το Σχήμα 4), το οποίο αναφέρεται στην αριθμητική διερεύνηση του προηγούμενου κεφαλαίου αλλά με τη χρήση του σχηματισμού PARX(1) του μοντέλου. Θα πρέπει να σημειωθεί ότι η μέθοδος δεν χρειάζεται καμιά υπόθεση σχετικά με τη δομή της ακολουθίας των αθροιστικών μεταβλητών και τον τρόπο γέννησης τους. Με αυτή την έννοια είναι δομικά διαφορετική από μια πρόσφατη μέθοδο του *Lin* [1990], η οποία εξαρτάται από το ειδικό μοντέλο που χρησιμοποιήθηκε για τη γέννηση των αθροιστικών μεταβλητών και αποσκοπεί στην τροποποίηση του τρόπου εκτίμησης παραμέτρων του μοντέλου Mejia-Rousselle έχοντας στόχο τη διατήρηση των συνδιασπορών των μεταβλητών χαμηλότερου επιπέδου.





Σχήμα 5: Ενδεικτικό διάγραμμα της επίδοσης του σχηματισμού PARX(1) του μοντέλου σε σχέση με τη διατήρηση της συσχέτισης (α) των μεταβλητών χαμηλότερου επιπέδου της τρέχουσας περιόδου με τις μεταβλητές υψηλότερου επιπέδου της επόμενης περιόδου, (β) των μεταβλητών χαμηλότερου επιπέδου και υψηλότερου επιπέδου της τρέχουσας περιόδου και (γ) διαδοχικών μεταβλητών χαμηλότερου επιπέδου της ίδιας ή διαδοχικών περιόδων. Οι υποθέσεις είναι ίδιες με αυτές του Σχήματος 4.

8. ΣΥΜΠΕΡΑΣΜΑΤΑ

Το πολυμεταβλητό δυναμικό μοντέλο επιμερισμού που αναπτύχθηκε βασίζεται σε μια όχι κατ' ανάγκη γραμμική διαδικασία διχασμού και συνδέεται με ένα αντίστοιχο σειριακό μοντέλο που στη συγκεκριμένη έκδοση είναι Μαρκοβιανό. Τα σημαντικά χαρακτηριστικά του μοντέλου είναι (α) η εξασφάλιση της διατήρησης της αθροιστικής ιδιότητας, (β) η βήμα-προς-βήμα πορεία του, (γ) η αρθρωτή δομή του (αποτελείται από δύο τμήματα που μελετώνται και εφαρμόζονται ξεχωριστά) και (δ) οι άμεσης επίλυσης αναλυτικές σχέσεις που χρησιμοποιεί.

Το μοντέλο είναι φειδωλό σε παραμέτρους, αφού χρησιμοποιεί το ελάχιστο δυνατό σύνολο στατιστικών παραμέτρων των μεταβλητών χαμηλότερου επιπέδου (περιθώριες ροπές πρώτης, δεύτερης και τρίτης τάξης, αυτοσυσχετίσεις μοναδιαίου βήματος και ετεροσυσχετίσεις μηδενικού βήματος). Η εκτίμηση των παραμέτρων γίνεται άμεσα από τα ιστορικά δεδομένα με τις συνήθεις μεθόδους που χρησιμοποιούνται και για το σειριακό Μαρκοβιανό μοντέλο.

Είναι δυνατοί διάφοροι σχηματισμοί του μοντέλου, οι οποίοι προκύπτουν από τη χρήση είτε διάφορων διαδικασιών διχασμού, είτε διάφορων διαδικασιών προσδιορισμού ροπών, είτε διάφορων αντίστοιχων σειριακών μοντέλων. Δύο ιδιαίτεροι σχηματισμοί του μοντέλου έχουν μελετηθεί εδώ, που έχουν το ίδιο αντίστοιχο σειριακό μοντέλο και συγκεκριμένα το Μαρκοβιανό. Ο πρώτος σχηματισμός (PAR(1)) χρησιμοποιεί το σειριακό Μαρκοβιανό μοντέλο στην αρχική του μορφή, ενώ ο δεύτερος (PARX(1)) προσθέτει ως "εξωγενή είσοδο" και τη γνωστή τιμή της μεταβλητής υψηλότερου επιπέδου της επόμενης περιόδου. Ο σχηματισμός PAR(1) είναι απλούστερος από τον PARX(1), ιδίως στο χειρισμό των τρίτων ροπών των μεταβλητών χαμηλότερου επιπέδου, και δίνει ικανοποιητικές προσεγγίσεις των συσχετίσεων μεταξύ μεταβλητών χαμηλότερου επιπέδου που ανήκουν σε διαφορετικές περιόδους. Ωστόσο, ο σχηματισμός PARX(1) παρουσιάζει καλύτερη συμπεριφορά σε σχέση με τη διατήρηση των κάθε είδους συνδιασπορών μεταξύ μεταβλητών χαμηλότερου και υψηλότερου επιπέδου.

Είναι προφανές ότι το μοντέλο έχει ένα αυστηρό περιορισμό που αφορά στην Μαρκοβιανή δομή που υποτίθεται για την ακολουθία των μεταβλητών χαμηλότερου επιπέδου· αυτό είναι το κόστος που εισάγει η απαίτηση της ελαχιστοποίησης του αριθμού των παραμέτρων. Ωστόσο, αυτός ο περιορισμός δεν είναι τόσο σημαντικός για ένα μοντέλο επιμερισμού, δεδομένου ότι τα τυχόν σφάλματα δεν αθροίζονται στις υψηλότερες κλίμακες, αφού η γέννηση των μεταβλητών υψηλότερου επιπέδου γίνεται ανεξάρτητα. Με την εξαίρεση ορισμένων ειδικών περιπτώσεων, το μοντέλο, όμοια με τα άλλα μοντέλα επιμερισμού, δεν είναι ακριβές με την αυστηρή μαθηματική έννοια, αλλά δίνει καλές προσεγγίσεις των στατιστικών χαρακτηριστικών που ενδιαφέρουν.

ΠΡΟΣΘΗΚΗ Α
ΕΞΑΓΩΓΗ ΤΩΝ ΠΑΡΑΜΕΤΡΩΝ
ΤΗΣ ΔΕΥΤΕΡΟΒΑΘΜΙΑΣ ΔΙΑΔΙΚΑΣΙΑΣ ΔΙΧΑΣΜΟΥ

Εισάγοντας στα δεξιά μέλη των (28), (29) και (31) τους όρους a_i και επιλύοντας το σύστημα παίρνουμε

$$a_1 = \frac{(\lambda_4 - \lambda_2^2)\zeta_{11} - \lambda_3 \zeta_{12}}{(\lambda_4 - \lambda_2^2)\lambda_2 - \lambda_3^2} \tag{A1}$$

$$a_2 = \frac{\lambda_2 \zeta_{12} - \lambda_3 \zeta_{11}}{(\lambda_4 - \lambda_2^2)\lambda_2 - \lambda_3^2} \tag{A2}$$

$$a_0 = -a_2 \lambda_2 \tag{A3}$$

Οι εξισώσεις των b_i, που προσδιορίζουν την f(*S*), εξάγονται από τις (30) και (32), μετά από ένα αριθμό πράξεων. Στο παρακάτω σύνολο εξισώσεων ο βαθμός ελευθερίας έχει αποδοθεί στο b_2, όπως επεξηγείται περαιτέρω στη συνέχεια.

$$b_0 = \sqrt{\frac{p}{\beta_2^2 \lambda_4 + 2\beta_1 \beta_2 \lambda_3 + (\beta_1^2 + 2\beta_2)\lambda_2 + 1}} \tag{A4}$$

$$b_1 = \beta_1 b_0 \tag{A5}$$

$$b_2 = \beta_2 b_0 \tag{A6}$$

όπου

$$\beta_2 = \begin{cases} 0 & \tau_0 \geq 0 \\ \left(-\tau_1 + \sqrt{\Delta_2}\right) / \tau_2 , & \tau_0 < 0,\ \tau_1 \geq 0 \\ \left(-\tau_1 - \sqrt{\Delta_2}\right) / \tau_2 , & \tau_0 < 0,\ \tau_1 < 0 \end{cases}$$

$$\beta_1 = \begin{cases} \dfrac{-(\lambda_4 - q\,\lambda_3)\beta_2 - \lambda_2 \pm \sqrt{\Delta_1}}{\lambda_3 - q\,\lambda_2} , & \lambda_3 - q\,\lambda_2 \neq 0 \\ \dfrac{1}{2}\,\dfrac{q - (\lambda_5 - q\,\lambda_4)\beta_2^2}{\lambda_2 + (\lambda_4 - q\,\lambda_3)\beta_2} , & \lambda_3 - q\,\lambda_2 = 0 \end{cases}$$

$$\tau_0 = (\lambda_3 - q\,\lambda_2)\,q + \lambda_2^2$$

$$\tau_1 = (\lambda_4 - q\,\lambda_3)\,\lambda_2 - (\lambda_3 - q\,\lambda_2)^2$$

$$\tau_2 = (\lambda_4 - q\,\lambda_3)^2 - (\lambda_3 - q\,\lambda_2)(\lambda_5 - q\,\lambda_4)$$

$$\Delta_2 = \tau_1^2 - \tau_0\,\tau_2$$

$$\Delta_1 = \tau_0 + 2\tau_1\,\beta_2 + \tau_2\,\beta_2^2$$

$$q = \left[\zeta_{21} - a_2^2\,\lambda_5 - 2a_1\,a_2\,(\lambda_4 - \lambda_2^2) - (a_1^2 - 2a_2^2\,\lambda_2)\,\lambda_3\right] / p$$

$$p = \eta_2 - a_2^2\,(\lambda_4 - \lambda_2^2) - 2a_1 a_2 \lambda_3 - a_1^2\,\lambda_2$$

Το νόημα της εξίσωσης (Α7) είναι ότι το β_2 (άρα και το b_2) τίθεται ίσο με μηδέν αν αυτό είναι δυνατό (δηλαδή αν αυτό δεν συνεπάγεται μη πραγματικές ρίζες του b_1) με αποτέλεσμα την μετατροπή της δευτεροβάθμιας εξίσωσης σε γραμμική. Διαφορετικά παίρνει την ελάχιστη (κατ' απόλυτη τιμή) τιμή που εξασφαλίζει $\Delta_1 \geq 0$ και συνεπώς πραγματική ρίζα του b_1.

Σημειώνεται ότι η (Α8) κανονικά δίνει δύο διαφορετικές τιμές του β_1 και κατά συνέπεια δύο ζεύγη (b_0, b_1). Από αυτά επιλέγεται εκείνο που δίνει τη μικρότερη απόλυτη τιμή του συντελεστή ασυμμετρίας θ_3. Το θ_3 υπολογίζεται από την (33) η οποία μπορεί να γραφεί με τη μορφή

$$\theta_3 = \left[\eta_3 - \varphi(a,a,a) - 3\varphi(a,b,b)\right] / \varphi(b,b,b) \tag{A16}$$

όπου το σύμβολο φ(.,.,.) αποτελεί συντομογραφία της ακόλουθης έκφρασης, στην οποία τα k, l, m συμβολίζουν τριάδες παραμέτρων. π.χ. $k = (k_0, k_1, k_2)$ κτλ.

$$\begin{aligned}\varphi(k,l,m) = {} & k_0 l_0 m_0 + \\ & + (k_0 l_0 m_2 + k_0 l_1 m_1 + k_0 l_2 m_0 + k_1 l_0 m_1 + k_1 l_1 m_0 + k_2 l_0 m_0)\, \lambda_2 \\ & + (k_0 l_1 m_2 + k_0 l_2 m_1 + k_1 l_0 m_2 + k_1 l_1 m_1 + k_1 l_2 m_0 + k_2 l_0 m_1 + k_2 l_1 m_0)\, \lambda_3 \\ & + (k_0 l_2 m_2 + k_1 l_1 m_2 + k_1 l_2 m_1 + k_2 l_0 m_2 + k_2 l_1 m_1 + k_2 l_2 m_0)\, \lambda_4 \\ & + (k_1 l_2 m_2 + k_2 l_1 m_2 + k_2 l_2 m_1)\, \lambda_5 + k_2 l_2 m_2\, \lambda_6\end{aligned} \tag{A17}$$

Μπορεί να αποδειχτεί ότι

$$p = \mathrm{E}[X^2] - \mathrm{E}[\mathrm{g}^2(S)] \tag{A18}$$

και λόγω της (30) θα πρέπει

$$p > 0 \tag{A19}$$

Είναι δυνατό ορισμένοι συνδυασμοί των αρχικών παραμέτρων της (8) να δίνουν τιμές των a_i που δεν ικανοποιούν την (Α19). Στην περίπτωση αυτή, παρόλο που τα a_i υπολογίζονται, η g(S) όπως ορίζεται από τις (25) και (26) δεν υπάρχει· κατά συνέπεια η (Α19) είναι αναγκαία συνθήκη για την ύπαρξη της g(S). 'Ομοια, αναγκαία συνθήκη για την ύπαρξη της f(S) αποτελεί η

$$\beta_2^2 \lambda_4 + 2\beta_1 \beta_2 \lambda_3 + (\beta_1^2 + 2\beta_2) \lambda_2 + 1 > 0 \tag{A20}$$

που προκύπτει από την (A4). Περαιτέρω, η ύπαρξη της f(S) απαιτεί

$$\tau_0 \geq 0 \quad \text{ή} \ \Delta_2 \geq 0 \tag{A21}$$

όπως προκύπτει από μελέτη της (A7).

Αν οι μεταβλητές X και Y (ή, ισοδύναμα, οι X και S) ακολουθούν από κοινού κανονική κατανομή, τότε το δευτεροβάθμιο σχήμα μεταπίπτει σε γραμμικό και οι εξισώσεις (25) έως (27) μεταπίπτουν στην (9). Πράγματι, το παραπάνω σύνολο εξισώσεων γράφεται

$$a_0 = a_2 = b_1 = b_2 = 0 \tag{A22}$$

$$a_1 = \zeta_{11} / \lambda_2 \tag{A23}$$

$$b_0 = \sqrt{(\eta_2 - \zeta_{11}^2 / \lambda_2)} \tag{A24}$$

όπως θεωρητικά αναμένονταν. Αυτό ισχύει επί πλέον και σε κάθε περίπτωση όπου $\lambda_3 = \zeta_{12} = \zeta_{21} = 0$.

ΠΡΟΣΘΗΚΗ Β
ΒΑΣΙΚΕΣ ΕΞΙΣΩΣΕΙΣ ΤΟΥ ΑΛΓΟΡΙΘΜΟΥ ΠΡΟΣΔΙΟΡΙΣΜΟΥ ΡΟΠΩΝ ΣΤΟ ΣΧΗΜΑΤΙΣΜΟ PAR(1)

Ο αλγόριθμος περιλαμβάνει τον υπολογισμό των δεσμευμένων ροπών της (42), οι οποίες χρειάζονται για τη διαδικασία διχασμού. Ο υπολογισμός βασίζεται στις ακόλουθες εξισώσεις που προκύπτουν από τις (52) και (56) (τα αντίστοιχα σύμβολα των μεγεθών της διαδικασίας διχασμού, όπως αυτά ορίζονται στην (8), δίνονται σε παρένθεση στην αρχή κάθε γραμμής). Οι αποδείξεις δίνονται στο τέλος αυτού της προσθήκης.

(λ_1)
$$E\left[S_j^t \mid \Omega_j^t\right] = a_j^t \pi_j^t x_j^{t-1} + \tau_j^t \tag{B1}$$

(λ_2) $$\mathrm{Var}\left[S_j^t \mid \Omega_j^t\right] = \varphi_j^t \tag{B2}$$

(λ_3) $$\mu_3\left[S_j^t \mid \Omega_j^t\right] = \psi_j^t \tag{B3}$$

(η_1) $$\mathrm{E}\left[X_j^t \mid \Omega_j^t\right] = a_j^t\, x_j^{t-1} + \delta_j^t \tag{B4}$$

(η_2) $$\mathrm{Var}\left[X_j^t \mid \Omega_j^t\right] = \left(e_j^t\right)^2 \tag{B5}$$

(η_3) $$\mu_3\left[X_j^t \mid \Omega_j^t\right] = \left(e_j^t\right)^3 \gamma_j^t \tag{B6}$$

(ζ_{11}) $$\mathrm{Cov}\left[X_j^t,\ S_j^t \mid \Omega_j^t\right] = \pi_j^t \left(e_j^t\right)^2 \tag{B7}$$

(ζ_{12}) $$\mu_{12}\left[X_j^t,\ S_j^t \mid \Omega_j^t\right] = \left(\pi_j^t\right)^2 \left(e_j^t\right)^3 \gamma_j^t \tag{B8}$$

(ζ_{21}) $$\mu_{21}\left[X_j^t,\ S_j^t \mid \Omega_j^t\right] = \pi_j^t \left(e_j^t\right)^3 \gamma_j^t \tag{B9}$$

όπου

$$\delta_j^t = \mathrm{E}\left[U_j^t \mid \Omega_j^t\right] + \mathrm{E}\left[W_j^t\right] = \sum_{r=1}^{j-1} d_{jr}^t (x_r^t - a_r^t\, x_r^{t-1}) + e_j^t\, \beta_j^t \tag{B10}$$

και τα ακόλουθα σύμβολα ορίζονται από τις εξής αναδρομικές εξισώσεις

$$\pi_j^t = 1 + a_j^{t+1}\pi_j^{t+1}, \qquad \pi_j^{k+1} = 0 \tag{B11}$$

$$\tau_j^t = \delta_j^t\, \pi_j^t + \tau_j^{t+1}, \qquad \tau_j^{k+1} = 0 \tag{B12}$$

$$\varphi_j^t = \left(\pi_j^t\right)^2 \left(e_j^t\right)^2 + \varphi_j^{t+1}, \qquad \varphi_j^{k+1} = 0 \tag{B13}$$

$$\psi_j^t = \left(\pi_j^t\right)^3 \left(e_j^t\right)^3 \gamma_j^t + \psi_j^{t+1}, \qquad \psi_j^{k+1} = 0 \tag{B14}$$

Ας σημειωθεί ότι η (B11) είναι ισοδύναμη με την (57), αλλά υπολογιστικά απλούστερη.

Απόδειξη της (B11): Από την (57) έχουμε $\pi_j^t = 1 + a_j^{t+1} + a_j^{t+1} a_j^{t+2} + \ldots + a_j^{t+1} a_j^{t+2} \cdots a_j^k = 1 + a_j^{t+1}(1 + a_j^{t+2} + \ldots + a_j^{t+2} \cdots a_j^k) = 1 + a_j^{t+1} \pi_j^{t+1}$. Για $t = k$ (λόγω της (2)) πρέπει να είναι $S_j^k = X_j^k$ και (λόγω των (52) και (53)) $S_j^k = a_j^k X_j^{k-1} + U_j^k + W_j^k$. Από σύγκριση με την (56) παίρνουμε $\pi_j^k = 1$. Αυτό συμβαδίζει με την αναδρομική σχέση αν θέσουμε $\pi_j^{k+1} = 0$.

Απόδειξη της (B4): Από την (52) έχουμε $\mathrm{E}\left[X_j^t \mid \Omega_j^t\right] = \mathrm{E}\left[T_j^t \mid \Omega_j^t\right] + \mathrm{E}\left[U_j^t \mid \Omega_j^t\right] + \mathrm{E}\left[W_j^t\right]$. Υπενθυμίζουμε ότι τα T_j^t και U_j^t είναι γνωστά αφού αποτελούν εκφράσεις μεταβλητών που περιέχονται στην Ω_j^t. Από την (53) παίρνουμε $\mathrm{E}\left[T_j^t \mid \Omega_j^t\right] = a_j^t x_j^{t-1}$, από την (54) $\mathrm{E}\left[U_j^t \mid \Omega_j^t\right] = \sum_{r=1}^{j-1} d_{jr}^t(x_r^t - a_r^t x_r^{t-1})$ και από την (55) $\mathrm{E}\left[W_j^t\right] = e_j^t \beta_j^t$. Άρα, αν εισάγουμε το δ_j^t, όπως ορίζεται από την (B10) παίρνουμε την (B4).

Απόδειξη των (B1) *και* (B12): Από την (56) έχουμε $\mathrm{E}\left[S_j^t \mid \Omega_j^t\right] = a_j^t \pi_j^t x_j^{t-1} + \mathrm{E}\left[\sum_{r=t}^{k} \pi_j^r (U_j^r + W_j^r) \mid \Omega_j^t\right] = a_j^t \pi_j^t x_j^{t-1} + \sum_{r=t}^{k} \pi_j^r \delta_j^r$. Σημειώνεται ότι, λόγω της οριζόντιας πορείας που ακολουθούμε, όλα τα U_j^r για $r > t$ είναι γνωστά, αφού είναι εκφράσεις μεταβλητών χαμηλότερου επιπέδου προηγούμενων θέσεων. Επίσης σημειώνεται ότι $\mathrm{E}\left[(U_j^t + W_j^t) \mid \Omega_j^t\right] = \mathrm{E}\left[U_j^t \mid \Omega_j^t\right] + \mathrm{E}\left[W_j^t\right] = \delta_j^t$. Έτσι, αν συμβολίσουμε $\sum_{r=t}^{k} \pi_j^r \delta_j^r = \tau_j^t$, παίρνουμε την (B1). Κατόπιν, γίνεται προφανές ότι αυτός ο ορισμός του τ_j^t οδηγεί στην αναδρομική σχέση (B12).

Απόδειξη των (B5) *και* (B6): Αν αφαιρέσουμε τις δεσμευμένες μέσες τιμές από την (52) (για δεδομένη Ω_j^t) και μετά υψώσουμε στο τετράγωνο και πάρουμε δεσμευμένες αναμενόμενες τιμές, παίρνουμε $\mathrm{Var}\left[X_j^t \mid \Omega_j^t\right] = \mathrm{Var}\left[W_j^t\right] = (e_j^t)^2\, \mathrm{Var}\left[V_j^t\right]$ (λόγω της (55)). Δεδομένου ότι $\mathrm{Var}\left[V_j^t\right] = 1$, η (B5) έχει αποδειχτεί. Σημειώνεται ότι τα T_j^t και U_j^t δεν συνεισφέρουν στη δεσμευμένη διασπορά επειδή είναι γνωστά. Η απόδειξη της (B6) είναι όμοια.

Απόδειξη των (B2), (B13), (B3) *και* (B14): Αν αφαιρέσουμε τις δεσμευμένες μέσες τιμές από την (56) (για δεδομένη Ω_j^t) και μετά υψώσουμε στο τετράγωνο και πάρουμε δεσμευμένες αναμενόμενες τιμές, παίρνουμε $\mathrm{Var}\left[S_j^t \mid \Omega_j^t\right] = \sum_{r=t}^{k} (\pi_j^r)^2 \mathrm{Var}\left[W_j^r\right] = \sum_{r=t}^{k} (\pi_j^r)^2 (e_j^r)^2$. Αν συμβολίσουμε με φ_j^t το τελευταίο άθροισμα, εύκολα προκύπτει ότι το φ_j^t δίνεται επίσης και από την αναδρομική σχέση (B13). Σημειώνεται ότι τα X_j^{t-1} και U_j^r στην (56) δεν συνεισφέρουν στη δεσμευμένη διασπορά επειδή είναι γνωστά. Επίσης, υπενθυμίζεται ότι τα W_j^r, $r = t, ..., k$, είναι ανεξάρτητα. Οι αποδείξεις των (B3) και (B14) είναι πανομοιότυπες.

Αποδείξεις των (B7), (B8) *και* (B9): Αν αφαιρέσουμε τις δεσμευμένες μέσες τιμές από καθεμιά από τις (52) και (56) (για δεδομένη Ω_j^t) και μετά τις πολλαπλασιάσουμε και πάρουμε δεσμευμένες αναμενόμενες τιμές, παίρνουμε $\mathrm{Cov}\left[X_j^t, S_j^t \mid \Omega_j^t\right] = \pi_j^t \mathrm{Var}\left[W_j^t\right] = \pi_j^t (e_j^t)^2$. Πανομοιότυπη είναι η απόδειξη των (B8) και (B9).

ΠΡΟΣΘΗΚΗ Γ
ΠΑΡΑΓΩΓΗ ΤΩΝ ΕΞΙΣΩΣΕΩΝ ΤΟΥ ΣΧΗΜΑΤΙΣΜΟΥ PARX(1)

Αφαιρώντας τις μέσες τιμές από την (63) και μετά πολλαπλασιάζοντας διαδοχικά με $(X_j^{t-1} - \xi_j^{t-1})$ και $(Z_j^2 - \mathrm{E}[Z_j^2])$ παίρνουμε αντίστοιχα

$$\mathrm{Var}\left[X_j^{t-1}\right] f_j^t + \mathrm{Cov}\left[Z_j^2, X_j^{t-1}\right] g_j^t = \mathrm{Cov}\left[X_j^t, X_j^{t-1}\right] \qquad (\Gamma1)$$

$$\mathrm{Cov}\left[Z_j^2, X_j^{t-1}\right] f_j^t + \mathrm{Var}\left[Z_j^2\right] g_j^t = \mathrm{Cov}\left[Z_j^2, X_j^t\right] \qquad (\Gamma2)$$

Αφαιρώντας τις μέσες τιμές από την (62), μετά πολλαπλασιάζοντας (πίσω) με $(\underline{X}^t - \underline{\xi}^t)^{\mathrm{T}}$, παίρνοντας αναμενόμενες τιμές, και επίσης παίρνοντας υπόψη ότι $\mathrm{Cov}[\underline{Q}^t, \underline{X}^t] = (\underline{h}^t)^{\mathrm{T}}$, καταλήγουμε στην εξίσωση

$$\underline{h}^t(\underline{h}^t)^{\mathrm{T}} = \mathrm{Cov}\left[\underline{X}^t,\ \underline{X}^t\right] - \underline{f}^t\ \mathrm{Cov}\left[\underline{X}^{t-1},\ \underline{X}^t\right] - \underline{g}^t\ \mathrm{Cov}\left[\underline{Z}^2,\ \underline{X}^t\right] \tag{Γ3}$$

Ο όρος $\mathrm{Cov}\left[\underline{X}^{t-1},\ \underline{X}^t\right]$ στην (Γ3) (παρών επίσης και στην (Γ1)) είναι ίσος με $\underline{\sigma}^{t-1}\underline{a}^t$, όπως εύκολα προκύπτει από την (43) που εξακολουθεί να ισχύει. Η παρουσία όρων όπως ο $\mathrm{Cov}[\underline{Z}^2,\ \underline{X}^t]$ στις (Γ1) έως (Γ3) δεν προσθέτει καμιά παράμετρο στο σύνολο που συζητήθηκε στο Κεφάλαιο 4. Πράγματι, το μητρώο συνδιασποράς της $\underline{Z}^2 = \underline{X}^{k+1} + \ldots + \underline{X}^{2k}$ με την $\underline{X}^t$ μπορεί να υπολογιστεί με επαναληπτική χρήση της (43) βάσει των $\underline{a}^l$, $l = 1, \ldots, k$, και $\underline{\sigma}^t$ (σημειώνεται ότι $\underline{a}^{l+k} = \underline{a}^l$), αφού μπορεί να δειχτεί ότι

$$\mathrm{Cov}\left[\underline{Z}^2,\ \underline{X}^t\right] = \left\{ \sum_{l=1}^{k} (\underline{a}^l \ldots \underline{a}^1) \right\} \underline{a}^k \ldots \underline{a}^{t+1}\ \underline{\sigma}^t \tag{Γ4}$$

ή, με τη χρήση των $\underline{\alpha}^t$ που ορίζονται με την (67)

$$\mathrm{Cov}\left[\underline{Z}^2,\ \underline{X}^t\right] = \underline{\alpha}^t\ \underline{\sigma}^t \tag{Γ5}$$

Κάνοντας συμβολικές απλοποιήσεις στις (Γ1), (Γ2), (Γ3) και χρησιμοποιώντας και τις (48), (68) και (Γ5), παίρνουμε τις (64) έως (66).

ΑΝΑΦΟΡΕΣ

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

Charbeneau, R., Comparison of the two- and three-parameter lognormal distributions used in streamflow synthesis, *Water Resour. Res.,* 14(1) 149-150, 1978.

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

Grygier, J.C. & 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.

Johnson, N.L. and Kotz, S., *Distributions in Statistics, Continuous multivariate distributions,* John Wiley, New York, 1972

Kendall, M.G. and Stuart, A., *The advanced theory of Statistics, Vol.1, Distribution theory,* 2nd edition, C. Griffin & Co., London, 1963.

Kottegoda, N.T., *Stochastic Water Resources technology,* Macmillan Press, London, 1980.

Κουτσογιάννης, Δ, *Μοντέλο επιμερισμού σημειακής βροχόπτωσης,* Διδακτορική διατριβή, ΕΜΠ, Αθήνα, 1988.

Κουτσογιάννης, Δ., Μαμάσης, Ν. και Ναλμπάντης, Ι., *Διερεύνηση προσφερομένων δυνατοτήτων για την ενίσχυση της ύδρευσης μείζονος περιοχής Αθηνών, Τεύχος 13, Στοχαστική προσομοίωση υδρολογικών μεταβλητών,* ΕΜΠ, Αθήνα 1990.

Koutsoyiannis, D. A nonlinear disaggregation method with a reduced parameter set for simulation of hydrologic series, *Wat. Resour. Res.*, in press.

Koutsoyiannis, D. and Xanthopoulos, Th., A dynamic model for short-scale rainfall disaggregation, *Hydrol. Sci. J.*, 35(3) 303-322, 1990.

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

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

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.

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*, A.K. Biswas editor, McGraw Hill, 1976.

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

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.

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 Modeling of Hydrologic Time Series*, Water Resources Publications, Littleton, Colo. 1980.

Stedinger, J.R., Fitting lognormal distributions to hydrologic data, *Water Resour. Res.* **16**(4) 481-490, 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.

ΠΑΡΑΡΤΗΜΑ 3

ΔΟΜΗ ΤΩΝ ΔΥΑΔΙΚΩΝ ΑΡΧΕΙΩΝ ΤΩΝ ΠΡΟΓΡΑΜΜΑΤΩΝ

Όπως αναφέρεται και στο κυρίως κείμενο αυτού του τεύχους, δύο από τους τύπους αρχείων που χρησιμοποιούνται στα προγράμματα προσομοίωσης χρονοσειρών είναι δυαδικά. Πρόκειται για τα αρχεία συνθετικών δεδομένων μεταβλητών υψηλότερου επιπέδου (τύπου ouh) και χαμηλότερου επιπέδου (τύπου oul). Τα αρχεία αυτά είναι δομημένα με τη μορφή δυαδικών εγγραφών (binary records). Η μία ή οι δύο πρώτες εγγραφές περιέχουν πληροφορίες σχετικές με τη δομή των χρονοσειρών, ενώ οι υπόλοιπες περιέχουν συνθετικά δεδομένα. Αναλυτικότερα η δομή των αρχείων έχει ως ακολούθως:

Αρχεία χρονοσειρών υψηλότερου επιπέδου

- Μήκος εγγραφής: PDlength = dim * 4 (bytes), όπου dim είναι ο αριθμός των θέσεων (διάσταση) στις οποίες γίνεται προσομοίωση.
- Δεσμευμένες εγγραφές: 2 (συνολικό μήκος 2 * PDlength)
- Δεδομένα πρώτης δεσμευμένης εγγραφής
    - Πεδίο 1, Μήκος: 2 bytes, Τύπος: integer, Καταχώρηση: NumStages (= μήκος κάθε χρονοσειράς)
    - Πεδίο 2, Μήκος: 2 bytes, Τύπος: integer, Καταχώρηση: NumSeries (= αριθμός χρονοσειρών)
    - Πεδίο 3, Μήκος: (PDlength - 4) bytes, Τύπος: -, Καταχώρηση: - (αχρησιμοποίητο)
- Δεδομένα δεύτερης δεσμευμένης εγγραφής
    - Πεδίο 1, Μήκος: 1 byte, Τύπος: byte, Καταχώρηση: dim (= αριθμός θέσεων προσομοίωσης)
    - Πεδίο 2, Μήκος: (PDlength - 1) bytes, Τύπος: -, Καταχώρηση: - (αχρησιμοποίητο)
- Δεδομένα καθεμιάς από τις υπόλοιπες εγγραφές:
    - Μήκος: PDlength, Τύπος: array[1..dim] of real (single), Καταχώρηση: οι μεταβλητές υψηλότερου επιπέδου για τις dim θέσεις

Αρχεία χρονοσειρών υψηλότερου επιπέδου

- Μήκος εγγραφής: PSDlength = dim * (steps + 1) * 4 (bytes), όπου dim είναι ο αριθμός των θέσεων (διάσταση) στις οποίες γίνεται προσομοίωση και steps είναι ο αριθμός των υποδιαιρέσεων μιας περιόδου (= αριθμός μεταβλητών χαμηλότερου επιπέδου για κάθε μεταβλητή υψηλότερου επιπέδου).
- Δεσμευμένες εγγραφές: 1 (συνολικό μήκος PSDlength)

- Δεδομένα δεσμευμένης εγγραφής
  - Πεδίο 1, Μήκος: 2 bytes, Τύπος: integer, Καταχώρηση: NumStages (= μήκος κάθε χρονοσειράς)
  - Πεδίο 2, Μήκος: 2 bytes, Τύπος: integer, Καταχώρηση: NumSeries (= αριθμός χρονοσειρών)
  - Πεδίο 3, Μήκος: 1 byte, Τύπος: byte, Καταχώρηση: dim (= αριθμός θέσεων προσομοίωσης)
  - Πεδίο 4, Μήκος: 1 byte, Τύπος: byte, Καταχώρηση: steps (= αριθμός υποδιαιρέσεων μιας περιόδου.
  - Πεδίο 5, Μήκος: (PSDlength - 6) bytes, Τύπος: -, Καταχώρηση: - (αχρησιμοποίητο)
- Δεδομένα καθεμιάς από τις υπόλοιπες εγγραφές:
  - Μήκος: PSDlength, Τύπος: array[1..(steps +1)*dim] of real (single), Καταχώρηση: οι μεταβλητές υψηλότερου και χαμηλότερου επιπέδου για τις dim θέσεις (πρώτα καταχωρείται η μεταβλητή υψηλότερου επιπέδου και στη συνέχεια οι steps μεταβλητές χαμηλότερου επιπέδου για μια θέση· ακολουθούν οι μεταβλητές των επόμενων θέσεων με τον ίδιο τρόπο καταχώρησης).