

Εφαρμογή ενός Γενικευμένου Γραμμικού Μοντέλου (ΓΓΜ) για την Προσομοίωση της Βροχόπτωσης στην Περιοχή του Νέστου

Η. Ιορδανίδης, Π. Αναγνωστόπουλος

Τμήμα Πολιτικών Μηχανικών ΑΠΘ, Εργαστήριο Υδραυλικής, 54124 Θεσσαλονίκη

Περίληψη

Είναι πολύ συχνό φαινόμενο κατά την εκπόνηση μελετών ή τη διεξαγωγή έρευνας στην Ελλάδα να υπάρχει δυσκολία στην απόκτηση δεδομένων ή, όταν αυτά είναι διαθέσιμα, να μην καλύπτουν μεγάλο χρονικό εύρος, με αποτέλεσμα να υπάρχει δυσκολία κατά την εκτέλεση προσομοιώσεων. Σε μια προσπάθεια να καλυφθεί αυτή η έλλειψη, παρουσιάζεται η εφαρμογή ενός Γενικευμένου Γραμμικού Μοντέλου (ΓΓΜ) προσομοίωσης βροχόπτωσης. Η εργασία αυτή καταδεικνύει τις δυνατότητες αυτού του μοντέλου στην προσομοίωση βροχόπτωσης, ακόμη και σε περιπτώσεις όπου τα αρχικώς διαθέσιμα δεδομένα είναι περιορισμένα. Επίσης παρουσιάζονται οι δυνατότητές του στην συμπλήρωση ελλειπών δεδομένων. Η εφαρμογή έγινε για την περιοχή του ελληνικού τμήματος του ποταμού Νέστου.

Application of a Generalized Linear Model (GLM) for the Simulation of Rainfall in the Greek Part of the Nestos River Basin

I. Iordanidis, P. Anagnostopoulos

Dept. of Civil Engineering AUTH, Hydraulics Laboratory 54124 Thessaloniki

Abstract

It is very common in Greece to encounter difficulties in obtaining data necessary for research or design projects, or, when datasets are available, to cover a narrow time span. This often results in difficulties when conducting simulations of any kind. In an effort to compensate for this situation, we present the use of a Generalized Linear Model (GLM) in order to produce rainfall datasets. This paper displays the potential of this particular model in simulating rainfall, even in cases where the initial datasets are limited. It also demonstrates its ability to fill incomplete datasets. The method has been applied to the part of the Nestos River Basin which belongs to Greece.

1. Εισαγωγή

Το πρόβλημα της έλλειψης δεδομένων για τη διεξαγωγή έρευνας στην Ελλάδα αποτελεί πολύ συχνό πρόβλημα, είτε επειδή τα αναγκαία δεδομένα δεν υπάρχουν είτε επειδή οι αρμόδιοι φορείς αρνούνται να τα διαθέσουν. Άλλες φορές, τα δεδομένα έχουν μικρό χρονικό εύρος με αποτέλεσμα να δυσχεραίνεται η εξαγωγή ασφαλών συμπερασμάτων.

Η παρούσα εργασία έχει ως σκοπό να παρουσιάσει την εφαρμογή ενός Γενικευμένου Γραμμικού Μοντέλου (Generalized Linear Model, GLM) μέσω H/Y για την προσομοίωση της βροχόπτωσης σε μία περιοχή, ή για την συμπλήρωση τυχόντων κενών σε υπάρχουσες χρονοσειρές βροχόπτωσης.

Το λογισμικό που χρησιμοποιήθηκε είναι το GlimClim, όπως αναπτύχθηκε και διατίθεται από τον Dr. Richard Chandler σε συνεργασία με το Τμήμα Πολιτικών Μηχανικών του Imperial College London. Πρόκειται για ελεύθερο λογισμικό, το οποίο ενσωματώνει τη θεωρία των GLM's και επιτρέπει την εύκολη και γρήγορη δημιουργία στοχαστικών χρονοσειρών δεδομένων βροχόπτωσης. Είναι ένα ιδιαίτερα εύχρηστο και αξιόπιστο εργαλείο, με την αξιοπιστία βεβαίως να βρίσκεται σε άμεση συνάρτηση με τα αρχικώς διαθέσιμα δεδομένα.

2. Ανασκόπηση βιβλιογραφίας

Οι McCullagh et al. (1989) παραθέτουν μία εκτενή περιγραφή των Γενικευμένων Γραμμικών Μοντέλων. Ένα πολύ καλό παράδειγμα εφαρμογής δίνεται από τους Yang et al. (2005) όπου παρουσιάζεται αναλυτικά η εφαρμογή του μοντέλου. Στη συγκεκριμένη εργασία η περιοχή μελέτης ήταν προκαθορισμένη και με εκτενείς σειρές δεδομένων με αποτέλεσμα να προκύψουν πολύ ενθαρρυντικά αποτελέσματα. Εκτός όμως από την περιοχή αυτή, η χρήση του μοντέλου και σε άλλες περιπτώσεις (Segond, 2006) όπου υπήρχε έλλειψη δεδομένων παρήγαγε επίσης αποδεκτά αποτελέσματα.

Από την εργασία των Coe & Stern (1982) είχε διαφανεί η δυνατότητα χρήσης των GLM's στην Υδρολογία σε διάφορες εφαρμογές. Ουσιαστικά, πρόκειται για μοντέλα πιθανοτήτων και αυτό επιτρέπει τη χρήση τους σε διάφορες περιβαλλοντικές και υδρολογικές εφαρμογές. Οι Chandler & Wheeler (2002) πρότειναν τη χρήση των GLM's για την χωροχρονική προσομοίωση της βροχόπτωσης σε μια περιοχή. Οι Yang et al. (2005) επισημαίνουν ένα ακόμη σημαντικό πλεονέκτημα των GLM's, το οποίο είναι η ευκολία στο χειρισμό και στην επεξεργασία των αποτελεσμάτων καθώς και η μικρή απαίτηση σε υπολογιστική ισχύ.

3. Παρουσίαση του προγράμματος

Το GlimClim βασίζεται στην προσαρμογή δύο ενσωματωμένων μοντέλων στα υπάρχοντα δεδομένα. Το πρώτο μοντέλο είναι ένα μοντέλο Λογιστικής Κατανομής, το οποίο προβλέπει αν σε μία συγκεκριμένη μέρα θα εκδηλωθεί ή όχι βροχόπτωση (Μοντέλο Συμβάντων, Events Model). Η κατανομή που χρησιμοποιείται είναι η

$$\ln\left(\frac{p_i}{1-p_i}\right) = x_i^T \beta \quad (3.1)$$

όπου p_i είναι η πιθανότητα εκδήλωσης βροχόπτωσης σε μια συγκεκριμένη ημέρα, i βασισμένη σε ένα σύνολο (διάνυσμα) μεταβλητών x_i για κάποιο διάνυσμα παραγόντων β . Με x_i^T συμβολίζεται ο ανάστροφος πίνακας του διανύσματος x_i .

Για τον υπολογισμό του ύψους βροχής σε μια ημέρα με βροχόπτωση, i , χρησιμοποιείται ένα μοντέλο κατανομής Γάμα, με την υπόθεση ότι η τιμή του ύψους βροχόπτωσης για την ημέρα αυτή είναι βασισμένη σε ένα διάνυσμα μεταβλητών ξ_i ,

$$\ln(\mu_i) = \xi_i^T \gamma \quad (3.2)$$

όπου μ_i είναι η μέση τιμή της κατανομής Γάμα για τη συγκεκριμένη ημέρα,
 γ το διάνυσμα παραγόντων και
 ξ_i^T ο ανάστροφος πίνακας του διανύσματος ξ_i .

Ο χρήστης έχει τη δυνατότητα να επιλέξει την πολυπλοκότητα και το πλήθος των παραμέτρων που θα εισαγάγει στην προσομοίωσή του, καθώς και τις αλληλεπιδράσεις μεταξύ τους. Η συνήθης πρακτική προσομοίωσης, η οποία δίνει πολύ καλό έλεγχο στο χρήστη κατά τη διάρκεια της διαδικασίας, είναι να ξεκινάει η προσαρμογή (fitting) των μοντέλων εισάγοντας λίγες παραμέτρους και αυξάνοντας σταδιακά τον αριθμό τους καθώς και τις αλληλεπιδράσεις μεταξύ τους. Μετά από κάθε προσαρμογή, το πρόγραμμα υπολογίζει τον παράγοντα log-likelihood, ο οποίος χρησιμοποιείται ως μέτρο βελτίωσης της προσαρμογής του μοντέλου στα δεδομένα. Η διαδικασία σταματάει όταν μεταξύ δύο διαδοχικών αυξήσεων των παραγόντων της προσομοίωσης δεν επέρχεται βελτίωση (αύξηση) του log-likelihood.

Ένα χρήσιμο στοιχείο που παρέχει το συγκεκριμένο μοντέλο είναι η δυνατότητα εξέτασης της αλληλεπίδρασης των διαφόρων παραμέτρων μεταξύ τους. Ο χρήστης μπορεί να επιλέξει να εξετάσει κατά πόσο η βροχόπτωση προηγούμενων ημερών επηρεάζει την πιθανότητα βροχόπτωσης μία δεδομένη μέρα, αν υπάρχει εποχιακή ή ομοιόμορφη κατανομή των βροχοπτώσεων μέσα στο έτος κ.λπ. Αν μετά την εισαγω-

γή των παραμέτρων αλληλεπίδρασης βελτιώνεται η απόδοση του μοντέλου τότε συμπεραίνεται ότι η αλληλεπίδραση αυτή είναι σημαντική για την κατανομή των βροχοπτώσεων στην περιοχή.

Όσον αφορά στη χωρική κατανομή των βροχοπτώσεων, το GlimClim προσφέρει την ευελιξία να μην απαιτεί συγκεκριμένη μορφή συντεταγμένων για τους σταθμούς για τους οποίους γίνεται η μελέτη. Αρκεί οι συντεταγμένες να είναι εκφρασμένες στο ίδιο σύστημα, είτε επίσημο είτε ακόμη και ως σχετικές θέσεις σε έναν απλό χάρτη. Η δυνατότητα αυτή απαλλάσσει το χρήστη από το άγχος της αναζήτησης συντεταγμένων σε ορισμένη μορφή και με τη χρήση κατάλληλων μετασχηματισμών να περιγράψει με ακρίβεια την τοπογραφία της περιοχής μελέτης.

Η προσομοίωση της βροχόπτωσης πραγματοποιείται με την χρήση των βέλτιστων παραμέτρων που προέκυψαν από την προσαρμογή του Λογιστικού και Γάμμα μοντέλου. Στη συνέχεια ο χρήστης επιλέγει τη χρονική περίοδο της προσομοίωσης, τον αριθμό των προσομοιώσεων και τη μορφή των αποτελεσμάτων. Στην περίπτωση που το μοντέλο χρησιμοποιηθεί για τη συμπλήρωση υπάρχοντων χρονοσειρών, θα επιστρέφει πάντα την μετρημένη τιμή, όταν αυτή υπάρχει (σε συγκεκριμένη μέρα), εφόσον η μετρημένη τιμή είναι εξ ορισμού η καλύτερη που μπορεί να βρεθεί για τη μέρα εκείνη.

Εκτός από τις παραμέτρους και τις αλληλεπιδράσεις τους που υπολογίζονται εντός του μοντέλου δίνεται και η δυνατότητα εισαγωγής εξωτερικών μεταβλητών που μπορεί να επηρεάζουν το σύστημα όπως η μεταβολή της θερμοκρασίας λόγω αλλαγής του κλίματος, η αλλαγή της στάθμης της θάλασσας σε κάποιες περιοχές κ.λπ. Η χρήση αυτής της επιλογής πρέπει να γίνεται με προσοχή και μετά από την πλήρη κατανόηση του μοντέλου καθώς ενδέχεται να οδηγήσει σε λανθασμένα συμπεράσματα. Αποτελεί όμως μια πολύ σημαντική προσθήκη η οποία διευρύνει σημαντικά τις δυνατότητες της προσομοίωσης.

Μειονέκτημα του προγράμματος (σε σχέση με την ελληνική πραγματικότητα κι όχι με την ίδια την απόδοσή του) είναι ότι απαιτεί τη χρήση ημερήσιων δεδομένων ως δεδομένων εισόδου. Αν αυτά δεν υπάρχουν, έστω και σε μικρό χρονικό εύρος, τότε η προσομοίωση είναι πρακτικώς αδύνατη, καθώς η χρήση του μέσου όρου της μηνιαίας βροχόπτωσης (μια συχνά εφαρμοζόμενη πρακτική) θα οδηγήσει σε τελείως εσφαλμένα αποτελέσματα.

4. Δεδομένα και περιοχή μελέτης

Ως περιοχή εφαρμογής επιλέχθηκε η περιοχή της λεκάνης του Νέστου και πιο συγκεκριμένα το ελληνικό τμήμα της. Ο Νέστος πηγάζει στην οροσειρά Momina Kula της Βουλγαρίας, εισέρχεται στην Ελλάδα στο Ν. Δράμας (οικισμός Δέλτα), συνεχίζει στο

Ν. Ξάνθης και εκβάλλει στο Θρακικό Πέλαγος. Έχει συνολικό μήκος 250 km, από τα οποία τα 150 διατρέχουν ελληνικό έδαφος. Κατά μήκος του ρου του στο ελληνικό τμήμα υπάρχουν 2 Υ/Η φράγματα στις θέσεις Θησαυρός και Πλατανόβρυση, ενώ είναι πιθανή η κατασκευή και του φράγματος Τεμένους. Το Δέλτα του ποταμού Νέστου αποτελεί προστατευόμενο βιότοπο από την UNESCO. Ο ποταμός, εκτός από την παραγωγή ενέργειας, χρησιμοποιείται για την εξυπηρέτηση αρδευτικών αναγκών, για οικότουρισμό και φυσικά συντηρεί το οικοσύστημα στο Δέλτα του. Κύριες χρήσεις του ποταμού παραμένουν η παραγωγή ενέργειας και σε δεύτερο βαθμό η άρδευση. Σε μελέτες των τοπικών Διευθύνσεων Εγγείων Βελτιώσεων (ΔΕΒ), και συγκεκριμένα της ΔΕΒ Δράμας (2002), διατυπώνεται το αίτημα να εκτραπεί ένα τμήμα της ροής του ποταμού προς την πεδιάδα της Δράμας, με σκοπό να καλύψει τις αρδευτικές ανάγκες της περιοχής. Το αίτημα αυτό δεν έχει γίνει δεκτό, είναι ενδεικτικό όμως της μεγάλης σημασίας που αποδίδουν οι τοπικές κοινωνίες στον ποταμό. Αυτή τη στιγμή, τα νερά του ποταμού χρησιμοποιούνται για την άρδευση των πεδιάδων Καβάλας και Ξάνθης.

Στον ενεργειακό τομέα, τα υδροηλεκτρικά φράγματα του ποταμού χρησιμοποιούνται για την κάλυψη απαιτήσεων αιχμής, κυρίως κατά τους καλοκαιρινούς μήνες.

Εξαιτίας της μεγάλης σημασίας του σε πολλαπλά επίπεδα (ενεργειακό, οικονομικό, περιβαλλοντικό) ο ποταμός αποτελεί αντικείμενο εκτεταμένης μελέτης και έρευνας. Ένα από τα προβλήματα που αντιμετωπίζουν όσοι ασχολούνται με θέματα υδρολογίας του ποταμού είναι η περιορισμένη διαθεσιμότητα δεδομένων σχετικών με την κατανομή των βροχοπτώσεων στην περιοχή. Η παρούσα εργασία αποσκοπεί στην παρουσίαση και την επίδειξη εφαρμογής ενός επιπλέον εργαλείου το οποίο θα βοηθήσει τους ερευνητές στην προσομοίωση της βροχόπτωσης στην περιοχή.

Για την εφαρμογή που παρουσιάζεται χρησιμοποιήσαμε δεδομένα από 10 βροχομετρικούς σταθμούς στην ευρύτερη περιοχή του Νέστου. Οι σταθμοί αυτοί βρίσκονται κοντά στους οικισμούς Αχλαδιά, Ποταμοί, Πρασινάδα, Μεσοχώρι, Μικροκλεισούρα, Βάλακας, Σιδηρόνερο, Παρανέστι, Τοξότες και Σκαλωτή, όπως φαίνεται στο Σχήμα 1. Το χρονικό εύρος που καλύπτουν τα δεδομένα των σταθμών αντιστοιχεί στην περίοδο Ιανουάριος 1983-Αύγουστος 1997, χωρίς όμως η περίοδος αυτή να καλύπτεται από τα δεδομένα όλων των σταθμών. Η συνολική μηνιαία βροχόπτωση για το σταθμό Ποταμών απεικονίζεται στο Σχήμα 2.

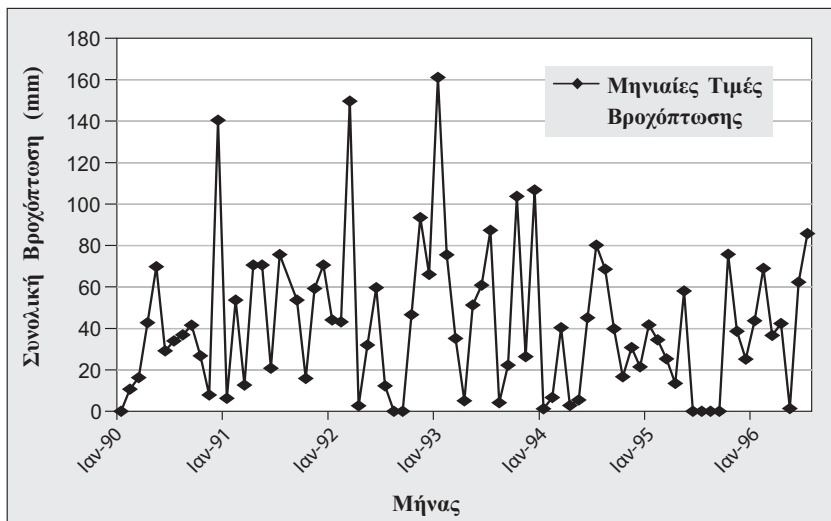
Διαθέσιμα δεδομένα υπήρχαν και για το βροχομετρικό σταθμό της Πτελέας, όμως μετά από έλεγχο διαπιστώθηκε ότι ο σταθμός είχε υποστεί βλάβη για μια μεγάλη περίοδο με αποτέλεσμα οι μετρήσεις να θεωρηθούν αναξιόπιστες. Για το λόγο αυτό, ο συγκεκριμένος σταθμός αποκλείστηκε εντελώς από τη συγκεκριμένη εργασία.

Ένα ακόμη μειονέκτημα των δεδομένων είναι ότι η περίοδος που καλύπτουν χαρακτηρίστηκε σε μεγάλο ποσοστό της από ξηρασία, με αποτέλεσμα οι μετρήσεις να μην αποτελούν πλήρως αντιπροσωπευτικό δείγμα των βροχοπτώσεων στην περιοχή.

Δυστυχώς όμως, αυτά είναι και τα μοναδικά διαθέσιμα επομένως η αβεβαιότητα που εισάγουν έπρεπε αναγκαστικά να γίνει αποδεκτή.



Σχήμα 1. Χάρτης της περιοχής μελέτης του Νέστου ποταμού.

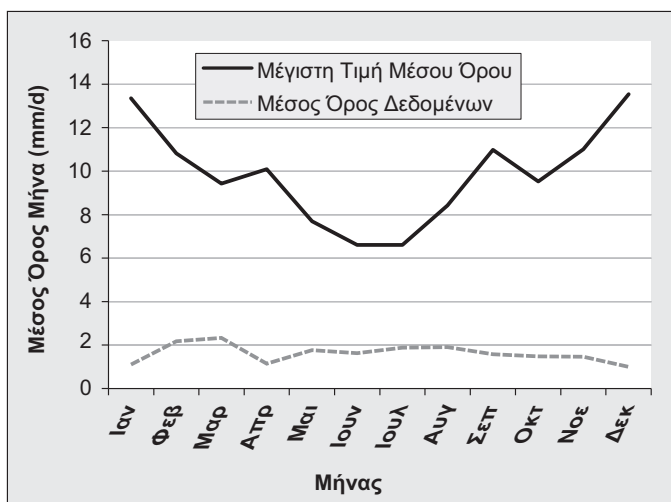


Σχήμα 2. Τιμές συνολικής μηνιαίας βροχόπτωσης για τον σταθμό στον οικισμό Ποταμοί.

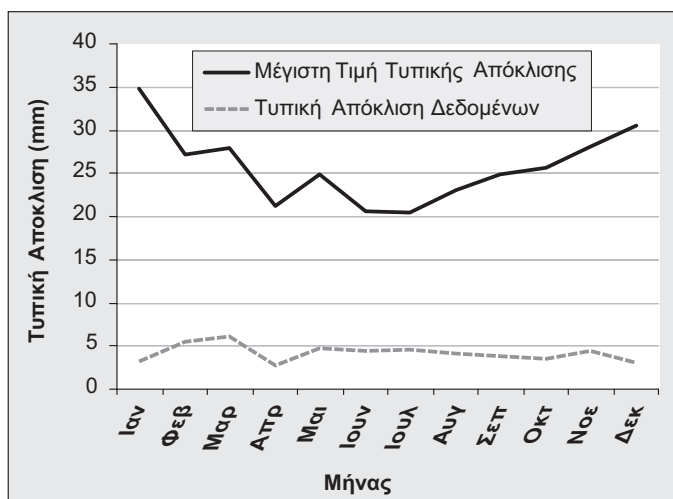
5. Προσομοίωση

Η τελική προσομοίωση καλύπτει την περίοδο 01/2000 μέχρι 12/2099. Ως αρχεία εισόδου χρησιμοποιήθηκαν τα δεδομένα από όλους τους σταθμούς, για την περίοδο που κάλυπτε ο καθένας. Το μοντέλο έχει τη δυνατότητα να αποδέχεται δεδομένα ανά διαφορετικό σταθμό τα οποία δεν καλύπτουν την ίδια περίοδο, κάνοντας έτσι ευκολότερη την εισαγωγή περισσότερων δεδομένων. Φυσικά, αν αυτά έχουν μεγάλη χρονική διασπορά, τότε η αβεβαιότητα της προσομοίωσης αυξάνεται.

Στο πρώτο στάδιο έγινε η προσαρμογή του Λογιστικού και Γάμα μοντέλου στα δεδομένα. Στη συνέχεια πραγματοποιήθηκε η προσομοίωση για την περίοδο 2000 έως 2099. Πραγματοποιήθηκε μεγάλος αριθμός επαναλήψεων ώστε να είναι εφικτός ο στατιστικός έλεγχος καθώς και η εξέταση της απόδοσης του μοντέλου. Για να γίνει αυτό, υπολογίστηκαν ορισμένα στατιστικά μεγέθη ώστε να δημιουργηθεί ένα πεδίο μέγιστης και ελάχιστης τιμής για καθένα από αυτά. Τα μεγέθη που επιλέχθηκαν είναι η μέση τιμή και η τυπική απόκλιση. Υπολογίστηκε η τιμή κάθε στατιστικού μεγέθους για κάθε μήνα για καθεμία προσομοίωση χωριστά. Στη συνέχεια επιλέχθηκε η μέγιστη και ελάχιστη τιμή για κάθε μέγεθος. Αν η προσομοίωση είναι σωστή, τότε οι τιμές των αντίστοιχων στατιστικών μεγεθών για την περίοδο που έχουμε διαθέσιμα δεδομένα θα πρέπει να βρίσκονται μεταξύ της μέγιστης και ελάχιστης τιμής (Σχήματα 3 και 4).



Σχήμα 3. Εύρος Μέσου Όρου Βροχόπτωσης για τον σταθμό στον οικισμό Αχλαδιά (Η ελάχιστη τιμή είναι μηδέν).



Σχήμα 4. Εύρος Τυπικής Απόκλισης Βροχόπτωσης για τον σταθμό στον οικισμό Αχλαδιά (Η ελάχιστη τιμή είναι μηδέν).

Παρατηρούμε ότι η απόδοση του μοντέλου είναι ιδιαίτερα ικανοποιητική. Όσον αφορά τις μέσες τιμές, παρατηρούμε ότι οι τιμές των δεδομένων βρίσκονται σχετικά χαμηλά, κάτι που επιβεβαιώνει την παρατήρησή μας όσον αφορά την ξηρασία της περιόδου που καλύπτουν τα δεδομένα. Αυτό είναι επίσης μία καλή ένδειξη όσον αφορά την απόδοση του μοντέλου στην περιοχή των ακραίων τιμών.

6. Συμπεράσματα

Η εργασία αυτή αποσκοπεί στο να παρουσιάσει για πρώτη φορά τη χρήση του συγκεκριμένου μοντέλου (GlimClim) στον ελληνικό χώρο ώστε να χρησιμοποιηθεί για να καλύψει το πρόβλημα της έλλειψης δεδομένων. Το μοντέλο μπορεί να χρησιμοποιηθεί για να καλύψει κενά σε υπάρχοντα δεδομένα ή να δημιουργήσει στοχαστικές χρονοσειρές για μελλοντικές περιόδους.

Από τα αποτελέσματα της εργασίας φαίνεται ότι το μοντέλο ανταποκρίνεται αρκετά ικανοποιητικά στο πρόβλημα. Σε προηγούμενες εργασίες (Yang et al., 2005) το μοντέλο είχε χρησιμοποιηθεί σε περιοχές όπου υπήρχαν διαθέσιμα δεδομένα για μεγάλες χρονικές περιόδους και ήταν σχεδιασμένες ακριβώς για την εκτέλεση σχετικών πειραμάτων.

Από την άλλη πλευρά, παρατηρούμε ότι το μοντέλο έχει ικανοποιητική απόδοση και στην περίπτωση μιας πραγματικής λεκάνης, με ελλιπή δεδομένα, όπως στην περί-

πρωσή μας. Όπως είναι αυτονόητο, όσο μεγαλύτερο χρονικό εύρος καλύπτουν τα δεδομένα τόσο καλύτερη και με μικρότερη αβεβαιότητα θα είναι η εκάστοτε προσομοίωση. Παρ' όλα αυτά όμως, ακόμη και με ελλιπή δεδομένα, η προσομοίωση μπορεί να δώσει πολύ καλά αποτελέσματα ειδικά αν χρησιμοποιηθούν σε αναλύσεις τύπου Monte Carlo που απαιτούν μεγάλο πλήθος προσομοιώσεων.

Καταλήγοντας, θα μπορούσαμε να σημειώσουμε ότι το συγκεκριμένο μοντέλο, αν εφαρμοστεί σωστά και στα πλαίσια των περιορισμών που υπάρχουν, μπορεί να αντεπεξέλθει στο πρόβλημα των ελλιπών δεδομένων. Αυτό όμως δεν πρέπει να μειώσει την απαίτηση των ερευνητών για διάθεση περισσότερων δεδομένων με σκοπό την εξαγωγή ασφαλέστερων συμπερασμάτων.

Βιβλιογραφία

1. Chandler, R. and Wheater, H.S., 2002. *Analysis of Rainfall Variability Using GLMs: A case study from the west of Ireland*. Water Resources Research, 38: 1192.
2. Chandler, R., 2002. *GLIMCLIM: Generalized Linear Modelling for daily climate time series (software and user guide)*. Tech Rep. 227, Dept. of Stat. Sci., UCL, London.
3. Coe, R. and Stern, R.D., 1982. *Fitting Models to daily rainfall*. Journal of Applied Meteorology, 21: 1024-1031.
4. McCullagh, P. and Nelder, J., 1989. *Generalised Linear Models*. CRC Press, Boca Raton, Florida.
5. Segond M.L., 2006. *Stochastic Modelling of Space-Time Rainfall and the Significance of Spatial Data for Flood Runoff Generation*. PhD Thesis, Imperial College, London.
6. Yang, C., Chandler, R.E. Isham, V.S. and Wheater H.S., 2005. *Spatial-Temporal rainfall simulation using generalized linear models*. Water Resources Research, 41: W11415.
7. Κισσούδης, Δ., 2002. *Διαχείριση-αξιοποίηση υδατικών πόρων στη γεωργία του Ν. Δράμας*. Μελέτη της ΔΕΒ Δράμας.