



**ΤΜΗΜΑ ΠΟΛΙΤΙΚΩΝ ΜΗΧΑΝΙΚΩΝ**

**ΔΠΜΣ Εφαρμοσμένες Πολιτικές και Τεχνικές Προστασίας του  
Περιβάλλοντος**

**ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ**

**Προβλέψεις ποιοτικών και ποσοτικών παραμέτρων νερού καρστικής  
υφάλμυρης πηγής με τη χρήση νευρωνικών δικτύων. Εφαρμογή στην  
πηγή του Αλμυρού (Ηράκλειο, Κρήτη).**

**Δρ ΓΕΩΡΓΙΟΣ Κ. ΜΠΕΚΑΣ**

**ΠΟΛΙΤΙΚΟΣ ΜΗΧΑΝΙΚΟΣ ΒEng., MEng.**

**ΕΠΙΒΛΕΠΩΝ: Δρ ΔΗΜΗΤΡΙΟΣ Ε. ΑΛΕΞΑΚΗΣ, ΑΝΑΠΛΗΡΩΤΗΣ  
ΚΑΘΗΓΗΤΗΣ**

**ΑΘΗΝΑ, 2020**



Το παρόν κείμενο αποτελεί εκτεταμένη περίληψη του κειμένου της μεταπτυχιακής διπλωματικής εργασίας. Το πλήρες κείμενο έχει κατατεθεί στην βιβλιοθήκη του Ιδρύματος.

## ΕΙΣΑΓΩΓΗ

Σκοπός της παρούσας μεταπτυχιακής διπλωματικής εργασίας είναι η κατανόηση του φυσικού συστήματος της καρστικής πηγής του Αλμυρού στον Νομό Ηρακλείου Κρήτης μέσω της δημιουργίας ενός νευρωνικού δικτύου το οποίο θα δομηθεί με όρους χρονοσειράς δεδομένων.

Τα νευρωνικά δίκτυα έχουν -μεταξύ άλλων χρήσεων- ως στόχο να δημιουργήσουν μια συνάρτηση, η οποία θα δέχεται ως μεταβλητές εισόδου, (ή ανεξάρτητες μεταβλητές) οι οποίες θα μπορούν να προβλέπουν μια μεταβλητή εξόδου (εξαρτημένη μεταβλητή).

Στην περίπτωση των περιβαλλοντικών δεδομένων με τη μορφή χρονοσειράς, η λογική που επικρατεί είναι ότι η όποια πρόβλεψη του χρονικού δεδομένου του επόμενου βήματος βασίζεται σε μια αλληλουχία από προηγούμενους χρονικούς κύκλους, με συγκεκριμένο αριθμό βημάτων. Τα προηγούμενα χρονικά βήματα (ένα ή περισσότερα), ονομάζονται καθυστερήσεις και χρησιμοποιούνται άρα ως μεταβλητές πρόβλεψης.

Η μεταπτυχιακή διπλωματική εργασία θα αξιοποιήσει δεδομένα περιορισμένης έκτασης και θα δημιουργηθεί μεταμοντέλο με υψηλό βαθμό στατιστικής ανεξαρτησίας, μέσω του οποίου θα δίνεται η δυνατότητα για πρόβλεψη μελλοντικών χρονικών βημάτων παροχών. Μέσω του μαθηματικού μοντέλου αυτού, θα επιχειρηθεί οι ήδη υπάρχουσες παρατηρήσεις να δύνανται να προβλεφθούν με υψηλό βαθμό ακρίβειας.

Το πρώτο κεφάλαιο της εργασίας εστιάζει στην περιγραφή του Αλμυρού ως φυσικού συστήματος. Γίνεται αναφορά σε προηγούμενες μελέτες με υδρολογικά, γεωλογικά, κλιματολογικά, ανθρωπογεωγραφικά και λοιπά περιβαλλοντικά δεδομένα.

Το δεύτερο κεφάλαιο της εργασίας περιγράφει σχετικές στατιστικές τεχνικές υψηλού ενδιαφέροντος και σχετιζόμενες με την παρούσα εργασία. Το τρίτο κεφάλαιο της εργασίας αναφέρεται στη διαδικασία και στις λεπτομέρειες κατασκευής του μεταμοντέλου. Για την κατασκευή του μεταμοντέλου, χρησιμοποιήθηκε η γλώσσα προγραμματισμού MATLAB. Η κατασκευή αφορά

ένα μη-γραμμικό νευρωνικό αυτοπαλινδρομούμενο δίκτυο (NAR). Η αρχιτεκτονική του, τα χαρακτηριστικά του, καθώς και δείκτες ποιότητας ως προς το βαθμό της αξιοπιστίας του, συζητούνται στο κεφάλαιο αυτό. Τέλος, στο κεφάλαιο 4 γίνεται αναφορά στα συμπεράσματα της μελέτης και σε προτάσεις για μελλοντική έρευνα.

## **ABSTRACT**

The purpose of the present master's thesis is to understand the natural system of the karstic source of Almyros in the Prefecture of Heraklion, Crete through the creation of a neural network which will be structured in the form of time series data.

Neural networks aim -among other uses- to create a function, which will make use of a series of input variables (or independent variables) that can predict an output variable (dependent variable).

The master's thesis will utilize rare data and will create a metamodel with a high degree of statistical independence, through which it will be possible to predict future time steps of water discharges. Through this mathematical model, it will be attempted to simulate existing observations with a high degree of accuracy.

The first chapter of the thesis focuses on the description of Almyros as a natural system. Reference is made to previous studies with hydrological, geological, climatological, anthropogeographical and other environmental data. The second chapter of the paper describes relevant statistical techniques that are of high interest and are related to the present study. The third chapter of the work refers to the process and the details of the construction of the metamodel. The MATLAB programming language was used to construct the metamodel. The construction involves a non-linear self-regulating neural network (NAR). Its architecture, its features, as well as quality indicators in terms of its degree of reliability, are discussed in this

chapter. Finally, chapter 4 refers to the conclusions of the study and makes suggestions for future research.

## **ΚΑΤΑΣΚΕΥΗ ΜΕΤΑΜΟΝΤΕΛΟΥ, ΑΞΙΟΛΟΓΗΣΗ ΤΗΣ ΠΟΙΟΤΗΤΑΣ ΤΟΥ ΤΕΛΙΚΩΝ ΕΠΙΛΕΧΘΕΝΤΟΣ ΝΕΥΡΩΝΙΚΟΥ ΔΙΚΤΥΟΥ ΚΑΙ ΔΟΚΙΜΕΣ ΠΟΥ ΕΓΙΝΑΝ**

Στην παρούσα έρευνα, ένα νευρωνικό δίκτυο εφαρμόζεται για σκοπούς πρόβλεψης σε μελλοντικές τιμές των συνολικών απορρίψεων του νερού στην καρστική πηγή του Αλμυρού (Ηράκλειο Κρήτη) με βάση προηγούμενες παρατηρήσεις, που εκφράζονται σε μηνιαία βάση (Σταθόπουλος, 2010).

Η εργαλειοθήκη νευρωνικών δικτύων του MATLAB απαιτεί τον προσδιορισμό των ποσοστών στα οποία, οι μετρούμενες παρατηρήσεις χωρίζονται σε ένα υποσέτ εκπαίδευσης (training set), ένα υποσέτ επικύρωσης (validation set) και σε ένα υποσέτ ελέγχου (test set). Αυτό συμβάλλει στην στατιστική ανεξαρτησία ενός μεταμοντέλου. Διάφορες τιμές ελέγχθηκαν για τα ποσοστά των υποσέτ ελέγχου, επικύρωσης και εκπαίδευσης. Το μέσο απόλυτο σφάλμα (MAE) και η συνολική τιμή του συντελεστή προσδιορισμού (R), χρησιμοποιήθηκαν, ως δείκτες για την αξιολόγηση της ποιότητας των κατασκευασμένων νευρωνικών δικτύων (Hagan, 2019).

Η αρχιτεκτονική των νευρωνικών δικτύων για την κατασκευή του μεταμοντέλου προσδιορίστηκε σύμφωνα με τις μεθόδους που περιγράφονται σε έγκυρη επιστημονική δημοσίευση που αποτελεί έγγραφο ανασκόπησης (Sheela Gnana et al., 2013). Οι κρυμμένοι νευρώνες αποφασίστηκαν να έχουν μια τιμή μεταξύ του μεγέθους στρώματος εισόδου και του μεγέθους του στρώματος εξόδου (1 έξοδος). Αυτό γίνεται για να αποφευχθεί η τυχόν υπερεκπαίδευση του νευρωνικού δικτύου (Sheela Gnana et al., 2013).

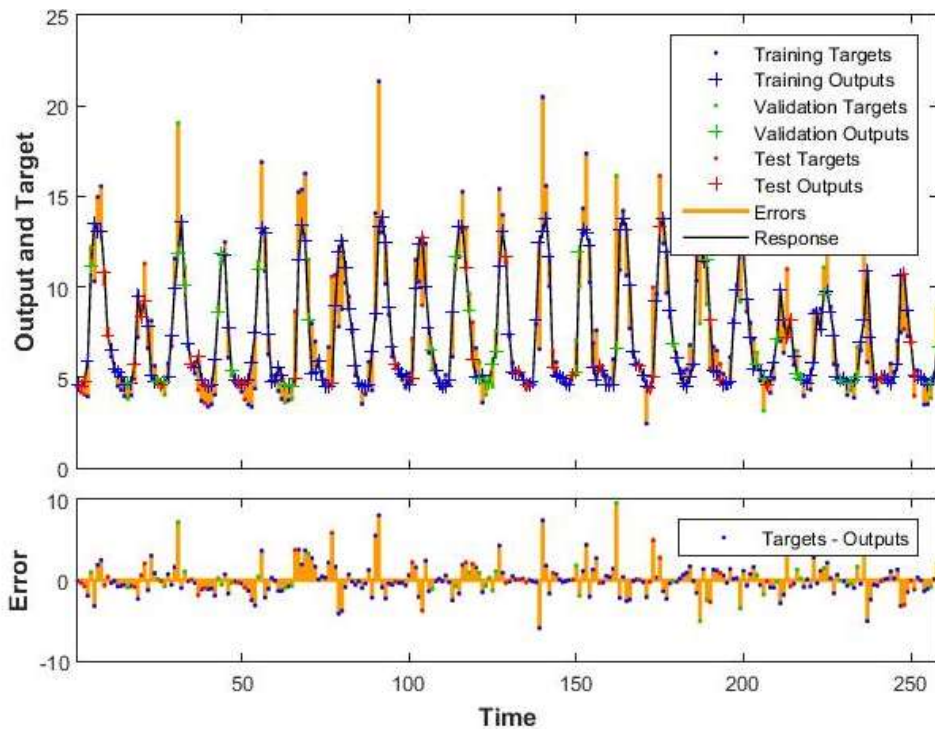
Αυτό συνεπάγεται ότι στην περίπτωση μας, ο προτεινόμενος αριθμός κρυφών νευρώνων, θα πρέπει να έχει μια τιμή μεταξύ του 1 και του συνολικού αριθμού

καθυστερήσεων, αφού το δίκτυο θα χρησιμοποιηθεί για να προβλέψει μία μόνο μεταβλητή.

Ο όρος «καθυστερήσεις» (Yazeed et al., 2017), αναφέρεται στον συνολικό αριθμό διαδοχικών μηνιαίων διαστημάτων, που χρησιμοποιείται ως μέσο πρόβλεψης του μήνα που είναι ένα διάστημα μετά από την προαναφερθείσα ακολουθία. Η διαδικασία επαναλαμβάνεται για κάθε μήνα, που περιέχεται στις συγκεντρωμένες παρατηρήσεις και αυτό καθιστά δυνατή την κατασκευή μεταμοντέλων, που ταιριάζουν καλά στα δεδομένα του δείγματος και συνοψίζει τις ιδιότητές του καθώς το προκύπτον μοντέλο προβλέπει κάθε μήνα. Μετά την κατασκευή ενός μεταμοντέλου, είναι δυνατό αυτό να χρησιμοποιηθεί για σκοπούς πρόβλεψης που μπορούν να εκτιμήσουν τις τιμές των διαστημάτων σε πολλά βήματα μπροστά από τις τρέχουσες παρατηρήσεις, με ικανοποιητική ακρίβεια.

Το νευρωνικό δίκτυο που τελικά επελέχθη ύστερα από 10 δοκιμές (με σταδιακές βελτιώσεις στην υπερπαραμέτρους του (ο αριθμός των κρυφών νευρώνων, ο αριθμός των καθυστερήσεων, τα ποσοστά των υποσυνόλων) και με χρήση όλων των 258 παρατηρήσεων) (Σταθόπουλος, 2010), έχει τις ακόλουθες ιδιότητες: Ο αλγόριθμος επίλυσης που χρησιμοποιήθηκε ήταν ο Levenberg-Marquardt (Hagan et al., 2019), οι κρυφοί νευρώνες ήταν ίσοι με 1, ο επιλεγμένος αριθμός των καθυστερήσεων ήταν ίσος με 25, η αναλογία του υποσύνολου εκπαίδευσης του δικτύου ήταν ίση με 70%, η αναλογία του υποσέτ επικύρωσης δικτύου ήταν ίση με 15%, και η αναλογία υποσέτ ελέγχου δικτύου ήταν ίση με το 15% (Mathworks, 2019).

Τα αποτελέσματα των επιδόσεων για το νευρωνικό δίκτυο που θεωρήθηκε ότι έχει την καλύτερη αρχιτεκτονική, είχαν ως εξής: Συντελεστής προσδιορισμού ( $R$ ) = 0.82949, μέσο απόλυτο σφάλμα (MAE) = 1.34. Η χρονοσειρά με τις εξόδους του προαναφερθέντος νευρωνικού δικτύου NAR, και τα σφάλματα ανταπόκρισης σε κάθε βήμα, δείχνονται παρακάτω (Εικ. 1):

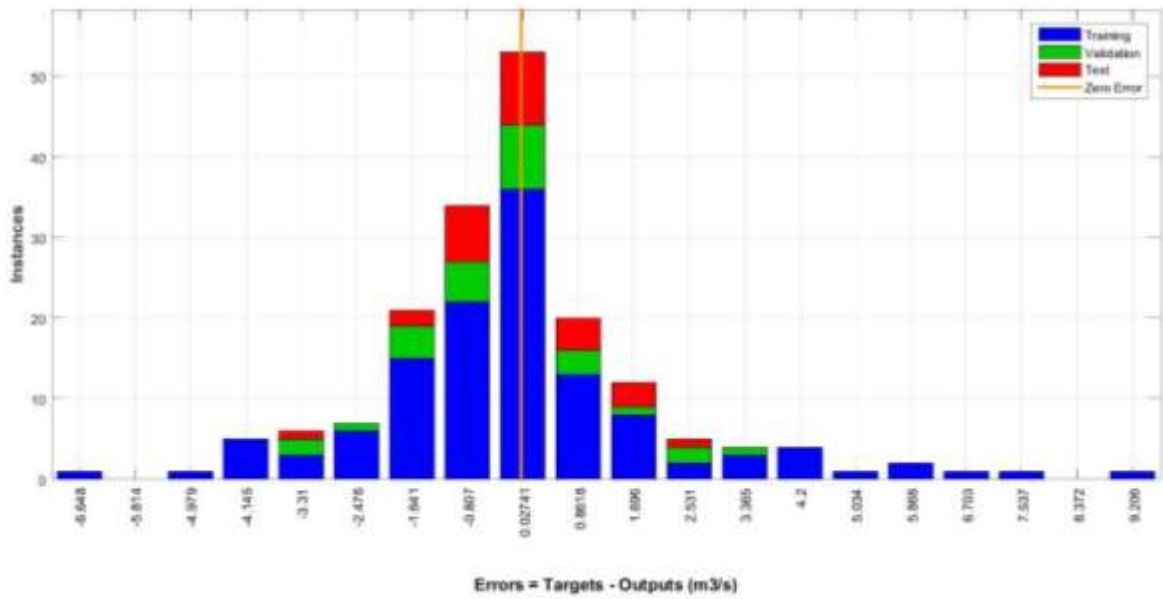


Εικόνα 1: Απόκριση νευρωνικού δικτύου NAR και σφάλματα για τα 258 εξεταζόμενα μηνιαία χρονικά διαστήματα.

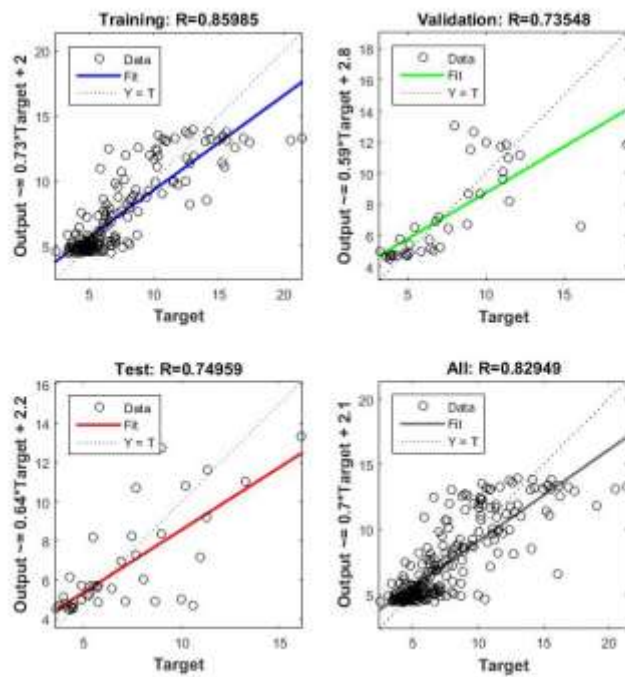
Έχει νόημα να σημειωθεί ότι ένα (φαινομενικά) αρκετά καλής απόδοσης νευρωνικό δίκτυο, με μια αρχική αρχιτεκτονική που βελτιώθηκε με δοκιμή και σφάλμα (Yazeed et al., 2017), χρησιμοποιήθηκε για να προβλέψει τις τιμές που έλειπαν, μέσω προεκβολής. Η προεκβολή έγινε μόνο για μερικά βήματα μπροστά και για ένα σχετικά μικρό αριθμό μηνιαίων βημάτων, που κρατήθηκε χαμηλότερος από τον αριθμό των καθυστερήσεων. Συνολικά 32 τιμές που έλειπαν, προβλέφθηκαν μέσω αυτής της διαδικασίας.

Το ιστόγραμμα των λαθών και οι συντελεστές προσδιορισμού, για το βέλτιστο νευρωνικό δίκτυο, καταδεικνύονται κατωτέρω (Εικ. 2-3).





Εικόνα 2: Ιστόγραμμα σφαλμάτων για τα 3 υποσύνολα ομάδων παρατηρήσεων.



Εικόνα 3: Συντελεστής προσδιορισμού για τα 3 υποσύνολα παρατηρήσεων και συνολικός συντελεστής προσδιορισμού.

## ΣΥΜΠΕΡΑΣΜΑΤΑ

Μια σύγχρονη και εύρωστη τεχνική για τη μελλοντική πρόβλεψη απορρίψεων υδάτινων όγκων, στην πηγή του Αλμυρού Κρήτης, συζητήθηκε και αναπτύχθηκε στην παρούσα μεταπτυχιακή διπλωματική εργασία. Λόγω της μικρής χρονοσειράς των διαθέσιμων δεδομένων για την πηγή του Αλμυρού Ηρακλείου Κρήτης και της πολυπλοκότητάς της, η επίτευξη ενός συνολικού συντελεστή προσδιορισμού ίσου με 0,83 θεωρείται ικανοποιητική. Εκτός από αυτό, με παρόμοιο τρόπο το μέσο απόλυτο σφάλμα μειώθηκε με διαδοχικές δοκιμές και ανάλυση συσχέτισης, ενώ άλλοι ποιοτικοί δείκτες όπως είναι ο χωρισμός των παρατηρήσεων στις οποίες βασίστηκε το νευρωνικό δίκτυο σε κατάλληλες αναλογίες υποσύνολων, ο αριθμός των κρυμμένων νευρώνων, τα μικρά σφάλματα αυτοσυσχέτισης, καθώς και άλλες τεχνικές για να βελτιωθεί η αρχιτεκτονική του δικτύου, τέθηκαν σε εφαρμογή για να εξασφαλιστεί ένα βελτιωμένο αποτέλεσμα.

Ο αλγόριθμος επίλυσης Levenberg-Marquardt, προτιμήθηκε έναντι άλλων γιατί κρίθηκε ότι επιτυγχάνει γρηγορότερη σύγκλιση, χωρίς όμως να γίνεται κάποιος συμβιβασμός στην ποιότητα του τελικού μεταμοντέλου.

Τα προβλεπόμενα μέσα σφάλματα είναι συγκριτικά πολύ χαμηλότερα, κατά τη διάρκεια της θερινής περιόδου. Τα αποτελέσματα υποδηλώνουν επίσης ότι οι όποιες μελλοντικές προσπάθειες, που θα συμπεριλάβουν έναν μεγαλύτερο αριθμό μεταβλητών (που εγγενώς λείπουν και επομένως εξηγούν τα υψηλότερα σφάλματα που παρατηρήθηκαν στις δυνατότητες πρόβλεψης του μοντέλου μας, ιδιαίτερα κατά τη χειμερινή περίοδο) είναι πολύ πιθανό να επιτύχουν υψηλότερο βαθμό ακρίβειας. Ως εκ τούτου, μια λογική πρόταση για το μελλοντική έρευνα, είναι η δημιουργία ενός NARX νευρωνικού δικτύου, το οποίο θα χρησιμοποιήσει ως μεταβλητές πρόβλεψης περισσότερες από μία μεταβλητές εκφραζόμενες (π.χ. βροχόπτωση, θερμοκρασία, δεδομένα εκμετάλλευσης) με τη μορφή χρονοσειράς.

## ΔΗΛΩΣΗ ΣΥΓΓΡΑΦΕΑ ΔΙΠΛΩΜΑΤΙΚΗΣ ΕΡΓΑΣΙΑΣ

Ο κάτωθι υπογεγραμμένος ΓΕΩΡΓΙΟΣ ΜΠΕΚΑΣ του ΚΩΝΣΤΑΝΤΙΝΟΥ φοιτητής του ΠΜΣ Εφαρμοσμένες Πολιτικές και Τεχνικές Προστασίας Περιβάλλοντος

του Α.Ε.Ι Πειραιά Τ.Τ, πριν αναλάβω την εκπόνηση της Διπλωματικής Εργασίας μου, δηλώνω ότι ενημερώθηκα για τα παρακάτω:


«Η Διπλωματική Εργασία (Δ.Ε) αποτελεί προϊόν πνευματικής ιδιοκτησίας τόσο του συγγραφέα, όσο και του Ιδρύματος και θα πρέπει να έχει μοναδικό χαρακτήρα και πρωτότυπο περιεχόμενο.

Απαγορεύεται αυστηρά οποιοδήποτε κομμάτι κειμένου της να εμφανίζεται αυτούσιο ή μεταφρασμένο από κάποια άλλη δημοσιευμένη πηγή. Κάθε τέτοια πράξη αποτελεί προϊόν λογοκλοπής και εγείρει θέμα Ηθικής Τάξης για τα πνευματικά δικαιώματα του άλλου συγγραφέα. Αποκλειστικός υπεύθυνος είναι ο συγγραφέας της Π.Ε, ο οποίος φέρει και την ευθύνη των συνεπειών, ποινικών και άλλων, αυτής της πράξης.

Πέραν των όποιων ποινικών ευθυνών του συγγραφέα, σε περίπτωση που το Ίδρυμα του έχει απονείμει Πτυχίο, αυτό ανακαλείται με απόφαση της Συνέλευσης του Τμήματος. Η Συνέλευση του Τμήματος με νέα απόφασή της, μετά από αίτηση του ενδιαφερόμενου, του αναθέτει εκ νέου την εκπόνηση Π.Ε με άλλο θέμα και διαφορετικό επιβλέποντα καθηγητή. Η εκπόνηση της εν λόγω Π.Ε πρέπει να ολοκληρώσει εντός τουλάχιστον ενός ημερολογιακού 6μήνου από την ημερομηνία ανάθεσής της. Κατά τα λοιπά εφαρμόζονται τα προβλεπόμενα στο άρθρο 18. παρ.5 του ισχύοντος Εσωτερικού Κανονισμού».

Ο Δηλών

Ημερομηνία



03/08/2020