



**ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΕΡΙΒΑΛΛΟΝΤΟΣ**

**ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ ΕΙΔΙΚΕΥΣΗΣ
“ΠΕΡΙΒΑΛΛΟΝΤΙΚΗ ΚΑΙ ΥΓΕΙΟΝΟΜΙΚΗ ΜΗΧΑΝΙΚΗ”**

ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΑΤΡΙΒΗ

**ΠΡΟΣΟΜΟΙΩΣΗ ΠΑΡΟΧΩΝ ΚΑΡΣΤΙΚΩΝ ΠΗΓΩΝ ΜΕ
ΧΡΗΣΗ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ**

ΕΙΡΗΝΗ ΣΚΙΤΖΗ

Διπλ. Πολιτικός Μηχανικός

Εξεταστική Επιτροπή:

**Παλαιολόγος Ευάγγελος (Επιβλέπων)
Κατσιφαράκης Κωνσταντίνος
Καρατζάς Γεώργιος**

Χανιά, 2010

Περίληψη

Στόχος της συγκεκριμένης μεταπτυχιακής εργασίας είναι η προσομοίωση των παροχών καρστικών πηγών με τη βοήθεια των Τεχνητών Νευρωνικών Δικτύων. Περιοχή μελέτης της εφαρμογής αποτέλεσε ο Δήμος Ρούβα, στο νομό Ηρακλείου Κρήτης, του οποίου το υπέδαφος χαρακτηρίζεται από καρστικότητα. Η εφαρμογή των αλγορίθμων πραγματοποιήθηκε βάσει υδρολογικών δεδομένων που ελήφθησαν από το δήμο Ρούβας για τα έτη 2007, 2008 και 2009 (έως 15/07/09), ενώ οι αλγόριθμοι που αναπτύχθηκαν, υλοποιήθηκαν σε γλώσσα προγραμματισμού Matlab.

Σε πρώτο στάδιο πραγματοποιήθηκε μία προεπεξεργασία των δεδομένων έτσι ώστε να είναι σε θέση να εισαχθούν στο νευρωνικό δίκτυο. Στα πλαίσια αυτά υπολογίσθηκε ο συντελεστής συσχέτισης της βροχόπτωσης και της απορροής για διαφορετικές χρονικές καθυστερήσεις (time lags), από 1 έως 30 μέρες. Ο στόχος ήταν, να βρεθεί ο μέγιστος συντελεστής συσχέτισης, ο οποίος θα αντιστοιχεί στο χρονικό διάστημα που απαιτείται προκειμένου να φτάσει το ύψος βροχής στην πηγή.

Κατόπιν, επιλέχθηκε νευρωνικό δίκτυο με τρία επίπεδα (ένα στρώμα εισόδου, ένα ενδιάμεσο στρώμα και ένα στρώμα εξόδου). Πρόκειται για δίκτυο προώθησης με ανάδραση (feed-forward neural network with Backpropagation). Δοκιμάστηκαν διάφορες συναρτήσεις δημιουργίας δικτύου, συναρτήσεις εκπαίδευσης και συναρτήσεις εκμάθησης, για διαφορετικό αριθμό νευρώνων του ενδιάμεσου στρώματος, μέχρι να εντοπιστεί το βέλτιστο αποτέλεσμα. Κριτήριο επιλογής ήταν το μέσο τετραγωνικό σφάλμα ανάμεσα στις αναμενόμενες και παραχθείσες από το δίκτυο τιμές. Τέλος, μελετήθηκαν εφαρμογές χρησιμοποιώντας διαφορετικά διανύσματα εισόδου και εξόδου κάθε φορά, με έμφαση την χρήση της διαφοράς της παροχής ως απόκριση του δικτύου, αντί για την ίδια την παροχή. Όλοι οι αλγόριθμοι υλοποιήθηκαν και για τις δύο πηγές Μάη και Πέρα Βρύση του Δήμου Ρούβα, για τις οποίες δόθηκαν τα απαιτούμενα στοιχεία.

Τέλος συνοψίζονται τα συμπεράσματα στα οποία κατέληξε η μεταπτυχιακή αυτή εργασία. Σημαντικό είναι ότι λόγω της χωρικής ποικιλομορφίας των γεωλογικών και υδραυλικών χαρακτηριστικών των καρστικών υδροφορέων, είναι αδύνατον να προσομοιωθούν με κοινά αποδεκτά μοντέλα. Λόγω της απλότητας της εφαρμογής και της φιλικότητας προς τον χρήστη, ένα σωστά εκπαιδευμένο νευρωνικό δίκτυο, είναι σε θέση να αποδώσει χρήσιμα συμπεράσματα στον τομέα της διαχείρισης των υδατικών πόρων και να προσφέρει κίνητρα περαιτέρω έρευνας και βαθύτερης κατανόησης.

Στην εργασία που ακολουθεί αναλύονται θεωρητικά όλα τα εργαλεία που χρησιμοποιήθηκαν καθώς και οι βασικές έννοιες της θεωρίας των Τεχνητών Νευρωνικών Δικτύων, προτού παρουσιαστεί η πρακτική εφαρμογή τους.

.

Πρόλογος

Η παρούσα διπλωματική εργασία εκπονήθηκε στα πλαίσια του Μεταπτυχιακού Προγράμματος Σπουδών «Περιβαλλοντική και Υγειονομική Μηχανική» του Τμήματος Μηχανικών Περιβάλλοντος του Πολυτεχνείου Κρήτης.

Αντικείμενο της μεταπτυχιακής διατριβής είναι η προσομοίωση παροχών καρστικών πηγών μέσω της χρήσης των Τεχνητών Νευρωνικών Δικτύων, ενώ η εφαρμογή των αλγορίθμων πραγματοποιήθηκε σε περιβάλλον Matlab, πάνω σε υδρολογικά δεδομένα του Δήμου Ρούβα, στον νομό Ηρακλείου Κρήτης.

Στο σημείο αυτό θα ήθελα να ευχαριστήσω θερμά τον επιβλέποντα καθηγητή κ. Ε. Παλαιολόγο για τις πολύτιμες γνώσεις του και την αμέριστη βοήθεια του καθ' όλη τη διάρκεια της έρευνας μου. Παράλληλα, τον καθηγητή κ. Κ. Κατσιφαράκη, του Αριστοτελείου Πανεπιστημίου Θεσσαλονίκης, για την σημαντική του καθοδήγηση και τις συμβουλές του σε αυτήν την προσπάθεια. Επίσης θα ήθελα να ευχαριστήσω τον καθηγητή κ. Γ. Καρατζά, μέλος της εξεταστικής επιτροπής της μεταπτυχιακής μου διατριβής. Τέλος, πρέπει να ευχαριστήσω τον υποψήφιο διδάκτορα κ. Ι. Τριχάκη για τη βοήθεια του κατά τη διάρκεια της εκπόνησης της συγκεκριμένης εργασίας.

Τέλος, θέλω να αφιερώσω την συγκεκριμένη εργασία στην οικογένεια μου, που με στηρίζει σε κάθε βήμα μου.

Χανιά, Μάρτιος 2010

Ειρήνη Σκιτζί

Περιεχόμενα

Περίληψη	i
Πρόλογος.....	iii
Κεφάλαιο 1.....	1
Εισαγωγή.....	1
1.1 Γενικά.....	1
1.2 Δομή της εργασίας	2
Κεφάλαιο 2.....	4
Διάκριση Υπόγειων Υδροφορέων	4
2.1 Εισαγωγή	4
2.2 Κοκκώδεις υδροφορείς	4
2.3 Ρηγματωμένοι υδροφορείς	6
2.4 Καρστικοί υδροφορείς	8
2.4.1 Καρστικός κύκλος	10
2.4.2 Χαρακτηριστικές καρστικές μορφές.....	11
2.4.3 Εμφάνιση καρστικών στον Ελλαδικό χώρο.	11
Κεφάλαιο 3.....	13
Περιοχή Μελέτης	13
3.1 Γεωγραφικά στοιχεία	13
3.2 Γεωλογική και υδρολογική δομή του Δήμου Ρούβα	16
3.3 Επιφανειακά νερά.....	17
3.3.1 Πηγές Δήμου Ρούβα (Γέργερης – Νυβρίτου- Πανασού)	18
3.3.2 Υδρολογική λεκάνη Δήμου Ρούβα.....	19
3.3.2.1 Λεκάνη τροφοδοσίας πηγών Γέργερης-Ζαρού.....	19
3.3.2.2 Υδρολογική Λεκάνη Λιθαίου	19
Κεφάλαιο 4.....	20

Νευρωνικά Δίκτυα	20
4.1 Ορισμός	20
4.2 Η έμπνευση από τη βιολογία	23
4.3 Από τα Βιολογικά Νευρωνικά Δίκτυα στα Τεχνητά.....	23
4.3.1 Ο Βιολογικός Νευρώνας	24
4.4 Αναλυτική περιγραφή ενός απλού τεχνητού νευρώνα.....	26
4.4.1 Το μοντέλο του τεχνητού νευρώνα	26
4.4.2 Συναρτήσεις ενεργοποίησης νευρώνα	29
4.4.3 Τοπολογία - Αρχιτεκτονικές των Τεχνητών Νευρωνικών Δικτύων	32
4.4.4 Μέθοδοι Εκπαίδευσης	33
4.4.5 Κανόνες Εκμάθησης	34
4.5 Ταξινόμηση Τεχνικών Νευρωνικών Δικτύων.....	37
4.6 Το μοντέλο του αισθητήρα – (Perceptron)	37
4.7 Το δίκτυο ADALINE (Adaptive Linear Element)	40
4.8 Δίκτυο Multi-Layer Perceptron - Κανόνας Backpropagation.....	41
4.8.1 Παραγωγή του αλγορίθμου Πίσω Διάδοσης (Back Propagation).....	42
4.9 Τεχνητά Νευρωνικά Δίκτυα Συνάρτησης Ακτινικής Βάσης (Radial-basis Function Neural Networks)	48
4.10 Μαθηματικό υπόβαθρο τεχνητών νευρωνικών δικτύων	51
4.10.1 Το μοντέλο McCulloch – Pitts.....	51
4.10.2 Θεωρία Διανυσμάτων	53
4.10.2.1 Μήκος διανύσματος.....	54
4.10.3 Ο ρόλος του εσωτερικού γινομένου.....	54
4.10.4 Εσωτερικό Γινόμενο και Τεχνητοί Νευρώνες	56
4.10.5 Μαθηματικό υπόβαθρο στην εκπαίδευση του νευρώνα.....	59
4.10.5.1 Προσαρμογή του διανύσματος των βαρών	60
4.11 Πλεονεκτήματα των Νευρωνικών Δικτύων.....	61
4.12 Εφαρμογές των Νευρωνικών Δικτύων	62
Κεφάλαιο 5.....	64
Ανάπτυξη Αλγορίθμων Νευρωνικών Δικτύων σε Περιβάλλον Matlab	64
5.1 Γενικά.....	64
5.2 Matlab και Τεχνητά Νευρωνικά Δίκτυα	66

5.2.1 Υποστηριζόμενες αρχιτεκτονικές Τεχνητών Νευρωνικών Δικτύων από το toolbox της Matlab	66
5.2.2 Η αρχιτεκτονική των Δικτύων Προώθησης με Ανάδραση (Backpropagation Feed-forward Networks) σε περιβάλλον Matlab.....	67
5.2.2.1 Συνάρτηση δημιουργίας δικτύου	68
5.2.2.2 Συνάρτηση αρχικοποίησης των συναπτικών βαρών και των πολώσεων του δικτύου	69
5.2.2.3 Συνάρτηση εκπαίδευσης και συνάρτηση εκμάθησης	69
5.2.2.4 Συνάρτηση εξομοίωσης.....	71
5.2.2.5 Το πρόβλημα του “Overfitting”	71
Κεφάλαιο 6.....	73
Βιβλιογραφική Ανασκόπηση	73
6.1 Εισαγωγή	73
6.2 Εφαρμογή Τεχνητών Νευρωνικών Δικτύων στην εκτίμηση της σχέσης βροχοπτώσεων - απορροής	74
6.3 Εφαρμογή Τεχνητών Νευρωνικών Δικτύων στα υπόγεια ύδατα	77
6.4 Εφαρμογή Τεχνητών Νευρωνικών Δικτύων σε καρστικούς υδροφορείς	81
Κεφάλαιο 7.....	83
Ανάπτυξη Αλγορίθμων με Χρήση Νευρωνικών Δικτύων – Ανάλυση Αποτελεσμάτων	83
7.1 Επεξεργασία Δεδομένων	83
7.2 Περιγραφή και γενικές αρχές του νευρωνικού δικτύου	84
7.3 Μεθοδολογία - Επιλογή βέλτιστου συνδυασμού τύπου εκπαίδευσης και πλήθους νευρώνων κρυφού στρώματος.....	85
7.4 Αναλυτική παρουσίαση βέλτιστων αποτελεσμάτων	112
7.4.1 Αποτελέσματα εκτέλεσης νευρωνικού δικτύου για την πηγή Μάη Βρύσης.	112
7.4.2 Αποτελέσματα εκτέλεσης νευρωνικού δικτύου για την πηγή Πέρα Βρύσης.	113
7.5 Έλεγχος επίδρασης γραμμικής παρεμβολής - Αναλυτική παρουσίαση αποτελεσμάτων	114
7.5.1 Αποτελέσματα εκτέλεσης νευρωνικού δικτύου, χωρίς την εφαρμογή γραμμικής παρεμβολής στα δεδομένα πηγής Μάη Βρύσης.	116
7.5.2 Αποτελέσματα εκτέλεσης νευρωνικού δικτύου, χωρίς την εφαρμογή γραμμικής παρεμβολής στα δεδομένα πηγής Πέρα Βρύσης.....	117
7.6 Βελτιωμένο νευρωνικό δίκτυο - Εισαγωγή στοιχείων θερμοκρασίας - Αναλυτική παρουσίαση αποτελεσμάτων	118

7.6.1 Αποτελέσματα βελτιωμένου νευρωνικού δικτύου για την πηγή Μάη Βρύσης.	119
7.6.2 Αποτελέσματα βελτιωμένου νευρωνικού δικτύου για την πηγή Πέρα Βρύσης.	120
7.7 Εισαγωγή των βροχοπτώσεων των ημερών όπου μεγιστοποιείται ο συντελεστής συσχέτισης ως δεδομένων εισόδου - Αναλυτική παρουσίαση αποτελεσμάτων	121
7.8 Δημιουργία νευρωνικού δικτύου με παραμέτρους εξόδου τη διαφορά της παροχής ανά χρονικό βήμα ημέρας - Αναλυτική παρουσίαση αποτελεσμάτων.....	124
7.8.1 Αποτελέσματα νευρωνικού δικτύου με παράμετρο εξόδου τη διαφορά παροχής για την πηγή Μάη Βρύσης.	128
7.8.2 Αποτελέσματα νευρωνικού δικτύου με παράμετρο εξόδου τη διαφορά παροχής για την πηγή Πέρα Βρύσης.	130
Κεφάλαιο 8.....	133
Συμπεράσματα και Συζήτηση.....	133
Βιβλιογραφία.....	137

Κεφάλαιο 1

Εισαγωγή

1.1 Γενικά

Αποτελεί γεγονός ότι τις τελευταίες δεκαετίες μαζί με την εμφάνιση των μαζικά παράλληλων ηλεκτρονικών υπολογιστών, έχει προκύψει ένα αυξανόμενο ενδιαφέρον σε αλγόριθμους που στηρίζουν τη λειτουργία τους στις αρχές φυσικών διαδικασιών. Πρόκειται για αλγόριθμους Υπολογιστικής Νοημοσύνης (Computational Intelligence), που μέσω της Τεχνητής Νοημοσύνης (Artificial Intelligence) και της Εκπαίδευσης των Μηχανών (Machine Learning) έχουν συμβάλει ουσιαστικά στην επίλυση πρακτικών προβλημάτων. Χαρακτηριστικά παραδείγματα αποτελούν τα νευρωνικά δίκτυα, οι γενετικοί αλγόριθμοι, ο εξελικτικός προγραμματισμός, οι εξελικτικές στρατηγικές, οι μέθοδοι τοπικής αναζήτησης, και τα συστήματα ταξινόμησης.

Τεχνητή Νοημοσύνη (T.N.) ονομάζεται η μελέτη των τεχνικών και των διεργασιών που δίνουν σε έναν υπολογιστή τη δυνατότητα να αποκτά διανοητικές ικανότητες, ανάλογες με αυτές που διαθέτει ο ανθρώπινος εγκέφαλος, στοχεύοντας να ανταποκρίνεται πολύ καλύτερα στις ανάγκες και τις επιθυμίες του ανθρώπου (M.T.Hagan et al, 1996).

Τα Τεχνητά Νευρωνικά Δίκτυα (ΤΝΔ) αποτελούν μια προσπάθεια προσέγγισης της λειτουργίας του ανθρώπινου εγκεφάλου από μια μηχανή. Έχουν την ικανότητα να εκτελούν υπολογισμούς με μαζικό παράλληλο τρόπο και αρχιτεκτονική τους βασίζεται στην αρχιτεκτονική των Βιολογικών Νευρωνικών Δικτύων. Τα ΤΝΔ είναι μια συλλογή από νευρώνες (Processing Units-PU) που συνδέονται μεταξύ τους. Κάθε νευρώνας έχει πολλές εισόδους αλλά μόνο μία έξοδο η οποία με τη σειρά της μπορεί να αποτελέσει είσοδο για άλλους νευρώνες. Η συνδέσεις μεταξύ των νευρώνων διαφέρουν ως προς τη σημαντικότητά τους, η οποία και προσδιορίζεται από το συντελεστή βάρους (σύναψη), ενώ η επεξεργασία τους καθορίζεται από τη συνάρτηση μεταφοράς, που ορίζει την κάθε έξοδο σε σχέση με τις εισόδους και τους συντελεστές βάρους.

Τα Νευρωνικά δίκτυα, τα τελευταία χρόνια, έχουν εφαρμοστεί σε μεγάλο βαθμό σε θέματα Υδρολογίας και Υδραυλικής, καθώς είναι σε θέση να δώσουν αξιόπιστες λύσεις. Βασικός στόχος της παρούσας εργασίας είναι η προσομοίωση των παροχών καρστικών πηγών μέσω της χρήσης των νευρωνικών δικτύων. Προς την επίτευξη του συγκεκριμένου στόχου αρχικά ερευνήθηκε η βέλτιστη δομή του δικτύου καθώς και ο βέλτιστος τύπος εκπαίδευσης του. Κατόπιν, υλοποιήθηκαν διάφοροι

αλγόριθμοι, προσεγγίζοντας το θέμα μέσω διαφορετικών περιπτώσεων εφαρμογών, αναζητώντας την καλύτερη δυνατή απόκριση του νευρωνικού δικτύου.

1.2 Δομή της εργασίας

Η παρούσα εργασία συνίσταται στα παρακάτω οκτώ κεφάλαια:

Στο πρώτο εισαγωγικό κεφάλαιο, πραγματοποιείται αναφορά στην έννοια της Τεχνητής Νοημοσύνης και ειδικότερα στα Νευρωνικά Δίκτυα, τα οποία αποτελούν το «εργαλείο» της συγκεκριμένης εργασίας. Κατόπιν παρουσιάζεται εν συντομία το αντικείμενο και η διάρθρωση των κεφαλαίων της διπλωματικής εργασίας.

Στο δεύτερο κεφάλαιο αναφέρεται ο διαχωρισμός των εδαφών σε τρεις κατηγορίες: στα κοκκώδη, τα ρηγματωμένα και τέλος τα καρστικά εδάφη, τα οποία αποτελούν το κύριο αντικείμενο της μελέτης αυτής, όπως προαναφέρθηκε. Κατόπιν αναλύεται η καρστικότητα ως φαινόμενο και τέλος παρουσιάζονται ορισμένες χαρακτηριστικές καρστικές μορφές στον Ελλαδικό χώρο.

Στο τρίτο κεφάλαιο παρουσιάζεται η περιοχή μελέτης, καθώς επίσης αναφέρονται και ορισμένα βασικά γεωλογικά και υδρολογικά στοιχεία της ευρύτερης περιοχής.

Στο τέταρτο κεφάλαιο πραγματοποιείται μία εκτενής αναφορά του θεωρητικού και μαθηματικού υποβάθρου των Τεχνητών Νευρωνικών Δικτύων, καθώς και η έμπνευση τους από τη βιολογία. Παρουσιάζεται η ταξινόμηση των Τεχνητών Νευρωνικών Δικτύων σε κατηγορίες, ενώ δίνεται ιδιαίτερη έμφαση στα Τεχνητά Νευρωνικά Δίκτυα Ανάδρασης, καθώς στην παρούσα εργασία επιλέχτηκε ο συγκεκριμένος τύπος δικτύου. Στο τέλος του κεφαλαίου παραθέτονται τα πλεονεκτήματα, αλλά και οι εφαρμογές των νευρωνικών δικτύων.

Στο πέμπτο κεφάλαιο περιγράφεται με συντομία η γλώσσα προγραμματισμού Matlab, καθώς επίσης αναφέρονται επιγραμματικά τα βασικότερα εργαλεία για τη δημιουργία των Τεχνητών Νευρωνικών Δικτύων της Matlab. Τέλος, αναφέρονται οι βασικές συναρτήσεις δημιουργίας, εκπαίδευσης και εκμάθησης, οι οποίες είναι απαραίτητες για την υλοποίηση ενός νευρωνικού δικτύου.

Στο έκτο κεφάλαιο γίνεται μία εκτενής παρουσίαση της εφαρμογής των νευρωνικών δικτύων σε θέματα υδραυλικής και υδρολογίας, όπως αυτή συνίσταται στη διεθνή βιβλιογραφία. Αναφέρονται έρευνες που εφήρμοσαν τη φιλοσοφία των Τεχνητών Νευρωνικών Δικτύων προς εκτίμηση της σχέσης βροχόπτωσης – απορροής, στα υπόγεια ύδατα και σε καρστικούς υδροφορείς.

Στο έβδομο κεφάλαιο παρουσιάζεται αναλυτικά η επεξεργασία των δεδομένων, καθώς και η επιλογή της βέλτιστης δομής και τύπου εκπαίδευσης του νευρωνικού δικτύου. Στη συνέχεια παραθέτονται οι διάφορες περιπτώσεις εφαρμογών που αναπτύχθηκαν, αναλύοντας κάθε φορά τα αποτελέσματα απόκρισης του νευρωνικού δικτύου.

Στο όγδοο και τελευταίο κεφάλαιο, διατυπώνονται τα γενικότερα συμπεράσματα που προέκυψαν, καθώς επίσης συζητείται η αναγκαιότητα της χρήσης των Τεχνητών Νευρωνικών Δικτύων σε περιοχές με καρστικότητα.

Κεφάλαιο 2

Διάκριση Υπόγειων Υδροφορέων

2.1 Εισαγωγή

Οι υδροφορείς διακρίνονται ως προς τη δομή και τον τρόπο δημιουργίας των κενών χώρων σε κοκκώδεις, ρηγματωμένους και καρστικούς. Η περιοχή μελέτης της συγκεκριμένης μεταπτυχιακής εργασίας, χαρακτηρίζεται από έντονη καρστικότητα και για αυτό τον λόγο θα παρατεθεί εκτενής ανάλυση και περιγραφή των χαρακτηριστικών των καρστικών υδροφορέων γενικότερα.

2.2 Κοκκώδεις υδροφορείς

Οι κοκκώδεις υδροφορείς χαρακτηρίζονται από εδαφικό υλικό με κόκκους ακανόνιστου σχήματος. Το σημαντικό είναι το ότι κατά την απόθεση των κόκκων παραμένουν κενοί χώροι, που επιτρέπουν την αποθήκευση και την κίνηση του υπόγειου νερού (Κατσιφαράκης, 2006).

Η εξίσωση της κίνησης του υπόγειου νερού στους κοκκώδεις υδροφορείς περιγράφεται από το νόμο του Darcy, ο οποίος ισχύει για στρωτές ροές στα πορώδη μέσα και ορίζει μια γραμμική σχέση για την ταχύτητα ροής.

Χαρακτηριστικό είναι το ότι ο αριθμός Reynolds (Re) καθορίζει το είδος της ροής και δίνεται από την σχέση:

$$Re = Vd/\nu \quad (2.1)$$

, όπου

ν = κινηματικό ιξώδες του ρευστού (L^2/T),

d = αντιπροσωπευτική διάσταση των κόκκων του εδάφους,

V = ταχύτητα Darcy ή ειδική παροχή κατά τη διεύθυνση της ροής.

Για να είναι στρωτή η ροή, ο Re πρέπει να είναι μικρότερος του 1 (αν και δεχόμαστε την ισχύ του νόμου του Darcy και για τιμές του Re μικρότερες του 10).

Ο νόμος του Darcy δίνεται από την σχέση:

$$\mathbf{V} = -K\nabla (p/\rho g + z) = -K\nabla h \quad (2.2)$$

Για ένα ομογενές και ισότροπο μέσο η πειραματική σχέση του Darcy μπορεί να επεκταθεί και να γενικευθεί σε πεδίο τριών διαστάσεων με τις σχέσεις:

$$V_x = -K \frac{\partial h}{\partial x} \quad (2.3)$$

$$V_y = -K \frac{\partial h}{\partial y} \quad (2.4)$$

$$V_z = -K \frac{\partial h}{\partial z} \quad (2.5)$$

, όπου:

V_x, V_y, V_z = η ταχύτητα Darcy κατά τους άξονες x, y, z ,

K = υδραυλική αγωγιμότητα (L/T),

h = πιεζομετρικό φορτίο (L).

Όσον αφορά τον προσδιορισμό της κατανομής του πιεζομετρικού φορτίου, χρησιμοποιείται ο συνδυασμός του νόμου του Darcy με την εξίσωση συνέχειας. Θεωρώντας τα πηγάδια ως σημειακούς φορτιστικούς όρους, η εξίσωση συνέχειας περιγράφεται από την σχέση:

$$S \frac{\partial h}{\partial t} = \nabla(T \nabla h) + \sum_{i=1}^n Q_i \cdot \delta_i \quad (2.6)$$

, όπου:

S = αποθηκευτικότητα (αδιάστατο μέγεθος),

T = μεταφορικότητα, ($T=b \cdot K$, όπου b το πάχος του υδροφορέα και K η υδραυλική αγωγιμότητα), (L^2/T),

Q_i = παροχή του πηγαιδιού i (L^3/T),

δ_i = Dirac δέλτα (L^{-2}).

Όταν πρόκειται για μόνιμο φαινόμενο, η (2.6) παίρνει την ακόλουθη μορφή:

$$\nabla(T \nabla h) = - \sum_{i=1}^n Q_i \cdot \delta_i \quad (2.7)$$

Στην περίπτωση ομογενούς υδροφορέα, χωρίς πηγάδια, τότε η σχέση (2.7) καταλήγει στην εξίσωση Laplace:

$$\nabla^2 h = 0 \quad (2.8)$$

Επίσης εάν ο υδροφορέας είναι ομογενής κατά ζώνες, η (2.8) ισχύει στο εσωτερικό κάθε ζώνης, ενώ στην επιφάνεια διαχωρισμού ζωνών διαφορετικής μεταφορικότητας ισχύουν οι ακόλουθες σχέσεις:

$$h^{(k)} = h^{(l)} \quad (2.9)$$

$$V_n^{(k)} = V_n^{(l)} \quad (2.10)$$

, όπου

V_n = η ειδική παροχή κάθετου στην επιφάνεια διαχωρισμού των ζωνών k και l, (Κατσιφαράκης, 2006).

2.3 Ρηγματωμένοι υδροφορείς

Οι ρηγματωμένοι υδροφορείς χαρακτηρίζονται από την παρουσία ρηγμάτων, δηλαδή κενών χώρων με μήκος πολύ μεγαλύτερο από το πλάτος τους. Υπάρχουν δύο κατηγορίες μαθηματικών ομοιωμάτων όσον αφορά τη ροή στους ρηγματωμένους υδροφορείς (Κατσιφαράκης, 2006):

- i. η παραδοχή του συνεχούς μέσου
- ii. η εξέταση της ροής αυτόνομα σε κάθε ρήγμα

Η κίνηση του υπόγειου νερού, όπως και στην κατηγορία των κοκκωδών εδαφών, περιγράφεται από τον νόμο του Darcy. Παρόλα αυτά στην περίπτωση που αναμένεται η ανάπτυξη τυρβώδους ροής μέσα στα ρήγματα, είναι ακριβέστερο να χρησιμοποιείται η σχέση του Forchheimer:

$$gradh = cV + dV^2 \quad (2.11)$$

, όπου οι συντελεστές c, d έχουν διαστάσεις (T/L) και (T²/L²) αντιστοίχως.

Στην περίπτωση που ο υδροφορέας προσομοιαστεί με δυο επάλληλα συνεχή μέσα, το ένα αντιστοιχεί στη ροή στο εσωτερικό των πόρων και το άλλο αντιστοιχεί στη ροή μέσα στα ρήγματα. Σε κάθε ένα από τα δύο συνεχή μέσα θεωρείται ότι ισχύει ο νόμος του Darcy. Όμως, η τιμή του K είναι διαφορετική σε κάθε μέσο.

Η εξίσωση συνέχειας για κάθε μέσο δίνεται αντίστοιχα από τις σχέσεις:

$$\frac{\partial(\rho_p \phi_p)}{\partial t} + \nabla(\rho_p \phi_p V_p) + M = 0 \quad (2.12)$$

$$\frac{\partial(\rho_f \phi_f)}{\partial t} + \nabla(\rho_f \phi_f V_f) - M = 0 \quad (2.13)$$

, όπου

ϕ = πορώδες,

t = χρόνος,

M = διακίνηση μάζας ρευστού μεταξύ ρηγμάτων και πόρων,
 f και p = δείκτες για τη ροή στα ρήγματα και τους πόρους, αντιστοίχως.

Σημαντικό είναι βέβαια να τονιστεί το ότι η εφαρμογή των μαθηματικών ομοιωμάτων στην εξέταση της ροής χωριστά σε κάθε ρήγμα προϋποθέτει καλή γνώση της γεωμετρίας και της τραχύτητας των ρηγμάτων.

Στο εσωτερικό των ρηγμάτων η ροή μπορεί να θεωρηθεί μονοδιάστατη, γιατί το εύρος τους είναι πολύ μικρό. Επομένως η εξίσωση διατήρησης της μάζας μπορεί να ολοκληρωθεί κάθετα στον άξονα του ρήγματος. Για μόνιμη ροή, η ολοκλήρωση αυτή και η παραδοχή ότι η μέση ταχύτητα είναι ανάλογη της βαθμίδας του υδραυλικού φορτίου οδηγούν στην εξίσωση:

$$\frac{d}{dx'} \left(T_f \frac{dh}{dx'} \right) + Q_f^{(k)} + Q_f^{(l)} = 0 \quad (2.14)$$

όπου

x' = συντεταγμένη που συμπίπτει με τον άξονα του ρήγματος,

T_f = μεταφορικότητα του ρήγματος,

$Q_f^{(l)}$ = εισροές στο ρήγμα από τα τοιχώματά του, δηλαδή από τις ζώνες k και l .

Η σχέση (2.14) μπορεί να ολοκληρωθεί κατά μήκος του άξονα του ρήγματος. Αν μεταξύ δυο σημείων x'_i και x'_{i+1} η τιμή της T_f και των πλευρικών εισροών παραμένουν σταθερές, η ολοκλήρωση της σχέσης (3.14) δίνει:

$$\left(T_f \frac{dh}{dx'} \right) \Big|_i = \frac{T_{fi}(h_{i+1} - h_i)}{x'_{i+1} - x'_i} - \frac{1}{2} (Q_{fi}^{(k)} + Q_{fi}^{(l)})(x'_{i+1} - x'_i) \quad (2.15)$$

Στα σημεία τομής των ρηγμάτων η εξίσωση συνέχειας παίρνει την ακόλουθη μορφή:

$$\sum_{j=1}^N \left(T_{fj} \frac{dh}{dx_j} \right) \Big|_i + Q_i = 0 \quad (2.16)$$

όπου

i = σημείο τομής,

N = αριθμός των τεμνομένων ρηγμάτων,

Q_i = παροχή πηγαδιού που βρίσκεται στο σημείο i , (Κατσιφαράκης, 2006).

2.4 Καρστικοί υδροφορείς

Η τρίτη κατηγορία πετρωμάτων είναι τα λεγόμενα καρστικά, τα οποία κατέχουν σημαντικά ποσοστά της συνολικής έκτασης των εδαφών στον κόσμο και συναντώνται πολύ συχνά στη Μεσόγειο και ιδιαίτερα στην Ελλάδα.

Η καρστικότητα αποτελεί γεωμορφολογικό φαινόμενο που συμβαίνει κυρίως στις ασβεστολιθικές περιοχές. Τα ασβεστολιθικά πετρώματα αποτελούνται βασικά από ανθρακικό ασβέστιο, το οποίο ενώ δε διαλύεται από το καθαρό νερό στη μορφή του αυτή, μετατρέπεται σε διαλυτό δισανθρακικό ασβέστιο, όταν δεχτεί την επίδραση του νερού της βροχής, που εμπλουτίζεται στην ατμόσφαιρα με διοξείδιο του άνθρακα και γίνεται έτσι όξινο. Αυτό συμβαίνει σύμφωνα με την ακόλουθη αντίδραση:



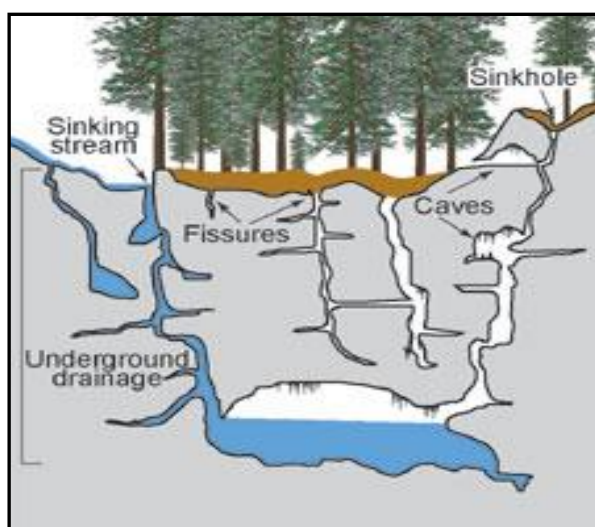
Τα πετρώματα δηλαδή αυτά, διαρρηγμένα καθώς είναι συνήθως, επιτρέπουν στο νερό της βροχής να διεισδύσει βαθύτερα και να προκαλέσει τη χημική αυτή διάβρωσή τους. Όταν οι ασβεστόλιθοι είναι δολομιτικοί, τότε η διάλυση αυτή είναι εντονότερη. Η έννοια της καρστικότητας έχει επεκταθεί και σε άλλα πετρώματα που παρουσιάζουν παρόμοια μορφολογία, όπως οι μάργες, οι γύψοι, το ορυκτό αλάτι κλπ.

Τα μορφολογικά χαρακτηριστικά των τοπίων που έχουν υποστεί το φαινόμενο της καρστικότητας είναι κυρίως:

- I. Η απουσία μόνιμου επιφανειακού δικτύου και, αντίθετα, η αφθονία υπόγειας υδροφορίας. Αυτό έχει ως αποτέλεσμα άλλοτε την εκδήλωση πηγών στα σημεία επαφής του υδροπερατού στρώματος με ένα αδιαπέρατο, άλλοτε την ανάπτυξη υπόγειων ποταμών, που οφείλονται στην καταβύθιση των επιφανειακών μέσα σε υπόγειους οχετούς και σπήλαια και άλλοτε τη δημιουργία υπόγειων λιμνών κλπ.
- II. Η φανερή παρουσία ειδικών επιφανειακών μορφών, τις οποίες προκαλεί το φαινόμενο της καρστικότητας, λιγότερο ή περισσότερο εκτεταμένων, βαθιά αυλακωμένων, που ονομάζονται αμαξοτροχίες, αυλακίες, δακτυλογλυφές, δολίνες, ουβάλες και πόλγες. Οι καρστικές αυτές περιοχές δίνουν τοπία που μοιάζουν με ερειπωμένες πόλεις. Οι δολίνες είναι κλειστές λεκάνες, κυκλικές ή ελλειπτικές, συνήθως συγκεντρωμένες σε ομάδες. Οι ουβάλες είναι ένα προχωρημένο στάδιο δολινών, που οφείλεται στη συνένωση πολλών δολινών, με αποτέλεσμα τη δημιουργία πολυπλοκότερων μορφών. Πόλγη σημαίνει μία καρστική πεδιάδα με μεγάλο εμβαδόν και ισχυρές κλίσεις στην περιφέρεια. Οι ποταμοί, που διαρρέουν συνήθως τις πόλγες, αποχετεύουν τα νερά τους μέσα από τις καταβόθρες, που πάντα υπάρχουν σε αυτές. Πολλές φορές όμως οι καταβόθρες παθαίνουν παροδική έμφραξη από τα αδιάλυτα φερτά υλικά που αποθέτονται στον πυθμένα τους, με αποτέλεσμα να σχηματίζονται περιοδικά λίμνες. Γι' αυτό, στις δειναρικές χώρες παρατηρείται το φαινόμενο στην ίδια θέση το καλοκαίρι να καλλιεργείται σιτάρι ενώ το χειμώνα να ψαρεύουν. Ο σχηματισμός της πόλγης όσο και της

δολίνης οφείλεται σε μια επιφανειακή ταπείνωση από τεκτονικά αίτια ή από κατάρρευση της οροφής υπόγειων σπηλαίων. Παράδειγμα πόλγης στην Ελλάδα είναι η λεκάνη της Κωπαΐδας, που παλιότερα μεταβαλλόταν περιοδικά πότε σε λίμνη και πότε σε τέλμα, έως ότου, στο τέλος του 19^{ου} αιώνα, αποστραγγίστηκε τελείως με τεχνικά έργα, για να καλλιεργηθεί.

- III. Πολυποίκιλα είναι επίσης και τα υπόγεια αποτελέσματα της καρστικότητας. Το όξινο νερό της βροχής, περνώντας μέσα από τις ρωγμές και τις διακλάσεις των πετρωμάτων, δημιουργεί με τη διαλυτική του δράση υπόγειους αγωγούς, στοές, σπήλαια με λαβυρινθώδη μορφή, των οποίων το σύνολο αποτελεί τις λεγόμενες καταβόθρες, οι οποίες σχηματίζονται κατά την κατακόρυφη έννοια, ενώ κατά την οριζόντια έννοια η δράση αυτή του νερού δημιουργεί υπόγεια έγκοιλα (σπήλαια), (Bonacci, 2001 και Σκιτζί κ.ά, 2007).



Σχήμα 2.1: Καρστική πηγή

Οι καταβόθρες, όταν αποτελούνται από μια πινακοειδή κοιλότητα, φράζουν εύκολα και έτσι μετατρέπεται η κλειστή λεκάνη σε έλος ή λίμνη. Αντίθετα, όταν παρουσιάζουν χάσματα και μεγάλες οπές το νερό χύνεται προς άγνωστες κατευθύνσεις ρέοντας μέσα από λαβυρινθώδεις οχετούς. Σπάνια βρίσκεται η υπόγεια ροή και το σημείο όπου αναβλύζουν και πάλι τα νερά αυτά.

Τα σπήλαια μπορεί να έχουν σημαντικές διαστάσεις και πολυσύνθετη δομή, με σχηματισμό λιμνών, καταρρακτών, σιφώνων, φυσικών γεφυρών κλπ. Άλλες μορφές που συναντάμε μέσα στα σπήλαια αυτά είναι οι σταλακτίτες και οι σταλαγμίτες, που προέρχονται από την απόθεση ανθρακικού ασβεστίου, καθώς το νερό που το μεταφέρει εξατμίζεται ή χάνει το διοξείδιο του άνθρακα. Οι σταλακτίτες κρέμονται από την οροφή του σπηλαίου, ως κρύσταλλοι πάγου, με μεγάλη ποικιλία μορφών, ενώ οι σταλαγμίτες ανυψώνονται από το έδαφος προς τα επάνω.

Ο πλούτος μορφών, που δημιουργείται πολλές φορές μέσα στα σπήλαια αυτά, ζωντανεύει ένα μαγευτικό υπόγειο κόσμο, ιδίως όταν υπάρχει και υπόγεια λίμνη. Γι' αυτό τα σπήλαια αυτά παρουσιάζουν μεγάλο τουριστικό ενδιαφέρον. Τα σπήλαια εξετάζονται από ιδιαίτερο κλάδο της γεωλογίας, τη σπηλαιολογία (Bakalowicz, 2005).

Το φαινόμενο αυτό της καρσικότητας, εκτός από το ίδιο το Καρστ της Σλοβενίας, εμφανίζεται και σε πολλές άλλες χώρες, όπου υπάρχουν οι προϋποθέσεις που το καθορίζουν. Υπάρχουν πάμπολλες περιπτώσεις του και στη Ελλάδα. Πασίγνωστα και ωραιότατα είναι τα ελληνικά σπήλαια της Αντιπάρου, της Γλυφάδας Δυρού Λακωνίας, των Πετραλώνων Χαλκιδικής και πολλά άλλα (περίπου 4000), που έχουν μελετηθεί ως σήμερα, (Σκιτζί κ.ά, 2007).



Σχήμα 2.2: Κυριότερες καρστικές περιοχές στον Ελλαδικό χώρο

2.4.1 Καρστικός κύκλος

Κατά τον Davies η εξέλιξη της καρστικής διάβρωσης μπορεί να περιγραφεί ως ένας κύκλος που περιλαμβάνει τρία στάδια. Τα στάδια νεότητας, ωριμότητας και γήρατος.

Το **στάδιο νεότητας** αρχίζει με την εμφάνιση των πρώτων τοπογραφικών καρστικών χαρακτηριστικών. Αν τα ευδιάλυτα πετρώματα καλύπτονται από αδιάλυτα, το στάδιο νεότητας μπορεί να αρχίσει πριν απομακρυνθεί τελείως το αδιάλυτο κάλυμμα. Το στάδιο νεότητας περιλαμβάνει την εμφάνιση καρστικών φαινομένων στα πετρώματα, όπως αμαξοτροχιών, καταβοθρών, δολινών. Οι δολίνες αυξάνουν σε μέγεθος και αριθμό κάθε μία χωριστά και κάθε επιφανειακό ρεύμα αποκτά υπόγεια ροή. Το στάδιο αυτό τελειώνει όταν οι κλιτύες των δολινών γίνουν ομαλές πλευρές κοιλάδας και οι κοιλάδες μετατραπούν σε κοιλάδες καταβύθισης.

Το **στάδιο ωριμότητας** περιλαμβάνει τη δημιουργία ουβαλών από τη συνένωση των δολινών. Η μορφολογία της πόλγης παραμένει σχεδόν αμετάβλητη από την έλλειψη επιφανειακής απορροής. Η διάβρωση έχει φτάσει στο υποκείμενο αδιάλυτο πέτρωμα, ενώ παραμένουν υπολειμματικοί σχηματισμοί του καρστικοποιημένου πετρώματος.

Το **στάδιο γήρατος** περιλαμβάνει τη σμίκρυνση των ράχων μεταξύ των ουβαλών. Η διάβρωση έχει φτάσει στο βασικό επίπεδο που είναι η στάθμη του υδροφόρου ορίζοντα ή το υποκείμενο αδιάλυτο πέτρωμα. Στο τέλος του σταδίου αυτού αρχίζει η επιφανειακή ροή των ρευμάτων. Η επιφάνεια στην περιοχή είναι σχεδόν επίπεδη και υψώνονται μερικοί υπολειμματικοί σχηματισμοί που ονομάζονται hums. Παράδειγμα hums στην Αττική αποτελεί η Ακρόπολη και ο Λυκαβηττός, (Davies et al,1968)

2.4.2 Χαρακτηριστικές καρστικές μορφές

Καταβόθρες	Δακτυλιογλυφές
Καρστικές κοιλάδες	Αμαξοτροχιές
Καρστικές σήραγγες, στοές, αγωγοί	Μικρά έγκοιλα, φρέατα, αυλοί, κ.λ.π.
Καρστικές γέφυρες, αψίδες, κ.λ.π.	Δολίνες (χοανοειδής μορφή)
Καρστικά σπήλαια (συνήθως με σταλακτίτες και σταλαγμίτες)	Πόλγες (πολύ μεγάλα έγκοιλα, πεδιάδες)
Καρστικά δίκτυα	Ουβάλες (πολλές δολίνες ενωμένες)

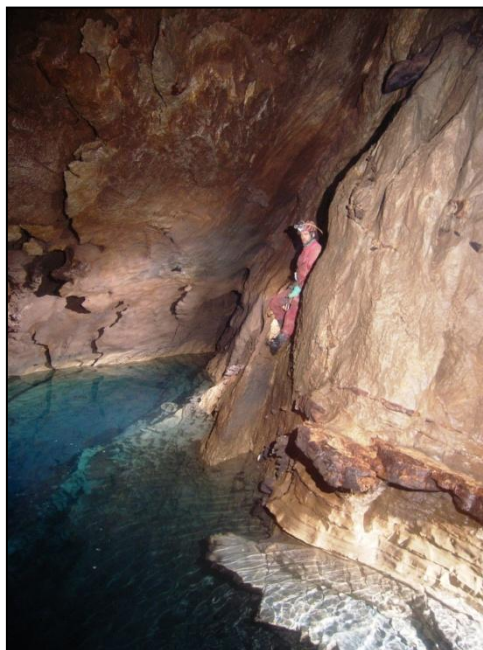
2.4.3 Εμφάνιση καρστικών στον Ελλαδικό χώρο.

Η Ελλάδα καλύπτεται κατά το 55% περίπου από καρστικά πετρώματα επιδεκτικά στις διαβρώσεις και τις χημικές διεργασίες.

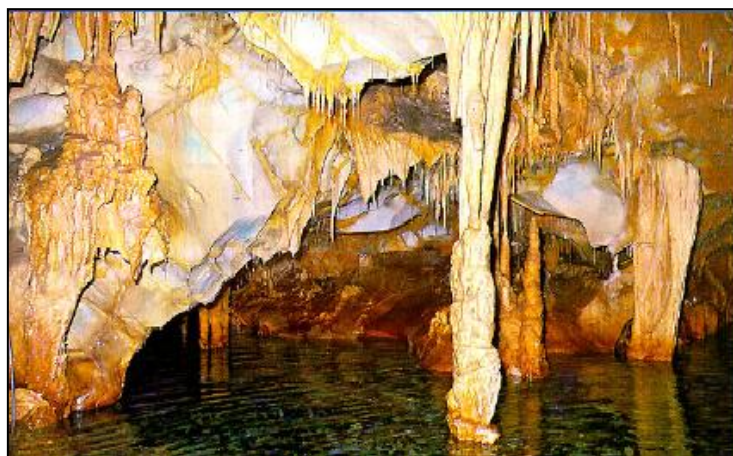
Χαρακτηριστικά παραδείγματα περιοχών στον Ελλαδικό χώρο αποτελούν:

- **Κεφαλονιά (σπήλαιο Δρογαράτης):** Ανήκει στην κοινότητα Χαλιωτάτων και με τη Σάμη το συνδέει ασφαλωμένος δρόμος. Βρίσκεται σε υψόμετρο 120 μέτρων και έχει μήκος 95 μέτρων. Το μεγαλύτερο πλάτος του είναι 45 μέτρα και το βάθος του 21 μέτρα. Το κυρίως σπήλαιο αποτελείται από ένα θάλαμο χωρισμένο σε δύο τμήματα. Στο πρώτο τμήμα, πάνω στους ογκόλιθους που αποσπάστηκαν από την οροφή, έχουν αναπτυχθεί διαφανείς σταλαγμίτες.
- **Πελοπόννησος (Γλυφάδα Δυρού Λακωνίας):** Ανακαλύφθηκε στο τέλος του δέκατου ένατου αιώνα. Βρίσκεται στα δυτικά παράλια της Λακωνικής Χερσονήσου στο μυχό του όρμου Δυρού σε απόσταση 40 χιλιομέτρων από το Γύθειο και 5 χιλιόμετρα από τον Πύργο Δυρού. Το σπήλαιο είναι ένας υπόγειος ποταμός που αναπτύσσεται κυρίως σε δύο παράλληλους διαδρόμους και πολλούς δευτερεύοντες. Καλύπτει έκταση 33,200 τ.μ. και είναι άγνωστο ακόμα το τελικό μήκος του σπηλαίου.
- **Κρήτη:** Από ειδικούς συχνά χαρακτηρίζεται ως ένας σπηλαιολογικός παράδεισος, αφού από το 1962 έχουν καταγραφεί περισσότερα από 3500 σπήλαια, από τα οποία μονάχα τα 800 έχουν εξερευνηθεί. Σήμερα γνωρίζουμε περίπου 3500 καρστικές μορφές στην Κρήτη. Ιδιαίτερο

ενδιαφέρον στο υδατικό διαμέρισμα της Κρήτης, από υδρολογική άποψη, παρουσιάζουν οι τρεις μεγάλες σε έκταση ανθρακικές ενότητες που αναπτύσσονται στους ορεινούς όγκους των Λευκών Ορέων, του Ψηλορείτη και της Δείκτης-Σελένας στα ανατολικά και δευτερευόντως οι καρστικές ενότητες της Σητείας. Οι ασβεστιλιθικοί αυτοί όγκοι τροφοδοτούν μεγάλο αριθμό αξιόλογων πηγών στην περίμετρό τους. Ιδιαίτερο χαρακτηριστικό στην Κρήτη είναι η ύπαρξη μεγάλων παράκτιων και υποθαλάσσιων καρστικών πηγών αυξημένης περιεκτικότητας σε χλωριόντα (Cl^-) μαζί με τις αντίστοιχες μεγάλες καρστικές πηγές με καλής ποιότητας υπόγειο νερό. Η ανυπαρξία γεωλογικών φραγμών στις εκτεταμένες ανθρακικές εμφανίσεις προς τη θάλασσα οδηγεί στην εκτεταμένη υφαλμύρωση των υπόγειων υδροφορέων με μηχανισμούς που ποικίλλουν από υδροφορέα σε υδροφορέα (Σκιτζί κ.ά, 2007).



Σχήμα 2.3: Καρστικές μορφές στα Λευκά Όρη Χανίων (Γαλλο-Ελληνική αποστολή)



Σχήμα 2.4: Σπήλαιο Γλυφάδας Δυρού

Κεφάλαιο 3

Περιοχή Μελέτης

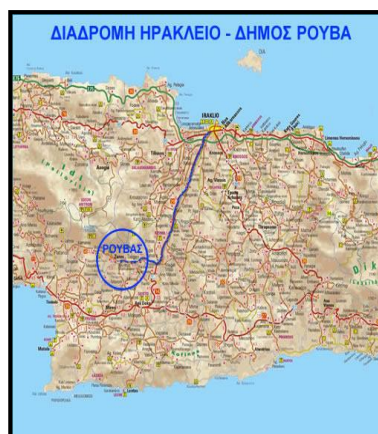
3.1 Γεωγραφικά στοιχεία

Η περιοχή μελέτης της συγκεκριμένης μεταπτυχιακής εργασίας είναι ο δήμος Ρούβα, ο οποίος ανήκει στην περιφέρεια Κρήτης και στο νομό Ηρακλείου. Σύμφωνα με την απογραφή του 2001 ο δήμος έχει συνολικά 2.324 κατοίκους, μειωμένος κατά 400 κατοίκους από την περασμένη απογραφή του 1991. Από το γεγονός αυτό συμπεραίνεται ότι το ανθρώπινο ενεργό δυναμικό φθίνει με ανησυχητικούς ρυθμούς. Το κύριο πλεονέκτημα του Ρούβα είναι οι φυσικές πηγές από όπου αναβλύζουν μεγάλες ποσότητες πόσιμου νερού και το δάσος του Ρούβα, το οποίο βρίσκεται στα βόρεια του Δήμου στον Ψηλορείτη.



Σχήμα 3.1: Χάρτης νήσου Κρήτης

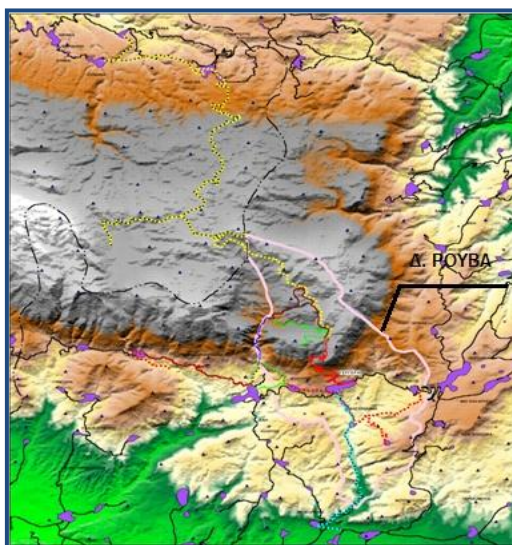
Η έκταση ανέρχεται σε 62.725 στρέμματα, από τα οποία περίπου τα 26.000 είναι καλλιεργήσιμα και τα υπόλοιπα αποτελούν δασική έκταση. Αποτελεί αγροτικό Δήμο, καθώς η πλειοψηφία των ανθρώπων ασχολείται με τη γεωργία και την κτηνοτροφία.



Σχήματα 3.2 και 3.3: Θέση Δήμου Ρούβα στο νομό Ηρακλείου

Ο δήμος Ρούβα αποτελείται από τρία δημοτικά διαμερίσματα με τους οικισμούς:

1. Δημοτικό διαμέρισμα Γέργερης: Γέργερη, Απομαρμά, Καρδαμιανά, Μαστραχιανά, Ράπτης, Τζανιανά και Ψαλίδα
2. Δημοτικό διαμέρισμα Νυβρίτου: Νύβριτος και Δρόσοι
3. Δημοτικό διαμέρισμα Πανάσου



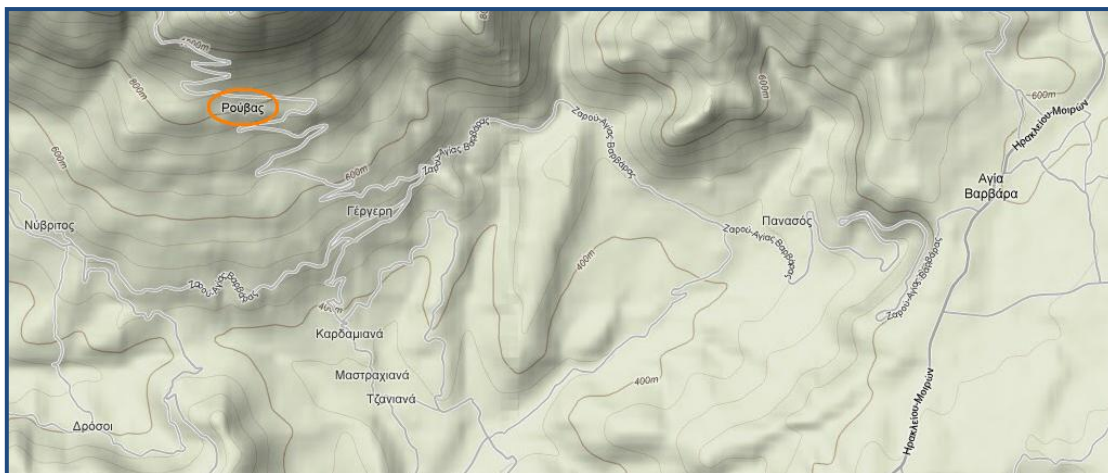
Σχήμα 3.4: Τμήμα γεωγραφικού χάρτη Δήμου Ρούβα



Σχήμα 3.5: Δορυφορική άποψη περιοχής Δήμου Ρούβα (Πηγή: Google Earth)



Σχήμα 3.6: Δορυφορική άποψη περιοχής Δήμου Ρούβα (Πηγή: Google Earth)



Σχήμα 3.7: Άποψη εδάφους περιοχής (Πηγή: Google Earth)

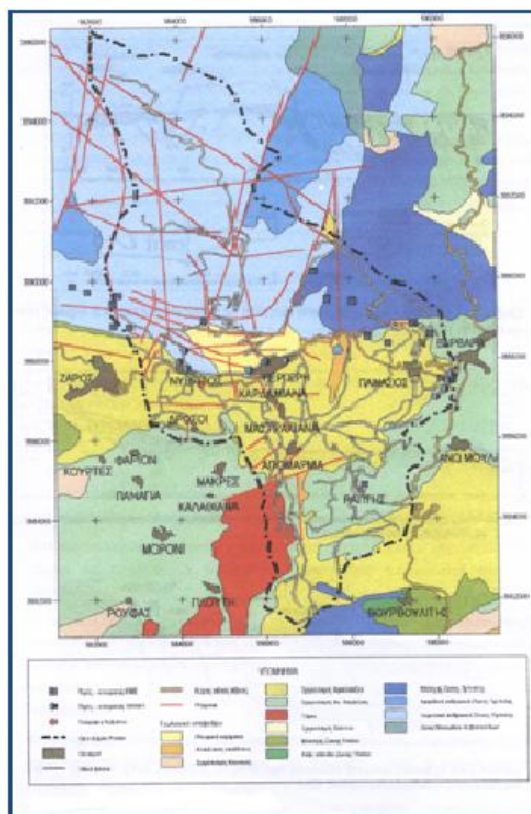
Υπάρχουν πέντε πηγές στο Δ.Δ. Γέργερης, αξιοποιήσιμες εκ των οποίων η Φουντάνα (ονομασία της πηγής), έχει τις μεγαλύτερες παροχές νερού. Αυτό το νερό, που εμφανίζεται στο ομώνυμο εμφιαλωτήριο νερού «ΡΟΥΒΑΣ», πρέπει να σημειωθεί ότι θεωρείται το δεύτερο καλύτερο νερό στην Ευρώπη.

Στο Δ.Δ. Νυβρίτου υπάρχουν τρεις πηγές, μικρότερης αξιοποίησης, σε σχέση με τις πηγές της Γέργερης, οι οποίες όμως καλύπτουν το σύνολο των αναγκών των κατοίκων σε ύδρευση και σχεδόν το 100% τις ανάγκες σε άρδευση της περιοχής. Τους θερινούς μήνες συνεπικουρεί και μία μικρή γεώτρηση, που υπάρχει στο νότια τμήμα του οικισμού, με παροχή $10^3 \text{ m}^3/\text{h}$ (Πηγή: Δήμος Ρούβας).

3.2 Γεωλογική και υδρολογική δομή του Δήμου Ρούβα

Η γεωλογική δομή της περιοχής μελέτης αποτελείται από τεταρτογενείς και νεογενείς σχηματισμούς στο νότιο ήμισυ του Δ. Ρούβα και από φλύσχη και ανθρακικά της ζώνης της Τρίπολης στο βόρειο τμήμα. Πρόκειται για περιοχή, η οποία χαρακτηρίζεται από την καρστικότητα του υπεδάφους της.

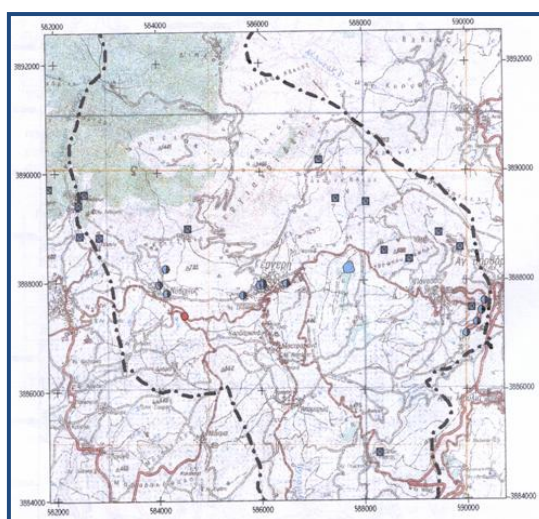
Στον γεωλογικό χάρτη του Δήμου Ρούβα που ακολουθεί με τη μαύρη διακεκομμένη γραμμή εμφανίζονται τα όρια του δήμου Ρούβα. Επίσης με τους μικρούς μπλε κύκλους φαίνονται οι πηγές καταγεγραμμένες από τον Οργανισμό Ανάπτυξης Ανατολικής Κρήτης (ΟΑΝΑΚ) και με τα μικρά τετράγωνα οι πηγές που έχουν καταγραφεί από το Ινστιτούτο Γεωλογικών και Μεταλλευτικών Ερευνών (ΙΓΜΕ) (Σχήματα 3.8 και 3.9).



Σχήμα 3.8: Γεωλογικός χάρτης Δήμου Ρούβα (Πηγή: Δήμος Ρούβας)

3.3 Επιφανειακά νερά

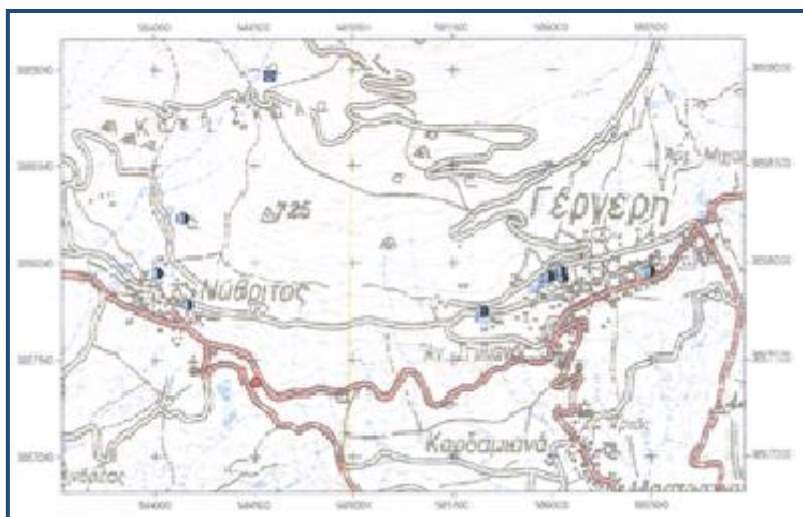
Στον Δήμο Ρούβα υπάρχουν 23 πηγές εκ των οποίων οι 12 έχουν καταγραφεί από τον Οργανισμό Ανάπτυξης Ανατολικής Κρήτης (ΟΑΝΑΚ).



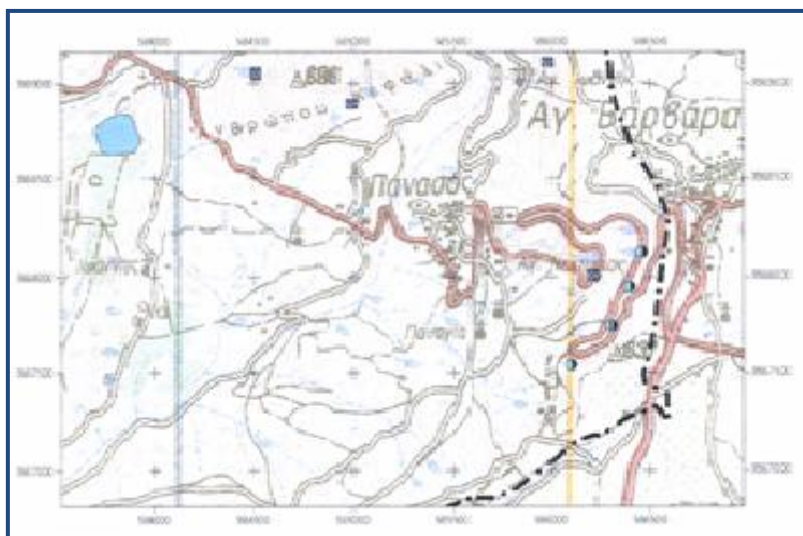
Σχήμα 3.9: Πηγές περιοχής Δήμου Ρούβα (Πηγή: Δήμος Ρούβας)

Με τους μικρούς μπλε – μαύρους κύκλους συμβολίζονται οι πηγές καταγεγραμμένες από τον ΟΑΝΑΚ, με τα μπλε τετράγωνα οι πηγές που καταγράφηκαν από το ΙΓΜΕ και με κόκκινο κύκλο η γεώτρηση της Νυβρίτου. Με ανοιχτό μπλε συμβολίζεται η λιμνοδεξαμενή της Γέργερης.

3.3.1 Πηγές Δήμου Ρούβα (Γέργερης – Νυβρίτου- Πανασού)



Σχήμα 3.10: Πηγές περιοχής Γέργερης – Νυβρίτου (Πηγή: Δήμος Ρούβας)

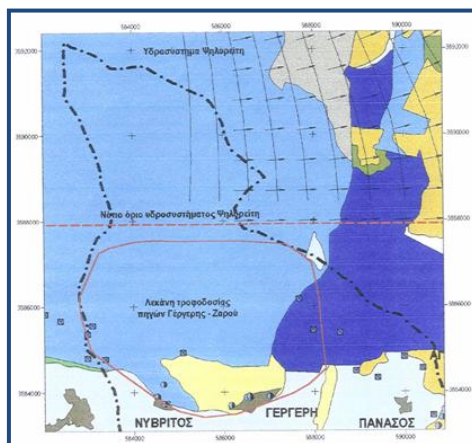


Σχήμα 3.11: Πηγές περιοχής Πανασού (Πηγή: Δήμος Ρούβας)

3.3.2 Υδρολογική λεκάνη Δήμου Ρούβα

3.3.2.1 Λεκάνη τροφοδοσίας πηγών Γέργερης-Ζαρού

Με την διακεκομμένη κόκκινη γραμμή ορίζεται το νότιο όριο του υδροσυστήματος του Ψηλορείτη. Σημειώνεται επίσης ότι με την κόκκινη γραμμή οροθετείται η λεκάνη τροφοδοσίας των πηγών Γέργερης-Ζαρού (Πηγή: Δήμος Ρούβας).

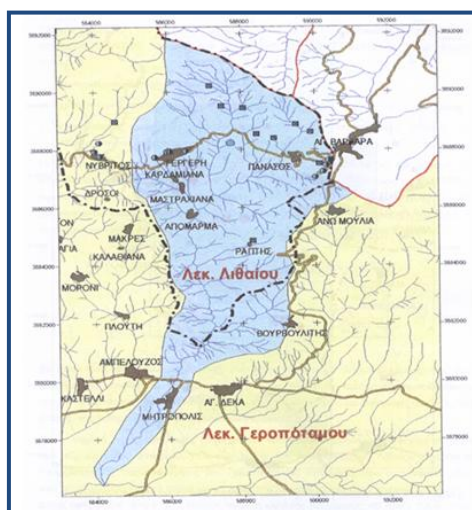


Σχήμα 3.12: Όρια υδρογεωλογικών συστημάτων περιοχής (Πηγή: Δήμος Ρούβας)

Με την κόκκινη διακεκομμένη γραμμή συμβολίζεται το νότιο όριο του υδροσυστήματος του βόρειο – ανατολικού τομέα του Ψηλορείτη που εκφορτίζει στην πηγή του Αλμυρού ποταμού. Η κόκκινη γραμμή οριοθετεί τη λεκάνη τροφοδοσίας των πηγών Γέργερης – Ζαρού.

3.3.2.2 Υδρολογική Λεκάνη Λιθαίου

Ο Δήμος Ρούβα βρίσκεται εξ ολοκλήρου στην υδρολογική λεκάνη του Γεροπόταμου με τη νότια χαμηλή ζώνη του να καταλαμβάνει την λεκάνη του παραπόταμου Λιθαίου και τη βόρεια υψηλή ζώνη να καταλαμβάνει μέρος του βόρειου τμήματος της λεκάνης του παραπόταμου Κουτσουλίδη (Πηγή: Δήμος Ρούβας).



Σχήμα 3.13: Υδρολογικές λεκάνες Δήμου Ρούβα (Πηγή: Δήμος Ρούβας)

Κεφάλαιο 4

Νευρωνικά Δίκτυα

4.1 Ορισμός

Τα τεχνητά νευρωνικά δίκτυα έχουν αναπτυχθεί σε διεθνές επίπεδο μόνο κατά τις τελευταίες δεκαετίες. Ένας μεγάλος αριθμός επιστημόνων ασχολείται με την συγκεκριμένη ερευνητική περιοχή, σημειώνοντας σημαντικά επιτεύγματα, προωθώντας έτσι την εφαρμογή των νευρωνικών δικτύων στις τεχνολογικές επιστήμες.

Χαρακτηρίζονται από τις πρώτες αρχές και λειτουργίες τους, οι οποίες εμπνέονται από το νευρικό σύστημα των ζώντων οργανισμών, αλλά η μελέτη και η χρήση τους έχει προχωρήσει πολύ πέρα από τους βιολογικούς οργανισμούς. Τα νευρωνικά δίκτυα κατέχουν μια τελείως διαφορετική φιλοσοφία από τον τρόπο με τον οποίο δουλεύουν οι κλασικοί υπολογιστές. Συνδυάζουν τον τρόπο σκέψης του ανθρώπινου εγκεφάλου με τον αφηρημένο μαθηματικό τρόπο σκέψης. Επομένως στα νευρωνικά δίκτυα το δίκτυο μαθαίνει και εκπαιδεύεται, θυμάται ή ξεχνά μια αριθμητική τιμή κτλ., λειτουργεί δηλαδή ακριβώς όπως η ανθρώπινη σκέψη. Σημαντικό είναι επίσης το γεγονός ότι χρησιμοποιούν επί πλέον και περίπλοκες μαθηματικές συναρτήσεις και κάθε είδους εργαλεία από την μαθηματική ανάλυση (Haykin, 1994 και Rojas, 1996).

Αρχικά, όμως θα πρέπει να δώσουμε μία απάντηση στο ερώτημα: *«Τι είναι τελικά το τεχνητό νευρωνικό δίκτυο;»*.

Προσδοκώντας να δώσουμε έναν ορισμό για το νευρωνικό δίκτυο, θα μπορούσαμε να πούμε ότι πρόκειται για ένα δίκτυο απλών μονάδων επεξεργασίας (νευρώνων) που εμφανίζουν πολύπλοκη, συλλογική και παράλληλη συμπεριφορά, η οποία καθορίζεται από τις συνδέσεις μεταξύ των μονάδων επεξεργασίας και από τις παραμέτρους αυτών. Το δίκτυο περιλαμβάνει αλγόριθμους σχεδιασμένους να τροποποιούν την ισχύ των εν λόγω συνδέσεων μεταξύ των νευρώνων, ώστε να επιτευχθεί η επιθυμητή ροή σήματος (Καραμπερίδου, 2007).

Ειδικότερα, το τεχνητό νευρωνικό δίκτυο είναι μια αρχιτεκτονική δομή αποτελούμενη από ένα πλήθος διασυνδεδεμένων μονάδων –τους τεχνητούς νευρώνες. Κάθε μονάδα χαρακτηρίζεται από εισόδους και εξόδους και υλοποιεί τοπικά έναν απλό υπολογισμό. Κάθε σύνδεση μεταξύ δύο μονάδων χαρακτηρίζεται από μία τιμή βάρους. Οι τιμές των βαρών συνδέσεων αποτελούν τη γνώση

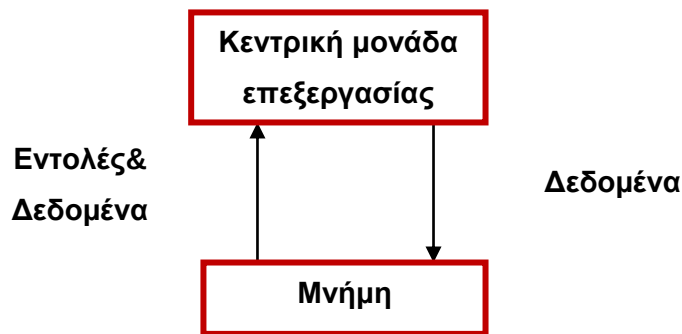
που είναι αποθηκευμένη στο δίκτυο και καθορίζουν τη λειτουργικότητά του. Η έξοδος κάθε μονάδας καθορίζεται από τον τύπο της μονάδας, τη διασύνδεση με τις υπόλοιπες μονάδες και πιθανώς κάποιες εξωτερικές εισόδους. Πέρα από μία πιθανή δεδομένη εκ κατασκευής λειτουργική ικανότητα ενός δικτύου, συνήθως ένα δίκτυο αναπτύσσει μια συνολική λειτουργικότητα μέσω μιας μορφής εκπαίδευσης.

Ο όρος «τεχνητό», αναφέρεται στο ότι οι υπολογισμοί υλοποιούνται με τη χρήση ηλεκτρονικών και όχι βιολογικών στοιχείων. Η διαρκώς αναπτυσσόμενη τεχνολογία των τεχνητών νευρωνικών δικτύων αποτελεί ένα ευρύ σύνολο τεχνικών, η οποία προσφέρει μία πρακτική, εναλλακτική των συμβατικών προσέγγιση και σε πολλές περιπτώσεις παρέχει τη δυνατότητα της αντιμετώπισης δύσκολων προβλημάτων. Οι συμβατικοί ηλεκτρονικοί υπολογιστές υπερτερούν των ανθρώπων στην ταχύτητα των πράξεων, αλλά είναι ακατάλληλοι για την επεξεργασία δεδομένων που δεν έχουν ακρίβεια ή πληρότητα. Τα τεχνητά νευρωνικά δίκτυα κατάφεραν να δώσουν λύσεις σε τέτοια δύσκολα προβλήματα. Είναι γεγονός ότι οι συμβατικοί ηλεκτρονικοί υπολογιστές χρειάζονται μία βήμα – βήμα λύση ενός προβλήματος και τον προγραμματισμό τους με μία σειρά εντολών την οποία πρέπει να ακολουθήσουν για να λύσουν το πρόβλημα. Αντιθέτως τα τεχνητά νευρωνικά δίκτυα δεν χρειάζονται μια συγκεκριμένη περιγραφή της λύσης του προβλήματος, αλλά ο χρήστης επιτρέπει στο τεχνητό νευρωνικό δίκτυο να εκπαιδευτεί και να προσαρμοστεί κατά τη διάρκεια μιας περιόδου εκπαίδευσης, όπως θα περιγραφεί αναλυτικά παρακάτω σε επόμενες παραγράφους.

Αξίζει να σημειωθεί, πως αυτό που χαρακτηρίζει τα τεχνητά νευρωνικά δίκτυα είναι ότι οι επιστήμονες, που μελετούν το συγκεκριμένο ερευνητικό αντικείμενο, προέρχονται από πλήθος ερευνητικών περιοχών, όπως η Ιατρική, οι διάφοροι κλάδοι των Μηχανικών, η Φυσική, η Χημεία, τα Μαθηματικά, η επιστήμη Υπολογιστών και άλλα. Το γεγονός αυτό απορρέει από τις πολλαπλές απαιτήσεις γνώσεων, αλλά και τεχνικών που απαιτούν τα νευρωνικά δίκτυα για την ανάπτυξή τους. Επομένως, είναι άκρως κατανοητή η προσφορά των νευρωνικών δικτύων στον άνθρωπο, λόγω των συνδυαστικών τους εφαρμογών.

Το νευρωνικό δίκτυο αποτελεί μία αλληλοσύνδεση απλών στοιχείων προς επεξεργασία, μονάδων ή κόμβων, των οποίων η λειτουργία βασίζεται και εμπνέεται από τους βιολογικούς νευρώνες. Η ικανότητα επεξεργασίας του δικτύου πηγάζει από την «εσωτερική ισχύ» της ένωσης των μονάδων ή τα βάρη, μέσω διαδικασιών εκμάθησης και προτύπων εκπαίδευσης (Anderson, 1988).

Η λειτουργία των συμβατικών ηλεκτρονικών υπολογιστών ακολουθεί το μοντέλο που παρουσιάζεται στο σχήμα 4.1:



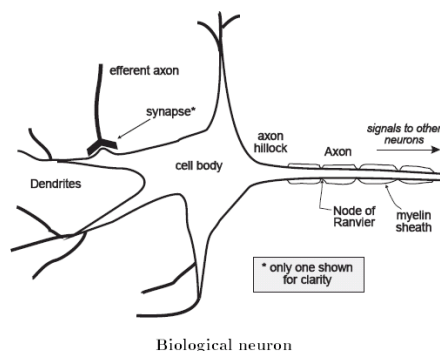
Σχήμα 4.1: Μηχανή von – Neumann (Πηγή: Gurney, Routledge, (1997)).

Ο συμβατικός ηλεκτρονικός υπολογιστής εκτελεί επανειλημμένα την ακόλουθη σειρά γεγονότων:

1. Φορτώνει μία εντολή από τη μνήμη.
2. Φορτώνει όλα τα απαραίτητα δεδομένα τα οποία απαιτούνται σύμφωνα με τις εντολές που φορτώθηκαν από τη μνήμη.
3. Εκτελεί την εντολή, επεξεργαζόμενος τα δεδομένα.
4. Αποθηκεύει τα αποτελέσματα στη μνήμη.
5. Επανέρχεται στο βήμα 1.

Προκύπτει ότι πολλά από τα καθημερινά ζητήματα είναι πολύ δύσκολο να τυποποιηθούν με αυτόν τον τρόπο, προκειμένου να υπάρξει επίλυση, ενώ τα νευρωνικά δίκτυα δίνουν ήδη ικανοποιητικά αποτελέσματα. Τέτοια παραδείγματα αποτελούν η επεξεργασία εικόνας, η αναγνώριση προτύπων και η τεχνητή νοημοσύνη (π.χ. υλοποίηση έμπειρων συστημάτων).

Οι σημαντικές ιδιότητες των βιολογικών συστημάτων, όπως η προσαρμοστικότητα, η ικανότητα αναγνώρισης από τα συμφραζόμενα, η ανοχή στα λάθη, η μεγάλη χωρητικότητα μνήμης, η ικανότητα επεξεργασίας βιολογικών πληροφοριών σε πραγματικό χρόνο (κυρίως από τον εγκέφαλο) μας κατευθύνουν στη μελέτη και την προσπάθεια προσομοίωσης αυτών των εναλλακτικών βιολογικών αρχιτεκτονικών. Σε πλήρη αντιστοιχία με το απλοποιημένο μοντέλο του βιολογικού νευρώνα αναπτύχθηκε το μοντέλο του τεχνητού νευρώνα. Στο σχήμα 4.2 παρουσιάζεται ένας τυπικός βιολογικός νευρώνας.



Σχήμα 4.2: Βιολογικός νευρώνας, (Πηγή: Gurney, Routledge, (1997)).

4.2 Η έμπνευση από τη βιολογία

Κάθε ζωντανός οργανισμός κατέχει ένα νευρικό σύστημα, το οποίο αποτελείται από πλήθος νευρωνικών δικτύων, εξειδικευμένο σε λειτουργίες, όπως η μνήμη και η μάθηση. Οι λεγόμενοι νευρώνες ή νευρώνια αποτελούν τις κύριες και ανεξάρτητες μονάδες του νευρωνικού δικτύου, καθώς έχουν τη δυνατότητα να επεξεργάζονται δεδομένα και κατόπιν να μεταφέρουν ηλεκτρικά σήματα σε άλλους νευρώνες (Αργυράκης, 2001). Αξίζει να αναφερθεί ότι ο εγκέφαλος αποτελεί την κεντρική μονάδα κάθε νευρικού συστήματος.

4.3 Από τα Βιολογικά Νευρωνικά Δίκτυα στα Τεχνητά

Η μελέτη των τεχνητών νευρωνικών δικτύων, κοινώς γνωστά σαν «νευρωνικά δίκτυα» υποκινήθηκε άμεσα από την αναγνώριση ότι οι εγκεφαλικοί υπολογιστές είναι ένα εξ' ολοκλήρου διαφορετικός δρόμος από τους συμβατικούς ψηφιακούς υπολογιστές. Τυπικά οι νευρώνες του εγκεφάλου είναι 5 - 6 τάξεις μεγέθους αργότεροι από τις λογικές πύλες σιλικόνης. Ο εγκέφαλος είναι ένας πολύ πολύπλοκος, μη-γραμμικός και παράλληλος υπολογιστής. Έχει τη δυνατότητα να οργανώνει τους νευρώνες έτσι ώστε να εκτελεί συγκεκριμένους υπολογισμούς πολύ πιο γρήγορα από τους πιο γρήγορους ψηφιακούς υπολογιστές που υπάρχουν.

Πώς τα καταφέρνει όμως ο εγκέφαλος; Κατά τη γέννησή του ο εγκέφαλος έχει την ικανότητα να κατασκευάζει τους δικούς του κανόνες, κοινώς «εμπειρία», η οποία μεγαλώνει με την πάροδο του χρόνου. Κατά τα 2 πρώτα χρόνια ζωής, έχουμε τη μέγιστη ανάπτυξη, όπου περίπου 1 εκατομμύριο συνάψεις (synapses) δημιουργούνται στο δευτερόλεπτο.

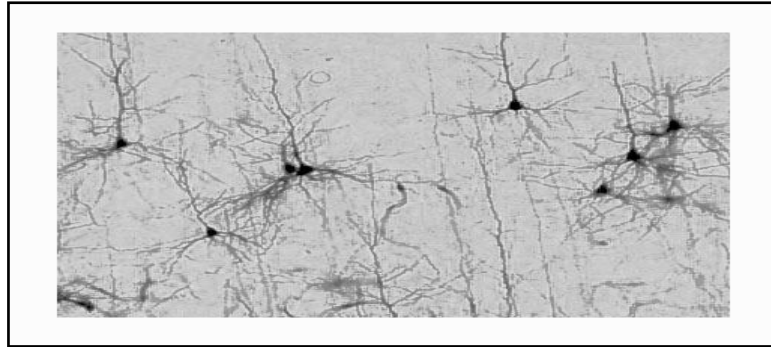
Οι συνάψεις είναι οι βασικές δομικές και λειτουργικές μονάδες που μεσολαβούν στην ενδοεπικοινωνία των νευρώνων, όπως φαίνεται στο παρακάτω σχήμα που δείχνει τη βασική δομή ενός νευρώνα. Άρα:

Ένα νευρωνικό δίκτυο είναι ένας συμπαγής παράλληλος κατανεμημένος επεξεργαστής, που έχει τη φυσική κλίση να αποθηκεύει εμπειριστατωμένη γνώση και να την κάνει διαθέσιμη για χρήση.

Ένα Τ.Ν.Δ. μοιάζει με τον εγκέφαλο στα εξής:

- Η γνώση αποκτάται από το δίκτυο μέσα από διαδικασία μάθησης.
- Οι δυνάμεις σύνδεσης των νευρώνων, γνωστές σαν συναπτικά (synaptic) βάρη, χρησιμοποιούνται για την αποθήκευση γνώσης.

Η διαδικασία για την εκμάθηση ονομάζεται «αλγόριθμος μάθησης». Ο καθορισμός του τρόπου μεταβολής των συναπτικών (synaptic) βαρών, αποτελεί την παραδοσιακή μέθοδο για τον σχεδιασμό των νευρωνικών δικτύων. Ένα παράδειγμα βιολογικού νευρωνικού δικτύου παρουσιάζεται στο σχήμα 4.3 που ακολουθεί.



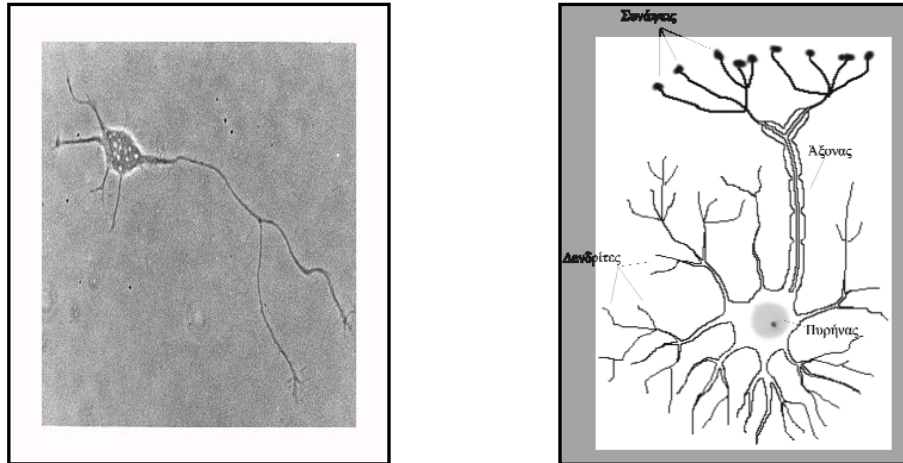
Σχήμα 4.3: Ένα βιολογικό νευρωνικό δίκτυο

4.3.1 Ο Βιολογικός Νευρώνας

Η λειτουργία του βιολογικού συστήματος βασίζεται στη διασύνδεση εξειδικευμένων φυσικών κυττάρων που ονομάζονται νευρώνες. Η βασική λειτουργία που πραγματοποιεί ένας νευρώνας είναι η συσσώρευση των σημάτων που δέχεται από τους νευρώνες με τους οποίους συνδέεται η είσοδός του, το φιλτράρισμα και η ενίσχυση αυτών των σημάτων και η παραγωγή ενός σήματος εξόδου, το οποίο στη συνέχεια μεταδίδεται μέσω των συνάψεων προς τους νευρώνες με τους οποίους συνδέεται η έξοδος του. Ένα πολύ σημαντικό στοιχείο είναι ότι η επίδραση ενός νευρώνα στους γειτονικούς του, μπορεί να είναι είτε διεγερτική (θετική σύνδεση), είτε ανασταλτική (αρνητική σύνδεση).

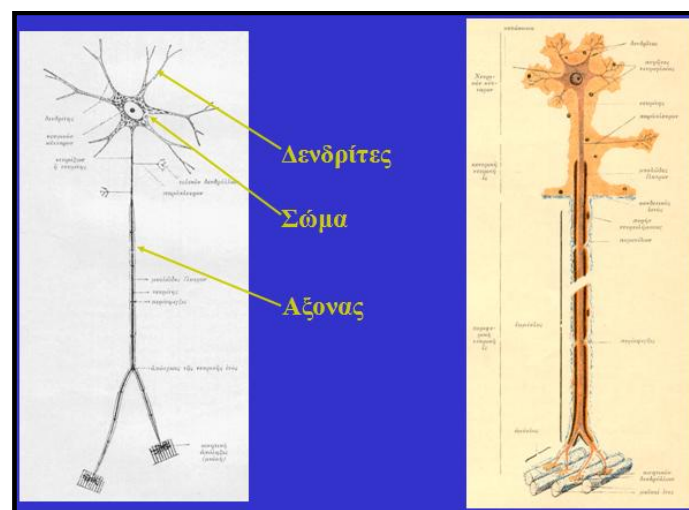
Ο νευρώνας αποτελείται από τα παρακάτω συστατικά στοιχεία. Το κεντρικό μέρος ονομάζεται σώμα του κυττάρου (cell body ή soma). Από το σώμα εξέρχονται κάποιες επεκτάσεις σαν ρίζες που ονομάζονται δενδρίτες (dendrites) όπως επίσης και ένα επίμηκες σωληνοειδές και λεπτό νεύρο, ο άξονας (axon) ο οποίος διαχωρίζεται στο τέλος του σε έναν αριθμό μικρών κλάδων. Το μέγεθος του σώματος ενός τυπικού νευρώνα είναι 10-80μm ενώ οι δενδρίτες και ο άξονας έχουν διάμετρο λίγων μm. Οι δενδρίτες, μαζί με το σώμα του κυττάρου, λειτουργούν ως λήπτες χημικών σημάτων. Ο άξονας, αντίθετα, μεταδίδει το ηλεκτρικό σήμα από το νευρώνα σε άλλα κύτταρα ή μυϊκές ίνες.

Η ένωση μεταξύ του τέλους ενός αξονικού κλάδου και ενός άλλου νευρώνα ονομάζεται σύναψη (synapse). Ωστόσο, η σύνδεση των νευρώνων μεταξύ τους δε γίνεται απευθείας, αλλά υπάρχει ένα μικρό κενό (συναπτικό κενό) πλάτους περίπου 200nm μεταξύ του κόμβου στην άκρη του αξονικού κλάδου και του δενδρίτη ή του σώματος του επόμενου νευρώνα. Ο νευρώνας του οποίου ο άξονας συνδέεται σε άλλον νευρώνα ονομάζεται προσυναπτικός (presynaptic), ενώ ο νευρώνας στον οποίο συνδέεται ο πρώτος ονομάζεται μετασυναπτικός (postsynaptic). Οι συνάψεις μπορεί να βρίσκονται είτε στο σώμα του κυττάρου είτε στους δενδρίτες είτε στο σώμα των επόμενων νευρώνων. Η επίδραση γενικά ελαττώνεται καθώς αυξάνεται η απόσταση από το σώμα. Το συνολικό μήκος των νευρώνων ποικίλει από 0.01mm (για νευρώνες στον εγκέφαλο) μέχρι 1m (για νευρώνες στα άκρα του σώματος). Στα σχήματα 4.4 και 4.5 παρουσιάζεται ο βιολογικός νευρώνας.

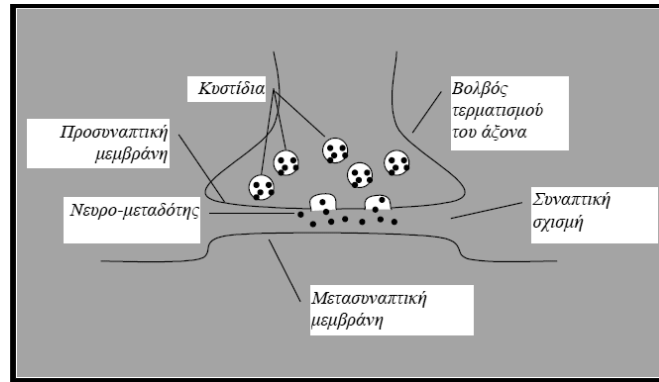


Σχήμα 4.4 και 4.5: Βιολογικός νευρώνας

Στον ανθρώπινο εγκέφαλο, ένας τυπικός νευρώνας συλλέγει σήματα από άλλους νευρώνες μέσω των δενδριτών. Όταν ο προσυναπτικός νευρώνας διεγείρεται στέλνει παλμούς ηλεκτρικής δραστηριότητας μέσω του άξονα ο οποίος, όπως είπαμε, διαχωρίζεται σε πολλά παρακλάδια. Στο τέλος καθενός από αυτά τα παρακλάδια, βρίσκεται η σύναψη η οποία μετατρέπει την δραστηριότητα από τον άξονα σε ηλεκτρικά αποτελέσματα τα οποία διεγείρουν ή αποτρέπουν την δραστηριότητα στους διασυνδεδεμένους νευρώνες. Τα ηλεκτρικά σήματα, φτάνουν στους δενδρίτες ή το κυτταρικό σώμα του μετασυναπτικού νευρώνα, μέσω του συναπτικού κενού, σε χρόνο περίπου 10 μs. Όταν ένας νευρώνας δεχθεί είσοδο διέγερσης που είναι αρκετά μεγαλύτερη συγκρινόμενη με την αποτρεπτική, στέλνει έναν παλμό ηλεκτρικής δραστηριότητας μέσω του άξονά του. Η εκπαίδευση λαμβάνει χώρα τροποποιώντας την αποδοτικότητα των συνάψεων, ώστε να αλλάξει η επίδραση ενός νευρώνα σε κάποιον άλλο (Haykin, 1994). Στα σχήματα 4.6 και 4.7 που ακολουθούν περιγράφεται αναλυτικά τόσο η ανατομία του νευρικού κυττάρου, όσο και ο τύπος των συνάψεων των νευρώνων.



Σχήμα 4.6: Ανατομία του νευρικού κυττάρου, (Πηγή: Διαμαντάρας, (2006)).



Σχήμα 4.7: Τύπος συνάψεως νευρώνων

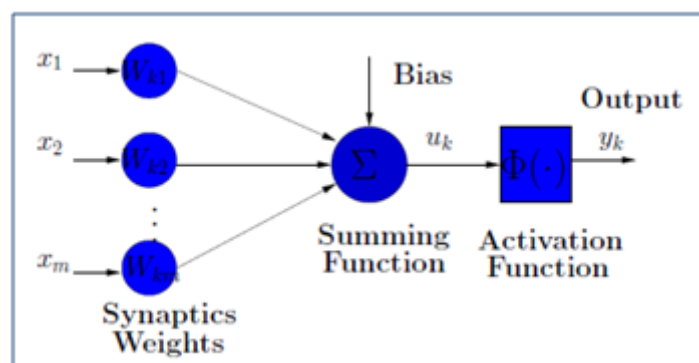
4.4 Αναλυτική περιγραφή ενός απλού τεχνητού νευρώνα

4.4.1 Το μοντέλο του τεχνητού νευρώνα

Ένας νευρώνας είναι μια μονάδα επεξεργασίας πληροφορίας, που είναι θεμελιακή για την λειτουργία ενός νευρωνικού δικτύου. Το μοντέλο του τεχνητού νευρώνα αποτελείται από τρία βασικά στοιχεία (Haykin, 1994):

1. Ένα σύνολο από συνάψεις ή συνδετικούς κρίκους. Κάθε μια από τις οποίες χαρακτηρίζεται από ένα βάρος” ή “δύναμη”. Συγκεκριμένα, ένα σήμα x_j στην είσοδο της σύναψης j που συνδέεται στον νευρώνα k , πολλαπλασιάζεται με το συναπτικό βάρος w_{kj} . Το βάρος w_{kj} είναι θετικό αν η σύναψη είναι διεγερτική (δηλαδή ωθεί τον νευρώνα να αποκριθεί στη διέγερση) και αρνητικό αν σύναψη είναι απαγορευτική (δηλαδή αποτρέπει την νευρώνα να παράγει μια απόκριση).
2. Έναν αθροιστή για την πρόσθεση των σημάτων εισόδου, που παίρνουν βάρος από την αντίστοιχη σύναψη. Αυτές οι λειτουργίες αποτελούν το γραμμικό συνδυαστή.
3. Μια συνάρτηση ενεργοποίησης για τη μείωση του εύρους της εξόδου του νευρώνα. Συνήθως το κανονικοποιημένο εύρος της εξόδου ενός νευρώνα είναι το κλειστό σύνολο $[0,1]$ ή $[-1,1]$.

Σύμφωνα με τα παραπάνω, στο σχήμα 4.8 παρουσιάζεται το μοντέλο ενός τεχνητού νευρώνα.



Σχήμα 4.8: Μοντέλο τεχνητού νευρώνα.

Το μοντέλο επίσης περιλαμβάνει μία εξωτερικά εφαρμοζόμενη τιμή που έχει επίδραση στην ελάττωση της εισόδου στην εφαρμοζόμενη συνάρτηση ενεργοποίησης, που ακολουθεί. Στην βιβλιογραφία, η τιμή αυτή αναφέρεται ως πόλωση (bias). Μερικές φορές θεωρούμε ότι εκτός από τα εισερχόμενα σήματα και τα αντίστοιχα βάρη w , ο νευρώνας έχει και ένα εσωτερικό βάρος που τον χαρακτηρίζει (bias), το οποίο εκφράζει τη προδιάθεση του νευρώνα. Είναι τελείως ξεχωριστό από τα άλλα βάρη, παρ' όλα αυτά όμως δρα με τον ίδιο ακριβώς τρόπο. Θεωρείται ως ένα εξωτερικό ερέθισμα και αποτελεί μία παράμετρο η οποία ορίζεται κατά την εκπαίδευση του δικτύου.

Η παραπάνω διαδικασία περιγράφεται μαθηματικά από τις παρακάτω συναρτήσεις:

$$u_k = \sum_{j=1}^p w_{kj} \cdot x_j \quad (4.1)$$

$$y_k = \varphi(u_k - \theta_k) \quad (4.2)$$

,όπου:

x_j : τα σήματα εισόδου,

w_{kj} : τα συναπτικά βάρη του k ,

u_k : η γραμμική συνδυαστική έξοδος του νευρώνα,

θ_k : τιμή πόλωσης,

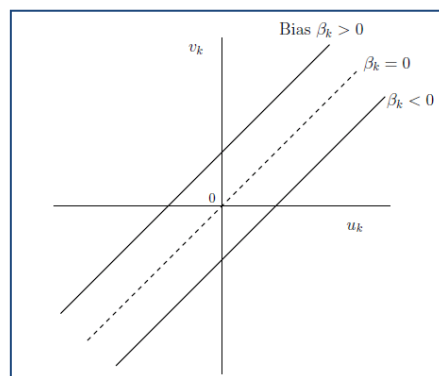
$\varphi(\cdot)$: η συνάρτηση ενεργοποίησης και

y_k : το σήμα εξόδου του νευρώνα, που αναφέρεται και σαν πραγματική έξοδος.

Η χρήση του θ_k έχει σαν αποτέλεσμα την εφαρμογή ενός εγγενούς (affine) μετασχηματισμού της τιμής u_k , όπως φαίνεται από την παρακάτω σχέση:

$$v_k = u_k - \theta_k \quad (4.3)$$

Η τιμή u_k , ανάλογα με το θ_k αν είναι θετικό ή αρνητικό μεταβάλλεται όπως φαίνεται στο σχήμα 4.9.



Σχήμα 4.9: Εγγενής μετασχηματισμός που παράγεται από την παρουσία κατωφλίου.

Η τιμή πόλωσης θ_k είναι εξωτερική παράμετρος του νευρώνα k . Έτσι μπορούμε να διαμορφώσουμε το συνδυασμό των εξισώσεων (4.1) και (4.2) ως εξής:

$$v_k = \sum_{j=0}^p w_{kj} x_j \quad (4.4)$$

και

$$y_k = \varphi(v_k) \quad (4.5)$$

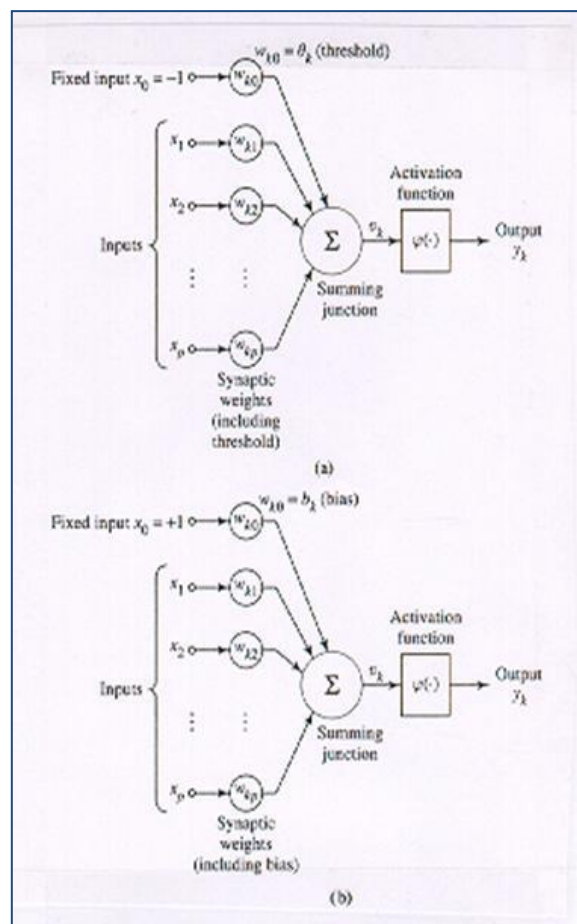
Στην εξίσωση (4.4) έχουμε προσθέσει μια νέα σύναψη, της οποίας η είσοδος είναι:

$$x_0 = -1 \quad (4.6)$$

και της οποίας το βάρος είναι:

$$w_{k0} = \theta_k \quad (4.7)$$

Έτσι μπορούμε να αναδιαμορφώσουμε το μοντέλο του νευρώνα k , όπως φαίνεται στο σχήμα 4.10.

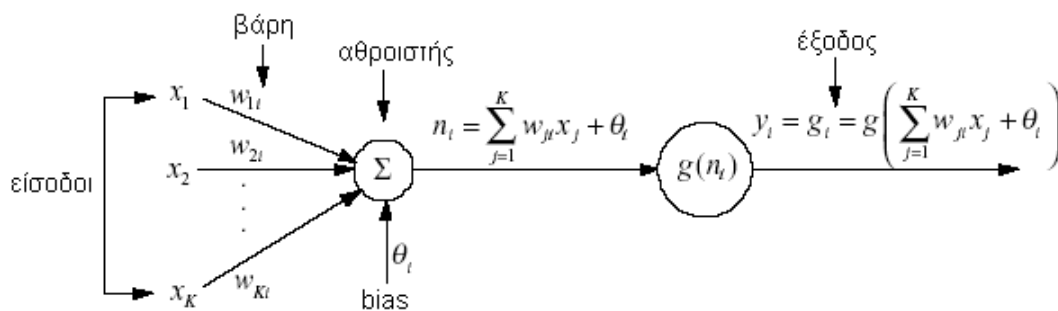


Σχήμα 4.10: Αναδιαμόρφωση μοντέλου νευρώνα.

Αξίζει να αναφέρουμε ότι το Τεχνητό Νευρωνικό δίκτυο μπορεί να οριστεί από οκτώ βασικά στοιχεία:

1. Το σύνολο των μονάδων επεξεργασίας.
2. Την κατάσταση ενεργοποίησης της κάθε μονάδας.
3. Την συνάρτηση εξόδου για κάθε μονάδα.
4. Την τοπολογία του δικτύου (τρόπος σύνδεσης μονάδων).
5. Έναν κανόνα διάδοσης (συνάρτηση συνδυασμού), ο οποίος ορίζει τον τρόπο με τον οποίο οι τιμές εξόδου από άλλες μονάδες, συνδυάζονται για να δημιουργηθεί μια μόνο τιμή, που είναι η τελική είσοδος της μονάδας.
6. Έναν κανόνα ενεργοποίησης για να αναπροσαρμόζει τις ενέργειες κάθε μονάδας χρησιμοποιώντας την τρέχουσα τιμή ενεργοποίησης και τις εισόδους που έρχονται από άλλες μονάδες.
7. Έναν κανόνα μάθησης για να τροποποιεί τα βάρη του δικτύου χρησιμοποιώντας πληροφορίες από το εξωτερικό περιβάλλον.
8. Το εξωτερικό περιβάλλον που δίνει πληροφορίες στο δίκτυο αλληλεπιδρώντας με αυτό.

Ένα παράδειγμα τεχνητού νευρώνα παρουσιάζεται στο σχήμα 4.11 που ακολουθεί.



Σχήμα 4.11: Απεικόνιση στοιχείων τεχνητού νευρώνα

4.4.2 Συναρτήσεις ενεργοποίησης νευρώνα

Η συνάρτηση ενεργοποίησης $\varphi(\cdot)$, ορίζει την έξοδο ενός νευρώνα συναρτήσει του επιπέδου ενεργοποίησης της εισόδου, (Haykin, 1994). Έχουμε 3 βασικούς τύπους συναρτήσεων ενεργοποίησης:

1. Συνάρτηση Κατωφλιού (Threshold function)

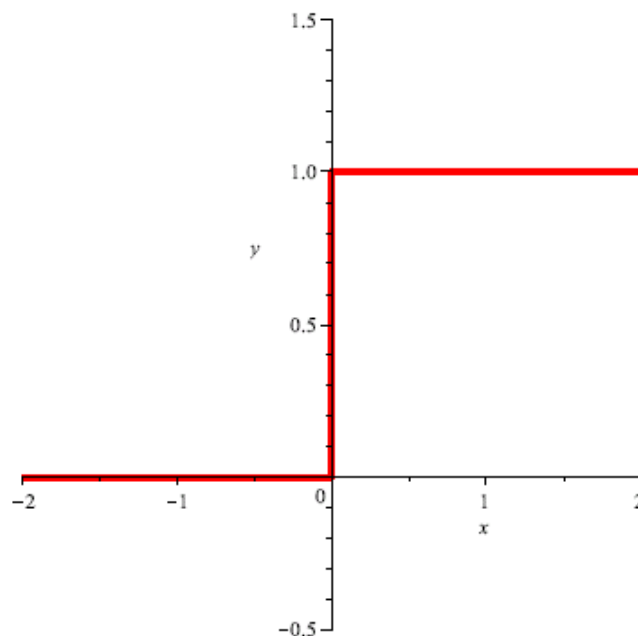
$$\varphi(v) = \begin{cases} 1, & v \geq 0 \\ 0, & v < 0 \end{cases} \quad (4.8)$$

Επομένως η έξοδος ενός νευρώνα k , που έχει την παραπάνω συνάρτηση ενεργοποίησης, έχει τη μορφή:

$$y_k = \begin{cases} 1, & v_k \geq 0 \\ 0, & v_k < 0 \end{cases} \quad (4.9)$$

, όπου v_k είναι το εσωτερικό επίπεδο ενεργοποίησης του νευρώνα και δίνεται από τη σχέση:

$$v_k = \sum_{j=1}^p w_{kj} x_j - \theta_k \quad (4.10)$$

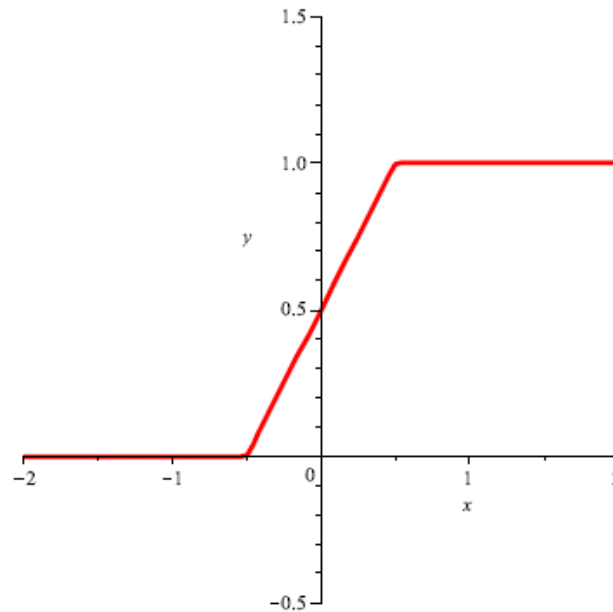


Σχήμα 4.12(α): Συνάρτηση κατωφλίου

Η συνάρτηση κατωφλίου δίνει τιμές που ανήκουν στο διάστημα $[0,1]$. Ένας νευρώνας που χρησιμοποιεί τέτοιου είδους συνάρτηση ενεργοποίησης ακολουθεί το μοντέλο των McCulloch-Pitts. Σε αυτό το μοντέλο η έξοδος του νευρώνα παίρνει την τιμή 1 αν η ολική δυνατότητα ενεργοποίησης του νευρώνα είναι μη αρνητική διαφορετικά παίρνει την τιμή 0.

2. Τμηματικά Γραμμική Συνάρτηση

$$\varphi(v) = \begin{cases} 1 & , & v \geq 1/2 \\ v & , & 1/2 > v > -1/2 \\ 0 & , & v \leq -1/2 \end{cases} \quad (4.11)$$



Σχήμα 4.12(b): Τμηματικά γραμμική συνάρτηση

Στη γραμμική συνάρτηση ξανά το πεδίο τιμών είναι $[0,1]$.

3. Σιγμοειδής

Είναι η πιο συνηθισμένη μορφή συνάρτησης ενεργοποίησης που χρησιμοποιείται στην κατασκευή τεχνητών νευρωνικών δικτύων. Ορίζεται σαν αυστηρά αύξουσα συνάρτηση, η οποία παρουσιάζει εξομάλυνση και ασυμπτωτικές ιδιότητες. Ένα παράδειγμα είναι η *λογιστική συνάρτηση*, που ορίζεται από τη σχέση:

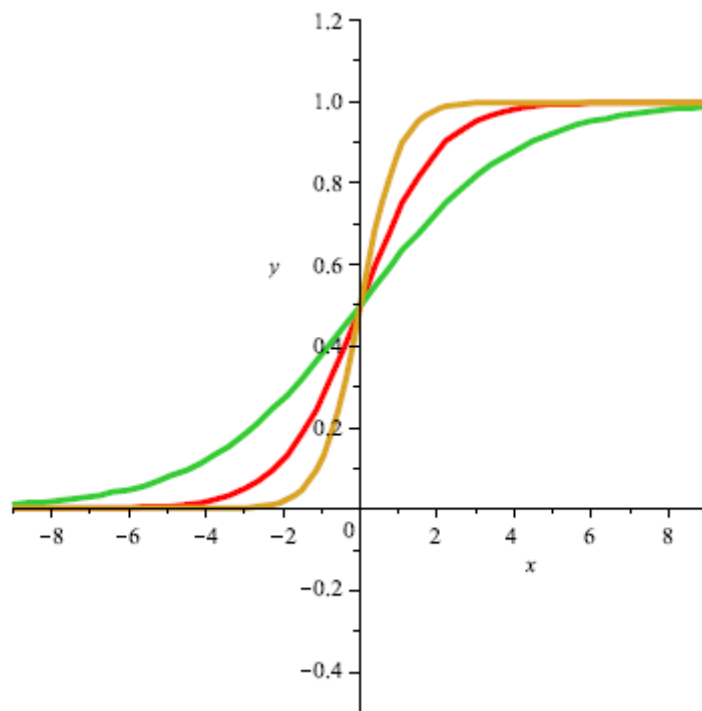
$$\varphi(v) = 1/[1 + \exp(-a \cdot v)] \quad (4.12)$$

Μεταβάλλοντας την παράμετρο a παίρνουμε διαφορετικές συναρτήσεις, όπως φαίνεται στο σχήμα 4.12c.

Οι προηγούμενες συναρτήσεις ενεργοποίησης κυμαίνονται από 0 ως +1. Αν θέλουμε συνάρτηση που να κυμαίνεται από -1 ως +1, επαναπροσδιορίζουμε, για παράδειγμα τη συνάρτηση κατωφλιού ως:

$$\varphi(v) = \begin{cases} 1, & v \geq 0 \\ 0, & v = 0 \\ -1, & v < 0 \end{cases} \quad (4.13)$$

, που ονομάζεται συνάρτηση *προσήμου* (*signum*) και συμβολίζεται σαν $\text{sgn}(\cdot)$.



Σχήμα 4.12(c): Σιγμοειδής συνάρτηση.

4.4.3 Τοπολογία - Αρχιτεκτονικές των Τεχνητών Νευρωνικών Δικτύων

Ο τρόπος με τον οποίο οι νευρώνες ενός νευρωνικού δικτύου συνδέονται μεταξύ τους, έτσι ώστε το δίκτυο που δημιουργείται να έχει τη δυνατότητα σωστής εκπαίδευσης, αποτελεί την αρχιτεκτονική του δικτύου και είναι άρρηκτα συνδεδεμένη με τον αλγόριθμο εκπαίδευσης. Όπως προαναφέρθηκε οι δομικές μονάδες του τεχνητού νευρωνικού δικτύου είναι οι νευρώνες, οι οποίοι συνδέονται μεταξύ τους και δημιουργούν στρώματα (layers), σχηματίζοντας το νευρωνικό δίκτυο. Σύμφωνα με το πλήθος των στρωμάτων διακρίνουμε τις εξής δύο κατηγορίες νευρωνικών δικτύων (Hecht-Nielsen, 1990):

- **Μονοστρωματικά Τεχνητά Νευρωνικά Δίκτυα (Single-layer Artificial Neural Networks)**
- **Πολυστρωματικά Τεχνητά Νευρωνικά Δίκτυα (Multi-layer Artificial Neural Networks)**

Στην κατηγορία των Μονοστρωματικών υπάγονται τα δίκτυα με ένα στρώμα νευρώνων ή ακόμα και τα δίκτυα που αποτελούνται από έναν μοναδικό νευρώνα. Όταν το δίκτυο αποτελείται από παραπάνω από ένα στρώματα νευρώνων, τότε ανήκει στην κατηγορία των Πολυστρωματικών Τεχνητών Νευρωνικών Δικτύων. Το δίκτυο διαμορφώνεται μέσω της ιδιότητας ότι τα στοιχεία εξόδου κάθε προηγούμενου επιπέδου νευρώνων, αποτελούν τα στοιχεία εισόδου του επόμενου επιπέδου. Βασική εξαίρεση αποτελεί το πρώτο στρώμα, καθώς τα αρχικά στοιχεία εισόδου δίνονται από τον χρήστη. Τα επίπεδα που διαμορφώνουν το δίκτυο είναι δυνατόν να διαφέρουν όχι μόνο

στον αριθμό των νευρώνων που τα απαρτίζουν αλλά και στην συνάρτηση ενεργοποίησης που χρησιμοποιούν. Οι νευρώνες όμως του ίδιου στρώματος απαιτείται να έχουν την ίδια συνάρτηση ενεργοποίησης.

4.4.4 Μέθοδοι Εκπαίδευσης

Η εκπαίδευση των νευρώνων ενός δικτύου στοχεύει στην μεταβολή της αποδοτικότητας των συνάψεων, με αποτέλεσμα την αλλαγή της επίδρασης κάθε προσυναπτικού νευρώνα σε κάποιον μετασυναπτικό.

Οι μέθοδοι εκπαίδευσης μπορούν να διαχωριστούν σε τρεις διακριτούς τύπους:

1. Επιβλεπόμενη ή συσχετιζόμενη εκπαίδευση (supervised/associative learning).

Η εκπαίδευση του τεχνητού νευρωνικού δικτύου γίνεται με συγκεκριμένα ζεύγη εισόδων εξόδων, τα οποία εισάγονται από τον χρήστη ή από το σύστημα που περιέχει το Τεχνητό Νευρωνικό Δίκτυο (αυτοεπιβλεπόμενο δίκτυο). Χρησιμοποιείται κυρίως ο κανόνας Widrow – Hoff, δηλαδή πραγματοποιείται υπολογισμός του σφάλματος και προσαρμογή των βαρών, ενώ η εκτίμηση του σφάλματος γίνεται μέσω του μέσου τετραγωνικού σφάλματος.

2. Μη επιβλεπόμενη εκπαίδευση ή αυτοοργάνωση (unsupervised learning/self-organization).

Στην περίπτωση της εκπαίδευσης χωρίς επίβλεψη δίνεται στο σύστημα πλήθος προτύπων (δεδομένων εισόδου), χωρίς να δίνεται η επιθυμητή έξοδος για καθένα από αυτά. Επομένως, το εξωτερικό περιβάλλον δεν παρέχει την επιθυμητή έξοδο του δικτύου, ούτε την πληροφορία για το εάν αυτή είναι καλή ή κακή. Αντίθετα από την επιβλεπόμενη εκμάθηση, εδώ δεν υπάρχουν εκ των προτέρων καθορισμένα σύνολα κατηγοριών στα οποία θα ταξινομηθούν τα πρότυπα. Το σύστημα πρέπει να ομαδοποιήσει τα δεδομένα εισόδου σύμφωνα με κάποια κριτήρια ομοιότητας και να ανακαλύψει στατιστικά αξιοπρόσεκτα χαρακτηριστικά των διανυσμάτων εισόδου. Άρα, χρησιμοποιώντας τη συσχέτιση (correlation) του διανύσματος εισόδου, ο κανόνας μάθησης αλλάζει τα βάρη του δικτύου με σκοπό να ομαδοποιήσει το διάνυσμα εισόδου σε ομάδες (clusters), έτσι ώστε παρόμοια διανύσματα εισόδου να παράγουν παρόμοιες εξόδους του δικτύου, εφόσον ανήκουν στην ίδια ομάδα. Έτσι, η διαδικασία εκπαίδευσης βγάζει τις στατιστικές ιδιότητες του σετ εκπαίδευσης και ομαδοποιεί τα παρόμοια διανύσματα. Εφαρμόζοντας ένα διάνυσμα από δοσμένη ομάδα στην είσοδο, θα παράγει συγκεκριμένο διάνυσμα εξόδου, αλλά δεν υπάρχει κανένας τρόπος ώστε να καθοριστεί πριν από την εκπαίδευση ποιο συγκεκριμένο πρότυπο εξόδου θα παραχθεί από δοσμένη ομάδα διανύσματος εισόδου. Για αυτό οι έξοδοι ενός τέτοιου δικτύου πρέπει, γενικά, να μετατραπούν σε μια κατανοητή μορφή που είναι επακόλουθο της διαδικασίας εκπαίδευσης. Η μέθοδος αυτή μάθησης λέγεται επίσης “αυτό-οργάνωση” (self

organization). Η συγκεκριμένη μέθοδος εκπαίδευσης χρησιμοποιεί τον κανόνα του Hebb ή τον κανόνα εκμάθησης βάσει ανταγωνισμού.

3. Εκπαίδευση δύο σταδίων (semi-supervised learning).

Πρόκειται για συνδυασμό των δύο μεθόδων που προαναφέρθηκαν και εκτελείται όπως είναι ευνόητο σε δύο στάδια. Στο πρώτο στάδιο, επιτυγχάνεται ομαδοποίηση των δεδομένων εισόδου σύμφωνα με κάποια κριτήρια ομοιότητας, ενώ στο δεύτερο στάδιο το δίκτυο εκπαιδεύεται με τη μέθοδο της επίβλεψης, δημιουργώντας μία συνάρτηση απεικόνισης εισόδου – εξόδου, (Haykin, 1994).

4.4.5 Κανόνες Εκμάθησης

Η μάθηση ενός τεχνητού νευρωνικού δικτύου πραγματοποιείται με τη βοήθεια αλγορίθμων που είναι γνωστοί ως κανόνες εκμάθησης ή αλγόριθμοι εκπαίδευσης. Ουσιαστικά, η μάθηση επιτυγχάνεται μέσω μιας επαναληπτικής διαδικασίας ρυθμίσεων της τιμής των συναπτικών βαρών και των κατωφλίων. Η επαναληπτική διαδικασία που τροποποιεί τα βάρη σε κάθε βήμα, ονομάζεται Κανόνας Εκμάθησης (learning rule) και αποδίδεται μέσω της ακόλουθης εξίσωσης:

$$w_{ij}(t_{n+1}) = w_{ij}(t_n) + \Delta w_{ij}(t_n) \quad (4.14)$$

Η παραπάνω εξίσωση μεταφράζεται ως εξής: Μία ποσότητα $\Delta w_{ij}(t_n)$ προστίθεται στα βάρη τη χρονική στιγμή t_n , για να καθορίσει την νέα τιμή των συναπτικών βαρών την επόμενη χρονική στιγμή t_{n+1} . Στη συνέχεια αναφέρονται οι συνηθέστερες προσεγγίσεις για τον υπολογισμό της ποσότητας Δw_{ij} , που διαφοροποιούν και τους κανόνες εκμάθησης.

1. Κανόνας του Hebb

Η φυσική παρατήρηση στην οποία στηρίχτηκε είναι ότι: «όταν ο άξονας ενός νευρώνα Α είναι αρκετά κοντά ώστε να διεγείρει ένα νευρώνα Β και συμμετέχει είτε διαρκώς είτε περιστασιακά στην ενεργοποίησή του, τότε παρατηρείται κάποια διαδικασία ανάπτυξης ή μεταβολικής αλλαγής είτε στον ένα είτε και στους δυο νευρώνες, τέτοια ώστε η αποτελεσματικότητα με την οποία Α διεγείρει τον Β, αυξάνεται» (Hebb, 1949).

Ωστόσο τα χεμπιανά μοντέλα εκμάθησης βασίζονται και σε μία εναλλακτική διατύπωση του κανόνα Hebb: «Αν ο προσυναπτικός και ο μετασυναπτικός νευρώνας ενεργοποιούνται ταυτόχρονα, τότε η σύναψη τους επιλέγεται προς ενίσχυση. Αντίθετα, αν οι νευρώνες ενεργοποιούνται σε διαφορετικές χρονικές στιγμές, η σύναψη επιλέγεται προς αποδυνάμωση ή καταστροφή».

Ο νόμος του Hebb μπορεί να εκφραστεί μαθηματικά ως εξής:

$$\Delta w_{ij}(t_n) = \eta \cdot y_i(t_n) \cdot x_j(t_n) \quad (4.15)$$

,όπου:

$y_i(t_n)$: σήμα εξόδου νευρώνα i , την χρονική στιγμή t_n , (μετά-συναπτικό σήμα)

$x_j(t_n)$: σήμα εισόδου στην j -οστή σύναψη του νευρώνα i , την χρονική στιγμή t_n (προσυναπτικό σήμα)

$\eta > 0$: ρυθμός μάθησης (ελέγχει το μέγεθος κάθε βήματος μάθησης),

Άρα η ποσότητα Δw_{ij} υπολογίζεται συναρτήσει των προσυναπτικών και μετασυναπτικών σημάτων, δηλαδή των σημάτων εισόδου και εξόδου αντίστοιχα.

Η παραπάνω σχέση εκφράζεται και ως εξής (Haykin, 1994):

$$\Delta w_{ij}(t_n) = F(y_i(t_n), x_j(t_n)) \quad (4.16)$$

, όπου $y_i(t_n)$: σήμα εξόδου νευρώνα i , την χρονική στιγμή t_n ,

$x_j(t_n)$: σήμα εισόδου νευρώνα j , την χρονική στιγμή t_n .

2. Κανόνας δέλτα ή Κανόνας Widrow – Hoff – Αλγόριθμος των ελαχίστων μέσων τετραγώνων (1960)

Ο κανόνας Δέλτα εκτός από κανόνας εκπαίδευσης είναι και κανόνας διόρθωσης σφάλματος, είναι δηλαδή ένας εποπτευόμενος κανόνας μάθησης. Ο κανόνας Δέλτα αλλάζει τα βάρη του δικτύου προς την κατεύθυνση που μειώνει το τετραγωνικό σφάλμα εξόδου, για κάθε παρουσίαση του προτύπου εισόδου/επιθυμητής εξόδου από το σετ εκπαίδευσης. Αυτό έχει σαν αποτέλεσμα η ταχύτητα μάθησης να είναι μεγάλη.

$$\text{Το τετραγωνικό σφάλμα εξόδου είναι: } \varepsilon(t_n) = \frac{1}{2} [e_i(t_n)]^2 \quad (4.17)$$

$$\text{,όπου: } e_i(t_n) = d_i(t_n) - y_i(t_n) \quad (4.18)$$

με : $e_i(t_n)$: σήμα λάθους στην έξοδο του νευρώνα i , την χρονική στιγμή t_n ,

$d_i(t_n)$: επιθυμητή έξοδο του νευρώνα i , την χρονική στιγμή t_n ,

$y_i(t_n)$: πραγματική έξοδο του νευρώνα i , την χρονική στιγμή t_n .

Σύμφωνα με τον κανόνα δέλτα η ποσότητα $\Delta w_{ij}(t_n)$ δίνεται από την σχέση:

$$\Delta w_{ij}(t_n) = \eta \cdot e_i(t_n) \cdot x_j(t_n) \quad (4.19)$$

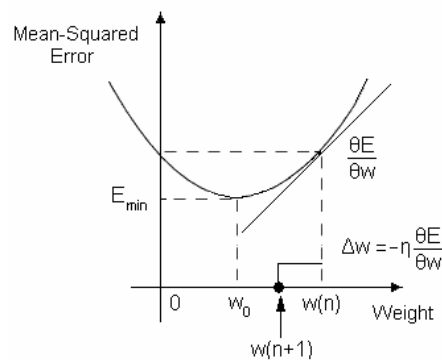
,όπου:

η : σταθερά αναλογικότητας που αναπαριστά τον ρυθμό εκμάθησης,

$e_i(t_n)$: σήμα λάθους στην έξοδο του νευρώνα i , την χρονική στιγμή t_n ,

$x_j(t_n)$: σήμα εισόδου νευρώνα j , την χρονική στιγμή t_n .

Ο κανόνας Δέλτα θεωρούμε ότι είναι μια διαδικασία ανίχνευσης κατηφορικής κλίσης (gradient-descent search procedure) η οποία εκτελείται σε κάθε επανάληψη. Σύμφωνα με αυτήν, η αλλαγή του βάρους γίνεται σε κατεύθυνση αντίθετη με αυτήν που έχει η παράγωγος (ως προς το βάρος) της επιφάνειας του σφάλματος, έχοντας σαν αποτέλεσμα την απότομη κατάβαση σφάλματος προς το ελάχιστο σημείο του όπως φαίνεται και στο παρακάτω σχήμα.



Σχήμα 4.14: Κανόνας Δέλτα

Αν επαναλάβουμε την διαδικασία αυτή σε όλο το σετ των προτύπων εκπαίδευσης, καταλήγουμε στην ελαχιστοποίηση του "μέσου" σφάλματος εξόδου.

Συμπερασματικά, η βασική ομοιότητα μεταξύ των παραπάνω δύο κανόνων, εντοπίζεται στο γεγονός ότι και οι δύο περιέχουν τον παράγοντα $\eta \cdot x_j(t_n)$. Ωστόσο ο κανόνας Δέλτα πολλαπλασιάζει το $\eta \cdot x_j(t_n)$ με την τιμή σφάλματος $e_i(t_n)$, ενώ ο κανόνας Hebb το πολλαπλασιάζει με την τιμή εξόδου $y_i(t_n)$. Επίσης θα μπορούσαμε να συμπληρώσουμε ότι ο κανόνας δέλτα χρειάζεται την γνώση της επιθυμητής εξόδου ενώ κάτι παρόμοιο δεν απαιτείται στον κανόνα Hebb.

4.5 Ταξινόμηση Τεχνικών Νευρωνικών Δικτύων

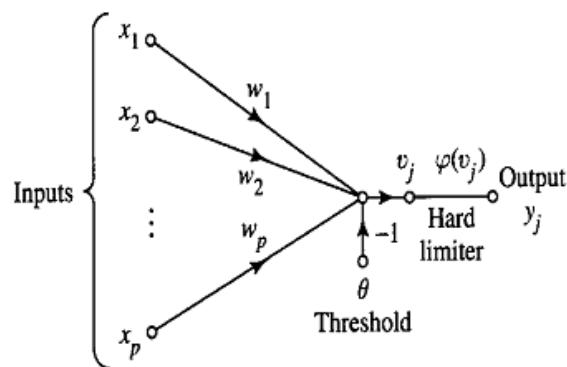
Ανάλογα τον τύπο του νευρώνα, την αρχιτεκτονική του δικτύου και τη μέθοδο κπαίδευσης, τα νευρωνικά μοντέλα ταξινομούνται σε κατηγορίες, όπως φαίνεται και από τον παρακάτω πίνακα(Διαμαντάρας, 2006).



Σχήμα 4.15: Ταξινόμηση Τεχνικών Νευρωνικών Δικτύων,(Πηγή: Διαμαντάρας, (2006)).

4.6 Το μοντέλο του αισθητήρα – (Perceptron)

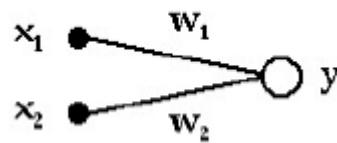
Το μοντέλο του αισθητήρα (Perceptron) είναι από τα πρώτα μοντέλα νευρωνικών δικτύων που αναπτύχθηκαν. Σήμερα, υπάρχουν πολλές παραλλαγές νευρωνικών δικτύων που βασίζονται στον αισθητήρα και έχουν διαφορετικές δομές, άλλες απλές και άλλες πιο περίπλοκες. Η πιο απλή μορφή είναι ο λεγόμενος στοιχειώδης αισθητήρας (elementary perceptron), γιατί αποτελείται από ένα μόνο νευρώνα και είναι το πιο απλό, αυτοδύναμο σύστημα που υπάρχει και επιτελεί μία ορισμένη διεργασία (Rosenblatt, 1958). Ένα παράδειγμα του μοντέλου του αισθητήρα φαίνεται στο παρακάτω σχήμα 4.16:



Σχήμα 4.16: Το διάγραμμα ροής σήματος του Perceptron.

Ο στόχος του απλού αισθητήρα είναι να μάθει να λύνει προβλήματα ταξινόμησης, να αντιστοιχεί δηλαδή κάθε σει εισόδων που δέχεται, στη σωστή κλάση. Όμως τα προβλήματα ταξινόμησης που μπορεί να λύσει ο αισθητήρας είναι εκείνα τα οποία είναι *γραμμικά διαχωρίσιμα* και μόνο. Το γεγονός αυτό αποδεικνύεται παρακάτω.

Για να γίνει κατανοητή η λειτουργία της ταξινόμησης, έστω αισθητήρας με δύο εισόδους και έναν νευρώνα εξόδου, όπως φαίνεται στο ακόλουθο σχήμα:



Σχήμα 4.17: Μοντέλο Αισθητήρα Perceptron

Έστω λοιπόν ότι ο στόχος του αισθητήρα είναι να δίνει σωστά αποτελέσματα σύμφωνα με τη λογική πράξη XOR*, (βλέπε σημείωση):

x_1	x_2	$x_1 \text{ XOR } x_2$
0	0	0
0	1	1
1	0	1
1	1	0

Σχήμα 4.18: Η πράξη XOR

Τότε έξοδος του παραπάνω αισθητήρα δίνεται από τις σχέσεις:

$$y = \begin{cases} 0, & \text{αν } w_1 \cdot x_1 + w_2 \cdot x_2 \leq \theta \\ 1, & \text{αν } w_1 \cdot x_1 + w_2 \cdot x_2 \geq \theta \end{cases} \quad (4.20)$$

,όπου : w_1, w_2 : τα συναπτικά βάρη και

θ : το κατώφλι

Τα τέσσερα σετ εισόδων του αισθητήρα αντιστοιχούν σε τέσσερα σημεία στο επίπεδο. Στην συγκεκριμένη περίπτωση τα σημεία (0,0) και (1,1) ανήκουν στην κλάση 0, ενώ τα (0,1) και (1,0) στην κλάση 1.

Η εξίσωση: $w_1 \cdot x_1 + w_2 \cdot x_2 = \theta$, είναι μία ευθεία που χωρίζει το επίπεδο (x, y) σε δύο ημιεπίπεδα. Αρκεί λοιπόν να βρεθεί ένα ζεύγος τιμών για τα w_1 και w_2 , τέτοιες ώστε να χωρίζουν τις δύο αυτές κλάσεις. Το γεγονός όμως αυτό δεν είναι εφικτό. Δεν υπάρχει καμία ευθεία τέτοια ώστε τα σημεία των δύο κλάσεων να ανήκουν σε διαφορετικά ημιεπίπεδα. Για να το πετύχουμε αυτό χρειαζόμαστε δύο ευθείες. Το πρόβλημα αυτό επομένως δεν είναι γραμμικά διαχωρίσιμο και άρα δεν είναι επιλύσιμο από τον αισθητήρα (Minsky M.L. and Papert S.A., 1969).

Η αδυναμία του αισθητήρα να επιλύσει το πρόβλημα της XOR, δεν είναι αμελητέα, καθώς πολλά προβλήματα βασίζονται στην XOR. Παρόλα αυτά, προσθέτοντας απλά ένα ή περισσότερα επίπεδα νευρώνων μεταξύ του επιπέδου εισόδου και αυτό της εξόδου, ο τροποποιημένος αισθητήρας που προκύπτει μπορεί πλέον να επιλύσει και μη γραμμικά διαχωριζόμενα προβλήματα.

Όσον αφορά τη διαδικασία εκπαίδευσης του perceptron, είναι απαραίτητα τόσο τα πρότυπα όσο και οι στόχοι, καθώς ακολουθείται η μέθοδος εκπαίδευσης με επίβλεψη. Το σήμα εξόδου y_k του νευρώνα k, συγκρίνεται με την επιθυμητή τιμή του στόχου d_k με αποτέλεσμα να επιτυγχάνεται διόρθωση των συνάψεων με τον κανόνα Δέλτα:

$$w_k(t_n + 1) = w_k(t_n) + \eta \cdot e_k(t_n) \cdot x_k(t_n) \quad (4.21)$$

, όπου:

$w_k(t_n + 1)$: τιμή συναπτικού βάρους του μοναδικού νευρώνα k, την χρονική στιγμή $t_n + 1$

$w_k(t_n)$: τιμή συναπτικού βάρους του μοναδικού νευρώνα k, την χρονική στιγμή t_n

η : σταθερά αναλογικότητας που αναπαριστά τον ρυθμό εκμάθησης,

$e_k(t_n)$: σήμα λάθους στην έξοδο του νευρώνα k, την χρονική στιγμή t_n ,

$x_k(t_n)$: σήμα εισόδου νευρώνα k, την χρονική στιγμή t_n .

Η εκπαίδευση λαμβάνει τέλος όταν δεν πραγματοποιείται πλέον καμία διόρθωση στις συνάψεις. Όταν δηλαδή ο στόχος ταυτίζεται με την έξοδο.

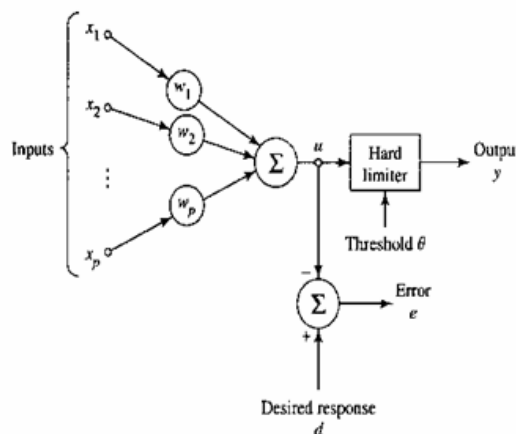
***Σημείωση:**

Μία λογική πύλη είναι ηλεκτρονικό κύκλωμα το οποίο πραγματοποιεί μία λογική πράξη στις εισόδους της και παράγει μία ή περισσότερες εξόδους. Οι λογικές πύλες έχουν δημιουργηθεί για να δουλεύουν στο δυαδικό σύστημα. Η πύλη XOR εκτελεί την λογική πράξη XOR μεταξύ των εισόδων της. Η πράξη XOR στην άλγεβρα συμβολίζεται με ένα συν μέσα σε ένα κύκλο ($\oplus$). Για παράδειγμα εάν η πύλη έχει 2 εισόδους (a και b) και μία έξοδο (c) θα γίνει η πράξη: $c=ab'+a'b$

4.7 Το δίκτυο ADALINE (Adaptive Linear Element)

Το Adaline (Adaptive linear element - γραμμικό προσαρμοζόμενο στοιχείο), αρχικά χρησιμοποιήθηκε από τους Widrow και Hoff και αποτελεί μία προσαρμοζόμενη μηχανή ταξινόμησης προτύπων.

Ένα παράδειγμα του Adaline φαίνεται στο σχήμα 4.19. Αποτελείται από έναν νευρώνα τύπου McCulloch–Pitts (ο οποίος θα περιγραφεί αναλυτικά σε επόμενη παράγραφο), με p εισόδους και μία έξοδο. Οι εισοδοί $x_1, x_2, \dots, x_p$ παίρνουν την τιμή ± 1 . Μια μεταβλητή κατωφλίου θ ($\theta \in [0, 1]$) εφαρμόζεται στη συσκευή κατωφλίου. Η επιθυμητή έξοδος d παίρνει επίσης τιμές ± 1 .



Σχήμα 4.19: Το λειτουργικό διάγραμμα του Adaline.

Η έξοδος y του Adaline αποδίδεται ως:

$$y = \begin{cases} +1, & u \geq \theta \\ -1, & u < \theta \end{cases} \quad (4.22)$$

Ο στόχος της προσαρμοζόμενης διαδικασίας στο Adaline είναι ο εξής: «Δεδομένου ενός συνόλου προτύπων εισόδου και των σχετικών επιθυμητών εξόδων, να βρεθεί το βέλτιστο σύνολο των συναπτικών βαρών και του κατωφλίου θ , έτσι ώστε να ελαχιστοποιηθεί το μέσο τετραγωνικό σφάλμα ανάμεσα στη διέγερση u και στον στόχο d » (Artificial Intelligence Technologies Tutorial).

$$, \text{ όπου: Μέσο τετραγωνικό σφάλμα: } J_{MSE} = E\{(u - d)^2\} = \frac{1}{p} \sum (u_k(t_n) - d_k(t_n))^2 \quad (4.23)$$

με: E: σύμβολο μέσης τιμής,

$u_k(t_n) = \sum w_k(t_n) \cdot x_k(t_n)$: τιμή διέγερσης του μοναδικού νευρώνα k, την χρονική στιγμή t_n ,

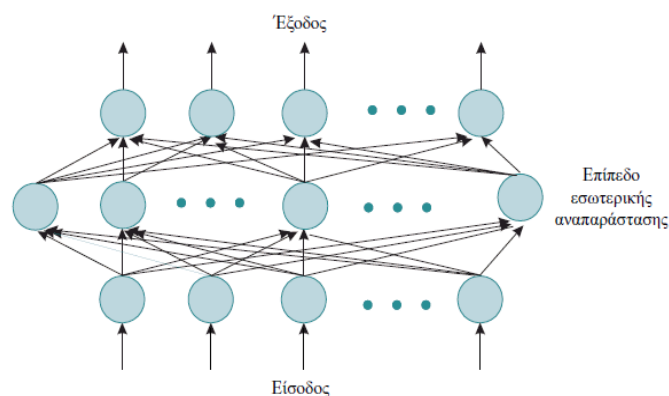
$w_k(t_n)$: τιμή στόχου του μοναδικού νευρώνα k, την χρονική στιγμή t_n ,

$x_k(t_n)$: τιμή εισόδου του μοναδικού νευρώνα k, την χρονική στιγμή t_n και

$d_k(t_n)$: τιμή στόχου του μοναδικού νευρώνα k, την χρονική στιγμή t_n

4.8 Δίκτυο Multi-Layer Perceptron - Κανόνας Backpropagation

Σε αυτή την ενότητα θα μελετήσουμε μία σπουδαία κλάση νευρωνικών δικτύων, τα δίκτυα εμπρός τροφοδότησης πολλών επιπέδων. Τυπικά ένα τέτοιο δίκτυο αποτελείται από ένα σύνολο αισθητήρων, που αποτελούν το επίπεδο εισόδου, ένα ή περισσότερα κρυφά επίπεδα (hidden layers) υπολογιστικών κόμβων και ένα επίπεδο υπολογιστικών κόμβων εξόδου. Το σήμα εισόδου διαδίδεται μέσα στο δίκτυο σε μία προς τα εμπρός κατεύθυνση, από επίπεδο σε επίπεδο. Αυτά τα νευρωνικά δίκτυα αναφέρονται σαν Perceptrons πολλών επιπέδων (Multi Layer Perceptrons- MLPs) τα οποία είναι μια γενίκευση του απλού Perceptron. Ένα παράδειγμα νευρωνικού δικτύου με πολλά επίπεδα παρουσιάζεται στο σχήμα 4.20.



Σχήμα 4.20: Δίκτυο με πολλαπλά επίπεδα

Ως συντομογραφία ενός πολυεπιπέδου νευρωνικού δικτύου συχνά χρησιμοποιείται ο εξής: $p-m1-m2-...-mn-n$, όπου p είναι ο αριθμός των εισόδων, n είναι ο αριθμός των εξόδων, m ο αριθμός των κρυμμένων επιπέδων με $m1$ κόμβους το πρώτο, $m2$ κόμβους το δεύτερο, ... και mn το τελευταίο. Πρώτα, υπάρχει ένα επίπεδο εισόδου το οποίο αποτελείται από μία ομάδα νευρώνων οι

οποίοι δεν κάνουν ουσιαστικά τίποτα άλλο παρά να δέχονται το σήμα εισόδου. Κατόπιν υπάρχει ένας αριθμός εσωτερικών επιπέδων, καθένα από τα οποία έχει έναν αριθμό νευρώνων, και τα οποία δέχονται το σήμα από το επίπεδο εισόδου, το επεξεργάζονται και κατόπιν το προωθούν προς την έξοδο. Στο παραπάνω σχήμα υπάρχει ένα μόνο κρυμμένο επίπεδο, αλλά θα μπορούσε να ήταν δύο, τρία ή οποιοσδήποτε άλλος αριθμός επιπέδων. Τέλος, υπάρχει ένα επίπεδο εξόδου που έχει επίσης έναν αριθμό νευρώνων, οι οποίοι δέχονται σήμα από τα εσωτερικά επίπεδα και το προωθούν προς την έξοδο του δικτύου. Γενικά, δεν υπάρχει κανόνας ως προς τον αριθμό τόσο των εσωτερικών επιπέδων όσο και ως προς τον αριθμό των νευρώνων που περιλαμβάνει κάθε επίπεδο (εισόδου, εξόδου ή εσωτερικό). Η απάντηση σ' αυτό είναι διαφορετική σε κάθε πρόβλημα.

Τα MLPs έχουν εφαρμοστεί με επιτυχία στην επίλυση δύσκολων και ποικίλων προβλημάτων, εκπαιδεύοντας τα με έναν επιβλεπόμενο τρόπο (supervised manner), με ένα πολύ δημοφιλή αλγόριθμο γνωστό σαν αλγόριθμο πίσω διάδοσης του λάθους (error Back Propagation algorithm - BP). Αυτός ο αλγόριθμος βασίζεται στον κανόνα μάθησης διόρθωσης του λάθους (error correction learning rule).

Βασικά η διαδικασία της πίσω διάδοσης του λάθους αποτελείται από δυο περάσματα διαμέσου των διαφορετικών επιπέδων του δικτύου ένα προς τα εμπρός πέρασμα (forward pass) και ένα προς τα πίσω πέρασμα (backward pass) (Haykin, 2004).

- Στο εμπρός πέρασμα ένα διάνυσμα εισόδου (input vector) εφαρμόζεται στους νευρώνες εισόδου του δικτύου, και η επίδραση του διαδίδεται μέσα στο δίκτυο από επίπεδο σε επίπεδο (layer by layer). Τελικά ένα σύνολο από εξόδους παράγεται ως η πραγματική απόκριση του δικτύου. Κατά τη διάρκεια του εμπρός περάσματος τα βάρη του δικτύου είναι σταθερά.
- Από την άλλη μεριά κατά τη διάρκεια της πίσω διάδοσης η πραγματική απόκριση του δικτύου αφαιρείται από την επιθυμητή απόκριση για την παραγωγή ενός σήματος λάθους, το οποίο διαδίδεται προς τα πίσω στο δίκτυο, αντίθετα από την κατεύθυνση των συνδέσεων, από όπου προκύπτει και το όνομα πίσω διάδοσης του λάθους. Τα βάρη προσαρμόζονται σε συμφωνία με τον κανόνα διόρθωσης λάθους.

Τα συναπτικά βάρη προσαρμόζονται έτσι ώστε να κάνουν την πραγματική απόκριση του δικτύου να πλησιάσει την επιθυμητή απόκριση.

Στην βιβλιογραφία ο αλγόριθμος πίσω διάδοσης του λάθους συχνά αναφέρεται και σαν αλγόριθμος πίσω διάδοσης (Back Propagation Algorithm) ή πιο απλά σαν Back Prop.. Η διαδικασία μάθησης που εκτελείται με αυτόν τον αλγόριθμο ονομάζεται μάθηση πίσω διάδοσης.

4.8.1 Παραγωγή του αλγορίθμου Πίσω Διάδοσης (Back Propagation)

Το σήμα λάθους στην έξοδο του νευρώνα j στην επανάληψη n ορίζεται από την σχέση:

$$e_j = d_j - y_j \quad (4.24)$$

,όπου: ο νευρώνας j είναι κόμβος εξόδου.

Ορίζουμε την στιγμιαία τιμή του τετραγωνικού λάθους για τον νευρώνα j ως:

$$E_j(n) = 1/2 \cdot e_j^2(n) \quad (4.25)$$

Έτσι το στιγμιαίο άθροισμα των τετραγωνικών λαθών του δικτύου γράφεται ως εξής:

$$E(n) = \frac{1}{2} \sum_{j \in C}^N e_j^2(n) \quad (4.26)$$

,όπου το σύνολο C περιλαμβάνει όλους τους νευρώνες του επιπέδου εξόδου του δικτύου. Έστω N ο συνολικός αριθμός διανυσμάτων στο σύνολο εκπαίδευσης. Το μέσο τετραγωνικό λάθος για όλο το σύνολο εκπαίδευσης είναι:

$$E_{av} = \frac{1}{N} \sum_{n=1}^N E(n) \quad (4.27)$$

Προφανώς, τόσο η τιμή του E όσο και του E_{av} εξαρτώνται από τις ελεύθερες παραμέτρους του δικτύου, δηλαδή από τα συναπτικά βάρη (synaptic weights) των συνδέσεων μεταξύ των νευρώνων, όπως και από τις τιμές των πολώσεων των νευρώνων (bias). Ο σκοπός της διαδικασίας εκπαίδευσης είναι η ελαχιστοποίηση της μέσης τιμής των σφαλμάτων E_{av} με την κατάλληλη προσαρμογή των ελεύθερων παραμέτρων του δικτύου. Για να συμβεί αυτό χρησιμοποιείται μια συγκεκριμένη λογική κατά την οποία τα βάρη ενημερώνονται μετά από το πέρασμα κάθε ζεύγους προτύπου $[x, t]$, σύμφωνα με τα σφάλματα που παρουσιάζονται κατά την εφαρμογή καθενός ζεύγους προτύπου στο δίκτυο. Με άλλα λόγια, η μέση τιμή των μεταβολών όλων των ελευθέρων παραμέτρων του δικτύου είναι μια προσεκτική εκτίμηση της πραγματικής μεταβολής των ελεύθερων παραμέτρων που προκύπτει από την διαδικασία ελαχιστοποίησης του E_{av} (Haykin, 1994).

Για τον νευρώνα εξόδου j έχουμε:

$$v_j(n) = \sum_{i=1}^p w_{ji}(n) y_i(n) \quad (4.28)$$

, όπου:

p : είναι το σύνολο των νευρώνων του προηγούμενου επιπέδου

$w_{ji}(n)$: είναι τα συναπτικά βάρη μεταξύ των νευρώνων j και i

$y_i(n)$: είναι τα σήματα εισόδου των νευρώνων του προηγούμενου επιπέδου

Πρέπει να σημειωθεί ότι $y_0(n) = -1$ είναι η τιμή της εξωτερικής εισόδου που αντιστοιχεί στην τιμή της πόλωσης που αντιπροσωπεύεται από το βάρος $w_{j0}(n)$.

Επίσης, θέτουμε $w_{j0} = \theta_j$.

Επομένως το λειτουργικό σήμα $y_j(n)$ στην έξοδο του νευρώνα j θα δίνεται από την σχέση:

$$y_j(n) = \varphi_j(v_j(n)) \quad (4.29)$$

Ο αλγόριθμος πίσω διάδοσης εφαρμόζει μια διόρθωση $\Delta w_{ji}(n)$ στο συναπτικό βάρος $w_{ji}(n)$, η οποία είναι ανάλογη της στιγμιαίας κλίσης $\partial E(n)/\partial w_{ji}(n)$. Σύμφωνα με τον αλυσιδωτό κανόνα (Chain Rule) μπορούμε να εκφράσουμε την κλίση ως εξής:

$$\frac{\partial E(n)}{\partial w_{ji}(n)} = \frac{\partial E(n)}{\partial e_j(n)} \cdot \frac{\partial e_j(n)}{\partial y_j(n)} \cdot \frac{\partial y_j(n)}{\partial v_j(n)} \cdot \frac{\partial v_j(n)}{\partial w_{ji}(n)} \quad (4.30)$$

Η κλίση $\frac{\partial E(n)}{\partial w_{ji}(n)}$ αντιπροσωπεύει ένα παράγοντα ευαισθησίας, καθορίζοντας την κατεύθυνση έρευνας στο χώρο των βαρών για το συναπτικό βάρος w_{ji} . Παραγωγίζοντας και τις δύο πλευρές της εξίσωσης ως προς $e_j(n)$ παίρνουμε:

$$\frac{\partial E(n)}{\partial w_{ji}(n)} = e_j(n) \quad (4.31)$$

Παραγωγίζοντας τώρα και τις δυο πλευρές της εξίσωσης 4.25 ως προς $y_j(n)$ παίρνουμε:

$$\frac{\partial e_j(n)}{\partial y_j(n)} = -1 \quad (4.32)$$

Τελικά, παραγωγίζοντας την 4.29 και 4.28 με το $v_j(n)$ και $w_{ji}(n)$ αντίστοιχα παίρνουμε:

$$\frac{\partial y_j(n)}{\partial v_j(n)} = \varphi'_j(v_j(n)) \quad (4.33)$$

$$\frac{\partial v_j(n)}{\partial w_{ji}(n)} = y_j(n) \quad (4.34)$$

Αντικαθιστώντας τις εξισώσεις 4.31, 4.32, 4.33, 4.34 στην εξίσωση 4.30 παίρνουμε:

$$\frac{\partial E(n)}{\partial w_{ji}(n)} = -e_j(n) \varphi'_j(v_j(n)) y_j(n) \quad (4.35)$$

Η διόρθωση $\Delta w_{ji}(n)$ που εφαρμόζεται στο $w_{ji}(n)$ καθορίζεται από τον «κανόνα δέλτα»:

$$\Delta w_{ji}(n) = -\eta \frac{\partial E(n)}{\partial w_{ji}(n)} \quad (4.36)$$

, όπου:

η : παράμετρος ρυθμού μάθησης (learning rate parameter).

Η χρήση του αρνητικού συμβόλου (-) στην 4.36 ερμηνεύεται σαν πτώση της κλίσης στο χώρο των βαρών.

Από τις 4.35 και 4.36 και παίρνουμε:

$$\Delta w_{ji}(n) = \eta \delta_j(n) y_i(n) \quad (4.37)$$

, όπου η τοπική κλίση $\delta_j(n)$ ορίζεται από την σχέση:

$$\delta_j(n) = -\frac{\partial E(n)}{\partial e_j(n)} \cdot \frac{\partial e_j(n)}{\partial y_j(n)} \cdot \frac{\partial y_j(n)}{\partial v_j(n)} = e_j(n) \varphi'_j(v_j(n)) \quad (4.38)$$

Η τοπική κλίση δείχνει τις απαιτούμενες αλλαγές στα βάρη. Σύμφωνα με την εξίσωση 4.38 η τοπική κλίση $\delta_j(n)$ για τον νευρώνα εξόδου j είναι ίσο με το γινόμενο του αντίστοιχου σήματος λάθους $e_j(n)$ και της παραγώγου $\varphi'_j(v_j(n))$ της συνάρτησης ενεργοποίησης του. Από τις εξισώσεις 4.37 και 4.38 φαίνεται ότι ένας παράγοντας κλειδί που εμπλέκεται στον υπολογισμό της προσαρμογής (του βάρους) $\Delta w_{ji}(n)$ είναι το σήμα στην έξοδο του νευρώνα j .

Στο σημείο αυτό μπορούμε να διακρίνουμε δυο περιπτώσεις, ανάλογα με το που είναι τοποθετημένος ο νευρώνας j , στο δίκτυο. Στην περίπτωση (I) ο j είναι ένας νευρώνας εξόδου. Την περίπτωση αυτή είναι απλό να την χειριστούμε, διότι κάθε νευρώνας εξόδου στο δίκτυο εφοδιάζεται με την επιθυμητή απόκρισή του, κάνοντας τον υπολογισμό του αντίστοιχου σήματος λάθους, μια εύκολη υπόθεση. Στην περίπτωση (II) ο νευρώνας j είναι ένας κρυφός νευρώνας.

1. Περίπτωση I : Ο νευρώνας j είναι ένας κόμβος εξόδου.

Όταν ο νευρώνας j βρίσκεται στο επίπεδο εξόδου του δικτύου, τροφοδοτείται με την επιθυμητή του έξοδο. Επομένως μπορούμε να χρησιμοποιήσουμε την εξίσωση 4.25 για να υπολογίσουμε το σήμα λάθους $e_j(n)$ που σχετίζεται με τον νευρώνα. Αφού έχουμε καθορίσει το σήμα λάθους, μετά είναι μια απλή διαδικασία να υπολογίσουμε την τοπική κλίση $\delta_j(n)$ χρησιμοποιώντας την εξίσωση 4.38.

2. Περίπτωση II : Ο νευρώνας j είναι ένας κρυφός κόμβος.

Όταν ένας νευρώνας j βρίσκεται σε ένα κρυφό επίπεδο του δικτύου, δεν υπάρχει κάποια καθορισμένη επιθυμητή απόκριση γι' αυτόν τον νευρώνα. Ανάλογα, το σήμα λάθους για ένα κρυφό νευρώνα θα έπρεπε να καθοριστεί επαναληπτικά σε όρους των σημάτων λάθους από

όλους τους νευρώνες με τους οποίους αυτός ο κρυφός νευρώνας συνδέεται άμεσα. Σε αυτήν την περίπτωση αυξάνει σημαντικά η πολυπλοκότητα του αλγορίθμου πίσω διάδοσης.

Σύμφωνα με την εξίσωση 4.38 μπορούμε να ορίσουμε πάλι την τοπική κλίση $\delta_j(n)$ για κρυφό νευρώνα j ως εξής:

$$\delta_j(n) = -\frac{\partial E(n)}{\partial y_j(n)} \cdot \frac{\partial y_j(n)}{\partial v_j(n)} = -\frac{\partial E(n)}{\partial y_j(n)} \phi'_j(v_j(n)) \quad (4.39)$$

, όπου ο j είναι κρυφός νευρώνας

Για να υπολογίσουμε την μερική παράγωγο $\frac{\partial E(n)}{\partial y_j(n)}$ μπορούμε να προχωρήσουμε ως εξής :

Το στιγμιαίο άθροισμα των τετραγώνων του λάθους στην επανάληψη n , δίνεται από τη σχέση:

$$E(n) = \frac{1}{2} \sum_{k \in C} e_k^2(n) \quad (4.40)$$

, όπου ο νευρώνας k είναι ένας κόμβος εξόδου

Καταλήγουμε:

$$\frac{\partial E(n)}{\partial y_j(n)} = \sum_k e_k \frac{\partial e_k(n)}{\partial y_j(n)} \quad (4.41)$$

Στη συνέχεια χρησιμοποιούμε τον αλυσιδωτό κανόνα για τον υπολογισμό της μερικής παράγωγου.

$$\frac{\partial E(n)}{\partial y_j(n)} = \sum_k e_k \frac{\partial e_k(n)}{\partial v_j(n)} \cdot \frac{\partial v_k(n)}{\partial y_j(n)} \quad (4.42)$$

Αλλά από την εξίσωση 4.28 έχουμε ότι:

$$e_k(n) = d_k(n) - y_k(n) \quad (4.43)$$

$$\text{ή } e_k(n) = d_k(n) - \phi_k(v_k(n)) \quad (4.44)$$

, αν ο νευρώνας k είναι κόμβος εξόδου

Επομένως:

$$\frac{\partial e_k(n)}{\partial v_k(n)} = -\phi'(v_k(n)) \quad (4.45)$$

Για τον νευρώνα k , το εσωτερικό επίπεδο ενεργοποίησης του δικτύου είναι:

$$v_k(n) = \sum_{j=0}^q w_{kj}(n) - y_j(n) \quad (4.46)$$

, όπου q είναι ο συνολικός αριθμός εισόδων (εξαιρούμε το κατώφλι), που εφαρμόζονται στον νευρώνα k . Και εδώ, επίσης το $w_{k0}(n)$ ισούται με το $\theta_k(n)$, που εφαρμόζεται στον νευρώνα k και η αντίστοιχη είσοδος y_0 έχει σταθερή τιμή -1 .

Από την εξίσωση 4.28:

$$\frac{\partial v_k(n)}{\partial y_j(n)} = w_{kj}(n) \quad (4.47)$$

Έτσι, χρησιμοποιώντας τις εξισώσεις 4.45 και 4.47 στην 4.42 παίρνουμε την επιθυμητή μερική παράγωγο.

$$\frac{\partial E(n)}{\partial y_j(n)} = \sum_k e_k(n) \phi'_k(v_k(n)) w_{kj}(n) = \sum_k \delta_k(n) w_{kj}(n) \quad (4.48)$$

Εδώ έχουμε χρησιμοποιήσει τη σχέση $\delta_k(n) = e_k(n) \phi'_k(v_k(n))$ (δηλ. τη σχέση 4.38, όπου ο k είναι νευρώνας εξόδου).

Τελικά χρησιμοποιώντας την εξίσωση 4.48 στην 4.39, παίρνουμε την τοπική κλίση $\delta_j(n)$, για ένα κρυφό νευρώνα j , αφού επαναδιατάξουμε τους όρους, ως εξής:

$$\delta_j(n) = \phi'_j(v_j(n)) \sum_k \delta_k(n) w_{kj}(n) \quad (4.49)$$

, όπου ο νευρώνας j είναι κρυφός.

Ο παράγοντας $\phi'_j(v_j(n))$ στην 4.49 εξαρτάται αποκλειστικά από την συνάρτηση ενεργοποίησης που σχετίζεται με τον κρυφό νευρώνα j .

Ο παράγοντας του απομένει στο υπολογισμό του $\delta_j(n)$, δηλαδή ο $\sum_k \delta_k(n) w_{kj}(n)$, δηλαδή η πρόσθεση για όλα τα k , εξαρτάται από δύο σύνολα όρων :

- Το πρώτο σύνολο όρων, το $\delta_k(n)$, απαιτεί γνώση των σημάτων λάθους $e_k(n)$, για όλους εκείνους τους νευρώνες που βρίσκονται στο αμέσως δεξιό επίπεδο, από τον κρυφό νευρώνα j , οι οποίοι είναι άμεσα συνδεδεμένοι με το νευρώνα j .

- Το δεύτερο σύνολο όρων, το $w_{kj}(n)$, αποτελείται από τα συναπτικά βάρη που σχετίζονται μ' αυτές τις συνδέσεις.

Συνοψίζοντας τις σχέσεις, που έχουμε παράγει για τον αλγόριθμο Πίσω-Διάδοσης:

Πρώτον, η διόρθωση $\Delta w_{ji}(n)$ που εφαρμόζεται στο συναπτικό βάρος, που συνδέει τον νευρώνα i στο νευρώνα j καθορίζεται από τον δέλτα κανόνα:

$$\Delta w_{ji}(n) = \eta \cdot \delta_j(n) \cdot y_i(n) \quad (4.50)$$

Δεύτερον, η τοπική κλίση $\delta_j(n)$ εξαρτάται από το εάν ο νευρώνας j είναι ένας κόμβος εξόδου ή αν ανήκει σε κρυμμένο επίπεδο.

- Εάν ο νευρώνας j είναι ένας κόμβος εξόδου, τότε η τοπική κλίση $\delta_j(n)$ ισούται με το γινόμενο της παραγώγου $\phi'_j(u_j(n))$ και του σφάλματος $e_j(n)$. Δηλαδή αν ο j είναι κόμβος εξόδου, τότε:

$$\delta_j(n) = e_j(n) \cdot \phi'_j(v_j(n)) \quad (4.51)$$

- Εάν ο νευρώνας j είναι ένας κρυφός νευρώνας, τότε η τοπική κλίση $\delta_j(n)$ ισούται με το γινόμενο της σχετιζόμενης παραγώγου $\phi'_j(u_j(n))$ και του ζυγισμένου αθροίσματος (weighted sum) των δ_k , που υπολογίζονται για τους νευρώνες, στο επόμενο κρυφό ή επίπεδο εξόδου, που είναι συνδεδεμένοι στον j . Δηλαδή, αν ο νευρώνας j είναι κρυφός κόμβος τότε:

$$\delta_j(n) = \phi'_j(v_j(n)) \cdot \sum_k \delta_k(n) w_{kj}(n) \quad (4.52)$$

4.9 Τεχνητά Νευρωνικά Δίκτυα Συνάρτησης Ακτινικής Βάσης (Radial-basis Function Neural Networks)

Μία ακόμη σημαντική κατηγορία τεχνητών νευρωνικών δικτύων πρόσθιας τροφοδότησης αποτελούν τα Δίκτυα Ακτινικών Συναρτήσεων Βάσης (Radial Basis Function Networks), τα οποία για συντομία ονομάζονται δίκτυα RBF.

Τα δίκτυα RBF βασίζονται στη μαθηματική θεωρία της προσέγγισης συναρτήσεων και ουσιαστικά υλοποιούν μία συνάρτηση παρεμβολής (interpolation function), η οποία προσεγγίζει την τιμή μιας συνάρτησης σε κάποιο σημείο ως το μέσο όρο των τιμών της συνάρτησης σε κοντινά σημεία.

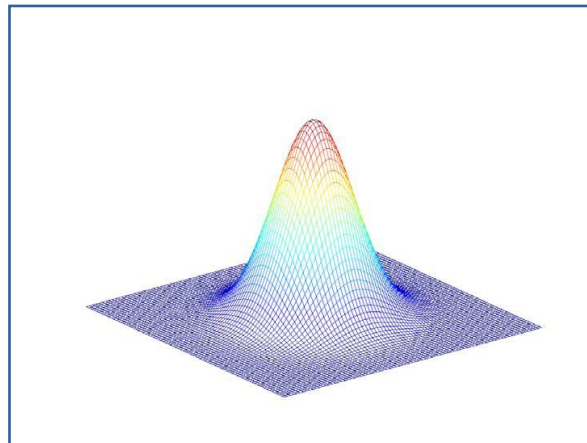
Η δομή ενός τέτοιου είδους δικτύου συνίσταται από ένα στρώμα εισόδου με n εισόδους, ένα ενδιάμεσο-κρυφό στρώμα με p νευρώνες και ένα στρώμα εξόδου με m νευρώνες. Οι νευρώνες του στρώματος εισόδου λαμβάνουν τα δεδομένα εισόδου και συνδέουν το δίκτυο με το περιβάλλον. Το επίπεδο εισόδου περιέχει τη συνάρτηση ενεργοποίησης f , η οποία είναι ακτινικού τύπου. Το ενδιάμεσο στρώμα αποτελείται από έναν μεγάλο αριθμό νευρώνων και εφαρμόζει ένα μη-γραμμικό μετασχηματισμό από το στρώμα εισόδου στο κρυμμένο στρώμα. Το στρώμα εξόδου είναι γραμμικό

και δίνει την απόκριση του δικτύου ως διάνυσμα που ενεργοποιείται από την εφαρμογή των διανυσμάτων εισόδου στο στρώμα εισόδου.

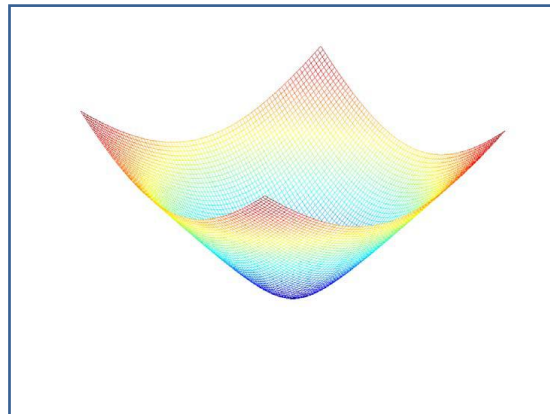
Οι νευρώνες που ανήκουν στα κρυφά στρώματα του νευρωνικού δικτύου παρέχουν ένα σύνολο συναρτήσεων ακτινικού τύπου (radial basis functions), οι οποίες προσεγγίζουν οποιαδήποτε συνεχή συνάρτηση και παρουσιάζονται στα επόμενα σχήματα. Στις ακτινικές συναρτήσεις βάσης υπάρχει κάποιο διάνυσμα c το οποίο καλείται κέντρο (centre) και η απόσταση του x από το κέντρο καθορίζει την τιμή της συνάρτησης.

$$f(x) = f(\|x - c\|) \quad (4.53)$$

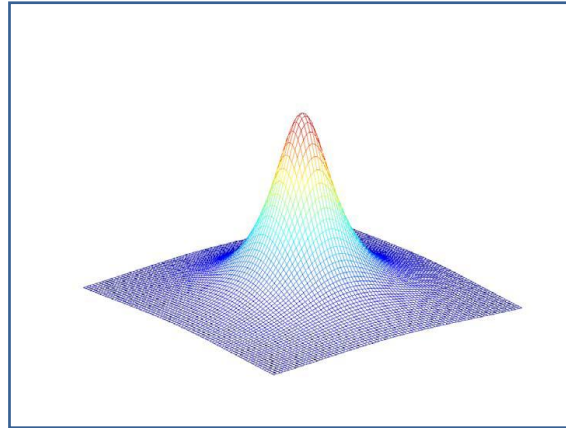
Έστω παραδείγματος χάριν ότι μία συνάρτηση έχει κέντρο c . Τότε αν x είναι η είσοδος (όρισμα) της συνάρτησης, η έξοδος της RBF είναι τόσο μεγαλύτερη όσο πιο κοντά είναι το x στο c . Δηλαδή όσο η διαφορά $x-c$ μικραίνει τόσο πιο μεγάλη έξοδος παράγεται. Η συνάρτηση RBF έχει δηλαδή την ενδογενή δυνατότητα να «αναγνωρίζει» εκείνα τα ορίσματα που είναι κοντά στο κέντρο της. Στα σχήματα που ακολουθούν παρουσιάζονται παραδείγματα συναρτήσεων ακτινικού τύπου.



(α) Συνάρτηση Gauss: $f(x) = e^{-\frac{\|x-c\|^2}{\sigma^2}}$



(β) Πολυτετραγωνική συνάρτηση: $f(x) = \frac{(\|x-c\|^2 + \sigma^2)^{-1}}{\sigma}$



(γ) Συνάρτηση Cauchy: $f(x) = \frac{(\|x-c\|^2 + \sigma^2)^{-1}}{\sigma}$

Σχήμα 4.21: Παραδείγματα συναρτήσεων ακτινικού τύπου (Πηγή: Διαμαντάρας, (2006)):

Αξίζει να σημειωθεί ότι ο νευρώνας RBF στο κρυφό στρώμα, χαρακτηρίζεται κυρίως από το κέντρο του c και το εύρος της συνάρτησης σ , και όχι από τα συναπτικά βάρη, τα οποία υφίστανται μονάχα μεταξύ του κρυφού στρώματος και του στρώματος εξόδου.

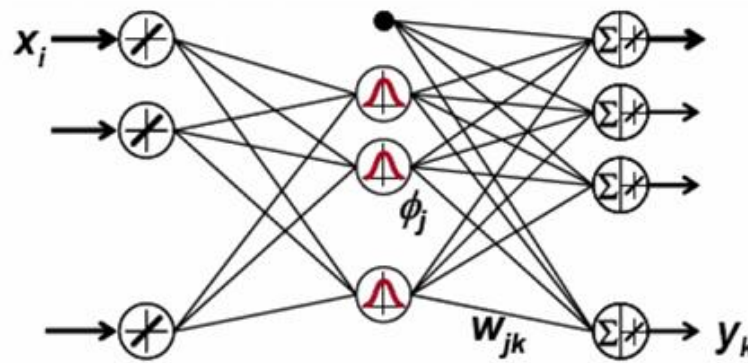
Το πλάτος της συνάρτησης (spread), δηλαδή το πόσο γρήγορα ή αργά θα «πέφτει» δεξιά και αριστερά από το κέντρο, μπορεί να καθοριστεί ελεύθερα. Συνήθως όμως φροντίζεται τα πλάτη να είναι ίσα μεταξύ τους.

$$\sigma = \frac{\max - d}{\sqrt{2n}} \quad (4.54)$$

, όπου d : το πλάτος του διαστήματος προτύπων και

n : το πλήθος των προτύπων

Το ενδιάμεσο στρώμα χρησιμοποιείται για την ομαδοποίηση των προτύπων, ανάλογα με τη θέση τους στο επίπεδο ή στο χώρο, με την εύρεση των κέντρων των ομάδων, του μέσου όρου, δηλαδή των διανυσμάτων της κάθε ομάδας. Το κάθε πρότυπο διάνυσμα εισόδου κατατάσσεται στην ομάδα εκείνη που απέχει την ελάχιστη απόσταση από το κέντρο της (Mathworks).



Σχήμα 4.22: Τοπολογία δικτύου RBF

Το τμήμα του δικτύου που αποτελείται από το ενδιάμεσο στρώμα και το στρώμα εξόδου, εκπαιδεύεται με τη μέθοδο της επίβλεψης. Ενώ για τους νευρώνες εξόδου χρησιμοποιείται μία γραμμική συνάρτηση ενεργοποίησης. Ο κανόνας εκπαίδευσης του κρυφού στρώματος(ο οποίος είναι εξαιρετικά αργός), διαφέρει από τον κανόνα εκπαίδευσης του εξωτερικού στρώματος(που πραγματοποιείται εύκολα και γρήγορα).

Κατά τη διάρκεια της εκπαίδευσης αρχικά υπολογίζεται για κάθε πρότυπο, η συνάρτηση της απόστασης του προτύπου από το κάθε κέντρο, η οποία αποτελεί και την είσοδο στο ενδιάμεσο στρώμα. Κατόπιν η εκπαίδευση πραγματοποιείται όπως το δίκτυο ADALINE. Η εκμάθηση ισοδυναμεί με την εύρεση μιας καμπύλης ή επιφάνειας στον πολυδιάστατο χώρο, η οποία να παρέχει την καλύτερη προσαρμογή στα διανύσματα εκπαίδευσης του δικτύου με βάση κάποια στατιστικής φύσης κριτήρια.

Τέλος, κατά την ανάκληση εισάγονται τα πρότυπα για την ταξινόμηση και πραγματοποιείται ο υπολογισμός της συνάρτησης της απόστασης του προτύπου από το κάθε κέντρο. Η ανάκληση γίνεται και σε αυτήν την περίπτωση όπως στο δίκτυο ADALINE.

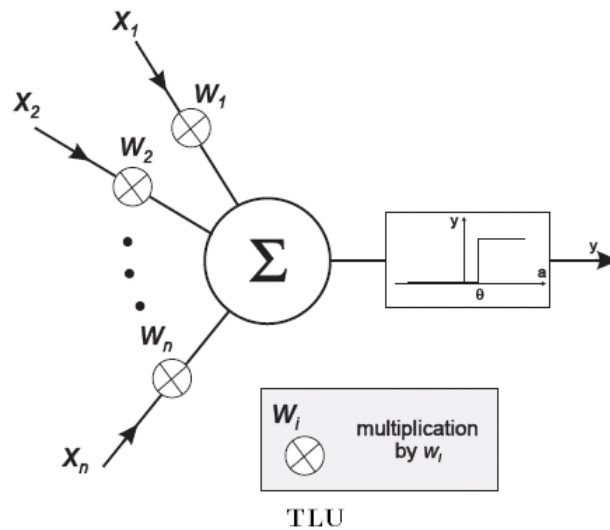
4.10 Μαθηματικό υπόβαθρο τεχνητών νευρωνικών δικτύων

4.10.1 Το μοντέλο McCulloch – Pitts

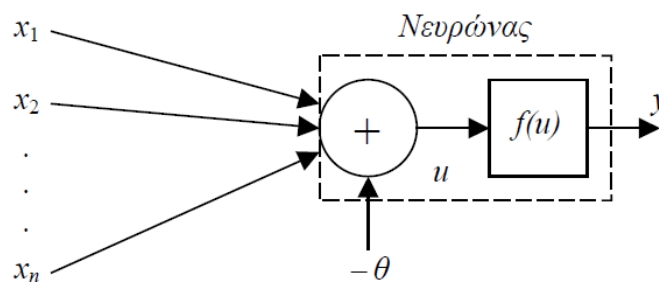
Προκειμένου να ερμηνευτεί το μαθηματικό υπόβαθρο των τεχνητών νευρωνικών δικτύων είναι απαραίτητο να περιγραφεί αρχικά το μοντέλο McCulloch – Pitts, καθώς και η λειτουργία Threshold Logic (Abdi, 2001).

Η επεξεργασία της πληροφορίας στο μοντέλο McCulloch – Pitts ακολουθεί την εξής διαδικασία: Τα σήματα αρχικά εμφανίζονται στην είσοδο της μονάδας, δηλαδή στις συνάψεις. Η τιμή του κάθε σήματος πολλαπλασιάζεται με το βάρος, δηλώνοντας έτσι την ισχύ της σύναψης. Τα πολλαπλασιασμένα με τα βάρη σήματα εισόδου στη συνέχεια αθροίζονται, με αποτέλεσμα να παράγεται μία συνολική μονάδα ενεργοποίησης. Εάν η συγκεκριμένη ενεργοποίηση υπερβαίνει ένα

προκαθορισμένο όριο, η μονάδα παράγει μία απόκριση εξόδου. Αυτή η λειτουργία είναι γνωστή ως Threshold Logic Unit (TLU) και προτάθηκε από τους McCulloch και Pitts (McCulloch and Pitts, 1943)



Σχήμα 4.23: Δομή Νευρωνικού Δικτύου (Πηγή: Gurney,Routledge, (1997)).



Σχήμα 4.24: Το μοντέλο McCulloch – Pitts

Έστω ότι εξετάζουμε τον νευρώνα k . Υποθέτουμε ότι έχουμε n εισόδους με σήματα εισόδου $x_1, x_2, \dots, x_n$ και βάρη $w_1, w_2, \dots, w_n$ αντίστοιχα. Το βάρος απεικονίζει το πόσο σημαντική είναι η προσφορά του συγκεκριμένου σήματος. Όπως περιγράφηκε αναλυτικά και σε προηγούμενη παράγραφο, ο νευρώνας αποτελείται από μία συνάρτηση αθροίσματος (Summing Function) και μία συνάρτηση ενεργοποίησης (Activation Function). Οι εισοδοί του νευρώνα (Input Signals) πολλαπλασιάζονται με τα συναπτικά βάρη w_i (Synaptic Weights) των συνάψεων του νευρώνα και το όρισμα της συνάρτησης αθροίσματος του νευρώνα δίνεται:

$$u_k = \sum_{i=1}^n w_{ki} x_i \quad (4.55)$$

Στη συνέχεια προστίθεται στην παραπάνω συνάρτηση αθροίσματος μία τιμή πόλωσης (bias) b του νευρώνα. Το αποτέλεσμα U που προκύπτει, αποτελεί το όρισμα της συνάρτησης ενεργοποίησης (Activation Function ή Transfer Function). Τέλος το αποτέλεσμα y_k είναι το σήμα εξόδου ή κατάσταση ενεργοποίησης (Output Signal) του νευρώνα.

Επομένως:

$$U_k = u_k + b_k \quad (4.56)$$

και

$$y_k = \varphi(U_k) \quad (4.57)$$

, όπου:

x_i : το διάνυσμα εισόδου, με $i=1$ έως n ,

w_{ki} : τα βάρη των συνάψεων i του νευρώνα k , με τα οποία πολλαπλασιάζονται τα στοιχεία x_i του διανύσματος εισόδου και

b_k : τιμή πόλωσης διαφορετική για κάθε νευρώνα

Μερικές φορές θεωρούμε ότι εκτός από τα εισερχόμενα σήματα και τα αντίστοιχα βάρη w , ο νευρώνας έχει και ένα εσωτερικό βάρος που τον χαρακτηρίζει (bias), το οποίο εκφράζει τη προδιάθεση του νευρώνα και θεωρείται ως ένα εξωτερικό ερέθισμα. Είναι τελείως ξεχωριστό από τα άλλα βάρη, παρ' όλα αυτά όμως δρα με τον ίδιο ακριβώς τρόπο. Θεωρούμε πάντα ότι η τιμή του σήματος που περνάει σε όλα τα εσωτερικά βάρη είναι 1. Έτσι οι μονάδες του b θα είναι ίδιες με τις μονάδες του γινομένου ($s \cdot w$). Επομένως θεωρείται ως ένα εξωτερικό ερέθισμα που προστίθεται στο υπόλοιπο άθροισμα για να δώσει το σωστό U . Τέλος το bias αποτελεί μία παράμετρο η οποία ορίζεται κατά την εκπαίδευση του δικτύου.

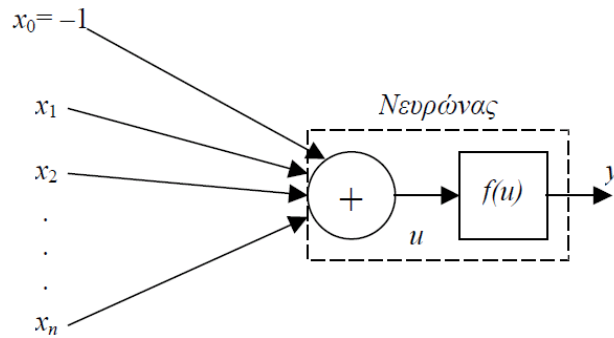
Στο μοντέλο των McCulloch και Pitts, που αναφέραμε παραπάνω τα σήματα εξόδου μπορούν να πάρουν μονάχα τις τιμές '0' και '1'. Τότε οι εξισώσεις παίρνουν την εξής μορφή:

$$u_k = \sum_{i=1}^n w_{ki} x_i \quad (4.58)$$

$$U_k = u_k + b_k \quad (4.59)$$

Και ως συνάρτηση ενεργοποίησης χρησιμοποιείται η συνάρτηση κατωφλιού, οπότε η έξοδος y_k δίνεται από:

$$y_k = \varphi(U_k) = \begin{cases} 1 & \text{εάν } U_k \geq 0 \\ 0 & \text{εάν } U_k < 0 \end{cases} \quad (4.60)$$



Σχήμα 4.25: Θέτοντας την τιμή πόλωσης ως συναπτικό βάρος

Στη συνέχεια θα χρησιμοποιήσουμε κάποια στοιχεία από τη θεωρία των διανυσμάτων προκειμένου να αποδείξουμε ότι ουσιαστικά υπάρχει ένα βασικό μαθηματικό υπόβαθρο στη δομή και λειτουργία των τεχνητών νευρώνων.

4.10.2 Θεωρία Διανυσμάτων

Τα διανύσματα ορίζονται ως μεγέθη, που για να προσδιοριστούν, εκτός από το μέτρο τους, χρειάζεται να ξέρουμε επιπλέον, τη διεύθυνσή τους στο χώρο και τη φορά τους.

Προκειμένου να γενικεύσουμε αργότερα σε n -διάστατα πρότυπα, θα περιγράψουμε τα διανύσματα βάσει του καρτεσιανού συστήματος συντεταγμένων. Δίνουμε λοιπόν τις προβολές ενός διανύσματος πάνω στους δύο κάθετους άξονες x_1 και x_2 .

Το διάνυσμα περιγράφεται από ένα ζεύγος αριθμών v_1 και v_2 , τα οποία αποτελούν τις συνιστώσες στο επιλεγμένο σύστημα συντεταγμένων. Οπότε το διάνυσμα περιγράφεται ως: $\vec{v} = (v_1, v_2)$. Ένα n -διάστατο διάνυσμα θα ορίζεται ως $\vec{v} = (v_1, v_2, \dots, v_n)$, ενώ το διάνυσμα των συναπτικών βαρών: $\vec{w} = (w_1, w_2, \dots, w_n)$ και το διάνυσμα εισόδων: $\vec{x} = (x_1, x_2, \dots, x_n)$. Αξίζει να σημειωθεί ότι η σειρά με την οποία αναγράφονται οι αριθμοί παίζει πολύ σημαντικό ρόλο, καθώς καθορίζει τη διεύθυνση του διανύσματος.

4.10.2.1 Μήκος διανύσματος

Όσον αφορά το 2-διάστατο επίπεδο, το μήκος του διανύσματος ισούται με το μήκος του στο επίπεδο. Το μήκος ορίζεται λοιπόν από το πυθαγόρειο θεώρημα και δίνεται από την σχέση:

$$\|\vec{v}\| = \sqrt{v_1^2 + v_2^2} \quad (4.61)$$

Για n -διαστάσεις το μήκος θα δίνεται από την σχέση:

$$\|\vec{v}\| = \left[\sum_{i=1}^n v_i^2 \right]^{1/2} \quad (4.62)$$

4.10.3 Ο ρόλος του εσωτερικού γινομένου

Θεωρούμε τα διανύσματα $\vec{v}$ και $\vec{w}$, των οποίων η μεταξύ τους γωνία είναι 45° . Το εσωτερικό γινόμενο των δύο παραπάνω διανυσμάτων θα ισούται με:

$$\vec{v} \cdot \vec{w} = v_1 w_1 + v_2 w_2 \quad (4.63)$$

Παρατηρούμε ότι το δεξί μέλος της εξίσωσης θυμίζει την σχέση:

$$u_k = \sum_{i=1}^n w_{ki} x_i \quad (4.64)$$

Στο σημείο αυτό, είναι σημαντικό να αναφερθούμε στην έννοια του εσωτερικού γινομένου διανυσμάτων.

Το εσωτερικό γινόμενο δύο διανυσμάτων $\vec{v}$ και $\vec{w}$, συμβολίζεται με $\vec{v} \cdot \vec{w}$ και ορίζεται ως:

$$\vec{v} \cdot \vec{w} = \|\vec{v}\| \|\vec{w}\| \cos \varphi \quad (4.65)$$

Όπου, με $\|\vec{v}\|$ και $\|\vec{w}\|$ συμβολίζουμε τα μέτρα των διανυσμάτων $\vec{v}$ και $\vec{w}$ αντίστοιχα, με φ τη γωνία που σχηματίζεται ανάμεσα στα δύο διανύσματα και $\cos \varphi$ το συνημίτονό της. Το εσωτερικό γινόμενο είναι ουσιαστικά το γινόμενο του πρώτου διανύσματος με τη προβολή του δεύτερου πάνω στο πρώτο. Είναι επίσης φανερό πως το εσωτερικό γινόμενο δύο διανυσμάτων θα είναι πάντα ένας αριθμός και όχι ένα νέο διάνυσμα όπως στην πρόσθεση και την αφαίρεση διανυσμάτων.

Όταν τα διανύσματα είναι κάθετα μεταξύ τους, το εσωτερικό γινόμενο είναι ίσο με το 0, ενώ όταν είναι παράλληλα (ή αντιπαράλληλα) το εσωτερικό γινόμενο ισούται με το θετικό (ή αρνητικό αντίστοιχα) γινόμενο των μέτρων τους. Αυτό είναι φανερό γιατί $\cos 90^\circ = 0$ και $\cos 0^\circ = 1$ (επίσης $\cos 180^\circ = -1$).

Το εσωτερικό γινόμενο ορίζεται επίσης και ως το άθροισμα των γινομένων των επιμέρους συνιστωσών των διανυσμάτων. Συγκεκριμένα, αν το διάνυσμα $\vec{a}_n$ διαστάσεων, είναι $\vec{a} = a_1 \hat{e}_1 + a_2 \hat{e}_2 + \dots + a_n \hat{e}_n$ και $\vec{b} = b_1 \hat{e}_1 + b_2 \hat{e}_2 + \dots + b_n \hat{e}_n$ με $\hat{e}_1, \hat{e}_2, \dots, \hat{e}_n$ να είναι τα μοναδιαία διανύσματα βάσης ενός διανυσματικού χώρου η διαστάσεων, τότε το εσωτερικό γινόμενο γράφεται ως:

$$\vec{a} \cdot \vec{b} = a_1 b_1 + a_2 b_2 + \dots + a_n b_n \quad (4.66)$$

Το εσωτερικό γινόμενο χρησιμοποιείται εκτεταμένα στις εξισώσεις της φυσικής, όπως και όλος ο διανυσματικός λογισμός άλλωστε.

Υπολογίζοντας τώρα το εσωτερικό γινόμενο των διανυσμάτων $\vec{v}$ και $\vec{w}$, θα έχουμε:

$$\vec{v} \cdot \vec{w} = \|\vec{v}\| \|\vec{w}\| \cos \varphi \quad (4.67)$$

Όπως προαναφέρθηκε στην περίπτωση των n -διαστάσεων, το εσωτερικό γινόμενο θα είναι δυνατόν να υπολογιστεί μέσω της έκφρασης:

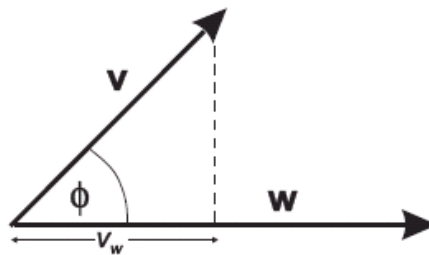
$$\vec{v} \cdot \vec{w} = v_1 w_1 + v_2 w_2 + \dots + v_n w_n \quad (4.68)$$

Επομένως καταλήγουμε στο ότι:

$$\vec{v} \cdot \vec{w} = \sum_{i=1}^n w_i v_i \quad (4.69)$$

Παρ' όλο που είναι αδύνατον να σχεδιαστούν τα διανύσματα για $n > 3$, θα ακολουθείται η ίδια διαδικασία γεωμετρικής ερμηνείας, αφού ο ορισμός τους είναι ο ίδιος. Για ευκολία στην κατανόηση, θα εξετάσουμε την σπουδαιότητα του εσωτερικού γινομένου στην περίπτωση των δύο διαστάσεων.

Έστω v_w η συνιστώσα του διανύσματος $\vec{v}$ πάνω στη διεύθυνση του διανύσματος $\vec{w}$, ή αλλιώς η προβολή του $\vec{v}$ πάνω στο $\vec{w}$.



Σχήμα 4.26: Προβολή διανύσματος

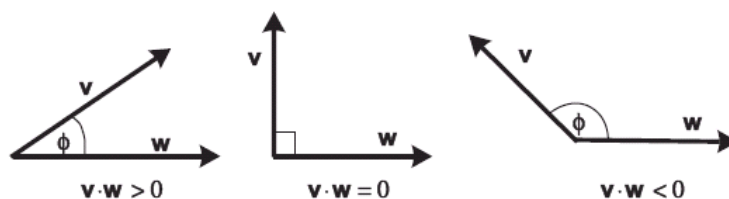
Τότε η προβολή v_w , σύμφωνα με το ορθογώνιο τρίγωνο που σχηματίζεται και χρησιμοποιώντας την έκφραση του εσωτερικού γινομένου, θα ισούται με:

$$v_w = \|\vec{v}\| \cos \varphi = \|\vec{v}\| \frac{\vec{v} \cdot \vec{w}}{\|\vec{v}\| \|\vec{w}\|} = \frac{\vec{v} \cdot \vec{w}}{\|\vec{w}\|} \quad (4.70)$$

Στην περίπτωση που η γωνία φ είναι μικρή, τότε η τιμή του συνημίτονου της είναι πολύ κοντά στη μονάδα. Επομένως το εσωτερικό γινόμενο, αλλά και η προβολή v_w θα πάρουν μέγιστη τιμή.

Στην περίπτωση που $\varphi = 90^\circ$, η τιμή του συνημίτονου είναι μηδενική, άρα και του εσωτερικού γινομένου. Η προβολή v_w είναι μηδενική καθώς τα διανύσματα είναι κάθετα μεταξύ τους.

Εάν $90^\circ < \varphi < 270^\circ$, η τιμή του συνημίτονου είναι αρνητική, όπως και η τιμή της προβολής v_w . Μπορεί το μέτρο του v_w να έχει μεγάλη τιμή, αλλά το αρνητικό πρόσημο υποδηλώνει ότι τα δύο διανύσματα $\vec{v}$ και $\vec{w}$ κατευθύνονται σε διαφορετικά ημιεπίπεδα.



Σχήμα 4.27: Σχηματική απεικόνιση περιπτώσεων εσωτερικού γινομένου

4.10.4 Εσωτερικό Γινόμενο και Τεχνητοί Νευρώνες

Έχοντας αναφέρει όλα τα παραπάνω θα συνδυάσουμε τη δράση ενός τεχνητού νευρώνα με την έκφραση του εσωτερικού γινομένου, όσον αφορά τα διανύσματα εισόδων και συναπτικών βαρών $\vec{x}$ και $\vec{w}$, αντίστοιχα.

Έτσι θα έχουμε:

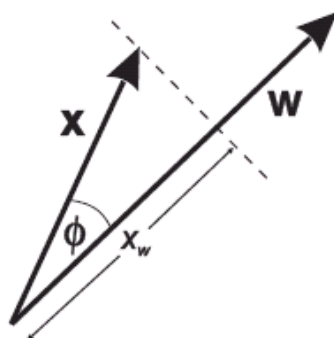
$$u = \vec{w} \cdot \vec{x} \quad (4.71)$$

Θέτοντας την ενεργοποίηση ίση με το σκαλοπάτι (threshold) προκύπτει:

$$\vec{w} \cdot \vec{x} = \theta \quad (4.72)$$

Εάν το $\vec{w}$ και το θ είναι σταθερές, τότε η προβολή του $\vec{x}$ πάνω στο $\vec{w}$ αποτελεί και αυτή μια σταθερά.

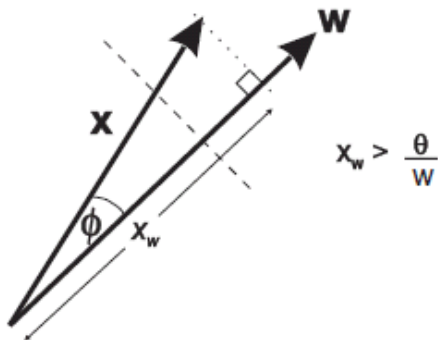
$$\text{Άρα: } x_w = \frac{\theta}{\|\vec{w}\|}$$



Σχήμα 4.28: Προβολή διανύσματος

Κατόπιν, διακρίνουμε τις εξής δύο περιπτώσεις:

1. Εάν $x_w > \frac{\theta}{\|\vec{w}\|}$:



Το $\vec{x}$ εκτείνεται σε μεγαλύτερη απόσταση από τη διακεκομμένη γραμμή.

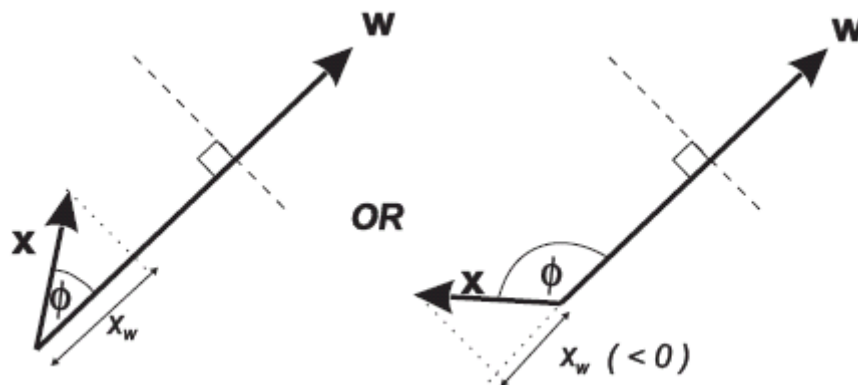
Έχουμε αποδείξει ότι:

$$x_w = \frac{\vec{x} \cdot \vec{w}}{\|\vec{w}\|} > \frac{\theta}{\|\vec{w}\|} \rightarrow \vec{x} \cdot \vec{w} > \theta \rightarrow u > \theta$$

Επομένως η ενεργοποίηση είναι μεγαλύτερη από το threshold.

Σχήμα 4.29: Περίπτωση $x > \theta$

2. Εάν $x_w < \frac{\theta}{\|\vec{w}\|}$:



Σχήμα 4.30: Περίπτωση $w \cdot x < \theta$

Το $\vec{x}$ εκτείνεται σε μικρότερη απόσταση από τη διακεκομμένη γραμμή.

Έχουμε αποδείξει ότι:

$$x_w = \frac{\vec{x} \cdot \vec{w}}{\|\vec{w}\|} < \frac{\theta}{\|\vec{w}\|} \rightarrow \vec{x} \cdot \vec{w} < \theta \rightarrow u < \theta$$

Επομένως η ενεργοποίηση είναι μικρότερη από το threshold (Πηγή: Gurney, Routledge, (1997), Wikipedia).

4.10.5 Μαθηματικό υπόβαθρο στην εκπαίδευση του νευρώνα

Η εκπαίδευση των τεχνητών νευρώνων ουσιαστικά αποτελεί μια διαδικασία προσαρμογής των βαρών. Η μάθηση επιτυγχάνεται μέσω μιας επαναληπτικής διαδικασίας ρυθμίσεων της τιμής των συναπτικών βαρών και των κατωφλίων. Όπως προείπαμε την τιμή πόλωσης την προσομοιάζουμε με βάρος w_0 το οποίο αντιστοιχεί στο μοναδιαίο σήμα $x_0=+1$.

Έχουμε τις εξής περιπτώσεις:

- $\vec{x} \cdot \vec{w} \geq 0 \rightarrow y = 1$
- $\vec{x} \cdot \vec{w} \leq 0 \rightarrow y = 0$

Υπενθυμίζουμε ότι θέτοντας $\vec{x} \cdot \vec{w} = 0$, ουσιαστικά παίρνουμε το υπερεπίπεδο απόφασης, όπως έχουμε ήδη εξηγήσει σε προηγούμενη παράγραφο.

Στο σημείο αυτό, είναι απαραίτητο να υπενθυμίσουμε ορισμένες βασικές ιδιότητες της θεωρίας των διανυσμάτων.

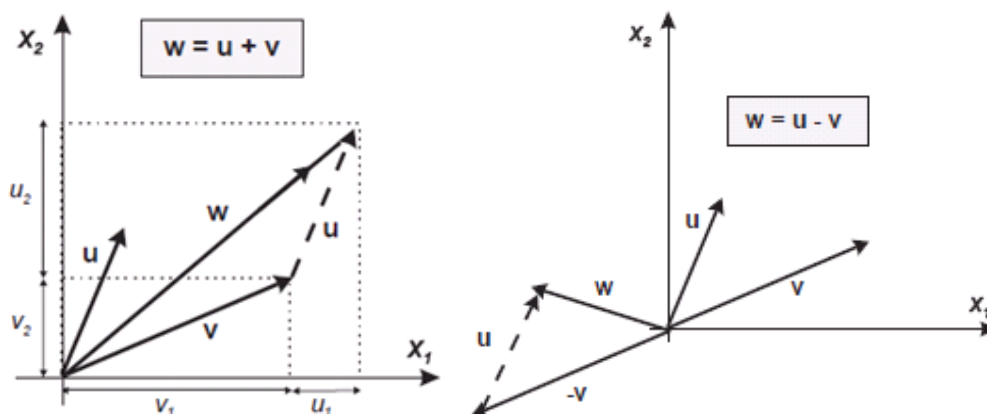
Έστω δύο διανύσματα: $\vec{u} = (u_1, u_2)$ και $\vec{v} = (v_1, v_2)$

Τότε προσθέτοντας τις αντίστοιχες συνιστώσες τους, προκύπτει ένα νέο διάνυσμα $\vec{w} = (w_1, w_2)$, όπου: $w_1 = u_1 + v_1$ και $w_2 = u_2 + v_2$.

Όσον αφορά την αφαίρεση διανυσμάτων, $(\vec{u} - \vec{v})$, ουσιαστικά προκύπτει από την πρόσθεση του $-\vec{v}$, δηλαδή του αντιθέτου.

Δύο μη μηδενικά διανύσματα ονομάζονται αντίθετα όταν έχουν αντίθετη κατεύθυνση και ίσα μέτρα.

Όλα τα παραπάνω συνοψίζονται στα δύο εξής σχήματα:



Σχήμα 4.31: Πρόσθεση & αφαίρεση διανυσμάτων (Πηγή: Gurney, Routledge, (1997), Wikipedia).

4.10.5.1 Προσαρμογή του διανύσματος των βαρών

Είναι ήδη γνωστό ότι στην επιβλεπόμενη ή συσχετιζόμενη εκμάθηση χρησιμοποιούνται δεδομένα εκπαίδευσης που αποτελούνται από πολλά ζεύγη προτύπων εισόδου-εξόδου, δηλαδή, για κάθε πρότυπο εισόδου, υπάρχει το αντίστοιχο πρότυπο εξόδου. Με αυτόν τον τρόπο το δίκτυο μαθαίνει μια συνάρτηση που προσεγγίζει την σχέση μεταξύ εισόδου – εξόδου και θεωρείται ότι μαθαίνει σωστά, αν τα σφάλματα πρόβλεψης τόσο για τα δεδομένα εκπαίδευσης, όσο και για τα δεδομένα αξιολόγησης είναι πολύ μικρά και επίσης αν η διαφορά μεταξύ των δύο σφαλμάτων είναι επίσης μικρή.

Έστω τώρα ότι $\vec{v}$ ένα διάνυσμα εισόδου και t είναι ο στόχος (έξοδος) και ότι πρόκειται για έναν νευρώνα τύπου McCulloch και Pitts, που όπως αναφέραμε παραπάνω, τα σήματα μπορούν να πάρουν μονάχα τις τιμές '0' και '1'. Υποθέτουμε λοιπόν ότι έχουμε ένα διάνυσμα εκπαίδευσης $\vec{v}$ του οποίου το πρόσφατο διάνυσμα βάρους είναι $\vec{w}$. Ο στόχος είναι $t = 1$, αλλά το συγκεκριμένο διάνυσμα βάρους, δίνει αποτέλεσμα $y = 0$. Κάτι το οποίο σημαίνει ότι η ενεργοποίηση βρέθηκε αρνητική ενώ για να δώσει το σωστό αποτέλεσμα θα έπρεπε να ήταν θετική. Εφ' όσον:

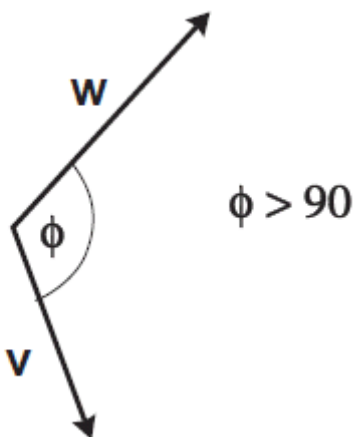
$$y = \begin{cases} 1 & \text{εάν } u \geq 0 \\ 0 & \text{εάν } u \leq 0 \end{cases}$$

1. Περίπτωση $t = 1$ ενώ $y = 0$:

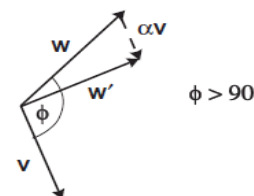
$$\hat{\phi} > 90^\circ \Rightarrow \vec{v} \cdot \vec{w} \leq 0 \Rightarrow u < 0 \Rightarrow y = 0$$

Στην περίπτωση αυτή θα πρέπει να στρέψουμε το διάνυσμα $\vec{w}$ δεξιόστροφα, έτσι ώστε να πλησιάσει το διάνυσμα $\vec{v}$ και να μειωθεί η μεταξύ τους γωνία. Ωστόσο η αλλαγή δε θα πρέπει να είναι ιδιαίτερα δραστική, αλλά σχετικά ομαλή σε σχέση με την προηγούμενη εκμάθηση.

Με τον τρόπο αυτό δημιουργούμε ένα νέο βάρος $\vec{w}'$, όπου: $\vec{w}' = \vec{w} + \alpha \vec{v}$

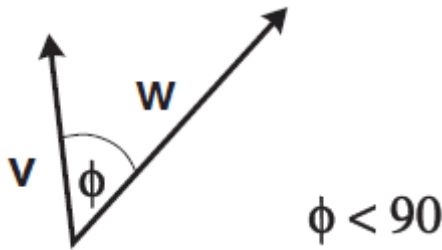


Σχήμα 4.32: Λάθος ταξινόμηση 1 -0



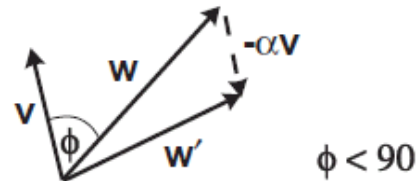
2. Περίπτωση $t = 0$ ενώ $y = 1$:

$$\hat{\phi} < 90^\circ \Rightarrow \vec{v} \cdot \vec{w} > 0 \Rightarrow u > 0 \Rightarrow y = 1$$



Στην περίπτωση αυτή θα πρέπει να στρέψουμε το διάνυσμα $\vec{w}$ δεξιόστροφα, έτσι ώστε να απομακρυνθεί από το διάνυσμα $\vec{v}$ και να αυξηθεί η μεταξύ τους γωνία.

Με τον τρόπο αυτό δημιουργούμε ένα νέο βάρος $\vec{w}'$, όπου: $\vec{w}' = \vec{w} - \alpha \vec{v}$



Σχήμα 4.33: Λάθος ταξινόμηση 0 - 1

Συνδυάζοντας τα παραπάνω:

$$\vec{w} = \vec{w} + \alpha(t - y)\vec{v} \quad (4.78)$$

, όπου $0 < \alpha < 1$

ή αλλιώς:

$$\Delta w_i = \alpha(t - y)v_i \quad (4.79)$$

Καταλήξαμε λοιπόν σε έναν κανόνα εκμάθησης ή εκπαίδευσης. Η παράμετρος α καλείται ρυθμός εκμάθησης. Ο συγκεκριμένος κανόνας χρησιμοποιείται ευρύτατα στο μοντέλο *Perceptron* (Πηγή: Gurney, Routledge, (1997), Wikipedia).

4.11 Πλεονεκτήματα των Νευρωνικών Δικτύων

Η συμπαγής παράλληλα κατανομημένη δομή και η ικανότητα μάθησης των νευρωνικών δικτύων, κάνουν δυνατή την επίλυση πολύπλοκων προβλημάτων, τα οποία σπάνε σε μικρότερα έργα, που αναλαμβάνουν τα νευρωνικά δίκτυα ανάλογα με τις δυνατότητές τους.

Η χρήση των Νευρωνικών Δικτύων παρέχει τις ακόλουθες χρήσιμες ιδιότητες :

1. **Μη-γραμμικότητα.** Αυτή οφείλεται στο γεγονός ότι ένα Νευρωνικό Δίκτυο δομείται από τη σύνδεση νευρώνων, οι οποίοι είναι μη-γραμμικές συσκευές. Η μη-γραμμικότητα είναι πολύ σημαντική ιδιότητα, ιδιαίτερα αν ο φυσικός μηχανισμός για την παραγωγή των σημάτων εισόδου είναι μη-γραμμικός.

2. **Σχεδιασμός Εισόδου-Εξόδου.** Ένα συνηθισμένο παράδειγμα μάθησης που καλείται επιβλεπόμενη μάθηση, εμπλέκει μεταβολή των συναπτικών βαρών του Νευρωνικού Δικτύου, εφαρμόζοντας ένα σύνολο δειγμάτων εξάσκησης ή παραδείγματα έργων. Κάθε παράδειγμα αποτελείται από ένα μοναδικό σήμα εισόδου και την επιθυμητή απόκριση. Η εξάσκηση του δικτύου επαναλαμβάνεται για πολλά παραδείγματα, μέχρι το δίκτυο να φτάσει σε μια σταθερή κατάσταση, όπου πλέον δεν γίνονται αλλαγές στα βάρη. Έτσι το δίκτυο μαθαίνει από τα παραδείγματα, κατασκευάζοντας ένα σχεδιασμό εισόδου-εξόδου για το εξεταζόμενο πρόβλημα.
3. **Προσαρμοστικότητα.** Τα Νευρωνικά Δίκτυα έχουν τη δυνατότητα να προσαρμόζουν τα βάρη τους στις αλλαγές του περιβάλλοντός τους. Μερικές φορές οι προσαρμογές οδηγούν σε μείωση της απόδοσης του συστήματος, γι' αυτό θα πρέπει να είναι επαρκείς δίλημμα σταθερότητας-πλαστικότητας.
4. **Συναφής Πληροφορία.** Η γνώση αναπαριστάνεται από την πολύ δομημένη και ενεργή κατάσταση του Νευρωνικού Δικτύου.
5. **Υλοποιησιμότητα σε VLSI** (Very-large-scale integration). Η συμπαγής παράλληλη φύση του Νευρωνικού Δικτύου, κάνει δυνατή την υλοποίηση του σε VLSI τεχνολογία, έτσι ώστε τα νευρωνικά δίκτυα να μπορούν να χρησιμοποιηθούν σε εφαρμογές πραγματικού χρόνου.
6. **Ομοιομορφία Ανάλυσης και Σχεδιασμού.** Η έννοια είναι ότι ο ίδιος συμβολισμός χρησιμοποιείται σε όλα τα πεδία που περιέχουν εφαρμογή των νευρωνικών δικτύων. Αυτό το χαρακτηριστικό υποδηλώνεται με διαφορετικούς τρόπους :
 - Οι νευρώνες παριστάνουν ένα συστατικό συνηθισμένο σε όλα τα νευρωνικά δίκτυα.
 - Αυτή η ιδιότητα κάνει δυνατή την εφαρμογή κοινών θεωριών και αλγορίθμων εκμάθησης σε διαφορετικές εφαρμογές των νευρωνικών δικτύων. Ρυθμιστικά (αναδρομικά) δίκτυα μπορούν να κατασκευαστούν μέσω μιας αμιγούς ολοκλήρωσης από modules.
7. **Αναλογία με Νευροβιολογία.** Ο σχεδιασμός νευρωνικών δικτύων γίνεται σε αναλογία με τον εγκέφαλο. Οι νευροβιολόγοι βλέπουν τα νευρωνικά δίκτυα σαν αντικείμενο έρευνας για την εξήγηση νευροβιολογικών φαινομένων. Ομοίως οι μηχανικοί βλέπουν στη νευροβιολογία για νέες ιδέες για την επίλυση πολύπλοκων προβλημάτων.

4.12 Εφαρμογές των Νευρωνικών Δικτύων

Η ευρεία εφαρμογή των νευρωνικών δικτύων σε μία ποικιλία προβλημάτων, σε πολλά διαφορετικά πεδία, τα κάνει πολύ ελκυστικά. Επίσης, οι πιο γρήγοροι υπολογιστές και αλγόριθμοι έχουν κάνει δυνατή την χρήση νευρωνικών δικτύων στην επίλυση πολύπλοκων βιομηχανικών προβλημάτων, τα οποία απαιτούσαν πάρα πολλούς υπολογισμούς. Από την πρώτη τους εφαρμογή σε

προσαρμοζόμενους εξισωτές καναλιών το 1984 μέχρι σήμερα έχουν εφαρμοστεί σε μεγάλη ποικιλία εφαρμογών όπως φαίνεται παρακάτω (Case, 1993):

1. **Αεροπορία:** Υψηλής απόδοσης αυτόματοι πιλότοι αεροπλάνων, προσομοιωτές πτήσης, συστήματα αυτομάτου ελέγχου αεροπλάνων, συστήματα ανίχνευσης βλαβών.
2. **Αυτοκίνηση:** Αυτοκινούμενα συστήματα αυτόματης πλοήγησης.
3. **Τραπεζικές εφαρμογές:** Αναγνώστες επιταγών και άλλων παραστατικών, συστήματα αξιολόγησης αιτήσεων δανειοδότησης .
4. **Ηλεκτρονική:** Πρόβλεψη ακολουθίας κωδίκων, μορφοποίηση ολοκληρωμένων κυκλωμάτων, έλεγχος διεργασιών, διάγνωση βλαβών ολοκληρωμένων κυκλωμάτων, μηχανική όραση, σύνθεση φωνής.
5. **Οικονομία:** Οικονομική ανάλυση, πρόβλεψη τιμών συναλλάγματος.
6. **Κοινωνική ασφάλιση:** Αξιολόγηση εφαρμοζόμενης πολιτικής, βελτιστοποίηση παραγωγής.
7. **Βιομηχανία:** Βιομηχανικός έλεγχος διεργασιών, ανάλυση και σχεδίαση προϊόντων, συστήματα ποιοτικού ελέγχου, διάγνωση βλαβών διεργασιών και μηχανών, ανάλυση σχεδιασμού χημικών προϊόντων, δυναμική προσομοίωση συστημάτων χημικών διεργασιών, σχεδιασμός και διοίκηση.
8. **Ιατρική:** Ανάλυση καρκινικών κυττάρων, ανάλυση Ηλεκτροεγκεφαλογραφήματος και Ηλεκτροκαρδιογραφήματος, βελτιστοποίηση χρόνου νοσηλείας, μείωση νοσοκομειακού κόστους, βελτίωση ποιότητας νοσοκομείων.
9. **Γεωλογικές έρευνες:** Εντοπισμός πετρελαίου και φυσικού αερίου.
10. **Ρομποτική:** Έλεγχος τροχιάς και σύστημα όρασης ρομπότ.
11. **Επεξεργασία φωνής:** Αναγνώριση φωνής, συμπίεση φωνής, σύνθεση φωνής από κείμενο.
12. **Χρηματιστηριακές εφαρμογές:** Ανάλυση αγοράς, πρόβλεψη τιμών μετοχών.
13. **Τηλεπικοινωνίες:** Συμπίεση εικόνας και δεδομένων, αυτοματοποιημένες υπηρεσίες πληροφοριών, μετάφραση πραγματικού χρόνου, συστήματα επεξεργασίας πληρωμών.
14. **Μεταφορές:** Συστήματα διάγνωσης βλαβών φρένων, χρονοπρογραμματισμός οχημάτων, συστήματα δρομολόγησης.

Από το παραπάνω πλήθος εφαρμογών, μπορούμε να συμπεράνουμε ότι τα νευρωνικά δίκτυα έχουν αναπτυχθεί ταχύτατα. Γι' αυτό το λόγο έχουν επενδυθεί πολλά χρήματα τόσο για την ανάπτυξη λογισμικού όσο και για την ανάπτυξη υλικού.

Κεφάλαιο 5

Ανάπτυξη Αλγορίθμων Νευρωνικών Δικτύων σε Περιβάλλον Matlab

5.1 Γενικά

Το όνομα της γλώσσας προγραμματισμού MATLAB προήλθε από τις λέξεις Matrix Laboratory και αποτελεί ένα σύστημα επεξεργασίας πινάκων και των συναρτήσεων τους για εφαρμογές αριθμητικής ανάλυσης και γραφικής παρουσίασης (Mathworks).

Τα βασικά χαρακτηριστικά της MATLAB είναι:

- Η επίλυση αριθμητικών προβλημάτων μικρού και μεσαίου μεγέθους, δίχως προγραμματισμό σε συμβατικές γλώσσες.
- Η γρήγορη ανάπτυξη και δοκιμή αλγορίθμων (πλήθος έτοιμων συναρτήσεων και απλουστευμένη αλγοριθμική γλώσσα).
- Η παρουσίαση αποτελεσμάτων σε γραφική μορφή.

Οι μήτρες και οι πίνακες γενικά αποτελούν τον πυρήνα του MATLAB, δεδομένου ότι όλα τα δεδομένα στο MATLAB αποθηκεύονται ως πίνακες. Εκτός από τις συνηθισμένες πράξεις της άλγεβρας μητρώων, παρέχεται η δυνατότητα επεξεργασίας συνόλων δεδομένων με διάφορους τρόπους. Σημαντικό είναι επίσης το ότι υπάρχουν εργαλεία διασύνδεσης με τον χρήστη μέσω γραφικών (graphical user interface – GUI) τα οποία αποτελούν εργαλεία ανάπτυξης. Ο συνδυασμός δομών δεδομένων πίνακα, δυνατοτήτων προγραμματισμού και εργαλείων GUI διαμορφώνουν το MATLAB ως ένα ισχυρό επιστημονικό εργαλείο (Αρναουτάκης, 2005).

Το MATLAB λειτουργεί επίσης ως διερμηνέας εντολών, οι οποίες δίνονται μέσω του παραθύρου εντολών της (MATLAB command window). Οι εντολές αυτές μπορεί να είναι:

- Ορισμοί μεταβλητών και πράξεις.
- Κλήση ενσωματωμένων συναρτήσεων της MATLAB και των εγκατεστημένων εργαλειοθηκών της (toolboxes).
- Κλήση συναρτήσεων (functions) ή αρχείων εντολών MATLAB (scripts) που κατασκευάζονται από τους χρήστες.

Το MATLAB (Math Works Inc.) παρέχει ένα δυναμικό, εύχρηστο και ανοικτό υπολογιστικό περιβάλλον για υλοποίηση επιστημονικών εφαρμογών σε ένα μεγάλο φάσμα πεδίων, όπως στη

Γραμμική Άλγεβρα, Στατιστική, Εφαρμοσμένα Μαθηματικά, Αριθμητική Ανάλυση και Επιστημονικό Υπολογισμό, Επεξεργασία Σημάτων και Εικόνας, Θεωρία Ελέγχου, Θεωρία Βελτιστοποίησης και Γραφικά. Έχει υλοποιηθεί σε πολλές λειτουργικές πλατφόρμες (όπως Windows, Macintosh OS και Unix) και δύο βασικές εκδόσεις, την επαγγελματική και την εκπαιδευτική (student edition).

Το περιβάλλον του MATLAB υποστηρίζει ένα μεγάλο αριθμό ενδογενών λειτουργιών και συναρτήσεων καθώς και εξωτερικές βιβλιοθήκες (Toolboxes) για εξειδικευμένες περιοχές εφαρμογών. Υποστηρίζει επίσης μια ευέλικτη, απλή και δομημένη γλώσσα προγραμματισμού (script language) με πολλές ομοιότητες με την Pascal και παρέχει δυνατότητες εύκολης δημιουργίας, διασύνδεσης και χρήσης βιβλιοθηκών σε κώδικα γραμμένο στη γλώσσα αυτή (M files).

Αξιοσημείωτο είναι το ότι υπάρχει η δυνατότητα εκτέλεσης απλών μαθηματικών υπολογισμών μέχρι και προγράμματα με εντολές παρόμοιες με αυτές που υποστηρίζει μια γλώσσα υψηλού επιπέδου. Συγκεκριμένα το MATLAB εκτελεί απλές μαθηματικές πράξεις, αλλά εξίσου εύκολα χειρίζεται μιγαδικούς αριθμούς, δυνάμεις, ειδικές μαθηματικές συναρτήσεις, πίνακες, διανύσματα και πολυώνυμα. Μπορεί επίσης να αποθηκεύει και να ανακαλεί δεδομένα, να δημιουργεί και να εκτελεί ακολουθίες εντολών που αυτοματοποιούν διάφορους υπολογισμούς και να σχεδιάζει γραφικά.

Οι λειτουργίες του MATLAB διακρίνονται στις τυποποιημένες, δηλαδή σε αυτές που χειρίζονται αριθμητικά δεδομένα και εξάγουν αριθμητικά αποτελέσματα, και στις συναρτήσεις του Symbolic Toolbox, οι οποίες χειρίζονται και υπολογίζουν συμβολικές εκφράσεις, δηλαδή επεξεργάζονται μαθηματικά σύμβολα.

Τα πλεονεκτήματα του λογισμικού MATLAB είναι τα εξής:

- Ευκολότερη εκμάθηση από οποιαδήποτε άλλη γλώσσα προγραμματισμού.
- Βελτιστοποιημένος κώδικας για διεξαγωγή υπολογισμών με πίνακες.
- Αποτελεί μία γλώσσα προγραμματισμού για ανάπτυξη εφαρμογών και ταυτόχρονα ένα λογισμικό υλοποίησης επιστημονικών υπολογισμών.
- Δυνατότητα εύκολου εντοπισμού και διόρθωσης λαθών.
- Διαθέτει ένα φιλικό περιβάλλον επικοινωνίας με τον χρήστη.

Σαφώς προκύπτουν και ορισμένα μειονεκτήματα:

- Αποτελεί μία εξειδικευμένη γλώσσα προγραμματισμού.
- Το λογισμικό MATLAB αναπτύχθηκε μόνο για διεξαγωγή επιστημονικών υπολογισμών και έτσι δεν υποστηρίζει την ανάπτυξη άλλου είδους εφαρμογών, όπως για παράδειγμα επεξεργασία κειμένου.
- Οι αναπτυσσόμενες εφαρμογές υστερούν σε απόδοση από την άποψη του χρόνου εκτέλεσης σε σχέση με αντίστοιχες που αναπτύσσονται με τις κλασσικές γλώσσες προγραμματισμού.

5.2 Matlab και Τεχνητά Νευρωνικά Δίκτυα

Τα τεχνητά νευρωνικά δίκτυα, εμπνευσμένα καθαρά από το ανθρώπινο βιολογικό νευρικό σύστημα, δομούνται από τους νευρώνες. Η βασική λειτουργία που επιτελεί ένας νευρώνας είναι η συσσώρευση των σημάτων που δέχεται από τους νευρώνες με τους οποίους συνδέεται η είσοδος του, το φιλτράρισμα και η ενίσχυση αυτών των σημάτων και η παραγωγή ενός σήματος εξόδου, το οποίο στη συνέχεια μεταδίδεται μέσω των συνάψεων προς τους νευρώνες με τους οποίους συνδέεται η έξοδος του. Τα συναπτικά βάρη των νευρώνων ουσιαστικά ορίζουν τη λειτουργία του δικτύου, καθώς έχουν τη δυνατότητα να ρυθμίζονται αυτόματα κατά την εκπαίδευση του δικτύου, σύμφωνα με τον κανόνα εκμάθησης που χρησιμοποιείται.

Τα τεχνητά νευρωνικά δίκτυα εκπαιδεύονται με σκοπό να διευκολύνουν την εκτέλεση πολύπλοκων εφαρμογών σε διάφορα ερευνητικά πεδία, όπως η αναγνώριση προτύπων, η ταυτοποίηση, η ταξινόμηση και τα συστήματα ελέγχου, κυρίως όμως συμβάλλουν στην επίλυση προβλημάτων τα οποία είναι ανέφικτο να απαντηθούν είτε από τους συμβατικούς υπολογιστές είτε από τον άνθρωπο. Όπως ειπώθηκε προηγουμένως, το περιβάλλον του MATLAB έχει τη δυνατότητα να υποστηρίζει ένα μεγάλο αριθμό από εξωτερικές βιβλιοθήκες (Toolboxes) για εξειδικευμένες περιοχές εφαρμογών, όπως είναι τα τεχνητά νευρωνικά δίκτυα. Η συμβολή της συγκεκριμένης βιβλιοθήκης είναι απόλυτα σημαντική καθώς επιτρέπει την χρήση γραφικών εργαλείων για την εκπαίδευση των νευρωνικών δικτύων. Σημαντικό είναι επίσης και το γεγονός ότι η Matlab προσφέρει έτοιμες δομές για τον σχεδιασμό και τη διαμόρφωση των νευρωνικών δικτύων, απλοποιώντας οποιαδήποτε διαδικασία πολύπλοκων υπολογισμών με αποτέλεσμα ο χρήστης να είναι σε θέση να μελετήσει με σαφήνεια την συμπεριφορά και τις εφαρμογές τους.

5.2.1 Υποστηριζόμενες αρχιτεκτονικές Τεχνητών Νευρωνικών Δικτύων από το toolbox της Matlab

Υπάρχουν δύο βασικές κατηγορίες στις οποίες διαχωρίζονται τα τεχνητά νευρωνικά δίκτυα, ανάλογα με τη μέθοδο που ακολουθείται κατά την εκπαίδευση του δικτύου. Αυτές είναι: α) τα Ελεγχόμενα Νευρωνικά Δίκτυα και β) τα Ανεπιτήρητα Νευρωνικά Δίκτυα. Το toolbox της Matlab για τα Τεχνητά Νευρωνικά Δίκτυα είναι σε θέση να υποστηρίζει τις παραπάνω κατηγορίες και πιο συγκεκριμένα τις αρχιτεκτονικές δομές που τους αντιστοιχούν, όπως θα αναφερθούν παρακάτω (Mathworks, tutorial).

Επομένως, όσον αφορά τα Ελεγχόμενα Νευρωνικά Δίκτυα, το toolbox της Matlab επεξεργάζεται:

- *Δίκτυα Προώθησης* (feedforward networks), στα οποία τα δεδομένα ρέουν αυστηρά και μόνο από τις μονάδες εισόδου προς τις μονάδες εξόδου. Τα υποστηριζόμενα από τη Matlab δίκτυα Προώθησης περιλαμβάνουν: τα δίκτυα προώθησης με ανάδραση (backpropagation), τα σειριακά δίκτυα προώθησης με ανάδραση (cascade-forward), τα δίκτυα προώθησης με ανάδραση και επιβράδυνση εισόδου (feedforward input-delay), τα γραμμικά δίκτυα (linear) και τα δισαισθητικά δίκτυα (perceptron).

- *Δίκτυα Ακτινικής Συνάρτησης Βάσης* (Radial Basis Networks), με αρχιτεκτονικές δικτύων: γενικευμένης παλινδρόμησης (generalized regression) και στοχαστικών νευρωνικών δικτύων (probabilistic neural networks).
- *Περιοδικά Δίκτυα* (Reccurent Networks), όπως τα δίκτυα Elman και Hopfield, τα οποία χρησιμοποιούν τη μέθοδο της ανάδρασης προκειμένου να αναγνωρίζουν χωρικά και χρονικά μεταβαλλόμενα πρότυπα.

Τέλος, το toolbox της Matlab υποστηρίζει δύο τύπους Ανεπιτήρητων Νευρωνικών Δικτύων:

- *Δίκτυα με ανταγωνιστικά στρώματα* (competitive layers), τα οποία αναγνωρίζουν και ομαδοποιούν τα παρεμφερή διανύσματα εισόδου των δεδομένων, ταξινομώντας αυτόματα τις εισόδους τους σε κατηγορίες.
- *Αυτορυθμιζόμενοι χάρτες* (self-organizing maps). Με τον τύπο αυτό ταξινομούνται τα διανύσματα εισόδου βάσει της ομοιότητάς τους.

5.2.2 Η αρχιτεκτονική των Δικτύων Προώθησης με Ανάδραση (Backpropagation Feed-forward Networks) σε περιβάλλον Matlab

Στην παρούσα εργασία επιλέχθηκε η χρήση της γλώσσας προγραμματισμού Matlab και η βιβλιοθήκη των Νευρωνικών Δικτύων που παρέχεται. Οι αλγόριθμοι υλοποίησης των νευρωνικών δικτύων δημιουργήθηκαν βάσει της αρχιτεκτονικής των Δικτύων Προώθησης με Ανάδραση (Backpropagation Feed-Forward Networks).

Υπάρχουν γενικά τέσσερα βήματα για τη διαδικασία εκπαίδευσης των Δικτύων Προώθησης με Ανάδραση:

1. Διαμόρφωση των δεδομένων εκπαίδευσης.
2. Δημιουργία της οντότητας του δικτύου.
3. Εκπαίδευση του δικτύου.
4. Προσομοίωση του δικτύου έτσι ώστε να ανταποκρίνεται σε νέες εισόδους.

Τα δίκτυα προώθησης με ανάδραση (πολυστρωματικά δίκτυα) συχνά χρησιμοποιούν: τη λογαριθμική-σιγμοειδή συνάρτηση μεταφοράς (log-sigmoid transfer fuction/ logsig), την υπερβολική εφαιπτομένη-σιγμοειδής συνάρτηση μεταφοράς (tan-sigmoid transfer fuction/ tansig) και τη γραμμική συνάρτηση μεταφοράς (linear/ purelin). Η επιλογή εξαρτάται από τη διακύμανση των επιθυμητών τιμών εξόδου σε κάθε στρώμα του δικτύου. Υπενθυμίζουμε ότι κάθε επίπεδο οφείλει να κατέχει την συνάρτηση μεταφοράς που του αντιστοιχεί σύμφωνα με τα στοιχεία που το διαμορφώνουν. Κατόπιν θα αναφερθούν αναλυτικά οι αναγκαίες δομικές συναρτήσεις της Matlab για τη δημιουργία Νευρωνικών Δικτύων με Ανάδραση.

5.2.2.1 Συνάρτηση δημιουργίας δικτύου

Το πρώτο βήμα για την εκπαίδευση ενός νευρωνικού δικτύου είναι η δημιουργία της οντότητας του δικτύου (network object). Αρχικά καλείται κάποια από τις τυποποιημένες συναρτήσεις της Matlab, όπου στην περίπτωση των νευρωνικών δικτύων με ανάδραση, οι πιο ευρέως χρησιμοποιούμενες είναι οι :

newff	δημιουργία Τεχνητού Νευρωνικού Δικτύου Προώθησης (feed-forward neural network) προς τα εμπρός, που εκπαιδεύεται βάσει του αλγορίθμου της ανάδρασης (backpropagation).
newcf	δημιουργία Τεχνητού Νευρωνικού Δικτύου Προώθησης με ρυθμό αλλαγής των συναπτικών βαρών και των πολώσεων μεγαλύτερο από εκείνο που παρέχει η newff (cascade-forward neural network), και το οποίο εκπαιδεύεται βάσει του αλγορίθμου της ανάδρασης (backpropagation).
newelm	δημιουργία Τεχνητού Νευρωνικού Δικτύου που χρησιμοποιεί τον αλγόριθμο Elman προς τα εμπρός, που εκπαιδεύεται, επίσης, βάσει του αλγορίθμου της ανάδρασης (backpropagation).
newrb	δημιουργία Τεχνητού Νευρωνικού Δικτύου Ακτινικής Συνάρτησης Βάσης.
newgrnn	δημιουργία Τεχνητού Νευρωνικού Δικτύου Γενικευμένης Παλινδρόμησης, που είναι επέκταση των Τεχνητών Νευρωνικών Δικτύων Ακτινικής Συνάρτησης Βάσης.
newlvq	δημιουργία Τεχνητού Νευρωνικού Δικτύου Κβαντοποίησης των δεδομένων κατά την εκμάθηση.
network	δημιουργία Τεχνητού Νευρωνικού Δικτύου που παραμετροποιείται και καθορίζεται από τις ανάγκες του χρήστη (customized neural network).

Πίνακας 5.1: Συναρτήσεις δημιουργίας νευρωνικού δικτύου

Οι παραπάνω συναρτήσεις απαιτούν τα παρακάτω ορίσματα:

- PR:** Πίνακας $R \times 2$ διαστάσεων που περιέχει τις ελάχιστες και τις μέγιστες τιμές των R διανυσμάτων εισόδου.
- [S1 S2 S3 ... Sn]:** Πίνακας που περιέχει των αριθμό S_i των νευρώνων κάθε στρώματος του δικτύου
- {TF1 TF2 TF3 ... TFn}:** Δηλώνεται η συνάρτηση ενεργοποίησης (transfer function) για κάθε στρώμα (layer) του δικτύου.
- BTF:** Δηλώνεται η συνάρτηση εκπαίδευσης του δικτύου με ανάδραση (backpropagation training function). (Ας σημειωθεί ότι οι συναρτήσεις που μπορούν να κληθούν για την εκπαίδευση του

δικτύου, δηλαδή οι συναρτήσεις που μπορεί να πάρει σαν τιμή το όρισμα BTF, παίζουν καθοριστικό ρόλο στον τρόπο που θα γίνει η εκπαίδευση του δικτύου, αλλά και πόσο σωστά μπορεί να εκπαιδευτεί σύμφωνα με τα δεδομένα που του δίνει ο χρήστης).

5. **BLF**: Δηλώνεται η συνάρτηση εκμάθησης των συναπτικών βαρών και των πολώσεων του δικτύου με ανάδραση (backpropagation learning function). (Ας σημειωθεί ότι με την συνάρτηση αυτή δηλώνεται από τον χρήστη ο τρόπος με τον οποίο το δίκτυο «μαθαίνει» κατά την διάρκεια της εκπαίδευσής του).
6. **PF**: Δηλώνεται η συνάρτηση απόδοσης σφάλματος ή performance function που χρησιμοποιεί το δίκτυο κατά την διαδικασία της ανάδρασης.

5.2.2.2 Συνάρτηση αρχικοποίησης των συναπτικών βαρών και των πολώσεων του δικτύου

Η διαδικασία της αρχικοποίησης των συναπτικών βαρών και των πολώσεων του δικτύου κρίνεται αυστηρά απαραίτητη πριν από την εκπαίδευση του δικτύου. Ουσιαστικά κατά το στάδιο αυτό αρχικοποιούνται οι ελεύθεροι παράμετροι του δικτύου (συναπτικά βάρη και πολώσεις) σε μικρές τυχαίες τιμές, ομοιόμορφα κατανεμημένες σε ένα μικρό διάστημα τιμών γύρω από το 0, π.χ. [-0.5, 0.5]. Στο σημείο αυτό πρέπει να επισημανθεί ότι η εντολή `newff`, όπως και όλες οι συναρτήσεις δημιουργίας της οντότητας του δικτύου, αρχικοποιούν αυτόματα τα βάρη, αλλά επιτρέπεται η επανάληψη της αρχικοποίησης μέσω κατάλληλης εντολής από τη Matlab. Αυτό επιτυγχάνεται μέσω της εντολής `init`, η οποία δέχεται σαν όρισμα την ονομασία της οντότητας του δικτύου (network object) και το επιστρέφει με αρχικοποιημένες τις τιμές των συναπτικών βαρών και των πολώσεων.

5.2.2.3 Συνάρτηση εκπαίδευσης και συνάρτηση εκμάθησης

Απαραίτητα στοιχεία για την έναρξη της εκπαίδευσης του δικτύου είναι η ύπαρξη ζευγών διανυσμάτων εισόδου και επιθυμητών εξόδων. Σε κάθε επανάληψη της εκπαίδευσης πραγματοποιείται μία ρυθμιστική μετατροπή των συναπτικών βαρών και των πολώσεων, στοχεύοντας στην μείωση της συνάρτησης σφάλματος, έως ότου προσεγγίσει μία ικανοποιητική τιμή. Συνήθως στα Τεχνητά Νευρωνικά Δίκτυα Προώθησης με ανάδραση η συνάρτηση απόδοσης σφάλματος που χρησιμοποιείται είναι το μέσο τετραγωνικό σφάλμα (mse). Ουσιαστικά πρόκειται για το μέσο τετραγωνικό σφάλμα της διαφοράς μεταξύ των διανυσμάτων εισόδου και των διανυσμάτων εξόδου που τους αντιστοιχούν.

Οι συναρτήσεις εκπαίδευσης συνοδεύονται και από ορισμένες παραμέτρους οι τιμές των οποίων είναι προκαθορισμένες από το πρόγραμμα, αλλά ο χρήστης έχει τη δυνατότητα μεταβολής τους, προκειμένου να τους δώσει τιμές που τον συμφέρουν ως προς την εκτέλεση του αλγορίθμου. Τέτοιες παράμετροι είναι οι παρακάτω: `epochs`, `show`, `goal`, `time`, `min_grad`, `max_fail`, `lr` κ.α. Αξίζει να σημειωθεί ότι κάθε συνάρτηση εκπαίδευσης συνοδεύεται και από διαφορετικές παραμέτρους ανάλογα με τη φύση της και τη λειτουργία της. Η διαδικασία της εκπαίδευσης λαμβάνει τέλος στις εξής περιπτώσεις: όταν ο αριθμός των επαναλήψεων υπερβεί την τιμή που έχει δοθεί στην

παράμετρο epochs, όταν η συνάρτηση απόδοσης πέφτει κάτω από την τιμή της παραμέτρου goal, όταν η κλίση είναι μικρότερη της τιμής min_grad, ή όταν ο χρόνος εκπαίδευσης είναι μεγαλύτερος από εκείνον που έχει προκαθοριστεί.

Οι υποστηριζόμενες από την Matlab συναρτήσεις για εκπαίδευση (training) και εκμάθηση (learning) των Νευρωνικών Δικτύων είναι μαθηματικές διαδικασίες που χρησιμοποιούνται για να ρυθμίζουν αυτόματα τα συναπτικά βάρη και τις πολώσεις του δικτύου. Οι συναρτήσεις εκπαίδευσης υπαγορεύουν έναν γενικό αλγόριθμο που επηρεάζει όλα τα συναπτικά βάρη και τα πολώσεις του Νευρωνικού Δικτύου, ενώ οι συναρτήσεις εκμάθησης μπορούν να εφαρμοστούν και σε ξεχωριστά συναπτικά βάρη και πολώσεις μέσα σε ένα δίκτυο. Στη συνέχεια παρατίθεται πίνακας με τις υποστηριζόμενες συναρτήσεις εκπαίδευσης και εκμάθησης από τη Matlab (Mathworks).

▪ Υποστηριζόμενες συναρτήσεις εκπαίδευσης:

trainb : Εκπαίδευση σε δεσμίδες δεδομένων με κανόνες εκμάθησης των συναπτικών βαρών και των bias (batch training with weight and bias learning rules)	trainbfg : BFGS quasi-Newton συνάρτηση εκπαίδευσης με ανάδραση (backpropagation)
trainbr : Bayesian regularization	trainc : Κυκλικής διαδοχής αυξητική ενημέρωση (cyclical order incremental update)
traincgb : Powell-Beale με σύζευξη συντελεστή και ανάδραση	traincgf : Fletcher-Powell με σύζευξη συντελεστή και ανάδραση
traincgp : Polak-Ribiere με σύζευξη συντελεστή και ανάδραση	traingd : Ελαττούμενος συντελεστής και ανάδραση
traingda : Ελαττούμενος συντελεστής με προσαρμόσιμο ρυθμό εκμάθησης και ανάδραση	traingdm : Ελαττούμενος συντελεστής με ορμή και ανάδραση
traingdx : Ελαττούμενος συντελεστής με ορμή & προσαρμόσιμη γραμμική ανάδραση	trainlm : Levenberg-Marquardt ανάδραση
trainoss : Αντίστροφου συνημίτονου με μοναδιαίο βήμα ανάδραση	trainr : Τυχαίας σειράς αυξητικής ενημέρωσης
trainrp : Ελαστική ανάδραση (resilient backpropagation (Rprop))	train : Σειριακή αυξητική ενημέρωση

Πίνακας 5.2: Συναρτήσεις εκπαίδευσης νευρωνικού δικτύου

▪ Υποστηριζόμενες συναρτήσεις εκμάθησης:

learncon : Συνάρτηση εκμάθησης των bias με «συνείδηση»	learngd : Ελαττούμενου συντελεστή weight/bias συνάρτηση εκμάθησης
learnngdm : Ελαττούμενου συντελεστή με ορμή weight/bias συνάρτηση εκμάθησης	learnh : Hebb weight συνάρτηση εκμάθησης
learnhd : Hebb με εξασθένηση των weights κανόνας εκμάθησης	learnis : Συνάρτηση εκμάθησης των weights με ενδιάμεσες φάσεις
learnk : Kohonen συνάρτηση εκμάθησης των weights	learnlv1 : LVQ1 συνάρτηση εκμάθησης των weights
learnlv2 : LVQ2 συνάρτηση εκμάθησης των weights	learnos : Outstar συνάρτηση εκμάθησης των weights
learnp : Perceptron συνάρτηση εκμάθησης των weights και των bias	Learnpn : Κανονικοποιημένη perceptron συνάρτηση εκμάθησης των weights και bias
learnsom : Συνάρτηση εκμάθησης αυτορυθμιζόμενων «χαρτών»	learnwh : Widrow-Hoff κανόνας εκμάθησης των weights και των bias

Πίνακας 5.3: Συναρτήσεις εκμάθησης νευρωνικού δικτύου

5.2.2.4 Συνάρτηση εξομοίωσης

Μετά το πέρας της εκπαίδευσης του νευρωνικού δικτύου έχει σειρά η εξομοίωση, μέσω της κλήσης της συνάρτησης **sim** της Matlab. Το όρισμα της συνάρτησης εξομοίωσης είναι πάντα ένα διάνυσμα εισόδου, το οποίο είτε έχει χρησιμοποιηθεί στην εκπαίδευση του δικτύου, είτε αποτελεί ένα νέο στοιχείο στο δίκτυο. Τέλος, το network object επιστρέφει την έξοδο του δικτύου.

5.2.2.5 Το πρόβλημα της υπερεκπαίδευσης (Overfitting)

Το πρόβλημα αυτό συναντάται όταν το σφάλμα στο σετ της εκπαίδευσης τείνει σε μία πολύ μικρή τιμή, αλλά με την εισαγωγή νέων δεδομένων στο δίκτυο το σφάλμα λαμβάνει μεγάλες τιμές. Πρόκειται λοιπόν για μία κατάσταση κατά την οποία το δίκτυο μπορεί να έχει απομνημονεύσει σωστά τα αποτελέσματα που προκύπτουν από την φάση της εκπαίδευσης, αλλά δεν έχει μάθει να γενικεύει σωστά πάνω στα καινούργια δεδομένα εισόδου. Για τον λόγο αυτό, παρουσιάζεται μικρή τιμή σφάλματος πάνω στις εξόδους κατά τη διαδικασία της εκπαίδευσης, αλλά και ένα αρκετά σημαντικό λάθος στις εξόδους κατά την είσοδο στο δίκτυο νέων δεδομένων.

Για την πρόληψη της παραπάνω κατάστασης τίθενται δύο λύσεις από τη Matlab:

1. **Η μέθοδος της κανονικοποίησης:** μέσω της χρήσης της συνάρτησης `msereg` ως συνάρτηση σφάλματος του δικτύου κατά την εκπαίδευση. Η συγκεκριμένη συνάρτηση εκτός από το σφάλμα των ελαχίστων τετραγώνων, περιλαμβάνει και το μέγεθος των αλλαγών στα συναπτικά βάρη και τις πολώσεις του δικτύου.
2. **Η μέθοδος του “early stopping”:** κατά την οποία τα δεδομένα που πρόκειται να εισαχθούν στο δίκτυο για εκπαίδευση χωρίζονται αυτόματα μέσω εντολών της Matlab σε τρεις ομάδες. Με τον τρόπο αυτό προκύπτουν τα εξής σετ δεδομένων: α) training set β) validation set και γ) test set. Είναι στην επιλογή του χρήστη το ποσοστό των συνολικών δεδομένων που θα αντιστοιχεί σε κάθε ομάδα. Τα δεδομένα του training και validation set, συμβάλλουν ουσιαστικά στο «στήσιμο» του δικτύου, ενώ κατά τη φάση του testing πραγματοποιείται μία εκτίμηση της απόδοσης του νευρωνικού δικτύου. Σημαντικό να αναφερθεί επίσης είναι το γεγονός ότι κατά το training τα συναπτικά βάρη και οι πολώσεις υφίστανται διορθώσεις ανά επανάληψη, ενώ κατά τη φάση του validation, διακόπτεται η διαδικασία της εκπαίδευσης μόλις παρουσιαστεί το πρόβλημα του overfitting (Mathworks, tutorial).

Κεφάλαιο 6

Βιβλιογραφική Ανασκόπηση

6.1 Εισαγωγή

Στον τομέα της επιφανειακής υδρολογίας, η θεώρηση των μοντέλων μαύρου κουτιού, είναι αρκετά διαδεδομένη. Ίσως ένα από τα πιο γνωστά εμπειρικά μοντέλα μαύρου κουτιού αποτελεί το μοναδιαίο υδρογράφημα, καθώς και τα μοντέλα που στηρίζουν τις αρχές τους στην θεωρία του (Sherman, 1932, Nash, 1957). Επίσης τα μοντέλα μαύρου κουτιού στην υδρολογία εμπλέκουν παραδοσιακές στατιστικές μεθόδους ανάλυσης, όπως τεχνικές γραμμικής παλινδρόμησης και συσχέτισης, που χρησιμοποιούνται για να καθορίσουν σχέσεις μεταξύ διαφορετικών ομάδων δεδομένων. Χαρακτηριστικό παράδειγμα τέτοιου τύπου μοντέλου αποτελεί το ARIMA (Autoregressive Integrated Moving Average), το οποίο χρησιμοποιείται για τον προσδιορισμό της σχέσης ανάμεσα στην βροχόπτωση και την απορροή. Ωστόσο στην υπόγεια υδραυλική η χρήση συστημάτων μέσω σύνδεσης δεδομένων εισόδου και εξόδου δεν έχει αναπτυχθεί σε ανάλογο βαθμό (Abbott and Refsgaard, 1996).

Στην υδρολογία, ως μοντέλα μαύρου κουτιού αναφέρονται και τα νευρωνικά δίκτυα, που διατυπώνουν έντονα μη γραμμικούς μετασχηματισμούς των μεταβλητών εισόδου σε εξόδους. Η δομή (αριθμός νευρώνων και κρυμμένων επιπέδων) και οι παράμετροί τους δεν έχουν φυσικό υπόβαθρο, και προκύπτουν μέσω μιας αυτοματοποιημένης διαδικασίας προσαρμογής, που βασίζεται στη χρήση γενετικών, συνήθως, αλγορίθμων και είναι γνωστή ως εκπαίδευση (training) του δικτύου. Στην πράξη, το νευρωνικό δίκτυο είναι ένας κρυφός υπολογιστικός κώδικας, στον οποίο δεν έχει πρόσβαση ο χρήστης.

Στα μοντέλα μαύρου κουτιού (black-box), οι εξισώσεις δεν αναπαριστούν τις ενδιάμεσες διεργασίες τον υδρολογικού κύκλου, όπως συμβαίνει με τα εννοιολογικά και τα φυσικής βάσης μοντέλα. Το πλεονέκτημά τους είναι η δυνατότητα περιγραφής οσοδήποτε πολύπλοκων συστημάτων, με πολύ μικρότερο (σχεδόν αμελητέο) υπολογιστικό φόρτο σε σχέση με τα αναλυτικά μοντέλα προσομοίωσης.

Η διάδοση των νευρωνικών δικτύων στις υδρολογικές επιστήμες ξεκίνησε στις αρχές της δεκαετίας του 1990. Ο γενετικός προγραμματισμός έχει μικρό χρόνο ζωής στις επιστήμες των υδατικών πόρων, και ακόμα μικρότερη εφαρμογή στην προσομοίωση υδρολογικών διεργασιών.

6.2 Εφαρμογή Τεχνητών Νευρωνικών Δικτύων στην εκτίμηση της σχέσης βροχοπτώσεων - απορροής

Η διαδικασία εκτίμησης της απορροής συναρτήσει των βροχοπτώσεων αποτελεί μία εξαιρετικά πολύπλοκη διαδικασία, καθώς τα κύρια χαρακτηριστικά της είναι η μη-γραμμικότητα, αλλά και η χωρική - χρονική μεταβλητότητα. Κατά το παρελθόν έχει γίνει προσπάθεια προσομοίωσης της συγκεκριμένης διαδικασίας με χρήση διάφορων απλών μοντέλων, τα οποία όμως απέτυχαν να απεικονίσουν τη δυναμική της μη-γραμμικότητας στον μετασχηματισμό βροχόπτωσης – απορροής. Μία νέα διάσταση στην προσέγγιση του θέματος δόθηκε μέσω της υιοθέτησης της τεχνικής των τεχνητών νευρωνικών δικτύων με σκοπό να επιλυθούν ορισμένα σοβαρά ζητήματα υδρολογίας αλλά και διαχείρισης υδατικών πόρων γενικότερα. Σημαντικό είναι το γεγονός, ότι κατά τον καθορισμό της σχέσης εισόδου-εξόδου, τα νευρωνικά δίκτυα έχουν τη δυνατότητα να μαθαίνουν μέσω παραδειγμάτων χωρίς να λαμβάνονται υπόψη οι νόμοι της φυσικής (Rajurkara et al., 2004).

Η εφαρμογή των τεχνητών νευρωνικών δικτύων όσον αφορά την διερεύνηση της σχέσης βροχόπτωσης-απορροής ξεκίνησε μέσω μίας προκαταρκτικής μελέτης των Halff, Halff και Azmoodeh (1993). Στην συγκεκριμένη εργασία αναπτύχθηκε ένα τεχνητό νευρωνικό δίκτυο τριών επιπέδων, το οποίο προέβλεπε υδρογραφήματα, χρησιμοποιώντας ως δεδομένα υετογραφήματα (Halff et al., 1993).

Αργότερα, το 1995, οι Hsu, Gupta και Sorooshian πρότειναν ένα νέο αλγόριθμο, πιο αξιόπιστο και αποτελεσματικό από τον κανόνα πίσω διάδοσης λάθους (backpropagation), κατά την εφαρμογή του νευρωνικού δικτύου. Ο αλγόριθμος αυτός ονομάζεται linear least squares simplex και χρησιμοποιείται για την εκπαίδευση του δικτύου, μέσω συνδυασμού των μεθόδων linear least squares και multi – start simplex optimization. Ο στόχος ήταν η επίτευξη της επιτάχυνσης του χρόνου εκπαίδευσης. Το δίκτυο προώθησης αποτελούνταν από τρία επίπεδα και είχε τη δυνατότητα να προσομοιάσει τη μη-γραμμική υδρολογική συμπεριφορά των λεκανών απορροής, δίνοντας ικανοποιητικά αποτελέσματα. Εν τέλει πραγματοποιήθηκε σύγκριση των αποτελεσμάτων των νευρωνικών δικτύων με τα αντίστοιχα γραμμικών και εννοιολογικών μοντέλων (Hsu et al., 1995).

Ο Shamseldin (1997) εφήρμοσε την τεχνική των νευρωνικών δικτύων προώθησης στοχεύοντας την μοντελοποίηση βροχόπτωσης – απορροής χρησιμοποιώντας δεδομένα από έξι επιφάνειες απορροής. Δημιουργήθηκαν τέσσερα διαφορετικά σενάρια εισόδου, χρησιμοποιώντας τρεις διαφορετικούς τύπους δεδομένων. Πραγματοποιήθηκε μια σύγκριση των αποτελεσμάτων της τεχνικής των νευρωνικών δικτύων με τρία απλούστερα μοντέλα. Τα μοντέλα που επιλέχθηκαν είναι: το απλό γραμμικό μοντέλο, ένα seasonally based linear perturbation model και ένα nearest neighbour linear perturbation model (μοντέλο πλησιέστερου γείτονα). Τα αποτελέσματα έδειξαν ότι τα νευρωνικά δίκτυα έχουν προοπτικές εφαρμογής σε θέματα υδρολογίας, καθώς έχουν τη δυνατότητα εξαγωγής ακριβέστερων προβλέψεων σε σχέση με άλλα παραδοσιακά μοντέλα (Shamseldin, 1997).

Το 2003 οι Lallahem και Mania πραγματοποίησαν μία μελέτη, προσπαθώντας να προσομοιώσουν την παροχή του υπόγειου νερού σε καρστικά μέσα και να καθορίσουν τους παράγοντες επιρροής της απορροής. Για τον σκοπό αυτό κατέληξαν στην συνδυαστική χρήση ενός μαθηματικού μοντέλου ονόματι MERO και της τεχνικής των νευρωνικών δικτύων. Η περιοχή της μελέτης βρίσκεται στο βόρειο τμήμα του Παρισιού της Γαλλίας και αποτελείται κυρίως από καρστικά και κρητιδικά πετρώματα. Αρχικά χρησιμοποιήθηκε το μοντέλο MERO προκειμένου να καθοριστούν οι απαιτούμενοι παράγοντες, οι οποίοι στη συνέχεια εισήχθησαν στο νευρωνικό δίκτυο. Το νευρωνικό δίκτυο που δημιουργήθηκε υπέστη δοκιμασία ελέγχου μέσω τεσσάρων διαφορετικών σεναρίων, ενώ στην εκπαίδευση χρησιμοποιήθηκε ο κανόνας της πίσω διάδοσης λάθους (backpropagation). Ο συνδυασμός των μοντέλων MERO και νευρωνικών δικτύων επέφερε καλά αποτελέσματα τα οποία όμως επιδέχονται βελτίωση μέσω μείωσης των κρυφών επιπέδων του νευρωνικού δικτύου, έτσι ώστε να επιτευχθεί μείωση του μέσου τετραγωνικού σφάλματος (RMSE) (Lallahem and Mania, 2003).

Στην εργασία των Riad, Mania, Bouchaou και Najjar (2004) αναπτύχθηκε ένα νευρωνικό δίκτυο τύπου Perceptron, το οποίο χρησιμοποιήθηκε για συσχέτιση απορροής – βροχόπτωσης σε μία επιφάνεια απορροής στο Μαρόκο, όπου το κλίμα έχει χαρακτηριστικά μερικής ξηρότητας. Το συμπέρασμα από την συγκεκριμένη εφαρμογή ήταν ότι τα τεχνητά νευρωνικά δίκτυα είναι σε θέση να δώσουν ικανοποιητικά αποτελέσματα συσχέτισης απορροής βροχόπτωσης, ακόμα και σε συνθήκες ξηρασίας, οι οποίες αποτελούν μία αρκετά ιδιάζουσα περίπτωση λόγω μη-ομαλότητας των δεδομένων (Riad et al., 2004).

Οι Rajurkara, Kothyarib και Chaubec (2004) παρουσίασαν μία προσπάθεια μοντελοποίησης ημερησίων παροχών κατά τη διάρκεια επεισοδίων πλημμύρας μέσω συνδυαστικής χρήσης ενός απλού γραμμικού μοντέλου και της τεχνικής των νευρωνικών δικτύων. Για την εφαρμογή χρησιμοποιήθηκαν δεδομένα από δύο μεγάλες επιφάνειες απορροής σε διαφορετικές γεωγραφικές περιοχές και τα αποτελέσματα έδειξαν ότι πραγματοποιήθηκε μία ικανοποιητική συσχέτιση απορροής – βροχόπτωσης, αποδεικνύοντας έτσι την πολλαπλή χρησιμότητα της μεθόδου (Rajurkara et al., 2004).

Σε εργασία των Chen και Adams (2006), πραγματοποιήθηκε μία ενοποίηση των τεχνητών νευρωνικών δικτύων με συμβατικά μοντέλα. Η χωρική μεταβλητότητα των δεδομένων βροχόπτωσης, η ετερογένεια των χαρακτηριστικών των εξεταζόμενων λεκανών απορροής, καθώς και η επιρροή τους στην απορροή είναι δυνατόν να ξεπεραστούν μέσω του παραπάνω εγχειρήματος ενοποίησης μοντέλων. Η δυνατότητα επίτευξης της συγκεκριμένης προσέγγισης αποδείχτηκε μέσω αυτής της μελέτης και βασίστηκε στην επιλογή τριών διαφορετικών τύπων μοντέλων εκτίμησης βροχόπτωσης – απορροής (Xinanjiang, SMAR και Tank). Τα αποτελέσματα ήταν αρκετά ενθαρρυντικά για την καθιέρωση των τεχνητών νευρωνικών δικτύων σε προβλήματα τέτοιας φύσης (Chen and Adams, 2006).

Το 2007 οι Srinivasulu και Jain συνέκριναν διάφορες μεθόδους εκπαίδευσης νευρώνα τύπου Perceptron, όπως ο κλασικός κανόνας διάδοσης λάθους, ένας real-coded γενετικός αλγόριθμος και αυτό-οργανωμένοι χάρτες S.O.M. (Self-organized maps). Η τελευταία μέθοδος χρησιμοποιήθηκε προκειμένου να ταξινομηθεί ο χώρος εισόδου-εξόδου σε διαφορετικές κατηγορίες και κατόπιν να αναπτυχθούν νευρωνικά δίκτυα προώθησης, ένα για κάθε κατηγορία, μέσω του κανόνα πίσω διάδοσης λάθους. Η διαδικασία αυτή αποδείχθηκε ότι έδωσε καλύτερα αποτελέσματα από την παραδοσιακή προσέγγιση της ανάπτυξης ενός μοναδικού νευρωνικού δικτύου για μοντελοποίηση βροχόπτωσης- απορροής, μέσω εκπαίδευσης backpropagation. Το δίκτυο που χρησιμοποίησε τον real-coded γενετικό αλγόριθμο για την εκπαίδευση του, έδωσε μία καλύτερη γενίκευση και κρίθηκε ως αρκετά αξιόπιστο (Srinivasulu and Jain, 2007).

Οι Goswami και O'Connor (2007) συνδύασαν μοντέλα πρόβλεψης απορροής μέσω δύο διαφορετικών σεναρίων σε έναν ποταμό της Γαλλίας και σε έναν της Ιρλανδίας. Η ειδοποιός διαφορά μεταξύ των δύο σεναρίων είναι ότι στο πρώτο εφαρμόστηκαν μοντέλα αυτοσυσχέτισης και τεχνητών νευρωνικών δικτύων, ενώ στο δεύτερο έγινε χρήση δύο γραμμικών μοντέλων μεταφοράς και δύο μοντέλων τεχνητών νευρωνικών δικτύων. Τα δύο σενάρια βασίστηκαν σε χρονοσειρές δεδομένων απορροής και σε σειρές βροχόπτωσης και απορροής, αντίστοιχα. Έπειτα πραγματοποιήθηκε συνδυασμός των αποτελεσμάτων από κάθε σενάριο με τεχνικές της απλής μέσης τιμής, της σταθμισμένης μέσης τιμής και της τεχνικής των νευρωνικών δικτύων. Το επόμενο βήμα ήταν η σύγκριση των μοντέλων και των συνδυασμών τους. Για τη διαδικασία αυτή χρησιμοποιήθηκαν ο δείκτης Nash-Sutcliffe, το μέσο τετραγωνικό σφάλμα και μέσο σχετικό σφάλμα. Στην περίπτωση της λεκάνης της Γαλλίας, όσον αφορά το πρώτο σενάριο εκτιμήθηκε ότι τα μοντέλα αυτοσυσχέτισης απέδωσε περισσότερο από τα τεχνητά νευρωνικά δίκτυα, εφόσον πρόκειται για χρονικό διάστημα λίγων ημερών. Σε περιπτώσεις μεγαλύτερου χρονικού διαστήματος αποδείχτηκε ότι τα νευρωνικά δίκτυα λειτουργούν αποτελεσματικότερα. Αντίθετα, τα γραμμικά μοντέλα μεταφοράς του δεύτερου σεναρίου είχαν υψηλότερη απόδοση. Για τη λεκάνη της Ιρλανδίας τα τεχνητά νευρωνικά δίκτυα εκτιμήθηκαν ως περισσότερο αξιόπιστα και στα δύο σενάρια που εφαρμόστηκαν. Αξίζει να σημειωθεί ότι η λειτουργία και των δύο σεναρίων ήταν καλύτερη όταν χρησιμοποιήθηκαν δεδομένα βροχόπτωσης και απορροής σε σχέση με την χρήση μονάχα δεδομένων απορροής (Goswami and O'Connor, 2007).

Σε μία πολύ πρόσφατη μελέτη, οι Sedki, Ouazar και Mazoudi (2009) διερεύνησαν την επιρροή του συνδυασμού των γενετικών αλγορίθμων με τα νευρωνικά δίκτυα για πρόβλεψη της απορροής στην υδρολογική λεκάνη Ouirika, που βρίσκεται σε περιβάλλον μερικής ξηρασίας στο Μαρόκο. Για την εκπαίδευση του δικτύου, χρησιμοποιήθηκαν δεδομένα βροχόπτωσης και απορροής από τέσσερις προηγούμενες χρονικές περιόδους (2000-2003). Τα δεδομένα των ετών 2000 ως 2002 χρησιμοποιήθηκαν για την εκπαίδευση του δικτύου, ενώ τα υπόλοιπα για τον έλεγχο της απόδοσής του. Οι γενετικοί τελεστές, στη μέθοδο των γενετικών αλγορίθμων, βελτιστοποιούν το νευρωνικό δίκτυο, αποφεύγοντας φαινόμενα πρόωρης σύγκλισης. Τα αποτελέσματα ήταν αρκετά

ικανοποιητικά και το συμπέρασμα από την συγκεκριμένη μελέτη είναι ότι η προτεινόμενη εμπλοκή των γενετικών αλγορίθμων δίνει πολύ καλύτερα αποτελέσματα σε σχέση με την παραδοσιακή εφαρμογή των νευρωνικών δικτύων με κανόνα backpropagation. Μοναδικό πρόβλημα στην παραπάνω εφαρμογή υπήρξε ο καθορισμός της τοπολογίας του νευρωνικού δικτύου (Sedki et al., 2009).

Από τα παραπάνω συνεπάγεται ότι τα Τεχνητά Νευρωνικά Δίκτυα είναι σε θέση να παρέχουν αξιόπιστες προβλέψεις όσον αφορά την συσχέτιση των βροχοπτώσεων με την απορροή. Είναι σαφές ότι υπάρχει πληθώρα βιβλιογραφίας στο συγκεκριμένο αντικείμενο καθώς η απόδοση της συγκεκριμένης τεχνικής, ειδικά αν συνδυαστεί και με άλλα υπολογιστικά μοντέλα, είναι υψηλού επιπέδου.

6.3 Εφαρμογή Τεχνητών Νευρωνικών Δικτύων στα υπόγεια ύδατα

Σημαντική είναι η προσφορά των τεχνητών νευρωνικών δικτύων στον τομέα της ποιότητας των υπόγειων υδάτων, καθώς και σε ζητήματα προστασίας και εξυγίανσης. Η τεχνολογία των νευρωνικών δικτύων, εμπνευσμένη από νευροβιολογικές θεωρίες μαζικής αλληλοσύνδεσης και παραλληλισμού, έχει εφαρμοστεί τα τελευταία χρόνια σε πληθώρα προβλημάτων βελτιστοποίησης. Το 1994 οι Rogers και Dowla σε μελέτη τους αρχικά χρησιμοποίησαν ένα μοντέλο μεταφοράς και εκπάιδευσαν ένα νευρωνικό δίκτυο προσδοκώντας να προβλέψουν την παροχή αντλούμενου νερού. Κατόπιν χρησιμοποίησαν το ήδη εκπαιδευμένο δίκτυο προκειμένου να εντοπιστεί το βέλτιστο σύστημα άντλησης για επιτυχή εξυγίανση. Δημιουργήθηκαν διάφορα σύνθετα σενάρια με μεταβλητούς παράγοντες μεταφοράς πλουμίου και προσομοιώθηκαν μέσω της θεωρίας των πεπερασμένων διαφορών/πεπερασμένων στοιχείων όσον αφορά την παροχή και τη μεταφορά του πλουμίου. Η συγκέντρωση του ρύπου όφειλε να διατηρηθεί κάτω από τα επιτρεπόμενα όρια, αλλά και η συνολική ποσότητα του νερού που αντλείται έπρεπε να ελαχιστοποιηθεί στοχεύοντας την εξυγίανση. Το στρώμα εισόδου χαρακτηρίζεται από το σύστημα αντλήσεων, καθώς κάθε στοιχείο εισόδου αντιπροσώπευε το επίπεδο άντλησης του πηγαδιού. Το στρώμα εξόδου απέδιδε την επιτυχία ή αποτυχία της εξυγίανσης. Για την εκπαίδευση του δικτύου χρησιμοποιήθηκε ο αλγόριθμος της πίσω διάδοσης λάθους, ενώ άλλες μέθοδοι, όπως των συζυγών κλίσεων και εξάλειψης βαρών, προσέφεραν επιτάχυνση της σύγκλισης και βελτίωση της απόδοσης του δικτύου. Σημαντικό είναι επίσης το ότι μέσω των γενετικών αλγορίθμων διερευνήθηκε η επιτυχία ή όχι των σεναρίων άντλησης. Τα αποτελέσματα ήταν αρκετά ενθαρρυντικά όσον αφορά την εμπλοκή των τεχνητών νευρωνικών δικτύων σε προβλήματα υπόγειας ροής (Rogers and Dowla, 1994).

Αξίζει να σημειωθεί ότι δύο χρόνια πριν την παραπάνω μελέτη, δηλαδή το 1992, ο Rogers, σε δημοσίευση του είχε παρουσιάσει μία καινοτομική υπολογιστική προσέγγιση για βελτιστοποίηση της

εξυγίανσης των υπόγειων υδάτων, μέσω συνδυαστικής χρήσης των τεχνητών νευρωνικών δικτύων και των γενετικών αλγορίθμων (Rogers, 1992).

Όσον αφορά την διερεύνηση των παραμέτρων των υπόγειων νερών οι Aziz και Wong το 1992, χρησιμοποιώντας δεδομένα αντλήσεων από πηγάδια παρατήρησης, εκπαιδύσαν ένα νευρωνικό δίκτυο, έτσι ώστε να είναι δυνατή η πρόβλεψη παραμέτρων του υδροφορέα, όπως τιμές της μεταφορικότητας, αποθηκευτικότητας και του λόγου της απόστασης από το πηγάδι παρατήρησης προς το πάχος του υδροφορέα. Οι τιμές που εισήχθησαν στο δίκτυο υπολογίστηκαν μέσω των εξισώσεων του Theis και Hantasu - Jacob για περιβάλλον περιορισμένου υδροφορέα και περιορισμένου υδροφορέα με διαρροή. Στη συγκεκριμένη μελέτη εφαρμόζεται ουσιαστικά η μέθοδος της υπόγειας υδραυλικής: «αντίστροφο πρόβλημα» (κατά την οποία πραγματοποιείται εκτίμηση των παραμέτρων ενός προβλήματος, εάν είναι γνωστή η λύση) και σαφώς αποδεικνύεται η καλή απόδοση των τεχνητών νευρωνικών δικτύων σε εφαρμογή αναγνώρισης προτύπων (Aziz and Wong, 1992).

Επιπλέον μελέτες χρησιμοποίησαν τα τεχνητά νευρωνικά δίκτυα προκειμένου να δοθεί λύση στο αντίστροφο πρόβλημα (Zio, 1997). Λόγω της δυνατότητας των νευρωνικών δικτύων να εκπαιδεύονται βάσει πολύπλοκων σχέσεων εισόδου-εξόδου και να παράγουν αξιόλογα αποτελέσματα, επιλέχθηκαν ως υπολογιστικό εργαλείο ώστε να δώσουν μία εκτίμηση σε φυσικές παραμέτρους του υδροφορέα βασισμένες σε τιμές παρατήρησης στο σύστημα. Μέσω ενός αναλυτικού μοντέλου προσομοιώθηκε η συμπεριφορά της υπόγειας ροής, ενώ για την εκπαίδευση του δικτύου προώθησης πολλαπλών επιπέδων χρησιμοποιήθηκε η μέθοδος με επίβλεψη.

Μία ακόμη ενδιαφέρουσα μελέτη πραγματοποιήθηκε το 2003, στην οποία προτάθηκε η χρήση των νευρωνικών δικτύων στη διαχείριση λεκάνης απορροής, προσδοκώντας την εξισορρόπηση ποιότητας πόσιμου νερού και περιορισμών στη χρήση γεωργικής γης. Η μέθοδος που προτάθηκε ουσιαστικά προσομοίωσε ένα υδρολογικό μοντέλο, προκειμένου να προσδιορίσει τα κατάλληλα σενάρια χρήσης γης έτσι ώστε να διατηρηθούν ικανοποιητικά επίπεδα ποιότητας υπόγειου νερού (Kralisch et al., 2003)

Τα υπόγεια συστήματα ροής περιγράφονται από στοιχεία πολυπλοκότητας, μη-γραμμικότητας και χωρικής – χρονικής μεταβλητότητας. Ελέγχονται από φυσικούς ή και ανθρωπογενείς παράγοντες, οι οποίοι περιπλέκουν τις δυναμικές προβλέψεις. Σε περίπτωση που δεν είναι διαθέσιμη μια βάση αξιόπιστων δεδομένων, και η ακριβής πρόβλεψη κρίνεται απαραίτητη, τότε τα εμπειρικά μοντέλα αποτελούν μία υπολογίσιμη εναλλακτική μέθοδο και είναι σε θέση να παράγουν χρήσιμες προβλέψεις. Σύμφωνα με μελέτη που πραγματοποιήθηκε το 2005, αναπτύχθηκε ένα τεχνητό νευρωνικό δίκτυο, το οποίο εκπαιδεύτηκε με τον αλγόριθμο Levenberg – Marquardt στοχεύοντας να προβλέπει τιμές της στάθμης του υδροφόρου ορίζοντα στην περιοχή της Μεσσαράς που βρίσκεται κεντρικά – νότια της Κρήτης. Τα διανύσματα εισόδου στο δίκτυο ήταν η βροχόπτωση, η θερμοκρασία, ο όγκος νερού που απορρέει στο χείμαρρο Γεροπόταμο καθώς και οι τιμές της στάθμης ενός πηγαιδιού που υπάρχει στην περιοχή. Με το μοντέλο που δημιουργήθηκε

πραγματοποιήθηκαν προβλέψεις 18 μηνών μπροστά από τα δεδομένα συλλογής, καταλήγοντας στο συμπέρασμα ότι η πεδιάδα της Μεσσαράς, αντιμετωπίζει σοβαρό πρόβλημα εξάντλησης των υδάτινων πόρων (Daliakopoulos et al., 2005).

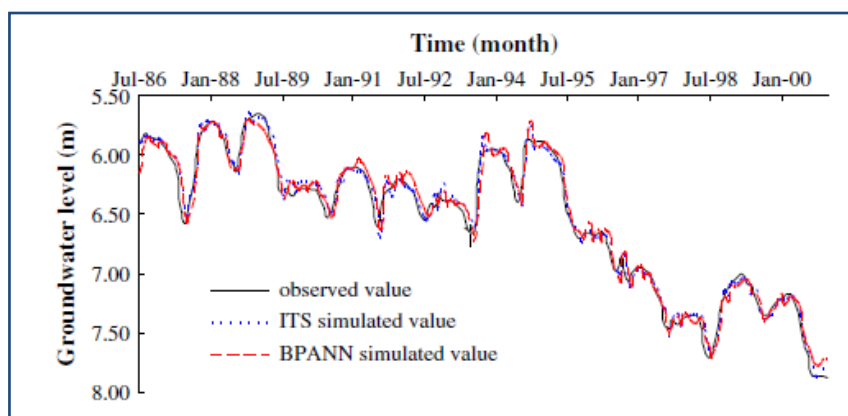
Τα τεχνητά νευρωνικά δίκτυα είναι δυνατόν να χρησιμοποιηθούν στην μοντελοποίηση προβλημάτων στα οποία η συσχέτιση εξαρτημένων και ανεξάρτητων μεταβλητών είναι σχετικά αδύναμη, καθώς αποτελούν ένα γενικής χρήσης εργαλείο προσέγγισης. Στην μελέτη των Garcia και Shigidi διερευνήθηκε η ικανότητα των νευρωνικών δικτύων να παρέχουν μία προσέγγιση δεδομένων για τον προσδιορισμό της σχέσης μεταξύ της μεταφορικότητας και του υδραυλικού φορτίου σε υπόγεια ροή. Αρχικά έπρεπε να εκπαιδευτεί καταλλήλως το νευρωνικό δίκτυο έτσι ώστε να είναι σε θέση να προσεγγίζει την σχέση μεταξύ οποιασδήποτε πιθανής μεταφορικότητας του υδροφορέα και τιμών υδραυλικού φορτίου, όπως περιγράφηκε στο μοντέλο ροής υπόγειου νερού. Στη συνέχεια προκειμένου να επιλυθεί και το αντίστροφο πρόβλημα, δημιουργήθηκε ένα πεδίο μεταφορικότητας, το οποίο θα αναγνωρίζει τις γνωστές τιμές μεταφορικότητας και θα αναπαράγει τις γνωστές τιμές του υδραυλικού φορτίου όταν χρησιμοποιείται μοντέλο ροής υπόγειου νερού. Η διαδικασία που ακολουθείται χρησιμοποιεί ένα υποθετικό διδιάστατο υδροφορέα, ενώ τα στοιχεία εισόδου και εξόδου θεωρούνται ήδη γνωστά και γι' αυτό τον λόγο είναι δυνατή η ποσοτικοποίηση της αντίστροφης διαδικασίας (Garcia et al., 2006).

Μία ακόμα μελέτη προσέγγισης των παραμέτρων του υδροφορέα μέσω της χρήσης των νευρωνικών δικτύων, πραγματοποιήθηκε το 2007 από τους Samani, Gohari-Moghadam και Safavi. Στην συγκεκριμένη εργασία χρησιμοποιείται ο αλγόριθμος εκπαίδευσης Levenberg–Marquardt αντί για τη μέθοδο της απότομης καθόδου (Lin and Chen, 2006), καθώς και η τεχνική Principal Component Analysis (PCA), η οποία μείωσε αισθητά την τοπολογία του δικτύου, έτσι ώστε το στρώμα εισόδου να αποτελείται από ένα μόνο νευρώνα και το ενδιάμεσο κρυφό στρώμα να απαρτίζεται από οχτώ νευρώνες. Το δίκτυο εκπαιδεύεται καταλλήλως και ο υπολογισμός των παραμέτρων του υδροφορέα πραγματοποιείται μέσω των εξισώσεων Theis. Τα αποτελέσματα που απέδωσε η συγκεκριμένη εφαρμογή συγκρίθηκαν με τα αντίστοιχα αποτελέσματα των Lin και Chen (Lin and Chen, 2006) καθώς επίσης και με γραφικές μεθόδους. Σύμφωνα με τα παραπάνω, επιβεβαιώθηκε ότι η εναλλακτική τεχνική των νευρωνικών δικτύων, όπως προτάθηκε στην συγκεκριμένη μελέτη, τείνει να είναι όχι μόνο η απλούστερη, αλλά και ακριβέστερη (Samani et al., 2007).

Ο Affandi και οι συνεργάτες του δημιούργησαν ένα τεχνητό νευρωνικό δίκτυο και το εκπαίδευσαν με πέντε διαφορετικούς αλγόριθμους της τεχνικής «πίσω διάδοσης λάθους» (backpropagation) στοχεύοντας την σύγκριση των αποτελεσμάτων πρόβλεψης τιμών της στάθμης του υπόγειου υδροφόρου ορίζοντα. Οι αλγόριθμοι που χρησιμοποιήθηκαν ήταν οι Gradient descent with momentum, Gradient descent with adaptive learning rate and momentum, The Fletcher-Reeves Conjugate gradient, Quasi–Newton και Levenberg–Marquardt. Για την ανάπτυξη του δικτύου χρησιμοποιήθηκε το πρόγραμμα MATLAB. Τα αποτελέσματα της σύγκρισης ανέδειξαν τον

αλγόριθμο Levenberg–Marquardt ως τον καλύτερο για την εκτίμηση πρόβλεψης (Affandi et al., 2007).

Μία ενδιαφέρουσα και αρκετά πρόσφατη εφαρμογή αποτέλεσε η έρευνα των Yang και συνεργατών του, οι οποίοι χρησιμοποίησαν και συνέκριναν δύο διαφορετικά μοντέλα πρόβλεψης όσον αφορά την στάθμη των υπόγειων υδάτων σε δύο περιοχές της Κίνας, με συνθήκες ξηρασίας και μερικής ξηρασίας αντίστοιχα. Στην συγκεκριμένη μελέτη έγινε χρήση χρονοσειρών (Integrated Time Series, ITS) και του αλγορίθμου ανάδρασης τεχνητών νευρωνικών δικτύων (Back-Propagation Artificial Neural Network, BPANN) βάσει μοντέρνων στατιστικών μοντέλων και τεχνικών, προκειμένου να προκύψει μία εμπειρική πρόβλεψη των υδρολογικών μεταβλητών. Οι μελετητές έθεσαν δύο στόχους: αρχικά έπρεπε να διερευνηθούν τα πλεονεκτήματα και μειονεκτήματα των δύο παραπάνω μοντέλων στην πρόβλεψη της στάθμης του υπόγειου νερού, σε περιοχές όπου το οριζόντιο στρώμα των υπόγειων υδάτων έχει υποστεί σταδιακή πτώση τη τελευταία δεκαετία, λόγω υπερεκμετάλλευσης. Στη συνέχεια σημαντικό είναι να διατεθούν αξιόπιστα δεδομένα καθώς και ένα ολοκληρωμένο σχέδιο βασισμένο σε τρόπους ορθολογικής χρήσης των υπόγειων υδάτων στην συγκεκριμένη περιοχή μελέτης. Για τη διερεύνηση της ακρίβειας των μεθόδων χρησιμοποιήθηκαν τα εξής: το μέσο τετραγωνικό σφάλμα (RMSE), το μέσο απόλυτο σφάλμα (MAE) και συντελεστής αξιοπιστίας (CE). Σύμφωνα με τα αποτελέσματα της προσομοίωσης και τα δύο μοντέλα είναι σε θέση να δώσουν εξίσου αξιόπιστα αποτελέσματα. Κατά τη φάση της επιβεβαίωσης όμως διαπιστώθηκε ότι η τεχνική των νευρωνικών δικτύων υπερέχει της μεθόδου των χρονοσειρών, σύμφωνα με τους δείκτες *RMSE*, *MAE* και *CE*. Στο σχήμα 6.1 παρουσιάζεται η σύγκριση των παρατηρούμενων τιμών στάθμης του υπόγειου νερού με τις αντίστοιχες τιμές που προέκυψαν υπολογιστικά κατά την περίοδο εκπαίδευσης (Yang et al., 2009).



Σχήμα 6.1: Αναμενόμενες και εκτιμώμενες τιμές (ITS και BPAAN) στάθμης υπόγειου νερού για το σετ δεδομένων εκπαίδευσης (Yang et al, 2009)

Το 2007 η Καραμπερίδου ανέπτυξε ένα τεχνητό νευρωνικό δίκτυο ανάδρασης, που εκτιμά την συνολική παροχή άντλησης υφάλμυρου νερού, από πηγάδια σε παράκτιο υδροφορέα της

Ανατολικής Μακεδονίας. Σημαντικό είναι το ότι το νευρωνικό δίκτυο εκπαιδεύτηκε χρησιμοποιώντας αποτελέσματα από την εφαρμογή του κώδικα οριακών στοιχείων, αναπτύσσοντας τον αλγόριθμο Quickprop. Στόχος της εργασίας ήταν η σύγκριση των αποκρίσεων του νευρωνικού δικτύου με τις τιμές πραγματοποίησης του κώδικα οριακών στοιχείων, προκειμένου να διαπιστωθεί η αξιοπιστία, αλλά και η απλότητα εφαρμογής της τεχνολογίας των τεχνητών νευρωνικών δικτύων (Καραμπερίδου κ.ά., 2007).

6.4 Εφαρμογή Τεχνητών Νευρωνικών Δικτύων σε καρστικούς υδροφορείς

Η παρούσα μεταπτυχιακή εργασία ερευνά την εφαρμογή των τεχνητών νευρωνικών δικτύων σε υδροφορείς με υψηλή καρστικότητα. Λόγω της φύσης και των ιδιοτεροτήτων της καρστικότητας ως φαινόμενο, έχει πραγματοποιηθεί πλήθος μελετών, που ερευνά την αποτελεσματικότητα και την αξιοπιστία των νευρωνικών δικτύων σε θέματα υδραυλικής και διαχείρισης υδατικών πόρων.

Οι Lallahem και Mania συνδύασαν ένα μαθηματικό μοντέλο (MMO8) και τα τεχνητά νευρωνικά δίκτυα, προκειμένου να προβλεφθεί η απορροή πηγής σε καρστική περιοχή (Lallahem and Mania, 2003). Επίσης μία ακόμη ενδιαφέρουσα μελέτη πραγματεύτηκε την εκτίμηση του βαθμού ερημοποίησης καρστικού – βραχώδους εδάφους, στην Κίνα. Στην συγκεκριμένη έρευνα υλοποιήθηκε συνδυαστική χρήση των νευρωνικών δικτύων ακτινικής βάσης (RBFN) και του GIS (Geographic information system) και η συσχέτιση των αποτελεσμάτων με τα αντίστοιχα ερευνών που εφήρμοσαν διαφορετικές μεθόδους, απέδειξε όχι μόνο την ορθότητα, αλλά και την δυνατότητα και ευκολία εφαρμογής των νευρωνικών δικτύων γενικότερα (Qiu-hao and Yun-long, 2006).

Στοχεύοντας την πρόβλεψη παροχών καρστικής πηγής στην περιοχή Αλμυρού του Ηρακλείου Κρήτης, πραγματοποιήθηκε συνδυασμός εφαρμογής τεχνητών νευρωνικών δικτύων και μη γραμμικών μεθόδων (Farmer's algorithm). Αφού αποδείχτηκε ότι η ύπαρξη υψηλής συχνότητας θορύβου στην χρονοσειρά της παροχής είναι ασήμαντη, επιβεβαιώθηκε, μέσω των αποτελεσμάτων, η εξαιρετική ικανότητα πρόβλεψης όχι μόνο της παραπάνω συνδυαστικής χρήσης των δύο εργαλείων τεχνολογίας, αλλά και ιδιαίτερα των νευρωνικών δικτύων (Lambrakis et al., 2000).

Στη συνέχεια αξίζει να αναφερθεί μια πρόσφατη έρευνα (Trichakis et al.) κατά την οποία χρησιμοποιήθηκε μία αρχιτεκτονική τεχνητού νευρωνικού δικτύου, ικανή να προβλέψει την απόκριση ενός καρστικού υδροφορέα, χρησιμοποιώντας την αλλαγή του υδραυλικού φορτίου, αντί για το ίδιο το υδραυλικό φορτίο ως έξοδο στο δίκτυο. Σημαντική είναι η εκτίμηση της χρονικής υστέρησης της βροχόπτωσης μέσω του συντελεστή συσχέτισης. Επίσης χρησιμοποιήθηκε η έννοια του Διαφορικού Εξελικτικού Αλγορίθμου στοχεύοντας την βελτιστοποίηση της αρχιτεκτονικής του δικτύου, όσον αφορά τον αριθμό των κόμβων σε κάθε κρυφό επίπεδο, των αριθμό των γενεών που τρέχει το νευρωνικό δίκτυο, καθώς και το σφάλμα αξιολόγησης. Τέλος, παρουσιάστηκε μία

πρόβλεψη του υδραυλικού ύψους για διάστημα 30 ως 90 ημερών, με πολύ αξιόλογα αποτελέσματα (Trichakis et al., 2009).

Συνοψίζοντας θα μπορούσε να πει κανείς πως τα τεχνητά νευρωνικά δίκτυα έχουν μεγάλες δυνατότητες και είναι σε θέση να εφαρμοστούν με άριστα αποτελέσματα στην υδρολογία, αλλά και σε άλλους τομείς στοχεύοντας την πρόβλεψη περιβαλλοντικών συνιστωσών. Αποτελεί γεγονός ότι ανατρέχοντας στη διεθνή βιβλιογραφία έχει σημειωθεί μεγάλος αριθμός μελετών, τις τελευταίες δεκαετίες, που μέσω του εργαλείου των νευρωνικών δικτύων αντιμετωπίζουν μείζονα θέματα υδραυλικής και υδρολογίας. Αναμφισβήτητα αποτελούν μία αποδεκτά καλή τεχνική πρόβλεψης, λόγω της φύσης της δομής και της λειτουργίας τους, η οποία σαφώς εξελίσσεται και είναι ικανή να δίνει αξιόλογα αποτελέσματα ειδικά σε περιοχές με καρστικότητα, όπου η εφαρμογή κοινών αποδεκτών μοντέλων δεν συνίσταται. Σημαντική βέβαια είναι κρίση του κάθε μελετητή, για το αν η χρήση της τεχνολογίας των νευρωνικών δικτύων αρμόζει στη φύση και τη λογική του διερευνούμενου προβλήματος.

Κεφάλαιο 7

Ανάπτυξη Αλγορίθμων με Χρήση Νευρωνικών Δικτύων – Ανάλυση Αποτελεσμάτων

7.1 Επεξεργασία Δεδομένων

Στην παρούσα ενότητα θα περιγραφεί αναλυτικά η διαδικασία προεπεξεργασίας των δεδομένων, προκειμένου να εισαχθούν στο τεχνητό νευρωνικό δίκτυο. Χρησιμοποιήθηκαν στοιχεία βροχομετρικών υψών και παροχών για δύο πηγές: α) πηγή Μάη Βρύσης και β) πηγή Πέρα Βρύσης, στην περιοχή του δήμου Ρούβας, για τα έτη 2007, 2008 και 2009 (έως 15/07/09). Τα στοιχεία αυτά είχαν συγκεντρωθεί από τον κ. Ν. Δαριβιανάκη, στο πλαίσιο της εκπόνησης της διδακτορικής του διατριβής (Δαριβιανάκης, 2007).

Η δυσκολία που αντιμετωπίστηκε ήταν ότι τα δεδομένα δεν ήταν πλήρη. Δηλαδή, υπήρχαν κενά στα στοιχεία των παροχών, τα οποία αντιμετωπίστηκαν με τρόπο που θα περιγραφεί παρακάτω.

Αρχικά σε φύλλο του excel δημιουργήθηκαν τρεις στήλες, οι οποίες αντιστοιχούσαν: 1) στην ημερομηνία (από 01/01/2007 έως 15/07/09), 2) στο ύψος βροχόπτωσης και 3) στην παροχή. Σε περίπτωση μη ύπαρξης τιμής της παροχής, καταγραφόταν η ένδειξη «Nan». Ουσιαστικά δηλαδή, δημιουργήθηκαν οι χρονοσειρές της βροχόπτωσης και της απορροής. Η παραπάνω διαδικασία ακολουθήθηκε για κάθε μία από τις δύο πηγές. Στόχος ήταν η πραγματοποίηση συσχέτισης του ύψους βροχόπτωσης και της τιμής της παροχής. Το ύψος βροχόπτωσης και η παροχή της πηγής που θα αντιστοιχεί στο συγκεκριμένο ύψος βροχής παρουσιάζουν ένα χρονικό διάστημα καθυστέρησης (time lag). Γι αυτόν τον λόγο χρησιμοποιήθηκε, σε περιβάλλον Matlab, ένα πρόγραμμα που υπολογίζει τον συντελεστή συσχέτισης βροχόπτωσης-απορροής. Υπολογίζεται ο συντελεστής συσχέτισης της βροχόπτωσης και της απορροής για διαφορετικές χρονικές καθυστερήσεις (time lags), από 0 έως 30 μέρες. Ο στόχος ήταν να βρεθεί ο μέγιστος συντελεστής συσχέτισης, ο οποίος θα αντιστοιχεί στο χρονικό διάστημα που απαιτείται προκειμένου να φτάσει το ύψος βροχής στην πηγή. Έπειτα, μετατέθηκε η χρονοσειρά της βροχόπτωσης, σύμφωνα με τον υπολογισμό του “time lag”.

Σημαντικό ήταν το πρόβλημα της μη ύπαρξης στοιχείων παροχής για ορισμένες ημερομηνίες. Τις ημέρες που η βροχόπτωση ήταν μηδενική ή εξαιρετικά μικρή και δεν υπήρχαν διαθέσιμα στοιχεία παροχής, έγινε παραδοχή ότι η μεταβολή στο πιεζομετρικό ύψος του υδροφορέα ήταν γραμμική και επομένως έγινε συμπλήρωση των στοιχείων της παροχής που έλειπαν, με βάση αυτή την σχέση. Το πλεονέκτημα από την χρήση της γραμμικής παρεμβολής ήταν η αύξηση των στοιχείων που πρόκειται να εισαχθούν στο τεχνητό νευρωνικό δίκτυο, γεγονός που συμβάλλει στην καλύτερη εκπαίδευσή του και επομένως στην εξαγωγή καλύτερων αποτελεσμάτων.

Τέλος, αγνοήθηκε οποιαδήποτε ημερομηνία στην οποία δεν υπήρχε στοιχείο απορροής και δεν δικαιολογεί την εφαρμογή της γραμμικής παρεμβολής. Το αποτέλεσμα όλων των παραπάνω ήταν η δημιουργία των τελικών στοιχείων βροχόπτωσης και απορροής που θα χρησιμοποιηθούν στο δίκτυο. Έτσι προέκυψαν οι τρεις ακόλουθες στήλες: 1) αύξοντας αριθμός που αντιστοιχεί στην ημερομηνία καταγραφής των στοιχείων βροχόπτωσης και απορροής (για ευκολία θα συναντάται ως «αριθμός ημέρας»), 2) ύψος βροχόπτωσης και 3) παροχή. Τα δύο πρώτα αποτέλεσαν τα διανύσματα εισόδου στο τεχνητό νευρωνικό δίκτυο που δημιουργήθηκε, ενώ η παροχή ήταν το διάνυσμα επιθυμητού στόχου. Και τα τρία αυτά διανύσματα εννοείται ότι θα πρέπει να έχουν ακριβώς τον ίδιο αριθμό στοιχείων. Στόχος ήταν η αναπαραγωγή των παροχών μέσω βροχομετρικών δεδομένων. Αξίζει να σημειωθεί ότι ερευνήθηκε μέσω των αποτελεσμάτων του δικτύου, η επίδραση της γραμμικής παρεμβολής στα στοιχεία της απορροής.

7.2 Περιγραφή και γενικές αρχές του νευρωνικού δικτύου

Όπως αναφέρθηκε στην παράγραφο που προηγήθηκε τα διανύσματα εισόδου στο νευρωνικό δίκτυο είναι ο αριθμός ημέρας και η βροχόπτωση, ενώ η απορροή είναι το διάνυσμα επιθυμητού στόχου. Το επόμενο βήμα ήταν η επιλογή του βέλτιστου νευρωνικού δικτύου. Αρχικά πρέπει να σημειωθεί ότι επιλέχθηκε δίκτυο με τρία επίπεδα (ένα στρώμα εισόδου, ένα ενδιάμεσο στρώμα και ένα στρώμα εξόδου). Η επιλογή των νευρώνων του κρυφού στρώματος έγινε βάσει δοκιμών, οι οποίες παρατίθενται παρακάτω. Επίσης το δίκτυο είναι δίκτυο προώθησης με ανάδραση (feed-forward neural network with Backpropagation), ενώ δοκιμάστηκαν διάφορες συναρτήσεις δημιουργίας δικτύου, συναρτήσεις εκπαίδευσης και συναρτήσεις εκμάθησης, μέχρι να εντοπιστεί το βέλτιστο αποτέλεσμα. Κριτήριο επιλογής ήταν το κανονικοποιημένο μέσο τετραγωνικό σφάλμα ανάμεσα στις αναμενόμενες και παραχθείσες από το δίκτυο τιμές. Η διαδικασία αυτή ακολουθήθηκε και για τις δύο πηγές: Μάη Βρύση και Πέρα Βρύση. Επομένως δοκιμάστηκαν διάφοροι αλγόριθμοι οι οποίοι παρατίθενται παρακάτω.

7.3 Μεθοδολογία - Επιλογή βέλτιστου συνδυασμού τύπου εκπαίδευσης και πλήθους νευρώνων κρυφού στρώματος

Προκειμένου να χρησιμοποιηθεί ο βέλτιστος τύπος εκπαίδευσης πραγματοποιήθηκε μία σειρά δοκιμών από συναρτήσεις δημιουργίας δικτύου, συναρτήσεις εκπαίδευσης και συναρτήσεις εκμάθησης, για διαφορετικό αριθμό νευρώνων του ενδιάμεσου στρώματος, μέχρι να εντοπιστεί το βέλτιστο αποτέλεσμα. Το κριτήριο επιλογής ήταν το κανονικοποιημένο μέσο τετραγωνικό σφάλμα ανάμεσα στις αναμενόμενες και παραχθείσες από το νευρωνικό δίκτυο τιμές. Βασικός στόχος ήταν η επιλογή του πλήθους H , των νευρώνων του ενδιάμεσου κρυφού στρώματος. Γι' αυτόν τον λόγο πραγματοποιήθηκαν δοκιμές για H από 1 έως 10 για κάθε τύπο νευρώνων εκπαίδευσης. Ο λόγος που δοκιμάστηκαν τιμές πλήθους νευρώνων από 1 έως 10, είναι ότι η φύση του προβλήματος μας είναι τέτοια, ώστε να μην απαιτείται πιο πολύπλοκη δομή νευρωνικού δικτύου.

▪ Καθορισμός της χρονικής υστέρησης της βροχόπτωσης:

Ο σωστός καθορισμός της χρονικής υστέρησης μεταξύ βροχόπτωσης και παροχής της πηγής αποτελεί έναν σημαντικό παράγοντα για τη διαμόρφωση των δεδομένων εισόδου του νευρωνικού δικτύου. Ο υπολογισμός της χρονικής υστέρησης πραγματοποιήθηκε μέσω του συντελεστή συσχέτισης των χρονοσειρών της βροχόπτωσης και της παροχής, για τιμές χρονικής υστέρησης από 0 έως 30 μέρες. Η σχέση που αποδίδει τον συντελεστή συσχέτισης δύο χρονοσειρών A και B , ορίζεται ως:

$$Correl(\alpha, \beta) = \frac{\sum(\alpha - \bar{\alpha}) \cdot (\beta - \bar{\beta})}{\sqrt{\sum(\alpha - \bar{\alpha})^2 \cdot (\beta - \bar{\beta})^2}} \quad (7.1)$$

, όπου α και β αποτελούν τα στοιχεία, ενώ $\bar{\alpha}$ και $\bar{\beta}$ τη μέση τιμή των χρονοσειρών A και B , αντίστοιχα.

Η χρονική υστέρηση που αντιστοιχούσε στον μέγιστο συντελεστή συσχέτισης, αποτέλεσε τη βέλτιστη επιλογή για τη διαμόρφωση των δεδομένων του δικτύου.

Η παραπάνω διαδικασία ακολουθήθηκε και για τις δύο πηγές Μάη και Πέρα βρύση, ενώ εφαρμόστηκε ξεχωριστά για κάθε μία από τις περιπτώσεις που θα παρουσιαστούν στη συνέχεια. Επομένως υπολογίστηκε ο συντελεστής συσχέτισης των χρονοσειρών βροχόπτωσης και παροχής κάθε πηγής, και για τις δύο περιπτώσεις εφαρμογής ή μη, της μεθόδου της γραμμικής παρεμβολής. Για την επιλογή της δομής του νευρωνικού δικτύου και του βέλτιστου τύπου εκπαίδευσης χρησιμοποιήθηκαν τα δεδομένα στα οποία εφαρμόστηκε η μέθοδος της γραμμικής παρεμβολής και υπολογίστηκε η χρονική υστέρηση μέσω της σχέσης που παρατέθηκε παραπάνω.

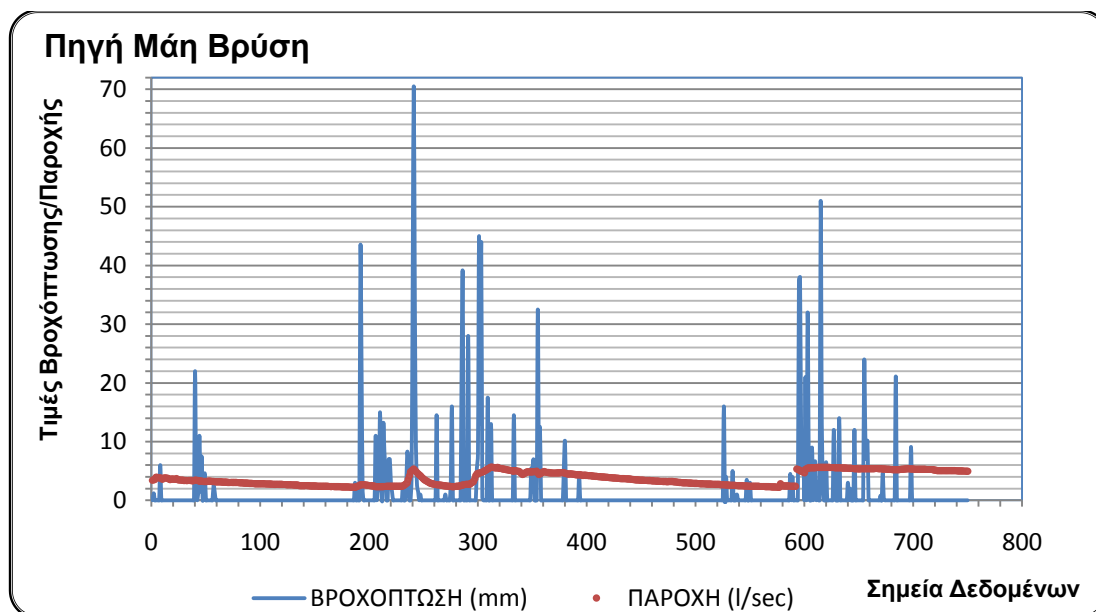
Οι τιμές του συντελεστή συσχέτισης για κάθε πηγή παρουσιάζονται στον πίνακα 7.1 που ακολουθεί:

TIME LAG (μέρες)	Συντελεστής Συσχέτισης		TIME LAG (μέρες)	Συντελεστής Συσχέτισης	
	Πηγή Μάη Βρύση	Πηγή Πέρα Βρύση		Πηγή Μάη Βρύση	Πηγή Πέρα Βρύση
0	-0,0402	0,0295	16	0,1099	0,2371
1	-0,0054	0,02	17	0,0767	0,2202
2	0,0284	0,0364	18	0,0729	0,1245
3	0,0613	0,0174	19	0,0683	0,1008
4	0,1147	0,098	20	0,0686	0,1559
5	0,1081	0,1684	21	0,066	0,1396
6	0,1386	0,135	22	0,0526	0,1722
7	0,1179	0,117	23	0,0687	0,1833
8	0,1051	0,1316	24	0,0723	0,1605
9	0,1148	0,098	25	0,08	0,1869
10	0,1196	0,1236	26	0,0903	0,1782
11	0,1204	0,122	27	0,0924	0,1381
12	0,1309	0,1321	28	0,0876	0,0839
13	0,1224	0,1602	29	0,1005	0,1229
14	0,1219	0,1534	30	0,1075	0,149
15	0,1166	0,2285			

Πίνακας 7.1 : Τιμές συντελεστή συσχέτισης για χρονική υστέρηση ημερών 0 έως 30

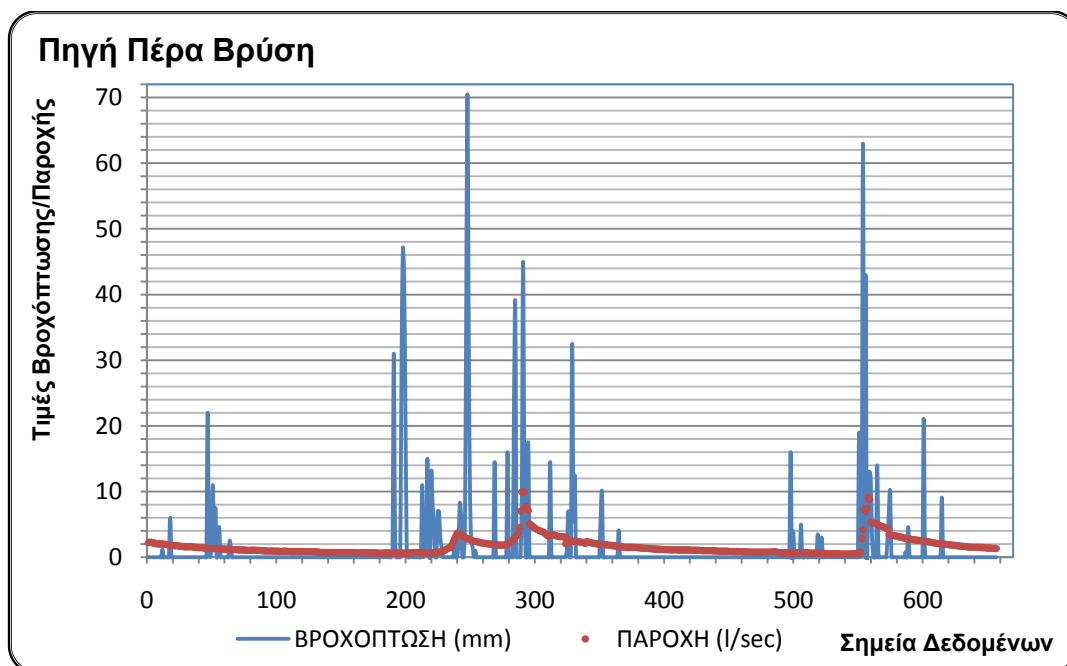
Παρατηρούμε ότι για την περίπτωση της πηγής Μάη Βρύση ο μέγιστος συντελεστής συσχέτισης προκύπτει για χρονική υστέρηση 6 ημερών, κάτι το οποίο δικαιολογείται από την γεωμορφολογική δομή του καρστικού υδροφορέα, ο οποίος αναμένεται να αντιδράσει σχετικά γρήγορα. Στην περίπτωση όμως της πηγής Πέρα Βρύση, παρατηρούμε ότι ο συντελεστής συσχέτισης μεγιστοποιείται την 16^η μέρα. Το γεγονός αυτό ενδεχομένως να οφείλεται στην ύπαρξη κάποιου γεωλογικού εμποδίου στο υπέδαφος, το οποίο έχει ως συνέπεια την καθυστέρηση της βροχόπτωσης. Άρα ενδεχομένως να υφίσταται μία ιδιομορφία, η οποία προκαλεί την συγκεκριμένη χρονική υστέρηση. Παρατηρούμε βέβαια ότι την 5^η μέρα παρουσιάζεται μία σχετικά υψηλή τιμή του συντελεστή συσχέτισης. Παρόλα αυτά, λόγω του ότι η έρευνα πραγματοποιήθηκε για την αντίδραση του υδροφορέα κατά το χρονικό διάστημα ενός μήνα, είναι αδύνατον να αγνοηθεί η μέγιστη τιμή της 16^{ης} μέρας καθώς είναι κατά πολύ μεγαλύτερη από την τιμή του συντελεστή για την 5^η μέρα. Γι αυτόν τον λόγο παρατίθεται η παραπάνω φυσική εξήγηση.

Προκειμένου να επιβεβαιωθεί η σωστή επιλογή της χρονικής υστέρησης δημιουργήθηκαν τα γραφήματα των χρονοσειρών της βροχόπτωσης και της παροχής. Σημειώνεται το ότι τα δεδομένα της βροχόπτωσης είναι τροποποιημένα σύμφωνα με την επιλογή του βέλτιστου συντελεστή συσχέτισης, όπως παρουσιάστηκε παραπάνω. Στόχος ήταν οι δύο χρονοσειρές να παρουσιάζουν όμοια συμπεριφορά.



Σχήμα 7.1: Χρονοσειρές βροχόπτωσης και παροχής για την πηγή Μάη Βρύση

Όντως από το παραπάνω διάγραμμα προκύπτει ότι η επιλογή του συντελεστή συσχέτισης για την πηγή Μάη Βρύσης είναι ικανοποιητική, καθώς η χρονοσειρά της βροχόπτωσης έχει όμοια συμπεριφορά με την χρονοσειρά της παροχής, παρουσιάζοντας σε κοινά σημεία τις μέγιστες τιμές τους. Ομοίως ακολουθεί η αντίστοιχη διαδικασία για την πηγή Πέρα Βρύση και προκύπτει το ακόλουθο διάγραμμα.



Σχήμα 7.2: Χρονοσειρές βροχόπτωσης και παροχής για την πηγή Πέρα Βρύση

Επομένως και στην περίπτωση αυτή διαπιστώνεται ότι οι δύο χρονοσειρές ακολουθούν όμοια συμπεριφορά.

Στη συνέχεια ακολουθούν αναλυτικά οι αλγόριθμοι και οι συνδυασμοί που υλοποιήθηκαν καθώς επίσης και τα αποτελέσματα που έδωσαν.

7.3.1 Ανάπτυξη αλγορίθμου με συνάρτηση δημιουργίας δικτύου την newff (feed- forward backpropagation network), συνάρτηση εκπαίδευσης την traingd και συνάρτηση εκμάθησης την learnnrm.

I. Συνάρτηση δημιουργίας δικτύου: "newff"

Η συνάρτηση δημιουργίας δικτύου που χρησιμοποιείται στον συγκεκριμένο αλγόριθμο είναι η newff. Μέσω της newff δημιουργείται ένα τεχνητό νευρωνικό δίκτυο το οποίο θα εκπαιδευτεί σύμφωνα με τον αλγόριθμο Backpropagation. Η συγκεκριμένη συνάρτηση δέχεται ως ορίσματα τα παρακάτω στοιχεία:

- i. Πίνακα minmax(pn) που αποτυπώνει την ελάχιστη και τη μέγιστη τιμή κάθε γραμμής του πίνακα pn των εισόδων του δικτύου.
- ii. Πίνακα [1, H, 1], ο οποίος εκφράζει τον αριθμό των νευρώνων σε κάθε στρώμα του δικτύου. Στη συγκεκριμένη εργασία επιλέχθηκαν ένα στρώμα εισόδου, ένα ενδιάμεσο κρυφό στρώμα και ένα στρώμα εξόδου. Κατά τη διαδικασία των δοκιμών συγκρίθηκαν τα αποτελέσματα για H=1 έως 10.
- iii. Συναρτήσεις ενεργοποίησης για κάθε επίπεδο. Η παράσταση: {'tansig', 'tansig', 'purelin'} δηλώνει ότι στο πρώτο στρώμα χρησιμοποιήθηκε η συνάρτηση ενεργοποίησης 'tansig', ομοίως και στο δεύτερο, ενώ στο στρώμα εξόδου χρησιμοποιήθηκε η συνάρτηση ενεργοποίησης 'purelin'.
- iv. Συνάρτηση εκπαίδευσης νευρωνικού δικτύου: 'traingd'.
- v. Συνάρτηση εκμάθησης νευρωνικού δικτύου: 'learnnrm'.

II. Συνάρτηση εκπαίδευσης δικτύου: "traingd"

Στον συγκεκριμένο αλγόριθμο χρησιμοποιήθηκε ως συνάρτηση εκπαίδευσης η traingd. Πρόκειται για συνάρτηση εκπαίδευσης με στοιβές δεδομένων (batch training). Με την μέθοδο της ροής των δεδομένων σε δέσμη, όλες οι εισοδοί εφαρμόζονται εξαρχής στο δίκτυο και οι τιμές των συναπτικών βαρών, των πολώσεων, καθώς και της βαθμωτής μεταβολής, αλλάζουν μετά την εφαρμογή των εισόδων. Η συνάρτηση traingd ενημερώνει τα συναπτικά βάρη και τις πολώσεις του δικτύου, σύμφωνα με τη μέθοδο της απότομης καθόδου (gradient descent) και απαιτεί τα συναπτικά βάρη, οι εισοδοί του δικτύου, καθώς και οι συναρτήσεις ενεργοποίησης να είναι παραγωγίσιμες. Ο

αλγόριθμος που δίνει την παράγωγο της συνάρτησης ενεργοποίησης perf ως προς τα συνναπτικά βάρη και τις πολώσεις του δικτύου X , δίνεται από την σχέση:

$$dX = lr * \frac{dperf}{dX} \quad (7.2)$$

, όπου:

$dperf/dX$: η παράγωγος της συνάρτησης ενεργοποίησης ως προς την μεταβολή των συνναπτικών βαρών dX και

lr : ο ρυθμός εκμάθησης του νευρωνικού δικτύου

Η κάθε συνάρτηση εκπαίδευσης συνοδεύεται από ορισμένες παραμέτρους. Η `traingd` χαρακτηρίζεται από τις εξής: `epochs`, `show`, `goal`, `time`, `min_grad`, `max_fail` και `lr`. Όταν μία παράμετρος δεν καλείται στον αλγόριθμο, αυτόματα δηλώνεται η υπάρχουσα προκαθορισμένη τιμή της παραμέτρου από την Matlab.

Η εκπαίδευση σταματά σε περίπτωση όπου:

- ο αριθμός των επαναλήψεων υπερβεί τον αριθμό των εποχών,
- η τιμή της συνάρτησης ενεργοποίησης πέσει κάτω από την τιμή που έχει τεθεί στην παράμετρο `goal`,
- η κλίση της συνάρτησης ενεργοποίησης πέσει κάτω από την τιμή `min_grad`
- μέγιστος χρόνος εκπαίδευσης υπερβεί το όριο `MAX_TIME` που έχει δοθεί από τον χρήστη (στον συγκεκριμένο αλγόριθμο το `MAX_TIME` είναι απεριόριστο και θέτεται από την Matlab ως `inf`).

III. Συνάρτηση εκμάθησης δικτύου: “`learnqdm`”

Όσον αφορά την συνάρτηση εκμάθησης αξίζει να σημειωθεί ότι η `newff` συνοδεύεται είτε από την συνάρτηση εκμάθησης `learnqd` είτε την `learnqdm`. Στην συγκεκριμένη εφαρμογή επιλέχθηκε η χρήση της `learnqdm`. Η `learnqdm` προτιμάται από την `learnqd` λόγω της παραμέτρου της «momentum», μέσω της οποίας αγνοεί μικρές ασυνέχειες στην πορεία μεταβολής των βαρών και έτσι το δίκτυο «μαθαίνει» πιο γρήγορα.

Με την `learnqdm` υπολογίζονται οι μεταβολές στα συνναπτικά βάρη και στις πολώσεις κάθε νευρώνα σε συνάρτηση με τις εισόδους κάθε νευρώνα, την τιμή της συνάρτησης ενεργοποίησης του νευρώνα, το ρυθμό εκμάθησης (`learning rate`) του δικτύου και μιας τιμής που συμβολίζει την ορμή στην μεταβολή των συνναπτικών βαρών και των πολώσεων, το `momentum`, `mc`. Η σχέση με την οποία μεταβάλλονται τα συνναπτικά βάρη και οι πολώσεις για κάθε νευρώνα με βάση την συνάρτηση εκμάθησης `learnqdm` δίνεται από την παρακάτω σχέση:

$$dW = MC * dW_{prev} + (1 - MC) * lr * gW \quad (7.3)$$

,όπου:

dWprev: η μεταβολή κατά την προηγούμενη επανάληψη του συγκεκριμένου συναπτικού βάρους ή πόλωσης, η οποία αποθηκεύεται προσωρινά στην μεταβλητή LS (learning state variable) και διαβάζεται από εκεί κατά την επόμενη επανάληψη,

MC: η παράμετρος της ορμής στην μεταβολή των συναπτικών βαρών και των πολώσεων.

lr: ο ρυθμός εκμάθησης του νευρωνικού δικτύου,

gW: κλίση αναφορικά με την απόδοση του δικτύου.

Στη συνέχεια ακολουθεί ο αλγόριθμος που υλοποιήθηκε με τα παραπάνω στοιχεία.

ΑΛΓΟΡΙΘΜΟΣ 1 (neural newff trainingd learnqdm)

```
p = [...]
t = [...]
%-----
Τα διανύσματα p και t αποτελούν το διάνυσμα εισόδου και το διάνυσμα
επιθυμητής εξόδου αντίστοιχα, όπως περιγράφηκαν σε προηγούμενη παράγραφο.
Ουσιαστικά πρόκειται για τους πίνακες:  $p = \begin{bmatrix} \text{ύψη βροχόπτωσης} \\ \text{αριθμός ημέρας} \end{bmatrix}$  και  $t =$ 
[τιμές παροχής]
%-----
[pn, ps]= mapminmax (p, 0, 1)
[tn, ts]= mapminmax (t, 0, 1)
%-----
Κανονικοποίηση των τιμών μέσα στο διάστημα τιμών [0,1]
%-----
[trainV, valV, testV]= dividevec (pn, tn, 0.15, 0.15);
%-----
Η συνάρτηση dividevec χρησιμοποιείται προκειμένου να χωρίσει ένα σετ
δεδομένων εισόδου και επιθυμητών στόχων σε ομάδες διανυσμάτων για
α)εκπαίδευση, β)επιβεβαίωση (validation) της απόδοσης του δικτύου κατά τη
διάρκεια της εκπαίδευσης, έτσι ώστε η εκπαίδευση να σταματήσει νωρίτερα
για να αποφύγει το overfitting των training data, και τέλος γ) δεδομένα
για έλεγχο ως ένα ανεξάρτητο μέτρο για το πώς το δίκτυο αναμένεται να
αποδώσει σε δεδομένα στα οποία δεν έχει εκπαιδευτεί.
%-----
net= newcf (minmax(pn), [1, H, 1], {'tansig', 'tansig', 'purelin'},
'trainingd', 'learnqdm')
%-----
Δημιουργία τεχνητού νευρωνικού δικτύου τριών επιπέδων, με συναρτήσεις
ενεργοποίησης τις tansig, tansig και purelin για κάθε επίπεδο αντίστοιχα,
ενώ η εκπαίδευση του δικτύου θα γίνει μέσω της συνάρτησης trainingd και η
εκμάθηση μέσω της learnqdm. Ο αλγόριθμος έτρεξε για H από 1 έως 10.
%-----
net= init(net)
%-----
```

Ανάπτυξη Αλγορίθμων με Χρήση Νευρωνικών Δικτύων – Ανάλυση Αποτελεσμάτων

Μέσω της συνάρτησης `init` πραγματοποιείται αρχικοποίηση των συναπτικών βαρών και των πολώσεων του νευρωνικού δικτύου. Αξίζει να σημειωθεί ότι η συνάρτηση δημιουργίας δικτύου `newff` αρχικοποιεί τις τιμές των συναπτικών βαρών και των πολώσεων αυτόματα, αλλά είναι πιο ασφαλές να κληθεί η συνάρτηση αρχικοποίησης `init` της Matlab. Το όρισμα της `init` είναι το `net` που έχει ήδη δημιουργηθεί μέσω της συνάρτησης `newff` και επιστρέφει το ίδιο `net`, μόνο που σε αυτό το σημείο έχει πραγματοποιηθεί η διαδικασία της αρχικοποίησης των τιμών των συναπτικών βαρών και των πολώσεων.

```
%-----
net. trainParam.epochs= 8000;
```

```
%-----
Ο αριθμός των επαναλήψεων εκπαίδευσης του δικτύου. (Προκαθορισμένη τιμή της Matlab: 10).
```

```
%-----
net. trainParam.show= 100;
```

```
%-----
Η κατάσταση της εκπαίδευσης θα απεικονίζεται στο command window της Matlab για κάθε 100 επαναλήψεις του αλγορίθμου. (Προκαθορισμένη τιμή της Matlab: 10).
```

```
%-----
net. trainParam.goal= 0.0001;
```

```
%-----
Η συνάρτηση mse που αποτελεί την performance function έχει ως κατώτατο όριο την τιμή 0.0001. (Προκαθορισμένη τιμή της Matlab: 0).
```

```
%-----
net. trainParam.min_grad= 1e-20;
```

```
%-----
Η κλίση της performance function έχει ως κατώτατο όριο την τιμή 1e-20. (Προκαθορισμένη τιμή της Matlab: 1e-10).
```

```
%-----
net. trainParam.lr= 0.01;
```

```
%-----
Ο ρυθμός εκμάθησης της συνάρτησης εκμάθησης learngdm. Όσο πιο μεγάλη είναι η τιμή που δίνεται στον ρυθμό εκμάθησης, τόσο πιο μεγάλο είναι το βήμα. (Προκαθορισμένη τιμή της Matlab: 0.01).
```

```
%-----
net= train(net, trainV.P, trainV.T, [], [], valV, testV);
```

```
%-----
Εκπαίδευση του δικτύου μέσω της συνάρτησης train.
```

```
%-----
y= sim(net, testV.P, [], [], testV.T);
```

```
%-----
Αποτελέσματα y του δικτύου.
```

```
%-----
e= testV.T-y;
```

```
%-----
Υπολογισμός σφάλματος ανάμεσα στον επιθυμητό στόχο και στα αποτελέσματα του νευρωνικού δικτύου.
```

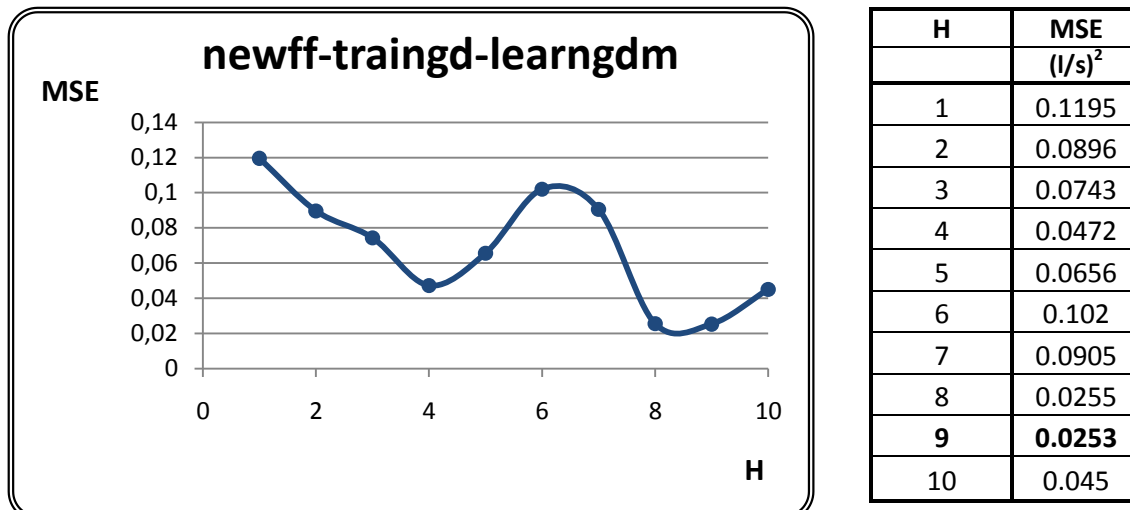
 %
Perf= mse (e)
 %

 Υπολογισμός μέσου τετραγωνικού σφάλματος (των κανονικοποιημένων τιμών) .
 %

End

Ο παραπάνω αλγόριθμος εφαρμόστηκε για H από 1 έως 10 και με κριτήριο το mse έδωσε τα εξής αποτελέσματα:

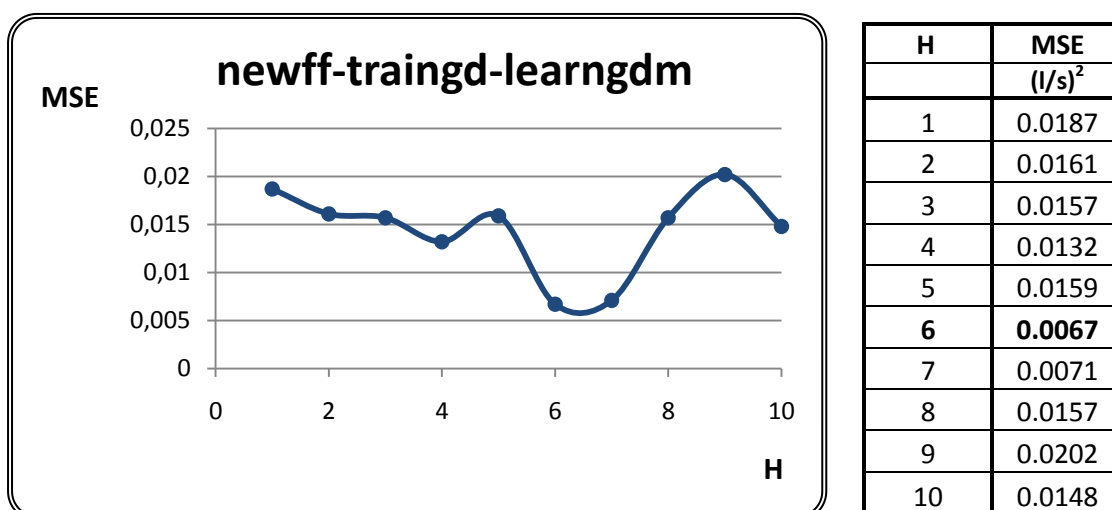
A) Πηγή Μάν Βρύση:



Σχήμα 7.3 : Γραφική απεικόνιση αποτελεσμάτων αλγορίθμου 1 για πηγή Μάν Βρύση

Παρατηρούμε ότι για τον συγκεκριμένο τύπο εκπαίδευσης, για H=9 δόθηκε η μικρότερη τιμή μέσου τετραγωνικού σφάλματος ίση με 0.0253 (l/s)².

B) Πηγή Πέρα Βρύση:



Σχήμα 7.4: Γραφική απεικόνιση αποτελεσμάτων αλγορίθμου 1 για πηγή Πέρα Βρύση

Στην πηγή Πέρα Βρύσης, για τον ίδιο τύπο εκπαίδευσης και για $H=6$ δόθηκε η μικρότερη τιμή μέσου τετραγωνικού σφάλματος ίση με 0.0067 (l/s)^2 .

7.3.2 Ανάπτυξη αλγορίθμου με συνάρτηση δημιουργίας δικτύου την newff (feed- forward backpropagation network), συνάρτηση εκπαίδευσης την traingdm και συνάρτηση εκμάθησης την learngdm.

Η μόνη διαφορά με τον προηγούμενο αλγόριθμο είναι η κλήση της συνάρτησης εκπαίδευσης traingdm, η οποία επίσης εκπαιδεύει με βάση τις στοίβες δεδομένων (batch training). Η ρύθμιση των βαρών και των πολώσεων πραγματοποιείται ομοίως με τη μέθοδο απότομης καθόδου (steepest descent), αλλά υπεισέρχεται μία ακόμη παράμετρος, η ορμή. Πρόκειται για την ορμή (momentum), με την οποία μεταβάλλονται τα συναπτικά βάρη και οι πολώσεις του δικτύου. Η παράμετρος αυτή επιτρέπει στο δίκτυο να αντιδρά όχι μόνο στην μεταβολή της κλίσης gradient, αλλά και σε προσωρινές αυξήσεις ή μειώσεις που ενδεχομένως να παρουσιάζει η τιμή συνάρτησης σφάλματος σε κάθε επανάληψη, αγνοώντας τις μικρές ασυνέχειες. Αξίζει να σημειωθεί ότι η προκαθορισμένη τιμή της παραμέτρου mc από την Matlab είναι 0.9. Μικρότερες τιμές της παραμέτρου αντιστοιχούν σε επιβράδυνση της διαδικασίας εκπαίδευσης.

Όπως και στην περίπτωση της traingd, έτσι και η συνάρτηση εκπαίδευσης traingdm, είναι σε θέση να εκπαιδεύσει κάθε νευρωνικό δίκτυο εφόσον τα συναπτικά βάρη, οι είσοδοι του δικτύου και οι συναρτήσεις ενεργοποίησης έχουν παράγωγες συναρτήσεις.

Η μέθοδος του backpropagation χρησιμοποιείται για να υπολογίζει τις παραγώγους της performance function ως προς τα βάρη και τις πολώσεις X . Η ρύθμιση των μεταβλητών γίνεται σύμφωνα με την μέθοδο απότομης καθόδου με ορμή (gradient descent with momentum),

$$dX = mc * dX_{prev} + lr(1 - mc) * \frac{dperf}{dX} \quad (7.4)$$

, όπου:

dXprev: η μεταβολή των συναπτικών βαρών και των πολώσεων κατά την προηγούμενη επανάληψη,

mc: παράμετρος της ορμής (momentum), με προκαθορισμένη τιμή από την Matlab: 0.9,

lr: ο ρυθμός εκμάθησης του νευρωνικού δικτύου,

dperf: η παράγωγος της απόδοσης του δικτύου σε σχέση με την μεταβολή των συναπτικών βαρών και των πολώσεων.

Ακολουθεί ο αλγόριθμος με συνάρτηση εκπαίδευσης την traingdm:

ΑΛΓΟΡΙΘΜΟΣ 2 (neural_newff_traingdm_learngdm)

```

p = [...]

t = [...]

[pn,ps]= mapminmax(p,0,1)

[tn,ts]= mapminmax(t,0,1)

[trainV, valV, testV]= dividevec(pn, tn, 0.15, 0.15);

net= newcf(minmax(pn), [1, H, 1], {'tansig', 'tansig', 'purelin'},
'traingdm', 'learngdm')
%-----
Δημιουργία τεχνητού νευρωνικού δικτύου τριών επιπέδων, με συναρτήσεις
ενεργοποίησης τις tansig, tansig και purelin για κάθε επίπεδο αντίστοιχα,
ενώ η εκπαίδευση του δικτύου θα γίνει μέσω της συνάρτησης traingdm και η
εκμάθηση μέσω της learngdm. Ο αλγόριθμος έτρεξε για H από 1 έως 10.
%-----
net= init(net)

net. trainParam.epochs= 8000;

net. trainParam.show= 100;

net. trainParam.goal= 0.0001;

net. trainParam.min_grad= 1e-20;

net. trainParam.lr= 0.01;

net. trainParam.mc= 0.9;

net= train(net, trainV.P, trainV.T, [], [], valV, testV);

y= sim(net, testV.P, [], [], testV.T);

e= testV.T-y;

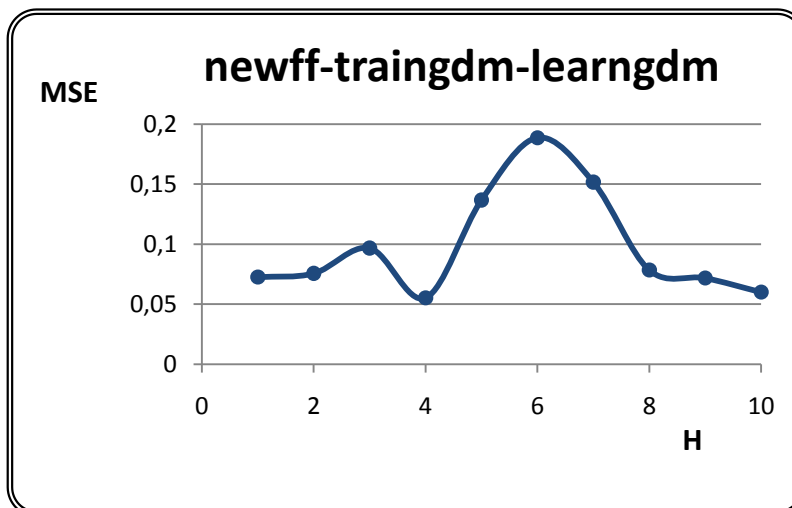
Perf= mse(e)

End

```

Ο παραπάνω αλγόριθμος εφαρμόστηκε για H από 1 έως 10 και με κριτήριο το mse έδωσε τα εξής αποτελέσματα:

Α) Πηγή Μή Βρύση:

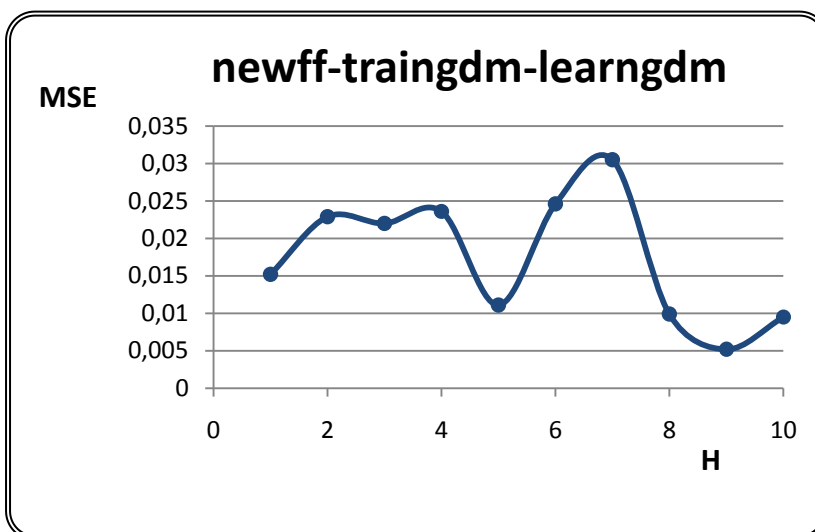


H	MSE (l/s) ²
1	0.0726
2	0.0757
3	0.0967
4	0.0553
5	0.1368
6	0.1886
7	0.1516
8	0.0785
9	0.0718
10	0.0601

Σχήμα 7.5: Γραφική απεικόνιση αποτελεσμάτων αλγορίθμου 2 για πηγή Μή Βρύση

Παρατηρούμε ότι για τον συγκεκριμένο τύπο εκπαίδευσης, για $H=4$ δόθηκε η μικρότερη τιμή μέσου τετραγωνικού σφάλματος ίση με 0.0553 (l/s)^2 .

Β) Πηγή Πέρα Βρύση:



H	MSE (l/s) ²
1	0.0152
2	0.0229
3	0.022
4	0.0236
5	0.0111
6	0.0246
7	0.0305
8	0.0099
9	0.0052
10	0.0095

Σχήμα 7.6: Γραφική απεικόνιση αποτελεσμάτων αλγορίθμου 2 για πηγή Πέρα Βρύση

Στην πηγή Πέρα Βρύσης, για τον ίδιο τύπο εκπαίδευσης και για $H=9$ δόθηκε η μικρότερη τιμή μέσου τετραγωνικού σφάλματος ίση με 0.0052 (l/s)^2 .

7.3.3 Ανάπτυξη αλγορίθμου με συνάρτηση δημιουργίας δικτύου την newcf (cascade - forward backpropagation network), συνάρτηση εκπαίδευσης την traingdx και συνάρτηση εκμάθησης την learngdm.

I. Συνάρτηση δημιουργίας δικτύου: “newcf”

Στην συγκεκριμένη συνάρτηση, το στρώμα εισόδου κατέχει βάρη, τα οποία προέρχονται από τα δεδομένα εισόδου. Κάθε επακόλουθο στρώμα, έχει συναπτικά βάρη, προερχόμενα όχι μόνο από το προηγούμενο στρώμα (όπως στην περίπτωση της newff), αλλά από έναν συνδυασμό των συναπτικών βαρών όλων των προηγούμενων στρωμάτων και των δεδομένων εισόδου. Η προσαρμογή πραγματοποιείται κατά τη διάρκεια της εκπαίδευσης, όπου ενημερώνονται τα συναπτικά βάρη με καθορισμένη συνάρτηση εκμάθησης (learning function). Η απόδοση αξιολογείται επίσης μέσω προκαθορισμένης συνάρτησης απόδοσης (performance function).

Η διαφοροποίηση της συνάρτησης newcf ως προς τον τρόπο προσαρμογής των συναπτικών βαρών, αυξάνει την αξιοπιστία της εκμάθησης στο δίκτυο και είναι σε θέση να αποδώσει καλύτερα. Αξίζει να σημειωθεί ότι η newcf συνδυάζεται με συγκεκριμένες συναρτήσεις εκπαίδευσης, όπως traingdx, trainlm, trainbr, trainrp, trainingd, κ.α.

II. Συνάρτηση εκπαίδευσης δικτύου: “traingdx”

Πρόκειται για συνάρτηση εκπαίδευσης που ενημερώνει τα συναπτικά βάρη και τις πολώσεις σύμφωνα με τη μέθοδο της απότομης καθόδου με ορμή και προσαρμόσιμο ρυθμό εκμάθησης (gradient descent momentum and an adaptive learning rate). Όπως και στις προηγούμενες συναρτήσεις εκπαίδευσης που αναφέρθηκαν, έτσι και η traingdx είναι σε θέση να εκπαιδεύσει ένα τεχνητό νευρωνικό δίκτυο εφόσον η σχέση που θα προκύψει από τα συναπτικά βάρη, τα δεδομένα εισόδου του δικτύου και οι συναρτήσεις μεταφοράς να είναι παραγωγίσιμη. Η ανάδραση συνίσταται στον υπολογισμό της παραγώγου της συνάρτησης απόδοσης (performance function) σε συνάρτηση με τα συναπτικά βάρη και τις πολώσεις X . Κάθε μεταβλητή μεταβάλλεται σύμφωνα με τη μέθοδο απότομης καθόδου με ορμή, στην οποία αναφερθήκαμε στην προηγούμενη παράγραφο και περιγράφεται με την σχέση:

$$dX = mc * dX_{prev} + lr(1 - mc) * \frac{dperf}{dX} \quad (7.4)$$

, όπου dX_{prev} : Η προηγούμενη μεταβολή σε συναπτικά βάρη ή πολώσεις.

Το βασικό χαρακτηριστικό της συνάρτησης εκπαίδευσης traingdx είναι το εξής: σε περίπτωση όπου η απόδοση μειωθεί κάτω από την τιμή που έχει τεθεί στην παράμετρο goal της συνάρτησης εκμάθησης, τότε ο ρυθμός εκμάθησης αυξάνεται κατά τον παράγοντα lr_inc . Εάν όμως η απόδοση

σε κάποια επανάληψη αυξηθεί περισσότερο από την τιμή που έχει δοθεί στην παράμετρο `max_perf_inc`, τότε ο ρυθμός εκμάθησης ρυθμίζεται από τον παράγοντα `lr_dec` και ακυρώνεται οποιαδήποτε μεταβολή, η οποία θα προκαλούσε αύξηση στη απόδοση. Σημαντικό είναι επίσης και το ότι εφόσον η τιμή της συνάρτησης απόδοσης αυξηθεί περισσότερες φορές από την τιμή που έχει δοθεί στην παράμετρο `max_fail`, τότε η εκπαίδευση του δικτύου διακόπτεται αυτόματα.

Από τα παραπάνω συμπεραίνουμε ότι η συγκεκριμένη συνάρτηση εκπαίδευσης στοχεύει στην ορθότερη εκπαίδευση του δικτύου, καθώς είναι σε θέση να προλαμβάνει και να διακόπτει μία μη σωστή εκπαίδευση. Μπορεί βέβαια η εκπαίδευση να χαρακτηρίζεται ως ελλιπής, αλλά σίγουρα τα αποτελέσματα που θα εξαχθούν από το δίκτυο θα είναι περισσότερα αξιόλογα από το να συνέχιζε η εκπαίδευση μέχρι τέλους.

Οι παράμετροι που ορίζουν την συνάρτηση `trainidx`, καθώς και οι προκαθορισμένες τιμές τους είναι:

- Αριθμός εποχών (επαναλήψεων) εκπαίδευσης (`net.trainParam.epochs=10`)
- Στόχος τιμής συνάρτησης απόδοσης (`net.trainParam.goal=0`)
- Ρυθμός εκμάθησης (`net.trainParam.lr=0.01`)
- Λόγος αύξησης ρυθμού εκμάθησης (`net.trainParam.lr_inc=1.05`)
- Λόγος μείωσης ρυθμού εκμάθησης (`net.trainParam.lr_dec=0.7`)
- Μέγιστος αριθμός αποτυχιών «validation» (`net.trainParam.max_fail=5`)
- Μέγιστη αύξηση απόδοσης (`net.trainParam.max_perf_inc=1.04`)
- Σταθερά της ορμής (`net.trainParam.mc=0.9`)
- Ελάχιστη κλίση της συνάρτησης απόδοσης (`net.trainParam.min_grad=1e-10`)
- Εποχές που μεσολαβούν για την απεικόνιση του αποτελέσματος (`net.trainParam.show=25`)
- Μέγιστος χρόνος εκπαίδευσης σε seconds (`net.trainParam.time=inf`)

Κατόπιν ακολουθεί ο αλγόριθμος που υλοποιήθηκε με τις συναρτήσεις `newcf`, `trainidx` και `learndgm`:

ΑΛΓΟΡΙΘΜΟΣ 3 (neural_newcf_trainidx_learndgm)

```
p = [...]
```

```
t = [...]
```

```
[pn,ps]= mapminmax(p,0,1)
```

```
[tn,ts]= mapminmax(t,0,1)
```

```
[trainV, valV, testV]= dividevec(pn, tn, 0.15, 0.15);
```

```

net= newcf(minmax(pn), [1, H, 1], {'tansig', 'tansig', 'purelin'},
'traingdx', 'learngdm')
%-----
Δημιουργία τεχνητού νευρωνικού δικτύου τριών επιπέδων(μέσω της newcf), με
συναρτήσεις ενεργοποίησης τις tansig, tansig και purelin για κάθε επίπεδο
αντίστοιχα, ενώ η εκπαίδευση του δικτύου θα γίνει μέσω της συνάρτησης
traingdx και η εκμάθηση μέσω της learngdm. Ο αλγόριθμος έτρεξε για H από
1 έως 10.
%-----
net= init(net)

net. trainParam.epochs= 8000;

net. trainParam.show= 100;

net. trainParam.min_grad= 1e-20;

net. trainParam.lr= 0.01;

net. trainParam.mc= 0.9;

net. trainParam.lr_inc= 1.05;

net. trainParam.lr_dec= 0.7;

net. trainParam.max_fail= 10;

net. trainParam.max_perf_inc= 1.04;

net= train(net, trainV.P, trainV.T, [], [], valV, testV);

y= sim(net, testV.P, [], [], testV.T);

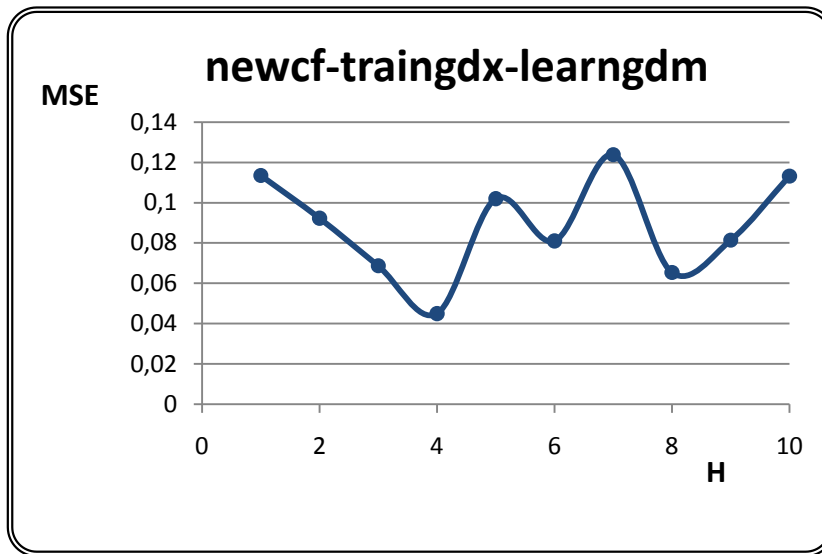
e= testV.T-y;

Perf= mse(e)

end

```

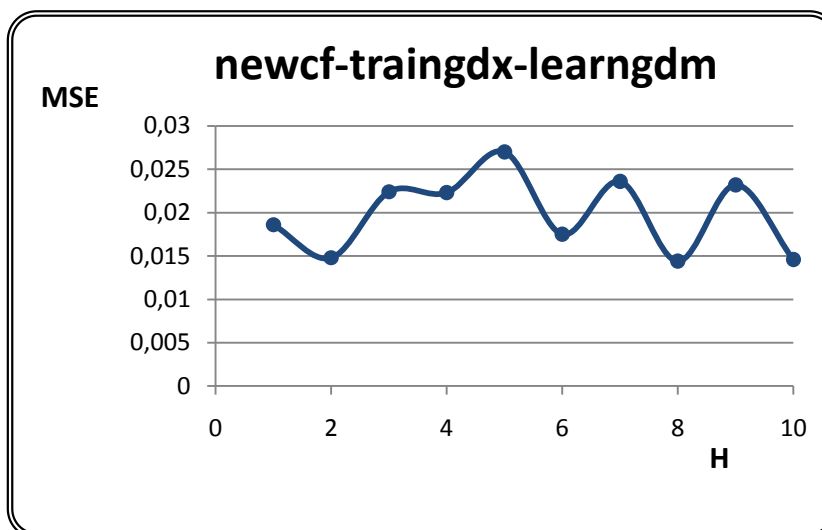
Ο παραπάνω αλγόριθμος εφαρμόστηκε για H από 1 έως 10 και με κριτήριο το mse έδωσε τα εξής αποτελέσματα:

A) Πηγή Μάη Βρύση:

H	MSE (l/s) ²
1	0.1135
2	0.0923
3	0.0687
4	0.0450
5	0.102
6	0.081
7	0.1238
8	0.0654
9	0.0814
10	0.1132

Σχήμα 7.7: Γραφική απεικόνιση αποτελεσμάτων αλγορίθμου 3 για πηγή Μάη Βρύση

Παρατηρούμε ότι για τον συγκεκριμένο τύπο εκπαίδευσης, για $H=4$ δόθηκε η μικρότερη τιμή μέσου τετραγωνικού σφάλματος ίση με $0.0450 (l/s)^2$.

B) Πηγή Πέρα Βρύση:

H	MSE (l/s) ²
1	0.0186
2	0.0148
3	0.0224
4	0.0223
5	0.027
6	0.0175
7	0.0236
8	0.0144
9	0.0232
10	0.0146

Σχήμα 7.8: Γραφική απεικόνιση αποτελεσμάτων αλγορίθμου 3 για πηγή Πέρα Βρύση

Στην πηγή Πέρα Βρύσης, για τον ίδιο τύπο εκπαίδευσης και για $H=8$ δόθηκε η μικρότερη τιμή μέσου τετραγωνικού σφάλματος ίση με $0.0144 (l/s)^2$.

7.3.4 Ανάπτυξη αλγορίθμου με συνάρτηση δημιουργίας δικτύου την *newcf*, συνάρτηση εκπαίδευσης την *trainlm* και συνάρτηση εκμάθησης την *learnqdm*.

Ο αλγόριθμος εκπαίδευσης *trainlm* θεωρείται αρκετά γρήγορος και αξιόπιστος. Τα υπολογιζόμενα συναπτικά βάρη και πολώσεις υπολογίζονται μέσω του αλγορίθμου “Levenberg-Marquardt”.

Όπως και στις συναρτήσεις που προηγήθηκαν η σχέση που θα προκύπτει από τα συναπτικά βάρη, τα δεδομένα εισόδου του δικτύου και τις συναρτήσεις ενεργοποίησης κάθε στρώματος, θα πρέπει να είναι παραγωγίσιμη.

Αντίθετα με άλλες συναρτήσεις εκπαίδευσης, η *trainlm* υποθέτει ότι το δίκτυο κατέχει ως συνάρτηση απόδοσης το μέσο τετραγωνικό σφάλμα, *mse* (mean square error). Η υπόθεση αυτή αποτελεί βασικό χαρακτηριστικό του αλγορίθμου Levenberg-Marquardt.

Η ανάδραση συνίσταται στον υπολογισμό του Ιακωμβιανού πίνακα της συνάρτησης απόδοσης σε σχέση με τα συναπτικά βάρη και τις πολώσεις *X*. Κάθε μεταβλητή ρυθμίζεται μέσω του αλγορίθμου Levenberg-Marquardt,

$$jj = jX * jX \quad (7.5)$$

$$je = jX * E \quad (7.6)$$

$$dX = -(jj + I * mu)/je \quad (7.7)$$

, όπου:

E: πίνακας όλων των σφαλμάτων,

I: μοναδιαίος πίνακας,

jX: Ιακωμβιανός πίνακας των συναπτικών βαρών και των πολώσεων,

je: Ιακωμβιανός πίνακας της συνάρτησης απόδοσης,

mu: παράμετρος προσαρμοστικότητας δικτύου και

dX: πίνακας της μεταβολής των συναπτικών βαρών και των πολώσεων.

Πιο αναλυτικά, ο αλγόριθμος Levenberg-Marquardt χρησιμοποιεί την κατεύθυνση εκείνη για να επιλυθεί που αποτελεί την λύση του γραμμικού συστήματος των εξισώσεων:

$$(J(x_k)^T J(x_k) + \lambda_k I) d_k = -J(x_k) F(x_k) \quad (7.8)$$

, όπου ο αριθμός λ_k ρυθμίζει την ποσότητα και την διεύθυνση του d_k . Όταν το λ_k είναι μηδέν, το d_k είναι το ίδιο με εκείνο στην μέθοδο Gauss-Newton. Όσο το λ_k τείνει στο άπειρο, το d_k τείνει σε μηδενικό διάνυσμα. Αυτό σημαίνει ότι για ένα αρκετά μεγάλο λ_k ισχύει η ανισότητα:

$$F(x_k + d_k) < F(x_k) \quad (7.9)$$

Η τιμή της παραμέτρου μ αυξάνεται κατά τον παράγοντα μ_inc έως ότου η αλλαγή των παραπάνω τριών σχέσεων να καταλήξει σε μειωμένη τιμή της συνάρτησης απόδοσης. Αφού πραγματοποιηθεί η αλλαγή στα συναπτικά βάρη και τις πολώσεις, η παράμετρος μ μειώνεται κατά τον παράγοντα μ_dec .

Η παράμετρος mem_reduc ουσιαστικά δείχνει πως θα χρησιμοποιηθεί η μνήμη και η ταχύτητα του υπολογιστή στον υπολογισμό του Ιακωμβιανού πίνακα JX . Εάν η τιμή της παραμέτρου mem_reduc τεθεί ίση με τη μονάδα, τότε η συνάρτηση `trainlm` εκπαιδεύει το δίκτυο να μην με μεγαλύτερη ταχύτητα, αλλά με την απαίτηση περισσότερης μνήμης. Στην περίπτωση που η τιμή αυξηθεί από 1 σε 2, τότε η απαίτηση μνήμης μειώνεται στο μισό, αλλά μειώνεται επίσης και η ταχύτητα εκπαίδευσης. Γενικά, όσο υψηλότερες τιμές θέτονται στην συγκεκριμένη παράμετρο, τόσο μειώνεται η απαίτηση μνήμης και ταυτόχρονα μειώνεται και η ταχύτητα εκπαίδευσης του δικτύου.

Οι παράμετροι που ορίζουν την συνάρτηση `traindxdx`, καθώς και οι προκαθορισμένες τιμές τους είναι:

- Αριθμός εποχών (επαναλήψεων) εκπαίδευσης (`net.trainParam.epochs=100`)
- Στόχος τιμής συνάρτησης απόδοσης (`net.trainParam.goal=0`)
- Παράγοντας ανταλλαγής μνήμης/ταχύτητας (`net.trainParam.mem_reduc=1`)
- Αρχική τιμή μ (`net.trainParam.mu=0.001`)
- Παράγοντας μείωσης μ (`net.trainParam.mu_dec=0.1`)
- Παράγοντας αύξησης μ (`net.trainParam.mu_inc=10`)
- Μέγιστη τιμή μ (`net.trainParam.mu_max=1e10`)
- Μέγιστος αριθμός αποτυχιών «validation» (`net.trainParam.max_fail=5`)
- Ελάχιστη κλίση της συνάρτησης απόδοσης (`net.trainParam.min_grad=1e-10`)
- Εποχές που μεσολαβούν για την απεικόνιση του αποτελέσματος (`net.trainParam.show=25`)
- Μέγιστος χρόνος εκπαίδευσης σε seconds (`net.trainParam.time=inf`)

Στη συνέχεια ακολουθεί ο αλγόριθμος που χρησιμοποιήθηκε για την επιλογή του βέλτιστου αριθμού νευρώνων του ενδιάμεσου στρώματος με τον συγκεκριμένο συνδυασμό συναρτήσεων, καθώς και τα αποτελέσματα που προέκυψαν για την κάθε πηγή.

ΑΛΓΟΡΙΘΜΟΣ 4 (neural newcf trainlm learnngdm)

```

p = [...]

t = [...]

[pn,ps]= mapminmax(p,0,1)

[tn,ts]= mapminmax(t,0,1)

[trainV, valV, testV]= dividevec(pn, tn, 0.15, 0.15);

net= newcf(minmax(pn), [1, H, 1], {'tansig', 'tansig', 'purelin'},
'trainlm', 'learnngdm')
%-----
Δημιουργία τεχνητού νευρωνικού δικτύου τριών επιπέδων(μέσω της newcf), με
συναρτήσεις ενεργοποίησης τις tansig, tansig και purelin για κάθε επίπεδο
αντίστοιχα, ενώ η εκπαίδευση του δικτύου θα γίνει μέσω της συνάρτησης
trainlm και η εκμάθηση μέσω της learnngdm. Ο αλγόριθμος έτρεξε για H από 1
εως 10.
%-----
net= init(net)

net. trainParam.epochs= 8000;

net. trainParam.show= 100;

net. trainParam.min_grad= 1e-20;

net. trainParam.max_fail= 10;

net. trainParam.mem_reduc= 1;

net. trainParam.mu_inc= 10;

net. trainParam.mu_dec= 0.10;

net. trainParam.mu_max= 1e+010;

net= train(net, trainV.P, trainV.T, [], [], valV, testV);

y= sim(net, testV.P, [], [], testV.T);

e= testV.T-y;

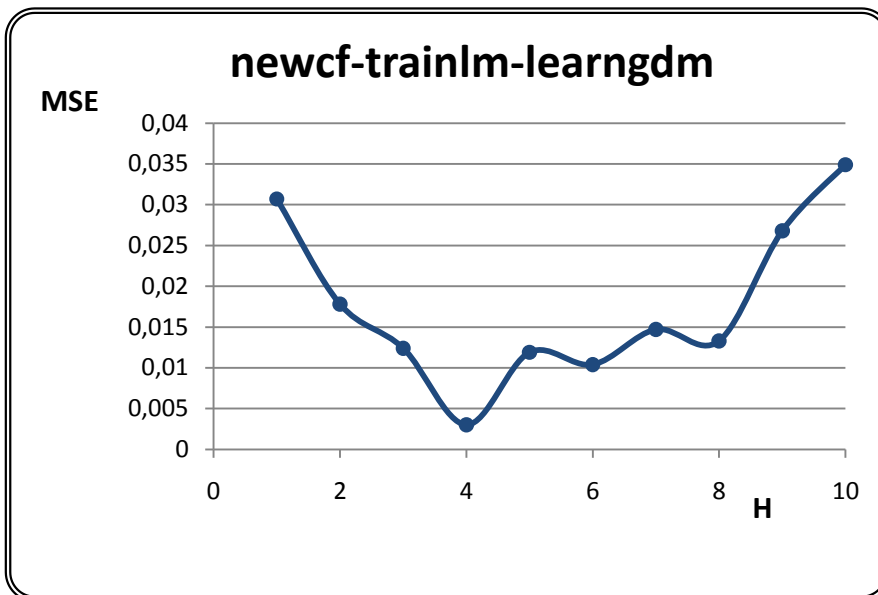
Perf= mse(e)

end

```

Ο παραπάνω αλγόριθμος εφαρμόστηκε για H από 1 έως 10 και με κριτήριο το mse έδωσε τα εξής αποτελέσματα:

A) Πηγή Μάη Βρύση:

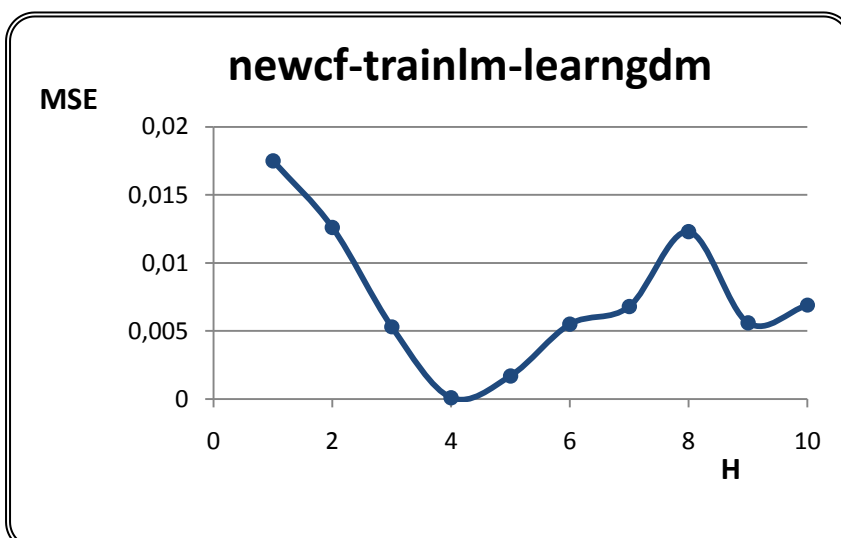


H	MSE
	$(l/s)^2$
1	0.0307
2	0.0178
3	0.0124
4	0.0030
5	0.0119
6	0.0104
7	0.0147
8	0.0133
9	0.0268
10	0.0349

Σχήμα 7.9: Γραφική απεικόνιση αποτελεσμάτων αλγορίθμου 4 για πηγή Μάη Βρύση

Παρατηρούμε ότι για τον συγκεκριμένο τύπο εκπαίδευσης, για $H=4$ δόθηκε η μικρότερη τιμή μέσου τετραγωνικού σφάλματος ίση με $0.0030 (l/s)^2$.

B) Πηγή Πέρα Βρύση:



H	MSE
	$(l/s)^2$
1	0.0175
2	0.0126
3	0.0053
4	0.0001
5	0.0017
6	0.0055
7	0.0068
8	0.0123
9	0.0056
10	0.0069

Σχήμα 7.10: Γραφική απεικόνιση αποτελεσμάτων αλγορίθμου 4 για πηγή Πέρα Βρύση

Στην πηγή Πέρα Βρύσης, για τον ίδιο τύπο εκπαίδευσης και για $H=4$ δόθηκε η μικρότερη τιμή μέσου τετραγωνικού σφάλματος ίση με 0.0001 (l/s)^2 .

7.3.5 Ανάπτυξη αλγορίθμου με συνάρτηση δημιουργίας δικτύου την newcf, συνάρτηση εκπαίδευσης την trainlm, συνάρτηση εκμάθησης την learnsgdm και συνάρτηση απόδοσης σφάλματος την msereg.

Οι συναρτήσεις “newcf”, “learnsgdm” και “trainlm” που αναλύθηκαν προηγουμένως χρησιμοποιούνται και σε αυτόν τον αλγόριθμο.

Συνάρτηση απόδοσης σφάλματος: “msereg”

Η συνάρτηση msereg ουσιαστικά προκύπτει από δύο άλλες συναρτήσεις: την mse και msw. Η συνάρτηση μέσου τετραγωνικού σφάλματος, mse ορίζεται από την σχέση:

$$mse = \frac{1}{N} \cdot \sum_{i=1}^N (e_i)^2 = \frac{1}{N} \cdot \sum_{i=1}^N (t_i - p_i)^2 \quad (7.10)$$

, όπου e_i : διαφορά μεταξύ διανύσματος εισόδου p_i και του διανύσματος επιθυμητής εξόδου t_i .

Η συνάρτηση μέσου τετραγωνικού αθροίσματος των συναπτικών βαρών και των πολώσεων του δικτύου, msw δίνεται από την παρακάτω σχέση:

$$msw = \frac{1}{n} \cdot \sum_{j=1}^n (w_j)^2 \quad (7.11)$$

, όπου w_j : τα συναπτικά βάρη και οι πολώσεις του δικτύου.

Η σχέση που ορίζει την συνάρτηση απόδοσης σφάλματος msereg δίνεται από την σχέση:

$$msereg = \gamma \cdot mse + (1 - \gamma) \cdot msw \quad (7.12)$$

, όπου γ : η κλίση ratio της msereg και παίρνει τιμές στο διάστημα $[0,1]$.

Για τιμή του ratio μικρότερη από 0.5 αποδίδεται μεγαλύτερη βαρύτητα στην συνάρτηση msw σε σχέση με την mse και άρα μικρότερη μεταβολή στα συναπτικά βάρη και στις πολώσεις. Αντιθέτως, μεγαλύτερη τιμή του 0.5, αποδίδει μεγαλύτερη βαρύτητα στην mse και συνεπώς μεγαλύτερη μεταβολή στις τιμές των συναπτικών βαρών και των πολώσεων.

Η χρήση της msereg ως συνάρτηση απόδοσης, οδηγεί σε μικρότερες τιμές συναπτικών βαρών και πολώσεων του δικτύου, καθώς και σε πιο «ήπια» αντίδραση του δικτύου σε νέες εισόδους δεδομένων.

Ο αλγόριθμος που ακολουθεί διαφέρει από τον προηγούμενο μονάχα ως προς την χρήση της msereg ως συνάρτηση απόδοσης και της δήλωσης της παραμέτρου ratio.

ΑΛΓΟΡΙΘΜΟΣ 5 (neural newcf trainlm learngdm msereg)

```

p = [...]

t = [...]

[pn,ps]= mapminmax(p,0,1)

[tn,ts]= mapminmax(t,0,1)

[trainV, valV, testV]= dividevec(pn, tn, 0.15, 0.15);

net= newcf(minmax(pn), [1, H, 1], {'tansig', 'tansig', 'purelin'},
'trainlm', 'learngdm')

net= init(net)

net. performFcn= 'msereg';

net. performParam.ratio= 0.5;

net. trainParam.epochs= 8000;

net. trainParam.show= 100;

net. trainParam.min_grad= 1e-20;

net. trainParam.max_fail= 10;

net. trainParam.mem_reduc= 1;

net. trainParam.mu_inc= 10;

net. trainParam.mu_dec= 0.10;

net. trainParam.mu_max= 1e+010;

net= train(net, trainV.P, trainV.T, [], [], valV, testV);
y= sim(net, testV.P, [], [], testV.T);

e= testV.T-y;

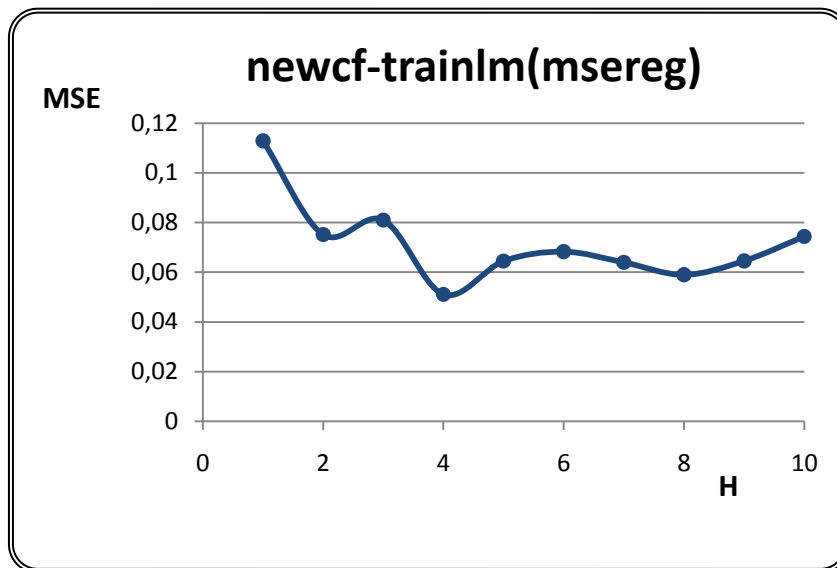
Perf= msereg(e)

end

```

Ο παραπάνω αλγόριθμος εφαρμόστηκε για H από 1 έως 10 και με κριτήριο το mse έδωσε τα εξής αποτελέσματα:

A) Πηγή Μάη Βρύση:

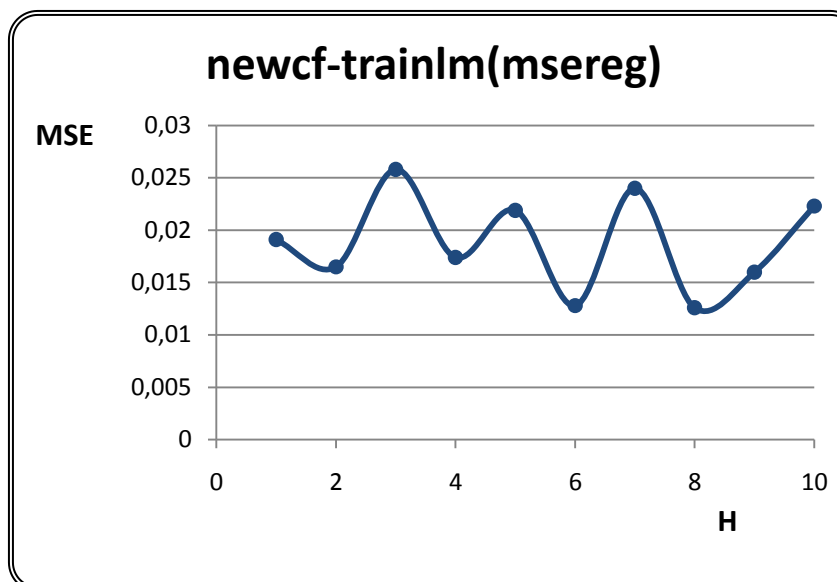


H	MSE (l/s) ²
1	0.1129
2	0.0752
3	0.081
4	0.0511
5	0.0645
6	0.0683
7	0.064
8	0.059
9	0.0646
10	0.0744

Σχήμα 7.11: Γραφική απεικόνιση αποτελεσμάτων αλγορίθμου 5 για πηγή Μάη Βρύση

Παρατηρούμε ότι για τον συγκεκριμένο τύπο εκπαίδευσης, για $H=4$ δόθηκε η μικρότερη τιμή μέσου τετραγωνικού σφάλματος ίση με 0.0511 (l/s)^2 .

B) Πηγή Πέρα Βρύση:



H	MSE (l/s) ²
1	0.0191
2	0.0165
3	0.0258
4	0.0174
5	0.0219
6	0.0128
7	0.024
8	0.0126
9	0.016
10	0.0223

Σχήμα 7.12: Γραφική απεικόνιση αποτελεσμάτων αλγορίθμου 5 για πηγή Πέρα Βρύση

Στην πηγή Πέρα Βρύσης, για τον ίδιο τύπο εκπαίδευσης και για $H=8$ δόθηκε η μικρότερη τιμή μέσου τετραγωνικού σφάλματος ίση με 0.0126 (l/s)^2 .

7.3.6 Ανάπτυξη αλγορίθμου με συνάρτηση δημιουργίας δικτύου την *newcf*, συνάρτηση εκπαίδευσης την *trainbr* και συνάρτηση εκμάθησης την *learn_gdm*.

Συνάρτηση εκπαίδευσης δικτύου: “trainbr”

Η συνάρτηση *trainbr* ενημερώνει τα συναπτικά βάρη και τις πολώσεις σύμφωνα με την βελτιστοποίηση Levenberg-Marquardt. Ουσιαστικά ελαχιστοποιεί έναν γραμμικό συνδυασμό των τετραγωνικών σφαλμάτων και των συναπτικών βαρών και έπειτα τροποποιεί τον γραμμικό συνδυασμό έτσι ώστε κατά το τέλος της εκπαίδευσης το προκύπτον δίκτυο να κατέχει καλές ιδιότητες γενίκευσης. Η παραπάνω διαδικασία καλείται κανονικοποίηση κατά Bayesian, καθώς η συνάρτηση εκπαίδευσης *trainbr* ακολουθεί το πρότυπο της κατανομής Bayesian. Στην περίπτωση αυτή τα συναπτικά βάρη και οι πολώσεις ορίζονται σαν τυχαίες μεταβλητές με καθορισμένες κατανομές. Σημαντικό είναι το ότι οι παράμετροι κανονικοποίησης του νευρωνικού δικτύου συνδέονται με τις διακυμάνσεις των κατανομών αυτών και εκτιμούνται μέσω στατιστικών τεχνικών.

Η ανάδραση συνίσταται στον υπολογισμό του Ιακωβιανού πίνακα jX της συνάρτησης απόδοσης βάσει των συναπτικών βαρών και των πολώσεων X . Κάθε μεταβλητή προσαρμόζεται σύμφωνα με τον αλγόριθμο Levenberg-Marquardt,

$$jj = jX * jX \quad (7.5)$$

$$je = jX * E \quad (7.6)$$

$$dX = -(jj + I * \mu) / je \quad (7.7)$$

, όπου:

E: πίνακας όλων των σφαλμάτων,

I: μοναδιαίος πίνακας,

jX: Ιακωβιανός πίνακας των συναπτικών βαρών και των πολώσεων,

je: Ιακωβιανός πίνακας της συνάρτησης απόδοσης,

mu: παράμετρος προσαρμοστικότητας δικτύου και

dX: πίνακας της μεταβολής των συναπτικών βαρών και των πολώσεων

Αξίζει να σημειωθεί ότι όσο μεγαλύτερο το πλήθος των ιστορικών στοιχείων που εισάγονται στο δίκτυο ως δεδομένα προς εκπαίδευση, τόσο μεγαλύτερος είναι ο αριθμός των παραμέτρων που απαιτεί η *trainbr* για τον καθορισμό της εκπαίδευσης.

Οι παράμετροι που ορίζουν την συνάρτηση `trainbr`, καθώς και οι προκαθορισμένες τιμές τους είναι:

- Αριθμός εποχών (επαναλήψεων) εκπαίδευσης (`net.trainParam.epochs=100`)
- Στόχος τιμής συνάρτησης απόδοσης (`net.trainParam.goal=0`)
- Παράγοντας ανταλλαγής μνήμης/ταχύτητας (`net.trainParam.mem_reduc=1`)
- Ρύθμιση παραμέτρου Marquardt (`net.trainParam.mu=0.005`)
- Παράγοντας μείωσης μ (`net.trainParam.mu_dec=0.1`)
- Παράγοντας αύξησης μ (`net.trainParam.mu_inc=10`)
- Μέγιστη τιμή μ (`net.trainParam.mu_max=1e10`)
- Μέγιστος αριθμός αποτυχιών «validation» (`net.trainParam.max_fail=5`)
- Ελάχιστη κλίση της συνάρτησης απόδοσης (`net.trainParam.min_grad=1e-10`)
- Εποχές που μεσολαβούν για την απεικόνιση του αποτελέσματος (`net.trainParam.show=25`)
- Μέγιστος χρόνος εκπαίδευσης σε seconds (`net.trainParam.time=inf`)

Κατόπιν ακολουθεί ο αλγόριθμος που υλοποιήθηκε με τις συναρτήσεις `newcf`, `trainbr` και `learndgm`:

ΑΛΓΟΡΙΘΜΟΣ 6 (neural newcf trainbr learngdm)

```
p = [...]  
t = [...]  
  
[pn,ps]= mapminmax(p,0,1)  
[tn,ts]= mapminmax(t,0,1)  
  
[trainV, valV, testV]= dividevec(pn, tn, 0.15, 0.15);  
  
net= newcf(minmax(pn), [1, H, 1], {'tansig', 'tansig', 'purelin'},  
'trainbr', 'learngdm')  
  
net= init(net)  
  
net.trainParam.epochs= 8000;  
  
net.trainParam.show= 100;  
  
net.trainParam.min_grad= 1e-20;  
  
net.trainParam.lr= 0.01;  
  
net.trainParam.mu_max= 1e+030;
```

```

net= train(net, trainV.P, trainV.T, [], [], valV, testV);

y= sim(net, testV.P, [], [], testV.T);

e= testV.T-y;

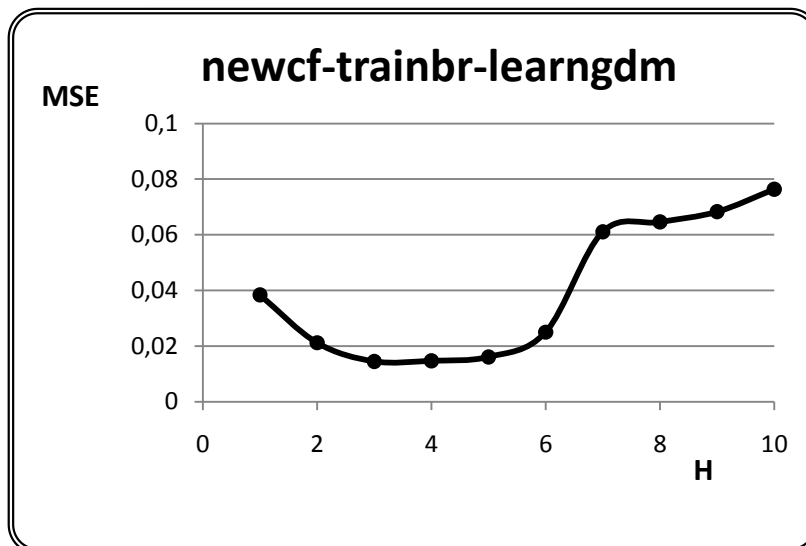
Perf= mse(e)

end

```

Ο παραπάνω αλγόριθμος εφαρμόστηκε για H από 1 έως 10 και με κριτήριο το mse έδωσε τα εξής αποτελέσματα:

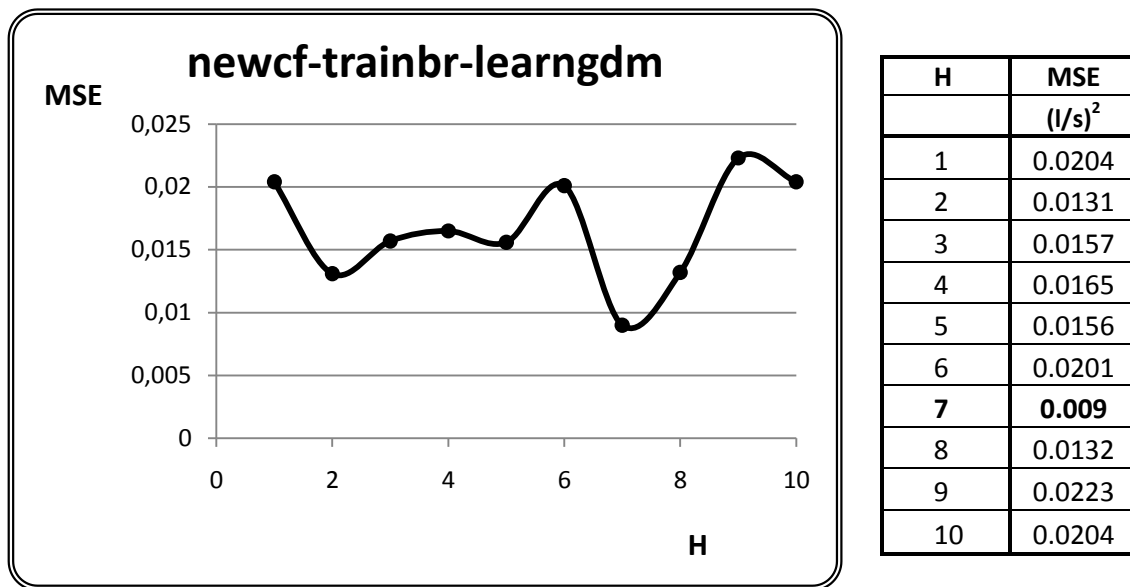
A) Πηγή Μάη Βρύση:



H	MSE (l/s) ²
1	0.0384
2	0.0212
3	0.0145
4	0.0147
5	0.0161
6	0.025
7	0.0611
8	0.0646
9	0.0683
10	0.0764

Σχήμα 7.13: Γραφική απεικόνιση αποτελεσμάτων αλγορίθμου 6 για πηγή Μάη Βρύση

Παρατηρούμε ότι για τον συγκεκριμένο τύπο εκπαίδευσης, για $H=3$ δόθηκε η μικρότερη τιμή μέσου τετραγωνικού σφάλματος ίση με $0.0145 (l/s)^2$.

Β) Πηγή Πέρα Βρύση:

Σχήμα 7.14: Γραφική απεικόνιση αποτελεσμάτων αλγορίθμου 6 για πηγή Πέρα Βρύση

Στην πηγή Πέρα Βρύσης, για τον ίδιο τύπο εκπαίδευσης και για $H=7$ δόθηκε η μικρότερη τιμή μέσου τετραγωνικού σφάλματος ίση με 0.009 (l/s)^2 .

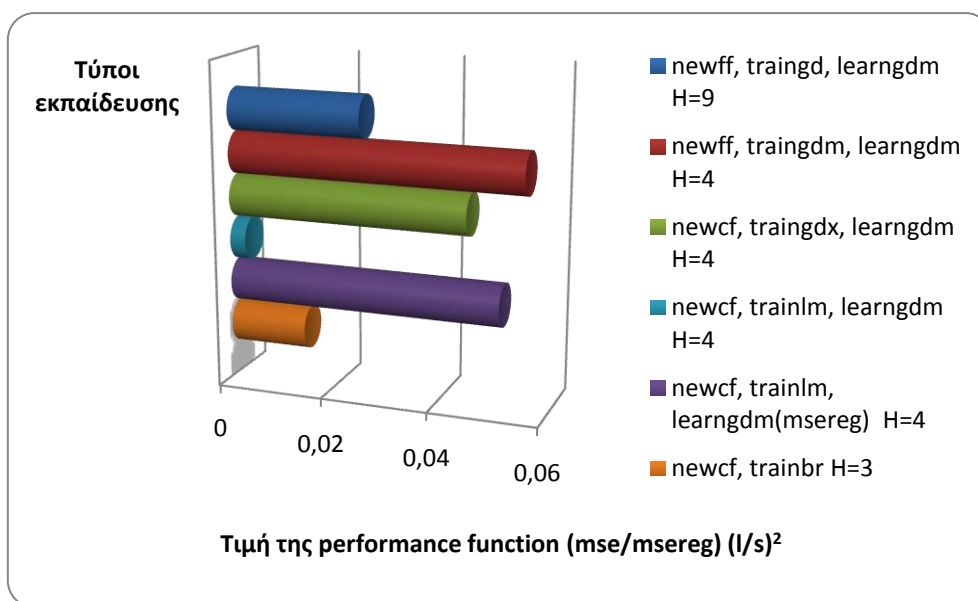
7.3.7 Συγκεντρωτικά αποτελέσματα αλγορίθμων – Σύγκριση και επιλογή βέλτιστου τύπου εκπαίδευσης και πλήθους νευρώνων ενδιάμεσου στρώματος.

Όσον αφορά την πηγή Μάη Βρύσης ακολουθεί συγκεντρωτικός πίνακας με τις βέλτιστες τιμές μέσου τετραγωνικού σφάλματος ανά τύπο εκπαίδευσης, καθώς και ο αριθμός των νευρώνων του ενδιάμεσου στρώματος που αντιστοιχεί στην βέλτιστη αυτή τιμή.

Α) Πηγή Μάη Βρύση:

1	newff, traingd, learnngdm	H=9	MSE=	0.0253 (l/s) ²
2	newff, traingdm, learnngdm	H=4	MSE=	0.0553 (l/s) ²
3	newcf, traingdx, learnngdm	H=4	MSE=	0.0450 (l/s) ²
4	newcf, trainlm, learnngdm (mse)	H=4	MSE=	0.003 (l/s) ²
5	newcf, trainlm, learnngdm (msereg)	H=4	MSEREG=	0.0511 (l/s) ²
6	newcf, trainbr	H=3	MSE=	0.0145 (l/s) ²

Πίνακας 7.2: Συγκεντρωτικά αποτελέσματα αλγορίθμων για την πηγή Μάη Βρύση



Σχήμα 7.15: Συγκεντρωτικά αποτελέσματα αλγορίθμων για την πηγή Μάη Βρύση

Παρατηρούμε επομένως ο τύπος εκπαίδευσης που έδωσε την μικρότερη τιμή μέσου τετραγωνικού σφάλματος πραγματοποιήθηκε μέσω της συνδυαστικής χρήσης των συναρτήσεων: newcf, trainlm και learngdm. Το κανονικοποιημένο μέσο τετραγωνικό σφάλμα πήρε την τιμή $mse = 0.0030 (l/s)^2$ για πλήθος νευρώνων κρυφού στρώματος $H=4$.

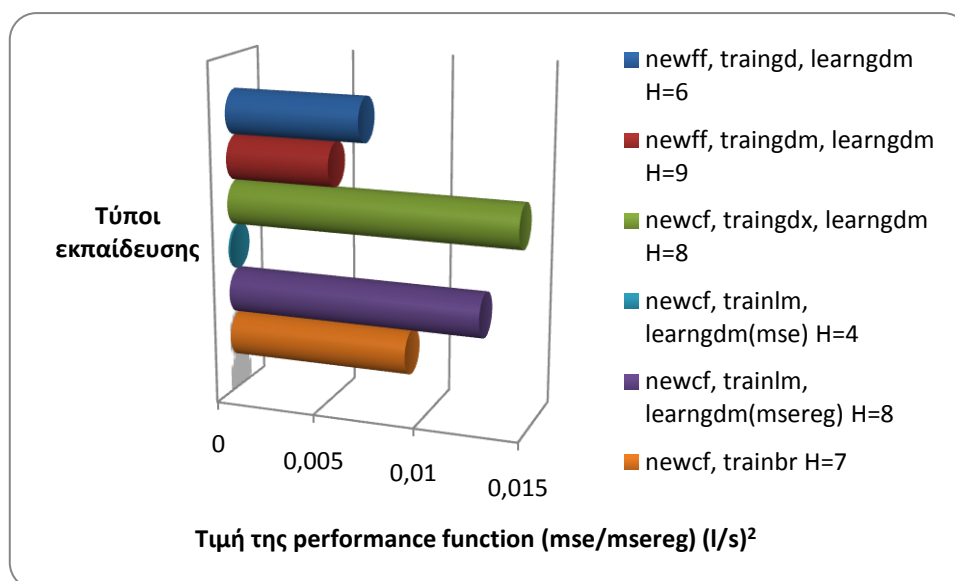
Αξιοσημείωτο είναι το ότι σε τέσσερις από τους έξι αλγόριθμους που υλοποιήθηκαν, σημειώθηκε το βέλτιστο mse για πλήθος νευρώνων ενδιάμεσου στρώματος ίσο με τέσσερα. Το γεγονός αυτό ενδεχομένως να οφείλεται στο ότι για την συγκεκριμένη πηγή ήταν διαθέσιμα περισσότερα στοιχεία δεδομένων, τα οποία εισήχθησαν στο νευρωνικό δίκτυο. Μεγαλύτερο πλήθος δεδομένων εισόδου, συνήθως εξασφαλίζει καλύτερη εκπαίδευση του δικτύου, η οποία συγκλίνει στην χρήση τεσσάρων νευρώνων στο ενδιάμεσο, κρυφό στρώμα του νευρωνικού δικτύου.

Ομοίως και για την δεύτερη πηγή:

Β) Πηγή Πέρα Βρύση:

1	newff, traingd, learngdm	H=6	MSE=	0.0067 (l/s) ²
2	newff, traingdm, learngdm	H=9	MSE=	0.0052 (l/s) ²
3	newcf, traingdx, learngdm	H=8	MSE=	0.0144 (l/s) ²
4	newcf, trainlm, learngdm (mse)	H=4	MSE=	0.0001 (l/s) ²
5	newcf, trainlm, learngdm (msereg)	H=8	MSEREG=	0.0126 (l/s) ²
6	newcf, trainbr	H=7	MSE=	0.009 (l/s) ²

Πίνακας 7.3: Συγκεντρωτικά αποτελέσματα αλγορίθμων για την πηγή Πέρα Βρύση



Σχήμα 7.16: Συγκεντρωτικά αποτελέσματα αλγορίθμων για την πηγή Πέρα Βρύση

Στην συγκεκριμένη πηγή το βέλτιστο μέσο τετραγωνικό σφάλμα (των κανονικοποιημένων τιμών) αποδόθηκε για τύπο εκπαίδευσης με συναρτήσεις: newcf, trainlm και learnsgdm, με τιμή $mse=0.001$ (l/s)² και αριθμό νευρώνων ενδιάμεσου στρώματος ίσο με τέσσερα.

Παρατηρούμε ότι και για τις δύο πηγές ο αλγόριθμος, που υλοποιήθηκε με συναρτήσεις τις newcf, trainlm και learnsgdm έδωσε το βέλτιστο μέσο τετραγωνικό σφάλμα. Το γεγονός αυτό ενδεχομένως να οφείλεται στην ίδια τη φύση του αλγορίθμου, καθώς η συνάρτηση εκπαίδευσης trainlm θεωρείται αρκετά πιο αξιόπιστη και γρήγορη, σε σχέση με πολλές από τις υπόλοιπες συναρτήσεις που δοκιμάστηκαν στην συγκεκριμένη εφαρμογή.

7.4 Αναλυτική παρουσίαση βέλτιστων αποτελεσμάτων

Στην παρούσα παράγραφο παρουσιάζονται και αναλύονται τα αποτελέσματα της εκτέλεσης του βέλτιστου τύπου εκπαίδευσης του νευρωνικού δικτύου και επιχειρείται εκτίμηση της συμπεριφοράς του. Η διαδικασία αυτή ακολουθήθηκε και για τις δύο πηγές Μάη και Πέρα Βρύση του Δήμου Ρούβα.

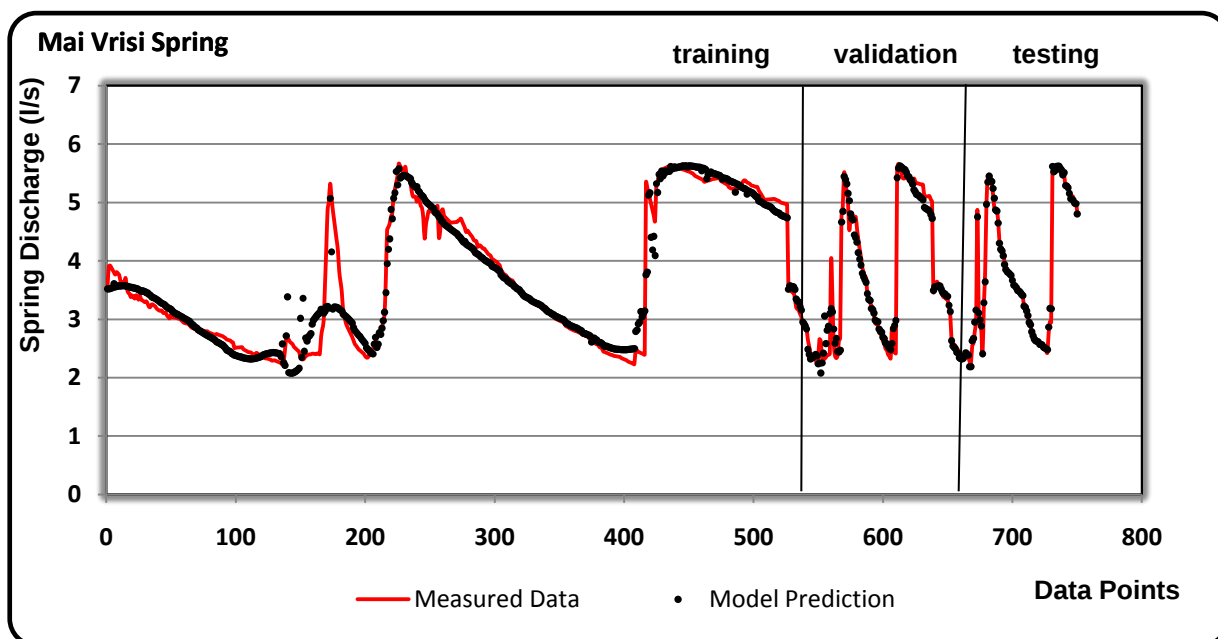
7.4.1 Αποτελέσματα εκτέλεσης νευρωνικού δικτύου για την πηγή Μάη Βρύσης.

Για την πηγή Μάη Βρύση το βέλτιστο αποτέλεσμα μέσω χρήσης των συναρτήσεων newcf, trainlm και learnsgdm και αριθμό νευρώνων ενδιάμεσου στρώματος μετά τις δοκιμές βρέθηκε ίσο με τέσσερα.

Το δίκτυο όπως έχει ήδη περιγραφεί, διαθέτει τρία επίπεδα. Ένα στρώμα εισόδου με δύο νευρώνες (βροχόπτωση και αριθμός ημέρας), ένα κρυφό στρώμα με τέσσερις νευρώνες και ένα στρώμα εξόδου με έναν νευρώνα.

Ο αλγόριθμος χώρισε με τυχαίο τρόπο το 15% των διαθέσιμων δεδομένων για αξιολόγηση, το 15% επίσης των δεδομένων για έλεγχο και το υπόλοιπο ποσοστό χρησιμοποιήθηκε για την εκπαίδευση του νευρωνικού δικτύου.

Στη συνέχεια παρουσιάζεται το γράφημα των αποτελεσμάτων που απέδωσε το νευρωνικό δίκτυο κατά τη φάση της εκπαίδευσης (training), της αξιολόγησης (validation) και του ελέγχου (testing) για την πηγή Μάη Βρύσης, με τιμή κανονικοποιημένου μέσου τετραγωνικού σφάλματος (κατά τον έλεγχο) ίση με $mse=0.0030$ (l/s)².



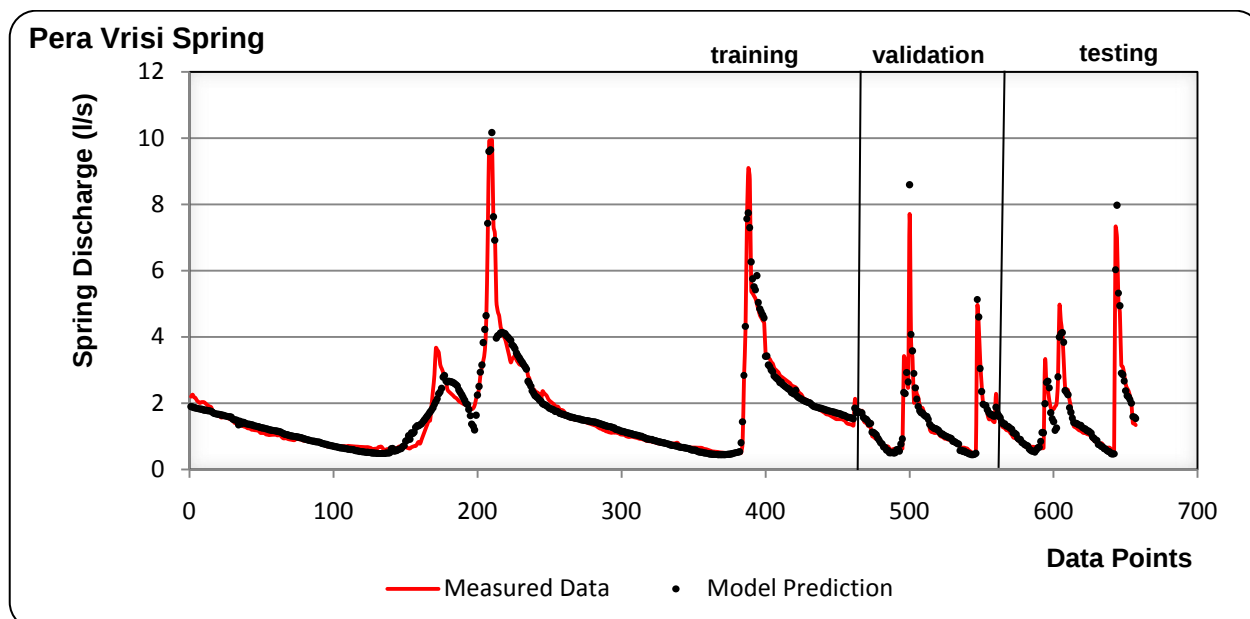
Σχήμα 7.17: Αποτελέσματα εκπαίδευσης, αξιολόγησης και ελέγχου του νευρωνικού δικτύου

Από το παραπάνω γράφημα παρατηρούμε ότι το νευρωνικό δίκτυο απέδωσε πολύ καλά αποτελέσματα, καθώς στη φάση του ελέγχου οι παραχθείσες τιμές παροχής προσεγγίζουν κατά πολύ τις μετρήσεις που συγκεντρώθηκαν. Σημαντικό είναι επίσης το ότι το νευρωνικό δίκτυο είναι σε θέση να αντιληφθεί τις κορυφές χωρίς να τις υπερεκτιμήσει. Επομένως αξιολογούμε ότι ο τύπος του νευρωνικού δικτύου για την συγκεκριμένη εφαρμογή δίνει αποτελέσματα τα οποία ακολουθούν την τάση των δεδομένων ελέγχου.

7.4.2 Αποτελέσματα εκτέλεσης νευρωνικού δικτύου για την πηγή Πέρα Βρύσης.

Ομοίως και για την πηγή Πέρα Βρύση δόθηκε το βέλτιστο αποτέλεσμα μέσω χρήσης των συναρτήσεων `newcf`, `trainlm` και `learnngdm` και αριθμό νευρώνων ενδιάμεσου στρώματος ίσο με τέσσερα.

Ακολουθεί το γράφημα των αποτελεσμάτων του νευρωνικού δικτύου για τη φάση της εκπαίδευσης, αξιολόγησης και ελέγχου της συγκεκριμένης πηγής με τιμή κανονικοποιημένου μέσου τετραγωνικού σφάλματος (κατά τον έλεγχο) ίση με $mse=0.0001 \text{ (l/s)}^2$.



Σχήμα 7.18: Αποτελέσματα εκπαίδευσης, αξιολόγησης και ελέγχου του νευρωνικού δικτύου

7.5 Έλεγχος επίδρασης γραμμικής παρεμβολής - Αναλυτική παρουσίαση αποτελεσμάτων

Στην ενότητα αυτή στόχος ήταν η εξακρίβωση της επιρροής της γραμμικής παρεμβολής στα δεδομένα. Γι αυτόν τον λόγο, εισήχθησαν στο νευρωνικό δίκτυο τα αρχικά δεδομένα, χωρίς την εφαρμογή της γραμμικής παρεμβολής και σκοπός ήταν η σύγκριση των αποτελεσμάτων του νευρωνικού δικτύου. Ουσιαστικά θέλουμε να διαπιστώσουμε αν η χρήση της γραμμικής παρεμβολής ήταν ευεργετική ή όχι για το νευρωνικό δίκτυο. Η διαδικασία αυτή ακολουθήθηκε και για τις δύο πηγές Μάη και Πέρα Βρύση του Δήμου Ρούβα.

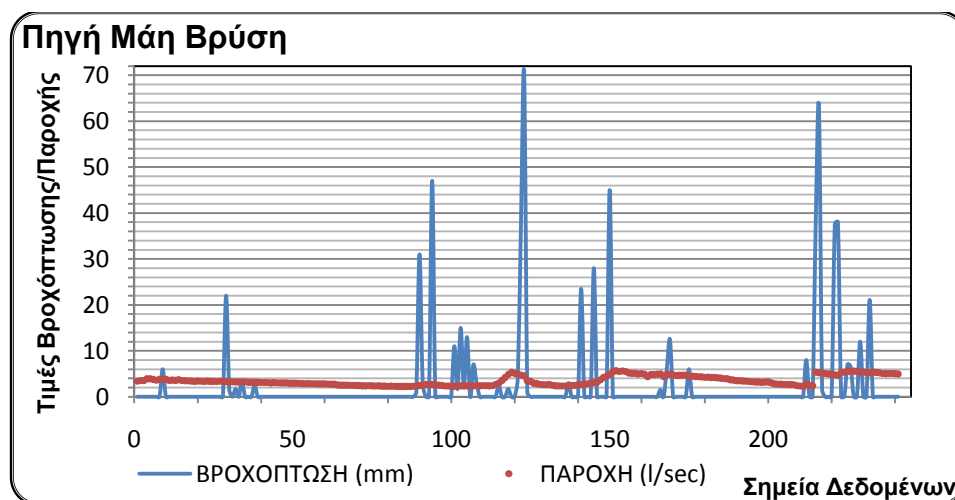
Αρχικά έπρεπε να επαναληφθεί η διαδικασία εύρεσης του μέγιστου συντελεστή συσχέτισης των χρονοσειρών της βροχόπτωσης και παροχής, με τον ίδιο ακριβώς τρόπο που περιγράφηκε σε προηγούμενη παράγραφο.

TIME LAG (μέρες)	Συντελεστής Συσχέτισης		TIME LAG (μέρες)	Συντελεστής Συσχέτισης	
	Πηγή Μάη Βρύση	Πηγή Πέρα Βρύση		Πηγή Μάη Βρύση	Πηγή Πέρα Βρύση
0	-0,0935	0,0167	16	0,1832	0,2737
1	-0,0534	-0,001	17	0,1447	0,2682
2	0,0352	0,0425	18	0,1509	0,1704
3	0,0617	-0,0052	19	0,0962	0,1119
4	0,1114	0,1373	20	0,0734	0,1835
5	0,1588	0,1792	21	0,0121	0,122
6	0,1438	0,1206	22	-0,0103	0,159
7	0,1454	0,1256	23	0,0842	0,2587
8	0,0933	0,098	24	0,1141	0,2349
9	0,1311	0,1096	25	0,161	0,2927
10	0,1712	0,1558	26	0,0153	0,1606
11	0,2153	0,196	27	-0,0275	0,1434
12	0,1842	0,2088	28	-0,0745	0,0355
13	0,0947	0,1991	29	-0,0169	0,045
14	0,1138	0,0767	30	-0,0153	0,01
15	0,1709	0,2938			

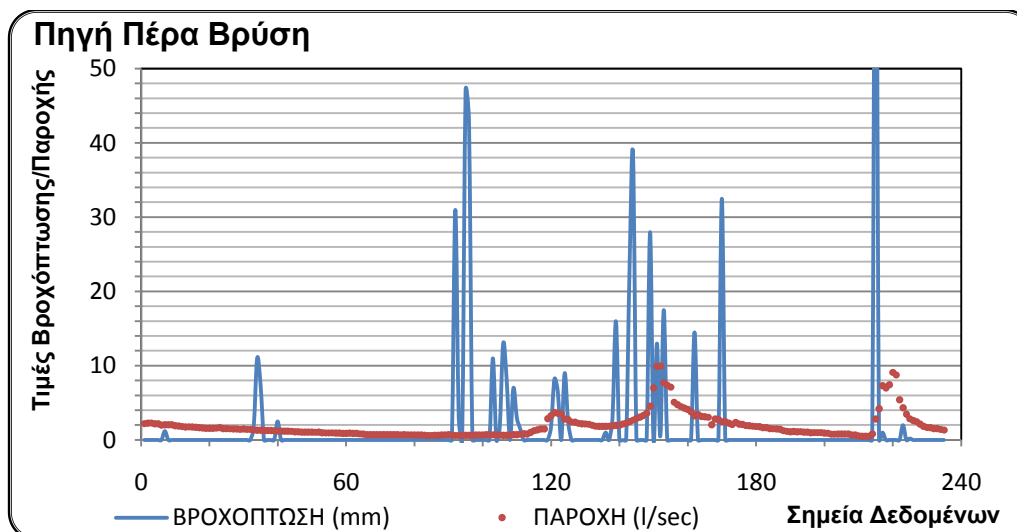
Πίνακας 7.4: Τιμές συντελεστή συσχέτισης για χρονική υστέρηση ημερών 0 έως 30

Από τον παραπάνω πίνακα παρατηρούμε ότι στην περίπτωση της πηγής Μάη Βρύση, ο συντελεστής συσχέτισης μεγιστοποιείται την 11^η μέρα, ενώ για τη δεύτερη πηγή Πέρα Βρύση, προκύπτει μία χρονική υστέρηση 15 ημερών.

Και σε αυτήν την περίπτωση παρατίθενται τα γραφήματα των χρονοσειρών της βροχόπτωσης και της παροχής προκειμένου να επαληθευτεί η συμπεριφορά τους.



Σχήμα 7.19: Χρονοσειρές βροχόπτωσης και παροχής για την πηγή Μάη Βρύση



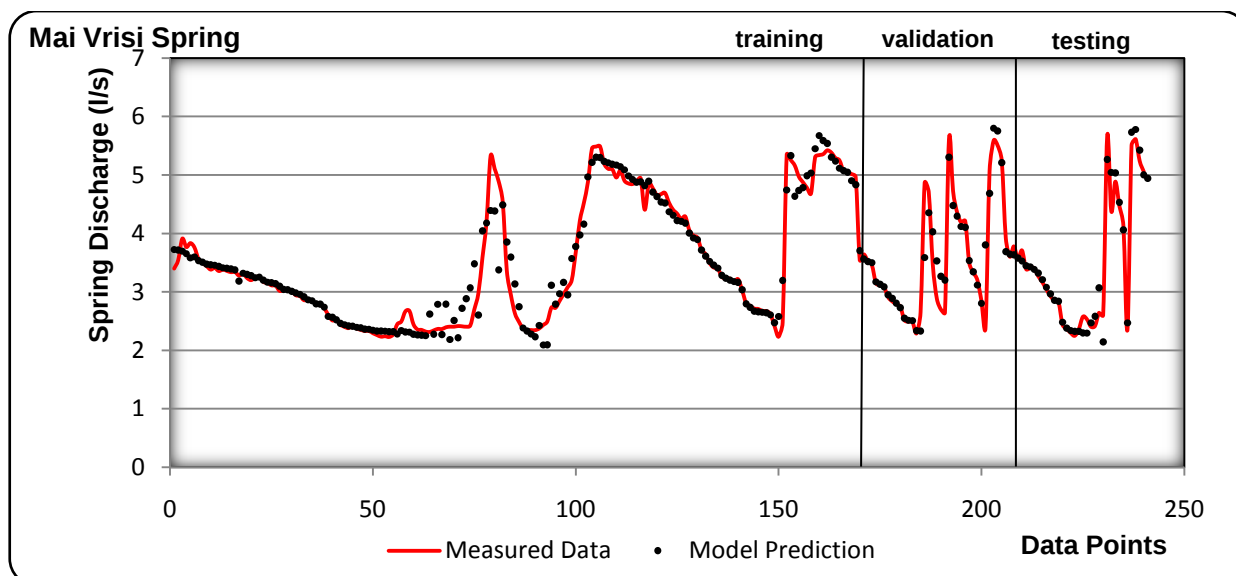
Σχήμα 7.20: Χρονοσειρές βροχόπτωσης και παροχής για την πηγή Πέρα Βρύση

Σε αυτό το σημείο πρέπει να σημειωθεί ότι λόγω της μη εφαρμογής της γραμμικής παρεμβολής, το πλήθος των δεδομένων των χρονοσειρών της βροχόπτωσης και της παροχής ήταν κατά πολύ μικρότερο και ενδεχομένως λιγότερο αντιπροσωπευτικό, γεγονός που προφανώς επηρέασε τα αποτελέσματα του συντελεστή συσχέτισης.

7.5.1 Αποτελέσματα εκτέλεσης νευρωνικού δικτύου, χωρίς την εφαρμογή γραμμικής παρεμβολής στα δεδομένα πηγής Μάη Βρύσης.

Για την συγκεκριμένη εφαρμογή χρησιμοποιήθηκε όμοιος τύπος δικτύου με τον προηγούμενο, καθώς αποδείχτηκε βέλτιστος. Επομένως πρόκειται ξανά για νευρωνικό δίκτυο τριών επιπέδων, με αριθμό νευρώνων ενδιάμεσου στρώματος ίσο με τέσσερα. Ο αλγόριθμος που εκτελέστηκε ήταν ο ίδιος με πριν, έχοντας ως συνάρτηση δημιουργίας δικτύου την `newcf`, συνάρτηση εκπαίδευσης την `trainlm` και συνάρτηση εκμάθησης την `learnqdm`. Όπως επισημάνθηκε προηγουμένως η μοναδική διαφορά με την προηγούμενη εφαρμογή είναι η χρήση της μεθόδου της γραμμικής παρεμβολής στα δεδομένα.

Επομένως για την πηγή Μάη Βρύσης το νευρωνικό δίκτυο έδωσε τα εξής αποτελέσματα για τη φάση της εκπαίδευσης (training), της αξιολόγησης (validation) και του ελέγχου (testing):



Σχήμα 7.21: Αποτελέσματα εκπαίδευσης, αξιολόγησης και ελέγχου του νευρωνικού δικτύου χωρίς γραμμική παρεμβολή

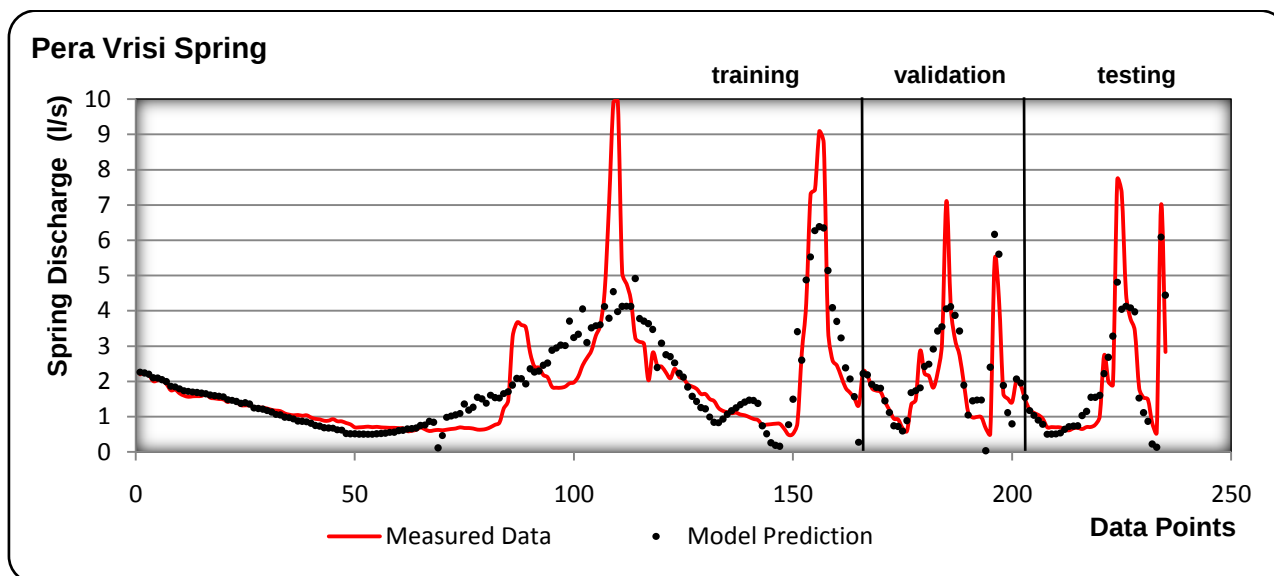
Το μέσο τετραγωνικό σφάλμα κατά τον έλεγχο ήταν ίσο με $mse=0.0034 \text{ (l/s)}^2$, δηλαδή ελάχιστα μεγαλύτερο από το μέσο τετραγωνικό σφάλμα που προέκυψε από τα αποτελέσματα του νευρωνικού δικτύου με χρήση της γραμμικής παρεμβολής, το οποίο ήταν ίσο με $mse=0.0030 \text{ (l/s)}^2$. Παρόλα αυτά μέσω της αποτύπωσης των αποτελεσμάτων σε γράφημα, παρατηρούμε ότι όντως η γραμμική παρεμβολή ευνόησε την διαδικασία και απέδωσε καλύτερα αποτελέσματα. Αξίζει να σημειωθεί ότι μέσω της μεθόδου της γραμμικής παρεμβολής ήταν διαθέσιμα 750 ζεύγη στοιχείων παροχής και βροχόπτωσης, ενώ χωρίς την εφαρμογή της γραμμικής παρεμβολής εισήχθησαν στο νευρωνικό δίκτυο μόλις 241 ζεύγη. Η αύξηση του πλήθους των στοιχείων ουσιαστικά συνέβαλε στην επίτευξη καλύτερης εκπαίδευσης του νευρωνικού δικτύου για την πηγή Μάη Βρύση.

Επομένως καταλήγουμε στο ότι η χρήση της γραμμικής παρεμβολής στα διαθέσιμα στοιχεία της πηγής Μάη Βρύση, ήταν όντως ευνοϊκή αποδίδοντας καλύτερα αποτελέσματα, τα οποία ακολουθούσαν σωστά την τάση των δεδομένων ελέγχου.

7.5.2 Αποτελέσματα εκτέλεσης νευρωνικού δικτύου, χωρίς την εφαρμογή γραμμικής παρεμβολής στα δεδομένα πηγής Πέρα Βρύσης.

Η ίδια διαδικασία ακολουθήθηκε και για τη δεύτερη πηγή Πέρα Βρύσης. Ο αλγόριθμος που εκτελέστηκε ήταν ο ίδιος με την προηγούμενη εφαρμογή της συγκεκριμένης πηγής, έχοντας ως συνάρτηση δημιουργίας δικτύου την `newcf`, συνάρτηση εκπαίδευσης την `trainlm` και συνάρτηση εκμάθησης την `learnngdm`.

Άρα για την πηγή Πέρα Βρύσης το νευρωνικό δίκτυο έδωσε τα εξής αποτελέσματα για τη φάση της εκπαίδευσης (training), της αξιολόγησης (validation) και του ελέγχου (testing), με δεδομένα χωρίς την εφαρμογή της μεθόδου της γραμμικής παρεμβολής:



Σχήμα 7.22: Αποτελέσματα εκπαίδευσης, αξιολόγησης και ελέγχου του νευρωνικού δικτύου χωρίς γραμμική παρεμβολή

Και σε αυτήν την περίπτωση το μέσο τετραγωνικό σφάλμα κατά τον έλεγχο ήταν μεγαλύτερο από το μέσο τετραγωνικό σφάλμα που προέκυψε από τα αποτελέσματα του νευρωνικού δικτύου με χρήση της γραμμικής παρεμβολής. Στην αρχική εφαρμογή προέκυψε $mse=0.0001 \text{ (l/s)}^2$, ενώ χωρίς την χρήση της γραμμικής παρεμβολής αποδόθηκε $mse=0.0094 \text{ (l/s)}^2$. Παρατηρούμε ότι όντως η γραμμική παρεμβολή ευνόησε κατά πολύ την διαδικασία και απέδωσε σαφώς καλύτερα αποτελέσματα. Σημαντικό είναι το ότι η αύξηση του πλήθους των στοιχείων, μέσω της γραμμικής παρεμβολής, ουσιαστικά συνέβαλε στην επίτευξη καλύτερης εκπαίδευσης του νευρωνικού δικτύου για την πηγή Πέρα Βρύση, καθώς τα ζεύγη στοιχείων παροχής και βροχόπτωσης από 235 αυξήθηκαν σε 657 ζεύγη.

Επομένως καταλήγουμε στο ότι η χρήση της γραμμικής παρεμβολής στα διαθέσιμα στοιχεία, απέδωσε καλύτερα αποτελέσματα, ακολουθώντας σωστά την τάση των δεδομένων ελέγχου.

7.6 Βελτιωμένο νευρωνικό δίκτυο - Εισαγωγή στοιχείων θερμοκρασίας - Αναλυτική παρουσίαση αποτελεσμάτων

Στη συνέχεια έγινε μία προσπάθεια βελτίωσης της απόκρισης του νευρωνικού δικτύου θέτοντας ως παραμέτρους εισόδου όχι μόνο τα ύψη βροχόπτωσης και τον αριθμό ημέρας, αλλά και στοιχεία θερμοκρασίας. Αξίζει να σημειωθεί ότι στην κοινότητα Γέργερη του δήμου Ρούβα, άρχισε να λειτουργεί από τον Απρίλιο του 2009, ένας νέος μετεωρολογικός σταθμός. Συνεπώς δεν ήταν

διαθέσιμα τα στοιχεία θερμοκρασίας για τα έτη 2007, 2008 και 2009 (έως 14/07), για τα οποία πραγματοποιείται η παρούσα εργασία. Γι αυτόν τον λόγο συλλέχθηκαν δεδομένα θερμοκρασίας του δήμου Ηρακλείου και κατόπιν πραγματοποιήθηκε σύγκριση των θερμοκρασιών των κοινών ημερών των δύο περιοχών (Γέργερης και Ηρακλείου) διαπιστώνοντας εν τέλει ότι η απόκλιση είναι ελάχιστη. Σύμφωνα με τα παραπάνω η χρήση των θερμοκρασιών του δήμου Ηρακλείου δεν πρόκειται να εισάγει κάποιο σφάλμα κι έτσι προστίθεται ένα ακόμη διάνυσμα στα δεδομένα εισόδου. Η παραπάνω διαδικασία ακολουθήθηκε και για τις δύο πηγές Μάη και Πέρα Βρύση του Δήμου Ρούβα.

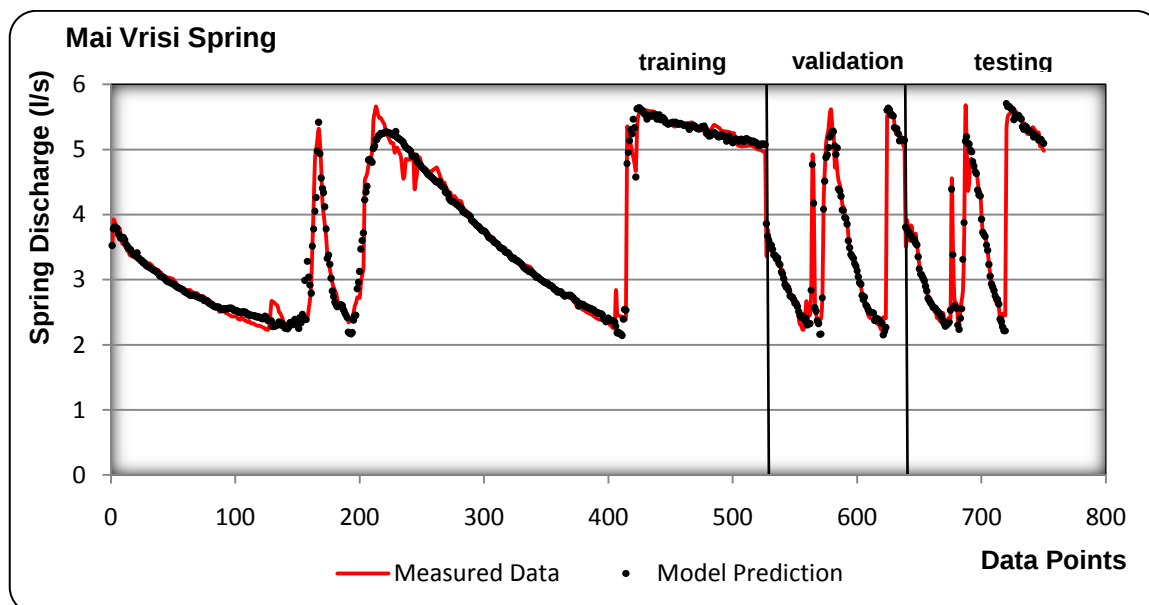


Σχήμα 7.23: Μετεωρολογικός σταθμός Γέργερης Δήμου Ρούβα

7.6.1 Αποτελέσματα βελτιωμένου νευρωνικού δικτύου για την πηγή Μάη Βρύσης.

Για την εκτέλεση του βελτιωμένου νευρωνικού δικτύου, ενισχυμένου με στοιχεία θερμοκρασίας, χρησιμοποιήθηκε νευρωνικό δίκτυο τριών επιπέδων, με αριθμό νευρώνων ενδιάμεσου στρώματος ίσο με τέσσερα. Ο αλγόριθμος είχε ως συνάρτηση δημιουργίας δικτύου την *newcf*, συνάρτηση εκπαίδευσης την *trainlm* και συνάρτηση εκμάθησης την *learnqdm*, καθώς ο συνδυασμός αυτός αποδείχτηκε βέλτιστος. Σε αυτήν την εφαρμογή το επίπεδο εισόδου κατείχε τρεις νευρώνες, καθώς είχαμε τρία διανύσματα εισόδου: το ύψος βροχόπτωσης, τον αριθμό ημέρας και τη θερμοκρασία. Αξίζει να σημειωθεί ότι χρησιμοποιήθηκαν τα δεδομένα που προέκυψαν μετά από την χρήση της γραμμικής παρεμβολής άρα και η χρονική υστέρηση, όπως υπολογίστηκε για την περίπτωση αυτή. Επομένως πρόκειται για ένα νευρωνικό δίκτυο όμοιο με την πρώτη εφαρμογή που παρουσιάστηκε, με μοναδική διαφορά την πρόσθεση ενός διανύσματος εισόδου, αυτού της θερμοκρασίας.

Επομένως για την πηγή Μάη Βρύσης το βελτιωμένο νευρωνικό δίκτυο έδωσε τα εξής αποτελέσματα για τη φάση της εκπαίδευσης (*training*), της αξιολόγησης (*validation*) και του ελέγχου (*testing*) με τιμή κανονικοποιημένου μέσου τετραγωνικού σφάλματος (κατά τον έλεγχο) ίση με $mse=0.0018$ (l/s)²:



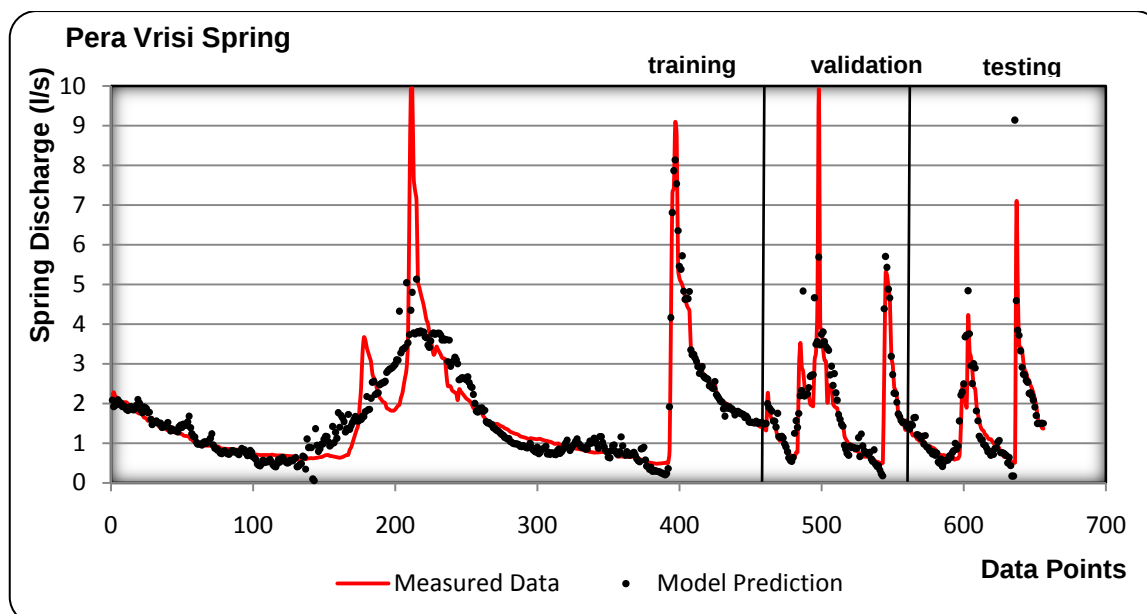
Σχήμα 7.24: Αποτελέσματα εκπαίδευσης, αξιολόγησης και ελέγχου του βελτιωμένου νευρωνικού δικτύου

Παρατηρούμε λοιπόν ότι η εισαγωγή των στοιχείων της θερμοκρασίας ως παράμετρο εισόδου στο δίκτυο, ουσιαστικά βελτίωσε την απόκριση του νευρωνικού δικτύου, καθώς το μέσο τετραγωνικό σφάλμα μειώθηκε από 0.0030 (l/s)^2 σε 0.0018 (l/s)^2 και όπως φαίνεται από το γράφημα το δίκτυο πραγματοποιεί μία πολύ καλή εκτίμηση.

7.6.2 Αποτελέσματα βελτιωμένου νευρωνικού δικτύου για την πηγή Πέρα Βρύσης.

Η ίδια διαδικασία με όμοιο τύπο εκπαίδευσης και δομή νευρωνικού δικτύου ακολουθήθηκε και για τη δεύτερη πηγή Πέρα Βρύσης. Ο αλγόριθμος λοιπόν εκτελέστηκε με συνάρτηση δημιουργίας δικτύου την `newcf`, συνάρτηση εκπαίδευσης την `trainlm` και συνάρτηση εκμάθησης την `learnqdm`. Επομένως χρησιμοποιήθηκε νευρωνικό δίκτυο όμοιο με την πρώτη εφαρμογή, με μοναδική διαφορά την πρόσθεση του διανύσματος εισόδου της θερμοκρασίας.

Για την πηγή Πέρα Βρύσης το νευρωνικό δίκτυο έδωσε τα εξής αποτελέσματα για τη φάση της εκπαίδευσης (training), της αξιολόγησης (validation) και του ελέγχου (testing), με τιμή κανονικοποιημένου μέσου τετραγωνικού σφάλματος (κατά τον έλεγχο) ίση με $mse=0.0017 \text{ (l/s)}^2$:



Σχήμα 7.25: Αποτελέσματα εκπαίδευσης, αξιολόγησης και ελέγχου του βελτιωμένου νευρωνικού δικτύου

Σύμφωνα με το γράφημα παρατηρούμε ότι και σε αυτήν την περίπτωση το νευρωνικό δίκτυο πραγματοποίησε μία εξίσου καλή εκτίμηση κατά τον έλεγχο.

7.7 Εισαγωγή των βροχοπτώσεων των ημερών όπου μεγιστοποιείται ο συντελεστής συσχέτισης ως δεδομένων εισόδου - Αναλυτική παρουσίαση αποτελεσμάτων

Όπως περιγράφηκε σε αρχική παράγραφο του κεφαλαίου, αρχικός στόχος κατά την επεξεργασία των δεδομένων ήταν η εύρεση του χρόνου υστέρησης της εμφάνισης κάθε επεισοδίου βροχόπτωσης στην παροχή της κάθε πηγής. Αξιοσημείωτο είναι το γεγονός ότι στην πηγή Πέρα Βρύση, ο συντελεστής συσχέτισης μεγιστοποιείται μεταξύ τριών διαδοχικών ημερών, με πολύ μικρές αποκλίσεις μεταξύ τους. Στις προηγούμενες εφαρμογές επιλέχθηκε ο μέγιστος συντελεστής συσχέτισης που αντιστοιχούσε στην 16^η μέρα με τιμή 0.2371 αγνοώντας την πολύ μικρή διαφορά που κατείχαν οι αντίστοιχοι συντελεστές της 15^{ης} και 17^{ης} ημέρας με τιμές 0.2285 και 0.2202 αντιστοίχως, όπως φαίνεται και στον πίνακα που ακολουθεί.

TIME LAG (μέρες)	Συντελεστής Συσχέτισης Πηγή Πέρα Βρύση	TIME LAG (μέρες)	Συντελεστής Συσχέτισης Πηγή Πέρα Βρύση
0	0,0295	16	0,2371
1	0,02	17	0,2202
2	0,0364	18	0,1245
3	0,0174	19	0,1008
4	0,098	20	0,1559
5	0,1684	21	0,1396
6	0,135	22	0,1722
7	0,117	23	0,1833
8	0,1316	24	0,1605
9	0,098	25	0,1869
10	0,1236	26	0,1782
11	0,122	27	0,1381
12	0,1321	28	0,0839
13	0,1602	29	0,1229
14	0,1534	30	0,149
15	0,2285		

Πίνακας 7.5: Τιμές συντελεστή συσχέτισης για χρονική υστέρηση ημερών 0 έως 30

Τα διανύσματα εισόδου στο δίκτυο ήταν τα εξής:

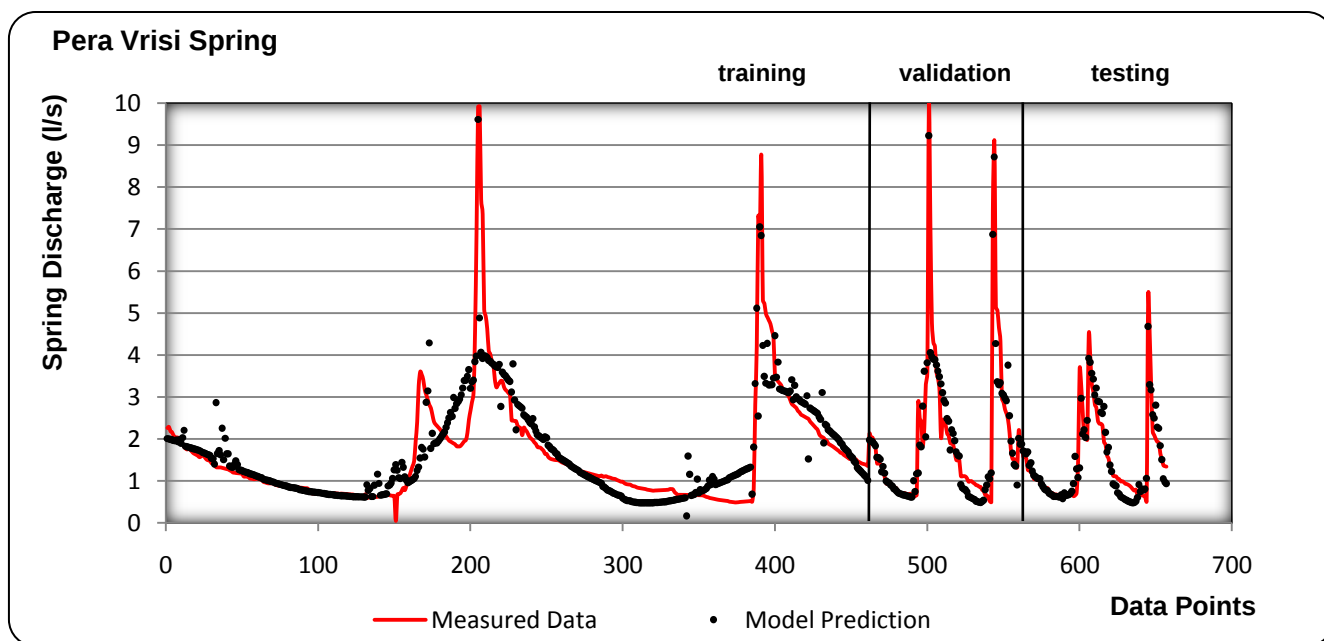
1. Ύψη βροχόπτωσης 15 ημερών χρονικής υστέρησης,
2. ύψη βροχόπτωσης 16 ημερών χρονικής υστέρησης,
3. ύψη βροχόπτωσης 17 ημερών χρονικής υστέρησης και
4. αριθμός ημέρας

Η συγκεκριμένη εφαρμογή απαιτούσε ειδική προεπεξεργασία των δεδομένων προτού εισαχθούν στο νευρωνικό δίκτυο. Αφού διαπιστώθηκε ότι ο συντελεστής συσχέτισης μεγιστοποιείται μεταξύ 15-17 ημερών, δημιουργήθηκαν σε φύλλο excel τρεις όμοιες διπλές στήλες με τα ζεύγη βροχόπτωσης – παροχών, με την ένδειξη Nan όπου δεν υπήρχε διαθέσιμη τιμή παροχής. Η κάθε μία από τις στήλες των βροχοπτώσεων μετατοπίστηκε προς τα κάτω κατά 15, 16 και 17 θέσεις αντίστοιχα. Κατόπιν σε οποιοδήποτε κελί παροχής υπήρχε η ένδειξη Nan, διαγραφόταν αυτόματα ολόκληρη η σειρά, αγνοώντας τελείως και τα τρία ζεύγη βροχοπτώσεων – παροχών. Το αποτέλεσμα όλων των παραπάνω ήταν η δημιουργία τριών ζευγών βροχόπτωσης – απορροής, τα οποία εισήχθησαν ως δεδομένα στο νευρωνικό δίκτυο. Αξιοσημείωτο είναι το ότι και τα τρία ζεύγη βροχοπτώσεων – παροχών, που αντιστοιχούν σε διαφορετικούς χρόνους υστέρησης, κατέχουν ακριβώς τον ίδιο αριθμό στοιχείων παροχής, καθώς η στήλη με τις παροχές δε μετατοπίστηκε καθόλου. Δηλαδή τα στοιχεία εισόδου και επιθυμητής εξόδου αντιστοιχούν ουσιαστικά στους εξής πίνακες:

$$\text{Είσοδος: } \begin{matrix} \overbrace{1 \dots n} \\ \boxed{\begin{matrix} \text{Βροχόπτωση 15} \\ \text{Βροχόπτωση 16} \\ \text{Βροχόπτωση 17} \\ \text{Αριθμός ημέρας} \end{matrix}} \end{matrix} \quad (7.8)$$

$$\text{Επιθυμητή έξοδος: } \begin{matrix} \overbrace{1 \dots n} \\ \boxed{\text{Παροχή}} \end{matrix} \quad (7.9)$$

Αφού λοιπόν ορίστηκαν τα διανύσματα εισόδου και επιθυμητής εξόδου στο νευρωνικό δίκτυο, χρησιμοποιήθηκε ο ίδιος τύπος αλγορίθμου που εφαρμόστηκε για την πηγή Πέρα Βρύση στις προηγούμενες εφαρμογές. Πρόκειται λοιπόν για νευρωνικό δίκτυο με συνάρτηση δημιουργίας δικτύου την `newcf`, συνάρτηση εκπαίδευσης την `trainlm` και συνάρτηση εκμάθησης την `learnqdm`. Στη συνέχεια παρουσιάζονται τα αποτελέσματα που έδωσε το δίκτυο για τη φάση της εκπαίδευσης (training), της αξιολόγησης (validation) και του ελέγχου (testing). Αξίζει να σημειωθεί ότι η τιμή του κανονικοποιημένου μέσου τετραγωνικού σφάλματος (κατά τον έλεγχο) βρέθηκε ίση με $mse=0.0026$ (l/s)².



Σχήμα 7.26: Αποτελέσματα εκπαίδευσης, αξιολόγησης και ελέγχου του νευρωνικού δικτύου

Επομένως παρατηρούμε ότι η συγκεκριμένη εφαρμογή δεν έδωσε ευνοϊκότερα αποτελέσματα, καθώς το μέσο τετραγωνικό σφάλμα αυξήθηκε από την τιμή 0.0001 (l/s)² (κατά την οποία είχε επιλεγεί μονάχα ο μέγιστος συντελεστής συσχέτισης στην επεξεργασία των δεδομένων) σε 0.0026 (l/s)². Το γεγονός ότι το νευρωνικό δίκτυο δεν πραγματοποίησε μία καλύτερη εκτίμηση φαίνεται και μέσω του παραπάνω γραφήματος, στο οποίο οι παραχθείσες τιμές δεν ακολουθούν απόλυτα την

τάση των δεδομένων ελέγχου, όπως στην αρχική περίπτωση επιλογής του μέγιστου συντελεστή συσχέτισης. Μπορεί βέβαια σε σύγκριση με τα προηγούμενα αποτελέσματα η απόκριση του νευρωνικού δικτύου να μη θεωρείται ευνοϊκότερη, αλλά παρόλα αυτά το νευρωνικό δίκτυο φαίνεται να πραγματοποιήσει μια σχετικά ορθή εκτίμηση.

7.8 Δημιουργία νευρωνικού δικτύου με παραμέτρους εξόδου τη διαφορά της παροχής ανά χρονικό βήμα ημέρας - Αναλυτική παρουσίαση αποτελεσμάτων

Στα πλαίσια της παρούσας εργασίας επιχειρήθηκε μία ακόμη δοκιμή εκτέλεσης του νευρωνικού δικτύου, θέτοντας διαφορετικές παραμέτρους εισόδου και δεδομένα εξόδου. Πιο συγκεκριμένα, τέθηκαν ως διανύσματα εισόδου τα ύψη βροχόπτωσης και η παροχή της προηγούμενης ημέρας, ενώ ως επιθυμητή έξοδος ορίστηκε η διαφορά της παροχής (σημερινής – προηγούμενης ημέρας). Η διαφορά της παροχής σαφώς κατέχει μικρότερο πεδίο τιμών, από ότι η ίδια η παροχή με αποτέλεσμα να επικεντρώνεται το νευρωνικό δίκτυο στις πολύ μικρές αλλαγές που προκύπτουν στην παροχή. Η συγκεκριμένη εφαρμογή απαιτούσε ειδική προεπεξεργασία των αρχικών διαθέσιμων δεδομένων, η οποία περιγράφεται παρακάτω.

Αρχικά δημιουργήθηκαν σε φύλλο του excel δύο στήλες με τα ύψη βροχόπτωσης και τις αντίστοιχες παροχές για τα έτη 2007, 2008 και 2009 (έως 14/07/09), με την ένδειξη “Nan”, όπου δεν υπήρχε διαθέσιμη τιμή παροχής. Κατόπιν εφαρμόστηκε η μέθοδος της γραμμικής παρεμβολής, μονάχα σε συνθήκες βασικής ροής, όπως ακριβώς περιγράφηκε και σε προηγούμενη παράγραφο. Αποτέλεσμα των παραπάνω είναι η δημιουργία ζευγών βροχόπτωσης – παροχής για τα μελετούμενα έτη, με ένδειξη “Nan”, σε οποιαδήποτε ημερομηνία δεν υπήρχαν διαθέσιμα στοιχεία παροχής (όπου δηλαδή δεν ήταν δυνατή η εφαρμογή της μεθόδου της γραμμικής παρεμβολής).

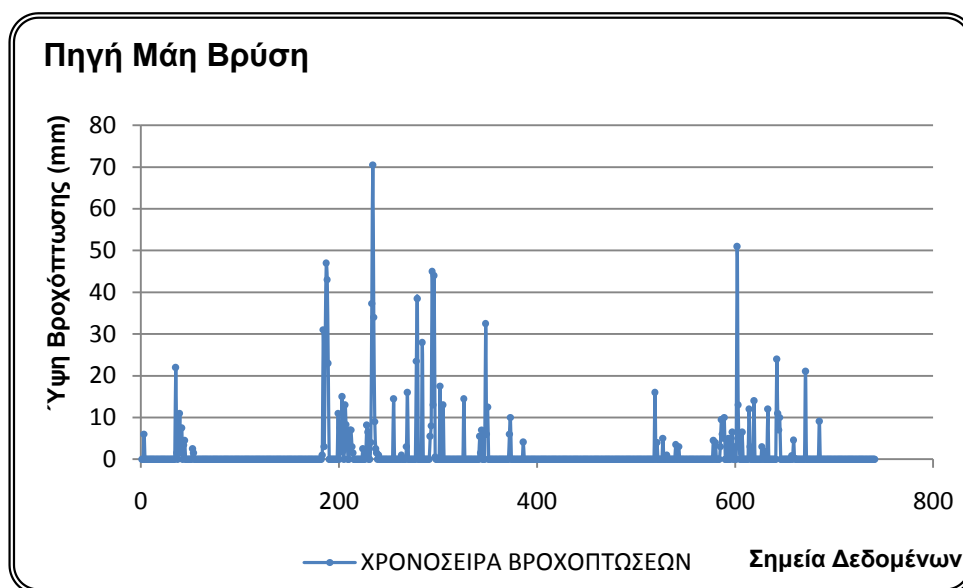
Επόμενο βήμα ήταν η δημιουργία στήλης με τη διαφορά της σημερινής με την προηγούμενη παροχή. Ουσιαστικά, για κάθε ημερομηνία για την οποία υπήρχε τιμή παροχής αφαιρούταν η παροχή που είχε σημειωθεί την προηγούμενη από αυτήν ημέρα. Στις ημερομηνίες για τις οποίες δεν υπήρχαν στοιχεία παροχών, καταγραφόταν η ένδειξη “NO DATA”. Στη συνέχεια δημιουργήθηκε μία ακόμη στήλη, η οποία περιείχε τις παροχές της προηγούμενης ημέρας κάθε ημερομηνίας για την οποία υπήρχε τιμή στην στήλη της διαφοράς των παροχών.

Κατόπιν για τα ζεύγη βροχόπτωσης και διαφοράς παροχής, που δημιουργήθηκαν, υπολογίστηκε εκ νέου ο συντελεστής συσχέτισης για χρονικό διάστημα 30 ημερών, προκειμένου να επιλεχθεί ο μέγιστος, ο οποίος θα αντιστοιχεί στην χρονική υστέρηση της εμφάνισης της βροχής στην παροχή της πηγής. Τα αποτελέσματα των τιμών του συντελεστή συσχέτισης και για τις δύο πηγές παρουσιάζονται στον πίνακα που ακολουθεί.

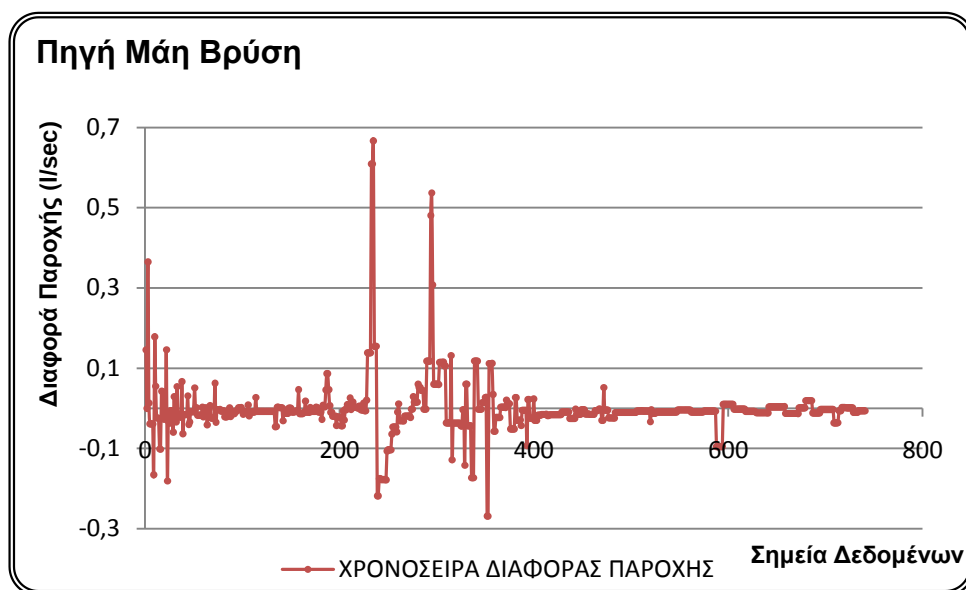
TIME LAG (μέρες)	Συντελεστής Συσχέτισης		TIME LAG (μέρες)	Συντελεστής Συσχέτισης	
	Πηγή Μάη Βρύση	Πηγή Πέρα Βρύση		Πηγή Μάη Βρύση	Πηγή Πέρα Βρύση
0	0,2432	0,1412	16	-0,0798	-0,0222
1	0,4007	0,1422	17	-0,0244	0,0382
2	0,4781	0,2274	18	-0,0028	0,017
3	0,407	0,1677	19	-0,0306	0,0794
4	0,2418	0,3172	20	-0,0655	0,0651
5	0,1003	0,2708	21	-0,0464	-0,1014
6	0,0273	0,3445	22	-0,0443	-0,0685
7	-0,0334	0,2761	23	0,0169	-0,0921
8	-0,0189	0,1965	24	0,014	-0,0343
9	-0,0062	0,0727	25	0,0353	-0,022
10	-0,0452	-0,0008	26	0,031	-0,0159
11	-0,0485	-0,0067	27	0,0523	0,0233
12	-0,0315	-0,0391	28	0,0618	0,0226
13	-0,0404	0,0003	29	0,0713	0,0307
14	-0,0461	0,0821	30	0,0413	-0,004
15	-0,0572	-0,0269			

Πίνακας 7.6: Τιμές συντελεστή συσχέτισης για χρονική υστέρηση ημερών 0 έως 30

Για την περίπτωση της πηγής Μάη Βρύση παρατίθενται τα γραφήματα των χρονοσειρών της βροχόπτωσης και της διαφοράς της παροχής. Λόγω της διαφορετικής τάξης των δεδομένων, οι χρονοσειρές παρουσιάζονται σε ξεχωριστά γραφήματα.



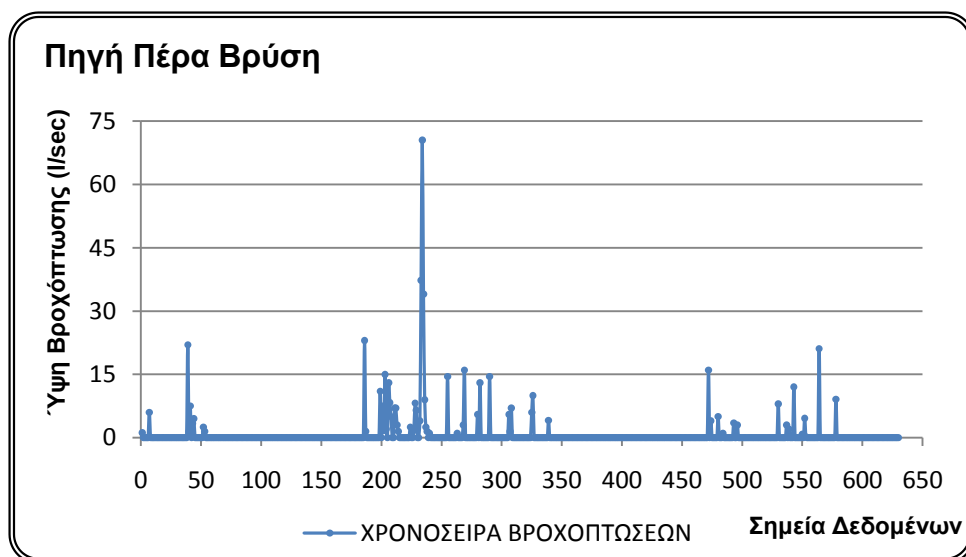
Σχήμα 7.27: Χρονοσειρά βροχόπτωσης για την πηγή Μάη Βρύση



Σχήμα 7.28: Χρονοσειρά παροχής για την πηγή Μάη Βρύση

Επομένως διαπιστώνεται όντως η ορθή επιλογή του συντελεστή συσχέτισης, καθώς οι δύο χρονοσειρές ακολουθούν την ίδια συμπεριφορά.

Ακολουθεί όμοια διαδικασία για την πηγή Πέρα Βρύση:



Σχήμα 7.29: Χρονοσειρά βροχόπτωσης για την πηγή Πέρα Βρύση



Σχήμα 7.30: Χρονοσειρά παροχής για την πηγή Πέρα Βρύση

Παρατηρούμε επίσης ότι στην συγκεκριμένη πηγή οι δύο χρονοσειρές, τείνουν στην ίδια συμπεριφορά.

Στην παρούσα παράγραφο θα μελετηθεί τεχνητό νευρωνικό δίκτυο τριών επιπέδων, με στρώμα εισόδου τεσσάρων νευρώνων, ενδιάμεσο στρώμα τεσσάρων νευρώνων και στρώμα εξόδου ενός νευρώνα.

Από τον παραπάνω πίνακα παρατηρούμε ότι στην περίπτωση της πηγής Μάη Βρύση, ο συντελεστής συσχέτισης μεγιστοποιείται την 2^η μέρα, ενώ για τη δεύτερη πηγή Πέρα Βρύση, προκύπτει μία χρονική υστέρηση 6 ημερών. Είναι προφανές ότι η συσχέτιση των χρονοσειρών της βροχόπτωσης με τη διαφορά της παροχής δίνει καλύτερα αποτελέσματα, ενώ και σε αυτήν την περίπτωση παρατηρούμε ότι στην πηγή Πέρα Βρύση, απαιτείται μεγαλύτερο χρονικό διάστημα σε σχέση με την πηγή Μάη Βρύση.

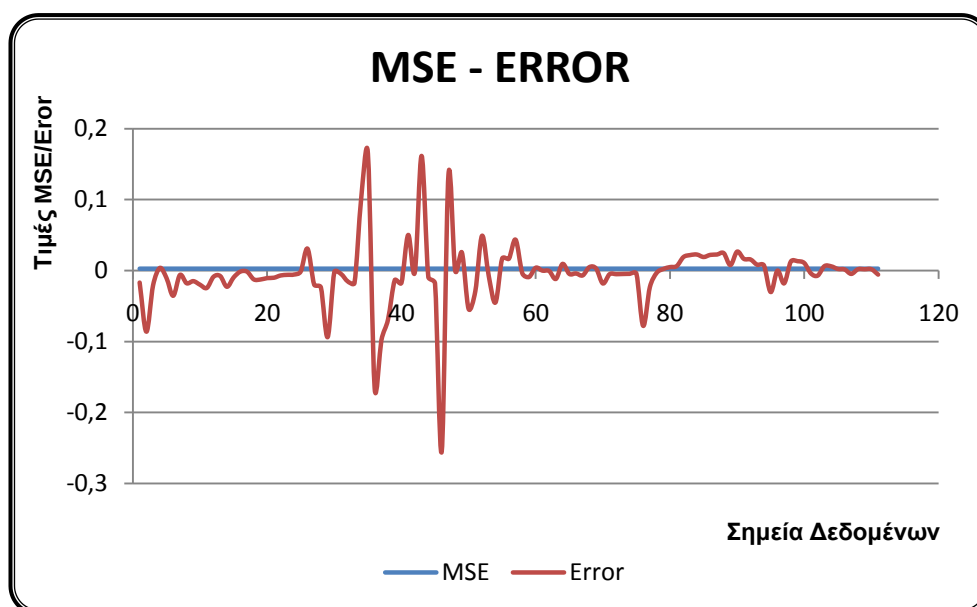
Στη συνέχεια, ακολουθήθηκε ο ίδιος τρόπος, όπως περιγράφηκε σε προηγούμενη εφαρμογή, με τον οποίο τροποποιήθηκαν οι χρονοσειρές των υψών βροχόπτωσης και δημιουργήθηκαν τα τελικά δεδομένα. Έτσι τα ύψη βροχόπτωσης και η παροχή προηγούμενης ημέρας αποτέλεσαν τα διανύσματα εισόδου στο νευρωνικό δίκτυο, ενώ διαφορά παροχής ήταν η επιθυμητή απόκριση του δικτύου.

Η διαδικασία που περιγράφηκε ακολουθήθηκε ομοίως για τα στοιχεία και των δύο πηγών Μάη και Πέρα Βρύση και τα αποτελέσματα που έδωσε το νευρωνικό δίκτυο παρουσιάζονται ακολούθως.

7.8.1 Αποτελέσματα νευρωνικού δικτύου με παράμετρο εξόδου τη διαφορά παροχής για την πηγή Μάη Βρύσης.

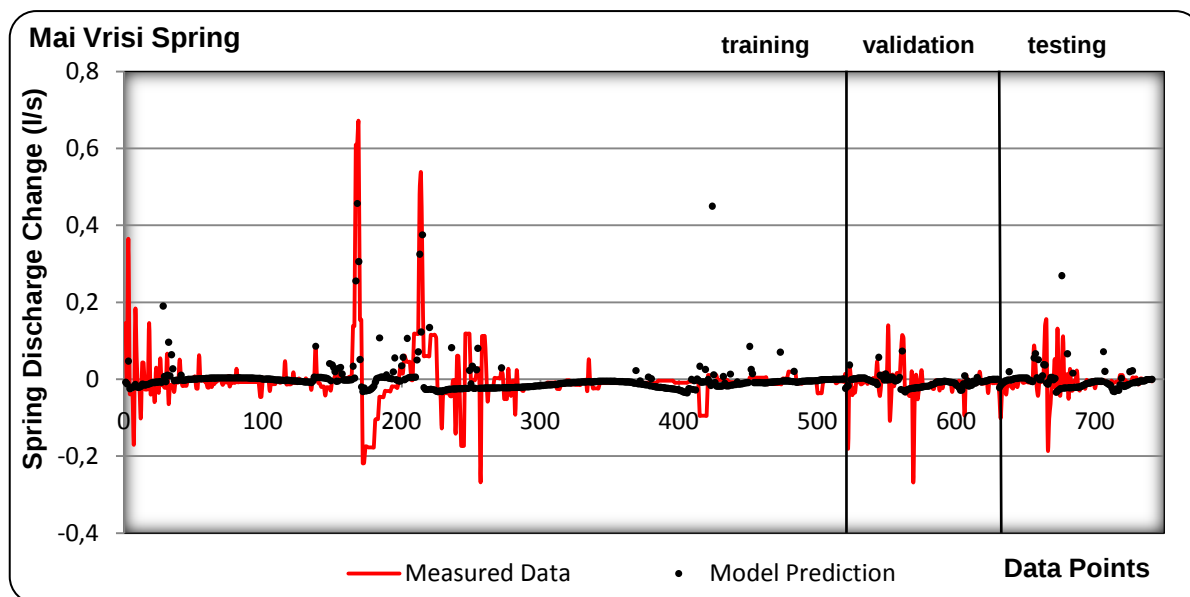
Για την εκτέλεση της συγκεκριμένης εφαρμογής, χρησιμοποιήθηκε νευρωνικό δίκτυο τριών επιπέδων, με αριθμό νευρώνων ενδιάμεσου στρώματος ίσο με τέσσερα. Ο αλγόριθμος έχει ως συνάρτηση δημιουργίας δικτύου την `newcf`, συνάρτηση εκπαίδευσης την `trainlm` και συνάρτηση εκμάθησης την `learnqdm`, καθώς ο συνδυασμός αυτός αποδείχτηκε βέλτιστος. Το επίπεδο εισόδου κατέχει τρεις νευρώνες, καθώς διαθέτει τρία διανύσματα εισόδου: το ύψος βροχόπτωσης, την παροχή της προηγούμενης ημέρας και τον αριθμό ημέρας. Τέλος το επίπεδο εξόδου αποτελείται από ένα νευρώνα, τη διαφορά της παροχής.

Επομένως για την πηγή Μάη Βρύσης το νευρωνικό δίκτυο έδωσε τα εξής αποτελέσματα για τη φάση της εκπαίδευσης (training), της αξιολόγησης (validation) και του ελέγχου (testing) με τιμή κανονικοποιημένου μέσου τετραγωνικού σφάλματος (κατά τον έλεγχο) ίση με $mse=0.0025 (l/s)^2$:



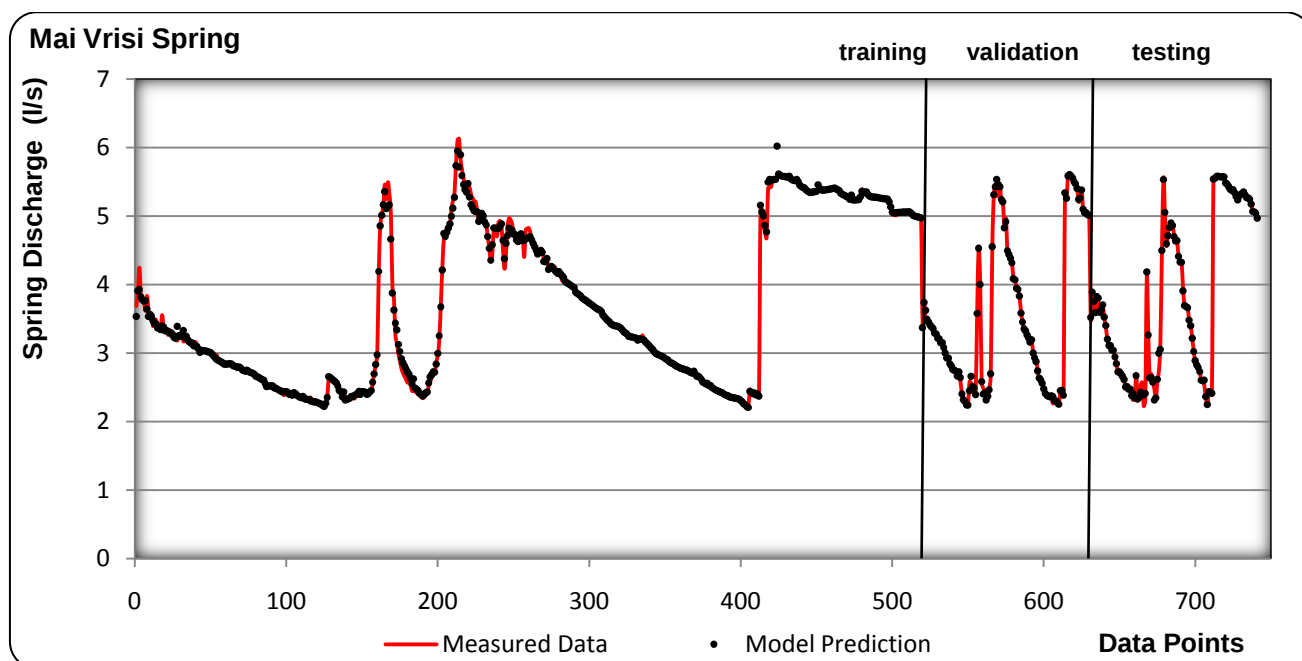
Σχήμα 7.31: Μέσο τετραγωνικό σφάλμα και σφάλμα αποτελεσμάτων ανάμεσα στις αναμενόμενες και παραχθείσες τιμές του νευρωνικού δικτύου.

Στο παραπάνω γράφημα αποτυπώνεται η «κίνηση» του σφάλματος πάνω στην τιμή του μέσου τετραγωνικού σφάλματος των αποτελεσμάτων του δικτύου.



Σχήμα 7.32: Αποτελέσματα εκπαίδευσης, αξιολόγησης και ελέγχου του νευρωνικού δικτύου

Η συγκεκριμένη εφαρμογή θεωρείται αρκετά δύσκολη ως προς την εκτέλεση, καθώς οι μεταβολές ΔQ είναι πάρα πολύ μικρές. Ωστόσο, όταν τα αποτελέσματα της διαφοράς της παροχής μετατραπούν σε τιμές παροχής, προσθέτοντας την τιμή της παροχής της προηγούμενης ημέρας, παρατηρείται μία σημαντική βελτίωση συγκριτικά με την προηγούμενη προσπάθεια.



Σχήμα 7.33: Αποτελέσματα παροχής κατά τη φάση της εκπαίδευσης, αξιολόγησης και ελέγχου του νευρωνικού δικτύου

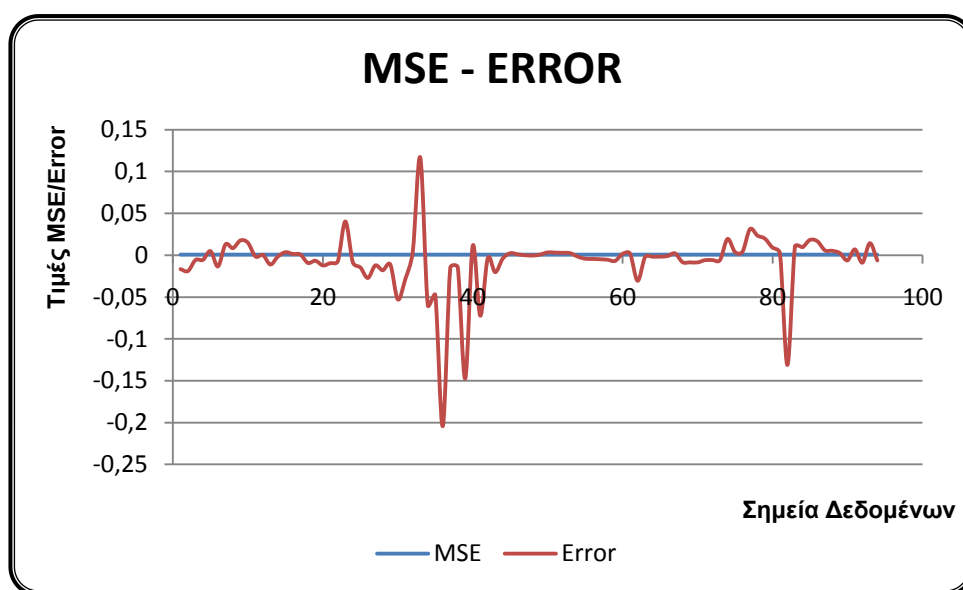
Το παραπάνω γράφημα της παροχής προέκυψε μέσω της πρόσθεσης των αποτελεσμάτων :

$$Q_{\text{προηγ.μέρας}} + \Delta Q = Q_{\text{προηγ.μέρας}} + (Q_{\text{σημερινή}} - Q_{\text{προηγ.μέρας}}) = Q_{\text{σημερινή}}$$

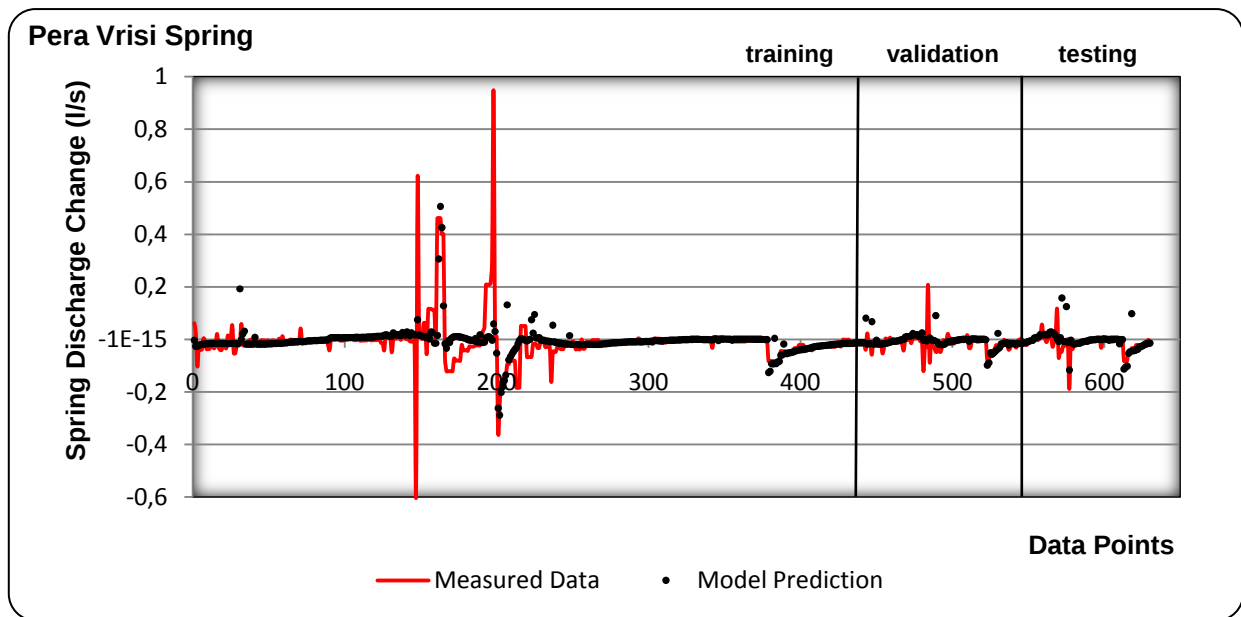
Παρατηρούμε ότι το νευρωνικό δίκτυο πραγματοποίησε μία σχετικά καλή εκτίμηση της παροχής, καθώς οι παραχθείσες τιμές ακολουθούν την τάση των παρατηρούμενων, ειδικά κατά τη φάση του ελέγχου του δικτύου.

7.8.2 Αποτελέσματα νευρωνικού δικτύου με παράμετρο εξόδου τη διαφορά παροχής για την πηγή Πέρα Βρύσης.

Η ίδια διαδικασία με όμοιο τύπο εκπαίδευσης και δομή νευρωνικού δικτύου ακολουθήθηκε και για τη δεύτερη πηγή Πέρα Βρύσης. Τα αποτελέσματα που δόθηκαν για τη φάση της εκπαίδευσης (training), της αξιολόγησης (validation) και του ελέγχου (testing) παρουσιάζονται παρακάτω. Αξίζει να σημειωθεί ότι η τιμή του κανονικοποιημένου μέσου τετραγωνικού σφάλματος (κατά τον έλεγχο) ήταν ίση με $\text{mse} = 0.0005 \text{ (l/s)}^2$:

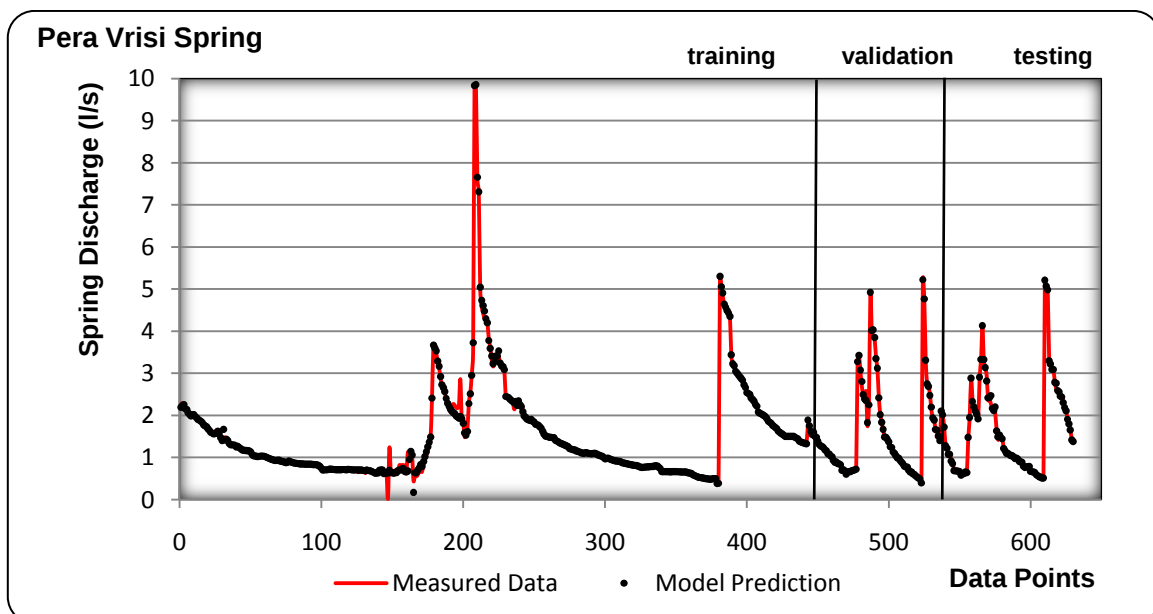


Σχήμα 7.34: Μέσο τετραγωνικό σφάλμα και σφάλμα αποτελεσμάτων ανάμεσα στις αναμενόμενες και παραχθείσες τιμές του νευρωνικού δικτύου.



Σχήμα 7.35: Αποτελέσματα εκπαίδευσης, αξιολόγησης και ελέγχου του νευρωνικού δικτύου

Στην περίπτωση εφαρμογής για την πηγή Πέρα Βρύση, σύμφωνα με το παραπάνω γράφημα διαπιστώνεται ότι το νευρωνικό δίκτυο, πραγματοποίησε μια σχετικά καλή εκτίμηση της απόκρισης. Τα αποτελέσματα τείνουν να ακολουθήσουν τα δεδομένα ελέγχου εκτός από μερικά σημεία όπου έχει πραγματοποιηθεί υπερεκτίμηση από το δίκτυο. Ωστόσο, όπως προαναφέρθηκε, η συγκεκριμένη εφαρμογή θεωρείται αρκετά δύσκολη και γι αυτόν τον λόγο η απόδοση του δικτύου κρίνεται από τα αποτελέσματα που προκύπτουν για την παροχή του νευρωνικού δικτύου, τα οποία παρουσιάζονται ακολούθως.



Σχήμα 7.36: Αποτελέσματα παροχής κατά τη φάση της εκπαίδευσης, αξιολόγησης και ελέγχου του νευρωνικού δικτύου

Το παραπάνω γράφημα της παροχής, όπως προαναφέρθηκε, προέκυψε μέσω της σχέσης:

$$Q_{\text{προηγ.μέρας}} + \Delta Q = Q_{\text{προηγ.μέρας}} + (Q_{\text{σημερινή}} - Q_{\text{προηγ.μέρας}}) = Q_{\text{σημερινή}}$$

Επομένως, και στην πηγή Πέρα Βρύση, οι παραχθείσες τιμές του νευρωνικού δικτύου ακολουθούν την τάση των παρατηρούμενων, με αποτέλεσμα να θεωρείται ικανοποιητική η απόδοση του δικτύου.

Κεφάλαιο 8

Συμπεράσματα και Συζήτηση

Η ιδιομορφία των καρστικών εδαφών όσον αφορά τη χωρική ποικιλομορφία των υδραυλικών και γεωλογικών τους χαρακτηριστικών, έχει ως συνέπεια την αντιμετώπιση σοβαρής δυσκολίας προσομοίωσης τους με κοινά αποδεκτούς τρόπους. Το γεγονός αυτό οφείλεται στο ιδιαίτερο ανάγλυφο τους και τα ασυνήθιστα χαρακτηριστικά τους, καθώς και στις βασικές διαφορές που παρουσιάζουν σε σχέση με τα συνήθη εμφανιζόμενα κοκκώδη εδάφη. Σε οποιαδήποτε προσπάθεια προσομοίωσης ενός καρστικού υδροφορέα πρέπει να ληφθούν υπόψη πολλοί και διάφοροι παράγοντες, όπως το πλήθος και το εύρος των υπόγειων καναλιών που συγκροτούν τα υπόγεια αποχετευτικά δίκτυα, τα υδρολογικά δεδομένα, το ανάγλυφο, η δομή του εδάφους και άλλοι παράγοντες, οι οποίοι πολλές φορές δε γίνονται εύκολα αντιληπτοί. Επομένως η αβεβαιότητα που προκύπτει στους καρστικούς υδροφορείς για τα διάφορα επιμέρους στοιχεία τους, οδηγεί στην αδυναμία ύπαρξης ενός καθολικά αποδεκτού μοντέλου. Για τον λόγο αυτό απαιτούνται αναλύσεις και κατάλληλες μετρήσεις, βασισμένες σε μακροχρόνια χρονοδιαγράμματα, προκειμένου να προκύψουν ασφαλή συμπεράσματα.

Είναι ήδη γνωστό το ότι στα μεν κοκκώδη η κίνηση του νερού χαρακτηρίζεται ως αργή, συνήθως μονάχα μερικές δεκάδες μέτρα τον χρόνο, σε αντίθεση με τα καρστικά πετρώματα, όπου οι ταχύτητες ροής του υπόγειου νερού προσεγγίζουν τις τιμές των αντίστοιχων επιφανειακών ροών. Το φαινόμενο αυτό αποτελεί το κύριο χαρακτηριστικό γνώρισμα των καρστικών υδροφορέων και οφείλεται στα τελικά μεγάλα δίκτυα ροής και τα υπόγεια αποχετευτικά συστήματα που σχηματίζουν.

Επίσης ιδιαίτερα σημαντικό είναι το ότι στα κοκκώδη εδάφη, τα υπόγεια νερά διέπονται από τον νόμο του Darcy, με βάση τον οποίο προκύπτει η τελική ανάλυση της προσομοίωσης, προσεγγίζοντας σε ικανοποιητικό βαθμό τις πραγματικές τιμές. Το γεγονός αυτό δεν παρατηρείται στα καρστικά εδάφη, αφού κάθε καρστική περιοχή χαρακτηρίζεται από μοναδικές ιδιαιτερότητες και φυσικά γνωρίσματα, τα οποία, όπως είναι φυσικό, απαιτούν τη δημιουργία ενός ξεχωριστού και πρωτότυπου μοντέλου, το οποίο και θα αντιπροσωπεύει αποκλειστικά τη συγκεκριμένη περιοχή. Κάθε προσπάθεια για έμμεση χρησιμοποίηση μοντέλων, θεωρείται παρακινδυνευμένη, εάν δεν έχει πραγματοποιηθεί εκ των προτέρων εκτενής ανάλυση των εδαφών των δύο περιοχών, προκειμένου να προκύψουν κοινά στοιχεία και ομοιότητες.

Σε όλα τα παραπάνω καλούνται να δώσουν λύση τα τεχνητά νευρωνικά δίκτυα, καθώς είναι σε θέση να παρέχουν αξιόπιστα αποτελέσματα σε περιπτώσεις που η πληροφορία των γεωλογικών χαρακτηριστικών της ευρύτερης περιοχής μελέτης είναι ελλιπής.

Βασικό πλεονέκτημα είναι το ότι λόγω της φύσης τους και της δομής τους, τα τεχνητά νευρωνικά δίκτυα είναι σε θέση να ξεπεράσουν οποιαδήποτε προβλήματα προσομοίωσης του ακριβή μηχανισμού του φυσικού συστήματος του υδροφορέα δίνοντας αξιόλογα αποτελέσματα στηριζόμενα μονάχα σε υδρομετεωρολογικές παραμέτρους και όχι σε γεωλογικές και υδρογεωλογικές ιδιότητες του φυσικού συστήματος (Trichakis et al., 2009).

Στην παρούσα εργασία αναπτύχθηκαν αλγόριθμοι εκπαίδευσης τεχνητών νευρωνικών δικτύων, με στόχο την αναπαραγωγή παροχών μέσω βροχομετρικών δεδομένων. Αρχικά ερευνήθηκε η βέλτιστη και αποδοτικότερη δομή του δικτύου. Αναπτύχθηκαν έξι διαφορετικοί τύποι εκπαίδευσης, ενώ για τον κάθε ένα ξεχωριστά δοκιμάστηκαν τιμές πλήθους νευρώνων, του μοναδικού ενδιάμεσου στρώματος, από ένα ως δέκα.

Λόγω του ότι τα διαθέσιμα στοιχεία δεδομένων αφορούσαν δυόμιση έτη περίπου, επομένως το πλήθος ήταν σχετικά περιορισμένο, το αποτέλεσμα των τεσσάρων νευρώνων στο κρυφό στρώμα, θεωρήθηκε μία καλή προσέγγιση. Επίσης σημαντικό είναι το ότι ο αλγόριθμος με συνάρτηση εκπαίδευσης την *trainlm* έδωσε την βέλτιστη απόδοση και στις δύο υπό μελέτη πηγές Μάη και Πέρα Βρύση. Το γεγονός αυτό ενδέχεται να οφείλεται στην ίδια τη φύση του αλγορίθμου, καθώς κάθε μεταβλητή ρυθμίζεται με βάση τον αλγόριθμο Levenberg – Marquardt, ο οποίος έχει αποδειχτεί αρκετά αξιόλογος (Kermani et al., 2005; Hagan&Menhaj, 1994; Ngia&Sjoberg, 2000).

Επίσης στα πλαίσια της συγκεκριμένης εργασίας μελετήθηκαν ορισμένες προσπάθειες ενίσχυσης της απόδοσης του τεχνητού νευρωνικού δικτύου. Προτάθηκαν διάφορες παραλλαγές τόσο στα διανύσματα εισόδου όσο και στα διανύσματα εξόδου του δικτύου.

Αρχικά μελετήθηκε η επίδραση εφαρμογής της γραμμικής παρεμβολής στις τιμές της παροχής, για συνθήκες βασικής ροής, και τα αποτελέσματα έδειξαν ότι η αύξηση των διανυσμάτων εισόδου-εξόδου ευνοεί στην συγκεκριμένη περίπτωση το δίκτυο. Επομένως τα αποτελέσματα που έδωσε το δίκτυο για τα αρχικά διανύσματα χωρίς την εφαρμογή της γραμμικής παρεμβολής ήταν λιγότερο ικανοποιητικά, καθώς το μέσο τετραγωνικό σφάλμα αυξήθηκε και στις δύο πηγές λαμβάνοντας τις τιμές 0.0034 (l/sec)^2 και 0.0094 (l/sec)^2 αντίστοιχα, αλλά παρόλα αυτά ακολουθούσαν την τάση των δεδομένων ελέγχου. Επομένως συμπεραίνεται ότι σε περίπτωση που τα δεδομένα παροχών είναι ελλιπή, η εφαρμογή της γραμμικής παρεμβολής μονάχα σε συνθήκες βασικής ροής που τα αντίστοιχα ύψη βροχόπτωσης είναι μηδενικά, ευνοεί κατά πολύ την απόδοση του νευρωνικού δικτύου.

Στη συνέχεια αποπειράθηκε μία ακόμη προσπάθεια βελτίωσης της απόκρισης του νευρωνικού δικτύου προσθέτοντας στις παραμέτρους εισόδου στοιχεία θερμοκρασίας. Ο τύπος εκπαίδευσης και η δομή του νευρωνικού δικτύου παρέμειναν τα ίδια. Εν τέλει, και σε αυτήν την περίπτωση το νευρωνικό δίκτυο πραγματοποίησε μία εξίσου καλή εκτίμηση κατά τον έλεγχο, παρέχοντας αξιόλογα

αποτελέσματα. Στην περίπτωση της πρώτης πηγής βελτιστοποιήθηκε η απόκριση, μειώνοντας το μέσο τετραγωνικό σφάλμα κατά τον έλεγχο από 0.003 (l/sec)^2 σε 0.0018 (l/sec)^2 . Όσον αφορά τη δεύτερη πηγή, μπορεί να μην επιτεύχθηκε βέλτιστο αποτέλεσμα ($0.0017 \text{ (l/sec)}^2 > 0.0001 \text{ (l/sec)}^2$), καθώς δεν παρατηρήθηκε μείωση του μέσου τετραγωνικού σφάλματος, παρόλα αυτά είμαστε σε θέση να διακρίνουμε ότι το νευρωνικό δίκτυο πραγματοποίησε μία πολύ καλή εκτίμηση.

Εμπνεόμενοι από το γεγονός ότι στην πηγή Πέρα Βρύση, ο συντελεστής συσχέτισης μεγιστοποιείται μεταξύ τριών διαδοχικών ημερών (με πολύ μικρές αποκλίσεις μεταξύ τους) προτάθηκε μία ακόμη παραλλαγή στις παραμέτρους εισόδου του νευρωνικού δικτύου. Εισήχθησαν, λοιπόν, στο δίκτυο ως δεδομένα εισόδου οι βροχοπτώσεις των τριών ημερών όπου μεγιστοποιείται ο συντελεστής συσχέτισης, ως μεμονωμένα διανύσματα. Το φαινόμενο αυτό παρατηρήθηκε μονάχα στην πηγή Πέρα Βρύση και για τον λόγο αυτό εφαρμόστηκε μονάχα σε αυτήν. Σε σύγκριση με τα προηγούμενα αποτελέσματα, η απόκριση του συγκεκριμένου νευρωνικού δικτύου δε θεωρήθηκε ευνοϊκότερη, καθώς το μέσο τετραγωνικό σφάλμα αυξήθηκε, λαμβάνοντας την τιμή 0.0026 (l/sec)^2 , αλλά παρόλα αυτά, οι παραχθείσες από το νευρωνικό δίκτυο τιμές τείνουν να ακολουθούν την τάση των δεδομένων ελέγχου.

Τέλος, επιχειρήθηκε μία ακόμη δοκιμή εκτέλεσης του νευρωνικού δικτύου, θέτοντας διαφορετικές παραμέτρους εισόδου και δεδομένα εξόδου. Στόχος ήταν η εκτίμηση της αλλαγής των τιμών της παροχής μεταξύ δύο διαδοχικών ημερών. Πιο συγκεκριμένα, τέθηκαν ως διανύσματα εισόδου τα ύψη βροχόπτωσης, ο αριθμός ημέρας και η παροχή της προηγούμενης ημέρας, ενώ ως επιθυμητή έξοδος ορίστηκε η διαφορά της παροχής (σημερινής – προηγούμενης ημέρας). Η συγκεκριμένη εφαρμογή είναι αρκετά δύσκολη ως προς την εκτέλεση, καθώς το πεδίο τιμών της διαφοράς της παροχής ΔQ , είναι σαφώς μικρότερο από το αντίστοιχο πεδίο τιμών της ίδιας της παροχής Q . Έτσι το νευρωνικό δίκτυο επικεντρώνεται κατά την εκπαίδευση στις πολύ μικρές μεταβολές των τιμών των συναπτικών βαρών, με αποτέλεσμα οι αποκρίσεις του δικτύου για τιμές ΔQ να σημειώνουν μία αισθητή απόκλιση από τις μετρηθείσες αναμενόμενες τιμές. Οι τιμές του μέσου τετραγωνικού σφάλματος κατά τον έλεγχο για τις δύο πηγές προέκυψαν 0.0025 (l/sec)^2 και 0.0005 (l/sec)^2 αντίστοιχα. Η δυσκολία της συγκεκριμένης εφαρμογής αποτυπώνεται και μέσω των γραφημάτων των μετρηθείσων τιμών και των αντίστοιχων προβλέψεων του δικτύου. Όταν όμως οι τιμές της διαφοράς της παροχής μετατράπηκαν σε τιμές παροχής (προσθέτοντας την τιμή της παροχής της προηγούμενης ημέρας), παρατηρήθηκε μία σημαντική βελτίωση, καθώς οι μετρηθείσες τιμές και οι αντίστοιχες προβλέψεις του δικτύου σχεδόν ταυίζονταν. Επομένως η πρόταση θεώρησης επιθυμητής εξόδου τη διαφορά παροχής, λειτούργησε θετικά για το δίκτυο και ενίσχυσε σημαντικά την απόδοση του, καταλήγοντας σε βέλτιστα τροποποιημένα σε τιμές παροχής αποτελέσματα, παρά την πολυπλοκότητα της εφαρμογής.

Συμπερασματικά, είμαστε σε θέση να καταλήξουμε στο ότι επιτεύχθηκε μία επιτυχημένη εφαρμογή της τεχνικής των νευρωνικών δικτύων στις δύο αυτές καρστικές πηγές του Δήμου Ρούβα, Ηρακλείου Κρήτης. Επίσης η παραδοχή της γραμμικότητας ανάμεσα στα ύψη βροχόπτωσης και την

απορροή όχι μόνο αύξησε τα ζεύγη δεδομένων του δικτύου, αλλά ταυτοχρόνως απέδωσε μικρότερα σφάλματα κατά τον έλεγχο, με αποτέλεσμα να επιτευχθεί καλύτερη εκτίμηση από το νευρωνικό δίκτυο.

Στοχεύοντας στην βελτίωση της απόδοσης του νευρωνικού δικτύου, η θεώρηση παραμέτρων εισόδου άμεσα σχετιζόμενων με το φυσικό σύστημα οδηγεί σε πιο αξιόπιστες αποκρίσεις του δικτύου. Επομένως κάθε επιπλέον πληροφορία σε σχέση με υδρολογικά και κλιματολογικά δεδομένα ενισχύει σημαντικά τη διαδικασία εκπαίδευσης. Στην συγκεκριμένη εφαρμογή, εκτός από τα δεδομένα θερμοκρασίας που εισήχθησαν σε μία προσπάθεια βελτίωσης της απόκρισης του δικτύου, ενδεχομένως η θεώρηση στοιχείων εξατμισιοδιαπνοής, υδραυλικού φορτίου ή ακόμα και γεωχημικών δεδομένων και γεωφυσικών τεχνικών, να συνέβαλε στην εξαγωγή πιο αξιόλογων αποτελεσμάτων. Επομένως, ένα κρίσιμο ζήτημα για την αποτελεσματικότητα του τεχνητού νευρωνικού δικτύου αποτελεί η κατάλληλη επιλογή των παραμέτρων εισόδου και εξόδου, η οποία είναι σε θέση να αναβαθμίσει τη δυνατότητα πρόβλεψης.

Όσον αφορά τα αποτελέσματα της συγκεκριμένης μεταπτυχιακής διατριβής, ως βέλτιστη θα μπορούσε να θεωρηθεί η εφαρμογή με την πρόσθετη θεώρηση των στοιχείων θερμοκρασίας στα διανύσματα εισόδου του νευρωνικού δικτύου, κρίνοντας από τα αποτελέσματα μέσω των γραφημάτων και από τις μειωμένες τιμές της συνάρτησης απόδοσης σφάλματος. Η εφαρμογή του δικτύου με διάνυσμα επιθυμητής εξόδου τη διαφορά της παροχής, μπορεί να κατέληξε σε βέλτιστα τροποποιημένα σε τιμές παροχής αποτελέσματα, αλλά παρόλα αυτά η πολυπλοκότητα του δικτύου, ενδεχομένως να μειώνει την αξία των αποκρίσεων.

Σύμφωνα με τα παραπάνω, καταλήγουμε στο ότι τα τεχνητά νευρωνικά δίκτυα λόγω της φύσης της δομής και της λειτουργίας τους, είναι σε θέση να απλοποιήσουν θέματα φυσικής, καθώς υπεισέρχονται οι έννοιες των συναπτικών βαρών και των πολώσεων, τα οποία «αντικαταστούν» και «εκφυλίζουν» σταθερές παραμέτρους των φυσικών συστημάτων. Γι αυτό τον λόγο αποτελούν μία αξιόλογη πρόταση προσομοίωσης, ιδιαίτερα σε καρστικούς υδροφορείς, οι οποίοι χρήζουν ιδιαίτερης αντιμετώπισης λόγω ιδιομορφίας.

Τέλος αξίζει να σημειωθεί ότι έχει πραγματοποιηθεί πλήθος εφαρμογών των τεχνητών νευρωνικών δικτύων σε τομείς της υδρολογίας και της διαχείρισης των υδατικών πόρων, λόγω της ευελιξίας τους και της δυνατότητας απόδοσης χρήσιμων συμπερασμάτων, προσφέροντας κίνητρα περαιτέρω έρευνας και βαθύτερης κατανόησης.

Βιβλιογραφία

- Abbott, M. B. and Refsgaard J. C. (1996). “*Distributed Hydrological Modelling*”, Kluwer Academic Publishers.
- Abdi, H. (2001). “*Linear Algebra for Neural Networks*”, Encyclopedia of the Social and Behavioral Sciences, London: Elsevier Science.
- Affandi A., Watanabe K., Tirtomihardjo H. (2007). “Application of an artificial neural network to estimate groundwater level fluctuation”, *Journal of Spatial Hydrology*, 7(2), 23-46.
- Anderson, D.Z. (1988). “*Neural Information Processing Systems*”, American Institute of Physics, New York.
- Aziz, R. A., and Wong, K. F. V. (1992). “A Neural – Network Approach to the Determination of Aquifer Parameters”, *Ground Water*, 30(2), 164-166.
- Bakalowicz, M. (2005). “Karst groundwater: a challenge for new resources”, *Hydrogeology Journal*, 13, 148–160.
- Bonacci, O. (2001). “Analysis of the maximum discharge of karst springs”, *Hydrogeology Journal*, 9, 328–338.
- Bustami, R., Bessaih, N., Bong, C., Suhaili, S. (2007), “Artificial Neural Network for Precipitation and Water Level Predictions of Bedup River”, *IAENG International Journal of Computer Science*, 34:2.
- Case, M., Hall, J. W., Yerramareddy, S., Herman, A., Sun, R., Ranjithan, S. and Westervelt, J. (1993). “Engineering applications of neural networks”, *Journal of Intelligent Manufacturing*, 4, 1-21.
- Chen, J. and Adams, B. J. (2006). “Integration of artificial neural networks with conceptual models in rainfall-runoff modeling”, *Journal of Hydrology*, 318, 232–249.
- Cheng, A. (1984). “Darcy’s flow with variable permeability: a boundary integral solution”, *Water Resources Research*, 20(7), 980-984.
- Cirstea, M.N., Dinu, A., Khor, J.G., McCormick, M. (2002), “*Neural and Fuzzy Logic Control of Drives and Power Systems*”, Published by Elsevier Science.
- Daliakopoulos, I.N., Coulibaly P., Tsanis I.K. (2005). “Groundwater level forecasting using artificial neural networks”, *Journal of Hydrology*, 309, 229–240.

- Davies, William E., 1968, *The earth sciences and speleology*: Reply: National Speleological Society Bulletin, v. 30, p. 93.
- Dawson, C.W. and Wilby, R.L. (2001). "Hydrological modelling using artificial neural networks", *Progress in Physical Geography* 25, 80–108.
- Dayan, P. and Abbott, L. F. (2003). "Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems", Massachusetts Institute of Technology, *Journal of Cognitive Neuroscience* 15(1), 154-155.
- Dowla, F.U. and Rogers, L.L. (1994). "Optimization of groundwater remediation using artificial neural networks with parallel solute transport modeling", *Water Resources Research*, 30(2), 457-481.
- Garcia, L.A., Shigidi, A. (2006). "Using neural networks for parameter estimation in ground water", *Journal of Hydrology*, 318, 215–231.
- Goswami, M., O'Connor, K. M. (2007). "Real-time flow forecasting in the absence of quantitative precipitation forecasts: A multi-model approach", *Journal of Hydrology*, 334, 125–140.
- Gurney, K. "*Computers and Symbols versus Nets and Neurons*", Dept. Human Sciences, Brunel University Uxbridge, Middx.
- Hagan, M.T., Demuth, H.B. and Beale M. (1996). "*Neural Networks Design*", PWS Publishing Company, Boston.
- Hagan, M. T. and Menhaj B. M., (1994). "Training Feedforward Networks with the Marquardt Algorithm", *IEEE TRANSACTIONS ON NEURAL NETWORKS*, 5(6), 989-993.
- Hahn, B. D. and Valentine D. T. (2002), "*Essential MATLAB® for Engineers and Scientists*", Published by Elsevier.
- Halff, A.H., Halff, H.M., Azmoodeh, M., (1993). "Predicting runoff from rainfall using neural networks", *Proc. Engng. Hydrol., ASCE*, New York, 768–775.
- Haykin, S. (1994). "*Neural Networks: A Comprehensive Foundation*", Macmillan Publishing Company, N.Y.
- Hebb, D. O., (1949). "*The organization of behavior - A neurophysiological theory*", Wiley.
- Hecht-Nielsen, R. (1990). "*Neurocomputing*", HNC, Inc. and University of California, San Diego, Addison – Wesley Publishing Company, Inc.
- Hertz J., Krogh A., Palmer R.G. (1991). "*Introduction to the theory of Neural Computation*", Lecture Notes Volume I, Santa FE Institute, Studies in the Science of Complexity, Addison-Wesley Publishing Company.
- Holzbecher, E., (2007). "*Environmental Modeling Using MATLAB*", Springer.
- Hsu, K.L., Gupta, H.V., Sorooshian, S., (1995). "Artificial neural network modeling of the rainfall runoff process", *Water Resour. Res.* 31 (10), 2517–2530.

- Kermani, B. G., Schiffmanb, S. S., Nagle H. T., (2005). "Performance of the Levenberg–Marquardt neural network training method in electronic nose applications", *Sensors and Actuators*, B 110, 13–22
- Kralisch, S., Fink, M., Flu"gel, W.-A. and Beckstein C. (2003). "A neural network approach for the optimisation of watershed management", *Environmental Modelling & Software*, 18, 815–823.
- Krose, B. and Van der Smagt, P. (1996). "*An introduction to neural networks*", University of Amsterdam.
- Lallahem, S. and Mania, J., (2003). "A nonlinear rainfall-runoff model using neural network technique: example in fractured porous media", *Mathematical and Computer Modelling*, 37, 1047-1061.
- Lambrakis, N., Andreou, A.S., Polydoropoulos, P., Georgopoulos, E., Bountis, T., (2000). "Nonlinear analysis and forecasting of a brackish karstic spring", *Water Resour Res.*, 36(4), 875–84.
- Lin, C. J. and Chen, C. H., (2003). "Nonlinear System Control Using Compensatory Neuro-Fuzzy Networks", *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, 86-A, 2309-2316.
- McCulloch, W. S. and Pitts, W., (1943). "A logical calculus of the ideas immanent in nervous activity", *Bull. Math. Biophysics*, (5), 115-133.
- Mathworks, Matlab, R2008a, Tutorial.
- Minsky M. L. and Papert S. A. 1969. "*Perceptron*", Cambridge, MA: MIT Press.
- Nash, J. E., (1957). "The form of the instantaneous unit hydrograph", *Proc. General Assembly*, Toronto 3, IASH, 114-121.
- Ngia, L. S. H. and Sjoberg, J., (2000). "Efficient training of neural nets for nonlinear adaptive filtering using a recursive Levenberg-Marquardt algorithm", *Signal Processing, IEEE Transactions*, 48 (7), 1915-1927.
- Qinghai, G., Yanxin, W., Teng, M. and Luxiu, L. (2005). "Variation of karst spring discharge in the recent five decades as an indicator of global climate change: A case study at Shanxi, northern China", *Science in China Ser. D, Earth Sciences*, 48 (11), 2001-2010.
- Qiu-hao, H., and Yun-long, C. (2006). "Assessment of karst rocky desertification using the radial basis function network model and GIS technique: a case study of Guizhou Province, China", *Environmental Geology*, 49, 1173–1179.
- Raeisi, E. and Karami, G. (1997). "Hydrochemographs of Berghan Karst Spring as Indicators of Aquifer Characteristics", *Journal of Cave and Karst Studies*, 59(3), 112-118.
- Rajurkara, M.P., Kothyarib, U.C., and Chaubec, U.C., (2004). "Modeling of the daily rainfall-runoff relationship with artificial neural network", *Journal of Hydrology*, 285, 96–113.
- Riad, S., Mania J., Bouchaou L., Najjar Y., (2004). "Predicting catchment flow in a semi-arid region via an artificial neural network technique", *Hydrological Processes*, 18 (13), 2387 - 2393

- Rogers, L.L. (1992). “*Optimal groundwater remediation using artificial neural networks and the genetic algorithm*”, Stanford University CA, USA.
- Rogers, L.L. and Dowla, F.U., (1994), “Optimization of groundwater remediation using artificial neural networks with parallel solute transport modeling”, *Water Resources Research*, 30 (2), 457-481.
- Rojas, R., (1996). “*Neural Networks*”, Published by Springer-Verlag, Berlin.
- Rosenblatt, Frank (1958), The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain, Cornell Aeronautical Laboratory, *Psychological Review*, v65, No. 6, pp. 386-408.
- Samani, N., Gohari-Moghadam, M. and Safavi, A.A. (2007). “A simple neural network model for the determination of aquifer parameters”, *Journal of Hydrology*, 340, 1– 11.
- Sedki, A., Ouazar, D., Mazoudi E., (2009). “Evolving neural network using real coded genetic algorithm for daily rainfall-runoff forecasting”. *Expert Syst. Appl. (ESWA)*, 36 (3), 4523-4527.
- Shamseldin, A. Y. (1997) Application of a neural network technique to rainfall-runoff modeling.. *J. Hydrol.* 199, 272-294.
- Sherman, L. K., (1932). “Streamflow from rainfall by the unit-graph method”, *Eng. News Rec.* 108, 501-505.
- Sjöberg, J. (2005). “*Mathematica, Neural Networks, Train and analyze neural networks to fit your data*”, Published by Wolfram Research, Inc.
- Solaimani, K. (2009). “Rainfall-runoff Prediction Based on Artificial Neural Network (A Case Study: Jarahi Watershed)”, *American-Eurasian J. Agric. & Environ. Sci.*, 5 (6), 856-865.
- Srinivasulu, S. and Jain, A. (2006). “A comparative analysis of training methods for artificial neural network rainfall-runoff models”, *Applied Soft Computing*, 6, 295–306, Published by Elsevier.
- Trichakis, I., Nikolos, I., Karatzas, G. (2009). “Optimal selection of artificial neural network parameters for the prediction of a karstic aquifer’s response”, *Hydrological Processes*, 23, 2956-2969.
- Vashisht, A. K. and Sharma, H. C. (2007). “Study on hydrological behaviour of a natural spring”, *CURRENT SCIENCE*, 93(6), 837 – 840.
- Warner, B. and Misra, M. (1996). “Understanding Neural Networks as Statistical Tools”, *The American Statistician*, 50(4), 284-293
- White, W. B. (2002). “Karst hydrology: recent developments and open questions”, *Engineering Geology*, 65, 85–105.
- Yang, Z.P., Lu, W.X., Long, Y.Q., Li, P. (2009). “Application and comparison of two prediction models for groundwater levels: A case study in Western Jilin Province, China”, *Journal of Arid Environments*, 73, 487–492.
- Zio, E. (1997). “Approaching the inverse problem of parameter estimation in groundwater models by means of artificial neural networks”, *Progress in Nuclear energy*, 31(3), 301-315.

- Αργυράκης, Π. (2001). «*Νευρωνικά Δίκτυα & Εφαρμογές*», Ελληνικό Ανοικτό Πανεπιστήμιο, Σχολή Θετικών Επιστημών και Τεχνολογίας, Πάτρα.
- Αρναουτάκης Ν. (2005). «*Εισαγωγή στο Matlab – Simulink*», Τμήμα Μηχανικών Παραγωγής και Διοίκησης, Πολυτεχνείο Κρήτης.
- Γάλλο-Ελληνική σπηλαιολογική αποστολή στα Λευκά Όρη Χανίων, 11 Αυγούστου 2008 (CATAMARAN, CDS Doubs, CDS Ain, GSHL από Γαλλία και ΣΕΛΑΣ, ΦΟΣ, ΣΠΟΚ και ΕΟΣ ΧΑΝΙΩΝ από Ελλάδα).
- Δαριβιανάκης, Ν. (2007). «*Διαχείριση υπόγειων και επιφανειακών υδάτινων πόρων Δήμου Ρούβα, Ηρακλείου*», *Εσωτερική Εργασία*, Τμήμα Πολιτικών Μηχανικών Α.Π.Θ.
- Διαμαντάρας, Κ. (2007). «*Τεχνητά Νευρωνικά Δίκτυα*», Εκδόσεις Κλειδάριθμος, Θεσσαλονίκη.
- Καπαμπερίδου, Χ. (2007). «*Βελτιστοποίηση εκμετάλλευσης υπόγειων υδατικών πόρων με χρήση νευρωνικών δικτύων*», Μεταπτυχιακή Διπλωματική Εργασία, Τμήμα Πολιτικών Μηχανικών, Α.Π.Θ.
- Κατσιφάρακης, Κ. (2006), «*Ειδικά κεφάλαια διαχείρισης υδατικών πόρων*», Τμήμα Πολιτικών Μηχανικών, Τμήμα Υδραυλικής και Τεχνικής Περιβάλλοντος, Α.Π.Θ.
- Μωραΐτης, Δ. (2009), Φωτογραφίες καρστικών μορφών, Προσωπική επικοινωνία.
- Πάνου, Ε. (2003). «*Υλοποίηση αλγορίθμου πρόβλεψης ζήτησης με χρήση τεχνητών νευρωνικών δικτύων*», Διπλωματική εργασία, Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών, Ε.Μ.Π.
- Πεταλά, Ζ.Ι. (2004). «*Βελτιστοποίηση της Διαχείρισης Παράκτιων Υδροφορέων με τη Χρήση Γενετικών Αλγορίθμων*», Διδακτορική Διατριβή, Τμήμα Πολιτικών Μηχανικών, Α.Π.Θ.
- Σκιτζή, Ε., Φοιτός, Σ., Χρονάκης, Ν. (2007). «*Κατανομή των βροχοπτώσεων σε καρστικές πηγές με τη βοήθεια των γενετικών αλγορίθμων*», Διπλωματική Εργασία, Τμήμα Πολιτικών Μηχανικών, Τμήμα Υδραυλικής και Τεχνικής Περιβάλλοντος, Α.Π.Θ.
- Τσεκούρας, Γ. (2004). «*Εφαρμογή της μεθόδου των τεχνητών νευρωνικών δικτύων σε θέματα σηράγγων*», Διπλωματική Εργασία, Τμήμα Πολιτικών Μηχανικών, Ε.Μ.Π.

Διαδικτυακοί Σύνδεσμοι:

- www.wolfram.com (18/06/09, 11:41)
- www.cs.cmu.edu/~awm (04/05/09, 14:01)
- <http://www.willamette.edu/~gorr/classes/cs449/intro.html> (01/12/09, 12:56)
- <http://en.wikipedia.org/wiki/> (29/11/09, 15:45)
- <http://www.dimosrouva.gr/rouvas/frontend/index.jsp> (29/11/09, 16:32)
- http://creteweather.blogspot.com/2009/04/blog-post_22.html (30/11/09, 09:44)
- <http://www.learnartificialneuralnetworks.com/perceptronadaline.html> (12/12/09, 18:46)
- www.selas.org (15/03/10, 17:08)
- <http://www.googleearth.gr/> (29/11/09, 17:54)