



Πολυτεχνείο Κρήτης
Σχολή Μηχανικών Περιβάλλοντος

ΕΚΤΙΜΗΣΗ ΤΗΣ ΣΤΑΘΜΗΣ ΥΠΟΓΕΙΩΝ ΥΔΑΤΩΝ ΣΤΗΝ ΕΥΡΥΤΕΡΗ ΠΕΡΙΟΧΗ ΤΟΥ ΠΟΤΑΜΟΥ ΔΟΥΝΑΒΗ ΜΕ ΧΡΗΣΗ ΤΕΧΝΗΤΩΝ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Λάνδρος Ι. Ηλίας

ΤΡΙΜΕΛΗΣ ΕΠΙΤΡΟΠΗ:

Καθηγητής Καρατζάς Γεώργιος (ΕΠΙΒΛΕΠΩΝ)

Καθηγητής Νικολαΐδης Νικόλαος

Δρ. Τριχάκης Ιωάννης

ΧΑΝΙΑ, ΙΟΥΛΙΟΣ, 2019

Περίληψη

Μέσα στο πλαίσιο της συγκεκριμένης διπλωματικής εργασίας εξετάζεται η χρήση και η εκπαίδευσή τεχνητών νευρωνικών δικτύων, με σκοπό την προσομοίωση της υπόγειας στάθμης των πηγαδιών στην ευρύτερη περιοχή του ποταμού Δούναβη. Ο ποταμός Δούναβης εκτείνεται σε 10 διαφορετικές χώρες, αλλά στην περίπτωση μας τα τυχαία πηγάδια παρατήρησης ως προς έρευνα βρίσκονται στην Αυστρία, Βουλγαρία, Γερμανία, Κροατία, Ουγγαρία, Ρουμανία, Σερβία και Σλοβενία.

Τα τεχνητά νευρωνικά δίκτυα προσπαθούν να προσομοιώσουν την υπόγεια στάθμη των πηγαδιών μέσα από το υδατικό ισοζύγιο. Αρχικά χρειαζόταν η κατάλληλη προεπεξεργασία των δεδομένων για την δημιουργία του πίνακα εισόδου και διανύσματος στόχου στο νευρωνικό δίκτυο. Ο πίνακας εισόδου αποτελείται από μετεωρολογικά δεδομένα, συντεταγμένες των πηγαδιών, χρονολογία, δυνητική εξατμισοδιαπνοή, εξάτμιση από υδάτινες επιφάνειες και εξατμισοδιαπνοή από το έδαφος και τα φυτά. Το διάνυσμα στόχου περιέχει τις πραγματικές τιμές της υπόγειας στάθμης των πηγαδιών. Τα δεδομένα που χρησιμοποιούνται για την εκπαίδευση των τεχνητών νευρωνικών δικτύων περιέχουν τιμές από 01/01/2000-31/10/2014 για συνολικά 128 πηγάδια παρατήρησης.

Μετά την ολοκλήρωση της προ επεξεργασίας των δεδομένων, τεχνητά νευρωνικά δίκτυα εκπαιδεύτηκαν με τα εργαλεία Neural Fitting tool (nftool) και Neural Network tool (nntool). Οι δυο αλγόριθμοι εκπαίδευσης που χρησιμοποιήθηκαν είναι οι Levenberg-Marquardt και Bayesian Regularization. Η εκπαίδευση των τεχνητών νευρωνικών δικτύων πραγματοποιήθηκε με βάση τα παραπάνω για διαφορετικές παραμέτρους κάθε φορά ως προς τους κρυφούς κόμβους, τα ποσοστά εκπαίδευσης καθώς και τους αλγόριθμους εκπαίδευσης.

Κατά την διάρκεια εκπαίδευσης τεχνητών νευρωνικών δικτύων, προσπαθούμε να εντοπίσουμε το μοντέλο με τις παραμέτρους από το οποίο θα προκύψουν τα βέλτιστα αποτελέσματα. Κριτήρια επιλογής για την επιλογή του βέλτιστου μοντέλου ήταν η τετραγωνική ρίζα του μέσου τετραγωνικού σφάλματος, ο συντελεστής συσχέτισης καθώς υπολογίστηκαν και κάποιοι επιπλέον δείκτες που χρησιμοποιούνται συχνά στην υδρολογία.

Στην συγκεκριμένη διπλωματική εργασία τεχνητά νευρωνικά δίκτυα εκπαιδεύτηκαν με βάση όλα τα πηγάδια παρατήρησης. Μετά από την επιλογή του βέλτιστου τεχνητού νευρωνικού δικτύου, από εκείνα που εκπαιδεύτηκαν με τα δεδομένα όλων των πηγαδιών εξετάστηκε και το ενδεχόμενο της εκπαίδευσης ενός τεχνητού νευρωνικού δικτύου για ένα μόνο τυχαίο πηγάδι παρατήρησης για να συγκρίνουμε την συμπεριφορά του τεχνητού νευρωνικού δικτύου.

Τέλος, συνοψίζοντας τα αποτελέσματα μας επιτεύχθηκε σφάλμα της τάξεως $10^{-1}m$ με την χρήση του αλγόριθμου Bayesian Regularization με βάση όλα τα πηγάδια παρατήρησης. Όσον αφορά τα αποτελέσματα από την εκπαίδευση τεχνητού νευρωνικού δικτύου για ένα πηγάδι μόνο και όχι για όλα τα πηγάδια μαζί, έδωσε καλύτερα αποτελέσματα ως προς την προσομοίωση της υπόγειας στάθμης καθώς οι τιμές βρίσκονται πιο κοντά στις παρατηρημένες τιμές αλλά το τεχνητό νευρωνικό δίκτυο δεν μπορεί να εκτιμήσει τις ακραίες τιμές.

Abstract

In the context of this thesis, the use and training of artificial neural networks is examined to simulate the underground level of the wells in the wider Danube River region. The Danube River extends into 10 different countries, but in our case the accidental observation wells for research are in Austria, Bulgaria, Germany, Croatia, Hungary, Romania, Serbia and Slovenia.

Artificial neural networks try to simulate the underground level of the wells through the water balance. Initially, a proper pre-processing of the data needed to create the input and target vectors in the neural network was needed. The input vector consists of meteorological data, well coordinates, chronology, potential evapotranspiration, evaporation from water surfaces and evapotranspiration from soil and plants. The target vector contains the actual values of the underground level of the wells. The vectors contain values from 01/01/2000-31/10/2014 for a total of 128 observation wells.

After the data pre-processing was completed, artificial neural networks were trained with Neural Fitting tool (nftool) and Neural Network tool (nntool). The two training algorithms used are Levenberg-Marquardt and Bayesian Regularization. The training of artificial neural networks was based on the above for different parameters each time in terms of hidden nodes, training percentages and training algorithms.

During the training of artificial neural networks, we try to locate the model with the parameters from which the best results will emerge. Selection criteria for selecting the optimal model were the square root of the mean square error, the correlation coefficient as well as some additional indicators commonly used in hydrology.

In this dissertation, artificial neural networks were trained based on all observation wells. After choosing the optimal artificial neural network, from those trained with the data of all wells we also studied the possibility of training an artificial neural network for a single random observation well to compare the behavior of the artificial neural network.

Finally, summarizing our results, a fault of the order 10^{-1} m was achieved using the Bayesian Regularization algorithm based on all observation wells. As far as the results from the training of an artificial neural network for a single well and not for all wells together, it gave better results in simulation of the underground level as the values are closer to the observed values but the artificial neural network can not simulate the extreme values.

Ευχαριστίες

Μετά το πέρας 5 ετών και την επιτυχή ολοκλήρωση των προπτυχιακών σπουδών μου για την απόκτηση του διπλώματος Μηχανικού Περιβάλλοντος της Σχολής Μηχανικών Περιβάλλοντος του Πολυτεχνίου Κρήτης, θα ήθελα να ευχαριστήσω όλους όσους με στήριξαν και με καθοδηγούσαν καθ'όλη την διάρκεια των σπουδών μου.

Πρώτα, θα ήθελα να ευχαριστήσω την οικογένεια μου και τους φίλους που με παρότρυναν να σπουδάσω και να αποκτήσω το δίπλωμα του Μηχανικού Περιβάλλοντος.

Ιδιαίτέρως θέλω να ευχαριστήσω τον επιβλέποντα Καθηγητή Γεώργιο Καρατζά για την καθοδήγηση μέσα από τις συμβουλές του, που με ώθησαν να ασχοληθώ με τον κλάδο της Υπόγειας Υδραυλικής όπου με στήριξε για τις αποφάσεις μου και με βοήθησε να κατασταλάξω.

Πολύτιμη επίσης ήταν η βοήθεια του Ιωάννη Τριχάκη, του οποίου οφείλω ένα μεγάλο ευχαριστώ, που με βοήθησε να κατανοήσω την χρησιμότητα των εργαλείων τεχνητής νοημοσύνης και να μπορώ να εξηγήσω την συμπεριφορά τους.

Επίσης ένα ευχαριστώ στον Εμμανουήλ Βαρουχάκη για την πολύτιμη βοήθεια για την εξαγωγή των χαρτών καθώς και τα βοηθητικά σχόλια.

Επίσης ένα ευχαριστώ στα παιδιά του εργαστηρίου Χριστίνα Στυλιανουδάκη και Χρήστο Γκούμα.

Επίσης ένα μεγάλο ευχαριστώ οφείλω στην τριμελή επιτροπή που διέθεσαν χρόνο για την διόρθωση και τα επικοδομητικά σχόλια ως προς την βελτίωση της διπλωματικής εργασίας.

Περιεχόμενα

Περίληψη.....	ii
Abstract.....	iii
Κεφάλαιο 1	1
1.1 Εισαγωγή.....	1
1.2 Σκοπός	1
1.3 Διαχείριση υδάτινων πόρων.....	1
1.4 Υδρολογικός κύκλος	3
1.5 Περιοχή μελέτης.....	5
1.6 Μεταβλητές που χρησιμοποιήθηκαν	7
1.7 Εφαρμογές Τεχνητών Νευρωνικών Δικτύων στην υδρολογία.....	7
Κεφάλαιο 2.....	9
Θεωρητικό Υπόβαθρο	9
2.1 Ανθρώπινος εγκέφαλος	9
2.2 Τεχνητά Νευρωνικά Δίκτυα (ΤΝΔ).....	10
2.2.1 Multi-layer Perceptrons	12
2.3 Συναρτήσεις ενεργοποίησης	14
2.3.1 Συνάρτηση κατωφλίου	14
2.3.2 Ήμι-γραμμική συνάρτηση.....	15
2.3.3 Σιγμοειδής συνάρτηση	16
2.4 Αλγόριθμοι εκπαίδευσης	16
2.4.1 Levenberg-Marquardt	16
2.4.2 Bayesian Regularization Backpropagation.....	18
2.5 Κριτήρια αξιολόγησης	19
2.6 Πλεονεκτήματα ΤΝΔ	21
2.7 Κριτήρια Τερματισμού.....	24
2.7.1 Εποχές (Epochs)	24
2.7.2 Σφάλμα εκπαίδευσης (Training error).....	24
2.7.3 Πρώιμος τερματισμός εκπαίδευσης (Early stopping).....	24
2.7.4 Μη τροποποίηση βαρών	25
2.8 Υπερεκπαίδευση του δικτύου (Overtaining)	25
Κεφάλαιο 3.....	27
Μεθοδολογία	27
3.1 Βήματα εκπαίδευσης νευρωνικού δικτύου με Neural Fitting tool	29
3.2 Βήματα εκπαίδευσης νευρωνικού δικτύου με Neural Network tool	33
Κεφάλαιο 4.....	37

Αποτελέσματα	37
4.1 Χρήση αλγόριθμου Levenberg-Marquardt.....	37
4.2 Χρήση αλγορίθμου Bayesian Regularization.....	50
4.3 Χρήση της εντολής nntool στο Command Window του επιστημονικού προγράμματος της Matlab	55
4.3.1 Χρήση αλγόριθμου Levenberg-Marquardt.....	55
4.3.2 Χρήση αλγορίθμου Bayesian Regularization.....	57
4.4 Γραφική απεικόνιση σύγκρισης παρατηρούμενων-προσομοιωμένων τιμών υπόγειας στάθμης για κάποιο πηγάδι ανά χώρα.....	59
4.5 Εκπαίδευση νευρωνικού δικτύου μόνο για ένα πηγάδι.....	63
Κεφάλαιο 5.....	67
5.1 Χωρική Παρεμβολή με την τεχνική του Kriging για την δημιουργία χαρτών με τις ισοϋψείς μεταξύ παρατηρημένων και προσομοιωμένων τιμών της υπόγειας στάθμης για τις χρονιές 2000, 2005, 2010 και 2012.	67
Κεφάλαιο 6.....	72
Συζήτηση.....	72
Κεφάλαιο 7.....	76
Συμπεράσματα	76
ΒΙΒΛΙΟΓΡΑΦΙΑ	78

Κατάλογος Πινάκων

Πίνακας 1: Παράμετροι αλγορίθμου LM (https://www.mathworks.com).....	17
Πίνακας 2: Παράμετροι αλγορίθμου BR (https://www.mathworks.com).....	19
Πίνακας 3: Απαραίτητοι παράμετροι για την εκπαίδευση με τον αλγόριθμο Levenberg-Marquardt.....	44
Πίνακας 4: Αποτελέσματα για το μέσο τετραγωνικό σφάλμα (MSE) και τον συντελεστή συσχέτισης (R)	44
Πίνακας 5: Αποτελέσματα για τον αριθμό των εποχών (Iterations), τον χρόνο εκπαίδευσης (Running time), το Validation Checks και το ριζικό του μέσου τετραγωνικού σφάλματος (RMSE).....	44
Πίνακας 6: Απαραίτητοι παράμετροι για την εκπαίδευση με τον αλγόριθμο Levenberg-Marquardt με διαφορετικά Percentages.....	49
Πίνακας 7: Αποτελέσματα για το μέσο τετραγωνικό σφάλμα (MSE) και τον συντελεστή συσχέτισης (R)	49
Πίνακας 8: Αποτελέσματα για τον αριθμό των εποχών (Iterations), τον χρόνο εκπαίδευσης (Running time), του Validation Check και της τετραγωνικής ρίζας του μέσου τετραγωνικού σφάλματος (RMSE)	49
Πίνακας 9: Απαραίτητοι παράμετροι για την εκπαίδευση με τον αλγόριθμο Bayesian Regularization και με διαφορετικά Percentages.....	54
Πίνακας 10: Αποτελέσματα για το μέσο τετραγωνικό σφάλμα (MSE) και τον συντελεστή συσχέτισης (R)	54
Πίνακας 11: Αποτελέσματα για τον αριθμό των εποχών (Iterations), τον χρόνο εκπαίδευσης (Running time) και της τετραγωνικής ρίζας του μέσου τετραγωνικού σφάλματος (RMSE).....	54
Πίνακας 12: Απαραίτητοι παράμετροι για την εκπαίδευση με τον αλγόριθμο Levenberg-Marquardt με χρήση της εντολής nntool.....	56
Πίνακας 13: Αποτελέσματα για το μέσο τετραγωνικό σφάλμα (MSE) και τον συντελεστή συσχέτισης (R)	56
Πίνακας 14: Αποτελέσματα για τον αριθμό των εποχών (Iterations), τον χρόνο εκπαίδευσης (Running time), του Validation Check και της τετραγωνικής ρίζας του μέσου τετραγωνικού σφάλματος (RMSE)	56
Πίνακας 15: Απαραίτητοι παράμετροι για την εκπαίδευση με τον αλγόριθμο Bayesian Regularization με χρήση της εντολής nntool.....	58
Πίνακας 16: Αποτελέσματα για το μέσο τετραγωνικό σφάλμα (MSE) και τον συντελεστή συσχέτισης (R)	58
Πίνακας 17: Αποτελέσματα για τον αριθμό των εποχών (Iterations), τον χρόνο εκπαίδευσης (Running time), και της τετραγωνικής ρίζας του μέσου τετραγωνικού σφάλματος (RMSE).....	58
Πίνακας 18: Αποτελέσματα από την εκπαίδευση νευρωνικού για ένα μόνο πηγάδι με χρήση αλγορίθμου BR.....	64
Πίνακας 19: Αποτελέσματα σφάλματος RMSE και συντελεστή συσχέτισης από το ΤΝΔ που εκπαιδεύτηκε για ένα πηγάδι παρατήρησης και από το ΤΝΔ που εκπαιδεύτηκε για όλα τα πηγάδια παρατήρησης μαζί	65

Κατάλογος Διαγραμμάτων

Διάγραμμα 1: Παλινδρόμηση εκπαίδευσης (Training Regression) με 20 κρυφούς κόμβους (Hidden nodes), με ποσοστά 70% Training, 15% Validation & 15% Testing, διακοπή εκπαίδευσης λόγω μέγιστου αριθμού Validation Checks	37
Διάγραμμα 2: Παλινδρόμηση εκπαίδευσης (Training Regression) με 25 κρυφούς κόμβους (Hidden nodes), με ποσοστά 70% Training, 15% Validation & 15% Testing	38
Διάγραμμα 3: Παλινδρόμηση εκπαίδευσης (Training Regression) με 30 κρυφούς κόμβους (Hidden nodes), με ποσοστά 70% Training, 15% Validation & 15% Testing, διακοπή εκπαίδευσης λόγω μέγιστου αριθμού Validation Checks	39
Διάγραμμα 4: Παλινδρόμηση εκπαίδευσης (Training Regression) με 35 κρυφούς κόμβους (Hidden nodes), με ποσοστά 70% Training, 15% Validation & 15% Testing	40
Διάγραμμα 5: Παλινδρόμηση εκπαίδευσης (Training Regression) με 50 κρυφούς κόμβους (Hidden nodes), με ποσοστά 70% Training, 15% Validation & 15% Testing	41
Διάγραμμα 6: Παλινδρόμηση εκπαίδευσης (Training Regression) με 80 κρυφούς κόμβους (Hidden nodes), με ποσοστά 70% Training, 15% Validation & 15% Testing	42
Διάγραμμα 7: Παλινδρόμηση εκπαίδευσης (Training Regression) με 100 κρυφούς κόμβους (Hidden nodes), με ποσοστά 70% Training, 15% Validation & 15% Testing	43
Διάγραμμα 8: Παλινδρόμηση εκπαίδευσης (Training Regression) με 100 κρυφούς κόμβους (Hidden nodes), με ποσοστά 55% Training, 25% Validation & 20% Testing	45
Διάγραμμα 9: Παλινδρόμηση εκπαίδευσης (Training Regression) με 100 κρυφούς κόμβους (Hidden nodes), με ποσοστά 50% Training, 25% Validation & 25% Testing	46
Διάγραμμα 10: Παλινδρόμηση εκπαίδευσης (Training Regression) με 100 κρυφούς κόμβους (Hidden nodes), με ποσοστά 30% Training, 35% Validation & 35% Testing	47
Διάγραμμα 11: Παλινδρόμηση εκπαίδευσης (Training Regression) με 100 κρυφούς κόμβους (Hidden nodes), με ποσοστά 60% Training, 20% Validation & 20% Testing	48
Διάγραμμα 12: Παλινδρόμηση εκπαίδευσης (Training Regression) με 10 κρυφούς κόμβους (Hidden nodes), με ποσοστά 70% Training, 15% Validation & 15% Testing	50
Διάγραμμα 13: Παλινδρόμηση εκπαίδευσης (Training Regression) με 30 κρυφούς κόμβους (Hidden nodes), με ποσοστά 70% Training, 10% Validation & 20% Testing	51
Διάγραμμα 14: Παλινδρόμηση εκπαίδευσης (Training Regression) με 80 κρυφούς κόμβους (Hidden nodes), με ποσοστά 70% Training, 10% Validation & 20% Testing	52
Διάγραμμα 15: Παλινδρόμηση εκπαίδευσης (Training Regression) με 100 κρυφούς κόμβους (Hidden nodes), με ποσοστά 70% Training, 5% Validation & 25% Testing	53
Διάγραμμα 16: Παλινδρόμηση εκπαίδευσης (Training Regression) με 100 κρυφούς κόμβους (Hidden nodes)	55

Διάγραμμα 17: Παλινδρόμηση εκπαίδευσης (Training Regression) με 100 κρυφούς κόμβους (Hidden nodes)	57
Διάγραμμα 18: Παλινδρόμηση εκπαίδευσης (Training Regression) με 10 κρυφούς κόμβους (Hidden nodes), με ποσοστά 70% Training, 5% Validation & 25% Testing για ένα μόνο πηγάδι	63

Κατάλογος Γραφημάτων

Γράφημα 1: Σύγκριση μεταξύ παρατηρούμενης-προσομοιωμένης υπόγειας στάθμης νερού σε ένα πηγάδι της Αυστρίας	59
Γράφημα 2: Σύγκριση μεταξύ παρατηρούμενης-προσομοιωμένης υπόγειας στάθμης νερού σε ένα πηγάδι της Βουλγαρίας	59
Γράφημα 3: Σύγκριση μεταξύ παρατηρούμενης-προσομοιωμένης υπόγειας στάθμης νερού σε ένα πηγάδι της Γερμανίας.....	60
Γράφημα 4: Σύγκριση μεταξύ παρατηρούμενης-προσομοιωμένης υπόγειας στάθμης νερού σε ένα πηγάδι της Κροατίας	60
Γράφημα 5: Σύγκριση μεταξύ παρατηρούμενης-προσομοιωμένης υπόγειας στάθμης νερού σε ένα πηγάδι της Ουγγαρίας.....	61
Γράφημα 6: Σύγκριση μεταξύ παρατηρούμενης-προσομοιωμένης υπόγειας στάθμης νερού σε ένα πηγάδι της Ρουμανίας.....	61
Γράφημα 7: Σύγκριση μεταξύ παρατηρούμενης-προσομοιωμένης υπόγειας στάθμης νερού σε ένα πηγάδι της Σερβίας	62
Γράφημα 8: Σύγκριση μεταξύ παρατηρούμενης-προσομοιωμένης υπόγειας στάθμης νερού σε ένα πηγάδι της Σλοβενίας.....	62
Γράφημα 9: Σύγκριση μεταξύ παρατηρούμενης (μπλε γραμμή)-προσομοιωμένης με βάση όλα τα δεδομένα (κόκκινη γραμμή)-προσομοιωμένης με βάση τα δεδομένα ενός πηγαδιού μόνο (πράσινη γραμμή) υπόγειας στάθμης για ένα πηγάδι της Αυστρίας	64

Κατάλογος Εικόνων

Εικόνα 1.1 Κατανομή νερού στην γη (https://water.usgs.gov)	2
Εικόνα 1.2: Ποσοστό νερού ανά πηγή (https://water.usgs.gov)	2
Εικόνα 1.3: Ο υδρολογικός κύκλος (https://water.usgs.gov/edu/watercycle.html)	3
Εικόνα 1.4: Γένεση του Ποταμού Δούναβη (https://el.wikipedia.org/wiki/Δούναβης) ..	5
Εικόνα 1.5: Όρια περιοχής μελέτης και απεικόνιση πηγαδιών παρατήρησης.....	6
Εικόνα 1.6 Δέλτα του Δούναβη (https://el.wikipedia.org/wiki/Δέλτα_Δούναβη).....	6
Εικόνα 2.7: Σχηματική αναπαράσταση ανθρώπινου νευρικού συστήματος (Simon Haykin, Neural network & learning mechanics)	9
Εικόνα 2.8: Τεχνητό Νευρωνικό Δίκτυο (https://el.wikipedia.org/wiki/Νευρωνικό_δίκτυο).....	10
Εικόνα 2.9: Τεχνητό νευρωνικό δίκτυο με δύο κρυφά επίπεδα (Simon Haykin, Neural network & learning mechanics)	11
Εικόνα 2.10: Μη γραμμικό μοντέλο νευρώνα k (Simon Haykin, Neural network & learning mechanics)	12
Εικόνα 2.11: Κατευθύνσεις ροής των δύο βασικών σημάτων ενός Perceptron πολλών επιπέδων (Simon Haykin, Neural network & learning mechanics)	13
Εικόνα 2.12: Συνάρτηση κατωφλίου (Simon Haykin, Neural network & learning mechanics).....	15
Εικόνα 2.13: Ημι-γραμμική συνάρτηση (Ben Kröse & Patrick van der Smag, An Introduction to Neural Networks)	15
Εικόνα 2.14: Σιγμοειδής συνάρτηση για μεταβαλλόμενη παράμετρο α (Simon Haykin, Neural network & learning mechanics)	16
Εικόνα 2.15: Γραφική αναπαράσταση του κανόνα πρώιμου τερματισμού (Simon Haykin, Neural network & learning mechanics)	24
Εικόνα 3.16: Αρχιτεκτονική δικτύου όπου το κρυφό επίπεδο και το επίπεδο εξόδου ενεργοποιούνται με βάση σιγμοειδή συνάρτηση	35

Κατάλογος Χαρτών

Χάρτης 1: Χωρική παρεμβολή των παρατηρημένων τιμών της υπόγειας στάθμης για το έτος 2000 ^[1]	67
Χάρτης 2: Χωρική παρεμβολή των προσομοιωμένων τιμών της υπόγειας στάθμης για το έτος 2000 ^[2]	67
Χάρτης 3: Χωρική παρεμβολή των παρατηρημένων τιμών της υπόγειας στάθμης για το έτος 2005 ^[3]	68
Χάρτης 4: Χωρική παρεμβολή των προσομοιωμένων τιμών της υπόγειας στάθμης για το έτος 2005 ^[4]	68
Χάρτης 5: Χωρική παρεμβολή των παρατηρημένων τιμών της υπόγειας στάθμης για το έτος 2010 ^[5]	69
Χάρτης 6: Χωρική παρεμβολή των προσομοιωμένων τιμών της υπόγειας στάθμης για το έτος 2010 ^[6]	69
Χάρτης 7: Χωρική παρεμβολή των παρατηρημένων τιμών της υπόγειας στάθμης για το έτος 2012 ^[7]	70
Χάρτης 8: Χωρική παρεμβολή των προσομοιωμένων τιμών της υπόγειας στάθμης για το έτος 2012 ^[8]	70

Κεφάλαιο 1

1.1 Εισαγωγή

Η συγκεκριμένη διπλωματική εργασία πραγματοποιείται από τον φοιτητή Ηλία Λάνδρο του τμήματος Μηχανικών Περιβάλλοντος του Πολυτεχνείου Κρήτης με έδρα τα Χανιά. Αποτελεί τμήμα των υποχρεώσεων των προπτυχιακών σπουδών για την ανακήρυξη ως Διπλωματούχος Μηχανικός Περιβάλλοντος και την απόκτηση του Διπλώματος.

1.2 Σκοπός

Σκοπός της παρούσας Διπλωματικής εργασίας είναι η μάθηση και κατανόηση της λειτουργίας από την χρήση τεχνικών υπολογιστικής νοημοσύνης (Τεχνητά Νευρωνικά Δίκτυα-ΤΝΔ). Αυτά έχουν την ικανότητα να περιγράφουν πολύπλοκες σχέσεις μεταξύ διαφορετικών παραμέτρων ταυτόχρονα. Η χρήση του Τεχνητού Νευρωνικού Δικτύου θα χρησιμοποιηθεί για την εκτίμηση της υπόγειας στάθμης των πηγαδιών στις λεκάνες απορροής του ποταμού Δούναβη, με βάση υδρολογικά και άλλα δεδομένα που έχουν αναλυθεί παρακάτω. Στόχος του Τεχνητού Νευρωνικού Δικτύου είναι να εκπαιδευτεί με βάση αυτές τις τιμές εισόδου που έχουμε εισάγει με διαφορετικούς αλγόριθμους εκπαίδευσης. Οι αλγόριθμοι που θα χρησιμοποιηθούν είναι οι Levenberg-Marquardt και Bayesian Regularization. Σκοπός είναι να μπορέσει να εκτιμήσει την στάθμη των πηγαδιών με όσο το δυνατόν μικρότερο σφάλμα. Ως κριτήρια αξιολόγησης επιλέχθηκαν το μέσο τετραγωνικό σφάλμα και ο συντελεστής συσχέτισης καθώς και κάποιοι ακόμα δείκτες που χρησιμοποιούνται στην υδρολογία. Με βάση τα κριτήρια αξιολόγησης θα κριθεί ποια μέθοδος και με ποια χαρακτηριστικά ήταν η καλύτερη.

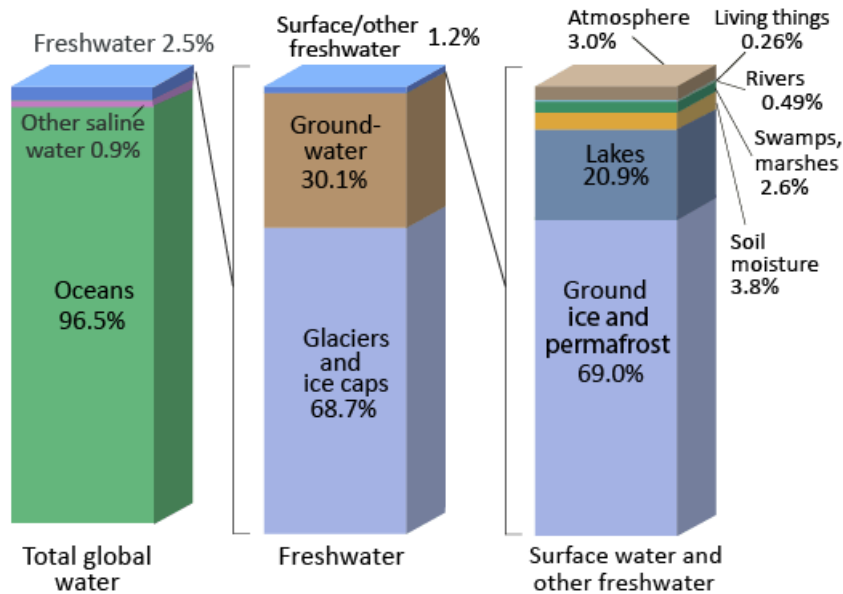
1.3 Διαχείριση υδάτινων πόρων

Το νερό θεωρείται ένα από τα βασικότερα αγαθά πάνω στην γη, καθώς συμβάλει στην ύπαρξη ζωής. Η ύπαρξη του νερού πολλές φορές είναι ένα αναπόσπαστο αγαθό στην διατροφή όλων των οργανισμών, καθώς συμμετέχει και σε κάθε παραγωγική μονάδα, είτε αυτή είναι αγροτική είτε είναι αστική.

Πρέπει να εφαρμοστεί η αειφορία του νερού όταν αυτό πρόκειται για γλυκό νερό μέσα σε ένα περιβάλλον όπου λαμβάνουν χώρα χωρικές και χρονικές κλιματικές αλλαγές. Έτσι ώστε να εξασφαλιστεί η μελλοντική ποιότητα και ισορροπία καθώς και η επιβίωση.

Η συνολική ποσότητα νερού πάνω στην γη είναι περίπου 1.34 δις κυβικά χιλιόμετρα ($\cdot 10^9 km^3$). Από την συνολική ποσότητα αυτού του νερού μόνο το 2.5% πρόκειται για γλυκό νερό. Το νερό που βρίσκεται στα υπόγεια ύδατα πρόκειται για γλυκό νερό όπου αντιστοιχεί ένα ποσοστό 0.76% όπως φαίνεται και στη **Εικόνα 1.2**(<https://water.usgs.gov>).

Where is Earth's Water?



Source: Igor Shiklomanov's chapter "World fresh water resources" in Peter H. Gleick (editor), 1993, *Water in Crisis: A Guide to the World's Fresh Water Resources*.
NOTE: Numbers are rounded, so percent summations may not add to 100.

Εικόνα 1.1 Κατανομή νερού στην γη (<https://water.usgs.gov>)

Water source	Water volume, in cubic miles	Water volume, in cubic kilometers	Percent of freshwater	Percent of total water
Oceans, Seas, & Bays	321,000,000	1,338,000,000	--	96.5
Ice caps, Glaciers, & Permanent Snow	5,773,000	24,064,000	68.7	1.74
Ground water	5,614,000	23,400,000	--	1.69
Fresh	2,526,000	10,530,000	30.1	0.76
Saline	3,088,000	12,870,000	--	0.93
Soil Moisture	3,959	16,500	0.05	0.001
Ground Ice & Permafrost	71,970	300,000	0.86	0.022
Lakes	42,320	176,400	--	0.013
Fresh	21,830	91,000	0.26	0.007
Saline	20,490	85,400	--	0.006
Atmosphere	3,095	12,900	0.04	0.001
Swamp Water	2,752	11,470	0.03	0.0008
Rivers	509	2,120	0.006	0.0002
Biological Water	269	1,120	0.003	0.0001

Source: Igor Shiklomanov's chapter "World fresh water resources" in Peter H. Gleick (editor), 1993, *Water in Crisis: A Guide to the World's Fresh Water Resources* (Oxford University Press, New York).

Εικόνα 1.2: Ποσοστό νερού ανά πηγή (<https://water.usgs.gov>)

Τα υπόγεια ύδατα αποτελούν ένα σημαντικό κομμάτι των διαθέσιμων υδατικών αποθεμάτων. Αποτελούν πηγή ύδρευσης και άρδευσης, προστατεύουν από την ξηρασία καθώς συνδέονται άμεσα με τις επιφανειακές πηγές υδάτων και τον υδρολογικό κύκλο.

Τα υπόγεια ύδατα είναι διαχειρίσιμα καθώς αποτελούν αξιόπιστη πηγή για χρήση σε ύδρευση και άρδευση. Η δε αξιόπιστη πηγή συνδέεται άμεσα με τα χαρακτηριστικά του νερού. Μερικά από τα χαρακτηριστικά του νερού όταν αυτό είναι διαχειρίσιμο είναι η χαμηλή θολερότητα, η απουσία ρύπων, το χαμηλό κόστος εκμετάλλευσης.

Έρευνες έχουν δείξει ότι η ποσοτικοποίηση του γλυκού νερού είναι πολύ δύσκολο να εκτιμηθεί. Λόγω της κλιματικής αλλαγής, καθώς και της μεταβαλλόμενης ζήτησης ως προς την ποσότητα ύδρευσης και της αύξησης του πληθυσμού αναμένεται αύξηση της απαιτούμενης ποσότητας μέσα στα επόμενα χρόνια [Vörösmarty, 2000].

1.4 Υδρολογικός κύκλος

Ο υδρολογικός κύκλος έχει να κάνει με την κυκλοφορία του νερού στην επιφάνεια της Γης. Περιγράφει την πορεία του νερού στα υπόγεια και επιφανειακά ύδατα. Ο υδρολογικός κύκλος είναι απαραίτητος καθώς η ζωή στη Γη εξαρτάται από αυτόν. Στην **Εικόνα 1.3** παρουσιάζεται ο κύκλος του νερού. Το νερό αλλάζει συνεχώς φυσική κατάσταση, από την στερεά μορφή (πάγοι) στην υγρή μορφή (ποτάμια, λίμνες και θάλασσες) και την αέρια κατάσταση (υδρατμοί).



Εικόνα 1.3: Ο υδρολογικός κύκλος (<https://water.usgs.gov/edu/watercycle.html>)

Σύμφωνα με την Γεωλογική Υπηρεσία των ΗΠΑ (USGS) υπάρχουν 16 συνιστώσες του υδρολογικού κύκλου οι οποίες φαίνονται στην **Εικόνα 1.3**.

- Αποθήκευση του νερού σε πάγους και χιόνια
- Αποθήκευση του νερού στην ατμόσφαιρα
- Συμπύκνωση
- Εξάχνωση
- Εξάτμιση
- Πηγή
- Αποθήκευση γλυκού νερού
- Αποθήκευση νερού στην θάλασσα
- Εκφόρτιση υπόγειου νερού
- Αποθήκευση υπόγειου νερού
- Διήθηση υπόγειου νερού
- Απορροή από λιώσιμο χιονιού
- Επιφανειακή απορροή
- Κατακρήμνιση
- Εξατμοδιαπνοή
- Ροή σε υδατορεύματα

Ασφαλώς ο υδρολογικός κύκλος δεν περιγράφεται μόνο εικονικά αλλά και μέσα από μαθηματικές εξισώσεις. Το υδατικό ισοζύγιο που περιγράφει τον υδρολογικό κύκλο φαίνεται στην **Εξίσωση 1.1** [Karatzas, 2004].

$$\Delta S = P - \Delta R - \Delta G - E - T - I$$

Εξίσωση 1.1

Όπου:

ΔS : Αλλαγή του ρυθμού αποθήκευσης επιφανειακών και υπόγειων υδάτων (Storage rate)

P: Κατακρήμνιση-Βροχόπτωση (Precipitation)

ΔR : Ο ρυθμός επιφανειακής ροής (Surface flow rate)

ΔG : Εισροές-Εκροές υδροφορέα

E: Εξάτμιση (Evaporation)

T: Αναπνοή (Transpiration)

I: Ο ρυθμός διήθησης (Infiltration)

1.5 Περιοχή μελέτης

Ο Δούναβης είναι ο δεύτερος μεγαλύτερος ποταμός της Ευρώπης μετά τον Βόλγα με μήκος 2.872 χιλιόμετρα (km). Έχει τις πηγές του στην πόλη Ντοναουεσίνγκεν, όπου βρίσκεται στο Μέλανα Δρυμό της Γερμανίας στη ένωση του ρέματος Μπριγκάχ και του μικρού ποταμού Μπρεγκ **Εικόνα 1.4**.

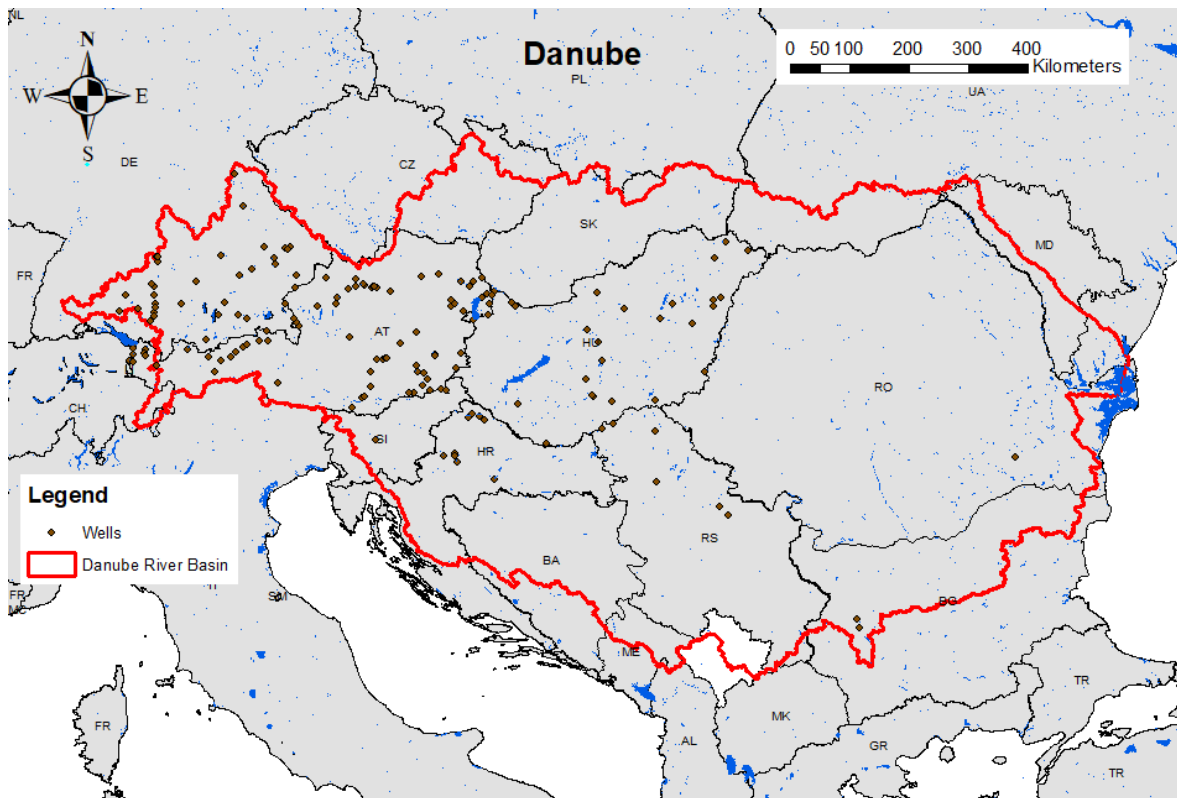


Εικόνα 1.4: Γένεση του Ποταμού Δούναβη
(<https://el.wikipedia.org/wiki/Δούναβης>)

Η συνολική έκταση της λεκάνης απορροής είναι περίπου 817.000 τετραγωνικά χιλιόμετρα (km²) όπου εκτείνεται σε δέκα χώρες καθώς περιλαμβάνει τμήματα 9 ακόμη χωρών Βοσνίας & Ερζεγοβίνη, Τσεχία, Σλοβενία, Μαυροβούνιο, Ελβετία, Ιταλία, Πολωνία, Βόρεια Μακεδονία και Αλβανία.

Η Γερμανία, Αυστρία, Σλοβακία, Ουγγαρία, Κροατία, Σερβία, Ρουμανία, Βουλγαρία, Μολδαβία και Ουκρανία είναι οι δέκα χώρες τις οποίες διασχίζει ο ποταμός Δούναβης **Εικόνα 1.5**.

Αξίζει να σημειωθεί ότι αποτελείται από πολλούς παραπόταμους όπου μερικοί από αυτούς είναι σημαντικοί καθώς είναι πλωτοί, όπου χρησιμοποιούνται για την διέλευση φορτηγίδων και άλλων σκαφών.



Εικόνα 1.5: Όρια περιοχής μελέτης και απεικόνιση πηγαδιών παρατήρησης

Οι εκβολές του ποταμού βρίσκονται στον Εύξεινο Πόντο μέσω του Δέλτα του σε Ρουμανία και Ουκρανία. Το Δέλτα του Δούναβη **Εικόνα 1.6** είναι το έκτο μεγαλύτερο στην Ευρώπη με συνολικό εμβαδόν 4.560 τετραγωνικά χιλιόμετρα (km²).



Εικόνα 1.6 Δέλτα του Δούναβη
(https://el.wikipedia.org/wiki/Δέλτα_Δούναβη)

1.6 Μεταβλητές που χρησιμοποιήθηκαν

Οι μεταβλητές που έχουν χρησιμοποιηθεί στην παρούσα Διπλωματική εργασία αποτελούνται από μια χρονοσειρά δεδομένων. Υπάρχουν διαθέσιμες χρονοσειρές για τις χώρες που διασχίζει ο ποταμός Δούναβης. Σαν δεδομένα εισόδου για την εκπαίδευση του Τεχνητού Νευρωνικού Δικτύου έχουν χρησιμοποιηθεί 10 παράμετροι. Οι δε 10 παράμετροι αποτελούν μετεωρολογικά δεδομένα, συντεταγμένες των πηγαδιών, χρονολογία (Έτος, Μήνας, Ημέρα), εξατμισοδιαπνοή και εξάτμιση. Ενώ οι πραγματικές στάθμες των πηγαδιών που υπήρχαν διαθέσιμες χρησιμοποιούνται για την εκπαίδευση του τεχνητού νευρωνικού δικτύου ως τιμές στόχοι (Targets).

Πιο αναλυτικά οι 10 παράμετροι που έχουν χρησιμοποιηθεί:

1. Έτος
2. Μήνας
3. Ημέρα
4. Βροχόπτωση
5. Μέση θερμοκρασία
6. Δυνητική εξατμισοδιαπνοή
7. Εξάτμιση από υδάτινες επιφάνειες
8. Εξατμισοδιαπνοή από έδαφος και φυτά
9. Συντεταγμένες κατά X
10. Συντεταγμένες κατά Y

1.7 Εφαρμογές Τεχνητών Νευρωνικών Δικτύων στην υδρολογία

Τα Τεχνητά Νευρωνικά Δίκτυα είναι μαθηματικά μοντέλα τα οποία προσπαθούν να προσομοιώσουν την λειτουργία του ανθρώπινου εγκεφάλου. Κατά καιρούς έχουν εφαρμογή σε πολλά προβλήματα που σχετίζονται με την υδρολογία.

Μια έρευνα που έγινε στην στα παράκτια ύδατα της Σιγκαπούρης, για την πρόβλεψη των ποσοτικών χαρακτηριστικών των υδάτινων σωμάτων με την χρήση Τεχνητών Νευρωνικών Δικτύου. Ως δεδομένα εισόδου για την εκπαίδευση του νευρωνικού χρησιμοποιήθηκαν η αλατότητα, η θερμοκρασία, το διαλυμένο οξυγόνο και η χλωροφύλλη. Τα αποτελέσματα έδειξαν ότι η ακρίβεια προσομοίωσης από τον συντελεστή συσχέτισης (R^2) κυμάνθηκε από 0.8-0.9 και ότι ένα νευρωνικό δίκτυο μπορεί να παρέχει προσομοιωμένες τιμές για τις θέσεις όπου είναι τα μετρημένα δεδομένα [Palani et al. 2008].

Σε μια άλλη μελέτη τα τεχνητά νευρωνικά δίκτυα συνδυάστηκαν μαζί με το Kriging για την προσομοίωση του υδραυλικού ύψους σε μια περιοχή της Γερμανίας στη Βαυαρία. Μετά από δοκιμές με διαφορετικές αρχιτεκτονικές δικτύου, τα καλύτερα αποτελέσματα επιτεύχθηκαν με τη χρήση νευρωνικών δικτύων με δύο κρυφά στρώματα με τη χρήση ασαφούς λογικής σε συνδυασμό με το ημιβαριόγραμμα power-law, πετυχαίνοντας έτσι ένα σφάλμα RMSE με τάξη μεγέθους 10^{-2} m. Η μεθοδολογία αυτή μπορεί να χρησιμοποιηθεί με επιτυχία σε υδροφορείς με ασαφή γεωλογικά χαρακτηριστικά [Tarloğlu et al., 2014].

Τα τεχνητά νευρωνικά δίκτυα χρησιμοποιήθηκαν για την πρόβλεψη του υδραυλικού ύψους σε πηγάδια στην περιοχή της Αγίας Χανίων. Σε εκείνη την εργασία, ο αλγόριθμος βελτιστοποίησης σμήνους σωματιδίων (PSO) χρησιμοποιήθηκε για την κατάρτιση ενός πολυστρωματικού ANN. Εξετάστηκαν τρεις παραλλαγές του αλγορίθμου PSO, ο κλασικός με βελτίωση βάρους αδράνειας, PSO με συντελεστές επιτάχυνσης χρόνου (PSO-TVAC) και ολικό βέλτιστο PSO (GLBest-PSO). Η καλύτερη απόδοση επιτεύχθηκε από το GLBest-PSO. Το εκπαιδευμένο ANN χρησιμοποιήθηκε στη συνέχεια για μεσοπρόθεσμη πρόβλεψη του υδραυλικού ύψους, καθώς και για τη μελέτη τριών σεναρίων κλιματικής αλλαγής. Τα διαθέσιμα δεδομένα εισόδου για την περιοχή μελέτης της Αγίας είναι χρονολογικές σειρές μετρήσεων ημερήσιας βροχόπτωσης από δύο μετεωρολογικούς σταθμούς, ημερήσια θερμοκρασία από έναν σταθμό και το υδραυλικό ύψος σε καθημερινή βάση [Tapoglou et al., 2014].

Τα τεχνητά νευρωνικά δίκτυα μπορούν να βρουν εφαρμογή σε πολλά προβλήματα υδρολογικού περιεχομένου. Σε μια έρευνα στην Ινδία, αναπτύχθηκε ένα νευρωνικό δίκτυο για την εκτίμηση των παραμέτρων του υδροφορέα δηλαδή του συντελεστή μεταβιβασιμότητας (Transmissivity) και τον συντελεστή αποθήκευσης από δεδομένα άντλησης σε φρέατα μεγάλης διαμέτρου [Khaled, 2002].

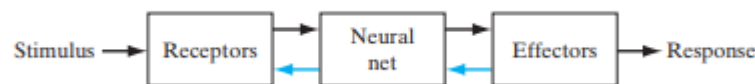
Κεφάλαιο 2

Θεωρητικό Υπόβαθρο

2.1 Ανθρώπινος εγκέφαλος

Το 1987 ο Arbib αντιμετώπισε το ανθρώπινο νευρικό σύστημα σαν ένα σύστημα τριών σταδίων όπως παρουσιάζεται στην **Εικόνα 2.7** [Arbib, 1987].

Αρχικά τα ερεθίσματα (Stimulus) που προέρχονται από το ανθρώπινο σώμα ή από το εξωτερικό περιβάλλον μετατρέπονται από τους υποδοχείς σε ηλεκτρικά σήματα (Impulses). Σαν ηλεκτρικό σήμα νοείται ένα αυτό-αναπαραγόμενο σήμα που μεταδίδεται κατά μήκος της μεμβράνης ενός νεύρου και διεγείρει την απελευθέρωση ενός νευροδιαβιβαστή. Ο οποίος στην συνέχεια διεγείρει μια άλλη ηλεκτρική ώση η δημιουργεί μια αίσθηση στον εγκέφαλο. Στην συνέχεια τα ηλεκτρικά σήματα λειτουργούν σαν όρισμα πληροφοριών στο νευρικό δίκτυο δηλαδή στον εγκέφαλο. Τέλος τα δραστικά κύτταρα μετατρέπουν τα ηλεκτρικά σήματα που προέρχονται από το νευρικό δίκτυο σε αισθητές αποκρίσεις. Οι τελικές αισθητές αποκρίσεις είναι και οι έξοδοι του συστήματος [Arbib, 1987].



Εικόνα 2.7: Σχηματική αναπαράσταση ανθρώπινου νευρικού συστήματος (Simon Haykin, Neural network & learning mechanics)

Σύμφωνα με τον Arbib το κέντρο του συστήματος είναι ο εγκέφαλος ο οποίος αναπαριστάται από το νευρωνικό δίκτυο το οποίο λαμβάνει συνεχώς πληροφορίες, τις επεξεργάζεται και λαμβάνει τις κατάλληλες αποφάσεις. Στην **Εικόνα 2.7** έχουμε την σχηματική αναπαράσταση του ανθρώπινου νευρικού συστήματος το οποίο αποτελείται από δύο ομάδες σύμφωνα με τα βέλη [Arbib, 1987].

1^η ομάδα: Υποδοχείς (Receptors) → Νευρωνικό δίκτυο (Neural net) → Δραστικά κύτταρα (Effectors).

Η 1^η ομάδα ονομάζεται πρόσθια τροφοδότηση του συστήματος, διότι η μετάδοση των σημάτων γίνεται προς τα εμπρός.

2^η ομάδα: Δραστικά κύτταρα (Effectors) → Νευρωνικό δίκτυο (Neural net) → Υποδοχείς (Receptors).

Η 2^η ομάδα ονομάζεται ανάδραση (Feedback), διότι γίνεται ανατροφοδότηση της εξόδου ενός συστήματος στην είσοδο του.

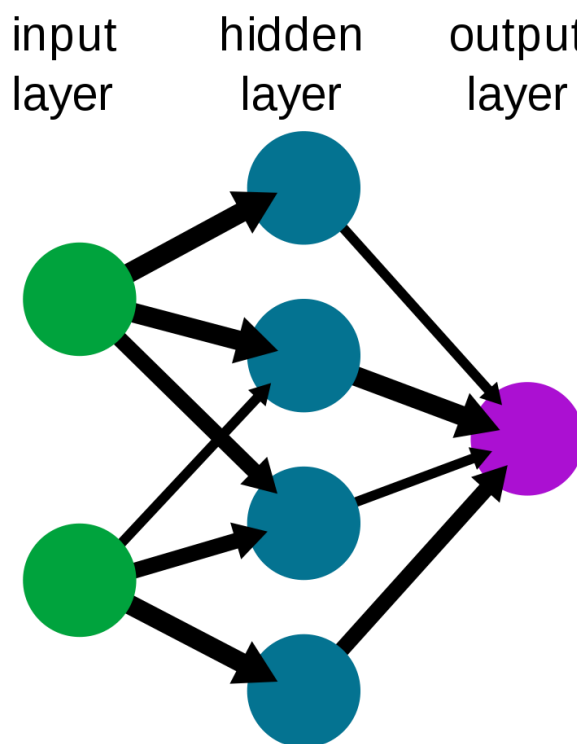
2.2 Τεχνητά Νευρωνικά Δίκτυα (ΤΝΔ)

Τα τεχνητά νευρωνικά δίκτυα αναπτύχθηκαν το 1943 από τους McCulloch & Pitts. Οι οποίοι προσπάθησαν να προσομοιώσουν και να κατανοήσουν την λειτουργία του ανθρώπινου εγκεφάλου μέσα από αυτά τα μοντέλα [McCulloch & Pitts, 1943].

Ένα Τεχνητό νευρωνικό δίκτυο αποτελείται από μια σειρά απλών μονάδων επεξεργασίας. Τα οποία έχουν την ικανότητα να επικοινωνούν, στέλνοντας σήματα μεταξύ τους, καθώς να αντιπροσωπεύουν τόσο γραμμικές όσο και μη γραμμικές σχέσεις [Rumelhart et al., 1986].

Στην **Εικόνα 2.8** φαίνεται ένα τυπικό νευρωνικό δίκτυο. Αποτελείται από διαφορετικά επίπεδα και συγκεκριμένα 3. Σε αυτά τα επίπεδα γίνεται η επεξεργασία πληροφοριών. Τα οποία επίπεδα αποτελούνται από ξεχωριστά στοιχεία που ονομάζονται κόμβοι ή νευρώνες.

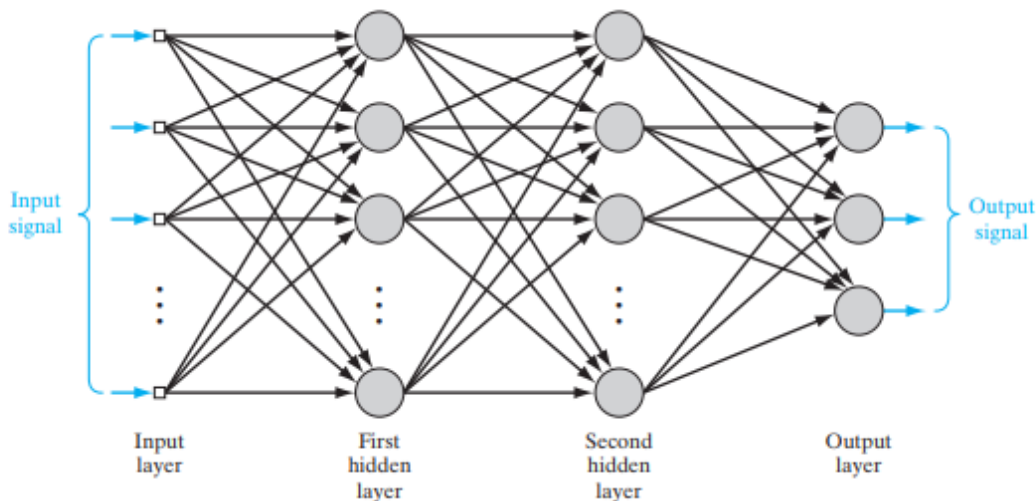
A simple neural network



Εικόνα 2.8: Τεχνητό Νευρωνικό Δίκτυο
(https://el.wikipedia.org/wiki/Νευρωνικό_δίκτυο)

Οι τρεις τύποι κόμβων είναι οι κόμβοι εισόδου στο επίπεδο εισόδου (Input Layer), οι κόμβοι εξόδου στο επίπεδο εξόδου (Output Layer) και οι κρυφοί κόμβοι στο κρυφό επίπεδο (Hidden Layer). Από αυτούς υπολογιστικοί κόμβοι είναι οι κόμβοι εξόδου και οι κρυφοί ενώ στους κόμβους εισόδου δεν γίνονται υπολογισμοί. Αρχικά οι υπολογιστικοί κόμβοι πολλαπλασιάζουν τις τιμές εισόδου με τα αντίστοιχα συναπτικά βάρη. Το άθροισμα του γινομένου χρησιμοποιείται ως είσοδο για την λειτουργία της συνάρτησης ενεργοποίησης. Τέλος οι προκύπτουσες τιμές είναι οι τιμές εξόδου για τις τρέχουσες τιμές εισόδου και τα αντίστοιχα συναπτικά βάρη για την δεδομένη χρονική στιγμή [Haykin, 1931].

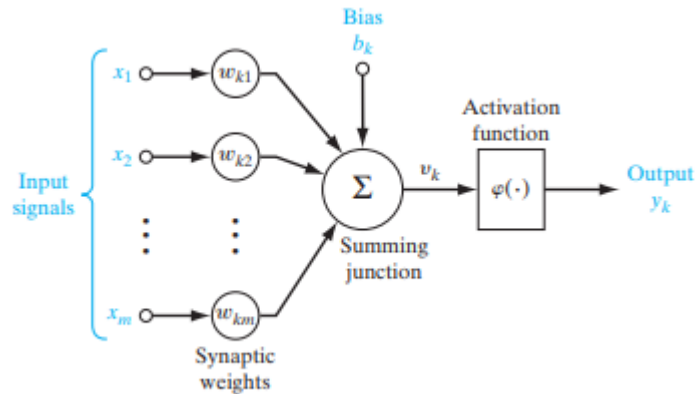
Αξίζει να σημειωθεί ότι υπάρχουν ΤΝΔ με περισσότερα από ένα κρυφό επίπεδο. Αυτό έχει σαν αποτέλεσμα, οι τιμές έπειτα από το άθροισμα των γινομένων να τροφοδοτούνται στο επόμενο κρυφό επίπεδο και να μην αποτελούν την απευθείας έξοδο του νευρωνικού δικτύου [Βλαχάβας, 2002]. Στην **Εικόνα 2.9** φαίνεται ένα Τεχνητό Νευρωνικό Δίκτυο με δύο κρυφά επίπεδα [Haykin, 1931].



Εικόνα 2.9: Τεχνητό νευρωνικό δίκτυο με δύο κρυφά επίπεδα (Simon Haykin, Neural network & learning mechanics)

Ο κάθε νευρώνας είναι μια μονάδα επεξεργασίας πληροφορίας. Στην **Εικόνα 2.10** παρουσιάζεται ένα μη γραμμικό μοντέλο του νευρώνα k . Το οποίο αποτελείται από τις τιμές εισόδου (Input signals) x_1, x_2, \dots, x_m με τα αντίστοιχα συναπτικά βάρη (Synaptic weights) του νευρώνα k $w_{k1}, w_{k2}, \dots, w_{km}$. Ο πρώτος δείκτης στο συναπτικό βάρος αναφέρεται στον εν λόγω νευρώνα και ο δεύτερος στην τιμή εισόδου της σύναψης στην οποία αντιστοιχεί το βάρος. Ο κόμβος άθροισης (Summing function) για την άθροιση των τιμών εισόδου, σταθμισμένων από τα αντίστοιχα συναπτικά βάρη. Το u_k συμβολίζει την έξοδο του γραμμικού συνδυαστή με βάση τις τιμές εισόδου. Το b_k είναι η πόλωση (Bias) του νευρώνα k δηλαδή ο σταθερός όρος της συνάρτησης ο οποίος δεν εξαρτάται από τις τιμές εισόδου. Η πόλωση b_k μπορεί να είναι θετική ή αρνητική, αντίστοιχα με το πρόσημο της αυξάνει ή μειώνει την δικτυακή διέγερση της συνάρτησης ενεργοποίησης. Με ϕ_k συμβολίζεται η συνάρτηση ενεργοποίησης (Activation function).

Σχετικά με την συνάρτηση ενεργοποίησης αναφέρεται και ως συνάρτηση περιορισμού (Squashing function). Έχει την ιδιότητα να περιορίζει το επιτρεπτό εύρος πλάτους για το σήμα εξόδου στο μοναδιαίο κλειστό διάστημα $[0,1]$ ή $[-1,1]$. Τέλος, έχουμε το σήμα εξόδου (Output) συμβολίζεται με y_k [Haykin, 1931].



Εικόνα 2.10: Μη γραμμικό μοντέλο νευρώνα k (Simon Haykin, Neural network & learning mechanics)

Οι εξισώσεις που μπορούν να περιγράψουν το μη γραμμικό μοντέλο του νευρώνα k είναι οι **Εξισώσεις 2.1 & 2.2** [Haykin, 1931].

$$u_k = \sum_{i=1}^m w_{ki} \cdot x_i \quad \text{Εξίσωση 2.1}$$

$$y_k = \varphi \left(\sum_{i=1}^m w_{ki} \cdot x_i + b_k \right) \quad \text{Εξίσωση 2.2}$$

2.2.1 Multi-layer Perceptrons

Το Perceptron εμφανίστηκε και αποδείχθηκε για πρώτη φορά το 1962 [Rosenblatt, 1962].

Στόχος του Perceptron είναι να μάθει τον μετασχηματισμό **Εξίσωση 2.3**.

$$d: \{1,1\}^N \rightarrow \{1,1\} \quad \text{Εξίσωση 2.3}$$

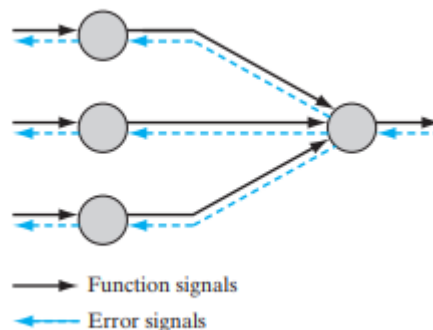
Χρησιμοποιώντας δείγματα μάθησης με είσοδο x και έξοδο $y=d(x)$ καθώς και να επιλέγει τα κατάλληλα τροποποιημένα βάρη σε περίπτωση που ισχύει $y \neq d(x)$ [Kröse, 1996].

Ένα πολυστρωματικό δίκτυο Perceptron (Multi-layer Perceptron) αποτελείται από έναν ή περισσότερους κόμβους εισόδου, έναν ή περισσότερους κρυφούς κόμβους και έναν ή περισσότερους κόμβους εξόδου όπως απεικονίζεται και στην **Εικόνα 2.9**.

Σε ένα Perceptron όλοι οι νευρώνες ενός επιπέδου συνδέονται με όλους του νευρώνες (κόμβους) του προηγούμενου επιπέδου. Η ροή του σήματος μέσα στο δίκτυο έχει την εξής κατεύθυνση, να προχωρά προς τα εμπρός δηλαδή από τα αριστερά προς τα δεξιά από επίπεδο προς επίπεδο. Οι κατευθύνσεις ροής σε ένα Perceptron αποτελούνται από τα λειτουργικά σήματα και τα σήματα σφάλματος. Τα λειτουργικά σήματα είναι εκείνα της εισόδου όπου διαδίδονται προς τα εμπρός σε όλο το δίκτυο (κρυφοί κόμβοι, κόμβοι εξόδου) έως ότου φτάσουν στην έξοδο του δικτύου. Τα σήματα σφάλματος προκύπτουν από τους νευρώνες εξόδους και διαδίδονται προς τα πίσω από επίπεδο προς επίπεδο. Λόγω του υπολογισμού του από κάθε νευρώνα του δικτύου απαιτείται η χρήση μια εξαρτώμενης από το σφάλμα συνάρτησης. Τα σήματα σφάλματος ορίζονται ως η διαφορά μεταξύ της επιθυμητής και της πραγματικής απόκρισης. Όπως έχουμε αναφέρει υπολογιστικοί κόμβοι στο δίκτυο είναι οι κρυφοί κόμβοι και οι κόμβοι εξόδου. Άρα κάθε κρυφός νευρώνας ή νευρώνας εξόδου εκτελεί δύο υπολογισμούς [Haykin, 1931]:

1. Υπολογισμός του λειτουργικού σήματος στην έξοδο κάθε νευρώνα, ο οποίος εκφράζεται ως μια συνεχής μη γραμμική συνάρτηση του σήματος εισόδου και των συναπτικών βαρών που σχετίζονται με τον κάθε νευρώνα
2. Υπολογισμός των σημάτων σφάλματος, δηλαδή μιας εκτίμησης του διανύσματος των κλίσεων σε σχέση με τα βάρη που είναι συνδεδεμένα στις εισόδους ενός νευρώνα

Στην **Εικόνα 2.11** απεικονίζονται οι κατευθύνσεις ροής των δύο βασικών σημάτων, Λειτουργικά σήματα (Function signals) και Σήματα σφάλματος (Error signals).



Εικόνα 2.11: Κατευθύνσεις ροής των δύο βασικών σημάτων ενός Perceptron πολλών επιπέδων
 (Simon Haykin, Neural network & learning mechanics)

2.3 Συναρτήσεις ενεργοποίησης

Οι συναρτήσεις ενεργοποίησης φ_k χρησιμοποιούν την συνολική είσοδο $s_k(t)$ και την τρέχουσα τιμή ενεργοποίησης $y_k(t)$. Παράγουν έτσι μια νέα τιμή ενεργοποίησης της μονάδας k . Η **Εξίσωση 2.4** αντιπροσωπεύει την νέα τιμή ενεργοποίησης [Kröse, 1996].

$$y_k(t+1) = \varphi_k(y_k(t), s_k(t)) \quad \text{Εξίσωση 2.4}$$

Οι πιο συχνά χρησιμοποιούμενες συναρτήσεις ενεργοποίησης $\varphi_k(u)$, είναι η συνάρτηση κατωφλίου, η ημι-γραμμική συνάρτηση και η σιγμοειδής συνάρτηση. Οι οποίες ορίζουν την έξοδο ενός νευρώνα βάσει του τοπικού πεδίου u .

2.3.1 Συνάρτηση κατωφλίου

Η συνάρτηση κατωφλίου (Threshold Function) καθώς και οι εξισώσεις που την περιγράφει προκύπτουν από το έργο των McCulloch-Pitts [McCulloch & Pitts, 1943].

Στην **Εξίσωση 2.5** περιγράφεται ο τύπος συνάρτησης για την συνάρτηση κατωφλίου [Haykin, 1931].

$$\varphi(u) = \begin{cases} 1 & \text{εάν } u \geq 0 \\ 0 & \text{εάν } u < 0 \end{cases} \quad \text{Εξίσωση 2.5}$$

Η έξοδος του νευρώνα k που χρησιμοποιεί η συνάρτηση κατωφλίου περιγράφεται από την **Εξίσωση 2.6** [Haykin, 1931].

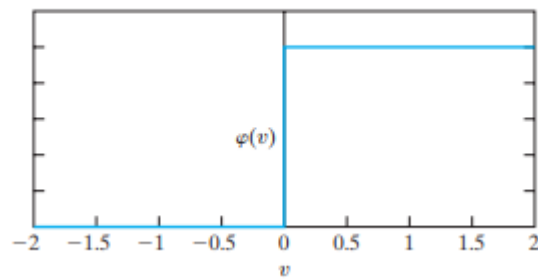
$$y_k = \begin{cases} 1 & \text{εάν } u_k \geq 0 \\ 0 & \text{εάν } u_k < 0 \end{cases} \quad \text{Εξίσωση 2.6}$$

Η έξοδος ενός νευρώνα k (y_k) για τοπικό πεδίο μη αρνητικό λαμβάνει τιμή 1 και για αρνητικό τιμή 0.

Το τοπικό πεδίο u_k του νευρώνα περιγράφεται από την **Εξίσωση 2.7** [Haykin, 1931].

$$u_k = \sum_{j=1}^m w_{jk} \cdot x_j + b_k \quad \text{Εξίσωση 2.7}$$

Στην **Εικόνα 2.12** γίνεται η γραφική αναπαράσταση της συνάρτησης κατωφλίου.



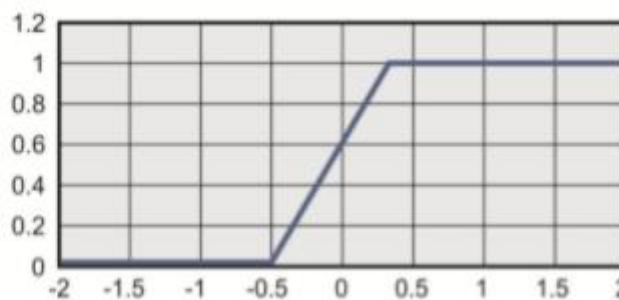
Εικόνα 2.12: Συνάρτηση κατωφλίου
(Simon Haykin, *Neural network & learning mechanics*)

2.3.2 Ήμι-γραμμική συνάρτηση

Η Ήμι-γραμμική συνάρτηση περιγράφεται από την **Εξίσωση 2.8** και η γραφική απεικόνισή της στην **Εικόνα 2.13**.

$$\varphi(u) = \begin{cases} 1 & \text{εάν } u \geq +\frac{1}{2} \\ u + \frac{1}{2} & \text{εάν } +\frac{1}{2} > u > -\frac{1}{2} \\ 0 & \text{εάν } u \leq -\frac{1}{2} \end{cases}$$

Εξίσωση 2.8



Εικόνα 2.13: Ήμι-γραμμική συνάρτηση
(Ben Kröse & Patrick van der Smag, *An Introduction to Neural Networks*)

2.3.3 Σιγμοειδής συνάρτηση

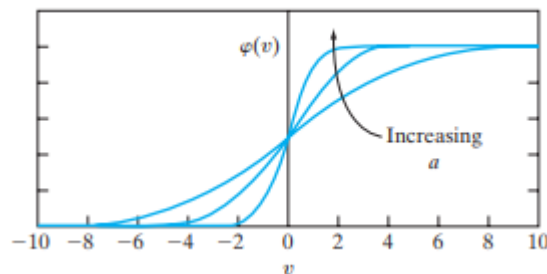
Η σιγμοειδής συνάρτηση είναι η πιο κοινή χρησιμοποιούμενη μορφή συνάρτηση ενεργοποίησης για την κατασκευή νευρωνικών δικτύων. Αυτό οφείλεται στο γεγονός ότι υπάρχει ισορροπία μεταξύ γραμμικής και μη γραμμικής συμπεριφοράς [Haykin, 1931].

Η λογιστική συνάρτηση είναι ένα παράδειγμα σιγμοειδούς συνάρτησης. Η λογιστική συνάρτηση περιγράφεται στην **Εξίσωση 2.9** και η γραφική της απεικόνιση στην **Εικόνα 2.14** [Haykin, 1931].

$$\varphi(u) = \frac{1}{1 + e^{(-a \cdot v)}}$$

Εξίσωση 2.9

Όπου a είναι η παράμετρος της κλίσης για την σιγμοειδή συνάρτηση. Όπως παρατηρούμε στην **Εικόνα 2.6** για μεταβαλλόμενο a έχουμε και διαφορετικές σιγμοειδής συναρτήσεις [Haykin, 1931].



Εικόνα 2.14: Σιγμοειδής συνάρτηση για μεταβαλλόμενη παράμετρο a (Simon Haykin, Neural network & learning mechanics)

2.4 Αλγόριθμοι εκπαίδευσης

2.4.1 Levenberg-Marquardt

Ο αλγόριθμος Levenberg-Marquardt (LM) εμφανίστηκε για πρώτη φορά το 1963 και προσπάθησε να εκτιμήσει την τιμή για τα ελάχιστα τετράγωνα μη γραμμικών παραμέτρων μέσα από δύο προσεγγίσεις. Ο αλγόριθμος μπορεί να επεκταθεί ως σειρά του Taylor και να κάνει τις απαραίτητες διορθώσεις στις παραμέτρους που υπολογίζονται, καθώς και ως μεθόδου της κλίσης (Gradient Descent) μετά από διαδοχικές επαναλήψεις [Marquardt, 1963].

Με την χρήση του αλγόριθμου LM το δίκτυο ενημερώνει τις τιμές βάρους και μεροληψίας (bias) σύμφωνα με την βελτιστοποίηση του Levenberg-Marquardt. Είναι ο ταχύτερος αλγόριθμος backpropagation και συνήθως επιλέγεται πρώτος για την εκπαίδευση νευρωνικών δικτύων αν και απαιτεί περισσότερη μνήμη από άλλους αλγόριθμους.

Η εκπαίδευση των νευρωνικών δικτύων σύμφωνα με τον αλγόριθμο LM πραγματοποιείται από τις παραμέτρους εκπαίδευσης του, που εμφανίζονται στον **Πίνακα 1** με τις προεπιλεγμένες τιμές (<https://www.mathworks.com> on search about trainlm).

Πίνακας 1: Παράμετροι αλγορίθμου LM (<https://www.mathworks.com>)

net.trainParam.epochs	1000	Maximum number of epochs to train
net.trainParam.goal	0	Performance goal
net.trainParam.max_fail	6	Maximum validation failures
net.trainParam.min_grad	1e-7	Minimum performance gradient
net.trainParam.mu	0.001	Initial mu
net.trainParam.mu_dec	0.1	mu decrease factor
net.trainParam.mu_inc	10	mu increase factor
net.trainParam.mu_max	1e10	Maximum mu
net.trainParam.show	25	Epochs between displays (NaN for no displays)
net.trainParam.showCommandLine	false	Generate command-line output
net.trainParam.showWindow	true	Show training GUI
net.trainParam.time	inf	Maximum time to train in seconds

Διανύσματα επικύρωσης χρησιμοποιούνται για να σταματήσουν την εκπαίδευση εάν η απόδοση του δικτύου δεν βελτιώνεται ή παραμένει ίδια για τις εποχές max_fail στην σειρά.

Η εκπαίδευση νευρωνικών με τον αλγόριθμο LM χρησιμοποιεί τον Jacobian για τους υπολογισμούς το οποίο υποθέτει ότι η απόδοση του δικτύου είναι ένας μέσος όρος ή άθροισμα των τετραγωνικών σφαλμάτων.

Όπως και οι μέθοδοι quasi-Newton, έτσι και ο αλγόριθμος LM σχεδιάστηκε για να προσεγγίσει την ταχύτητα εκπαίδευσης δεύτερης τάξης χωρίς να χρειάζεται να υπολογίσει τον πίνακα Hessian. Όταν η συνάρτηση απόδοσης έχει τη μορφή ενός αθροίσματος τετραγώνων όπως συμβαίνει κατά την εκπαίδευση feedforward δικτύων, τότε ο πίνακας Hessian μπορεί να προσεγγιστεί από την **Εξίσωση 2.10** και η κλίση από την **Εξίσωση 2.11**.

$$H = J^T \cdot J \quad \text{Εξίσωση 2.10}$$

$$g = J^T \cdot e \quad \text{Εξίσωση 2.11}$$

Όπου J είναι ο πίνακας Jacobian που περιέχει τις πρώτες παραγώγους των σφαλμάτων του δικτύου σε σχέση με τα βάρη και τους σταθερούς όρο (biases) και e είναι ο φορέας των σφαλμάτων του δικτύου. Ο πίνακας Jacobian μπορεί να υπολογιστεί μέσω μιας τυπικής τεχνικής backpropagation [Hagan and Menhaj, 1994] που είναι λιγότερο πολύπλοκη από τον υπολογισμό του πίνακα Hessian.

Ο αλγόριθμος LM χρησιμοποιεί αυτήν την προσέγγιση στο Hessian matrix στην ακόλουθη **Εξίσωση 2.12** που αντιπροσωπεύει μια νέα έκδοση τύπου Newton.

$$x_{k+1} = x_k - [J^T \cdot J + \mu \cdot I]^{-1} \cdot J^T \cdot e \quad \text{Εξίσωση 2.12}$$

Όταν η παράμετρος μ είναι μηδέν τότε έχουμε την μέθοδο του Νεύτωνα χρησιμοποιώντας τον κατά προσέγγιση Hessian matrix. Όταν το μ είναι μεγάλο αυτό γίνεται κλιμακωτή κάθοδος (Gradient descent) ή με μικρό μέγεθος βήματος. Η μέθοδος του Νεύτωνα είναι ταχύτερη και ακριβέστερη κοντά σε ένα ελάχιστο σφάλμα οπότε ο στόχος είναι να στραφεί προς τη μέθοδο του Νεύτωνα όσο το δυνατόν γρηγορότερα. Με αυτό τον τρόπο το μ μειώνεται μετά από κάθε επιτυχημένο βήμα (μείωση της αντικειμενικής συνάρτησης) και αυξάνεται μόνο όταν ένα δοκιμαστικό βήμα θα αυξήσει την αντικειμενική συνάρτηση. Με αυτόν τον τρόπο η αντικειμενική συνάρτηση μειώνεται πάντα σε κάθε επανάληψη (Iteration) του αλγορίθμου.

Η εκπαίδευση του νευρωνικού δικτύου με την χρήση του αλγορίθμου LM σταματάει όταν:

- Επιτυγχάνεται ο μέγιστος αριθμός των εποχών
- Γίνεται υπέρβαση του μέγιστου χρονικού ορίου
- Η απόδοση ελαχιστοποιείται στο στόχο
- Η κλίση της απόδοσης πέφτει κάτω από το `min_grad`
- `mu` υπερβαίνει το `mu_max`
- Το `max_fail` έχει αυξηθεί περισσότερο από την τελευταία φορά που μειώθηκε (όταν χρησιμοποιείται η επικύρωση (Validation))

2.4.2 Bayesian Regularization Backpropagation

Αποτελεί μια τεχνική εκπαίδευσης, η οποία ενημερώνει τις τιμές από τα βάρη και τα συστηματικά σφάλματα με την βοήθεια της μεθόδου βελτιστοποίησης Levenberg-Marquardt. Σκοπός του είναι να ελαχιστοποιήσει τον συνδυασμό τετραγωνικών σφαλμάτων και βαρών και στην συνέχεια να καθορίσει τον σωστό συνδυασμό τους έτσι ώστε να παραχθεί ένα δίκτυο το οποίου η γενίκευση να είναι καλύτερη (Βλέπε [MacKay,1992] & [Foresse and Hagan, 1997] για λεπτομερείς συζητήσεις σχετικά με τον αλγόριθμο Bayesian Regularization) (www.mathworks.com). Ο αλγόριθμος BR μπορεί να εκπαιδεύσει οποιοδήποτε δίκτυο, εφόσον οι συναρτήσεις βάρους, εισόδου και μεταφοράς έχουν παράγωγες συναρτήσεις. Η εκπαίδευση των νευρωνικών δικτύων σύμφωνα με τον αλγόριθμο BR πραγματοποιείται από τις παραμέτρους εκπαίδευσης του που εμφανίζονται στον **Πίνακα 2** με τις προεπιλεγμένες τιμές (<https://www.mathworks.com> on search about trainbr).

Πίνακας 2: Παράμετροι αλγορίθμου BR (<https://www.mathworks.com>)

net.trainParam.epochs	1000	Maximum number of epochs to train
net.trainParam.goal	0	Performance goal
net.trainParam.mu	0.005	Marquardt adjustment parameter
net.trainParam.mu_dec	0.1	Decrease factor for mu
net.trainParam.mu_inc	10	Increase factor for mu
net.trainParam.mu_max	1e10	Maximum value for mu
net.trainParam.max_fail	inf	Maximum validation failures
net.trainParam.min_grad	1e-7	Minimum performance gradient
net.trainParam.show	25	Epochs between displays (NaN for no displays)
net.trainParam.showCommandLine	false	Generate command-line output
net.trainParam.showWindow	true	Show training GUI
net.trainParam.time	inf	Maximum time to train in seconds

Η εκπαίδευση του νευρωνικού δικτύου με την χρήση του αλγορίθμου BR σταματάει όταν:

- Επιτυγχάνεται ο μέγιστος αριθμός των εποχών
- Γίνεται υπέρβαση του μέγιστου χρονικού ορίου
- Η απόδοση (Performance) ελαχιστοποιείται στο στόχο
- Η κλίση της απόδοσης πέφτει κάτω από το min_grad
- mu υπερβαίνει το mu_max

2.5 Κριτήρια αξιολόγησης

Η αποτελεσματικότητα κάθε δικτύου ορίζεται από δύο διαφορετικούς δείκτες. Οι δύο αυτοί δείκτες είναι το RMSE (Root Mean Square Error) και R που φαίνονται αντίστοιχα στις **Εξισώσεις 2.13 & 2.14**. Το RMSE υποδεικνύει την απόκλιση μεταξύ των παρατηρούμενων και υπολογισμένων τιμών. Ο συντελεστής συσχέτισης R δηλώνει την αναλογία της διακύμανσης στην εξαρτημένη μεταβλητή που είναι προβλέψιμη από την ανεξάρτητη μεταβλητή. Με βάση αυτούς τους δύο δείκτες θα αξιολογηθεί η αποτελεσματικότητα του κάθε δικτύου και η ικανότητά του να κάνει ακριβείς προβλέψεις.

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (O_t - S_t)^2}{N}}$$

Εξίσωση 2.13

$$R = \frac{(\sum (O_t - \bar{O}) \cdot (S_t - \bar{S}))}{\sqrt{\sum (O_t - \bar{O})^2 \cdot \sum (S_t - \bar{S})^2}}$$

Εξίσωση 2.14

Όπου S_t είναι η προσομοιωμένη τιμή, O_t είναι η παρατηρημένη τιμή, \bar{O} είναι ο μέσος όρος των παρατηρούμενων τιμών και \bar{S} είναι ο μέσος όρος των προσομοιωμένων τιμών..

Τέλος υπολογίστηκαν 3 επιπλέον δείκτες:

1. NSE (Nash–Sutcliffe model efficiency coefficient)

Ο συντελεστής απόδοσης NSE χρησιμοποιείται ευρέως για προβλήματα της υδρολογίας. Πιο συγκεκριμένα χρησιμοποιείται για την εκτίμηση της προγνωστικής ισχύος των υδρολογικών μοντέλων και περιγράφεται από την **Εξίσωση 2.15** [Nash & Sutcliffe, 1970]:

$$NSE = 1 - \frac{\sum_{t=1}^n (O_t - S_t)^2}{\sum_{t=1}^n (O_t - \bar{O})^2} \quad \text{Εξίσωση 2.15}$$

Όπου S_t είναι η προσομοιωμένη τιμή, O_t είναι η παρατηρούμενη τιμή και \bar{O} είναι ο μέσος όρος των παρατηρούμενων τιμών.

Ο συντελεστής απόδοσης NSE κυμαίνεται από $-\infty$ έως 1:

- Όταν $NSE=1$ αντιστοιχεί στην τέλεια προσομοίωση του μοντέλου στα παρατηρούμενα δεδομένα
- Όταν $NSE=0$ υποδεικνύει ότι οι προβλέψεις του μοντέλου είναι τόσο ακριβείς όσο ο μέσος όρος των παρατηρούμενων δεδομένων
- Όταν $NSE<0$ σημαίνει ότι ο παρατηρούμενος μέσος όρος είναι καλύτερος προγνωστικός παράγοντας από του μοντέλου

Όσο πιο κοντά είναι η απόδοση του μοντέλου στην μονάδα τόσο πιο ακριβές είναι το μοντέλο. Ένα μοντέλο θεωρείται ικανοποιητικό για τιμές τους συντελεστή απόδοσης $0,5 < NSE < 0,65$ [Moriassi et al., 2007] και [Ritter & Munoz-Carpena, 2013].

2. Bias

Η μέση απόκλιση μεταξύ προσομοιωμένης και πραγματικής τιμής δίνεται από την **Εξίσωση 2.16**:

$$Bias = \frac{1}{n} * \sum_{t=1}^n (S_t - O_t) \quad \text{Εξίσωση 2.16}$$

Η τιμή του Bias μπορεί να είναι είτε θετική είτε αρνητική. Αν είναι θετική σημαίνει πως το μοντέλο υπερεκτιμάει τις παρατηρούμενες τιμές ενώ αν είναι αρνητική σημαίνει πως τις υποεκτιμάει.

3. MAE (Mean Absolute Error)

Το μέσο απόλυτο σφάλμα μεταξύ προσομοιωμένης και πραγματικής τιμής δίνεται από την **Εξίσωση 2.17**. Το μέσο απόλυτο σφάλμα είναι απλώς η μέση απόλυτη κατακόρυφη ή οριζόντια απόσταση μεταξύ κάθε σημείου σε ένα διάγραμμα σκέδασης και της γραμμής $Y=X$ δηλαδή είναι η μέση απόλυτη διαφορά μεταξύ X και Y .

$$MAE = \frac{1}{n} \sum_{t=1}^n |S_t - O_t|$$

Εξίσωση 2.17

Όσο μικρότερο είναι το RMSE και όσο το R τείνει στην μονάδα τόσο πιο ακριβής θα είναι η πρόβλεψη. Το βέλτιστο ζευγάρι που υπάρχει μεταξύ των δύο δεικτών, το οποίο είναι απίθανο να συμβεί είναι να έχει $RMSE=0$ & $R=1$.

Σε περίπτωση που συνέβαινε κάτι τέτοιο το μοντέλο θα ήταν σε θέση να περιγράψει με απόλυτη ακρίβεια το φυσικό σύστημα.

2.6 Πλεονεκτήματα ΤΝΔ

Η υπολογιστική ισχύς ενός Τεχνητού Νευρωνικού Δικτύου οφείλεται στην παράλληλη, κατανεμημένη δομή του και στην ικανότητα να γενικεύει. Η γενίκευση του ΤΝΔ έχει να κάνει με τα λογικά αποτελέσματα εξόδου με βάση τις τιμές εισόδου που έχουν δοθεί τις οποίες δεν τις έχει συναντήσει ξανά το Τεχνητό Νευρωνικό Δίκτυο [Haykin, 1931].

Τα νευρωνικά δίκτυα αποτελούνται από μία ομάδα χρήσιμων ιδιοτήτων και δυνατοτήτων συγκεκριμένα 9 πλεονεκτημάτων, όπου περιλαμβάνεται [Haykin, 1931] :

1. Μη γραμμικότητα
2. Αντιστοίχιση Εισόδου-Εξόδου
3. Προσαρμοστικότητα
4. Ενδεικτική απόκριση
5. Πληροφορία σχετική με το περιεχόμενο
6. Ανοχή σε βλάβες
7. Δυνατότητα υλοποίησης σε VLSI
8. Ομοιομορφία ανάλυσης και σχεδίασης
9. Αναλογία με τη νευροφυσιολογία του εγκεφάλου

Παρακάτω θα αναλυθούν τα 9 πλεονεκτήματα των τεχνητών νευρωνικών δικτύων.

Μη γραμμικότητα: Ένα νευρωνικό δίκτυο μπορεί να είναι γραμμικό ή μη γραμμικό. Αποτελείται βέβαια από την σύνδεση μη γραμμικών νευρώνων, άρα είναι μη γραμμικό. Η μη γραμμικότητα του δικτύου είναι κατανεμημένη σε όλη την έκταση του δικτύου. Η μη γραμμικότητα είναι η σημαντικότερη ιδιότητα εάν ο φυσικός μηχανισμός παραγωγής του σήματος εισόδου είναι μη γραμμικός [Haykin, 1931].

Αντιστοίχιση Εισόδου-Εξόδου: Αυτή η ιδιότητα έχει σημαντικό χαρακτήρα καθώς κατά την επιβλεπόμενη εκπαίδευση, τα συναπτικά βάρη ενός νευρωνικού δικτύου τροποποιούνται, με την βοήθεια εφαρμογής ενός συνόλου χαρακτηρισμένων παραδειγμάτων εκπαίδευσης. Σε κάθε ένα παράδειγμα εκπαίδευσης αντιστοιχεί ένα μοναδικό σήμα εισόδου και μια αντίστοιχη επιθυμητή απόκριση. Κατά την διάρκεια της λειτουργίας το νευρωνικό δίκτυο παρουσιάζεται με ένα δείγμα επιλεγμένο τυχαία από το σύνολο των δειγμάτων και τα συναπτικά βάρη τροποποιούνται. Αυτό συμβαίνει για να ελαχιστοποιηθεί η διαφορά μεταξύ επιθυμητής και πραγματικής απόκρισης του δικτύου, όπου παράχθηκαν από το σήμα εισόδου, με την βοήθεια ενός στατιστικού κριτηρίου. Η εκπαίδευση του δικτύου επαναλαμβάνεται για πολλά παραδείγματα του συνόλου εκπαίδευσης ωστότου το δίκτυο φτάσει στην ιδανική κατάσταση, όπου δεν θα υπάρχουν σημαντικές διαφορές στα συναπτικά βάρη [Haykin, 1931]. Αυτή η προσέγγιση έχει εφαρμοστεί στον κλάδο της στατιστικής στη μελέτη *μη παραμετρικού στατιστικού συμπερασμού*, που ασχολείται με μη βασιζόμενες σε μοντέλα εκτιμήσεις ή από βιολογικής σκοπιάς, τη μάθηση από την κατάσταση *tabula rasa* [Geman et al. 1992].

Προσαρμοστικότητα: Μια εγγενή δυνατότητα που απαρτίζουν τα νευρωνικά δίκτυα είναι να προσαρμόζουν τα συναπτικά βάρη ανάλογα με τις μεταβολές που γίνονται στο περιβάλλον τους. Σε ένα νευρωνικό δίκτυο το οποίο είναι εκπαιδευμένο να λειτουργεί σε ένα συγκεκριμένο περιβάλλον, μπορεί να επανεκπαιδευτεί έτσι ώστε να μπορεί να διαχειριστεί ήσσονος σημασίας μεταβολές στο περιβάλλον λειτουργίας του. Επιπλέον σε περίπτωση που λειτουργεί σε ένα μη στατικό περιβάλλον (μεταβολή στοιχείων με το χρόνο) το νευρωνικό δίκτυο μπορεί να σχεδιαστεί έτσι ώστε να αλλάζει τα συναπτικά βάρη σε πραγματικό χρόνο. Η φυσική αρχιτεκτονική ενός νευρωνικού δικτύου για την ταξινόμηση προτύπων, επεξεργασία σήματος και εφαρμογές αυτόματου ελέγχου σε συνδυασμό με την προσαρμοστική ικανότητα το καθιστούν χρήσιμο εργαλείο για προσαρμοστική ταξινόμηση προτύπων, προσαρμοστική επεξεργασία και προσαρμοστικό έλεγχο συστημάτων [Haykin, 1931]. Το δίλλημα σταθερότητας-πλαστικότητας έχει αναπτυχθεί για την αξιοποίηση όλων των πλεονεκτημάτων της προσαρμοστικότητας, έτσι ώστε το σύστημα να αγνοεί τις πλασματικές διαταραχές και ταυτόχρονα να ανταποκρίνεται στις μεταβολές του περιβάλλοντος του που έχουν πραγματική σημασία [Grossberg, 1988].

Ενδεικτική Απόκριση: Σχετικά με την ταξινόμηση προτύπων, ένα νευρωνικό δίκτυο μπορεί να σχεδιαστεί ώστε να παρέχει πληροφορία από την επιλογή κατάλληλου προτύπου καθώς και το βαθμό εμπιστοσύνης στην ληφθείσα απόφαση. Η χρησιμότητα αυτής της πληροφορίας μπορεί να χρησιμοποιηθεί στην απόρριψη διφορούμενων προτύπων, με αποτέλεσμα να βελτιωθεί η απόδοση του δικτύου [Haykin, 1931].

Πληροφορία σχετική με το περιεχόμενο: Σε ένα νευρωνικό δίκτυο η γνώση προκύπτει από την δομή και την κατάσταση ενεργοποίησης. Ο κάθε νευρώνας μπορεί να επηρεάζεται από την συνολική δραστηριότητα όλων των άλλων νευρώνων. Άρα η κάθε σχετική με το περιεχόμενο πληροφορία διαχειρίζεται με φυσικό τρόπο από ένα νευρωνικό δίκτυο [Haykin, 1931].

Ανοχή σε βλάβες: Μία εγγενή δυνατότητα του νευρωνικού δικτύου είναι η ανθεκτικότητα σε λάθη, να μην η απόδοσή του θα μειωθεί αλλά θα μειωθεί βαθμιαία υπό αντίξοων συνθηκών [Haykin, 1931]. Σε περίπτωση που πρέπει να εξεταστεί αν ένα νευρωνικό δίκτυο είναι ανθεκτικό σε βλάβες, είναι αναγκαίο να ληφθούν διορθωτικά μέτρα στην σχεδίαση του αλγορίθμου που χρησιμοποιείται προς εκπαίδευση [Kerlirzin & Vallet, 1993].

Δυνατότητα υλοποίησης σε VLSI: Η παράλληλη αρχιτεκτονική ενός νευρωνικού δικτύου αποσκοπεί στον γρήγορο υπολογισμό συγκεκριμένων εφαρμογών. Η οποία καθιστά το νευρωνικό δίκτυο κατάλληλο για υλοποίηση με την χρήση τεχνολογίας VLSI (Very Large Scale Integration) [Haykin, 1931]. Το πλεονέκτημα της τεχνολογίας VLSI είναι ότι παρέχει ένα μέτρο σύλληψης με πολύπλοκη συμπεριφορά αλλά με εξαιρετικά ιεραρχικό τρόπο [Mead, 1989].

Ομοιομορφία ανάλυσης και σχεδίασης: Τα νευρωνικά δίκτυα απολαμβάνουν καθολικής αποδοχής ως επεξεργαστές πληροφοριών. Αυτό το χαρακτηριστικό μπορεί να περιγραφεί με διάφορους τρόπους [Haykin, 1931]:

1. Οι νευρώνες αντιπροσωπεύουν ένα συστατικό κοινό για όλα τα νευρωνικά δίκτυα
2. Ο κοινός χαρακτήρας καθιστά εφικτή την χρήση ίδιων θεωριών και αλγορίθμων μάθησης σε διαφορετικές εφαρμογές
3. Σπονδυλωτά δίκτυα μπορούν να κατασκευάζονται με απρόσκοπτη ενοποίηση επιμέρους λειτουργικών μονάδων

Αναλογία με τη νευροφυσιολογία του εγκεφάλου: Η σχεδίαση και η λειτουργία του νευρωνικού δικτύου επινοήθηκε από την αναλογία του με τον ανθρώπινο εγκέφαλο. Αυτό αποδεικνύει ότι η ανθεκτική σε σφάλματα παράλληλη επεξεργασία δεν είναι μόνο φυσικά εφικτή αλλά γρήγορη και ισχυρή. Τα τεχνητά νευρωνικά δίκτυα αντιμετωπίζονται από τους νευροβιολόγους ως ένα εργαλείο έρευνας για να μπορέσουν να περιγράψουν τα νευροβιολογικά φαινόμενα [Haykin, 1931]. Η σύγκριση γραμμικών μοντέλων του αιθουσο-οφθαλμικού αντανάκλαστικού VOR (Vestibulo-Ocular Reflex) με μοντέλα νευρωνικών δικτύων που βασίζονται σε ανατροφοδοτούμενα δίκτυα (Recurrent Network) έγινε στο έργο του Anastasio [Anastasio, 1993]. Ο αμφιβληστροειδής χιτώνας είναι το σημείο που αρχίζουν να δημιουργούνται σχέσεις μεταξύ του εξωτερικού κόσμου δια της όρασης και των πρώτων νευρωνικών εικόνων. Η λειτουργία του είναι να μετατρέπει την οπτική εικόνα σε νευρική εικόνα. Κατά τον Sterling ο μετασχηματισμός από οπτική εικόνα σε νευρική εικόνα περιλαμβάνει τρία στάδια [Sterling, 1990]:

1. Μετασχηματισμός της προσλαμβανόμενης εικόνας από ένα στρώμα νευρώνων (υποδοχείς)
2. Μετάδοση των σημάτων (αντίδραση στο φως) μέσω χημικών συνάψεων σε ένα στρώμα διπολικών κυττάρων
3. Μετάδοση αυτών των σημάτων μέσω χημικών συνάψεων, σε νευρώνες εξόδου (γαγγλιακά κύτταρα)

2.7 Κριτήρια Τερματισμού

2.7.1 Εποχές (Epochs)

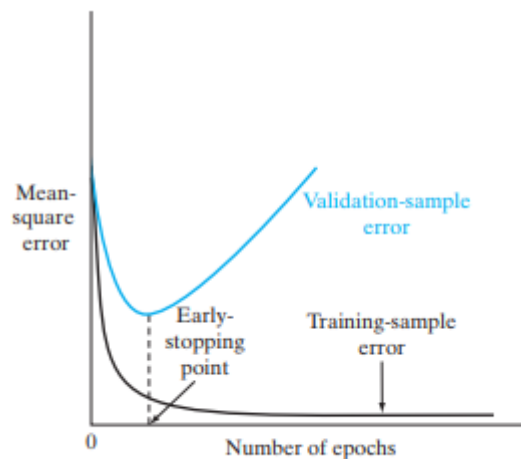
Στην περίπτωση που η εκπαίδευση ενός τεχνητού νευρωνικού δικτύου ολοκληρώσει τις δοσμένες εποχές η εκπαίδευση του σταματάει. Σε περίπτωση που ολοκληρωθεί και το σφάλμα δεν έχει φτάσει στο επιθυμητό αποτέλεσμα, μπορούμε να αυξήσουμε τις εποχές για την βελτίωση των αποτελεσμάτων.

2.7.2 Σφάλμα εκπαίδευσης (Training error)

Η επίτευξη του μικρότερου σφάλματος εκπαίδευσης (δηλαδή η παράμετρος **Performance goal** όπως φαίνετε και στους **Πίνακες 1 & 2**) κατά την εκπαίδευση ενός νευρωνικού δικτύου οδηγεί σε τερματισμό της εκπαίδευσης του νευρωνικού δικτύου.

2.7.3 Πρώιμος τερματισμός εκπαίδευσης (Early stopping)

Το σφάλμα της επικύρωσης (Validation check) χρησιμοποιείται για να ανιχνεύσει πότε ένα νευρωνικό δίκτυο ξεκινάει την υπερεκπαίδευση. Στην **Εικόνα 2.15** μπορούμε να διακρίνουμε ότι κατά την εκπαίδευση και με την παρέλευση των εποχών παρατηρούμε την καμπύλη του σφάλματος δείγματος επικύρωσης (Validation sample error) αρχικά να μειώνεται μονοτονικά μέχρι ένα τοπικό ελάχιστο. Αξίζει να σημειωθεί ότι κατά την εκπαίδευση μπορεί να υπάρξουν αρκετά τοπικά ελάχιστα πριν αρχίσει η καμπύλη σφάλματος δείγματος επικύρωσης να αυξάνεται. Επίσης η καμπύλη για το σφάλμα δείγματος εκπαίδευσης (Training sample error) μειώνεται μονοτονικά με την αύξηση του αριθμού των εποχών και συγκλίνει στον άξονα Χ (Number of epochs) η οποία μπορεί να βελτιστοποιηθεί αυξάνοντας τον αριθμός των εποχών. Στην συνέχεια μετά το τοπικό ελάχιστο ή αλλιώς στο σημείο πρώιμου τερματισμού (Early stopping point) παρατηρούμε την καμπύλη επικύρωσης να αυξάνεται καθώς συνεχίζεται η εκπαίδευση και εκεί η εκπαίδευση σταματάει, αυτό ονομάζεται πρόωρη διακοπή (Early stopping). Η επιλογή ενός τέτοιου κριτηρίου τερματισμού μπορεί να χρησιμοποιηθεί στην γρηγορότερη μάθηση ή στην βελτίωση της γενίκευσης του δικτύου [Prechelt, 1998] & [Haykin, 1931].



Εικόνα 2.15: Γραφική αναπαράσταση του κανόνα πρώιμου τερματισμού (Simon Haykin, Neural network & learning mechanics)

Σύμφωνα με τον Prechelt, σχετικά με μία έρευνα που διεξήγαγε το 1998, απέδειξε ότι υπάρχει μία συσχέτιση μεταξύ του χρόνου εκπαίδευσης και της γενίκευσης. Μέσα από ένα σετ δεδομένων 1296 μονάδων εκπαίδευσης, 12 διαφορετικών προβλημάτων και 24 διαφορετικών αρχιτεκτονικών δικτύου. Κατέληξε στο συμπέρασμα ότι όταν παρουσιάζονται δύο η περισσότερα τοπικά ελάχιστα γίνεται η επιλογή ενός αργότερου κριτηρίου τερματισμού το οποίο επιφέρει μια μικρή βελτίωση στην απόδοση της γενίκευσης (περίπου 4%) αλλά με πολύ μεγαλύτερο χρόνο εκπαίδευσης (περίπου τετραπλάσιο) [Prechelt, 1998].

2.7.4 Μη τροποποίηση βαρών

Κατά την εκπαίδευση ενός νευρωνικού δικτύου τα σήματα σφάλματος είναι εκείνα για τα οποία τα βάρη τροποποιούνται, σε περίπτωση που τα σήματα σφάλματος παραμένουν ίδια και για την επόμενη εποχή τα βάρη δεν θα τροποποιηθούν. Αυτό έχει σαν αποτέλεσμα να σταματήσει η εκπαίδευση του νευρωνικού δικτύου διότι δεν θα αλλάξουν ούτε τα βάρη ούτε οι ζητούμενες τιμές που προσπαθεί το νευρωνικό δίκτυο να εκτιμήσει.

2.8 Υπερεκπαίδευση του δικτύου (Overtaining)

Στην περίπτωση μας χρησιμοποιήθηκε η πρόωρη διακοπή (Early Stopping) [Morgan & Bourlard, 1990]. Εναλλακτικές τεχνικές για την αντιμετώπιση της υπερεκπαίδευσης μπορούν να χωριστούν σε δύο κατηγορίες. Αυτές που μειώνουν τον αριθμό των διαστάσεων του χώρου παραμέτρων και αυτές που μειώνουν το πραγματικό μέγεθος κάθε διάστασης.

Τεχνικές για την μείωση του αριθμού των παραμέτρων:

- Άπληστη επικοδομητική μάθηση (Greedy constructive learning) [Fahlman & Lebiere, 1990]
- Κλάση (Prunning) [Le Cun et. al, 1990], [Hassibi & Stork, 1993], [Levin et. al, 1994]
- Κατανομή βάρους (Weight sharing) [Nowlan & Hinton, 1992]

Τεχνικές για την μείωση μεγέθους κάθε διάστασης:

- Πρόωρη διακοπή (Early stopping) [Morgan & Bourlard, 1990]
- Αποσύνθεση βάρους (Weight decay) [Krogh & Hertz, 1992]
- Και άλλες [Weigend, 1991]

Κεφάλαιο 3

Μεθοδολογία

Αρχικά τα δεδομένα όσον αφορά τις στάθμες των πηγαδιών καθώς και τις 10 παραμέτρους εισόδου που προαναφέραμε, περιέχουν τιμές από 01/01/2000 έως 31/10/2014. Από τα συνολικά 161 πηγάδια παρατήρησης που επιλέξαμε αρχικά μόνο για τα 128 πηγάδια παρατήρησης έχουμε δεδομένα για τις τρέχουσες ημερομηνίες. Πιο συγκεκριμένα υπάρχουν 57 πηγάδια παρατήρησης στην Αυστρία, 2 πηγάδια παρατήρησης στην Βουλγαρία, 32 πηγάδια παρατήρησης στην Γερμανία, 9 πηγάδια παρατήρησης στην Κροατία, 20 πηγάδια παρατήρησης στην Ουγγαρία, 1 πηγάδι παρατήρησης στην Ρουμανία, 6 πηγάδια παρατήρησης στην Σερβία και 1 πηγάδι παρατήρησης στην Σλοβενία.

Με βάση τον παρακάτω κώδικα δημιουργήσαμε τον πίνακα εισόδου και το διάνυσμα στόχου (Target) τα οποία είναι απαραίτητα για την εκπαίδευση του νευρωνικού δικτύου.

Ο πίνακας εισόδου *neuralin* αποτελείται από 10 στήλες. Πιο αναλυτικά:

- 1^η, 2^η & 3^η στήλη περιέχουν δεδομένα για την χρονολογία (δλδ. Έτος, Μήνας, Ημέρα αντίστοιχα)
- 4^η στήλη περιέχει δεδομένα για την Βροχόπτωση
- 5^η στήλη περιέχει δεδομένα για την μέση θερμοκρασία
- 6^η στήλη περιέχει δεδομένα για δυνητική εξατμισοδιαπνοή
- 7^η στήλη περιέχει δεδομένα για εξατμισμό από υδάτινες επιφάνειες
- 8^η στήλη περιέχει δεδομένα για την εξατμισοδιαπνοή από το έδαφος και τα φυτά
- 9^η & 10^η στήλη περιέχει δεδομένα για τις κατά Χ και Υ συντεταγμένες αντίστοιχα

Το διάνυσμα στόχου *target* περιέχει τιμές για τις πραγματικές στάθμες των πηγαδιών και χρησιμοποιείται ως διάνυσμα εξόδου κατά την εκπαίδευση του νευρωνικού δικτύου για τον υπολογισμό του σήματος σφάλματος.

Τέλος το διάνυσμα *jrc* περιέχει τους κωδικούς των πηγαδιών που χρησιμοποιούνται.

Κώδικας δημιουργίας πίνακα εισόδου (neuralin) και διανύσματος στόχου (Target) για την εκπαίδευση του νευρωνικού δικτύου

```
[k1,m1]=size(danube);
[k2,m2]=size(pcp);
neuralin=zeros(k1,10);
target=zeros(k1,1);
nv=0;
for k=1:k1
    for n=1:k2
        if danube(k,1)==E0(n,1) && danube(k,2)==E0(n,2) &&
danube(k,3)==E0(n,3)
            neuralin(k,1:3)=E0(n,1:3); % year month day
            neuralin(k,4)=pcp(n,danube(k,8)+3); %precipitation
            neuralin(k,5)=tav(n,danube(k,8)+3); %temperature average
            neuralin(k,6)=E0(n,danube(k,8)+3); %E0
            neuralin(k,7)=ES(n,danube(k,8)+3); %ES
            neuralin(k,8)=ET(n,danube(k,8)+3); %ET
            neuralin(k,9)=danube(k,5); % X_Coord
            neuralin(k,10)=danube(k,6); % Y_Coord
            target(k,1)=danube(k,4); % h
            jrc{k,1}=jrccode{k}; %jrccode
            nv=nv+1;
        end
    end
end
```

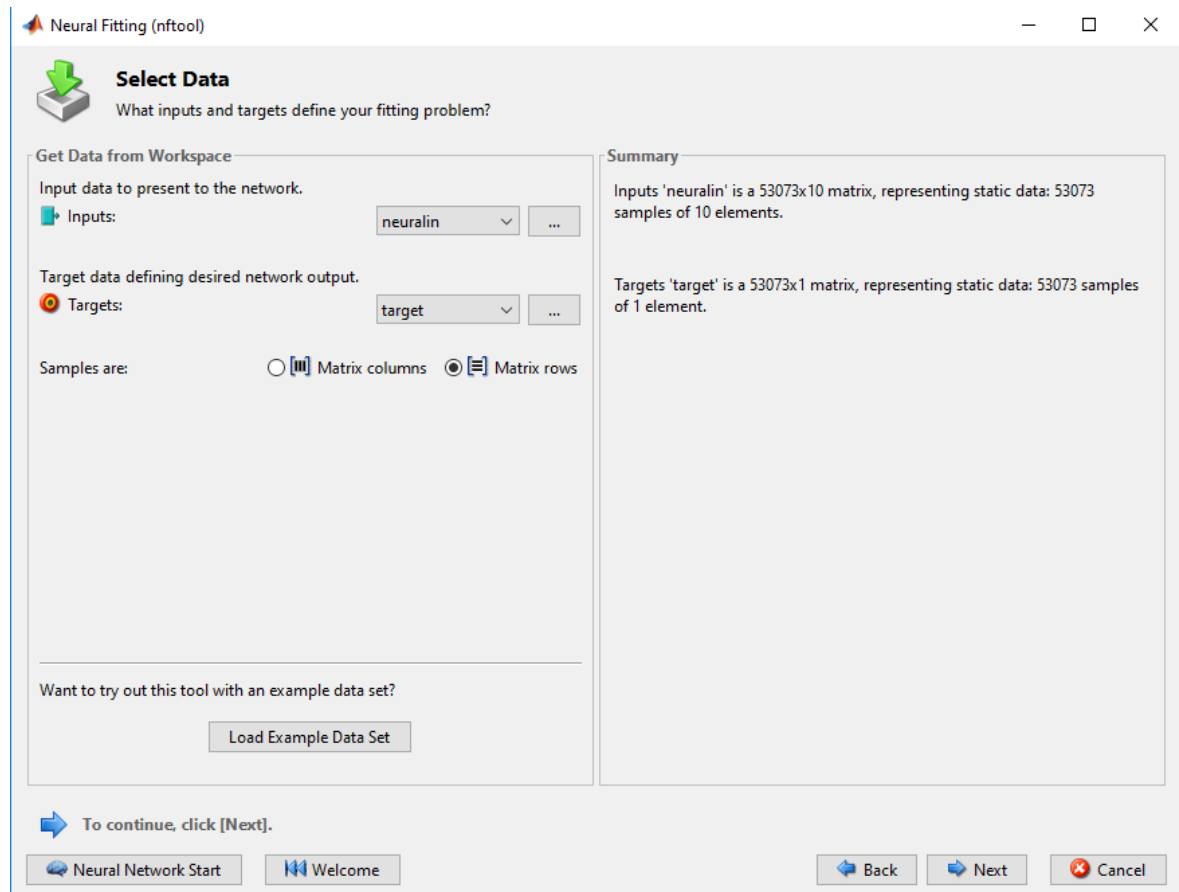
Στην συνέχεια αφού δημιουργήσαμε τα διανύσματα τα οποία είναι απαραίτητα για την εκπαίδευση του νευρωνικού δικτύου. Εκπαιδευτήκαν νευρωνικά δίκτυα με την χρήση δύο διαφορετικών αλγορίθμων:

- 1^{ος} αλγόριθμος→ Levenberg-Marquardt
- 2^{ος} αλγόριθμος→ Bayesian Regularization

3.1 Βήματα εκπαίδευσης νευρωνικού δικτύου με Neural Fitting tool

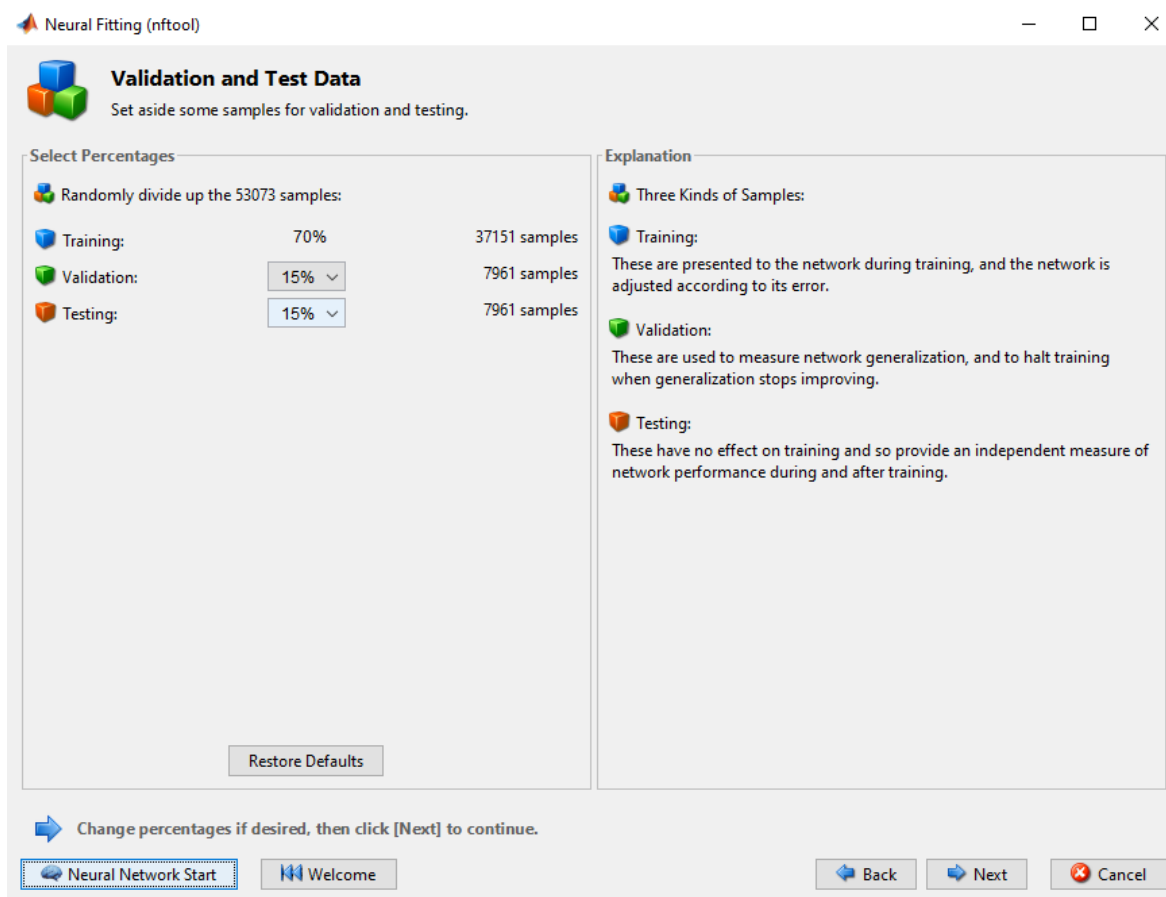
Κατά τη χρήση του Neural Fitting (nftool) σε περιβάλλον Matlab, ακολουθούνται τα παρακάτω βήματα.

Στο **Βήμα 1** ένα γίνεται η επιλογή του πίνακα εισόδου και διανύσματος στόχου.



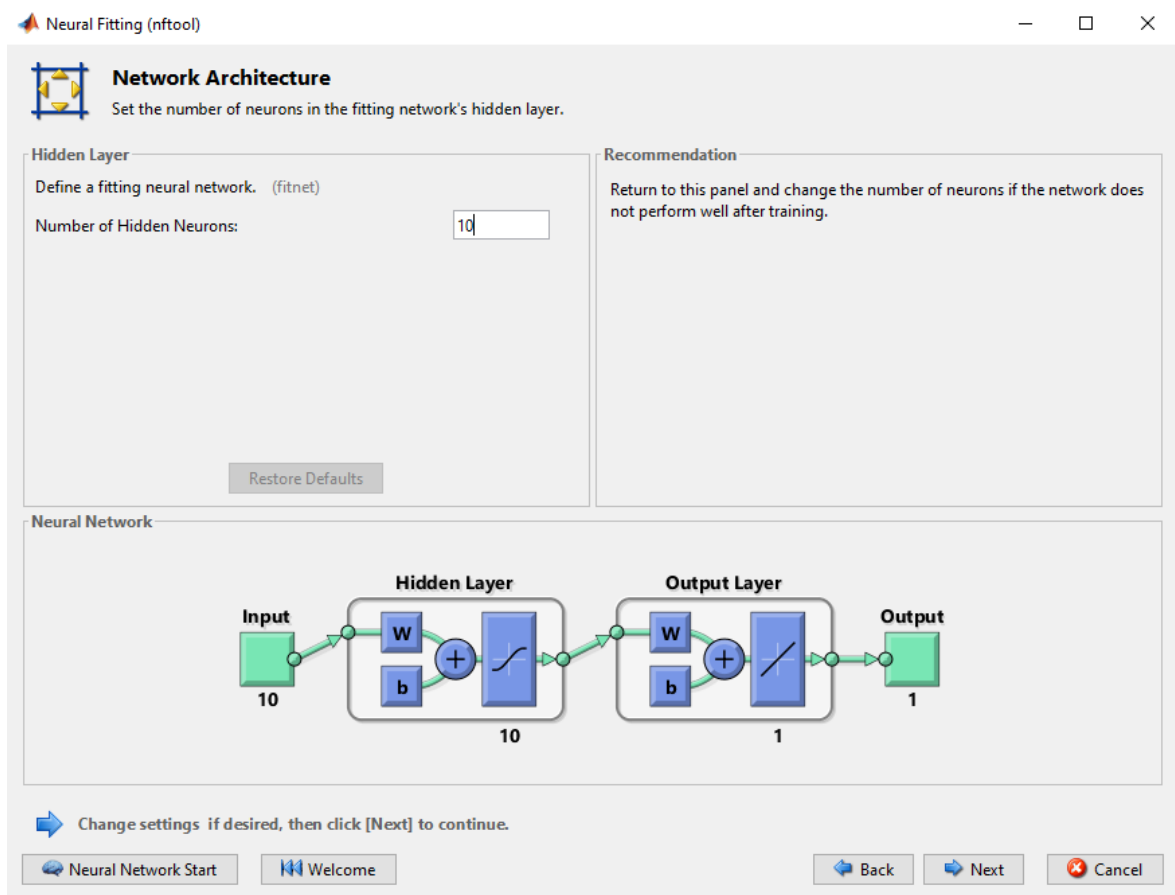
Βήμα 1: Εισαγωγή δεδομένων Εισόδου και Στόχου

Στο **Βήμα 2** γίνεται η τυχαία κατανομή δεδομένων που κρατούνται για εκπαίδευση (Training), για επικύρωση (Validation) και για έλεγχο (Testing). Τα δεδομένα που κρατούνται για εκπαίδευση παρουσιάζονται στο δίκτυο κατά την διάρκεια της εκπαίδευσης και το δίκτυο προσαρμόζεται ανάλογα με το σφάλμα του. Τα δεδομένα που κρατούνται για επικύρωση χρησιμοποιούνται για την μέτρηση της γενίκευσης του δικτύου και η εκπαίδευση του δικτύου σταματάει όταν η γενίκευση σταματάει να βελτιώνεται. Τέλος τα δεδομένα που κρατούνται για έλεγχο δεν επηρεάζουν την εκπαίδευση, παρέχουν όμως ένα ανεξάρτητο μέτρο της απόδοσης του δικτύου κατά τη διάρκεια και μετά την εκπαίδευση. Αξίζει να σημειωθεί ότι το ελάχιστο ποσοστό για εκπαίδευση είναι 30% και για επικύρωση και έλεγχο κυμαίνονται από 5%-35%.



Βήμα 2: Επιλογή ποσοστών για εκπαίδευση, επικύρωση και έλεγχο

Στο **Βήμα 3** γίνεται η επιλογή κρυφών κόμβων. Τα όρια των κρυφών κόμβων κυμαίνονται από 1-10000 κόμβους.



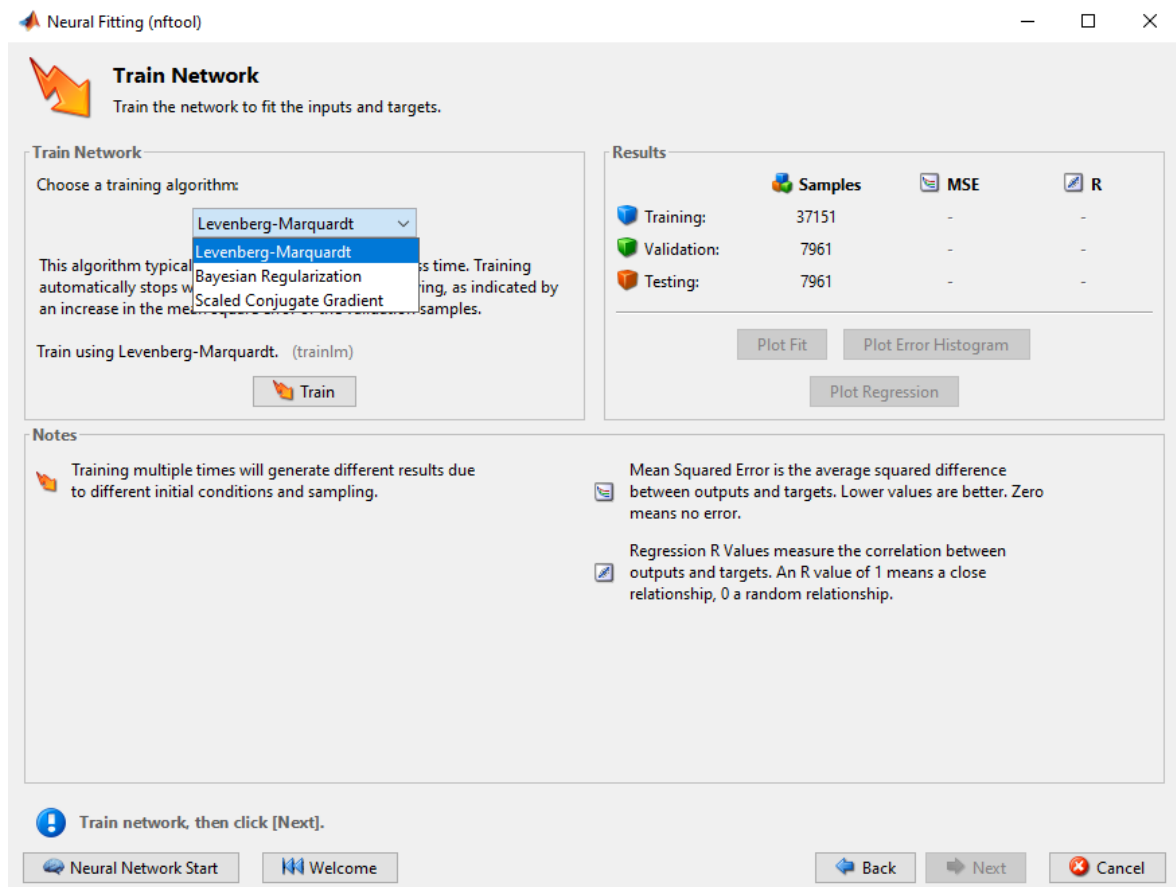
Βήμα 3: Επιλογή κρυφών κόμβων

Αξίζει να σημειωθεί ότι το κρυφό επίπεδο (Hidden Layer) ενεργοποιείται με βάση μια σιγμοειδή συνάρτηση ενώ το επίπεδο εξόδου (Output Layer) ενεργοποιείται με βάση μια γραμμική συνάρτηση. Νευρωνικά δίκτυα εκπαιδεύτηκαν όπου και τα δύο επίπεδα εκφράζονται από σιγμοειδή συνάρτηση.

Στο **Βήμα 4** γίνεται η επιλογή του αλγόριθμου εκπαίδευσης.

Ο αλγόριθμος Levenberg-Marquardt απαιτεί τυπικά περισσότερη μνήμη αλλά λιγότερο χρόνο. Η εκπαίδευση σταματάει αυτόματα όταν η γενίκευση σταματάει να βελτιώνεται, όπως φαίνεται από την αύξηση του μέσου τετραγωνικού σφάλματος των δειγμάτων επικύρωσης.

Ο αλγόριθμος Bayesian Regularization απαιτεί περισσότερο χρόνο, αλλά μπορεί να οδηγήσει σε καλή γενίκευση για δύσκολα, μικρά ή θορυβώδη σύνολα δεδομένων. Η εκπαίδευση σταματάει σύμφωνα με την ελαχιστοποίηση του προσαρμοστικού βάρους (κανονικοποίηση). Το ποσοστό του Validation συμπεριλαμβάνεται στο ποσοστό του Training κατά την εκπαίδευση του τεχνητού νευρωνικού δικτύου.



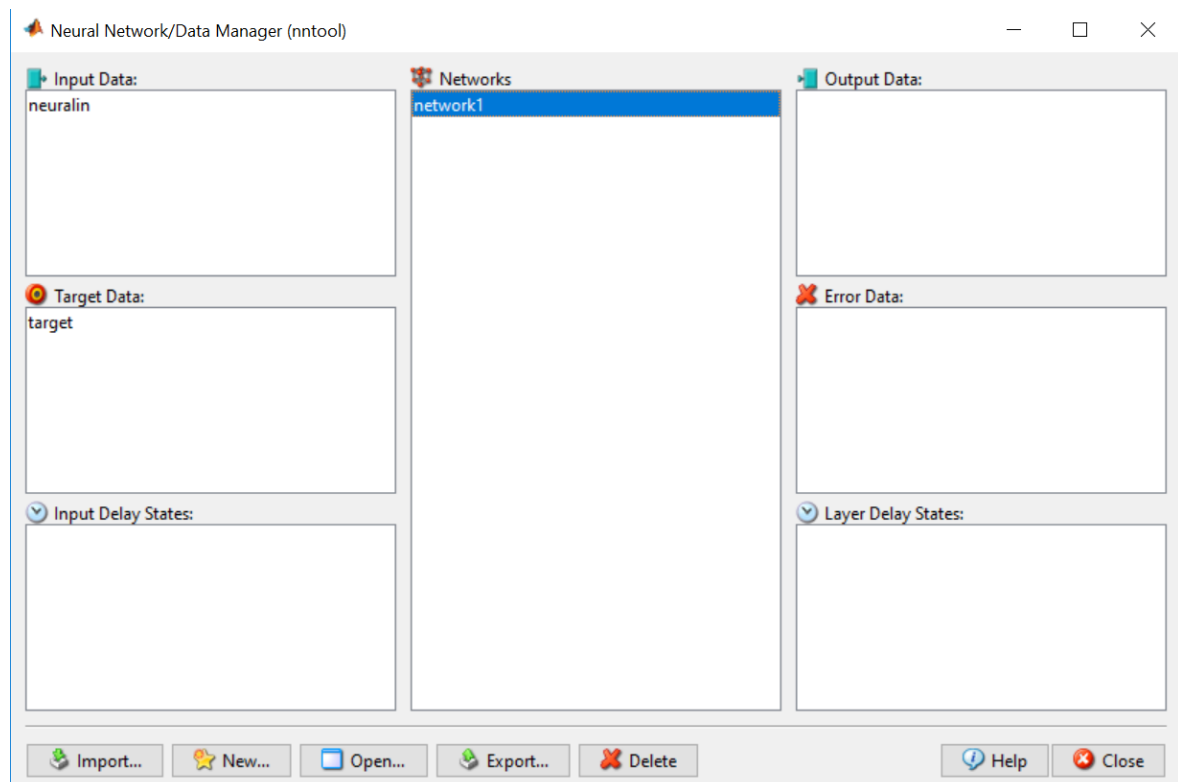
Βήμα 4: Επιλογή αλγόριθμου εκπαίδευσης

Όπως φαίνεται και στο **Κεφάλαιο 4** νευρωνικά δίκτυα εκπαιδεύτηκαν για διαφορετικές παραμέτρους. Οι διαφορετικές παράμετροι συμπεριλαμβάνουν τα ποσοστά για τα οποία εκπαιδεύεται το νευρωνικό δίκτυο, τους αλγόριθμους εκπαίδευσης που περιγράψαμε καθώς και τον αριθμό των εποχών που στην συγκεκριμένη περίπτωση παραμένει σταθερός με τιμή τις 1000 εποχές.

3.2 Βήματα εκπαίδευσης νευρωνικού δικτύου με Neural Network tool

Στην Ενότητα 4.3.1 και 4.3.2 εκπαιδεύτηκαν νευρωνικά δίκτυα με την χρήση της εντολής nntool στο Command Window του επιστημονικού προγράμματος Matlab. Με το εξής εργαλείο για τον αλγόριθμο Levenberg-Marquardt έχουμε το πλεονέκτημα να αυξήσουμε τις εποχές (Iterations) έτσι ώστε η εκπαίδευση του νευρωνικού δικτύου να μην σταματάει μετά το πέρας των εποχών καθώς επίσης να μπορεί ο χρήστης να δώσει μεγαλύτερο μέγιστο αριθμό Validation Checks με αποτέλεσμα όταν η γενίκευση σταματάει να βελτιώνεται για μια σειρά από σφάλματα να έχει μεγαλύτερο περιθώριο. Διότι υπάρχει περίπτωση να έχει μια σειρά από σφάλμα που δεν βελτιώνουν την γενίκευση αλλά μετά το πέρας των σφαλμάτων να αρχίσει να βελτιώνεται πάλι. Για τον αλγόριθμο του Bayesian Regularization έχουμε το πλεονέκτημα μόνο για την αύξηση των εποχών καθώς δεν κρατάει σφάλμα επικύρωσης.

Στο **Βήμα 1** γίνεται η επιλογή του πίνακα εισόδου και διανύσματος στόχου καθώς και η δημιουργία του Νευρωνικού δικτύου network1.



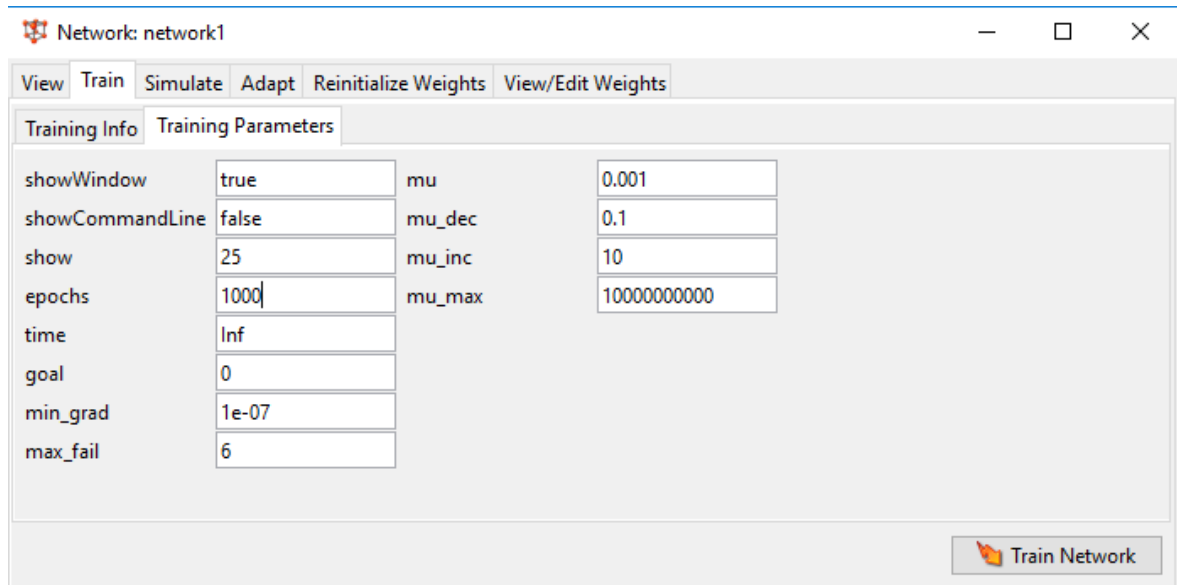
Βήμα 1: Επιλογή δεδομένων εισόδου και στόχου

Στο **Βήμα 2** γίνεται η επιλογή του τύπου του νευρωνικού (Network Type), στην συνέχεια επιλέγονται τα δεδομένα εισόδου και στόχου. Στην επιλογή Training Function επιλέγεται ο αλγόριθμος με τον οποίο θα εκπαιδευτεί το νευρωνικό δίκτυο, η επιλογή TRAINLM→Levenberg-Marquardt και η επιλογή TRAINBR→Bayesian Regularization. Επιλέγουμε 2 επίπεδα που αντιστοιχούν στο κρυφό επίπεδο και στο επίπεδο εξόδου και επιλέγουμε και στα δύο ως Transfer Function την επιλογή TANSIG. Καθώς ορίζουμε και των αριθμό των κρυφών κόμβων που θέλουμε να χρησιμοποιήσουμε.

The screenshot shows the 'Create Network or Data' dialog box with the 'Network' tab selected. The 'Name' field is 'network1'. The 'Network Properties' section shows 'Feed-forward backprop' for 'Network Type'. The 'Input data' is 'neuralin', 'Target data' is 'target', 'Training function' is 'TRAINLM', 'Adaption learning function' is 'LEARNGDM', and 'Performance function' is 'MSE'. The 'Number of layers' is set to 2. The 'Properties for: Layer 1' section shows 'Number of neurons' as 100 and 'Transfer Function' as 'TANSIG'. At the bottom, there are buttons for 'View', 'Restore Defaults', 'Help', 'Create', and 'Close'.

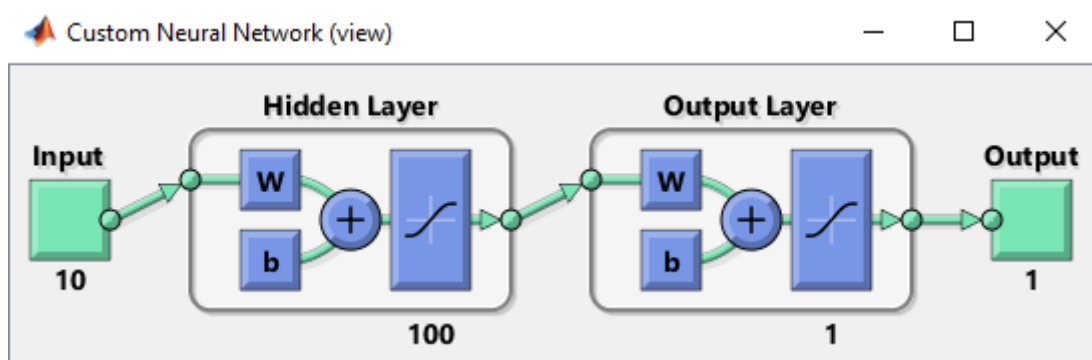
Βήμα 2: Επιλογή κατάλληλων παραμέτρων

Στο **Βήμα 3** στην επιλογή epochs ορίζουμε τον αριθμό των εποχών και για τον Αλγόριθμο Levenberg-Marquardt στην επιλογή max_fail ορίζουμε τον μέγιστο αριθμό Validation Checks.



Βήμα 3: Ορισμός αριθμού εποχών και Validation Checks (για τον αλγόριθμο LM)

Στην **Εικόνα 3.16** παρουσιάζεται η αρχιτεκτονική του δικτύου όπου το κρυφό επίπεδο και το επίπεδο εξόδου εκφράζονται από σιγμοειδή συνάρτηση.



Εικόνα 3.16: Αρχιτεκτονική δικτύου όπου το κρυφό επίπεδο και το επίπεδο εξόδου ενεργοποιούνται με βάση σιγμοειδή συνάρτηση

Στην περίπτωση μας, η εκπαίδευση νευρωνικού δικτύου για ένα πηγάδι αποτελεί ένα διαφορετικό μοντέλο, για το οποίο θέλαμε να μάθουμε την αποτελεσματικότητά του. Στην Ενότητα 4.5 εκπαιδεύτηκε νευρωνικό δίκτυο για ένα πηγάδι της Αυστρίας με κωδικό πηγαδιού 'at_300533'. Ως δεδομένα εισόδου παραμένουν τα ίδια με αυτά που προαναφέραμε στην μεθοδολογία τα οποία αντιστοιχούν στο συγκεκριμένο πηγάδι με μια μικρή διαφορά. Αφαιρέσαμε την 9η & 10η στήλη που αντιστοιχούν σε δεδομένα ως προς τις συντεταγμένες του πηγαδιού, διότι για ένα πηγάδι αυτές οι τιμές παραμένουν ίδιες και δεν είναι απαραίτητες για την εκπαίδευση. Το διάνυσμα στόχου περιέχει τις αντίστοιχες τιμές για τις πραγματικές στάθμες για το συγκεκριμένο πηγάδι.

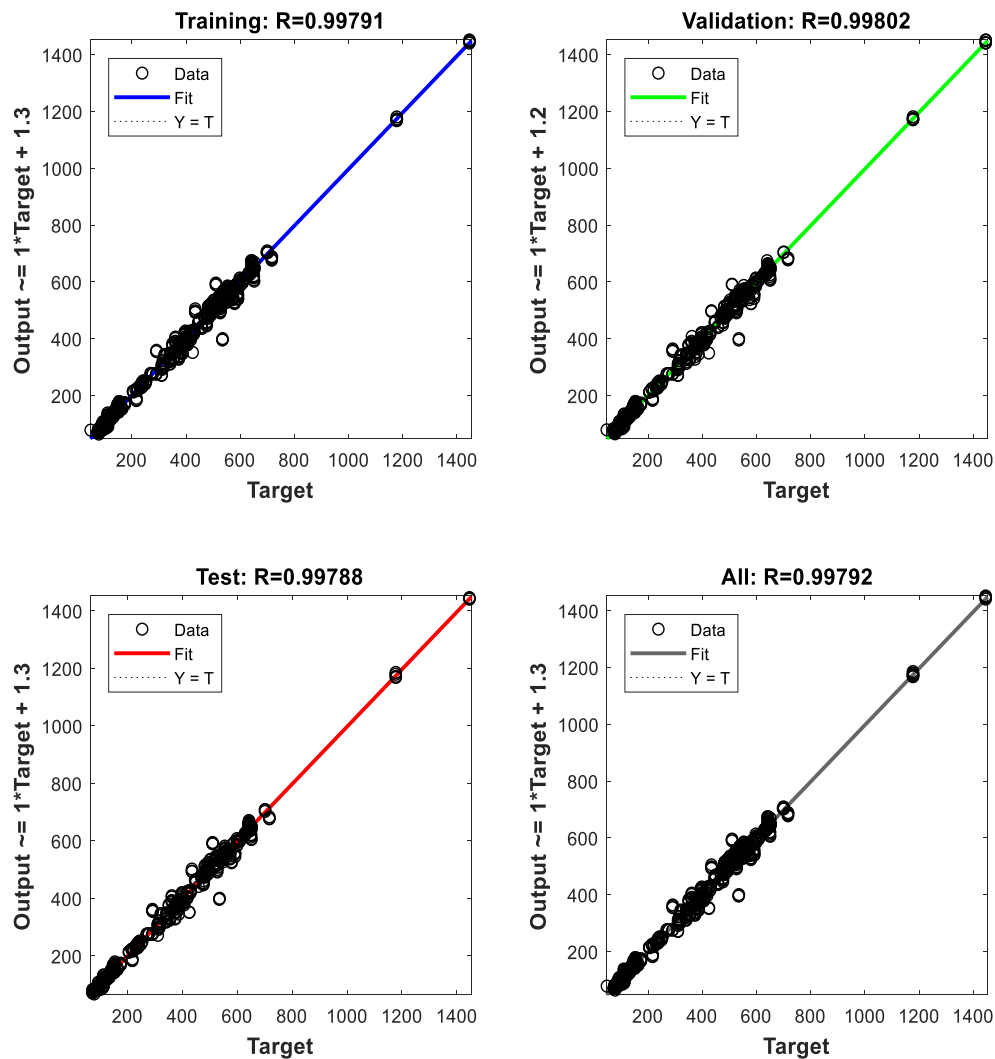
Το νευρωνικό δίκτυο εκπαιδεύτηκε με τον αλγόριθμο BR, για ποσοστά Training=70%, Validation=5% & Testing=25%, για 1000 εποχές και με 10 κρυφούς κόμβους.

Η σύγκριση στο **Γράφημα 9** γίνεται για ίδιους αλγόριθμους (BR) αλλά για διαφορετικές παραμέτρους. Ο λόγος για τον οποίο χρησιμοποιήθηκαν μόνο 10 κρυφοί κόμβοι είναι γιατί τα δεδομένα μας για την εκπαίδευση νευρωνικού δικτύου για ένα μόνο πηγάδι ήταν πολύ λιγότερα σε σχέση με τα συνολικά δεδομένα εισόδου για όλα τα πηγάδια (128 πηγάδια συνολικά).

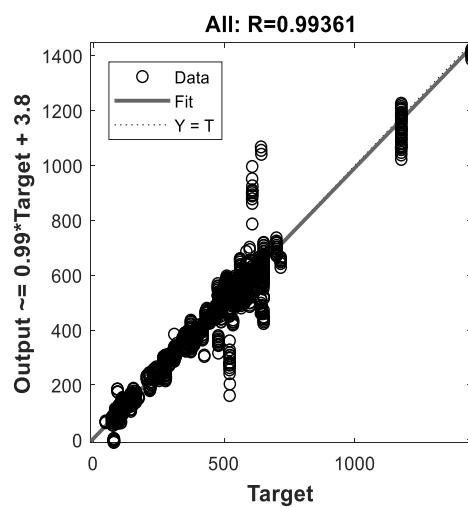
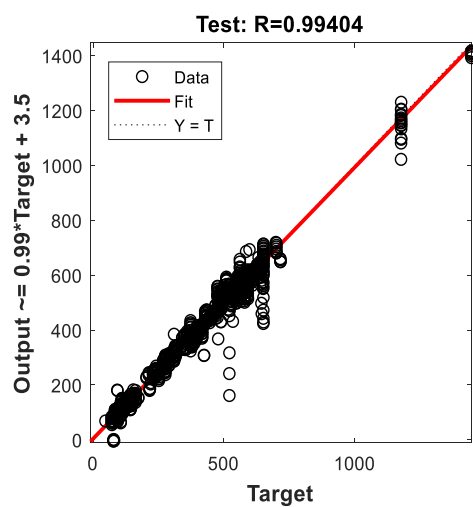
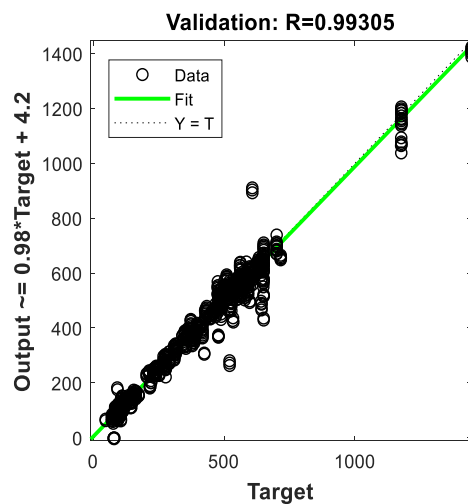
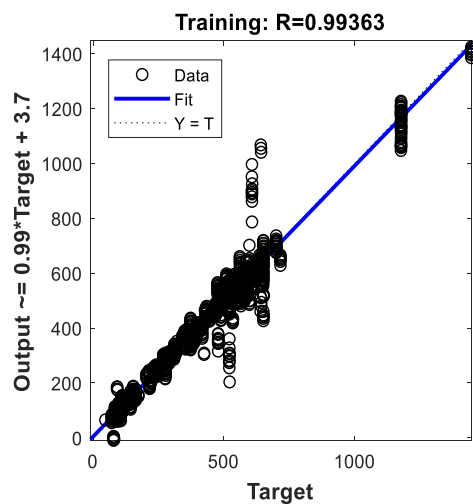
Κεφάλαιο 4

Αποτελέσματα

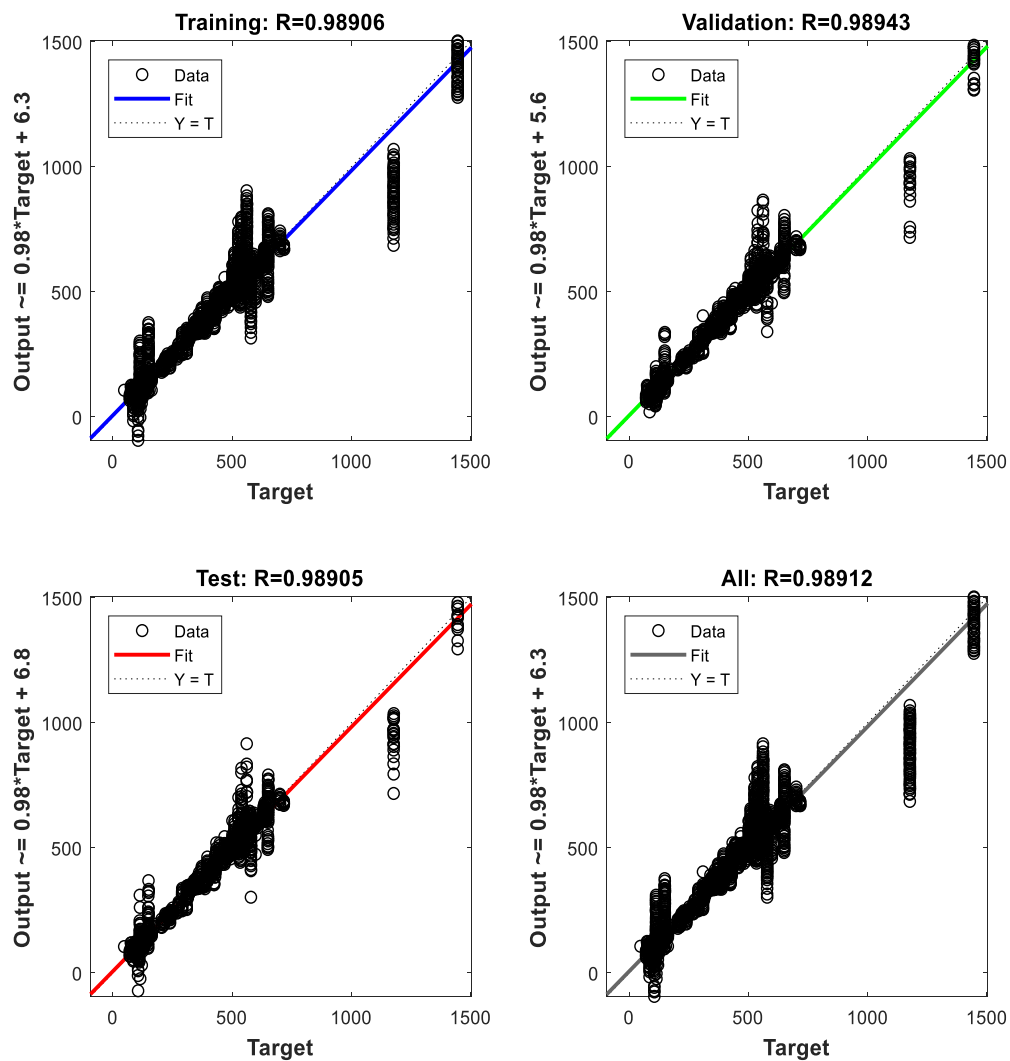
4.1 Χρήση αλγόριθμου Levenberg-Marquardt



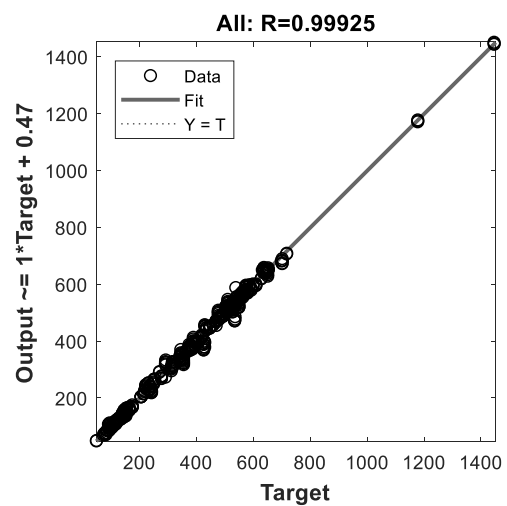
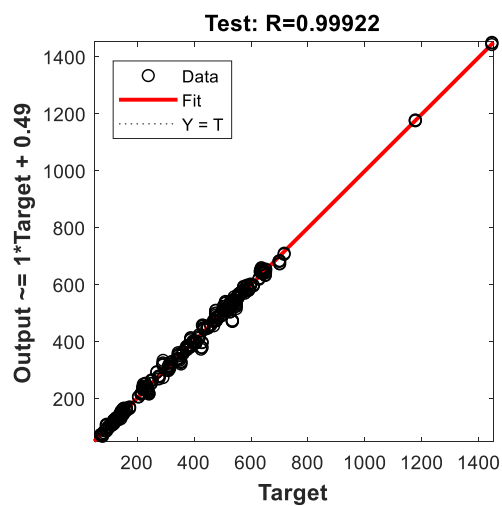
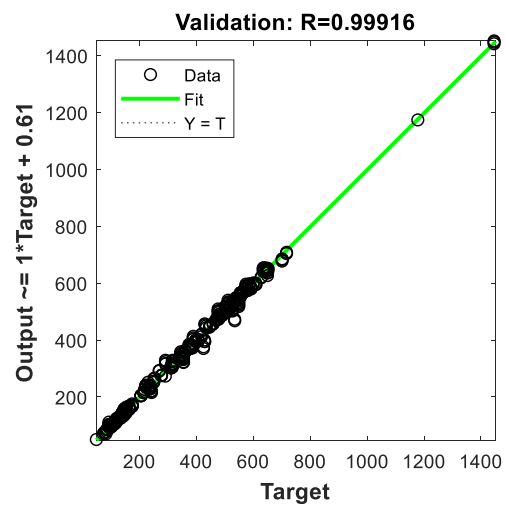
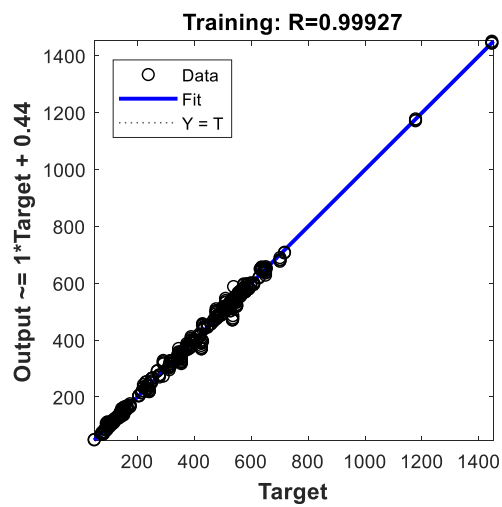
Διάγραμμα 1: Παλινδρόμηση εκπαίδευσης (Training Regression) με 20 κρυφούς κόμβους (Hidden nodes), με ποσοστά 70% Training, 15% Validation & 15% Testing, διακοπή εκπαίδευσης λόγω μέγιστου αριθμού Validation Checks



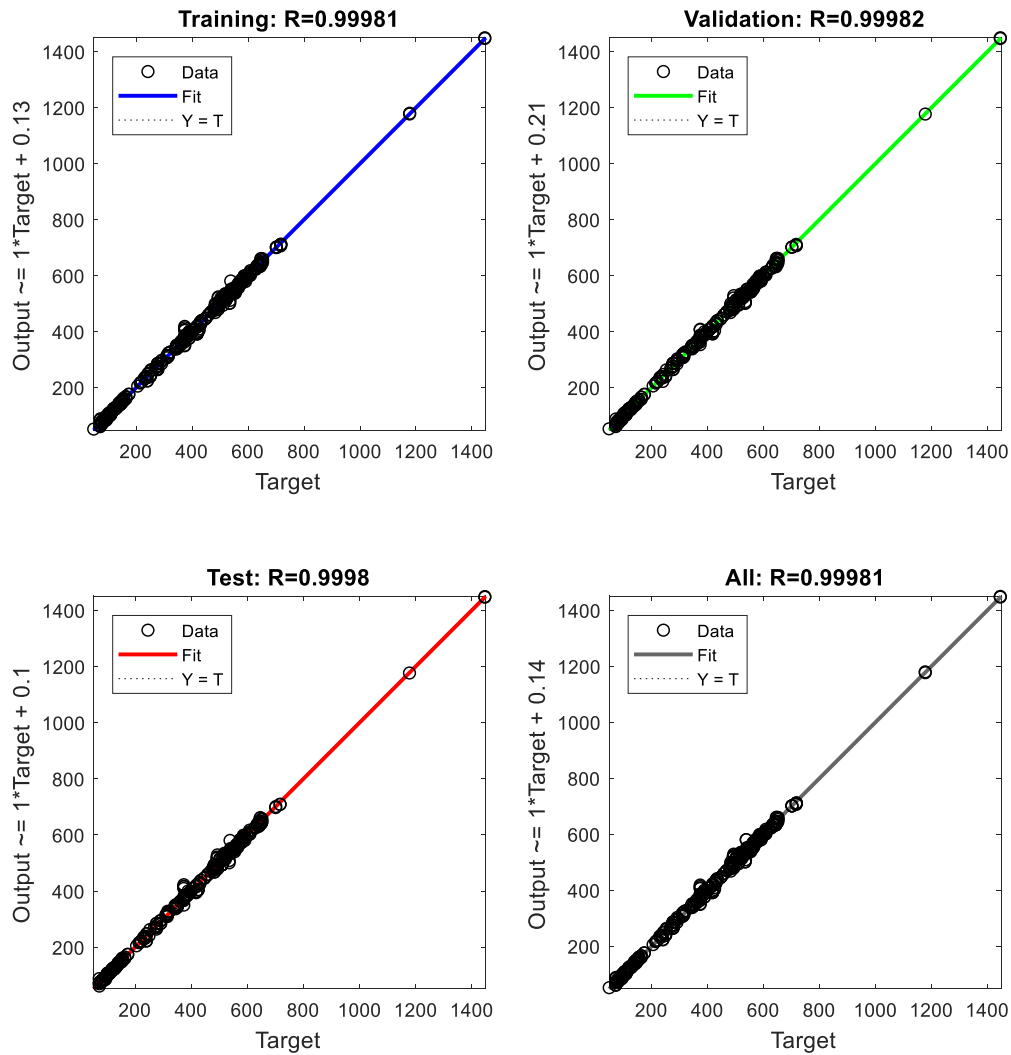
Διάγραμμα 2: Παλινδρόμηση εκπαίδευσης (Training Regression) με 25 κρυφούς κόμβους (Hidden nodes), με ποσοστά 70% Training, 15% Validation & 15% Testing



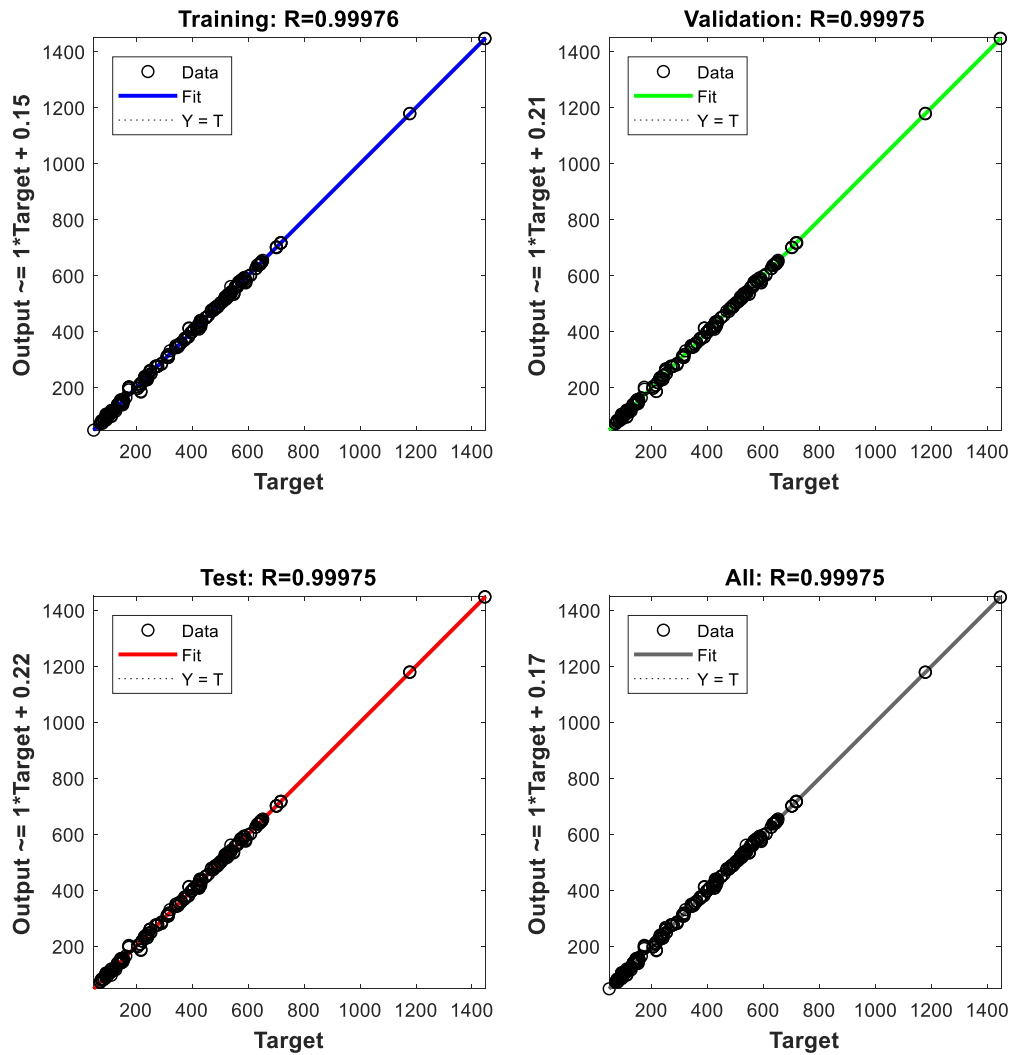
Διάγραμμα 3: Παλινδρόμηση εκπαίδευσης (Training Regression) με 30 κρυφούς κόμβους (Hidden nodes), με ποσοστά 70% Training, 15% Validation & 15% Testing, διακοπή εκπαίδευσης λόγω μέγιστου αριθμού Validation Checks



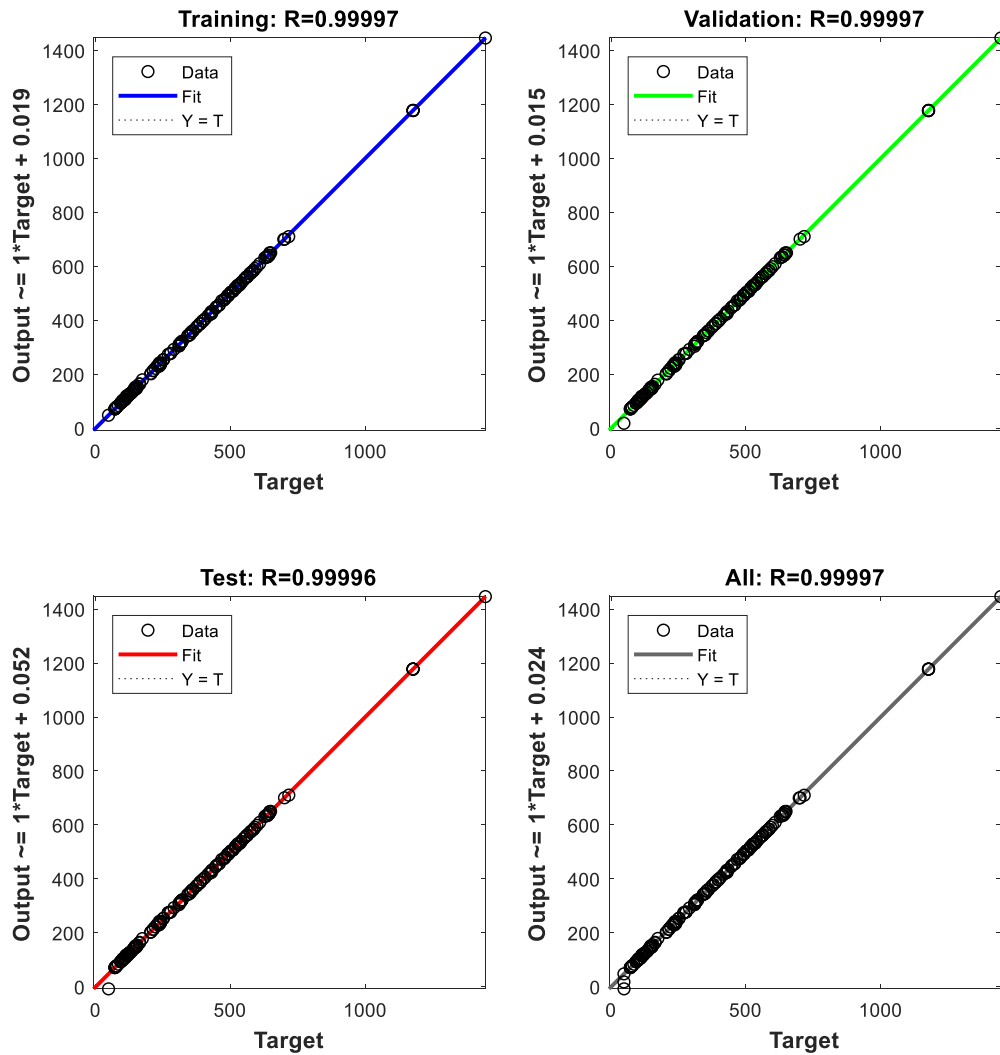
Διάγραμμα 4: Παλινδρόμηση εκπαίδευσης (Training Regression) με 35 κρυφούς κόμβους (Hidden nodes), με ποσοστά 70% Training, 15% Validation & 15% Testing



Διάγραμμα 5: Παλινδρόμηση εκπαίδευσης (Training Regression) με 50 κρυφούς κόμβους (Hidden nodes), με ποσοστά 70% Training, 15% Validation & 15% Testing



Διάγραμμα 6: Παλινδρόμηση εκπαίδευσης (Training Regression) με 80 κρυφούς κόμβους (Hidden nodes), με ποσοστά 70% Training, 15% Validation & 15% Testing



Διάγραμμα 7: Παλινδρόμηση εκπαίδευσης (Training Regression) με 100 κρυφούς κόμβους (Hidden nodes), με ποσοστά 70% Training, 15% Validation & 15% Testing

Πίνακας 3: Απαραίτητοι παράμετροι για την εκπαίδευση με τον αλγόριθμο Levenberg-Marquardt

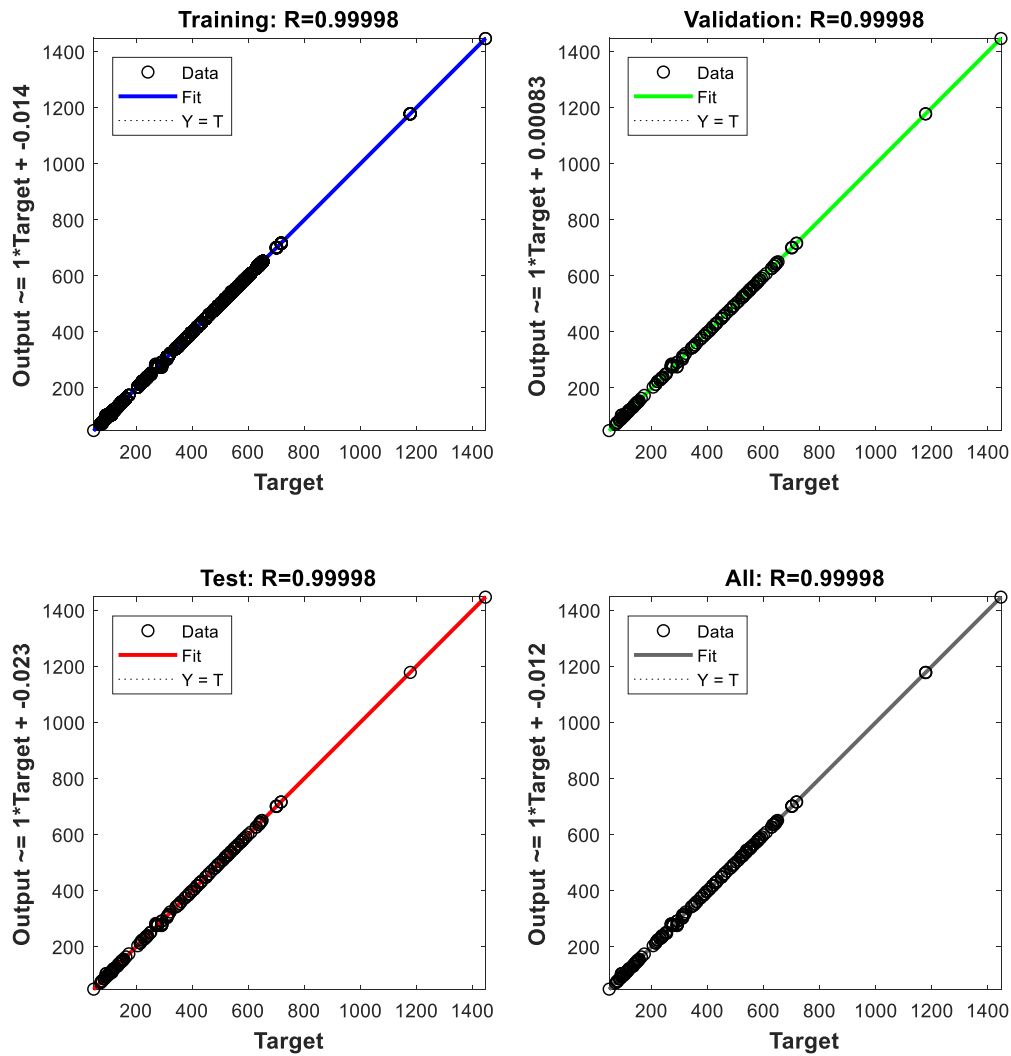
Διάγραμμα	Κρυφοί Κόμβοι	Percentages			Epochs
		Training	Validation	Testing	
1	20	70%	15%	15%	1000
2	25	70%	15%	15%	1000
3	30	70%	15%	15%	1000
4	35	70%	15%	15%	1000
5	50	70%	15%	15%	1000
6	80	70%	15%	15%	1000
7	100	70%	15%	15%	1000

Πίνακας 4: Αποτελέσματα για το μέσο τετραγωνικό σφάλμα (MSE) και τον συντελεστή συσχέτισης (R)

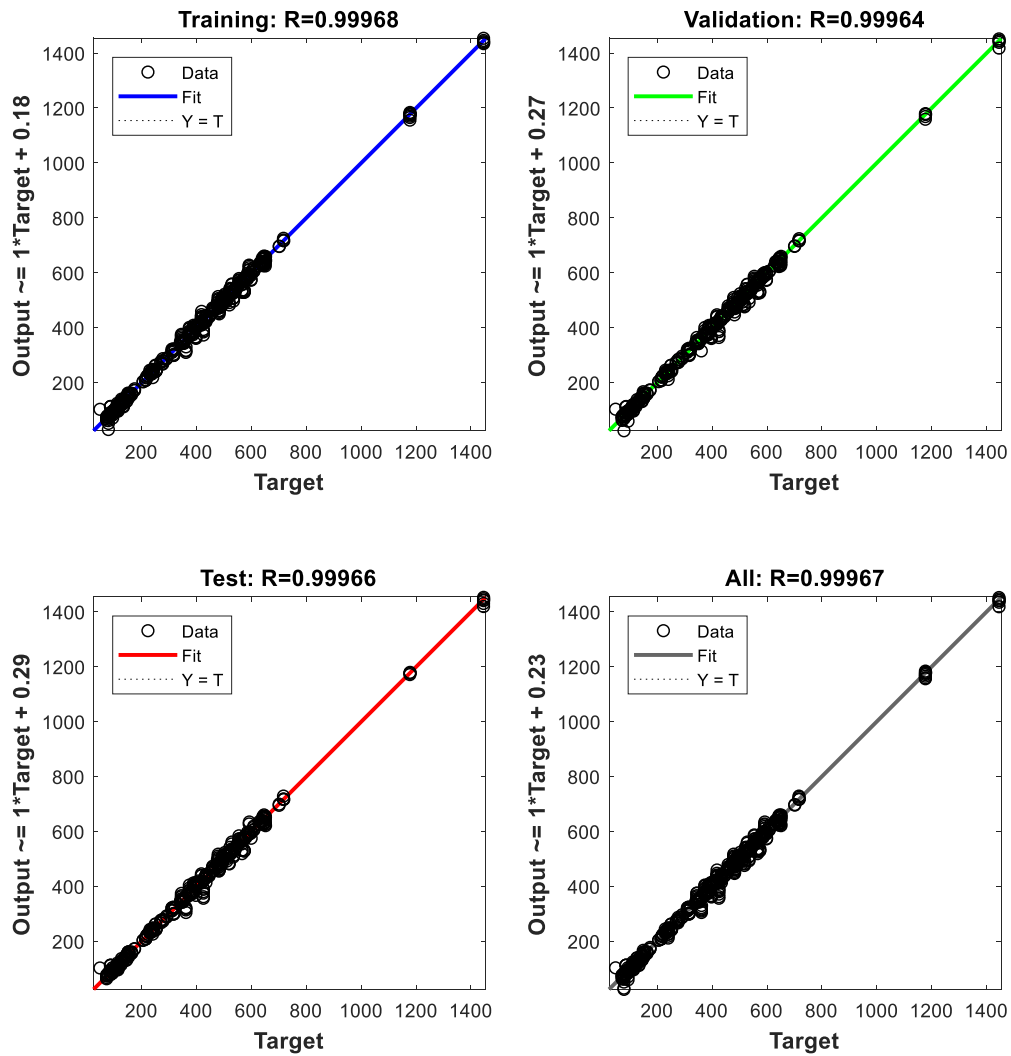
Διάγραμμα	Κρυφοί Κόμβοι	MSE			R		
		Training	Validation	Testing	Training	Validation	Testing
1	20	169.5067	158.9639	164.6003	0.997906	0.998019	0.997884
2	25	514.0580	539.6998	479.7194	0.993629	0.993054	0.994042
3	30	872.6055	856.6543	878.7909	0.989064	0.989428	0.989045
4	35	58.4799	65.5309	64.9114	0.999282	0.999163	0.999218
5	50	15.3789	15.0472	16.3239	0.999808	0.999815	0.999795
6	80	19.7691	19.7450	19.7334	0.999755	0.999750	0.999754
7	100	2.4289	2.6535	2.9343	0.999969	0.999966	0.999963

Πίνακας 5: Αποτελέσματα για τον αριθμό των εποχών (Iterations), τον χρόνο εκπαίδευσης (Running time), το Validation Checks και το ριζικό του μέσου τετραγωνικού σφάλματος (RMSE)

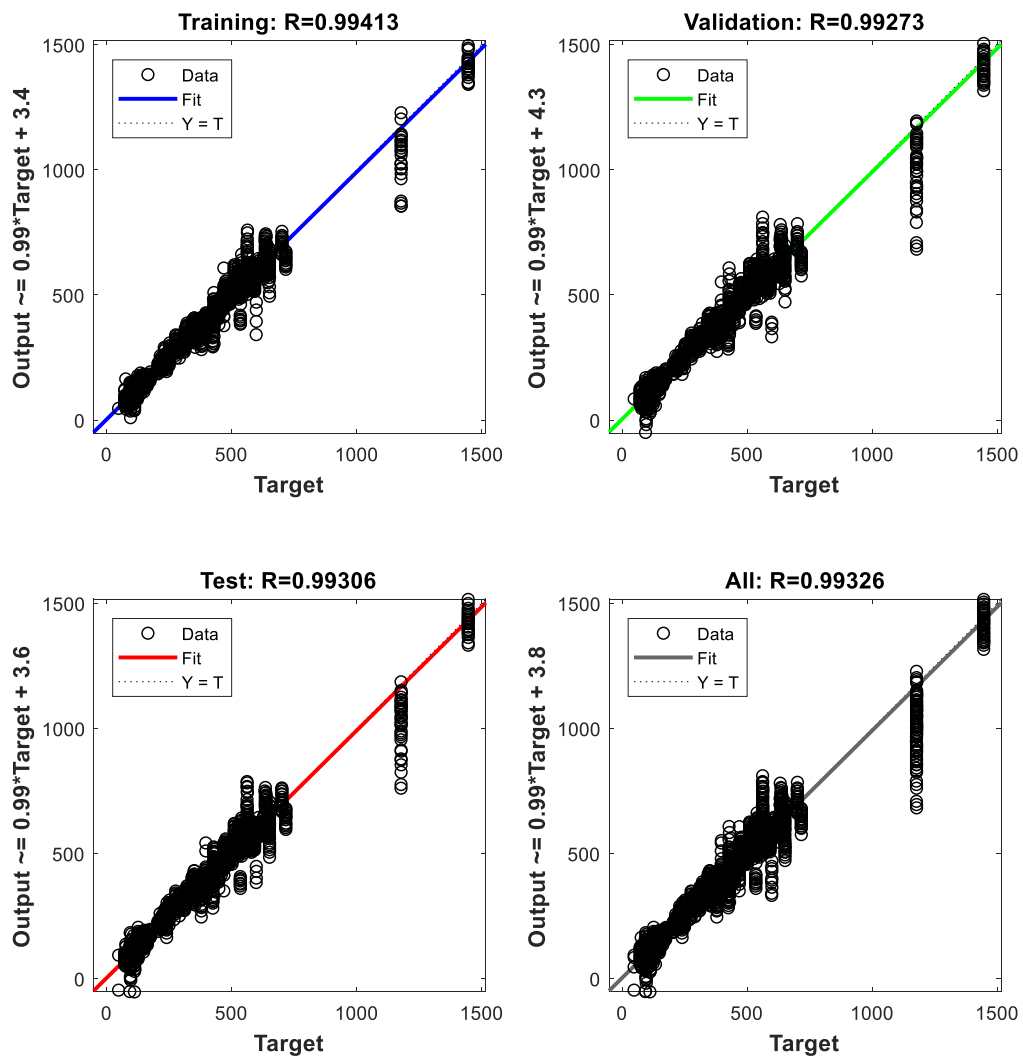
Διάγραμμα	Κρυφοί Κόμβοι	Reached Epochs (Iterations)	Running Time (hh/mm/ss)	Validation Checks (max=6)	RMSE		
					Training	Validation	Testing
1	20	170	00:04:10	6/6	13.0195	12.6081	12.8297
2	25	1000	00:05:34	1/6	22.6729	23.2314	21.9025
3	30	629	00:04:44	6/6	29.5398	29.2687	29.6444
4	35	1000	00:10:01	0/6	7.6472	8.0951	8.0568
5	50	1000	00:24:39	0/6	3.9216	3.8791	4.0403
6	80	1000	01:05:58	0/6	4.4462	4.4435	4.4422
7	100	1000	01:35:40	0/6	1.5585	1.6289	1.7130



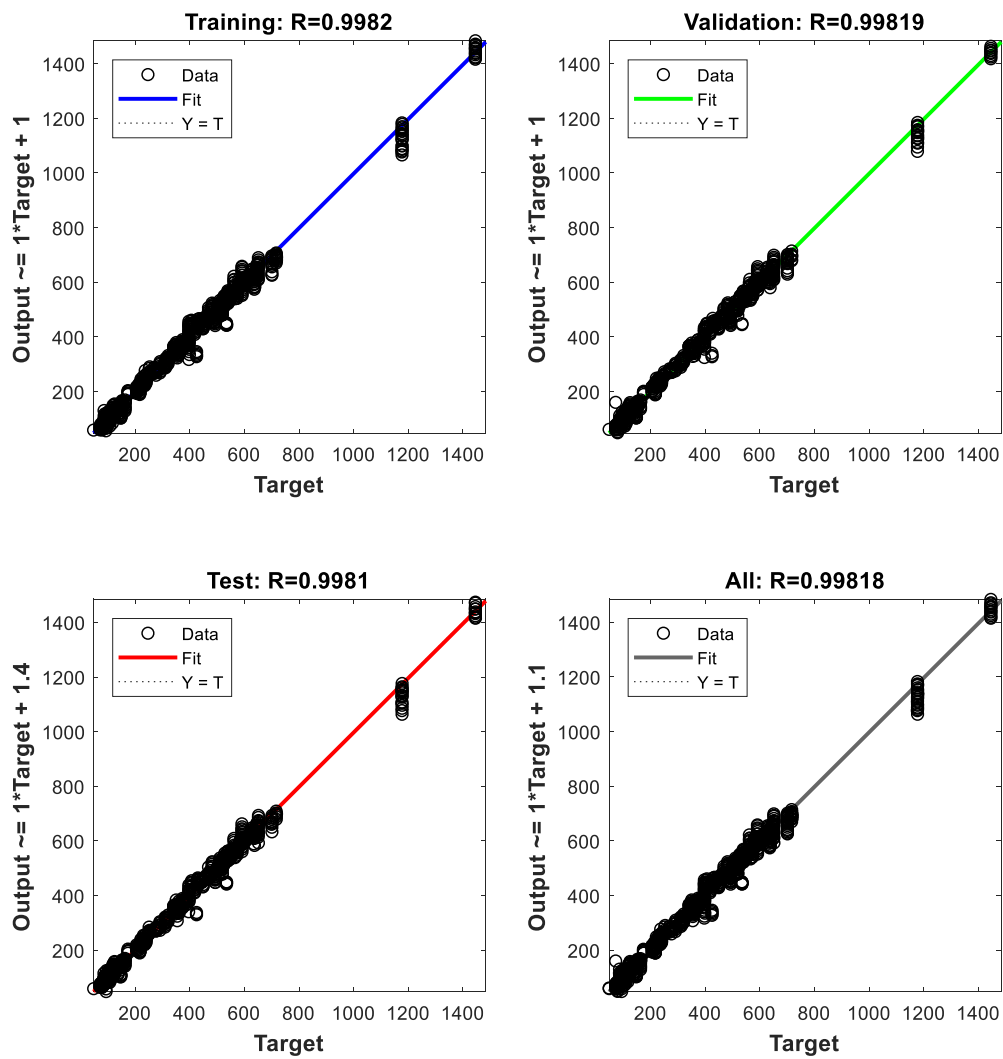
Διάγραμμα 8: Παλινδρόμηση εκπαίδευσης (Training Regression) με 100 κρυφούς κόμβους (Hidden nodes), με ποσοστά 55% Training, 25% Validation & 20% Testing



Διάγραμμα 9: Παλινδρόμηση εκπαίδευσης (Training Regression) με 100 κρυφούς κόμβους (Hidden nodes), με ποσοστά 50% Training, 25% Validation & 25% Testing



Διάγραμμα 10: Παλινδρόμηση εκπαίδευσης (Training Regression) με 100 κρυφούς κόμβους (Hidden nodes), με ποσοστά 30% Training, 35% Validation & 35% Testing



Διάγραμμα 11: Παλινδρόμηση εκπαίδευσης (Training Regression) με 100 κρυφούς κόμβους (Hidden nodes), με ποσοστά 60% Training, 20% Validation & 20% Testing

Πίνακας 6: Απαραίτητοι παράμετροι για την εκπαίδευση με τον αλγόριθμο Levenberg-Marquardt με διαφορετικά Percentages

Διάγραμμα	Κρυφοί Κόμβοι	Percentages			Epochs
		Training	Validation	Testing	
8	100	55%	25%	20%	1000
9	100	50%	25%	25%	1000
10	100	30%	35%	35%	1000
11	100	60%	20%	20%	1000

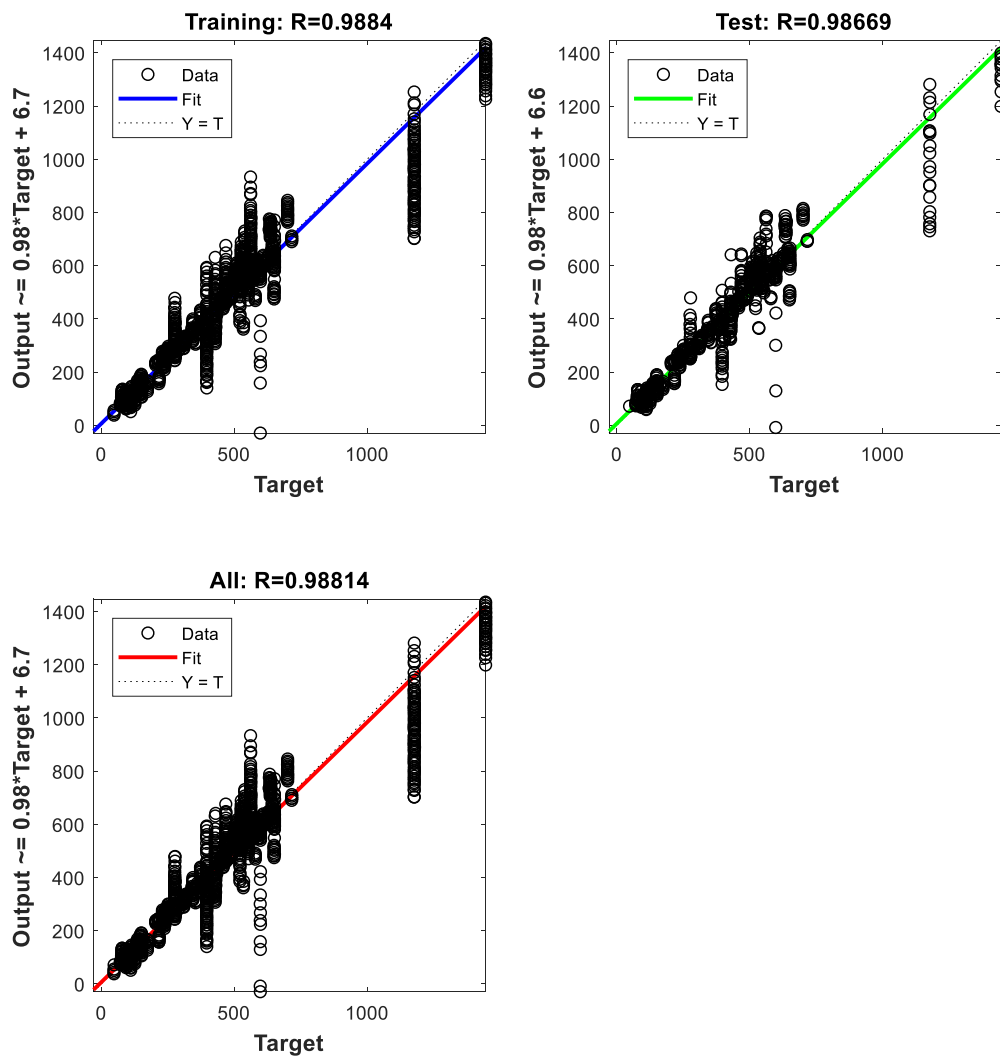
Πίνακας 7: Αποτελέσματα για το μέσο τετραγωνικό σφάλμα (MSE) και τον συντελεστή συσχέτισης (R)

Διάγραμμα	Κρυφοί Κόμβοι	MSE			R		
		Training	Validation	Testing	Training	Validation	Testing
8	100	1.3241	1.5303	1.47443	0.999983	0.999980	0.999981
9	100	25.6782	27.6421	28.60032	0.999681	0.999644	0.999655
10	100	467.9077	589.4731	554.14956	0.994132	0.992725	0.993061
11	100	143.7785	146.6999	154.67930	0.998198	0.998194	0.998097

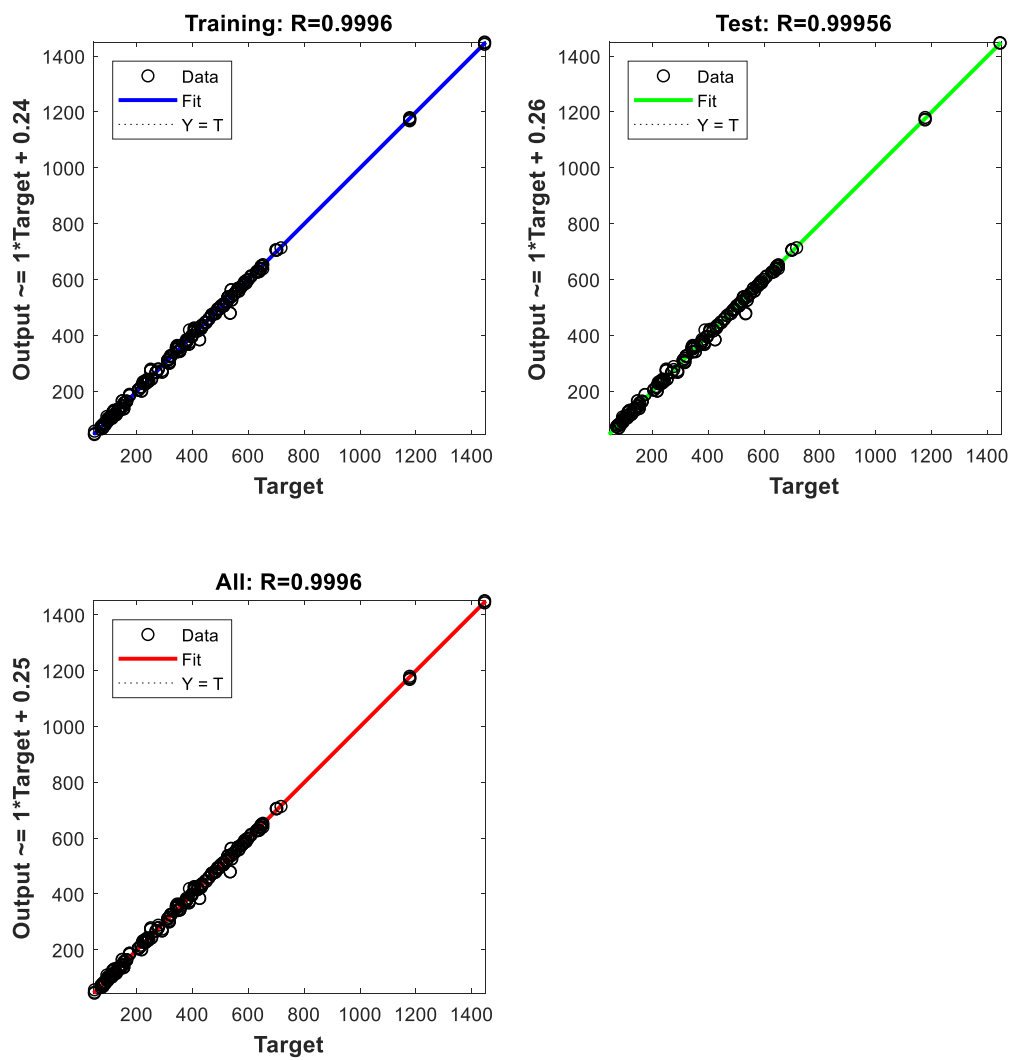
Πίνακας 8: Αποτελέσματα για τον αριθμό των εποχών (Iterations), τον χρόνο εκπαίδευσης (Running time), του Validation Check και της τετραγωνικής ρίζας του μέσου τετραγωνικού σφάλματος (RMSE)

Διάγραμμα	Κρυφοί Κόμβοι	Reached Epochs (Iterations)	Running Time (hh/mm/ss)	Validation Checks (max=6)	RMSE		
					Training	Validation	Testing
8	100	1000	01:34:16	0/6	1.1507	1.2371	1.2143
9	100	1000	01:35:16	0/6	5.0674	5.2576	5.3479
10	100	1000	01:34:10	0/6	21.6312	24.2791	23.5404
11	100	1000	01:34:10	0/6	11.9908	12.1120	12.4370

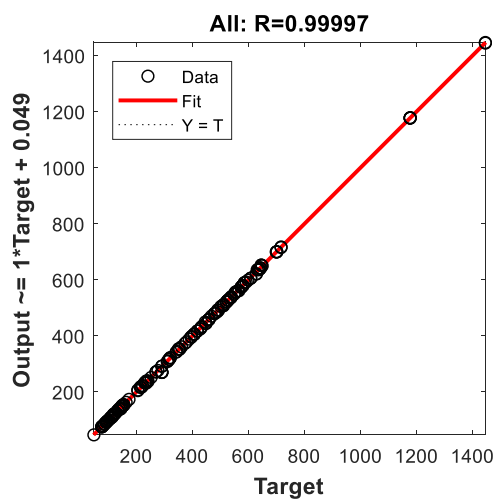
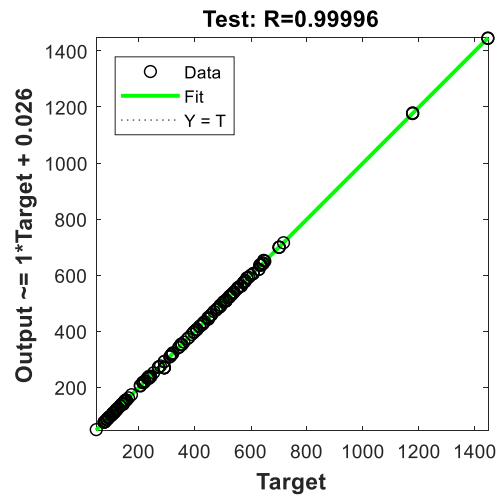
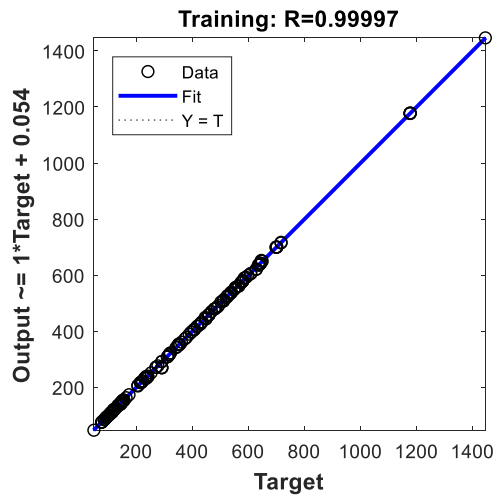
4.2 Χρήση αλγορίθμου Bayesian Regularization



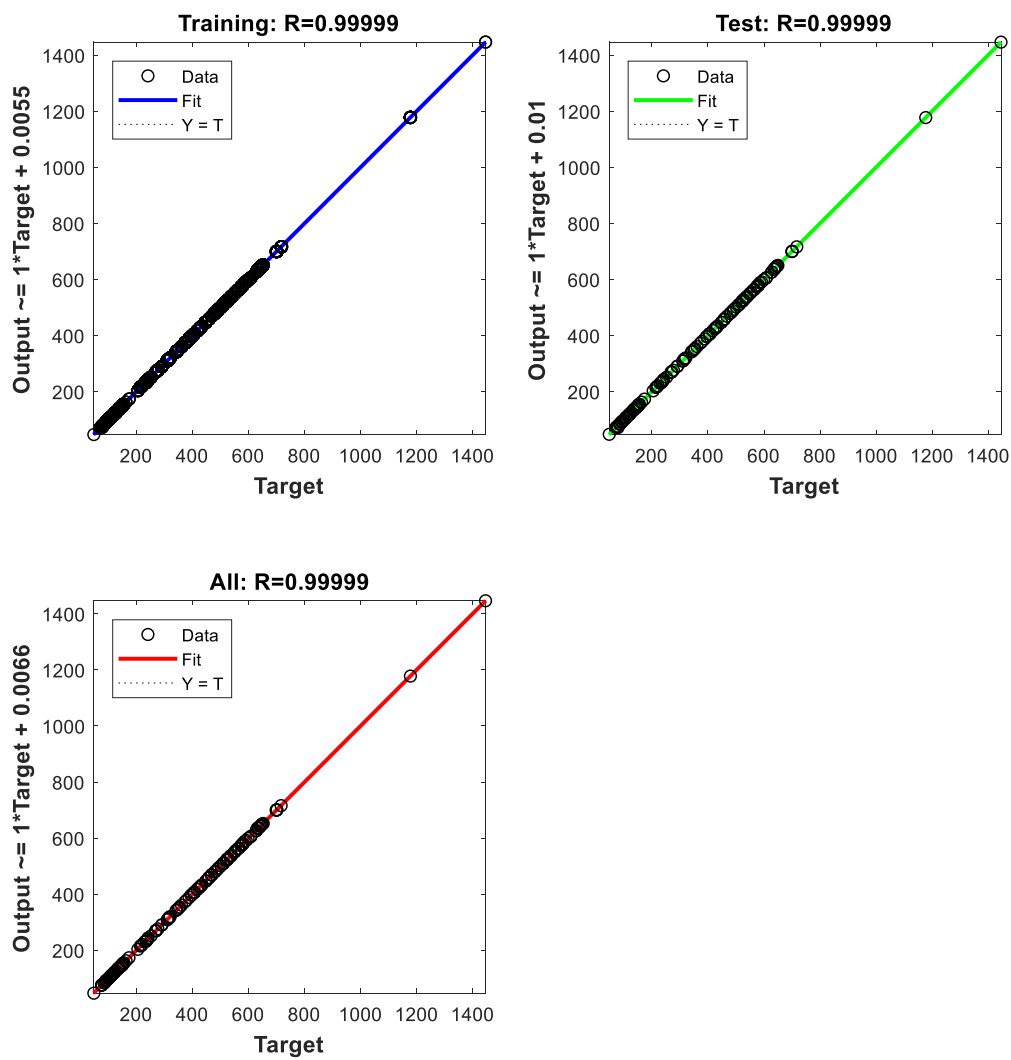
Διάγραμμα 12: Παλινδρόμηση εκπαίδευσης (Training Regression) με 10 κρυφούς κόμβους (Hidden nodes), με ποσοστά 70% Training, 15% Validation & 15% Testing



Διάγραμμα 13: Παλινδρόμηση εκπαίδευσης (Training Regression) με 30 κρυφούς κόμβους (Hidden nodes), με ποσοστά 70% Training, 10% Validation & 20% Testing



Διάγραμμα 14: Παλινδρόμηση εκπαίδευσης (Training Regression) με 80 κρυφούς κόμβους (Hidden nodes), με ποσοστά 70% Training, 10% Validation & 20% Testing



Διάγραμμα 15: Παλινδρόμηση εκπαίδευσης (Training Regression) με 100 κρυφούς κόμβους (Hidden nodes), με ποσοστά 70% Training, 5% Validation & 25% Testing

Πίνακας 9: Απαραίτητοι παράμετροι για την εκπαίδευση με τον αλγόριθμο Bayesian Regularization και με διαφορετικά Percentages

Διάγραμμα	Κρυφοί Κόμβοι	Percentages			Epochs
		Training	Validation	Testing	
12	10	70%	15%	15%	1000
13	30	70%	10%	20%	1000
14	80	70%	10%	20%	1000
15	100	70%	5%	25%	1000

Πίνακας 10: Αποτελέσματα για το μέσο τετραγωνικό σφάλμα (MSE) και τον συντελεστή συσχέτισης (R)

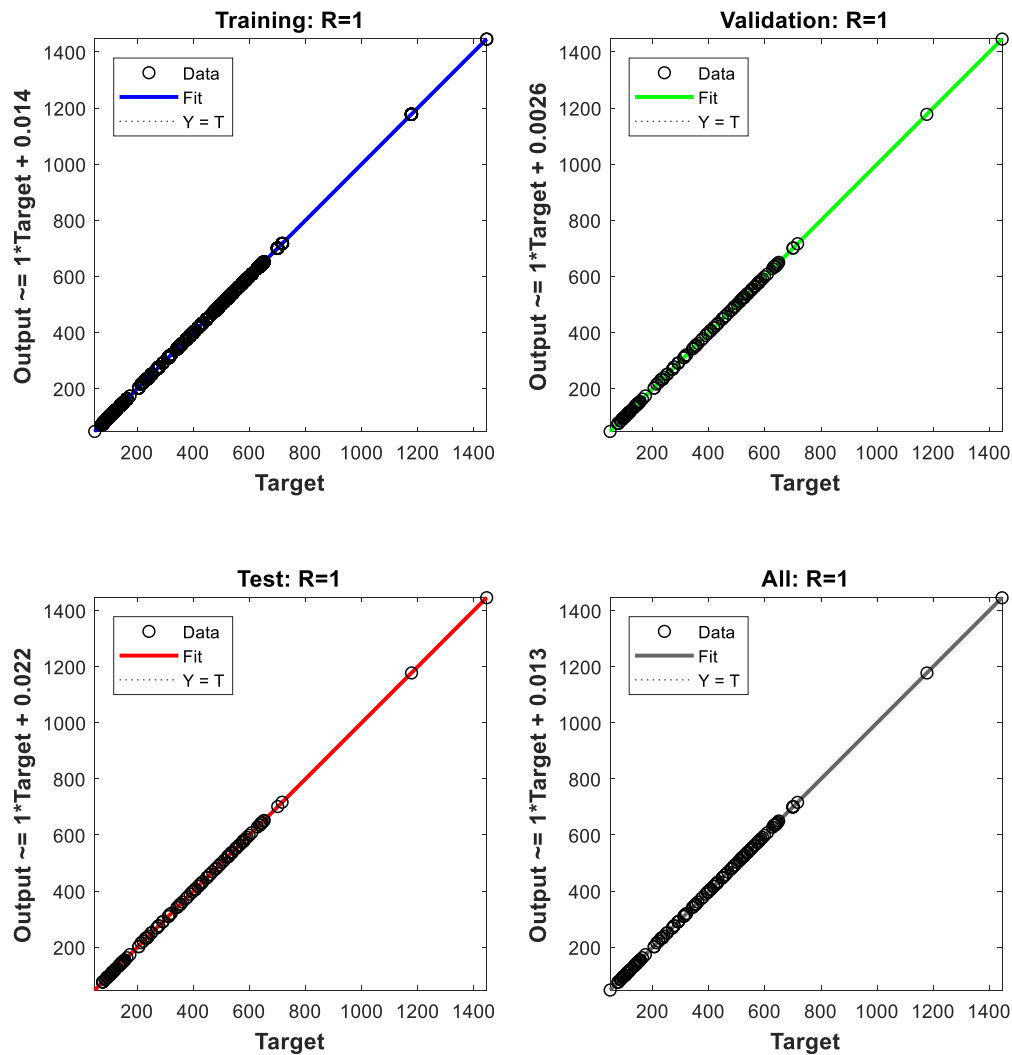
Διάγραμμα	Κρυφοί Κόμβοι	MSE			R		
		Training	Validation	Testing	Training	Validation	Testing
12	10	926.6624	-	1071.9875	0.988403	-	0.986686
13	30	31.9071	-	34.9472	0.999604	-	0.999562
14	80	2.6633	-	3.3785	0.999966	-	0.999958
15	100	0.4919	-	0.4924	0.999993	-	0.999993

Πίνακας 11: Αποτελέσματα για τον αριθμό των εποχών (Iterations), τον χρόνο εκπαίδευσης (Running time) και της τετραγωνικής ρίζας του μέσου τετραγωνικού σφάλματος (RMSE)

Διάγραμμα	Κρυφοί Κόμβοι	Reached Epochs (Iterations)	Running Time (hh/mm/ss)	RMSE		
				Training	Validation	Testing
12	10	1000	00:05:09	30.4411	-	32.7412
13	30	1000	00:12:40	5.6486	-	5.9116
14	80	1000	01:01:37	1.6320	-	1.8381
15	100	1000	05:30:53	0.7014	-	0.7017

4.3 Χρήση της εντολής `nntool` στο Command Window του επιστημονικού προγράμματος της Matlab

4.3.1 Χρήση αλγόριθμου Levenberg-Marquardt



Διάγραμμα 16: Παλινδρόμηση εκπαίδευσης (Training Regression) με 100 κρυφούς κόμβους (Hidden nodes)

Πίνακας 12: Απαραίτητοι παράμετροι για την εκπαίδευση με τον αλγόριθμο Levenberg-Marquardt με χρήση της εντολής nntool

Διάγραμμα	Κρυφοί Κόμβοι	Percentages			Epochs
		Training	Validation	Testing	
16	100	-	-	-	2000

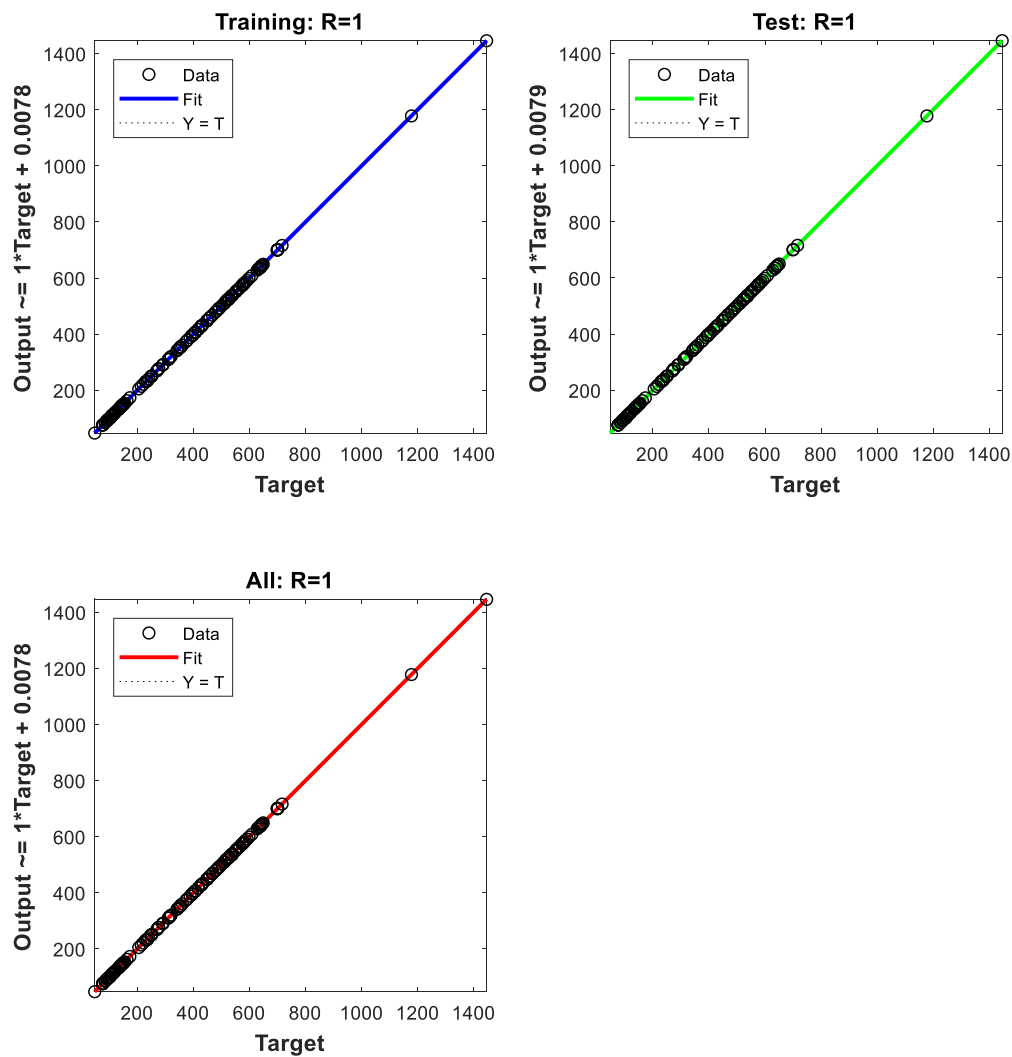
Πίνακας 13: Αποτελέσματα για το μέσο τετραγωνικό σφάλμα (MSE) και τον συντελεστή συσχέτισης (R)

Διάγραμμα	Κρυφοί Κόμβοι	MSE			R		
		Training	Validation	Testing	Training	Validation	Testing
16	100	0.304	-	-	1	1	1

Πίνακας 14: Αποτελέσματα για τον αριθμό των εποχών (Iterations), τον χρόνο εκπαίδευσης (Running time), του Validation Check και της τετραγωνικής ρίζας του μέσου τετραγωνικού σφάλματος (RMSE)

Διάγραμμα	Κρυφοί Κόμβοι	Reached Epochs (Iterations)	Running Time (hh/mm/ss)	Validation Checks	RMSE		
					Training	Validation	Testing
16	100	2000	03:07:39	0/60	0.551	-	-

4.3.2 Χρήση αλγορίθμου Bayesian Regularization



Διάγραμμα 17: Παλινδρόμηση εκπαίδευσης (Training Regression) με 100 κρυφούς κόμβους (Hidden nodes)

Πίνακας 15: Απαραίτητοι παράμετροι για την εκπαίδευση με τον αλγόριθμο Bayesian Regularization με χρήση της εντολής nntool

Γράφημα	Κρυφοί Κόμβοι	Percentages			Epochs
		Training	Validation	Testing	
17	100	-	-	-	2000

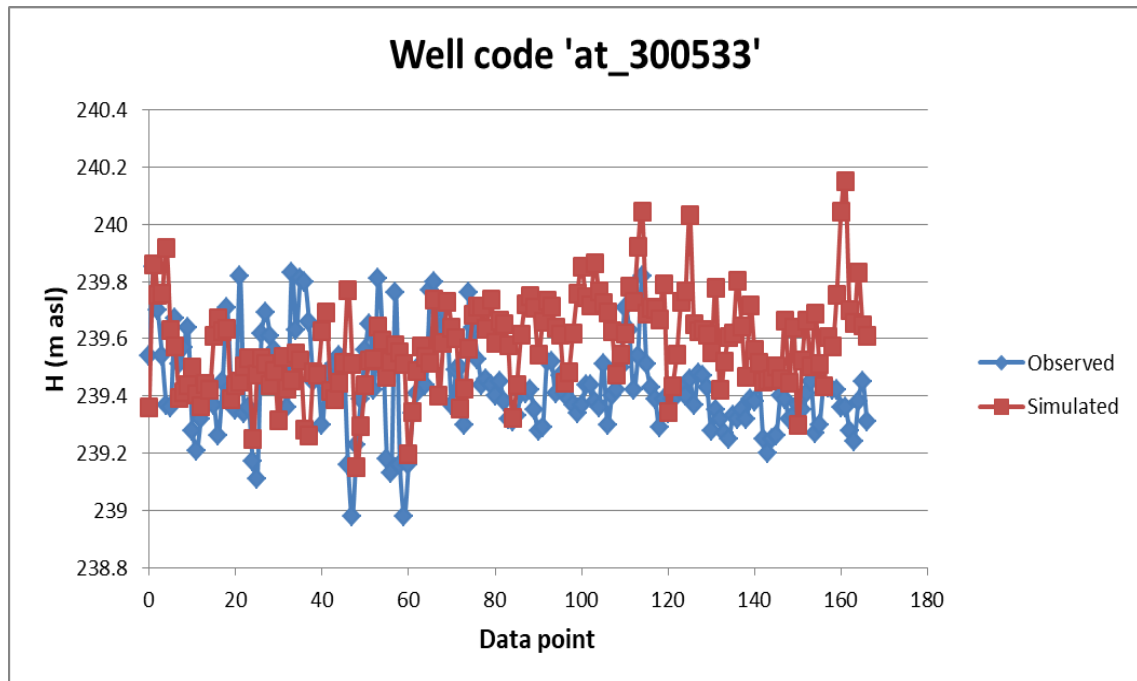
Πίνακας 16: Αποτελέσματα για το μέσο τετραγωνικό σφάλμα (MSE) και τον συντελεστή συσχέτισης (R)

Γράφημα	Κρυφοί Κόμβοι	MSE			R		
		Training	Validation	Testing	Training	Validation	Testing
17	100	0.208	-	-	1	-	1

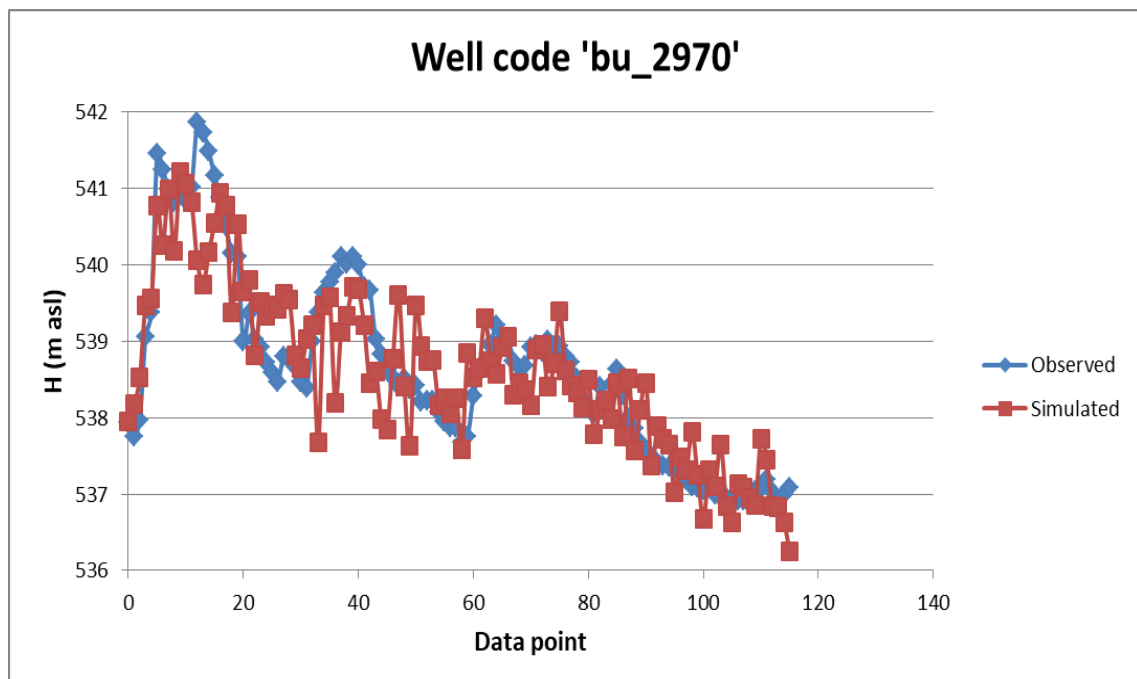
Πίνακας 17: Αποτελέσματα για τον αριθμό των εποχών (Iterations), τον χρόνο εκπαίδευσης (Running time), και της τετραγωνικής ρίζας του μέσου τετραγωνικού σφάλματος (RMSE)

Γράφημα	Κρυφοί Κόμβοι	Reached Epochs (Iterations)	Running Time (hh/mm/ss)	RMSE		
				Training	Validation	Testing
17	100	2000	06:20:44	0.456	-	-

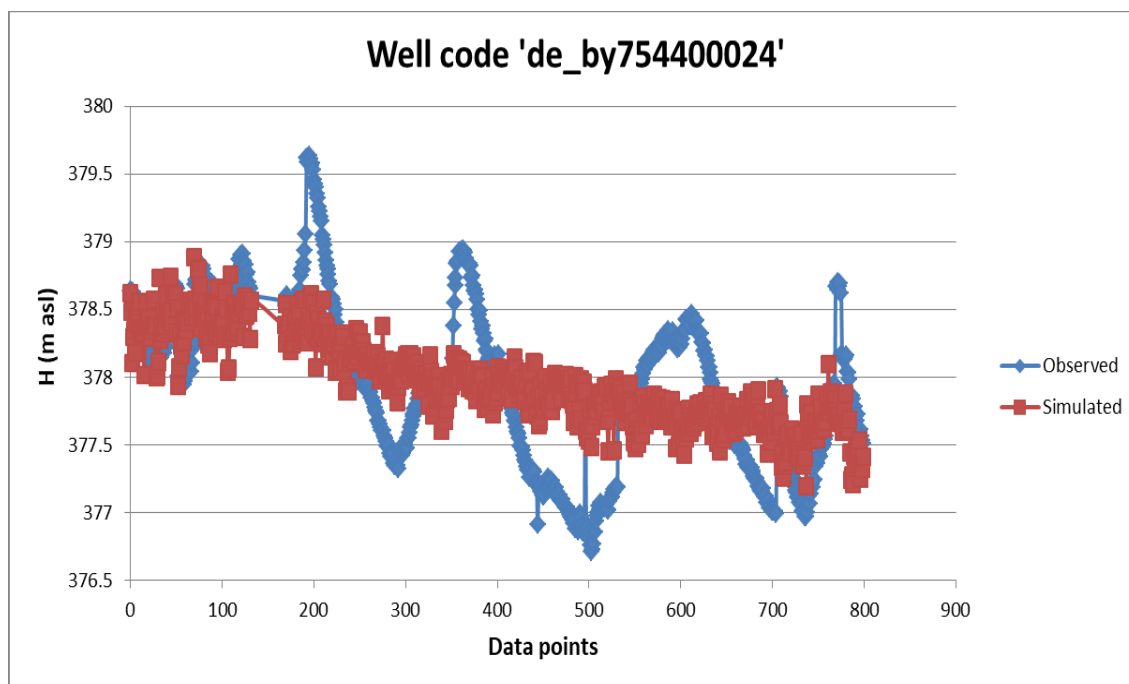
4.4 Γραφική απεικόνιση σύγκρισης παρατηρούμενων-προσομοιωμένων τιμών υπόγειας στάθμης για κάποιο πηγάδι ανά χώρα



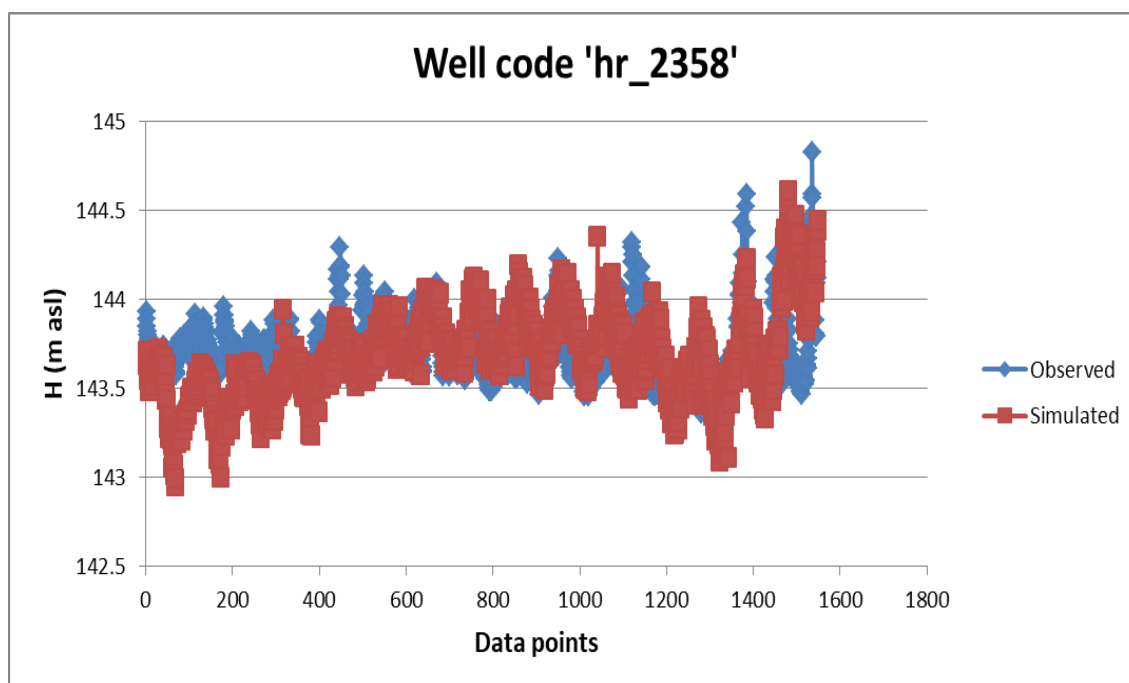
Γράφημα 1: Σύγκριση μεταξύ παρατηρούμενης-προσομοιωμένης υπόγειας στάθμης νερού σε ένα πηγάδι της Αυστρίας



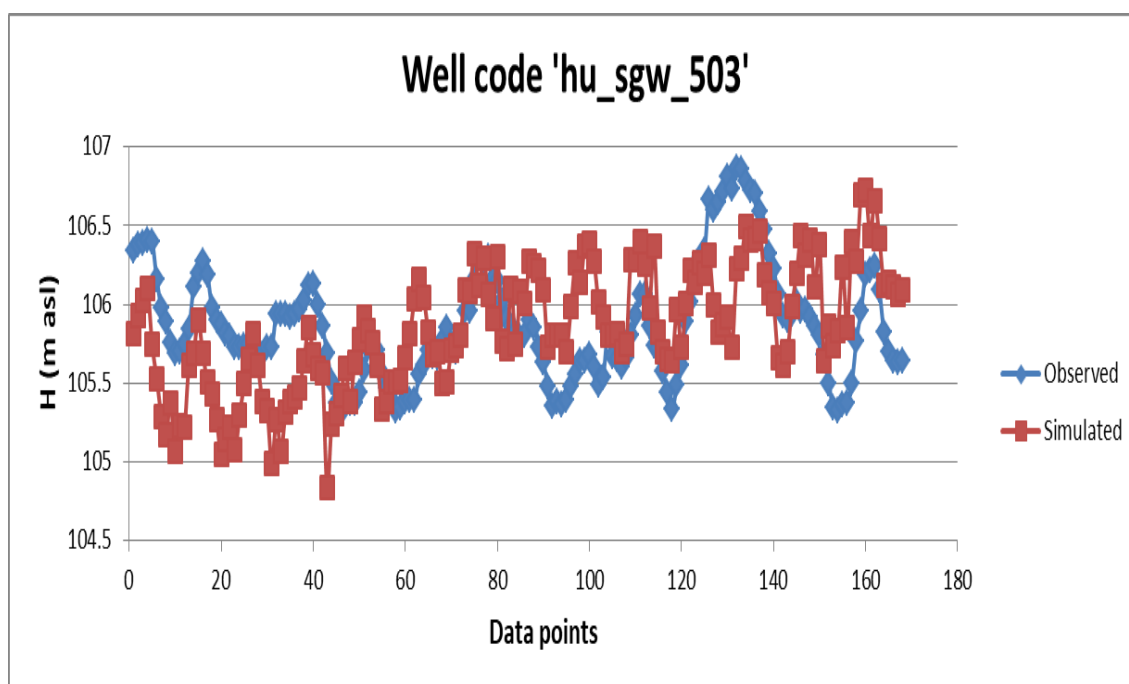
Γράφημα 2: Σύγκριση μεταξύ παρατηρούμενης-προσομοιωμένης υπόγειας στάθμης νερού σε ένα πηγάδι της Βουλγαρίας



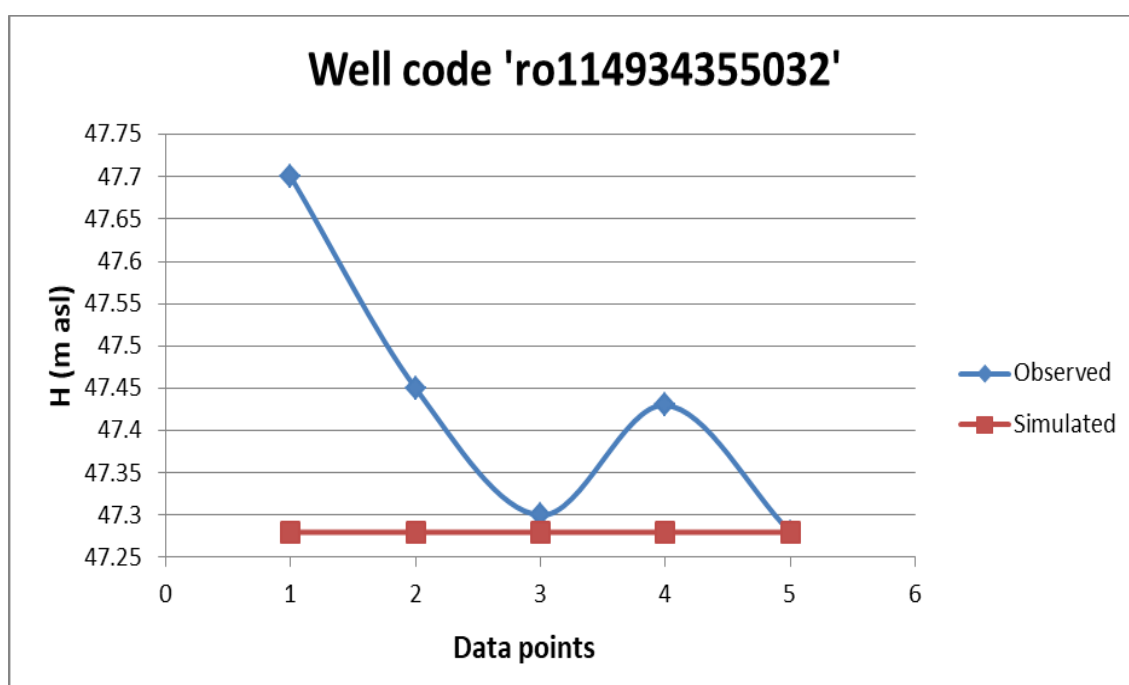
Γράφημα 3: Σύγκριση μεταξύ παρατηρούμενης-προσομοιωμένης υπόγειας στάθμης νερού σε ένα πηγάδι της Γερμανίας



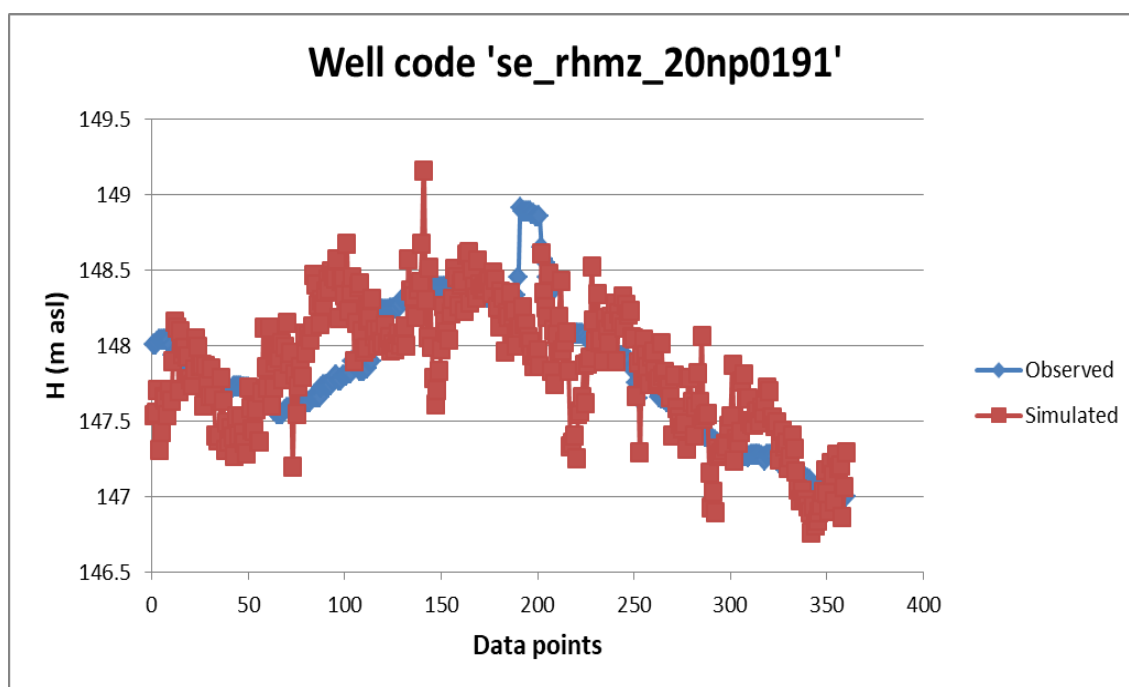
Γράφημα 4: Σύγκριση μεταξύ παρατηρούμενης-προσομοιωμένης υπόγειας στάθμης νερού σε ένα πηγάδι της Κροατίας



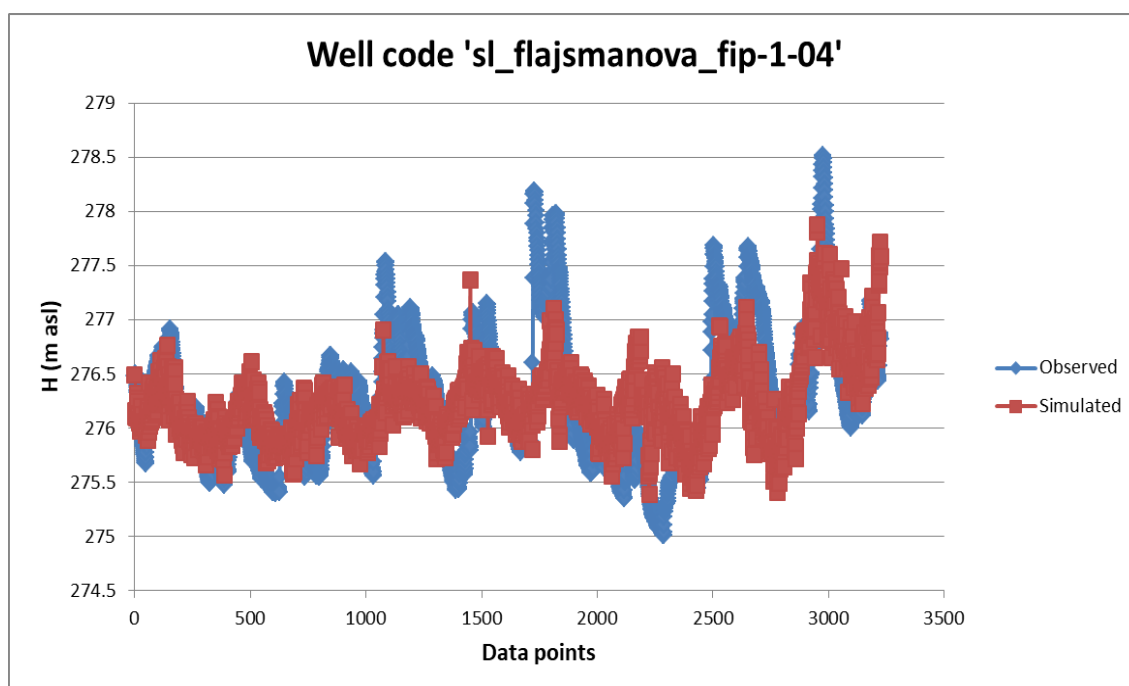
Γράφημα 5: Σύγκριση μεταξύ παρατηρούμενης-προσομοιωμένης υπόγειας στάθμης νερού σε ένα πηγάδι της Ουγγαρίας



Γράφημα 6: Σύγκριση μεταξύ παρατηρούμενης-προσομοιωμένης υπόγειας στάθμης νερού σε ένα πηγάδι της Ρουμανίας



Γράφημα 7: Σύγκριση μεταξύ παρατηρούμενης-προσομοιωμένης υπόγειας στάθμης νερού σε ένα πηγάδι της Σερβίας

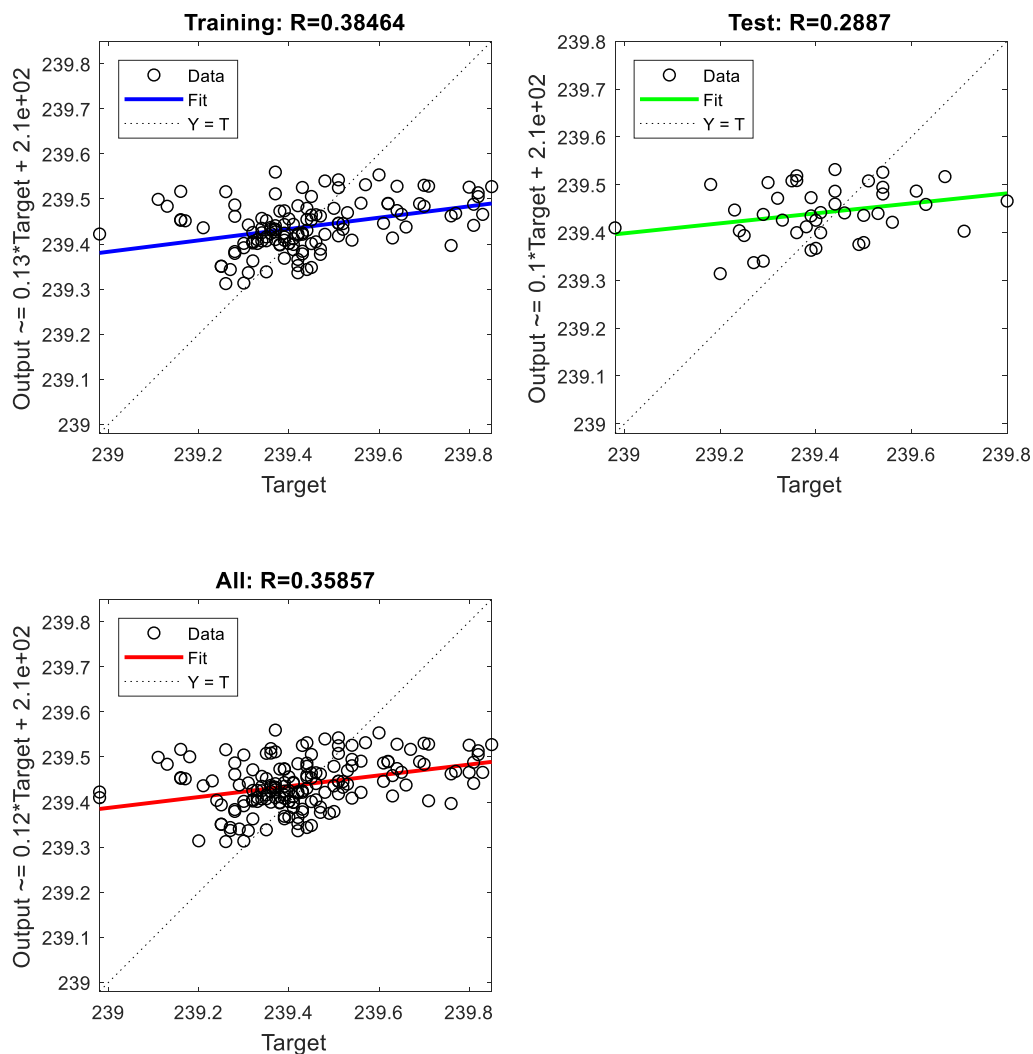


Γράφημα 8: Σύγκριση μεταξύ παρατηρούμενης-προσομοιωμένης υπόγειας στάθμης νερού σε ένα πηγάδι της Σλοβενίας

4.5 Εκπαίδευση νευρωνικού δικτύου μόνο για ένα πηγάδι

Για να ελεγχθεί το κατά πόσο η εκπαίδευση ενός νευρωνικού με τα δεδομένα ενός μόνο πηγαδιού θα έδινε καλύτερα αποτελέσματα σε σχέση με την εκπαίδευση του νευρωνικού με τα δεδομένα όλων των πηγαδιών όπως έχει γίνει μέχρι αυτό το σημείο, επιλέχθηκε ένα τυχαίο πηγάδι στην Αυστρία.

Στην **Ενότητα 4.5** εκπαιδεύτηκε ένα νευρωνικό δίκτυο για το πηγάδι στην Αυστρία με κωδικό πηγαδιού 'at_300533'. Ως δεδομένα εισόδου παραμένουν τα ίδια με αυτά που προαναφέραμε στην μεθοδολογία τα οποία αντιστοιχούν στο συγκεκριμένο πηγάδι με μια μικρή διαφορά. Αφαιρέσαμε την 9^η & 10^η στήλη που αντιστοιχούν σε δεδομένα ως προς τις συντεταγμένες του πηγαδιού, διότι για ένα πηγάδι αυτές οι τιμές παραμένουν ίδιες και δεν είναι απαραίτητες για την εκπαίδευση. Το διάγραμμα στόχου περιέχει τις αντίστοιχες τιμές για τις πραγματικές στάθμες για το συγκεκριμένο πηγάδι.



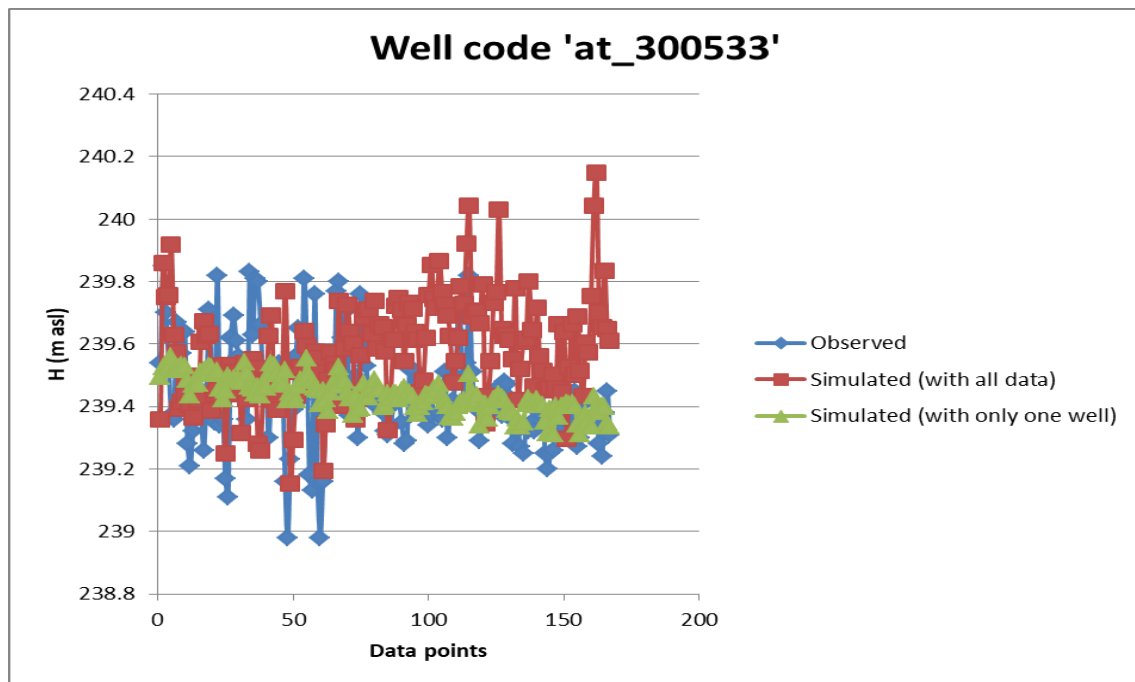
Διάγραμμα 18: Παλινδρόμηση εκπαίδευσης (Training Regression) με 10 κρυφούς κόμβους (Hidden nodes), με ποσοστά 70% Training, 5% Validation & 25% Testing για ένα μόνο πηγάδι

Πίνακας 18: Αποτελέσματα από την εκπαίδευση νευρωνικού για ένα μόνο πηγάδι με χρήση αλγορίθμου BR

Διάγραμμα	Κρυφοί Κόμβοι	Percentages			Epochs		
		Training	Validation	Testing			
18	10	70%	5%	25%	1000		
//		MSE			R ²		
18	10	Training	Validation	Testing	Training	Validation	Testing
		0.02339	-	0.02227	0.38464	-	0.2887
//		Reached Epochs (Iterations)	Running Time (hh/mm/ss)	RMSE			
				Training	Validation	Testing	
18	10	1000	00:00:50	0.153	-	0.149	

*Αξίζει να σημειωθεί ότι δεν εκπαιδεύτηκαν νευρωνικά δίκτυα με την χρήση του αλγορίθμου LM διότι η παράμετρος *Validation checks* έφτανε στο μέγιστο για πολύ μικρό αριθμό εποχών άρα είχαμε *Early stopping* διότι η γενίκευση δεν βελτιωνόταν*

Παρακάτω στο **Γράφημα 9** παρουσιάζονται τα αποτελέσματα από τα οποία τα νευρωνικά εκπαιδεύτηκαν για όλο το σετ δεδομένων και για ένα μόνο πηγάδι.



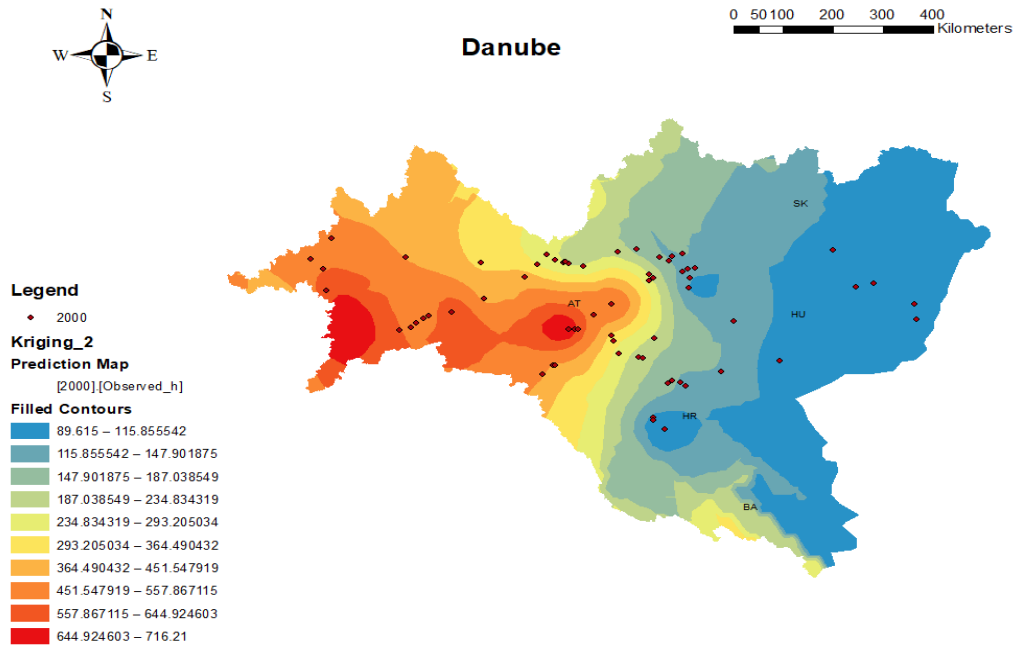
Γράφημα 9: Σύγκριση μεταξύ παρατηρούμενης (μπλε γραμμή)-προσομοιωμένης με βάση όλα τα δεδομένα (κόκκινη γραμμή)-προσομοιωμένης με βάση τα δεδομένα ενός πηγαδιού μόνο (πράσινη γραμμή) υπόγειας στάθμης για ένα πηγάδι της Αυστρίας

Πίνακας 19: Αποτελέσματα σφάλματος RMSE και συντελεστή συσχέτισης από το ΤΝΔ που εκπαιδεύτηκε για ένα πηγάδι παρατήρησης και από το ΤΝΔ που εκπαιδεύτηκε για όλα τα πηγάδια παρατήρησης μαζί

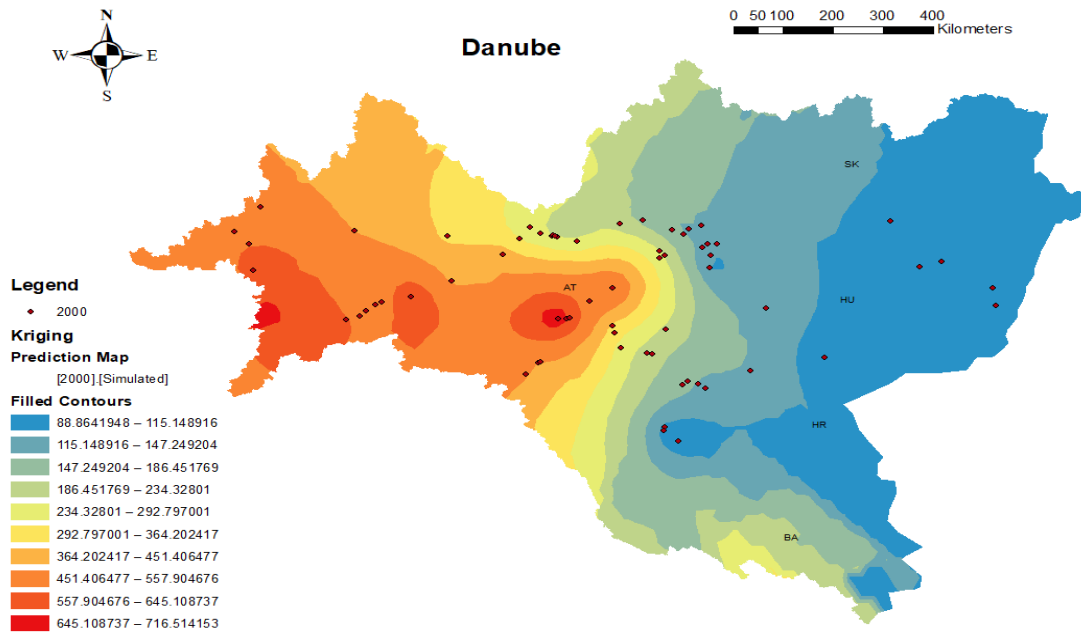
Results	RMSE	R
		All
Green line	0.153	0.35857
Red line	0.264	0.10191

Κεφάλαιο 5

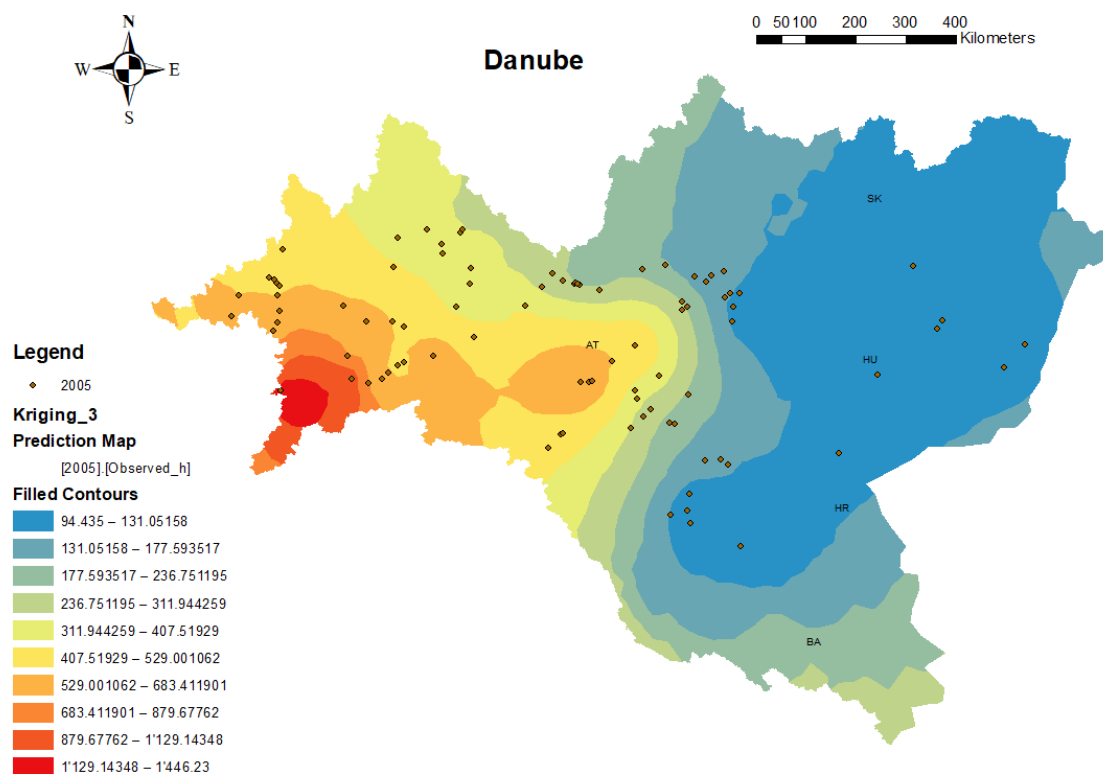
5.1 Χωρική Παρεμβολή με την τεχνική του Kriging για την δημιουργία χαρτών με τις ισοϋψείς μεταξύ παρατηρημένων και προσομοιωμένων τιμών της υπόγειας στάθμης για τις χρονιές 2000, 2005, 2010 και 2012.



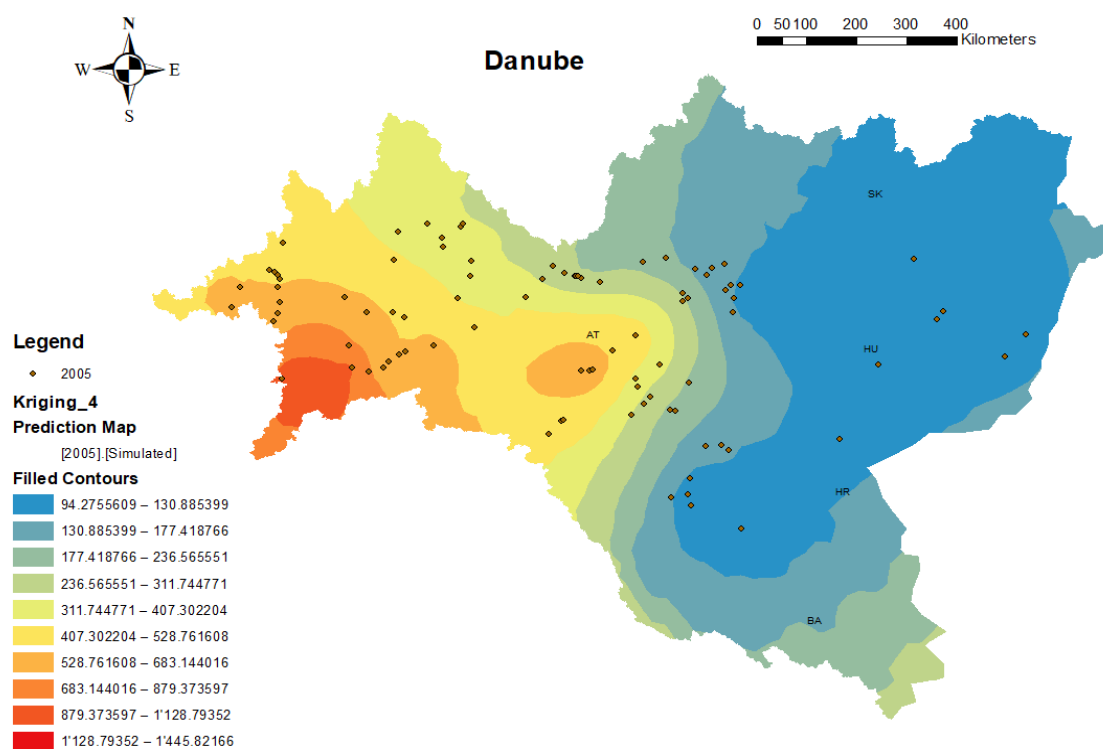
Χάρτης 1: Χωρική παρεμβολή των παρατηρημένων τιμών της υπόγειας στάθμης για το έτος 2000 [1]



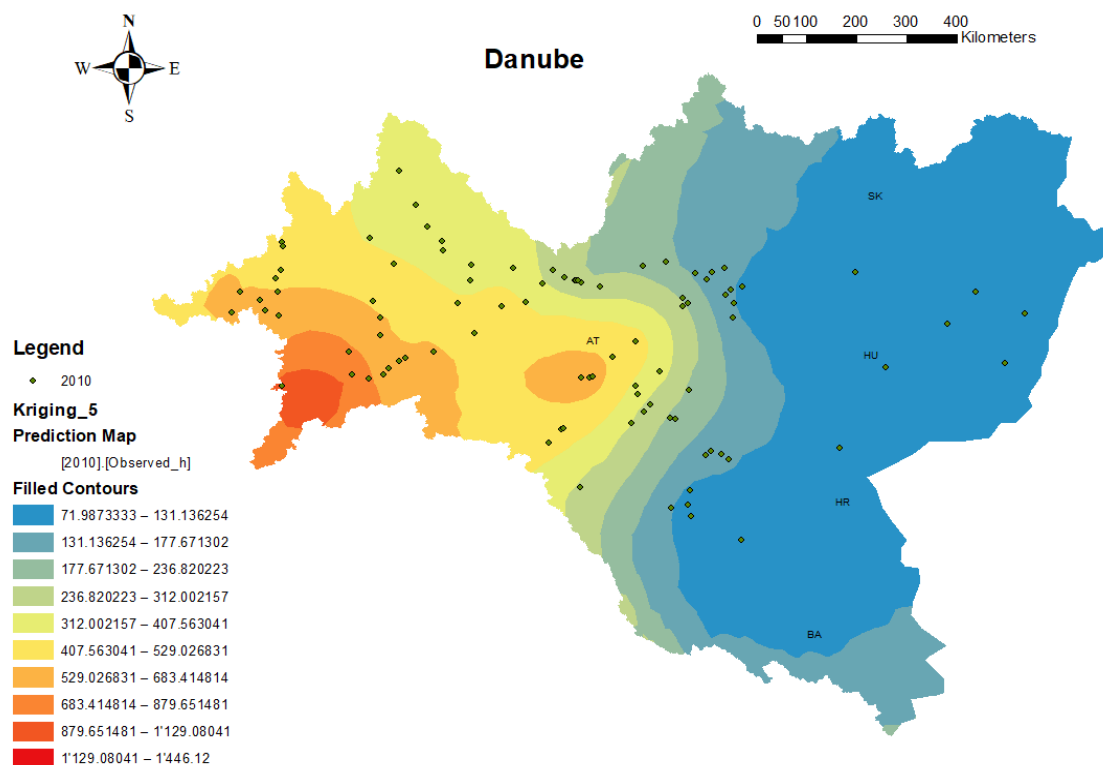
Χάρτης 2: Χωρική παρεμβολή των προσομοιωμένων τιμών της υπόγειας στάθμης για το έτος 2000 [2]



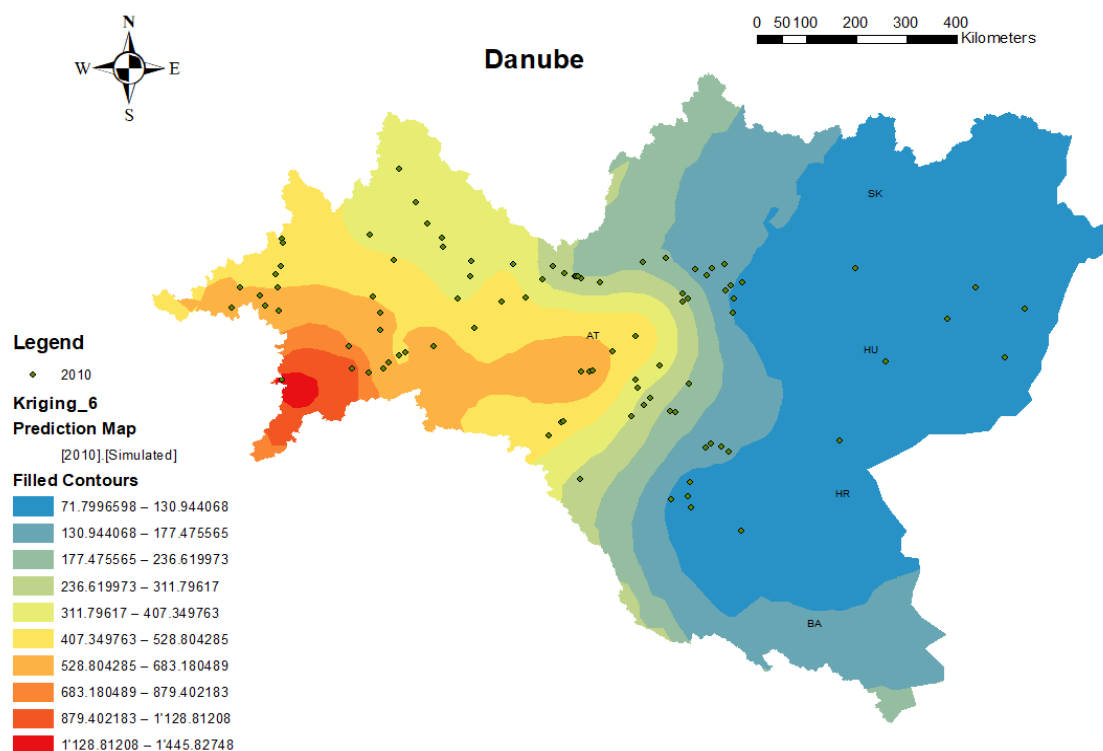
Χάρτης 3: Χωρική παρεμβολή των παρατηρημένων τιμών της υπόγειας στάθμης για το έτος 2005 ^[3]



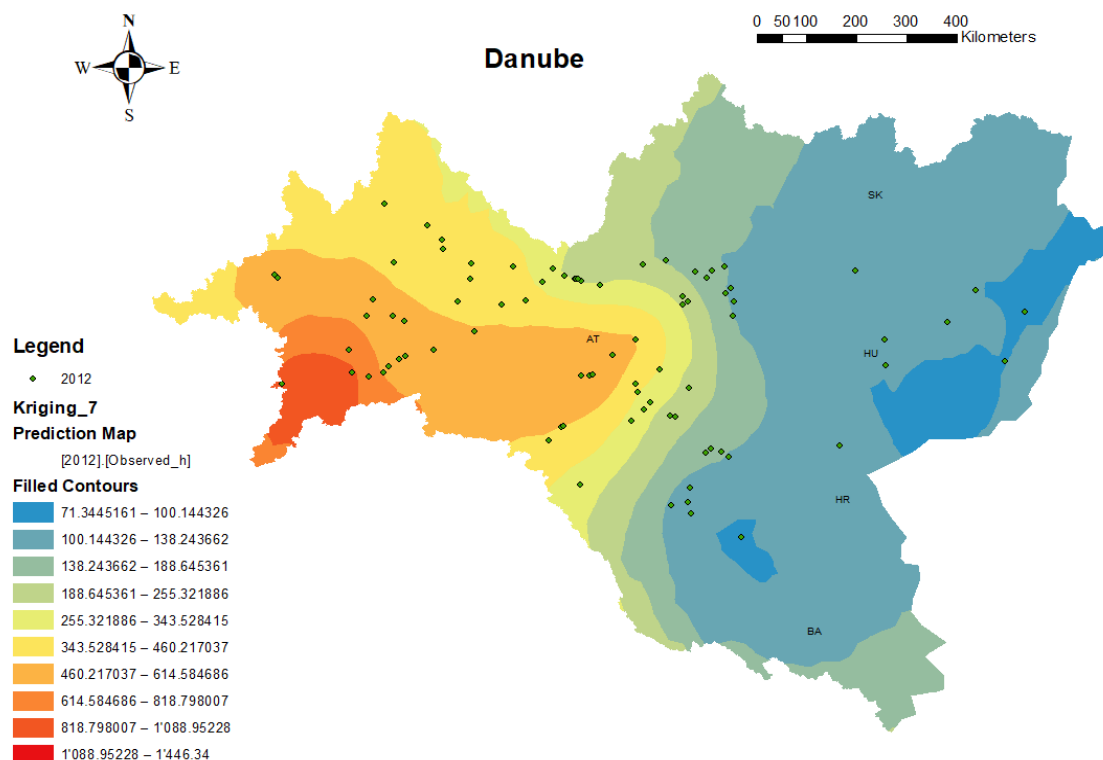
Χάρτης 4: Χωρική παρεμβολή των προσομοιωμένων τιμών της υπόγειας στάθμης για το έτος 2005 ^[4]



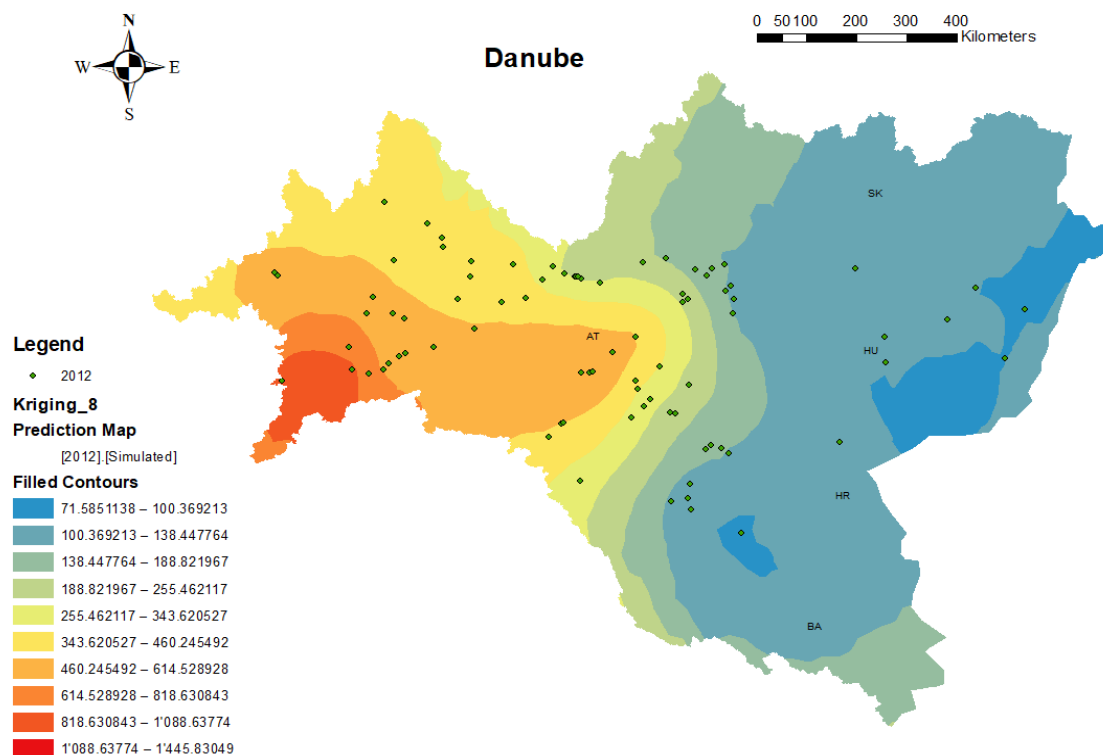
Χάρτης 5: Χωρική παρεμβολή των παρατηρημένων τιμών της υπόγειας στάθμης για το έτος 2010 ^[5]



Χάρτης 6: Χωρική παρεμβολή των προσομοιωμένων τιμών της υπόγειας στάθμης για το έτος 2010 ^[6]



Χάρτης 7: Χωρική παρεμβολή των παρατηρημένων τιμών της υπόγειας στάθμης για το έτος 2012 ^[7]



Χάρτης 8: Χωρική παρεμβολή των προσομοιωμένων τιμών της υπόγειας στάθμης για το έτος 2012 ^[8]

[1] & [2] : Μεταξύ του Χάρτη 1 που είναι για τις παρατηρημένες τιμές της υπόγειας στάθμης και του Χάρτη 2 που είναι για τις προσομοιωμένες τιμές που έχουν προκύψει από το τεχνητό νευρωνικό δίκτυο για το έτος 2000. Παρατηρούμε ότι στην δυτική πλευρά του Χάρτη 1 η ισοϋψής με τις μεγαλύτερες τιμές της υπόγειας στάθμης καλύπτουν πιο μεγάλη έκταση από ότι στον Χάρτη 2 που σημαίνει ότι το μοντέλο υποεκτιμά τις τιμές της υπόγειας στάθμης. Επίσης στα βορειοανατολικά του Χάρτη 1 καλύπτεται λίγο μεγαλύτερη έκταση για την ισοϋψή των 89.615-115.855542 m σε σχέση με τον Χάρτη 2 που είναι πολύ μικρότερη η έκταση όπου το μοντέλο υποεκτιμά τις τιμές. Επίσης στα νοτιοανατολικά του Χάρτη 1 η ισοϋψείς με τις μικρότερες στάθμες καλύπτουν μεγαλύτερη έκταση που σημαίνει ότι το μοντέλο υπερεκτιμά τις τιμές της υπόγειας στάθμης.

[3] & [4] : Μεταξύ του Χάρτη 3 που είναι για τις παρατηρημένες τιμές της υπόγειας στάθμης και του Χάρτη 4 που είναι για τις προσομοιωμένες τιμές που έχουν προκύψει από το τεχνητό νευρωνικό δίκτυο για το έτος 2005. Παρατηρούμε ότι στα βορειοδυτικά του Χάρτη 3 η ισοϋψής των 529.001062-683.411901 m καλύπτει μεγαλύτερη έκταση σε σχέση με την περίπου αντίστοιχη ισοϋψής του Χάρτη 4 (υποεκτιμά). Καθώς επίσης η ισοϋψής για τις μεγαλύτερες τιμές της υπόγειας στάθμης υπάρχει μόνο στον Χάρτη 3 για τις παρατηρημένες τιμές ενώ στον Χάρτη 4 δεν υπάρχει που σημαίνει ότι το μοντέλο υποεκτιμά τις τιμές της υπόγειας στάθμης.

[5] & [6] : Μεταξύ του Χάρτη 5 που είναι για τις παρατηρημένες τιμές της υπόγειας στάθμης και του Χάρτη 6 που είναι για τις προσομοιωμένες τιμές που έχουν προκύψει από το τεχνητό νευρωνικό δίκτυο για το έτος 2010. Παρατηρούμε ότι στα βορειοδυτικά του Χάρτη 6 η ισοϋψής των 528.804285-683.180489 m καλύπτει μεγαλύτερη έκταση σε σχέση με την περίπου αντίστοιχη ισοϋψής του Χάρτη 5 (υπερεκτιμά). Καθώς επίσης η ισοϋψής για τις μεγαλύτερες τιμές της υπόγειας στάθμης υπάρχει μόνο στον Χάρτη 6 για τις προσομοιωμένες τιμές που σημαίνει ότι το μοντέλο υπερεκτιμά τις τιμές της υπόγειας στάθμης.

[7] & [8] : Μεταξύ του Χάρτη 7 που είναι για τις παρατηρημένες τιμές της υπόγειας στάθμης και του Χάρτη 8 που είναι για τις προσομοιωμένες τιμές που έχουν προκύψει από το τεχνητό νευρωνικό δίκτυο για το έτος 2012. Στην συγκεκριμένη περίπτωση το τεχνητό νευρωνικό δίκτυο για το έτος 2012 κατάφερε να προσομοιώσει με καλή ακρίβεια την υπόγεια στάθμη των πηγαδιών και αυτό φαίνεται και από τους δύο Χάρτες. Οι οποίοι δεν έχουν κάποια σημαντική διαφορά μεταξύ τους, είναι σχεδόν όμοιοι με πολύ μικρή διαφορά στις κλάσεις των ισοϋψών.

Κεφάλαιο 6

Συζήτηση

Αρχικά τα νευρωνικά δίκτυα στα **Διαγράμματα 1-7** εκπαιδεύτηκαν με τον αλγόριθμο Levenberg-Marquardt, για ποσοστό Training 70%, Validation 15% και Testing 15%, ο αριθμός των εποχών παραμένει σταθερός με μέγιστο τις 1000 εποχές και για διαφορετικό αριθμό κρυφών κόμβων. Το σφάλμα για το Training θα είναι κοντά επίσης στα άλλα σφάλματα όπως φαίνεται και στον **Πίνακα 5** αφού στο Training γίνεται η αναπροσαρμογή των βαρών του νευρωνικού δικτύου ώστε να προσομοιώσει όσο καλύτερα γίνεται τα ζητούμενα αποτελέσματα.

Στο **Διάγραμμα 1** το νευρωνικό δίκτυο εκπαιδεύεται με 20 κρυφούς κόμβους, όπου παρατηρείται ότι το νευρωνικό δίκτυο έχει συγκλίνει αρκετά για μερικά πηγάδια ενώ για κάποια άλλα όχι, όμως το σφάλμα (RMSE) είναι μεγάλο για Training=13.0195 m, Validation=12.6080 m και Testing=12.8296 m όπως παρατηρείται και στον **Πίνακα 5** το συγκεκριμένο νευρωνικό δίκτυο σταμάτησε να εκπαιδεύεται στις 170/1000 εποχές λόγω υπερεκπαίδευσης του δικτύου. Αυτό σημαίνει ότι σε κάποια σημεία δεδομένων οι προσομοιωμένες τιμές του νευρωνικού δικτύου συγκλίνουν στις παρατηρημένες τιμές ενώ σε κάποια άλλα άρχισαν να αποκλίνουν.

Στην συνέχεια αυξάνοντας τους κρυφούς κόμβους τα σφάλματα αυξανόντουσαν, συγκεκριμένα για 25 και 30 κόμβους τα σφάλματα αυξήθηκαν κατά πολύ. Αξίζει να σημειωθεί ότι στους 30 κρυφούς κόμβους το νευρωνικό δίκτυο σταμάτησε την εκπαίδευση στις 629/1000 εποχές και με πολύ μεγαλύτερο σφάλμα από τα αποτελέσματα του **Διαγράμματος 1**.

Μετά από δοκιμές για 35, 50, 80 και 100 κόμβους τα σφάλματα ξεκίνησαν να μειώνονται σημαντικά. Η βέλτιστη εκπαίδευση μέχρι στιγμής ήταν αυτή για 100 κρυφούς κόμβους με σφάλματα (RMSE) Training=1.5585 m, Validation=1.6289 m, και Testing=1.7130 m. Ναι μεν ο χρόνος αυξήθηκε συγκεκριμένα 01:35:40 αλλά τα αποτελέσματα μειώθηκαν επίσης σημαντικά.

Ο λόγος για τον οποίο δεν εκπαιδεύτηκαν νευρωνικά δίκτυα για περισσότερους από 100 κρυφούς κόμβους είναι γιατί ο χρόνος εκπαίδευσης είχε ήδη αρχίσει να αυξάνεται σημαντικά. Προτιμήθηκε η αύξηση του αριθμού των εποχών δεδομένου του ότι τα περισσότερα νευρωνικά δίκτυα σταμάτησαν την εκπαίδευση επειδή έφτασαν το μέγιστο αριθμό εποχών. Αυτό δηλώνει ότι υπήρχε ακόμα περιθώριο εκπαίδευσης.

Καταλήγουμε ότι για τον αλγόριθμο Levenberg-Marquardt, στους 100 κόμβους ήταν η βέλτιστη εκπαίδευση για σταθερά ποσοστά που προαναφέραμε και με σχετικά μικρό χρόνο εκπαίδευσης. Η διακοπή του νευρωνικού δικτύου οφείλεται στα Epochs διότι έφτασαν το μέγιστο αριθμό.

Στην συνέχεια για τα **Διαγράμματα 8-11** τα νευρωνικά δίκτυα εκπαιδεύτηκαν για σταθερό αριθμό κόμβων στους 100, μιας και προηγουμένως παρατηρήσαμε ότι ήταν η βέλτιστη εκπαίδευση. Για σταθερό αριθμό εποχών και διαφορετικά ποσοστά (Percentages) για Training, Validation και Testing ελέγχεται η συμπεριφορά του νευρωνικού δικτύου.

Στο **Διάγραμμα 8** για ποσοστά Training=55%, Validation=25% και Testing=20% το σφάλμα μειώθηκε σε σχέση με τα αποτελέσματα από το **Διάγραμμα 7**. Στον **Πίνακα 6, 7**

& 8 φαίνονται τα ποσοστά για τα οποία εκπαιδεύτηκαν τα νευρωνικά δίκτυα και παρατηρείται ότι για τα αποτελέσματα του **Διαγράμματος 8** με ποσοστά Training=55%, Validation=25% και Testing=20% είχαμε καλύτερα αποτελέσματα σε σχέση με αυτά του **Διαγράμματος 7**. Τα αποτελέσματα για τα σφάλματα (RMSE) έχουν ως εξής, για Training=1.1507 m, Validation=1.2371 m και Testing=1.2143 m.

Επιπλέον εξετάστηκε η συμπεριφορά του νευρωνικού δικτύου για το μικρότερο ποσοστό εκπαίδευσης (Training=30%) που μπορεί να λάβει μέσα στο nftool και παρατηρήθηκε αύξηση του σφάλματος όπως φαίνεται και στο **Διάγραμμα 10** καθώς και στον **Πίνακα 8**. Αυτό οφείλεται στο γεγονός ότι το ποσοστό εκπαίδευσης ήταν αρκετά μικρό με αποτέλεσμα το νευρωνικό δίκτυο να μην μπορεί να εκπαιδευτεί σωστά και να δυσκολεύεται να εκτιμήσει τις ζητούμενες μεταβλητές.

Συνεπώς, η βέλτιστη λύση μέχρι στιγμής είναι αυτή από το **Διάγραμμα 8** όπως φαίνεται αναλυτικά και από τα αποτελέσματα στους **Πίνακες 6, 7 & 8**.

Τα νευρωνικά δίκτυα στα **Διαγράμματα 12-15** εκπαιδεύτηκαν με τον αλγόριθμο Bayesian Regularization. Αξίζει να σημειωθεί με τον συγκεκριμένο αλγόριθμο το νευρωνικό δίκτυο δεν πρόκειται να υπερεκπαιδευτεί αφού δεν κρατάει σφάλμα επικύρωσης (Validation Error). Επίσης για τον αλγόριθμο BR, το ποσοστό τιμών για Validation συμπεριλαμβάνεται στα ποσοστά του Training για την εκπαίδευση του τεχνητού νευρωνικού δικτύου.

Για το **Διάγραμμα 12** κάνοντας χρήση 10 κρυφών κόμβων και με ποσοστά για Training=70%, Validation=15%, Testing=15% τα σφάλματα ήταν αρκετά μεγάλα.

Όπως φαίνεται στους **Πίνακες 9, 10 & 11** για διαφορετικά ποσοστά (Percentages) στα Training, Validation και Testing και με την αύξηση των κρυφών κόμβων από 30 στους 80 και στους 100 τα σφάλματα (RMSE) μειώνονταν και ο χρόνος εκπαίδευσης αυξανόταν.

Στα αποτελέσματα του **Διαγράμματος 15** για ποσοστά Training=70%, Validation=5% και Testing=25%, με 100 κρυφούς κόμβους αξίζει να σημειωθεί ότι επιτεύχθηκε η καλύτερη εκπαίδευση του νευρωνικού δικτύου με σφάλμα (RMSE) Training=0.701 m το οποίο είναι αρκετά πιο μικρό από αυτό του **Διαγράμματος 8**.

Επειδή με τον αλγόριθμο BR η εκπαίδευση σταματούσε πάντα μετά την παρέλευση όλων των διαθέσιμων εποχών (1000/1000 epochs), αποφασίστηκε να αυξηθούν οι διαθέσιμες εποχές για να ελεγχθεί αν με αυτόν τον τρόπο θα μπορούσε να βελτιωθούν περαιτέρω τα αποτελέσματα του μοντέλου.

Συνεπώς, η βέλτιστη λύση πλέον είναι αυτή από το **Διάγραμμα 15** με σφάλμα της τάξεως 10^{-1} m.

Στην συνέχεια μέσω της εντολής nftool στο Command Window του επιστημονικού προγράμματος Matlab, εκπαιδεύτηκε νευρωνικό δίκτυο με τον αλγόριθμο Levenberg-Marquardt, για διαφορετικό αριθμό εποχών (2000) και 100 κρυφούς κόμβους καθώς και διαφορετικό μέγιστο αριθμό Validation Checks τα 60.

Πιο συγκεκριμένα στο **Διάγραμμα 16** με χρήση 100 κρυφών κόμβων, Epochs=2000 και μέγιστο αριθμό Validation Checks ίσο με 60, παρατηρούμε ότι μετά την παρέλευση

03:07:39 και με τα Validation Checks να μην φτάνουν στα μέγιστα το νευρωνικό δίκτυο εκπαιδεύτηκε με μεγάλη επιτυχία μιας και το σφάλμα (RMSE) μειώθηκε με τιμή Training=0.551 m έναντι 0.701 m στο **Διάγραμμα 15**.

Στην προκειμένη περίπτωση η μέχρι τώρα βέλτιστη λύση είναι αυτή του **Διαγράμματος 16** η οποία έχει μια πάρα πολύ καλή εμφάνιση με αποτελέσματα τα οποία είναι πολύ κοντά στις παρατηρημένες τιμές. Δύσκολα μπορεί να διακριθεί με το μάτι η απόκλιση των σημείων παρατηρημένη-προσομοιωμένη τιμή από την ευθεία $y=x$.

Έπειτα έγινε δοκιμή για τον αλγόριθμο Bayesian Regularization, με μέγιστο αριθμό εποχών στις 2000 και χρήση 100 κρυφών κόμβων. Η γραφική απεικόνιση είναι αυτή του **Διαγράμματος 17**, με εξίσου καλά αποτελέσματα. Επίσης το σφάλμα (RMSE) σε αυτήν την περίπτωση μειώθηκε περισσότερο με τιμή Training=0.456 m. Μπορεί ο χρόνος εκπαίδευσης να αυξήθηκε σημαντικά φτάνοντας τις 06:20:44 αλλά είχαμε τα βέλτιστα αποτελέσματα από όλες τις δοκιμές.

Δεν έγιναν εκπαιδεύσεις νευρωνικών δικτύων με περισσότερους από 100 κρυφούς κόμβους και 2000 εποχές, διότι ο χρόνος εκπαίδευσης θα αυξανόταν ακόμα περισσότερο με την αύξηση των 2 παραμέτρων. Με δεδομένο ότι για μια τόσο μεγάλη περιοχή μελέτης με εύρος τιμών υδραυλικού ύψους από 40-1440 μέτρα, ένα σφάλμα που κυμαίνεται μεταξύ των τάξεων $10^{-1} - 10^{-3}$ είναι μια αρκετά καλή εκτίμηση, η εκπαίδευση των νευρωνικών δικτύων σταμάτησε στο συγκεκριμένο σημείο. Μείωση του σφάλματος σε ακόμα μικρότερες τάξεις μεγέθους, κάτω από το επίπεδο των χιλιοστών του μέτρου θα σταματούσε πλέον να έχει φυσικό νόημα.

Στην **Ενότητα 4.5** εκπαιδεύτηκε νευρωνικό δίκτυο για ένα πηγάδι της Αυστρίας, με κωδικό πηγαδιού 'at_300533'. Το νευρωνικό δίκτυο εκπαιδεύτηκε με τον αλγόριθμο BR, για ποσοστά Training=70%, Validation=5% & Testing=25%, για 1000 εποχές και με 10 κρυφούς κόμβους. Η σύγκριση στο **Γράφημα 9** γίνεται για ίδιους αλγόριθμους (BR) αλλά για διαφορετικές παραμέτρους. Ο λόγος για τον οποίο χρησιμοποιήθηκαν μόνο 10 κρυφοί κόμβοι είναι γιατί τα δεδομένα μας για την εκπαίδευση τεχνητού νευρωνικού δικτύου για ένα μόνο πηγάδι ήταν πολύ λιγότερα σε σχέση με τα συνολικά δεδομένα εισόδου για όλα τα πηγάδια (128 πηγάδια συνολικά). Όπως παρατηρείται στο **Γράφημα 9** οι προσομοιωμένες τιμές του μοντέλου για την εκπαίδευση του νευρωνικού με δεδομένα για ένα μόνο πηγάδι (πράσινη γραμμή) είναι αρκετά καλύτερες από αυτές του νευρωνικού που εκπαιδεύτηκε με όλα τα δεδομένα μαζί (κόκκινη γραμμή). Αυτό το καταλαβαίνουμε από το γράφημα καθώς από ένα σημείο και μετά οι προσομοιωμένες τιμές της κόκκινης γραμμής αυξάνονται αντί να μειώνονται. Ενώ οι τιμές της πράσινης γραμμής για τα δεδομένα ενός πηγαδιού τείνουν πολύ κοντά στις αρχικές μας τιμές της υπόγειας στάθμης (μπλε γραμμή). Το μόνο αρνητικό είναι ότι οι τιμές της πράσινης γραμμής δεν είναι σε θέση να εκτιμήσουν τις ακραίες τιμές και περιορίζονται στην εκτίμηση της τάσης της στάθμης.

Από την άλλη πλευρά στον **Πίνακα 19**, το κύριο συμπέρασμα είναι ότι τα αποτελέσματα από το Τεχνητό Νευρωνικό Δίκτυο που εκπαιδεύτηκε με όλα τα δεδομένα για όλα τα πηγάδια είναι αρκετά χειρότερα από το Τεχνητό Νευρωνικό Δίκτυο που εκπαιδεύτηκε με ένα σύνολο δεδομένων για ένα πηγάδι παρατήρησης. Όπως φαίνεται στον **Πίνακα 19**, το σφάλμα RMSE είναι λίγο υψηλότερο (0.264 m) και ο συντελεστής συσχέτισης απέχει περισσότερο από την τέλεια τιμή. Σκοπός της μελέτης μας είναι να δημιουργήσουμε ένα

Τεχνητό Νευρωνικό Δίκτυο που θα μπορούσε να προσομοιώσει την στάθμη των υπογείων υδάτων με όλα τα δεδομένα και όχι για ένα μόνο πηγάδι.

Τέλος υπολογίστηκαν κάποιοι επιπλέον δείκτες για να δούμε αν το μοντέλο προσομοιώνει με καλή ακρίβεια την υπόγεια στάθμη των πηγαδιών. Αυτοί οι δείκτες είναι Μέσο Απόλυτο Σφάλμα (MAE), το Nash-Sutcliffe σφάλμα (NSE) και το Bias.

Κεφάλαιο 7

Συμπεράσματα

Συνοψίζοντας, έπειτα από αρκετές εκπαιδεύσεις νευρωνικών δικτύων για διαφορετικές παραμέτρους και για διαφορετικούς αλγόριθμους εκπαίδευσης, καταλήγουμε στο συμπέρασμα ότι και οι δύο αλγόριθμοι εκπαιδεύουν νευρωνικά δίκτυα με συγκρίσιμα καλά αποτελέσματα. Στις **Ενότητες 4.3.1** και **4.3.2** επιτεύχθηκαν τα μικρότερα σφάλματα τάξεως 10^{-1} . Πιο συγκεκριμένα είχαμε σφάλματα με τιμές 0,551 m και 0,456 m αντίστοιχα και με τους συντελεστές συσχέτισης (R) να είναι μονάδα. Η βέλτιστη εκπαίδευση νευρωνικού θεωρείται αυτή της **Ενότητας 4.3.2** με χρήση του αλγορίθμου BR.

Έπειτα εξετάσαμε την συμπεριφορά του νευρωνικού δικτύου χρησιμοποιώντας δεδομένα για ένα μόνο πηγάδι όπως είδαμε στην **Ενότητα 4.5**. Τα αποτελέσματα ήταν αρκετά καλά μιας και το σφάλμα μειώθηκε ακόμα περισσότερο φτάνοντας 0,153 m χρησιμοποιώντας τον αλγόριθμο BR. Στο **Διάγραμμα 18** έχουμε την παλινδρόμηση εκπαίδευσης για το συγκεκριμένο νευρωνικό δίκτυο και παρατηρούμε ότι οι συντελεστές συσχέτισης είναι πολύ μικρότεροι σε σχέση με τις δύο προηγούμενες ενότητες που αναφέραμε ακριβώς από πάνω. Όπως είδαμε και στο **Γράφημα 9** η εκπαίδευση νευρωνικού δικτύου για ένα πηγάδι μόνο και όχι για όλα τα πηγάδια μαζί αποσκοπεί σε καλύτερα αποτελέσματα ως προς την προσομοίωση της υπόγειας στάθμης καθώς οι τιμές (πράσινη γραμμή) βρίσκονται πιο κοντά στις παρατηρημένες τιμές (μπλε γραμμή). Εν αντιθέσει με την κόκκινη γραμμή από κάποιο σημείο και μετά αρχίζει να υπερεκτιμάει την υπόγεια στάθμη του πηγαδιού.

Σχετικά με την εκπαίδευση νευρωνικού δικτύου για ένα πηγάδι μόνο και λαμβάνοντας υπόψη τα αποτελέσματα που προέκυψαν καταλήγουμε στο συμπέρασμα ότι η εκπαίδευση νευρωνικού δικτύου για ένα πηγάδι θα ήταν προτιμητέα σε περίπτωση που μας ενδιέφερε να δούμε τη απόκριση της υπόγειας στάθμης του συγκεκριμένου πηγαδιού σε καθημερινή βάση. Η καλή απόκριση του μοντέλου για ένα πηγάδι, απαιτεί όμως την ύπαρξη ενός αρκετά μεγάλου πλήθους δεδομένων, κάτι που δεν είναι πάντα διαθέσιμο για όλα τα πηγάδια παρατήρησης. Όταν το πρόβλημα είναι η προσομοίωση ενός πολύ μεγάλου πλήθους πηγαδιών σε μεγάλη κλίμακα, π.χ. σε Πανευρωπαϊκή, τότε προτείνεται η χρήση ενός τεχνητού νευρωνικού δικτύου, εκπαιδευμένου με δεδομένα από όλα τα πηγάδια παρατήρησης. Όταν το πρόβλημα είναι για λίγα πηγάδια παρατήρησης σε μικρή κλίμακα, αν υπάρχει ικανοποιητικός αριθμός δεδομένων για κάθε πηγάδι, ενδεχομένως η εκπαίδευση χωριστών νευρωνικών δικτύων να καταλήξει σε καλύτερα αποτελέσματα. Από τα αποτελέσματα πάντως της παρούσας εργασίας, φαίνεται ότι ενώ η προσομοίωση της στάθμης βελτιώνεται, στην περίπτωση της εκπαίδευσης με τα δεδομένα ενός μόνο πηγαδιού, το νευρωνικό δίκτυο δυσκολεύεται να προσομοιώσει τις ακραίες τιμές και δίνει τιμές πιο κοντά στην τάση των παρατηρημένων δεδομένων.

Τέλος υπολογίστηκαν 3 επιπλέον δείκτες που χρησιμοποιούνται συχνά στην υδρολογία για να δούμε την απόδοση του μοντέλου και αν υπερεκτιμά η υποεκτιμά τα ζητούμενα δεδομένα, δηλαδή τις στάθμες των υπόγειων υδάτων σε διάφορα πηγάδια. Το αποτέλεσμα του δείκτη NSE με τιμή πολύ κοντά στην μονάδα ($NSE=0,9999$) αντιστοιχεί στην σχεδόν τέλεια προσομοίωση του μοντέλου στα παρατηρούμενα δεδομένα. Η τιμή του δείκτη NSE είναι μεγαλύτερη και εκτός των ορίων $0,5 < NSE < 0,65$ όπου για αυτά τα

όρια ένα μοντέλο θεωρείται ικανοποιητικό [Moriasi et al.,2007] & [Ritter & Munoz-Carpena, 2013]. Στην δική μας περίπτωση θα λέγαμε ότι το μοντέλο είναι κάτι αρκετά παραπάνω από ικανοποιητικό καθώς ο δείκτης φτάνει πολύ κοντά στην μονάδα. Η τιμή του δείκτη Bias όπως προαναφέραμε μπορεί να πάρει είτε θετικές είτε αρνητικές τιμές. Στην περίπτωση μας η τιμή του Bias είναι αρνητική με τιμή ($Bias = -3,7648 * 10^{-6}$). Εφόσον η τιμή του δείκτη Bias είναι αρνητική αυτό σημαίνει ότι το μοντέλο υποεκτιμάει τις ζητούμενες τιμές. Η τιμή του είναι ιδιαίτερα μικρή και σε καμία περίπτωση δεν προδιαθέτει για την ύπαρξη ενός σημαντικού συστηματικού σφάλματος του μοντέλου. Ο τελευταίος δείκτης αφορά το μέσο απόλυτο σφάλμα (MAE) του οποίου η τιμή είναι πάντα θετική. Ο δείκτης MAE με τιμή (MAE=0,3107) μας περιγράφει την μέση απόλυτη διαφορά μεταξύ παρατηρούμενης και προσομοιωμένης στάθμης.

ΒΙΒΛΙΟΓΡΑΦΙΑ

Διεθνείς

- Anastasio, T.J., 1993, Modeling vestibulo-ocular reflex dynamics: From classical analysis to neural networks in F. Eeckman, ed., *Neural Systems: Analysis and Modeling*, pages 407–430, Norwell, MA: Kluwer
- Arbib, M.A., 1987, *Brains, Machines and Mathematics*, Δεύτερη έκδοση, New York: Springer-Verlag
- Daliakopoulos, Ioannis N., Paulin Coulibaly, Ioannis K. Tsanis, 2005, Groundwater level forecasting using artificial neural networks, *Journal of Hydrology*, Volume 309 Issues 1–4, pages 229–240
- Fahlman, S. E. and Lebiere C., 1990, The Cascade-Correlation learning architecture, In [22f], pages 524–532
- Foresee, F. Dan, and Martin T. Hagan., 1997 June, Gauss-Newton approximation to Bayesian learning, *Proceedings of the International Joint Conference on Neural Networks*
- Geman, S., E. Bienenstock, and R. Doursat, 1992, Neural networks and the bias/variance dilemma, *Neural Computation*, volume 4, pages 1–58
- Grossberg, S., 1988, *Neural Networks and Natural Intelligence*, Cambridge, MA: MIT Press
- Hagan, M.T., and M. Menhaj, 1994, Training feed-forward networks with the Marquardt algorithm, *IEEE Transactions on Neural Networks*, Volume 5, No. 6, pages 989–993
- Hassibi, B., and Stork D. G., 1993, Second order derivatives for network pruning: Optimal brain surgeon, pages 164–171
- Kerlirzin, P., Vallet F., 1993, Robustness in multilayer perceptrons, *Neural Computation*, volume 5, pages 473–482
- Khaled, Balkhair S, 2002, Aquifer parameters determination for large diameter wells using neural network approach, *Journal of Hydrology*, Volume 265 Issues 1–4, pages 118–128
- Krogh, A. and Hertz J. A., 1992, A simple weight decay can improve generalization, pages 950–957
- Kröse, Ben, Patrick van der Smagt, November 1996, *An introduction to Neural-Networks*, Eighth edition
- Levin, A. U., Leen T. K., and Moody J. E., 1994, Fast pruning using principal components
- Le Cun, Y., Denker, J. S. and Solla S. A., 1990, Optimal brain damage. In [22], pages 598–605
- MacKay, David J. C., 1992, Bayesian interpolation, *Neural computation*, Volume 4 No. 3, pages 415–447

- Marquardt, Donald W., 1963, An Algorithm for Least-Squares Estimation of Non Linear Parameters, *Journal of the Society for Industrial and Applied Mathematics*, Volume 11 Issue 2, pages 431-441
- Mead, C.A., 1989, *Analog VLSI and Neural Systems*, Reading, MA: Addison-Wesley
- McCulloch, W.S., Pitts, W., 1943. A logical calculus of the ideas immanent in nervous activity, *Bull. Math. Biophysics*, pages 115-133
- Morgan N. and Bourlard H., 1990, Generalization and parameter estimation in feedforward nets: Some experiments, pages 630–637
- Nash J.E., J.V. Sutcliffe, 1970, River flow forecasting through conceptual models part I — A discussion of principles, *Journal of Hydrology*, Volume 10 Issue 3, pages 282-290
- Nowlan, S. J. and Hinton G. E., 1992, Simplifying neural networks by soft weight-sharing, *Neural Computation*, Volume 4 Issue 4, 473–493
- Moriasi, D. N.; Arnold, J. G.; Van Liew, M. W.; Bingner, R. L.; Harmel, R. D.; Veith, T. L., 2007, Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations, *Transactions of the ASABE*, pages 885–900
- Okut, Hayrettin, 2016, Bayesian Regularized Neural Networks for Small n Big p Data, Chapter 2, IntechOpen, pages 27-48
- Palani, Sundarambal, Shie-Yui Liong, Pavel Tkachik, 2008, An ANN application for water quality forecasting, *Marine Pollution Bulletin*, Volume 56 Issue 9, pages 1586-1597
- Prechelt, Lutz, 1998, *Neural Networks: Tricks of the Trade*, Springer Berlin Heidelberg, pages 55-69
- Ritter Axel, Rafael Muñoz-Carpena, 2013, Performance evaluation of hydrological models: Statistical significance for reducing subjectivity in goodness-of-fit assessments, *Journal of Hydrology*, Volume 480, pages 33-45
- Rumelhart, D. E., McClelland, J. L., Hinton, G. E., 1986, A general framework for parallel distributed processing, pages 46
- Sterling, P., 1990, “Retina,” in *The Synoptic Organization of the Brain*, G.M. Shepherd, ed., 3d ed., pages 170–213, New York: Oxford University Press
- Tapoglou Evdokia, George P. Karatzas, Ioannis C. Trichakis, Emmanouil A. Varouchakis, 2014, A spatio-temporal hybrid neural network-Kriging model for groundwater level simulation, *Journal of Hydrology*, Volume 519 Part D, pages 3193-3203.
- Tapoglou Evdokia, Ioannis C. Trichakis, Zoi Dokou, Ioannis K. Nikolos & George P. Karatzas, 2014, Groundwater-level forecasting under climate change scenarios using an artificial neural network trained with particle swarm optimization, *Hydrological Sciences Journal*.

Vörösmarty Charles, J., Pamela Green, Joseph Salisbury, Richard B. Lammers, Vulnerability from Climate Change and Population Growth, Global Water Resources, Volume 289 Issue 5477, pages 284-288

Weigend, A. S., Rumelhart D. E., and Huberman B. A., 1991, Generalization by weight elimination with application to forecasting, pages 875–882

Ελληνική

Γεώργιο Π. Καρατζάς, 2004, Το φαινομενολογικό υδρολογικό μοντέλο, Πανεπιστημιακές σημειώσεις: Ροή υπόγειων υδάτων και μεταφορά ρύπων, σελίδες 5-9

Βλαχάβας, Ι., Κεφαλάς, Π., Βασιλειάδης, Ν., Ρεφανίδης, Ι., Κόκκορας, Φ., Σακελλαρίου, Η., 2002. Τεχνητή Νοημοσύνη. Εκδόσεις Γαρταγάνη, Θεσσαλονίκη

Ιστοσελίδες

- United State Geological Survey, <https://water.usgs.gov> Accessed [1/03/2019]
- Wikipedia, https://el.wikipedia.org/wiki/Νευρωνικό_δίκτυο Accessed [8/03/2019]
- Wikipedia, <https://el.wikipedia.org/wiki/Δούναβης> Accessed [8/03/2019]
- Wikipedia, https://el.wikipedia.org/wiki/Δέλτα_Δούναβη Accessed [8/03/2019]
- Matlab Documentations, Search for Bayesian Regularization→trainbr https://www.mathworks.com/help/deeplearning/ref/trainbr.html?searchHighlight=bayesian%20regularization&s_tid=doc_srchtile Accessed [12/03/2019]