



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΗΛΕΚΤΡΙΚΗΣ ΙΣΧΥΟΣ
ΕΡΓΑΣΤΗΡΙΟ ΥΨΗΛΩΝ ΤΑΣΕΩΝ

Εκτίμηση της κρίσιμης τάσης υπερπήδησης σε μονωτήρες με τη χρήση Τεχνητών Νευρωνικών Δικτύων

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Αγγελική Αθ. Γιαλκέτση

Καθηγητής : Ιωάννης Αθ. Σταθόπουλος

Επιβλέπουσα: Βασιλική Θ. Κονταργύρη

Αθήνα, Μάρτιος 2005

*“Where is the wisdom we lost in knowledge,
where is the knowledge we lost in information?”*

T. S. Eliot

*“Επιστήμη χωρίς συνείδηση δεν είναι
παρά καταστροφή της ψυχής”*

Ραμπελέ



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΗΛΕΚΤΡΙΚΗΣ ΙΣΧΥΟΣ
ΕΡΓΑΣΤΗΡΙΟ ΥΨΗΛΩΝ ΤΑΣΕΩΝ

122

Εκτίμηση της κρίσιμης τάσης υπερπήδησης σε μονωτήρες με τη χρήση Τεχνητών Νευρωνικών Δικτύων

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Αγγελική Αθ. Γιαλκέτση

Καθηγητής : Ιωάννης Αθ. Σταθόπουλος

Επιβλέπουσα : Βασιλική Θ. Κονταργύρη

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 9^η Μαρτίου 2005.

Ιωάννης Αθ. Σταθόπουλος
Καθηγητής

Περικλής Δ. Μπούρκας
Καθηγητής

Φραγκίσκος Β. Τοπαλής
Αναπληρωτής Καθηγητής

Αθήνα, Μάρτιος 2005

Αγγελική Αθ. Γιαλκέτση

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Copyright © Εθνικό Μετσόβιο Πολυτεχνείο

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

Περίληψη

Σκοπός αυτής της διπλωματικής εργασίας είναι να αναδειχθούν οι δυνατότητες των Τεχνητών Νευρωνικών Δικτύων για εφαρμογές στις υψηλές τάσεις και ειδικότερα να εκτιμηθεί η κρίσιμη τάση υπερπήδησης σε ρυπασμένους μονωτήρες με τη χρήση ενός ΤΝΔ, εκπαιδευόμενου με τον αλγόριθμο ανάστροφης διάδοσης σφάλματος.

Για το λόγο αυτό κατασκευάστηκε ένα ΤΝΔ με έτοιμα εργαλεία του MATLAB, το οποίο εκπαιδεύτηκε με διάφορες μεθόδους εκπαίδευσης του MATLAB, ενώ δοκιμές έγιναν και ως προς το πλήθος νευρώνων, τον αριθμό εποχών, την τιμή του ρυθμού μάθησης και την προσθήκη και δεύτερου κρυφού επιπέδου, προκειμένου να βρεθεί για ποια αρχιτεκτονική δικτύου και ποια τιμή των υπόλοιπων παραμέτρων προκύπτει βέλτιστο αποτέλεσμα. Τα πρότυπα που χρησιμοποιήθηκαν ήταν συνολικά 118 και κάποια είχαν προκύψει από παλαιότερα πειράματα, ενώ κάποια άλλα προέκυψαν από εφαρμογή μαθηματικού τύπου βασισμένου σε απλοποιημένο μοντέλο για την εύρεση της κρίσιμης τάσης.

Τα αποτελέσματα του δικτύου αυτού συγκρίθηκαν με αυτά που έδωσε άλλο προσαρμοστικό ΤΝΔ κατασκευασμένο σε digital Fortran, στο οποίο προηγουμένως έγιναν κατάλληλες τροποποιήσεις. Τα αποτελέσματα κρίνονται ικανοποιητικά, αν και υπάρχει περιθώριο βελτίωσης.

Λέξεις κλειδιά

Μονωτήρες, Υπερπήδηση, Κρίσιμη τάση, Ρύπανση, Τεχνητά Νευρωνικά Δίκτυα, Επιβλεπόμενη μάθηση, Ανάστροφη διάδοση σφάλματος, Εκτίμηση, Ανταγωνιστική μάθηση, MATLAB.

Abstract

This diploma thesis attempts to elucidate the potentials of Artificial Neural Networks in high voltage applications and especially to estimate the flashover voltage on polluted insulators, using an ANN which is trained with the error backpropagation algorithm.

For this purpose, an ANN was constructed, using functions of MATLAB and has been trained with several MATLAB training functions, while tests regarding the number of neurons, the number of epochs, the value of learning rate and the addition of a second hidden layer have taken place, in order to find which net architecture and which value of the other parameters give the best result. The patterns used were 118 in total. Some of them had resulted from former experiments, while some other had resulted by applying a mathematical type based on a simplified model for the calculation of the flashover voltage.

The results of this first network were compared with those given by another adaptive ANN, implemented in digital Fortran, which had firstly been properly modified. The described method gave satisfactory results, although there is further work to be done in the future.

Key words

Insulators, Flashover, Critical Voltage, Pollution, Artificial Neural Networks, Supervised learning, Error backpropagation, Estimation, Competitive learning, MATLAB.

Περιεχόμενα

	Σελίδα
ΠΕΡΙΕΧΟΜΕΝΑ	i
ΠΡΟΛΟΓΟΣ	v
ΚΕΦΑΛΑΙΟ 1 Μονωτήρες	
1.1 Γενικά	1
1.1.1 Υλικά των μονωτήρων	2
1.1.2 Ιδιότητες των μονωτήρων	3
1.1.3 Τύποι μονωτήρων	5
1.2 Ρύπανση μονωτήρων	10
1.2.1 Το φαινόμενο της υπερπήδησης στους μονωτήρες	13
1.2.2 Στάδια της υπερπήδησης σε ένα μονωτήρα εξαιτίας της ρύπανσης	15
1.2.3 Τρόποι μετριασμού του φαινομένου της υπερπήδησης	19
1.2.4 Παρακολούθηση του ρεύματος διαρροής	21
1.3 Σύνοψη	22
ΚΕΦΑΛΑΙΟ 2 Νέες τεχνικές για ποιοτικό έλεγχο μονωτήρων	
2.1 Εισαγωγή	23
2.2 Συσκευές ελέγχου ρύπανσης (pollution monitors)	23
2.2.1 Συσκευές ελέγχου ρεύματος διαρροής	24
2.2.2 Λειτουργία των συσκευών ελέγχου ρεύματος διαρροής	25
2.2.3 Εφαρμογές των παραπάνω συστημάτων	25
2.3 UV κάμερα για την επιθεώρηση σύνθετων μονωτήρων	26
2.3.1 Αρχή λειτουργίας	28
2.3.2 Ανάλυση και επεξεργασία εικόνων	28
2.3.3 Πειραματική διάταξη	29
2.3.4 Αποτελέσματα	30
2.4 Τεχνική με λέιζερ για έλεγχο σύνθετων μονωτήρων	31
2.4.1 Γενικά	32
2.4.2 Μονωτήρες που εξετάστηκαν	32
2.4.3 Ηλεκτρικές δοκιμές	35
2.4.4 Φασματοσκοπία φθορισμού	38

2.5	Μετασχηματισμός Κύματος (Wavelet Transformation) για την πρόβλεψη αστοχιών στους σύνθετους μονωτήρες	41
2.6	Τεχνητά Νευρωνικά Δίκτυα για τον έλεγχο των μονωτήρων	44
2.6.1	TNΔ για την εκτίμηση του επιπέδου ρύπανσης σε μονωτήρες ΥΤ	44
2.6.2	Εκτίμηση του χρόνου έως την υπερπήδηση σε μονωτήρες που βρίσκονται σε συνθήκες ρύπανσης με χρήση TNΔ	46
2.6.3	Χρήση συσκευών ελέγχου ρύπανσης σε συνδυασμό με TNΔ για την πρόβλεψη υπερπήδησης σε μονωτήρες	48
2.6.4	TNΔ για την ανάλυση επιφανειακών ρωγμών σε συμπαγείς μονωτήρες	48
2.7	Σύνοψη	49

ΚΕΦΑΛΑΙΟ 3 Υπολογιστική Νοημοσύνη και Τεχνητά Νευρωνικά Δίκτυα

3.1	Υπολογιστική νοημοσύνη: ένας νέος κλάδος	50
3.1.1	Ασαφή σύνολα και ασαφής λογική	52
3.1.2	Γενετικοί Αλγόριθμοι	55
3.1.3	Τεχνητά Νευρωνικά Δίκτυα	62
3.2	Θεωρία Τεχνητών Νευρωνικών Δικτύων	63
3.2.1	Ένας πρώτος ορισμός του TNΔ	63
3.2.2	Ο βιολογικός νευρώνας	65
3.2.3	Τεχνητοί νευρώνες	67
3.2.4	Επικοινωνία σημάτων με μη δυαδικό τρόπο	68
3.2.5	Χαρακτηριστικά TNΔ – Σύγκριση με συμβατικούς υπολογιστές	69
3.2.6	Διαδικασία κατασκευής TNΔ	70
3.2.7	Κατηγοριοποίηση των TNΔ	71

ΚΕΦΑΛΑΙΟ 4 Δίκτυα με επίβλεψη

4.1	Γενικά	73
4.2	Το Perceptron	73
4.2.1	Ελαχιστοποιώντας το σφάλμα: ο κανόνας δέλτα	76
4.2.2	Πολυεπίπεδα δίκτυα	80
4.3	Ο αλγόριθμος Backpropagation	82
4.3.1	Μη γραμμικός διαχωρισμός προβλήματα	86
4.3.2	Τοπικά ακρότατα	92
4.3.3	Επιταχύνοντας τη μάθηση: ο ρόλος της ορμής (momentum)	93
4.3.4	Κριτήρια τερματισμού του αλγόριθμου Backpropagation	93
4.3.5	Τρόποι εκπαίδευσης	94
4.4	Σύνοψη δυνατοτήτων του Πολυστρωματικού Perceptron	95

4.4.1	Ταξινόμηση προτύπων	96
4.4.2	Προσέγγιση συναρτήσεων	96
4.4.3	Γενίκευση	97
4.4.4	Ανεκτικότητα σε βλάβες	98
ΚΕΦΑΛΑΙΟ 5 Βασικές Αρχές Ασαφούς Λογικής		
5.1	Γενικά	99
5.2	Ανταγωνιστική μάθηση	99
5.3	Μάθηση Hebb	106
5.4	Αυτό-οργανούμενοι χάρτες Kohonen	107
5.4.1	Ο αλγόριθμος	109
5.4.2	Ένα παράδειγμα	109
5.5	Αυτοσυσχετιστικές μνήμες – Δίκτυα Hopfield	111
5.6	Σύνοψη	115
ΚΕΦΑΛΑΙΟ 6 Ανάπτυξη αλγορίθμου για την πρόβλεψη της τάσης διάσπασης σε μονωτήρες ΥΤ		
6.1	Περιγραφή προβλήματος	116
6.2	Δεδομένα εκπαίδευσης	118
6.3	Πειραματική διάταξη	118
6.3.1	Θάλαμος ρύπανσης	119
6.3.2	Θάλαμος ομίχλης	120
6.3.3	Δοκιμές αντοχής	121
6.4	Μαθηματικό μοντέλο για την εύρεση της κρίσιμης τάσης	121
6.5	Επιλογή των συνόλων εκπαίδευσης και επιβεβαίωσης	123
6.6	Κατασκευή του νευρωνικού δικτύου	125
6.7	Σύνοψη	131
ΚΕΦΑΛΑΙΟ 7 Αποτελέσματα		
7.1	Εισαγωγή	133
7.1.1	Λειτουργία της <code>traingd</code>	133
7.1.2	Λειτουργία των <code>traingda</code> και <code>traingdx</code>	134
7.1.3	Λειτουργία της <code>trainlm</code>	135
7.2	Σενάρια 1-10 για μέθοδο εκπαίδευσης την <code>traingd</code>	136
7.3	Σενάρια 1-10 για μέθοδο εκπαίδευσης την <code>traingda</code>	144
7.4	Σενάρια 1-10 για μέθοδο εκπαίδευσης την <code>traingdx</code>	150
7.5	Σενάρια 1-10 για μέθοδο εκπαίδευσης την <code>trainlm</code>	157
7.6	Σχολιασμός των αποτελεσμάτων	163
7.7	Μεταβάλλοντας το ρυθμό μάθησης	164
7.8	Δοκιμή για δύο επίπεδα	176

7.9	Παρατηρήσεις	177
ΚΕΦΑΛΑΙΟ 8	Πρόγραμμα σε Fortran	
8.1	Εισαγωγή	178
8.2	Δοκιμές	183
ΚΕΦΑΛΑΙΟ 9	Συμπεράσματα – Η επόμενη μέρα	189
ΒΙΒΛΙΟΓΡΑΦΙΑ		196
ΠΑΡΑΡΤΗΜΑ		199

Πρόλογος

Η εργασία αυτή αποτελεί τη διπλωματική εργασία της φοιτήτριας Γιαλκέτση Αγγελικής για την απόκτηση του διπλώματος του Ηλεκτρολόγου Μηχανικού και Μηχανικού Υπολογιστών του Εθνικού Μετσόβιου Πολυτεχνείου. Αντικείμενο της εργασίας αυτής είναι η εφαρμογή των Τεχνητών Νευρωνικών Δικτύων (ΤΝΔ) για την εκτίμηση της κρίσιμης τάσης υπερπήδησης σε μονωτήρες Υψηλών Τάσεων. Για το σκοπό αυτό κατασκευάστηκε ΤΝΔ στο MATLAB, το οποίο εκπαιδεύτηκε με δεδομένα από πειραματικές μετρήσεις, αλλά και δεδομένα που προέκυψαν από προσεγγιστικό μαθηματικό τύπο. Επίσης τροποποιήθηκε υπάρχον ΤΝΔ, υλοποιημένο σε Fortran και εκπαιδεύτηκε με τα ίδια δεδομένα, για να γίνει σύγκριση των αποτελεσμάτων. Στα Κεφάλαια 1 έως 5 γίνεται αναφορά στη θεωρία τόσο των μονωτήρων, όσο και των νευρωνικών δικτύων, ενώ στα Κεφάλαια 6 έως 9 γίνεται επεξήγηση του αλγορίθμου που κατασκευάστηκε και της μεθοδολογίας που χρησιμοποιήθηκε και παράθεση των αποτελεσμάτων. Συγκεκριμένα:

Στο **Κεφάλαιο 1** γίνεται μια γενική αναφορά στους μονωτήρες. Αναλύονται τα υλικά απ' τα οποία κατασκευάζονται, οι ιδιότητες που προσδίδουν αυτά στους μονωτήρες, ενώ επιχειρείται και μια κατηγοριοποίηση των τύπων μονωτήρων. Επίσης περιγράφεται το πρόβλημα της ρύπανσης των μονωτήρων και η συνεισφορά του στο φαινόμενο της υπερπήδησης. Τέλος, αναλύονται οι μηχανισμοί, οι οποίοι λαμβάνουν χώρα ως τη δημιουργία του φαινομένου της υπερπήδησης, τόσο σε μονωτήρες πορσελάνης, όσο και σε πολυμερείς μονωτήρες.

Στο **Κεφάλαιο 2** παρουσιάζονται διάφορες τεχνικές με τις οποίες μπορεί να ελεγχθεί η κατάσταση των μονωτήρων, ώστε να μπορέσει να καθιερωθεί πρόγραμμα συντήρησής τους ή αντικατάστασής τους, για να αποφευχθούν περαιτέρω βλάβες στα συστήματα μεταφοράς ενέργειας.

Το **Κεφάλαιο 3** αναφέρεται στις βασικές αρχές της Υπολογιστικής Νοημοσύνης (ΥΝ) και των τριών κλάδων που τη συνθέτουν: την ασαφή λογική, τους γενετικούς αλγόριθμους και τα νευρωνικά δίκτυα. Μετά τη σύντομη περιγραφή των τριών αυτών μεθόδων της ΥΝ δίνεται μεγαλύτερη έμφαση στη θεωρία των Τεχνητών Νευρωνικών Δικτύων, γιατί αυτά θα απασχολήσουν στη συνέχεια.

Στο **Κεφάλαιο 4** αρχίζει η περιγραφή της μίας απ' τις δύο βασικές κατηγορίες Τεχνητών Νευρωνικών Δικτύων, των δικτύων με επίβλεψη. Παρουσιάζεται το perceptron, το πιο απλό απ' τα δίκτυα αυτά, η μάθηση που βασίζεται στον κανόνα ανάστροφης διάδοσης σφάλματος (error backpropagation) αλλά και τα προβλήματα που μπορούν να επιλυθούν με τα πολυεπίπεδα δίκτυα (multilayer perceptrons).

Το **Κεφάλαιο 5** ασχολείται με τη δεύτερη μεγάλη κατηγορία δικτύων, τα δίκτυα χωρίς επίβλεψη. Γίνεται μια γενική αναφορά στην ανταγωνιστική μάθηση και τη μάθηση Hebb, ενώ επίσης παρουσιάζονται, χωρίς να γίνει περιγραφή σε βάθος, οι χάρτες SOM και τα δίκτυα Hopfield.

Στο **Κεφάλαιο 6** περιγράφεται ο αλγόριθμος που αναπτύχθηκε με χρήση έτοιμων εργαλείων του MATLAB για την κατασκευή, εκπαίδευση και έλεγχο του ΤΝΔ. Εξηγείται ο τρόπος με τον οποίον αποκτήθηκαν τα δεδομένα εισόδου και ο τρόπος με τον οποίον αυτά χωρίστηκαν σε “πληθυσμούς” εκπαίδευσης και επιβεβαίωσης. Τέλος, παρατίθεται ο αλγόριθμος όπως γράφτηκε στο MATLAB.

Στο **Κεφάλαιο 7** παρουσιάζονται τα αποτελέσματα των διαφόρων δοκιμών που έγιναν με το πρόγραμμα σε MATLAB, προκειμένου να βρεθεί ο συνδυασμός που δίνει το μικρότερο σφάλμα μεταξύ πραγματικών και εκτιμώμενων τιμών.

Στο **Κεφάλαιο 8** παρουσιάζεται ένας εναλλακτικός τρόπος κατασκευής ΤΝΔ, σε digital Fortran, με το οποίο έγιναν επίσης δοκιμές για να βρεθεί το μικρότερο σφάλμα μεταξύ πραγματικών και εκτιμώμενων τιμών.

Τέλος, στο **Κεφάλαιο 9** γίνεται σύγκριση των παραπάνω δικτύων με βάση τα αποτελέσματά τους, αιτιολόγηση των διαφορών που παρουσιάζονται και εκτίμηση των προοπτικών που υπάρχουν για περαιτέρω μελέτη στο θέμα αυτό.

Στο σημείο αυτό θα ήθελα να εκφράσω τις ειλικρινείς ευχαριστίες μου στον **κ. Ιωάννη Αθ. Σταθόπουλο**, καθηγητή του τομέα Ηλεκτρικής Ισχύος του Εθνικού Μετσόβιου Πολυτεχνείου, για την ανάθεση της διπλωματικής εργασίας και για το άριστο κλίμα συνεργασίας καθ' όλη τη διάρκεια της συγγραφής της.

Επίσης θα ήθελα να ευχαριστήσω θερμά την **κα. Βασιλική Κονταργύρη**, υποψήφια διδάκτορα και μηχανικό του Εθνικού Μετσόβιου Πολυτεχνείου για το ότι ήταν δίπλα μου όλο αυτό το διάστημα, πρόθυμη να με βοηθήσει με κάθε δυνατό τρόπο. Την ευχαριστώ για το υλικό που μου παρείχε, στο οποίο συμπεριλαμβάνεται και τμήμα της ενδιάμεσης κρίσης της διδακτορικής διατριβής της, αλλά και για το χρόνο που μου αφιέρωσε.

Ευχαριστώ επίσης τον **κ. Γεώργιο Τσεκούρα**, υποψήφιο διδάκτορα και μηχανικό του Εθνικού Μετσόβιου Πολυτεχνείου, για τις πολύτιμες συμβουλές του πάνω στο θέμα των Τεχνητών Νευρωνικών Δικτύων και για την αμέριστη συμπαράστασή του, χωρίς την οποία η εργασία αυτή δεν θα ήταν εύκολο να ολοκληρωθεί. Τον ευχαριστώ ιδιαίτερα για το γεγονός ότι μου παραχώρησε το πρόγραμμα κατασκευής του ΤΝΔ που είχε υλοποιήσει σε digital Fortran, ώστε να το χρησιμοποιήσω ως μέτρο σύγκρισης για τα αποτελέσματα του δικού μου προγράμματος.

Θερμές ευχαριστίες εκφράζω επίσης προς τον **κ. Ιωάννη Φ. Γκόνο**, διδάκτορα μηχανικό του Εθνικού Μετσόβιου Πολυτεχνείου, που πρόθυμα προσέφερε τη βοήθειά του για την επίλυση των διαφόρων προβλημάτων που ανέκυψαν, αλλά και προς όλους τους ανθρώπους του εργαστηρίου Υψηλών Τάσεων για το ευχάριστο κλίμα και τη συμπαράστασή τους στις δύσκολες στιγμές αυτής της περιόδου.

Ιδιαίτερα ευχαριστώ και τους φίλους μου και συμφοιτητές μου για την κατανόηση που επιδεικνύουν όλα αυτά τα χρόνια.

Τέλος, ευχαριστώ τους γονείς μου, που σε όλα τα χρόνια της ζωής μου στηρίζουν ηθικά και υλικά τις επιλογές μου και πιστεύουν σε μένα και τις δυνατότητές μου.

Κεφάλαιο 1

Μονωτήρες

1.1 Γενικά

Οι μονωτήρες που θα μελετηθούν στην εργασία αυτή είναι οι λεγόμενοι μονωτήρες Υψηλής Τάσης (High Voltage Insulators) και χρησιμοποιούνται σε δίκτυα παροχής ηλεκτρισμού για να στηρίζουν, να διαχωρίζουν ή να εμπεριέχουν αγωγούς υψηλής τάσης [1].

Όμως οι πρώτες εφαρμογές των μονωτήρων δεν είχαν σχέση με τη μεταβίβαση ηλεκτρικής ισχύος. Αρχικά χρησιμοποιούνταν στα δίκτυα τηλεγράφων (1835), ενώ σε δίκτυο μεταφοράς ηλεκτρικής ισχύος χρησιμοποιήθηκαν για πρώτη φορά το 1882, όταν τέθηκε σε λειτουργία η πρώτη γραμμή μεταφοράς ηλεκτρικής ενέργειας. Αυτή η σύντομη γραμμή μεταφοράς ήταν συνεχούς τάσης (1343V) και συνέδεε το Miesbach με το Μόναχο. Στη συνέχεια η ανάπτυξη ήταν ταχύτατη, ενώ τα τελευταία 50 χρόνια η μεταφορά γίνεται με γραμμές υψηλών τάσεων μήκους πολλών χιλιομέτρων, όπως συμβαίνει στις Η.Π.Α και στην πρώην Σοβιετική Ένωση, και με πυκνά διασυνδεδεμένα δίκτυα στην Ευρώπη. Οι ηλεκτρικές και μηχανικές τάσεις στις οποίες έπρεπε να ανθίστανται οι τηλεγραφικοί μονωτήρες ήταν προφανώς αμελητέες σε σύγκριση με εκείνες που προκύπτουν από το φορτίο μιας εναέριας γραμμής μεταφοράς ηλεκτρικής ενέργειας. Οι νέες αυτές απαιτήσεις αποκάλυψαν σοβαρές ελλείψεις όχι μόνο στα υλικά, αλλά και στα σχέδια των μονωτήρων τα οποία αρχικά ήταν απλά αντίγραφα των τύπων τηλεγράφων [1].

Η εμπειρία της χρήσης και συνεπώς η ανάπτυξη των προϊόντων από γυαλί ή πορσελάνη για τους μονωτήρες γραμμών υψηλής τάσης συγκεντρώνεται εδώ και πάνω από 100 χρόνια. Οι μονωτήρες τύπου cap and pin από γυαλί ή πορσελάνη, μαζί με τους μονωτήρες πορσελάνης μεγάλου μήκους κυριαρχούσαν στην αγορά μέχρι την εμφάνιση των εναλλακτικών μονωτήρων από πολυμερή τη δεκαετία 1960.

Ένα από τα κρίσιμα σημεία για τους υπαίθριους μονωτήρες είναι η διεπιφάνεια μεταξύ του στερεού μονωτικού σώματος και του περιβάλλοντος αέρα. Τόσο το σώμα του στερεού, όσο και ο αέρας έχουν ορισμένες διηλεκτρικές ιδιότητες, που δεν επηρεάζονται από τη διεπιφάνεια. Οι λόγοι για τους οποίους η διεπιφάνεια προκαλεί προβλήματα προέρχονται από το γεγονός ότι η επιφάνεια του μονωτήρα μπορεί να ρυπανθεί. Σήμερα οι επιφάνειες των μονωτήρων αποτελούνται από στιλβωμένη πορσελάνη, γυαλί, οργανικά ή ημιοργανικά πολυμερή και συμπαγή πολυμερή (πολυμερή ‘τσιμέντα’) [2].

Με τα νέα υλικά, καθώς και με την εισαγωγή των πολυμερών μονωτήρων έρχονται νέες τεχνικές παραγωγής, αλλά και νέα σχέδια.

1.1.1 Υλικά των μονωτήρων

Τα τρία βασικά μέρη των μονωτήρων είναι το διηλεκτρικό, οι ακροδέκτες, που συνδέουν το διηλεκτρικό με την υπόλοιπη μηχανολογική δομή και τα ενδιάμεσα υλικά στο εσωτερικό του μονωτήρα, όπως τσιμέντο ή λιπαντικά.

Τα διηλεκτρικά είναι κυρίως κεραμικά (πορσελάνη, γυαλί), ή πολυμερή. Στην περίπτωση των κεραμικών διηλεκτρικών, το κυρίως επαναλαμβανόμενο άτομο είναι το πυρίτιο, ενώ στην περίπτωση των πολυμερών ο άνθρακας. Τα υλικά αυτά απαιτείται να διατηρούν την εφαρμοζόμενη διαφορά δυναμικού για αρκετές δεκαετίες χωρίς να αστοχήσουν. Πρέπει επίσης να ανθίστανται σε καταπονήσεις κρουστικών τάσεων που προκαλούνται από κεραυνούς ή χειρισμούς διακοπών του δικτύου χωρίς να διασπώνται. Επειδή οι επιφάνειες των διηλεκτρικών είναι πάντα, στην πράξη, ηλεκτρικά αγώγιμες εξαιτίας της υγρασίας και των επικαθίσεων, είναι απαραίτητη μία διαχρονική ανθεκτικότητα στις επιφανειακές ηλεκτρικές εκκενώσεις, τα ηλεκτροχημικά προϊόντα και τη διάβρωση.

Η βασική χημική δομή των διηλεκτρικών διαδραματίζει σπουδαίο ρόλο στον προσδιορισμό τόσο των ηλεκτρικών, όσο και των μηχανικών ιδιοτήτων των μονωτήρων. Τόσο τα κεραμικά, όσο και τα πολυμερή υλικά αποτελούνται από τετρασθενή άτομα, τα οποία έχουν την ικανότητα σχηματισμού εκτεταμένων δομών.

Η ηλεκτρική πορσελάνη αποτελείται από ένα υαλώδες πλέγμα (άμορφη δομή) που περιλαμβάνει κρυστάλλους διαφορετικών ειδών και κόκκους από τα αρχικά συστατικά των ορυκτών, ορισμένα από τα οποία έχουν διαλυθεί μερικώς. Το ηλεκτρικό γυαλί, στην ιδανική περίπτωση, είναι ένα μίγμα πυριτικών αλάτων υαλώδους μορφής χωρίς

προσμίξεις και ακαθαρσίες. Στην πράξη όμως πάντα περιέχονται παγιδευμένες φυσαλίδες αερίων και θραύσματα από πυρίμαχο υλικό τα οποία εισάγουν ατέλειες στη δομή του υλικού. Οι μονωτήρες πορσελάνης είναι πάντα άμορφα υλικά (υαλώδης κατάσταση).

Τα καθαρά πολυμερή αποτελούνται από ένα επαναλαμβανόμενο βασικό είδος μορίων και είναι απαλλαγμένα από ξένες ουσίες. Αυτά σπάνια χρησιμοποιούνται στην κατασκευή μονωτήρων, καθώς η τακτική παγκοσμίως είναι να χρησιμοποιούνται συμπολυμερή και ορυκτά πληρωτικά μέσα για την επίτευξη των επιθυμητών μηχανικών και ηλεκτρικών ιδιοτήτων των παραγόμενων προϊόντων [1].

Όλοι οι μονωτήρες πρέπει να καλύπτουν ηλεκτρικές και μηχανικές λειτουργικές ανάγκες, οι οποίες συνήθως εμφανίζονται στο σχεδιαστή αντιφατικές απαιτήσεις. Στις ηλεκτρικές και μηχανικές ιδιότητες των μονωτήρων θα αναφερθούμε στην επόμενη παράγραφο.

1.1.2 Ιδιότητες των μονωτήρων

Η αξιοπιστία είναι η σημαντικότερη ιδιότητα ενός μονωτήρα. Η μεγαλύτερη αστοχία σε έναν μονωτήρα γραμμής συμβαίνει όταν αυτός σπάσει μηχανικά και «κοπεί» η γραμμή. Ωστόσο, ακόμα κι αν αυτό συμβεί, οι μονωτήρες στους γειτονικούς πυλώνες θα πρέπει να μπορούν όχι μόνο να αναλάβουν το επιπλέον φορτίο, αλλά και να αντέξουν τη μηχανική καταπόνηση που σχετίζεται με ένα τέτοιο γεγονός, ώστε να ελαχιστοποιηθούν τα προβλήματα αποκατάστασης και το κόστος. Για να αποφύγουμε τέτοιου είδους διαδοχικά σφάλματα στη γραμμή, οι μονωτήρες πρέπει να είναι κατάλληλα σχεδιασμένοι. Επιπλέον, οι μηχανικές ιδιότητες των μονωτήρων αποκτούν ακόμη μεγαλύτερη σημασία σε περιοχές με υψηλή σεισμική δραστηριότητα [2].

Η θεμελιώδης διαφορά μεταξύ πολυμερών και κεραμικών υλικών βρίσκεται στη σταθερότητα. Οι ισχυροί ηλεκτροστατικοί δεσμοί πυριτίου – οξυγόνου που συγκρατούν τα κεραμικά υλικά (πορσελάνη και γυαλί) έχουν σαν αποτέλεσμα το υψηλό σημείο τήξεως, τη μεγάλη μηχανική αντοχή (αλλά και το εύθραυστο) και υψηλή αντίσταση διάβρωσης από χημικά μέσα [1]. Από αυτήν την άποψη, όταν συγκρίνουμε τις βασικές φυσικές ιδιότητες, τα ενισχυμένα με ίνες γυαλιού πολυμερή, που είναι περισσότερο ελαστικά, φαίνεται να πλεονεκτούν έναντι των στέρεων πλην όμως εύθραυστων κεραμικών υλικών.

Ένας δεύτερος ανεπιθύμητος τύπος αστοχίας είναι η υπερπήδηση (flashover), που μπορεί να προκληθεί από υπερτάσεις ή ρύπανση. Καθώς όμως οι μονωτήρες είναι πιο φτηνοί από όλα σχεδόν τα υπόλοιπα στοιχεία του συστήματος ενέργειας, μερικές φορές είναι αποδεκτό να επιτρέπουμε επιφανειακές εκκενώσεις στο μονωτήρα, αν αυτό προστατεύει έναν πιο ακριβό εξοπλισμό από μια πολύ μεγάλη υπέρταση. Ο μονωτήρας πρέπει βέβαια σ' αυτήν την περίπτωση να ανθίσταται σε όλες τις αναμενόμενες υπερτάσεις που προκαλούνται από συνηθισμένους χειρισμούς. Γι' αυτού του είδους τις καταπονήσεις, το μήκος τόξου στο μονωτήρα είναι η πιο σημαντική παράμετρος. Από την άλλη, μια αστοχία λόγω υπερπήδησης που οφείλεται σε συνθήκες ρύπανσης σχετίζεται κυρίως με το μήκος ερπυσμού του μονωτήρα και τις επιφανειακές του ιδιότητες [2].

Η ύπαρξη ισχυρών ηλεκτροστατικών χημικών δεσμών κοντά στην επιφάνεια των κεραμικών μονωτήρων συμβάλλει στην αύξηση της ελεύθερης ενέργειας, δηλαδή θερμοδυναμικής ιδιότητας που καθορίζει τη συνάφεια ή τη δύναμη συνάφειας των επιφανειακών προσμίξεων, και αποτελεί την κύρια αιτία για την οποία οι κεραμικοί μονωτήρες υγραίνονται και ρυπαίνονται εύκολα. Αντίθετα, οι δεσμοί των μορίων των πολυμερών είναι ασθενείς και κατά συνέπεια δεν παρατηρείται έντονη επιφανειακή ρύπανση.

Το σημαντικότερο πλεονέκτημα των πολυμερών είναι οι εξαιρετικά χαμηλές τιμές της ελεύθερης επιφανειακής τους ενέργειας. Θερμοδυναμικά, το εξωτερικό περιβάλλον των ατόμων υδρογόνου που είναι συνδεδεμένα με τα άτομα του άνθρακα, παρέχει τη δυνατότητα στα πολυμερή να ανθίστανται στην υγρασία και την ατμοσφαιρική ρύπανση περισσότερο απ' ό,τι τα κεραμικά υλικά. Όταν λοιπόν λαμβάνεται υπόψη και η αντοχή στις διασπάσεις λόγω ρύπανσης, οι μονωτήρες με ελαστομερή προστατευτικά καλύμματα σιλικόνης (silicon rubber housing) αποδεικνύεται ότι υπερτερούν έναντι των παραδοσιακών μονωτήρων από πορσελάνη και γυαλί. Με άλλα λόγια, τα πολυμερή υλικά φαίνεται να είναι ιδανικά για εξωτερικούς μονωτήρες. Όμως, οι περιβαλλοντικές συνθήκες είναι συχνά τέτοιες που οδηγούν στη διάβρωση της επιφάνειας του πολυμερούς. Κάτι τέτοιο οδηγεί στην απώλεια των άριστων επιφανειακών ιδιοτήτων. Οι διαβρωτικές διαδικασίες συμβάλλουν σε μια μη αναστρέψιμη και χρονικά εξαρτώμενη φθορά.

Όλα τα πολυμερή αποσυντίθενται σε θερμοκρασία μερικών εκατοντάδων βαθμών, ενώ επίσης υπόκεινται σε φωτόλυση, δηλαδή σε καταστροφή των επιφανειακών χημικών τους δεσμών από την πρόσπτωση της υπεριώδους ακτινοβολίας και είναι πιθανό να αντιδράσουν κατά τη λειτουργία τους με το ατμοσφαιρικό οξυγόνο, προκαλώντας επιφανειακές ηλεκτρικές εκκενώσεις. Το σημαντικότερο μειονέκτημα των πολυμερών πηγάζει από το γεγονός ότι το βασικό στοιχείο τους, ο άνθρακας, αποτελεί στις περισσότερες περιπτώσεις έναν καλό ηλεκτρικό αγωγό. Η πιθανή προσβολή της επιφάνειας, ιδιαίτερα από τις έρπουσες μερικές εκκενώσεις, αυξάνει τοπικά τη θερμοκρασία και δημιουργεί τελικά αγωγίμες επιφανειακές οδούς που είναι δυνατό να οδηγήσουν σε επιφανειακή διάσπαση και δημιουργία υψηλών θερμοκρασιών, οπότε τελικά καταστρέφεται ο μονωτήρας [1].

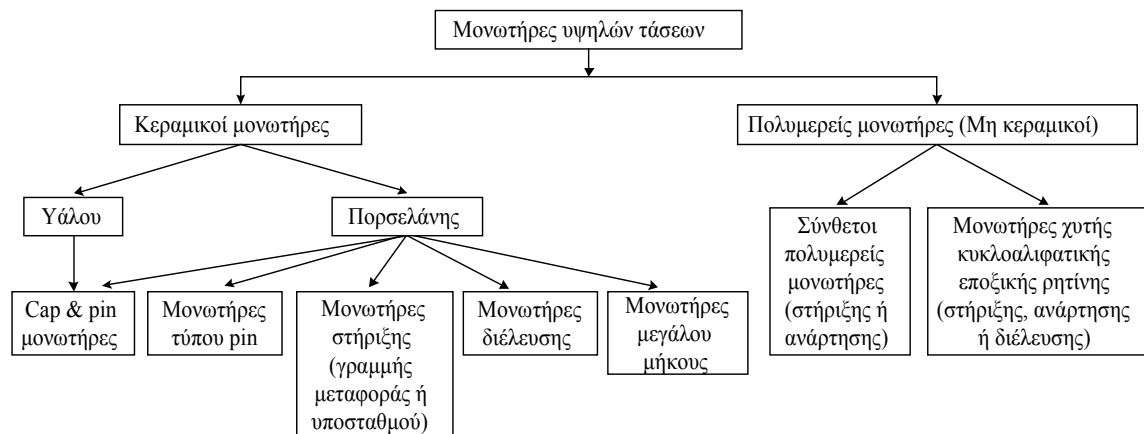
Η αντίσταση στην περιβαλλοντική φθορά μπορεί να αυξηθεί με τη χρήση υλικών που προστατεύουν από τη θερμοκρασία (thermal protective fillers), όπως είναι το ATH (aluminum tri hydrate) ή αντιοξειδωτικών. Στους μονωτήρες υψηλών τάσεων ο κυριότερος λόγος που χρησιμοποιούμε ATH (ATH-filler) είναι η ιδιότητά του να μειώνει την επιφανειακή θερμοκρασία στις περιπτώσεις εμφάνισης τόξου, μέσω της διάσπασής του σε οξείδιο του αλουμινίου και νερό που ακολουθείται από εξάτμιση του νερού.

Όσον αφορά τους κεραμικούς μονωτήρες, είναι αναγκαίο να γίνει εκμετάλλευση των καλών ιδιοτήτων τους, αλλά με ταυτόχρονη μείωση της επιφανειακής υγροσκοπικότητάς τους και της ανεπιθύμητης έλξης ρύπων. Η έρευνα για την επίτευξη αυτών των στόχων συνεχίζεται μέχρι και σήμερα.

1.1.3 Τύποι μονωτήρων

Είναι γνωστό ότι τόσο το υλικό, όσο και το γεωμετρικό σχέδιο των μονωτήρων παίζει έναν πολύ σημαντικό ρόλο για τις μονωτικές τους επιδόσεις, καθώς επηρεάζει τον αυτοκαθαρισμό τους (από τη βροχή), τα μονοπάτια διαρροής, αλλά και την κατανομή της έντασης του ηλεκτρικού πεδίου [2].

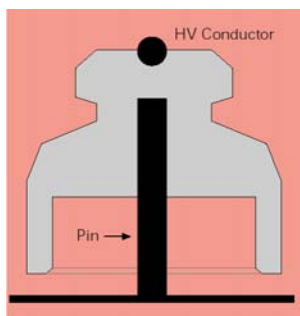
Μια ταξινόμηση των κύριων τύπων μονωτήρων σύμφωνα με το υλικό και το σχήμα τους παρουσιάζεται στο Σχήμα 1.1, ενώ στη συνέχεια αναλύονται οι βασικές κατηγορίες μονωτήρων.



Σχήμα 1.1: Κατάταξη των μονωτήρων υψηλής τάσης

Μονωτήρες πορσελάνης τύπου pin (porcelain pin type insulators)

Αυτός ο τύπος μονωτήρων χρησιμοποιήθηκε αρχικά για τις τηλεφωνικές γραμμές και τα αλεξικέραυνα, ενώ στη συνέχεια τροποποιήθηκε για να χρησιμοποιηθεί στις γραμμές μεταφοράς ηλεκτρικής ενέργειας. Μερικές παραλλαγές του είναι ακόμα σε χρήση στα συστήματα μέσης τάσης. Ο μονωτήρας αυτός παρουσιάζεται στο Σχήμα 1.2.

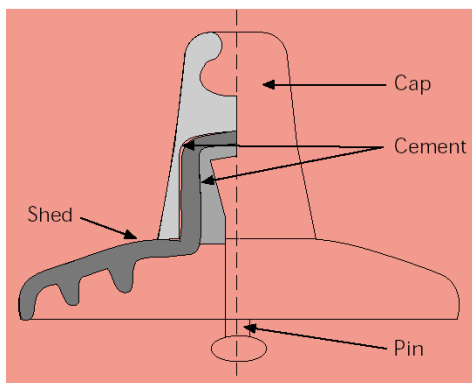


Σχήμα 1.2: Μονωτήρας τύπου pin

Μονωτήρες τύπου cap & pin (cap and pin insulators)

Αυτοί κατασκευάζονται από πορσελάνη ή γυαλί και είναι βασισμένοι στις ίδιες κατασκευαστικές αρχές με τους μονωτήρες τύπου pin. Ένας αριθμός τέτοιων μονάδων συνδέεται μέσω χαλύβδινων καλυμμάτων (caps) και πείρων (pins) σχηματίζοντας μια αλυσίδα μονωτήρων. Αυτές οι αλυσίδες χρησιμοποιούνται ως μονωτήρες ανάρτησης και στήριξης γραμμών μεταφοράς. Τα καλύμματα και ο πείρος είναι στερεωμένα στο γυάλινο ή πορσελάνινο δίσκο με τσιμέντο. Οι κωνικές μορφές των συναρμολογήσεων εξασφαλίζουν υψηλή μηχανική αντοχή κατά την άσκηση δυνάμεων εφέλκυσμού. Στους

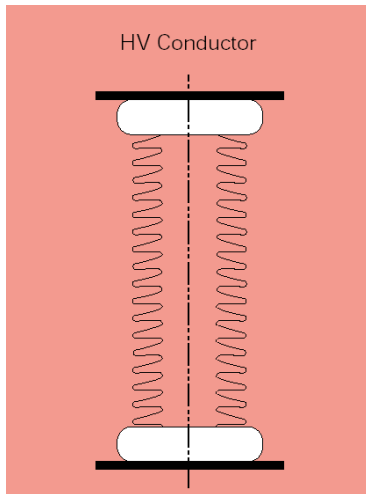
μονωτήρες τύπου pin και τύπου cap & pin η συντομότερη απόσταση μεταξύ των μεταλλικών ηλεκτροδίων μέσω της πορσελάνης ή του γυαλιού είναι λιγότερη από το 50% της συντομότερης απόστασης μεταξύ των ηλεκτροδίων μέσω του αέρα. Συνεπώς η πορσελάνη ή το γυαλί μπορεί να διατηρηθεί κατόπιν μεγάλης ηλεκτρικής καταπόνησης. Η διαδικασία κατασκευής των μονωτήρων υάλου περιλαμβάνει τη θερμική ψύξη, που εξασφαλίζει ότι οι γυάλινοι δίσκοι θρυμματίζονται σε περίπτωση διάτρησης, οπότε ο ελαττωματικός δίσκος είναι εμφανής.



Σχήμα 1.3: Μονωτήρας τύπου cap and pin

Μονωτήρες στήριξης (post type insulators)

Αυτοί οι μονωτήρες αποτελούνται από ένα συμπαγή κύλινδρο πορσελάνης με κυματοειδείς πτυχές, για να αυξάνεται το μήκος ερπυσμού, και από μεταλλικά ηλεκτρόδια σε κάθε άκρη. Οι μονωτήρες στήριξης έχουν μεγάλο μήκος. Χρησιμοποιούνται οριζόντια ή κάθετα για να στηρίζουν τον αγωγό υψηλής τάσης και τοποθετούνται στις βάσεις ή στους εγκάρσιους βραχίονες των γραμμών μεταφοράς (line post insulators). Στους μονωτήρες αυτούς η συντομότερη απόσταση μέσω της πορσελάνης υπερβαίνει το 50% της συντομότερης απόστασης μέσω του αέρα μεταξύ των ηλεκτροδίων, οπότε δεν διαρρηγνύονται εύκολα. Ένας μονωτήρας τύπου post εικονίζεται στο Σχήμα 1.4.



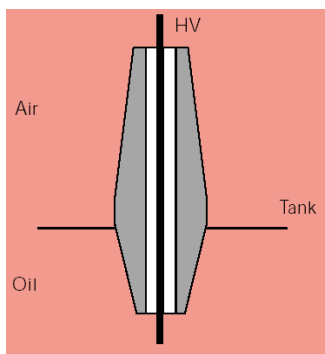
Σχήμα 1.4: Μονωτήρας τύπου post

Μονωτήρες πορσελάνης μεγάλου μήκους (longrod insulators)

Οι μονωτήρες μεγάλου μήκους είναι παρόμοιοι με τους μονωτήρες στήριξης, αλλά είναι ελαφρύτεροι, λεπτότεροι και χρησιμοποιούνται ως μονωτήρες ανάρτησης.

Μονωτήρες διέλευσης (bushings)

Οι μονωτήρες αυτού του τύπου χρησιμοποιούνται για να μονώσουν τους αγωγούς των ακροδεκτών υψηλής τάσης κυρίως μετασχηματιστών ή άλλων διατάξεων, όπως φαίνεται και στο Σχήμα 1.5. Συνήθως οι μονωτήρες διέλευσης κατασκευάζονται από πορσελάνη.



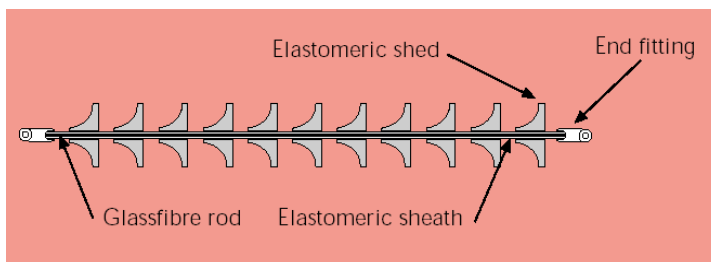
Σχήμα 1.5: Μονωτήρας διέλευσης

Σύνθετοι πολυμερείς μονωτήρες (composite polymeric insulators)

Αυτοί οι μονωτήρες είναι παρόμοιοι με τους πορσελάνινους μονωτήρες μεγάλου μήκους ως προς το σχήμα, αλλά διαφέρουν ως προς τα υλικά κατασκευής, τα οποία είναι τα εξής:

- Ένας πυρήνας ρητίνης ενισχυμένος με ίνες γυαλιού, ο οποίος παρέχει μηχανική αντοχή, ενώ ταυτόχρονα ανθίσταται και στην ηλεκτρική καταπόνηση.
- Ένα ελαστομερές περίβλημα για να προστατεύει τον πυρήνα από τις καιρικές συνθήκες.
- Ελαστομερείς δίσκοι, που αποσκοπούν στην αύξηση του μήκους ερπυσμού, ώστε να ανθίσταται ο μονωτήρας στις εντάσεις που επικρατούν στο σύστημα. Δύο ευρέως χρησιμοποιούμενα υλικά είναι το σιλικονούχο καουτσούκ (silicone rubber) και το αιθυλένιο-προπυλένιο-διένιο μονομερές καουτσούκ (ethylene propylene diene monomer – EDPM).

Οι μεταλλικοί ακροδέκτες αυτών των μονωτήρων είναι συνήθως στερεωμένοι πάνω στην εξωτερική πλευρά της ράβδου του πυρήνα και τα συγκεκριμένα σημεία επαφής μεταξύ του ελαστομερούς και των ακροδεκτών είναι πολύ σημαντικά. Ένα σημαντικό πλεονέκτημα των σύνθετων πολυμερών μονωτήρων είναι η μείωση βάρους μέχρι και 90% σε σύγκριση με τους αντίστοιχους κεραμικούς.



Σχήμα 1.6: Σύνθετος πολυμερής μονωτήρας

Μονωτήρες κυκλοαλιφατικής εποξικής ρητίνης (Cyclo-alifatic epoxy resin insulators)

Η κυκλοαλιφατική ρητίνη μπορεί να χρησιμοποιηθεί για να χυτευτούν μονωτήρες παρόμοιοι με τους πορσελάνης και τύπου line post για διανομή των τάσεων. Σε περιβάλλον με κακές καιρικές συνθήκες οι επιφάνειες των μονωτήρων αυτών γίνονται τραχιές, γεγονός που μπορεί να έχει επιπτώσεις στην αξιοπιστία του μονωτήρα, αν δεν ληφθεί υπόψη.

1.2 Ρύπανση μονωτήρων

Η έκθεση των μονωτικών υλικών στις διάφορες ατμοσφαιρικές συνθήκες ήταν πάντοτε αναπόφευκτη σε όλα τα συστήματα μεταφοράς ηλεκτρικής ενέργειας. Αν και η γνώση μας για τις υπερτάσεις που προκαλούνται από κεραυνικά πλήγματα ή εσωτερικούς χειρισμούς έχει αυξηθεί και οι συσκευές προστασίας από τις υπερτάσεις αυτές έχουν βελτιωθεί, το πρόβλημα της ρύπανσης των μονωτήρων παραμένει ιδιαίτερα οξύ. Το πρόβλημα αυτό εντείνεται και από την αυξανόμενη ατμοσφαιρική ρύπανση, αλλά και τον συνεχώς αυξανόμενο αριθμό των γραμμών μεταφοράς. Με την απελευθέρωση της αγοράς ενέργειας, οι παραδοσιακές δημόσιες επιχειρήσεις θα αντιμετωπίσουν αυξανόμενο ανταγωνισμό, που θα οδηγήσει σε πιέσεις για μείωση του κόστους λειτουργίας και συνεπώς μείωση των τιμών και αύξηση της αξιοπιστίας του συστήματος, που επιτυγχάνεται ως ένα σημείο και από τη βελτίωση της λειτουργίας των μονωτήρων [3]. Κάποιες από τις τεχνικές που χρησιμοποιούνται για την επίτευξη του παραπάνω στόχου περιλαμβάνουν την αύξηση του μήκους ερπυσμού (leakage length) στις πιο «ευάλωτες» τοποθεσίες, την επικάλυψη των επιφανειών των μονωτήρων με υδροφοβικά υλικά και τον περιοδικό καθαρισμό των μονωτήρων ή την αντικατάσταση των πορσελάνινων μονωτήρων από πολυμερείς. Όλες αυτές οι τεχνικές παρουσιάζουν όμως μειονεκτήματα. Οι μη κεραμικές επιφάνειες χάνουν την υδροφοβικότητά τους και υφίστανται φθορά της επιφάνειάς τους έπειτα από παρατεταμένη έκθεση σε ρύπανση και ηλεκτρικές εκκενώσεις. Η καθιέρωση ενός αξιόπιστου προγράμματος καθαρισμού και συντήρησης φαίνεται σε πρώτο επίπεδο να είναι μια ευρέως αποδεκτή πρακτική για τη βελτίωση της κατάστασης. Όμως, το υψηλό κόστος και η ανάγκη για εκπαιδευμένο εργατικό προσωπικό κάνουν το πλύσιμο των μονωτήρων ιδιαίτερο δαπανηρό. Στο Σχήμα 1.7 φαίνεται πως γίνεται το πλύσιμο των μονωτήρων σε μια γραμμή μεταφοράς [3].



Σχήμα 1.7: Πλύσιμο μονωτήρων σε μια γραμμή μεταφοράς

Η αξιοπιστία του συστήματος ενέργειας εξαρτάται λοιπόν κυρίως από τις περιβαλλοντικές και καιρικές συνθήκες που προκαλούν διασπάσεις σε μονωτήρες με ρύπανση, οδηγώντας έτσι σε βλάβες του συστήματος. Είναι γενικά παραδεκτό ότι οι κύριες αιτίες που οδηγούν στη ρύπανση των μονωτήρων είναι η θαλάσσια ρύπανση – που συναντάται στις παράκτιες περιοχές και τα στερεά κατάλοιπα της βιομηχανικής ρύπανσης – που συναντώνται στις βιομηχανικές ζώνες. Η συνύπαρξη ρύπανσης (θαλάσσιας, βιομηχανικής ή συνδυασμού αυτών) και υγρασίας (με τη μορφή δροσιάς, ομίχλης ή σιγανής βροχής) αποτελεί δυσμενέστατη συνθήκη λειτουργίας των μονωτικών υλικών και μπορεί να υποβιβάσει τη μονωτική τους ικανότητα σε μεγάλο βαθμό (40% έως 80%) με αποτέλεσμα τη δημιουργία ηλεκτρικού τόξου.

Στο πρότυπο IEC 815/1986, προκειμένου να υπάρχει μία κατηγοριοποίηση των περιοχών ανάλογα με το ποσοστό ρύπανσης, έχουν καθοριστεί ποιοτικά τέσσερα επίπεδα ρύπανσης. Τα επίπεδα αυτά παρουσιάζονται αναλυτικά στον Πίνακα 1.1.

Πίνακας 1. 1: Επίπεδα ρύπανσης σύμφωνα με το IEC 815 [4]

Επίπεδο ρύπανσης	Παραδείγματα τυπικού περιβάλλοντος
Ελαφριά	<ul style="list-style-type: none"> • Περιοχές χωρίς βιομηχανίες και με μικρή πυκνότητα κατοικιών εξοπλισμένων με εγκατάσταση θέρμανσης • Περιοχές με μικρή πυκνότητα βιομηχανιών ή κατοικιών, αλλά εκτεθειμένες σε ισχυρούς ανέμους και/ή βροχοπτώσεις • Αγροτικές περιοχές • Ορεινές περιοχές

	Όλες αυτές οι περιοχές πρέπει να βρίσκονται τουλάχιστον 10km με 20km από τη θάλασσα και δεν πρέπει να εκτίθενται σε ανέμους κατ' ευθείαν από τη θάλασσα
Μεσαία	<ul style="list-style-type: none"> • Περιοχές με βιομηχανίες, οι οποίες δεν παράγουν ιδιαίτερα ρυπασμένο καπνό, και/ή με μέση πυκνότητα κατοικιών εξοπλισμένων με εγκατάσταση θέρμανσης • Περιοχές με υψηλή πυκνότητα κατοικιών και/ή βιομηχανιών, αλλά εκτεθειμένες σε συχνούς ανέμους και/ή βροχοπτώσεις • Περιοχές εκτεθειμένες σε ανέμους από τη θάλασσα, αλλά όχι πολύ κοντά σε ακτή (τουλάχιστον σε απόσταση μερικών χιλιομέτρων ανάλογα με τη μορφολογία της παράκτιας περιοχής και τις συνθήκες ανέμου).
Βαριά	<ul style="list-style-type: none"> • Περιοχές με υψηλή πυκνότητα βιομηχανιών και προάστια μεγάλων πόλεων με υψηλή πυκνότητα εγκαταστάσεων θέρμανσης, οι οποίες παράγουν ρύπανση • Περιοχές κοντά στη θάλασσα ή κάθε περιοχή που εκτίθεται σε σχετικά ισχυρούς ανέμους από τη θάλασσα
Πολύ βαριά	<ul style="list-style-type: none"> • Περιοχές περιορισμένης έκτασης, εκτεθειμένες σε αγωγίμη σκόνη και βιομηχανικό καπνό που προκαλεί αγωγίμες εναποθέσεις σημαντικού πάχους • Περιοχές περιορισμένης έκτασης, οι οποίες βρίσκονται πολύ κοντά σε ακτή και εκτίθενται σε ψεκασμό θαλασσινού νερού ή σε πολύ δυνατό και ρυπασμένο αέρα από τη θάλασσα • Περιοχές ερήμου, οι οποίες χαρακτηρίζονται από έλλειψη βροχής για μεγάλες χρονικές περιόδους, εκτίθενται σε ισχυρούς ανέμους που μεταφέρουν άμμο και αλάτι και υποβάλλονται σε κανονική συμπίκνωση

Για κάθε επίπεδο ρύπανσης που περιγράφεται στον Πίνακα 1.1, το αντίστοιχο ελάχιστο ειδικό μήκος ερπυσμού (minimal nominal specific creepage distance), μετρούμενο σε

mm/kV (πολική τάση) της υψηλότερης εφαρμοζόμενης τάσης στο μονωτήρα δίνεται στον Πίνακα 1.2 [4].

Πίνακας 1. 2: Ελάχιστο ειδικό μήκος ερπυσμού συναρτήσει του επιπέδου ρύπανσης

Επίπεδο ρύπανσης	Ελάχιστο ειδικό μήκος ερπυσμού (mm/kV)
Ελαφριά	16
Μεσαία	20
Βαριά	25
Πολύ βαριά	31

Η ρύπανση μπορεί να μειώσει την τάση υπερπήδησης του μονωτήρα, οδηγώντας έτσι σε μη προγραμματισμένες βλάβες και μειωμένη αξιοπιστία. Η υπερπήδηση σε ένα μονωτήρα εξαιτίας της ρύπανσης είναι μια περίπλοκη διαδικασία που περνά από διάφορα στάδια. Αυτά ονομαστικά είναι τα εξής:

- Η δημιουργία στρώματος ρύπανσης στην επιφάνεια του μονωτήρα
- Ο σχηματισμός ξηρών ζωνών
- Οι μερικές εκκενώσεις και τέλος, αν οι συνθήκες το ευνοούν
- Η υπερπήδηση

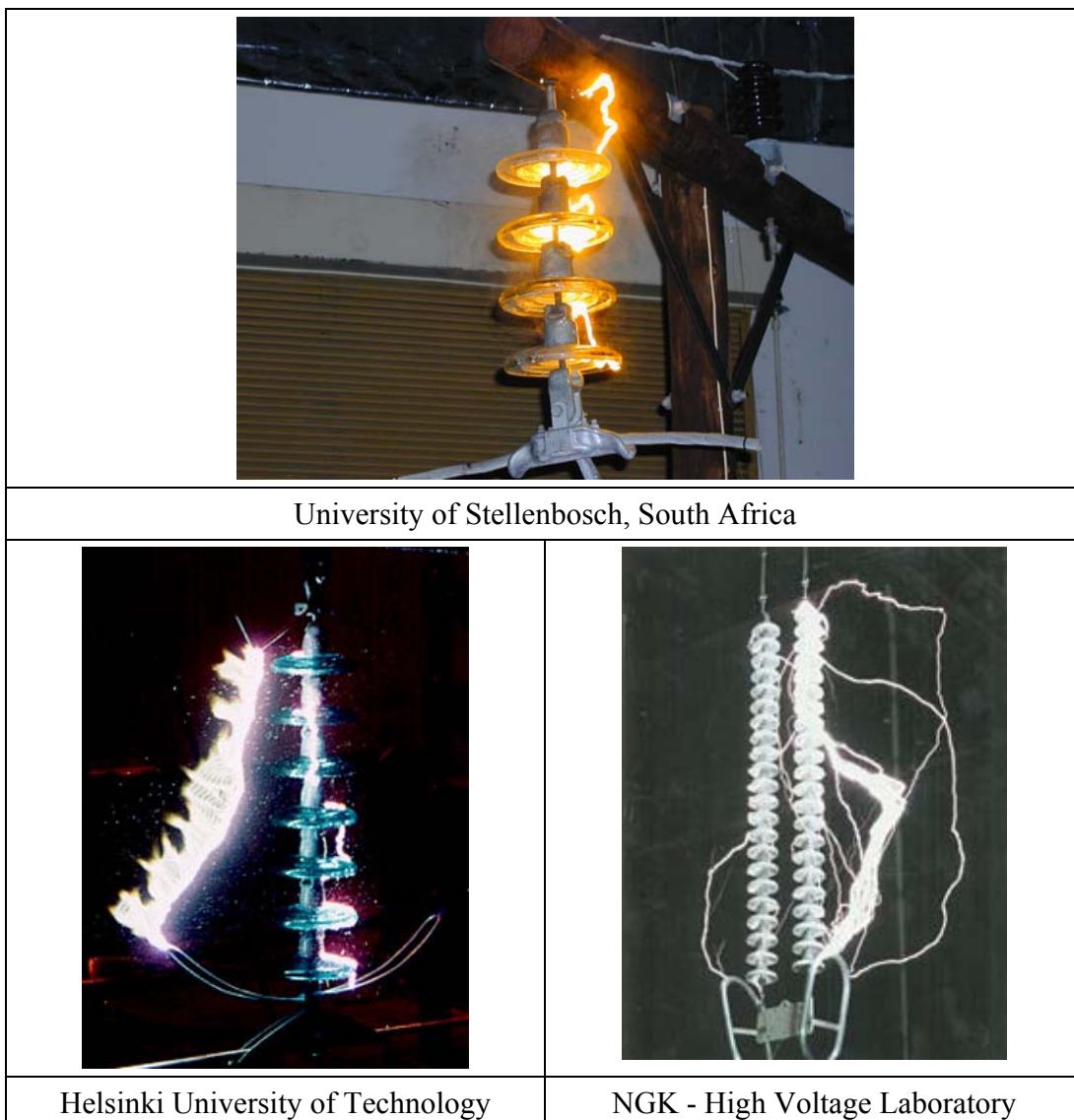
Παρακάτω θα δούμε αναλυτικά τι συμβαίνει στα στάδια αυτά.

1.2.1 Το φαινόμενο της υπερπήδησης στους μονωτήρες

Το φαινόμενο της υπερπήδησης μονωτήρων λόγω ρύπανσης αναφέρεται στο γεγονός ότι δημιουργείται γεφύρωση με ηλεκτρικό τόξο που οδεύει διαμέσου του αέρα του διακένου μεταξύ του σημείου πρόσδεσης του αγωγού της γραμμής στο μονωτήρα και του γειωμένου σημείου στήριξης ή ανάρτησης του μονωτήρα. Το φαινομενικό παράδοξο στην υπερπήδηση μονωτήρων λόγω ρύπανσης είναι ότι παράγονται καταστροφικές ηλεκτρικές εκκενώσεις, που επεκτείνονται σε μέτρα αέρα, από ηλεκτρικά δυναμικά τα οποία, σε κανονικές συνθήκες, θα μπορούσαν να αναχαιτιστούν από διάκενα αέρα μήκους λίγων εκατοστών. Με κάποιον τρόπο, η παρουσία δυσδιάκριτων αγωγίμων σωματιδίων, τα οποία επικάθονται σε μια επιφάνεια που αλλιώς θα ήταν ισχυρά

μονωτική, μειώνει την πραγματική ηλεκτρική αντοχή της κατά έναν παράγοντα όχι πολύ μικρότερο του 100.

Τα αίτια γι' αυτό είναι δύο: α) η εντοπισμένη αφυδάτωση ενός στρώματος ηλεκτρολύτη αυξάνει τις ασυνέχειες στο αγωγίμο στρώμα – οι γνωστές και ως ξηρές ζώνες – κατά μήκος των οποίων παράγονται ηλεκτρικές τάσεις (stresses) που επαρκούν για να ιονιστεί ο αέρας και β) τα τόξα σε ένα αέριο, που από τη στιγμή που δημιουργούνται μπορούν εύκολα να επεκταθούν χωρίς εξασθένηση με σχετικά αργό διαχωρισμό των ηλεκτροδίων μεταξύ των οποίων εμφανίζονται.



Σχήμα 1.8: Διάσπαση μονωτήρων

Για μεγάλο μέρος της ζωής του ένας μονωτήρας θα λειτουργεί με ξηρές ζώνες στην επιφάνειά του τις οποίες κατά καιρούς διαπερνούν ηλεκτρικές εκκενώσεις. Αυτές οι εκκενώσεις είναι ακίνδυνες, εκτός από τα προβλήματα πιθανής πρόκλησης παρεμβολών και επιφανειακών βλαβών για τα οποία ευθύνονται. Μόνο πολύ σπάνια ο συνδυασμός αγωγιμότητας και ηλεκτρικής τάσης θα είναι τέτοιος που θα επιτρέπει την ανάπτυξη ενός τόξου με τόσο ρεύμα, ώστε να αυτο-διατηρείται και να διαδίδεται: τότε προκαλείται διάσπαση. Το τεχνικό πρόβλημα είναι ότι η επιφανειακή αγωγιμότητα που προκαλεί τη διάσπαση παραμένει, ακόμα κι όταν το τόξο έχει εξαλειφθεί από τη λειτουργία της προστασίας, γεγονός που επιτρέπει να ακολουθήσουν κι άλλες διασπάσεις [5]. Παρακάτω (Σχήμα 1.8) παρουσιάζονται χαρακτηριστικές φωτογραφίες του φαινομένου της διάσπασης προερχόμενες από ιστοσελίδες Εργαστηρίων Υψηλών Τάσεων.

1.2.2 Στάδια της υπερπήδησης σε ένα μονωτήρα εξαιτίας της ρύπανσης

Δημιουργία στρώματος ρύπανσης στην επιφάνεια του μονωτήρα

Η ρύπανση, όπως αναφέρθηκε και προηγούμενα, μπορεί να προκληθεί από διάφορες πηγές, όπως είναι η αιωρούμενη τέφρα, το αλάτι της θάλασσας, η σκόνη από τις βιομηχανίες κ.ά. Η εναπόθεση των ρύπων ελέγχεται από την αλληλεπίδραση διαφόρων δυνάμεων που επιδρούν ταυτόχρονα στα σωματίδια που τους αποτελούν (π.χ. της βαρύτητας, του ανέμου και των ηλεκτροστατικών δυνάμεων). Τα αγωγιμα συστατικά των ρύπων επηρεάζουν την τάση διάσπασης του μονωτήρα παρέχοντας, όταν αυτός βραχεί, ένα αγωγίμο στρώμα στην επιφάνειά του. Τα αδρανή συστατικά απ' την άλλη είναι το ποσοστό του στερεού υλικού που δεν διαλύεται, αλλά σχηματίζει ένα μηχανικό περίβλημα στο οποίο ενσωματώνεται το αγωγίμο στρώμα [3].

Υγρασία μπορεί να παραχθεί υπό συνθήκες ομίχλης ή πάχνης κατά τις πρωινές ώρες. Επίσης το ψιχάλισμα και η βροχή μπορεί να έχουν το ίδιο αποτέλεσμα. Έτσι, καταστάσεις που προειδοποιούν για πιθανή διάσπαση του μονωτήρα είναι οι εξής:

1. Η πτώση σχεδόν καθαρού νερού, όπως σταγόνες, βροχή ή ομίχλη, σε ένα μονωτήρα, πάνω στον οποίο βρίσκεται ποσότητα ρύπανσης που περιλαμβάνει διαλυτά ιοντικά στοιχεία, όπως το κοινό αλάτι.
2. Η εναπόθεση σταγονιδίων θαλάσσιας ή βιομηχανικής ομίχλης, ή άλλου συνδυασμού νερού και ηλεκτρολύτη.

3. Η συγκέντρωση παγετού, παγετώδους ομίχλης ή πάγου στη βρώμικη επιφάνεια ενός μονωτήρα. Τότε, τα ιοντικά συστατικά των ρύπων ενεργούν, ώστε να μειώσουν το σημείο ψύξης του νερού κι έτσι επιτρέπουν την ύπαρξη υγρού διαλύματος στο σημείο επαφής των δύο επιφανειών.
4. Η ένταξη ενός κυκλώματος που περιέχει υγρούς και βρώμικους μονωτήρες.
5. Η έλευση μιας προσωρινής υπέρτασης, ή μιας μεταβατικής αιχμής, σε ένα μονωτήρα υγρό, βρώμικο και πιθανόν ήδη ενεργοποιημένο.

Από αυτές τις περιπτώσεις, η πρώτη είναι η πιο συνηθισμένη. Κυρίως σε ερημικές περιοχές, τα περιστατικά διασπάσεων σχετίζονται με περιόδους υγρασίας και πρωινής πάχνης, ενώ στις θαλάσσιες περιοχές οι επικίνδυνες ώρες είναι αυτές της ομίχλης με άπνοια. Ταυτόχρονη εναπόθεση νερού και διαλυμένων ουσιών συμβαίνει συνήθως σε καταιγίδες στην ξηρά.

Η περίπτωση της παγετώδους ομίχλης έχει προκαλέσει μερικά απ' τα πιο σοβαρά περιστατικά, όπως ήταν το 1962 τα πολλαπλά προβλήματα και η προσωρινή διακοπή στο δίκτυο μεταφοράς της Αγγλίας. Το στρώμα ηλεκτρολύτη που προκαλεί το πρόβλημα είναι κολλημένο πάνω στο μονωτήρα και απαιτείται η αφαίρεσή του με καθαρισμό. Η τέταρτη και η πέμπτη περίπτωση είναι λιγότερο συνηθισμένες [5].

Σχηματισμός ξηρών ζωνών

Το στρώμα ηλεκτρολύτη που σχηματίζεται κατά μήκος του μονωτήρα μειώνει την επιφανειακή αντίσταση και ένα υπολογίσιμο ρεύμα διαρροής μπορεί πλέον να κυκλοφορήσει. Το επιφανειακό αυτό ρεύμα έχει σαν αποτέλεσμα την ωμική θέρμανση της επιφάνειας. Η αγωγιμότητα αρχικά αυξάνεται με την αύξηση της θερμοκρασίας, αλλά όταν φτάσει μια συγκεκριμένη τιμή, η εξάτμιση του νερού γίνεται υπολογίσιμη και το διάλυμα γίνεται υπερκορεσμένο σε αλάτι. Η επιφάνεια αρχίζει να ξηραίνεται στις ζώνες με τη μεγαλύτερη διαφεύγουσα ισχύ και η αγωγιμότητα σ' αυτές τις ζώνες πέφτει γρήγορα, έως ότου μηδενιστεί. Η κατανομή της ροής του ρεύματος μεταβάλλεται, εντείνοντας την πλευρική ξήρανση και δημιουργώντας ξηρές ζώνες [3]. Καθώς οι ξηρές ζώνες είναι απομονωμένες, η επιφανειακή δραστηριότητα συνεχίζεται μέσα στην περιοχή της ζώνης.

Μερικές εκκενώσεις και υπερπήδηση

Μετά τη δημιουργία μια ξηρής ζώνης, το μεγαλύτερο μέρος της τάσης που εφαρμόζεται στο μονωτήρα επιβάλλεται στην ξηρή ζώνη, λόγω της υψηλότερης αντίστασης που αυτή έχει. Υπερπήδηση λαμβάνει χώρα εάν μια εκκένωση ξηρής ζώνης επεκταθεί στην εναπομένουσα υγρή επιφάνεια του μονωτήρα. Οι μερικές εκκενώσεις εξαλείφονται λίγο πριν το μηδενισμό της τάσης. Εάν παρ' όλ' αυτά η τιμή της τάσης και του ρεύματος διαρροής είναι αρκετά υψηλή, οι εκκενώσεις μπορεί να επεκταθούν σε ολόκληρο το μήκος του μονωτικού και να αρχίσει το φαινόμενο της υπερπήδησης. Η ορατή δραστηριότητα στην επιφάνεια δεν σημαίνει πάντα ότι θα συμβεί υπερπήδηση, καθώς το ρεύμα διαρροής μπορεί να εξαλειφθεί εάν υπάρχει ανεπαρκής διαφεύγουσα ισχύς. Αν όμως το τόξο επεκταθεί και καταφέρει να καλύψει ένα κρίσιμο μήκος, τότε η διάσπαση είναι πρακτικά αναπόφευκτη.

Τα παραπάνω ισχύουν για τους μονωτήρες πορσελάνης. Στους πολυμερείς μονωτήρες υπάρχουν κάποιες διαφοροποιήσεις στη διαδικασία της διάσπασης. Τα στάδια αυτής είναι:

Εναπόθεση ρύπανσης: Πραγματοποιείται με τον ίδιο τρόπο, όπως και στους μονωτήρες πορσελάνης.

Υγρανση: Επειδή οι πολυμερείς μονωτήρες έχουν υδρόφοβες επιφάνειες, η υγρασία σχηματίζει σταγονίδια πάνω σ' αυτές, τα οποία κυλούν εξαιτίας της βαρύτητας. Όταν η βαρύτητα δεν είναι ικανή να ενισχύσει την κίνηση των σταγονιδίων, αυτά παραμένουν διαχωρισμένα στις θέσεις τους. Άλατα και/ή αγωγίμη ρύπανση διαλύονται στις σταγόνες νερού αυξάνοντας την αγωγιμότητα του υγρού.

Σχηματισμός αγωγίμου στρώματος: Η παραμένουσα ξηρή επιφάνεια ρύπανσης σταδιακά υγραίνεται μέσω της μετακίνησης των σταγονιδίων. Αυτό σχηματίζει ένα αγωγίμο στρώμα ισχυρά ωμικής συμπεριφοράς με αποτέλεσμα το ρεύμα διαρροής να μεταβάλλεται από χωρητικό σε ωμικό.

Θερμότητα (ωμική): Πραγματοποιείται με τον ίδιο τρόπο, όπως και στους μονωτήρες πορσελάνης.

Επίδραση του ηλεκτρικού πεδίου σε υδροφοβική επιφάνεια: Το εφαρμοζόμενο ηλεκτρικό πεδίο προκαλεί την ένωση των σταγονιδίων, που βρίσκονται σε μικρή απόσταση μεταξύ τους, σε μια μεγάλη σταγόνα νηματοειδούς μορφής. Ο μηχανισμός της υπερπήδησης διαρκεί περισσότερο σε μια υδροφοβική επιφάνεια εξαιτίας του χρόνου που απαιτείται για να σχηματιστεί μια αγωγή διαδρομή με νηματοειδείς σταγόνες.

Μερικές εκκενώσεις σε υδροφοβικές επιφάνειες: Οι νηματοειδείς σταγόνες μειώνουν την απόσταση μεταξύ των άκρων του μονωτικού, αυξάνοντας το ηλεκτρικό πεδίο μεταξύ των γειτονικών νηματοειδών σταγόνων. Όταν η τάση αποκτήσει κατάλληλη τιμή, τότε μπορεί να συμβούν επιφανειακές εκκενώσεις.

Εξασθένιση της υδροφοβικότητας: Οι εκκενώσεις καταστρέφουν ένα λεπτό στρώμα πολυμερούς γύρω από τα σταγονίδια και μειώνουν την υδροφοβικότητα, περιστρέφοντας ή σπάζοντας τις πολυμερείς αλυσίδες. Η απώλεια ή η εξασθένιση της επιφανειακής υδροφοβικότητας έχει ως αποτέλεσμα την εξάπλωση των σταγόνων και το σχηματισμό ενός συνεχούς αγωγίμου στρώματος, επιτρέποντας τη ροή του ρεύματος διαρροής.

Σχηματισμός ξηρής ζώνης: Οι ξηρές ζώνες σχηματίζονται όπως και στους μονωτήρες πορσελάνης. Η προκύπτουσα δραστηριότητα προκαλεί επιφανειακή διάβρωση, της οποίας ο ρυθμός εξαρτάται από τον τρόπο σχηματισμού του συγκεκριμένου υλικού και συμβάλλει στη γήρανση.

Αποκατάσταση της υδροφοβικότητας: Η ολική ή μερική αποκατάσταση της υδροφοβικότητας είναι δυνατή μόνο αν το υλικό εκφορτιστεί ελεύθερα για ικανό χρονικό διάστημα. Η διαδικασία αποκατάστασης εξαρτάται από το υλικό, τον τρόπο σχηματισμού του, το σχεδιασμό του μονωτήρα και το περιβάλλον.

Επανάληψη του κύκλου γήρανσης: Η επανάληψη του κύκλου γήρανσης προκαλεί επιπλέον ύγρανση και πιθανόν διάβρωση της επιφάνειας, φαινόμενα τα οποία ενισχύονται από χημικές αντιδράσεις και τοπική άνοδο της θερμοκρασίας. Κατά τη διάρκεια εκκένωσης η θερμοκρασία σε κάποια σημεία μπορεί να φτάσει τους 400°C.

Υπερπήδηση: Η υπερπήδηση μπορεί να συμβεί, ακολουθώντας την ίδια διαδικασία όπως σε μονωτήρες πορσελάνης, εάν η επιφάνεια γίνει υδρόφιλη.

Οι υδρόφοβες επιφάνειες αντιστέκονται περισσότερο στη ροή του ρεύματος διαρροής από τις υδρόφιλες επιφάνειες και απαιτούν υψηλότερο ρεύμα διαρροής και αντίστοιχη έκλυση ενέργειας, προκειμένου να ξεκινήσει η υπερπήδηση. Αυτός είναι ο λόγος για τον οποίο οι πολυμερείς μονωτήρες παρουσιάζουν μεγαλύτερη τάση υπερπήδησης από τους μονωτήρες πορσελάνης. Όπως και στους πορσελάνινους μονωτήρες, όλοι οι παραπάνω μηχανισμοί πρέπει να συμβούν διαδοχικά για να οδηγήσουν σε υπερπήδηση. Εάν η διαδικασία διακοπεί, η υπερπήδηση δεν λαμβάνει χώρα.

1.2.3 Τρόποι μετριασμού του φαινομένου της υπερπήδησης

Για τους μονωτήρες πορσελάνης τα μέτρα που μπορούν να ληφθούν για να μειωθεί η συχνότητα των διασπάσεων είναι τα εξής:

1. **Επέκταση του μήκους:** Πολυμερείς δίσκοι εγκαθίστανται ακριβώς πάνω από τους δίσκους του πορσελάνινου μονωτήρα, προκειμένου να αυξηθεί το μήκος ερπυσμού και να προστεθεί η ιδιότητα της υδροφοβικότητας.
2. **Πλύσιμο:** Οι μονωτήρες είναι δυνατό να πλυθούν, είτε εντός λειτουργίας, είτε εκτός λειτουργίας, με τη χρήση νερού υπό υψηλή πίεση. Επειδή, όμως, το πλύσιμο των μονωτήρων αποτελεί μια δαπανηρή διαδικασία πρέπει να πραγματοποιείται κατόπιν σωστού προγραμματισμού.
3. **Πολύπλοκο σχήμα του μονωτικού υλικού:** Είναι διαθέσιμοι μονωτήρες, με υψηλότερο κόστος, οι οποίοι, λόγω του σχήματός τους αντιστέκονται στην εναπόθεση ρύπανσης.
4. **Επιφανειακές επιστρώσεις:** Γράσο ή πολυμερές επιστρώση εφαρμόζεται στην πορσελάνη προκειμένου να βελτιωθούν οι ιδιότητές της υπό συνθήκες ρύπανσης. Η βελτίωση των ιδιοτήτων της και η διάρκεια μέχρι να απαιτηθεί επανάληψη της

επίστρωσης εξαρτώνται από την ποιότητα του επιστρώματος, την εφαρμογή του και τις τοπικές συνθήκες ρύπανσης.

5. Υάλωση: Μία μη αγώγιμη ή ημιαγώγιμη υάλωση διατηρεί μια μεγάλη περιοχή της επιφάνειας ξηρή κατά τη διάρκεια της φυσικής ύγρανσης. Σε αυτού τους μονωτήρες, όμως, διαρκή ρεύματα διαρροής οδηγούν σε μεγάλες απώλειες ισχύος. Η φθορά της υάλωσης και η έκλυση θερμότητας είναι, επίσης, μεγάλα προβλήματα αυτών των μονωτήρων.

Στα πολυμερή υλικά σπάνια χρησιμοποιούνται πρόσθετα μέτρα για την αποφυγή διασπάσεων. Απαιτείται όμως ιδιαίτερη προσοχή κατά τη δημιουργία του υλικού και το σχεδιασμό του μονωτήρα. Τα πολυμερή μονωτικά υλικά είναι απαραίτητα να έχουν δύο αμυντικούς μηχανισμούς κατά της ρύπανσης και κατ' επέκταση της δημιουργίας διασπάσεων. Αυτοί είναι:

1. Έλεγχος του ρεύματος διαρροής: Επειδή είναι πιθανό να συμβεί φθορά του υλικού κατά τη διάρκεια της λειτουργίας, ένας αμυντικός μηχανισμός είναι ο περιορισμός του ρεύματος διαρροής. Εάν το ρεύμα διαρροής δεν έχει ικανή τιμή, ώστε να προκαλέσει ξήρανση του ηλεκτρολύτη και δημιουργία ξηρών ζωνών, τότε ο μηχανισμός της υπερπήδησης δεν μπορεί να εξελιχθεί. Το παραπάνω μπορεί να επιτευχθεί τόσο με την επιλογή του κατάλληλου υλικού, όσο και με τον κατάλληλο σχεδιασμό του μονωτήρα.
2. Αντοχή στη διάβρωση: Τα πολυμερή υλικά δεν θα πρέπει να υφίστανται σημαντική διάβρωση κατά τη διάρκεια αναπόφευκτων περιόδων, κατά τις οποίες η επιφανειακή δραστηριότητα μπορεί να προκαλέσει απώλεια της υδροφοβικότητας. Ειδικά, τα υλικά που έχουν την τάση να ανακτούν την υδροφοβικότητά τους, πρέπει να αντέχουν περιόδους επιφανειακής δραστηριότητας χωρίς να υφίστανται σημαντική διάβρωση. Ένα πολυμερές υλικό, το οποίο δεν έχει καλή αντοχή στη διάβρωση, δεν μπορεί να τοποθετηθεί σε περιοχή με έντονη ρύπανση, γιατί θα οδηγήσει σε πρόωρη βλάβη. Ένα υλικό με μεγάλη αντοχή στη διάβρωση, ακόμα κι αν χάσει την υδροφοβικότητά του, υπερτερεί ενός υλικού, το οποίο μπορεί να επανακτήσει την υδροφοβικότητά του, αλλά έχει μικρή αντοχή στη διάβρωση.

1.2.4 Παρακολούθηση του ρεύματος διαρροής

Το επίπεδο της ρύπανσης συνήθως καθορίζεται από τη μέτρηση της αγωγιμότητας ενός διαλύματος ρύπων, που αφαιρούνται από την επιφάνεια του μονωτήρα, σε αποσταγμένο νερό. Ένας τρόπος να ελεγχθεί το επίπεδο της ρύπανσης είναι μέσω του καθορισμού του ESDD (Equivalent Salt Deposit Density – Ισοδύναμη Πυκνότητα Εναπόθεσης Αλατιού) σε mg/cm^2 . Το μέγεθος αυτό εκφράζει την ισοδύναμη ποσότητα NaCl που πρέπει να διαλυθεί στην ίδια ποσότητα νερού, ώστε να επιτευχθεί η ίδια αγωγιμότητα. Το ESDD ερμηνεύει μόνο το διαλυτό μέρος της ρύπανσης. Η μέτρηση του ESDD είναι μια έμμεση και στατική μέθοδος για τον προσδιορισμό της “κατάστασης” του μονωτήρα και παρέχει μόνο ένα στιγμιότυπο γι’ αυτήν.

Μια καλύτερη επιλογή είναι να ελέγχεται το ρεύμα διαρροής στους μονωτήρες, μέγεθος που μπορεί να σχετιστεί άμεσα με το πόσο κοντά είναι ο εκάστοτε μονωτήρας στη διάσπαση. Γενικά, το στρώμα ρύπανσης στους μονωτήρες σχηματίζεται κατά τη διάρκεια μηνών ή ετών, με ενδιάμεσες περιόδους μεγάλης υγρασίας. Ο έλεγχος του ρεύματος διαρροής κατά τη διάρκεια αυτών των περιόδων παρέχει έναν πιο άμεσο και αξιόπιστο τρόπο για τον προσδιορισμό της συμπεριφοράς του μονωτήρα με ρύπανση. Μετά από αυτό θα είναι εφικτό να δούμε πώς η ρύπανση αυξάνει με το χρόνο, αλλά και την επίδραση της βροχής (φυσικός καθαρισμός) και να αποφασίσουμε για τον τεχνητό καθαρισμό ή πλύσιμο των μονωτήρων. Ο Πίνακας 1.3 δείχνει πώς το ρεύμα διαρροής αποτελεί ένδειξη για την κατάσταση του μονωτήρα [3].

Πίνακας 1. 3: Το ρεύμα διαρροής ως ένδειξη της κατάστασης του μονωτήρα

Ρεύμα διαρροής	Συμπτώματα
0-2 mA	Δεν υπάρχουν ορατά ή ακουστικά αποτελέσματα
3-5 mA	Ελαφρύς θόρυβος
5-10 mA	Σπινθήρες ορατοί τη νύχτα
10-15 mA	Κάψιμο των ξύλινων βραχιόνων στήριξης
15-100 mA	Έντονος σπινθηρισμός
Πάνω από 100 mA	Ακαριαία υπερπήδηση

Σε κάποιες περιπτώσεις που η ρύπανση συσσωρεύεται στην επιφάνεια του μονωτήρα και αμέσως μετά ακολουθεί βρέξιμο αυτής, δεν μεσολαβεί αρκετό χρονικό διάστημα για να γίνει καθαρισμός. Σ' αυτήν την περίπτωση ο real-time έλεγχος του ρεύματος διαρροής θα μπορούσε να προειδοποιεί για την έλευση της κρίσιμης αυτής κατάστασης και το φορτίο θα μπορούσε να μοιράζεται σε άλλες γραμμές, αν βεβαίως υπάρχει τέτοια εναλλακτική περίπτωση [3].

1.3 Σύνοψη

Στο κεφάλαιο αυτό έγινε μια εισαγωγή στη θεωρία των μονωτήρων. Στο πρώτο μέρος παρουσιάζονται τα υλικά απ' τα οποία αυτοί κατασκευάζονται, οι ιδιότητές τους και οι διάφοροι τύποι μονωτήρων που συναντώνται στην πράξη. Στο δεύτερο μέρος γίνεται αναφορά στο φαινόμενο της υπερπήδησης υπό συνθήκες ρύπανσης και στους τρόπους μετριάσμού του. Στο δεύτερο κεφάλαιο θα παρουσιαστούν κάποιες τεχνικές για τον ποιοτικό έλεγχο των μονωτήρων, όπως αυτές έχουν δημοσιευτεί σε περιοδικά και άρθρα.

Κεφάλαιο 2

Νέες τεχνικές για ποιοτικό έλεγχο μονωτήρων

2.1 Εισαγωγή

Τα τελευταία χρόνια παρατηρείται μια ολοένα αυξανόμενη ζήτηση ηλεκτρικής ενέργειας, γεγονός που έχει οδηγήσει στην αύξηση του επιπέδου της τάσεως στα δίκτυα μεταφοράς, ή σε σπανιότερες περιπτώσεις στη δημιουργία νέων δικτύων. Σκοπός είναι η καλύτερη και οικονομικότερη μεταφορά περισσότερης ενέργειας σε μεγαλύτερες αποστάσεις. Τα νέα όμως αυτά δεδομένα δημιουργούν και καινούριες απαιτήσεις σε ό,τι αφορά τους μονωτήρες. Καθώς το επίπεδο της τάσης στα δίκτυα μεταφοράς αυξάνει, οι εξωτερικές και εσωτερικές υπερτάσεις (υπερτάσεις χειρισμών), αλλά και η αντοχή των μονωτήρων σ' αυτές, κυρίως υπό συνθήκες ρύπανσης, έχουν γίνει πολύ σημαντικοί παράγοντες που επηρεάζουν ως ένα βαθμό την αξιοπιστία του συστήματος.

Όπως αναφέρθηκε και στο πρώτο κεφάλαιο, μια ευρέως διαδεδομένη τακτική είναι η καθιέρωση προγράμματος για το πλύσιμο και τη συντήρηση των μονωτήρων, που όμως παρουσιάζει το σημαντικό μειονέκτημα του αυξημένου κόστους. Είναι λοιπόν πολύ σημαντικό να υπάρχουν τρόποι ελέγχου της κατάστασης των μονωτήρων, ώστε να εξασφαλίζεται ότι ένα πρόγραμμα καθαρισμού είναι πράγματι αναγκαίο κι ότι δεν γίνεται πρόωρα. Επιπλέον, η αυξημένη ζήτηση ενέργειας φέρνει τους μονωτήρες στο προσκήνιο και νέες τεχνολογίες αναπτύσσονται για τον ποιοτικό έλεγχό τους. Στο κεφάλαιο αυτό θα αναφερθούν κάποιες από τις τεχνολογίες αυτές, μεταξύ των οποίων είναι και η εφαρμογή των νευρωνικών δικτύων για την εκτίμηση της κατάστασης ενός μονωτήρα και την αποφυγή του φαινομένου της υπερπήδησης που μπορεί να οδηγήσει σε σοβαρά προβλήματα.

2.2 Συσκευές ελέγχου ρύπανσης (Pollution Monitors)

Ένας αξιόπιστος έλεγχος για την κατάσταση του μονωτήρα γίνεται μέσω των συσκευών ελέγχου ρύπανσης (pollution monitors). Οι συσκευές αυτές χρησιμοποιούνται για την

επιλογή κατάλληλων περιόδων συντήρησης και την αποφυγή απρόβλεπτων διασπάσεων. Τα συστήματα αυτά κάνουν γενικά δυνατό το πέρασμα από τα προγράμματα συντήρησης που βασίζονται στο χρόνο, στα προγράμματα συντήρησης που βασίζονται στην κατάσταση του μονωτήρα, οπότε περιορίζουν τις προγραμματισμένες διακοπές του συστήματος για το πλύσιμο των μονωτήρων, εκλογικεύοντας έτσι το κόστος συντήρησης. Οι συσκευές ελέγχου ρύπανσης μπορεί να χρησιμοποιηθούν ώστε:

- Να μετρούν τη ρύπανση κάποιας τοποθεσίας, για να καθορίσουν τις πιο επιβαρημένες τοποθεσίες για τη λειτουργία των μονωτήρων.
- Να ελέγχουν την κατάσταση του μονωτήρα σχετικά με το επίπεδο ρύπανσης, για λόγους συντήρησης, ώστε να καθορίζεται πότε χρειάζεται πλύσιμο ή/και λίπανση των μονωτήρων για την αποφυγή υπερπηδήσεων εξαιτίας της ρύπανσης.
- Να συγκρίνουν την αποτελεσματικότητα των διαφορετικών σχεδίων μονωτήρων (σχήματα και μήκη) και/ή των υλικών των μονωτήρων υπό συγκεκριμένες συνθήκες ρύπανσης και να καθορίζουν την καλύτερη λύση.

Ο έλεγχος του ρεύματος διαρροής σε έναν μονωτήρα για την αποφυγή υπερπήδησης ή για τη σύγκριση της λειτουργίας του με άλλον τύπο μονωτήρα μπορεί να επιτευχθεί με την καταγραφή του πραγματικού πλάτους των παλμών ρεύματος σε μία χρονική περίοδο, ή με την αποθήκευση αυτών των τιμών σε μια μονάδα μνήμης της συσκευής ελέγχου ή στο σκληρό δίσκο ενός υπολογιστή που είναι μόνιμα συνδεδεμένος στη συσκευή, πράγμα που γενικά είναι δύσκολο να γίνει στους υποσταθμούς και σχεδόν αδύνατο στις περιπτώσεις γραμμών μεταφοράς.

Εναλλακτικά, η ίδια πληροφορία θα μπορούσε να εξαχθεί από την αποθήκευση της στατιστικής κατανομής των παλμών ρεύματος αντί του πλάτους. Το πλεονέκτημα είναι μια μεγαλύτερη περίοδος απόκτησης δεδομένων και μια μικρότερη, πιο οικονομική και εύκολη στη χρήση συσκευή ελέγχου του ρεύματος [3].

2.2.1 Συσκευές ελέγχου του ρεύματος διαρροής

Ένα σύστημα ελέγχου σαν κι αυτά που αναφέρθηκαν μπορεί να ελέγχει το ρεύμα διαρροής (βλ. παράγραφο 1.2.4, Κεφαλαίου 1). Ένα τέτοιο σύστημα αποτελείται από τρία μέρη: τη μονάδα συλλογής (acquisition unit), το μετατροπέα (transducer box), που

περιλαμβάνει κυκλώματα προστασίας και ένα φιλικό προς το χρήστη λογισμικό για την ανάκτηση και ανάλυση των δεδομένων.

2.2.2 Λειτουργία των συσκευών ελέγχου ρεύματος διαρροής

Οι παράμετροι συλλογής των συσκευών ελέγχου του ρεύματος διαρροής τίθενται με τη χρήση ενός ηλεκτρονικού υπολογιστή εκτός του χώρου λειτουργίας και πριν την εγκατάσταση της συσκευής. Το σύστημα στη συνέχεια εγκαθίσταται στην επιθυμητή τοποθεσία και ξεκινά η συλλογή των δεδομένων. Αυτές οι τιμές αναλύονται στη συνέχεια από τη μονάδα επεξεργασίας. Το σύστημα αποτυπώνει τα δεδομένα για ένα προκαθορισμένο διάστημα, σώνει τα δεδομένα σε μια μονάδα μνήμης και ξεκινά μια νέα παρτίδα συλλογής.

2.2.3 Εφαρμογές των παραπάνω συστημάτων

Τα συστήματα ελέγχου ρεύματος διαρροής μπορούν να χρησιμοποιηθούν στις ακόλουθες εφαρμογές:

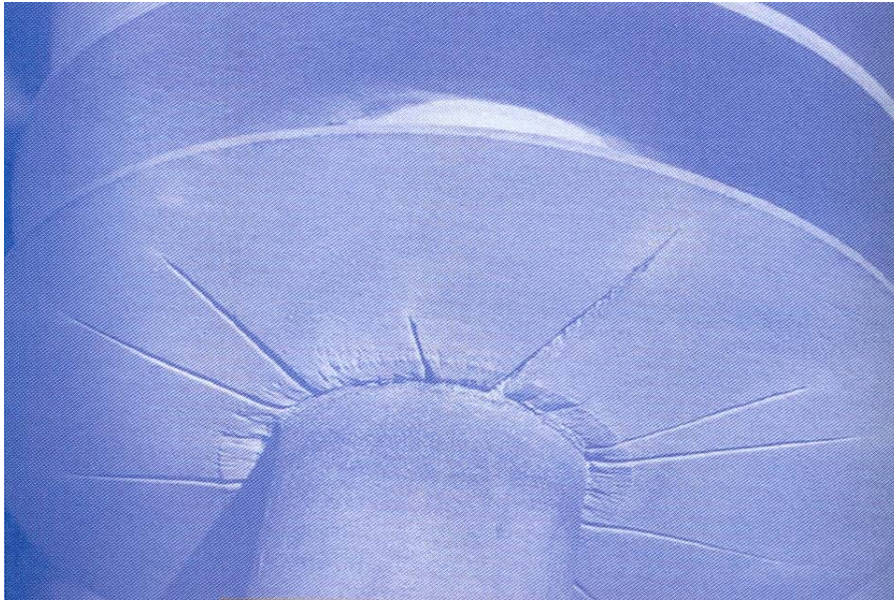
1. Τον καθορισμό των απωλειών ενός μονωτήρα. Για την ακρίβεια, τον καθορισμό των απωλειών ενέργειας σε kWh, κατά τη διάρκεια λειτουργίας της συσκευής.
2. Τον καθορισμό του επιπέδου ρύπανσης μιας περιοχής. Μπορούν να χρησιμοποιηθούν για τον προσδιορισμό των ζωνών ρύπανσης σε κάποιες περιοχές. Αυτό παρέχει μια δυναμική χαρτογράφηση της σοβαρότητας της ρύπανσης σε ένα συγκεκριμένο μέρος, άσχετα με το σχέδιο ή τον τύπο του μονωτήρα.
3. Τον έλεγχο της αξιοπιστίας διαφόρων σχεδίων μονωτήρων και τη συγκριτική μελέτη της λειτουργίας μονωτήρων που βρίσκονται στην ίδια περιοχή και κάτω από τις ίδιες συνθήκες ρύπανσης.
4. Την καθιέρωση προγραμμάτων συντήρησης που βασίζονται στην κατάσταση των μονωτήρων για την αντικατάσταση των προγραμμάτων συντήρησης που βασίζονται στο χρόνο. Αυτό εξασφαλίζει ότι η συντήρηση γίνεται όταν πραγματικά είναι απαραίτητη, γεγονός που περιορίζει τα έξοδα σε λογικά πλαίσια και μειώνει το κόστος.
5. Τέτοια συστήματα μπορούν να εγκατασταθούν σε μονωτήρες γραμμών μεταφοράς, ασφαλειοδιακόπτες, μετασχηματιστές, πυκνωτές κ.ά. [3].

2.3 UV κάμερα για την επιθεώρηση σύνθετων μονωτήρων

Η ανάπτυξη μιας UV κάμερας που επιτρέπει την καταγραφή των μερικών εκκενώσεων, ακόμα και κατά τη διάρκεια της ημέρας είναι μια πολύ θετική εξέλιξη για τις ηλεκτρικές εφαρμογές, κι αυτό γιατί οι κάμερες αυτές προσφέρουν ακόμη ένα πρακτικό εργαλείο για τις συνηθισμένες επιθεωρήσεις γραμμών και εξοπλισμού. Ωστόσο, δεν είναι συχνά δυνατό να φτάσουμε σε κάποιο συμπέρασμα σχετικά με την κατάσταση των μονωτήρων βασιζόμενοι μόνο σε οπτικά δεδομένα – κυρίως αν αυτοί είναι σύνθετοι / μη κεραμικοί μονωτήρες. Στο άρθρο του καθηγητή Ravi S. Gorur του Πανεπιστημίου της Αριζόνα [6] γίνεται αναφορά στην τρέχουσα έρευνα που αφορά πιθανούς τρόπους αποτίμησης της σοβαρότητας των μερικών εκκενώσεων που εμφανίζονται σε συνθετικούς μονωτήρες – με καλύμματα σιλικόνης, ή EPDM. Μια τέτοια πληροφορία έχει ιδιαίτερη σημασία όταν πρόκειται να προσδιοριστεί ποιοι συγκεκριμένοι τύποι εκκενώσεων είναι πιθανότερο να δημιουργήσουν προβλήματα στην ακεραιότητα ενός μονωτήρα.

Οι μερικές εκκενώσεις μπορεί να αποτελέσουν σημαντική απειλή για τους μη κεραμικούς μονωτήρες (non-ceramic insulators – NCIs), εξαιτίας της οργανικής φύσης των υλικών των περιβλημάτων τους. Για μονωτήρες μεταφοράς (69kV και άνω), οι μερικές εκκενώσεις μπορεί να εμφανιστούν όχι μόνο σε ρυπασμένο, αλλά και σε καθαρό περιβάλλον. Στους μη κεραμικούς μονωτήρες οι μερικές εκκενώσεις μπορεί να εγκατασταθούν για μεγάλα χρονικά διαστήματα εξαιτίας παραγόντων όπως ο κακός σχεδιασμός του εξοπλισμού, βλάβες στον εξοπλισμό ή επιφανειακές ατέλειες. Στην πράξη είναι δύσκολο να αποφευχθεί το φαινόμενο των μερικών εκκενώσεων στα δίκτυα μεταφοράς, ειδικά κάτω από συνθήκες υγρασίας και ρύπανσης. Εντούτοις είναι σημαντική η γνώση της έντασης του φαινομένου, καθώς και τα όρια αντοχής του υλικού του καλύμματος.

Στο Σχήμα 2.1 φαίνεται η φθορά που προκάλεσαν οι μερικές εκκενώσεις σε ένα μη κεραμικό μονωτήρα. Ο συγκεκριμένος μονωτήρας είναι προφανές ότι βρίσκεται σε προχωρημένο στάδιο φθοράς. Τέτοιου είδους βλάβες που οδηγούν σε σπάσιμο των καλυμμάτων μπορούν να εκθέσουν τον ενισχυμένο με ίνες γυαλιού πυρήνα σε υγρασία, οδηγώντας έτσι σε πρόωρες αστοχίες λόγω διάβρωσης. Έτσι, είναι αναγκαίο οι μη κεραμικοί μονωτήρες να επιθεωρούνται περιοδικά, ώστε να αντικαθίστανται οι φθαρμένοι μονωτήρες.



Σχήμα 2. 1: Φθορά μη κεραμικού μονωτήρα 230kV από μερικές εκκενώσεις.

Γενικά, οι σύγχρονες μέθοδοι που χρησιμοποιούνται για την επιθεώρηση συνθετικών μονωτήρων σε λειτουργία περιλαμβάνουν ελέγχους με [6]:

- ηχητική εκπομπή (acoustic emission),
- μέτρηση ραδιοφωνικών παρεμβολών (Radio interference voltage measurement- RIV),
- υπέρυθη (IR) θερμογραφία (Infrared thermography),
- μέτρηση ηλεκτρικού πεδίου,
- οπτική παρατήρηση.

Τόσο η ηχητική εκπομπή, όσο και η μέτρηση RIV είναι ευαίσθητες στο θόρυβο. Συνεπώς οι μέθοδοι αυτές δεν μπορούν να χρησιμοποιηθούν για τον αξιόπιστο καθορισμό της ακριβούς θέσης της εκκένωσης πάνω σ' ένα μονωτήρα. Αυτό αποτελεί σημαντικό μειονέκτημα, καθώς οι εκκενώσεις στον εξοπλισμό ενός μονωτήρα δεν μπορούν να αποφευχθούν πάντα και ενδεχομένως σε κάποιες περιπτώσεις να μην αποτελούν πρόβλημα, ωστόσο εκκενώσεις στο περίβλημα του μονωτήρα μπορεί να είναι προάγγελοι σοβαρών προβλημάτων.

Ο καθορισμός της επιφανειακής θερμοκρασίας με τη μέθοδο της υπέρυθρης θερμογραφίας έχει κάποια επιτυχία. Παρ' όλ' αυτά, η μέθοδος αυτή δεν μπορεί να χρησιμοποιηθεί για τον εντοπισμό μικρών ατελειών που δεν προκαλούν κάποια σημαντική αύξηση της επιφανειακής θερμοκρασίας. Η αξιοπιστία της μεθόδου αυτής

είναι επίσης μειωμένη όταν εκτελείται υπό συνθήκες υψηλής θερμοκρασίας ή έντονης ηλιοφάνειας.

Η μέτρηση του ηλεκτρικού πεδίου κατά μήκος ενός μονωτήρα εφαρμόζεται με μεγαλύτερη επιτυχία σε κεραμικούς παρά σε μη κεραμικούς μονωτήρες, εξαιτίας ενυπαρχόντων διαφορών στη δομή αυτών των μονωτήρων. Την ίδια στιγμή φαίνεται ότι αυτή η μέθοδος μπορεί να καταλήξει σε λανθασμένα αποτελέσματα υπό την παρουσία υγρασίας.

Για τους παραπάνω λόγους η οπτική παρατήρηση από έμπειρο προσωπικό είτε απ' το έδαφος, είτε από ελικόπτερα παραμένει ως τώρα ίσως η μόνη πιο αποτελεσματική μέθοδος για την επιθεώρηση των μη κεραμικών μονωτήρων.

2.3.1 Αρχή λειτουργίας

Η παρατήρηση των μερικών εκκενώσεων με γυμνό μάτι είναι δύσκολη διότι εκπέμπεται ασθενής ακτινοβολία κυρίως στο υπεριώδες (UV) φάσμα, η οποία δεν είναι ορατή απ' το ανθρώπινο μάτι. Η ηλιακή ακτινοβολία σ' αυτήν την περιοχή είναι πολύ πιο έντονη και συνεπώς καλύπτει τα σήματα των εκκενώσεων.

Στην περιοχή από 240-280nm (solar-blind region) η εκπομπή των μερικών εκκενώσεων είναι πολύ ασθενής. Ωστόσο η ηλιακή ακτινοβολία είναι πρακτικά αμελητέα. Είναι λοιπόν αυτά τα σήματα που ενεργοποιούνται από τις μερικές εκκενώσεις στη solar-blind περιοχή που χρησιμοποιούνται από ειδικές κάμερες για την ανίχνευση του φαινομένου.

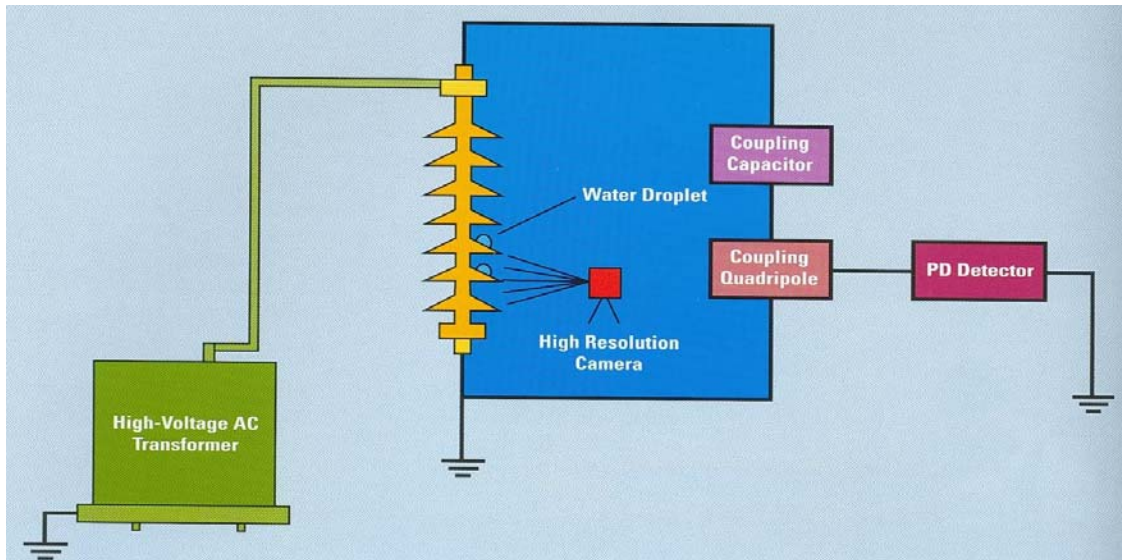
2.3.2 Ανάλυση και επεξεργασία εικόνων

Η εικόνα που “συλλαμβάνεται” από την κάμερα καταγραφής των μερικών εκκενώσεων αποθηκεύεται γενικά σε έναν υπολογιστή σαν ένας πίνακας με pixels σε ένα JPG αρχείο. Η εικόνα αυτή στη συνέχεια χρειάζεται να υποβληθεί σε μια διαδικασία επεξεργασίας και αφαίρεσης θορύβου για να διευκολυνθούν οι ακόλουθες μαθηματικές λειτουργίες και η ποσοτική ανάλυση. Βασικά, το αρχείο JPG αναπαριστά τη φωτεινότητα στις διάφορες θέσεις των στοιχείων του πίνακα της εικόνας. Με τη χρήση του προγράμματος MATLAB οι εντάσεις των pixels (που αναπαριστώνται από τα στοιχεία στον πίνακα της εικόνας) προστίθενται για να προκύψει η αθροιστική ένταση. Επίσης καθορίζεται το μέγιστο επίπεδο φωτεινότητας των pixels (I_m) για το φόντο της εικόνας.

Στην τελική εικόνα του φαινομένου των μερικών εκκενώσεων τα pixels που έχουν τιμή μεγαλύτερη απ' την I_m θεωρούνται φωτεινά (τιμή '1', ή υψηλή, αλλιώς παίρνουν τιμή '0', ή χαμηλή). Ο αριθμός των φωτεινών pixels είναι ένα μέτρο για τη φωτεινότητα της εικόνας. Η αθροιστική ένταση και η φωτεινότητα των pixels αναφέρονται ως μετασηματισμένες παράμετροι εικόνας.

2.3.3 Πειραματική διάταξη

Η πειραματική διάταξη φαίνεται στο Σχήμα 2.2. Χρησιμοποιήθηκε ένας μονωτήρας από σιλικονούχο καουτσούκ ονομαστικής τάσης 115kV. Ο μονωτήρας αυτός πριν το πείραμα ήταν σε λειτουργία και επελέγη για το πείραμα γιατί παρουσίαζε αυξημένη αντοχή. Είχε βέβαια κάποια σημάδια καψίματος από κεραυνοπληξία.



Σχήμα 2. 2: Πειραματική διάταξη.

Η εφαρμοζόμενη τάση αυξανόταν σταδιακά σε βήματα των 5kV, έως ότου φτάσει την ονομαστική. Οι εικόνες των μερικών εκκενώσεων και τα αντίστοιχα φορτία που εκλύονται (PD magnitudes) καταγράφονταν ταυτόχρονα. Η απόσταση μεταξύ του σημείου ενδιαφέροντος και της κάμερας ήταν 5m.

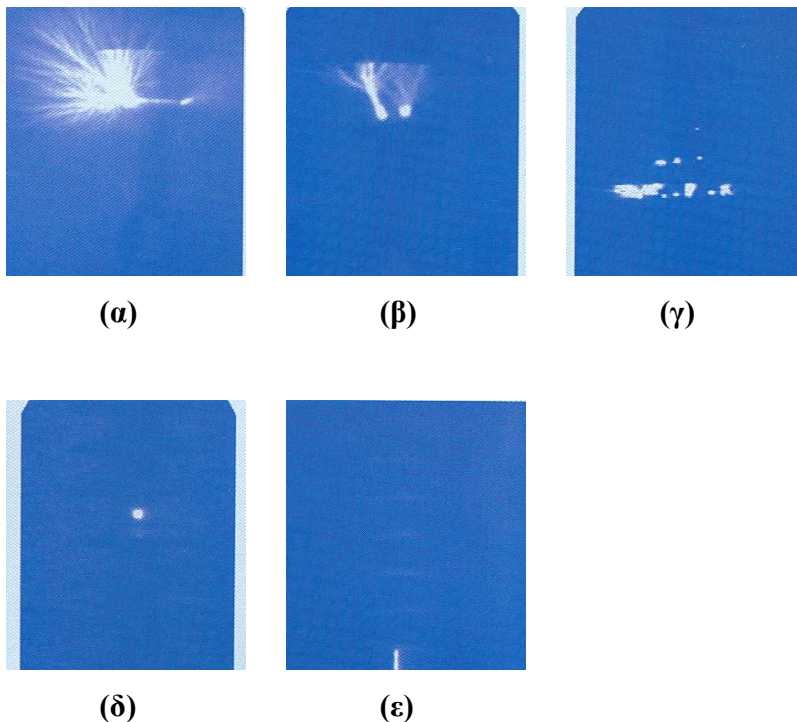
Οποιαδήποτε αύξηση της τάσης είχε σαν αποτέλεσμα μεγαλύτερη ένταση εκκένωσης (όπως αυτή μετρήθηκε από τον ανιχνευτή φορτίου - PD detector).

2.3.4 Αποτελέσματα

Με το παραπάνω πείραμα έγινε δυνατό να υπάρξει συσχέτιση ανάμεσα στις μετασηματισμένες παραμέτρους εικόνας και το μέγεθος της μερικής εκκένωσης. Το φορτίο της εκκένωσης φαίνεται να είναι ανάλογο με την ένταση της φωτεινότητας των pixels της εικόνας.

Υπάρχουν πολλά σημεία σε ένα μη κεραμικό μονωτήρα που μπορούν να λειτουργήσουν σαν εστίες έναρξης μερικών εκκενώσεων. Τέτοια είναι:

- Ο πυρήνας του μονωτήρα, ακόμα και σε ξηρές συνθήκες, όταν δεν υπάρχει τοροειδής (grading ring) (Σχήμα 2.3α).
- Εξαρτήματα, αλλά και το υδρόφοβο περίβλημα, κυρίως κατά την παρουσία σταγόνων νερού (Σχήματα 2.3β και 2.3γ).
- Επιφάνειες που βρίσκονται ανάμεσα σε δίσκους, κυρίως όταν υπάρχουν και σταγονίδια νερού (Σχήμα 2.3δ).
- Το τμήμα στήριξης μεταξύ του ακροδέκτη YT και του πρώτου πιάτου, κυρίως υπό συνθήκες υγρασίας (Σχήμα 2.3ε).



Σχήμα 2. 3: Παραδείγματα μερικών εκκενώσεων.

Η γνώση της θέσης που γίνεται η εκκένωση είναι πολύ σημαντική για την αποτίμηση του κινδύνου που διατρέχει ένας μονωτήρας λόγω των μερικών εκκενώσεων.

Οι εκκενώσεις που συμβαίνουν στον πυρήνα και/ή μακριά από το υλικό του περιβλήματος είναι λιγότερο καταστροφικές από εκείνες που συμβαίνουν πάνω ή κοντά στο περίβλημα. Ο λόγος της αθροιστικής έντασης ανά περιοχή φωτεινών pixels είναι ένα μέτρο για την ένταση της εκκένωσης και θα μπορούσε συνεπώς να είναι μια καλή παράμετρος για την πρόβλεψη της φθοράς στο μονωτήρα.

Συνοψίζοντας, οι εικόνες που παρέχει μια κάμερα καταγραφής των μερικών εκκενώσεων μπορούν να μετατραπούν σε μετρήσιμες παραμέτρους και υπάρχει καλή συσχέτιση μεταξύ της εικόνας και της έντασης της μερικής εκκένωσης. Τα αποτελέσματα της έρευνας εφαρμόστηκαν στην επιθεώρηση μονωτήρων που βρίσκονταν σε λειτουργία, επιβεβαιώνοντας έτσι ότι είναι δυνατό να χρησιμοποιηθούν τέτοιες κάμερες για την πρόβλεψη και αντιμετώπιση προβλημάτων [6].

2.4 Τεχνική με λέιζερ για τον έλεγχο σύνθετων μονωτήρων

Μια ακόμη απειλή για τους σύνθετους μονωτήρες είναι οι διάφοροι μικροοργανισμοί και μύκητες που μπορούν να αποικήσουν τα οργανικά τμήματά τους, αλλοιώνοντας έτσι την επιφανειακή τους δομή. Συνέπεια αυτού είναι να τίθενται ερωτήματα σχετικά με το αν μια τέτοια αλλοίωση μπορεί να επηρεάσει τη μονωτική λειτουργία των σύνθετων μονωτήρων και με ποιον ακριβώς τρόπο. Αν και η πιθανότητα αποίκησης των μονωτήρων (τόσο των κεραμικών, όσο και των μη κεραμικών) από μικροοργανισμούς σχετίζεται κυρίως με την εγκατάσταση των μονωτήρων αυτών σε τροπικές περιοχές, πρόσφατες έρευνες δείχνουν ότι αυτός ο τύπος “μόλυνσης” του μονωτήρα μπορεί να εμφανιστεί και σε άλλου τύπου κλιματικές συνθήκες. Δυστυχώς, παρά το γεγονός ότι το πρόβλημα είναι ευρέως διαδεδομένο, λείπει η γνώση για το πώς μπορεί να αποτιμηθεί και να αντιμετωπιστεί. Πάνω σε αυτόν τον τομέα ξεκίνησε η έρευνα στη Σουηδία, με σκοπό να προσφέρει καλύτερη κατανόηση για την επίδραση, αν υπάρχει, της “βιολογικής μόλυνσης” στη λειτουργία του μονωτήρα. Μελετήθηκαν επίσης διαγνωστικές μέθοδοι για την ανίχνευση μιας τέτοιου τύπου μόλυνσης, καθώς και τρόποι ελαχιστοποίησης του ρυθμού ανάπτυξης τέτοιων καλλιεργειών σε πολυμερείς επιφάνειες. Ο Andreas Dernfalk [7] σε εργασία του αναφέρεται στην επίδραση της ανάπτυξης καλλιεργειών στους

μονωτήρες, τόσο σε εργαστηριακό επίπεδο, όσο και σε συνθήκες λειτουργίας, αλλά και στη χρήση της μεθόδου LIF (laser-induced fluorescence, φθορισμός με διέγερση από λέιζερ) για τον εντοπισμό βιολογικής μόλυνσης σε μονωτήρες εν λειτουργία.

2.4.1 Γενικά

Αρχικά ο Derrfalk μελέτησε τις διάφορες μεθόδους για την αποίκηση των επιφανειών μονωτήρων με φύκη και μύκητες κάτω από εργαστηριακές συνθήκες. Την ίδια στιγμή ταυτοποιήθηκε ένας αριθμός ήδη μολυσμένων μονωτήρων που βρίσκονταν σε λειτουργία στη Σουηδία. Και στις δύο ομάδες μονωτήρων έγιναν μετρήσεις για το ρεύμα διαρροής υπό καθαρή ομίχλη και για την τάση υπερπήδησης υπό βροχή. Στη συνέχεια οι επιφάνειες των μονωτήρων μελετήθηκαν από απόσταση 60 μέτρων με τη μέθοδο LIF για την ανίχνευση βιολογικής μόλυνσης.

2.4.2 Μονωτήρες που εξετάστηκαν

Έξι μη κεραμικοί μονωτήρες, με αρίθμηση από #1 έως #6 χρησιμοποιήθηκαν κατά τη διάρκεια πειραμάτων που διεξήχθησαν στο εργαστήριο για την ανάπτυξη και καλλιέργεια μυκήτων. Αυτοί οι μονωτήρες ήταν τύπου διανομής (distribution type). Η καλλιέργεια ξεκίνησε σε τρεις απ' τους μονωτήρες (#4, #5 και #6), ψεκάζοντας πάνω σ' αυτούς ένα μίγμα από σπόρια μυκήτων που είχαν παρθεί από μονωτήρες που βρίσκονταν σε λειτουργία και μια κατάλληλη θρεπτική ουσία. Στη συνέχεια οι μονωτήρες τοποθετήθηκαν σε έναν ειδικά σχεδιασμένο θάλαμο (Σχήμα 2.4), στον οποίο η σχετική υγρασία διατηρούνταν σταθερή και περίπου 90% και η θερμοκρασία ήταν μεταξύ 20 και 22°C.



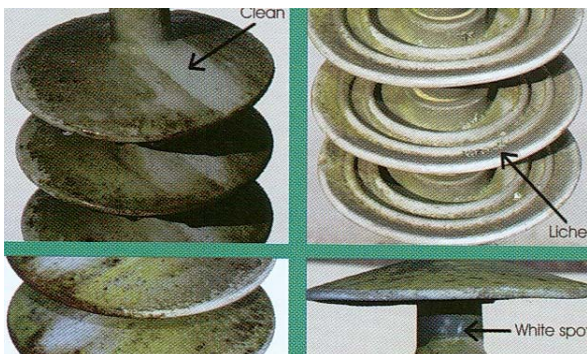
Σχήμα 2. 4: Ειδικός θάλαμος προσομοίωσης κλιματικών συνθηκών.

Μετά από τέσσερις εβδομάδες στο θάλαμο ένα καφετί στρώμα μπορούσε ήδη να παρατηρηθεί στους δίσκους του μονωτήρα. Δέκα μήνες αργότερα οι μύκητες είχαν εξαπλωθεί και οι άνω πλευρές των πιάτων είχαν χάσει την υδροφοβικότητά τους (Σχήμα 2.5)



Σχήμα 2. 5: Μονωτήρας που αποικήθηκε από μύκητες στο εργαστήριο.

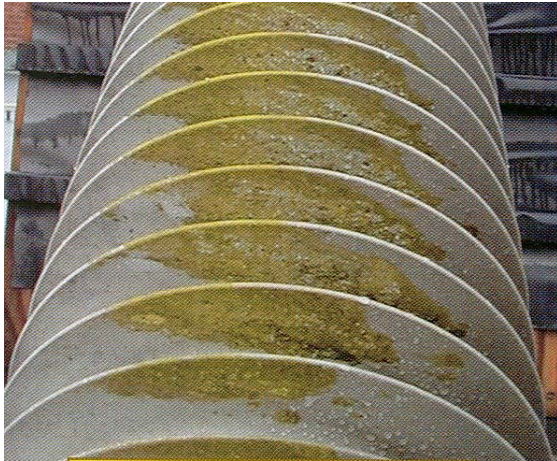
Άλλοι τρεις μονωτήρες ανάρτησης από σιλικονούχο καουτσούκ (με κωδικούς B1, B2 και B3) είχαν μολυνθεί από χλωροφύκη (μούχλα) και λειχήνες κάτω από φυσικές συνθήκες κι όχι στο εργαστήριο. Αυτοί οι μονωτήρες είχαν ένα ελαφρώς μεγαλύτερο μήκος ερπυσμού απ' ότι οι μονωτήρες της πρώτης ομάδας και ήταν εγκατεστημένοι για 12 χρόνια σε σύστημα των 11kV (Σχήμα 2.6).



Σχήμα 2. 6: Μονωτήρες ανάρτησης μετά από 12 χρόνια λειτουργίας.

Οι μονωτήρες αυτοί είχαν στηθεί οριζόντια σε γραμμή που περνούσε από πυκνό δάσος και συνεπώς τα χαρακτηριστικά της τοποθεσίας αυτής (υψηλή υγρασία, χαμηλή ηλιοφάνεια και παρουσία οργανικών ουσιών κατάλληλων για τροφή) ήταν ιδιαίτερα ευνοϊκά για την ανάπτυξη βιολογικής μόλυνσης. Τα πιάτα των μονωτήρων αυτών είχαν αποικηθεί σε μεγάλο βαθμό. Οι επιφάνειες των προβληματικών περιοχών είχε γίνει τραχιά, ενώ η καλλιέργεια είχε την τάση να σπάει σε μικρότερες “αποικίες”. Οι λειχήνες εντοπίζονταν σε λίγες και μικρές περιοχές των μονωτήρων, κυρίως σε επιφάνειες που δεν τις έβλεπε καθόλου ο ήλιος. Μερικά λευκά στίγματα παρατηρήθηκαν επίσης στο πολυμερές περίβλημα του πυρήνα σε όλους αυτούς τους μονωτήρες. Αυτό δεν ήταν αποτέλεσμα κάποια μόλυνσης του συγκεκριμένου σημείου, αλλά αποχρωματισμός του ίδιου του υλικού του περιβλήματος. Εφόσον τα στίγματα αυτά παρατηρήθηκαν σε σημεία που βρίσκονταν ανάμεσα στις προσβεβλημένες περιοχές θεωρήθηκε ότι ο αποχρωματισμός προκλήθηκε από τοπικές εκκενώσεις ή από θέρμανση εξαιτίας της ύπαρξης ρεύματος διαρροής. Οι υδροφοβικές ιδιότητες των μονωτήρων εκτιμήθηκαν στη συνέχεια με τη βοήθεια του οδηγού STRI για την κατάταξη με βάση την υδροφοβικότητα (HC, hydrophobicity classification). Βρέθηκε λοιπόν ότι οι προσβεβλημένες περιοχές είχαν δείκτη HC από 6 έως 7, δηλαδή ήταν σχεδόν πλήρως διαβρέξιμες (wetable). Σημεία που με γυμνό μάτι φαίνονταν να είναι καθαρά ήταν περισσότερο υδρόφοβα και γενικά είχαν δείκτες HC 3-4, ενώ μερικές μικρότερες περιοχές είχαν και HC 1.

Το τελευταίο σύνολο μονωτήρων που συμπεριλήφθηκε στη μελέτη αποτελούνταν από τέσσερα μεγάλα μονωτικά περιβλήματα σιλικόνης χωρίς πυρήνα (hollow core insulators) με κωδικούς H1-H4 που προηγούμενα ήταν εγκατεστημένοι στο σταθμό Anneberg κοντά στο Gothenburg και λειτουργούσαν υπό εναλλασσόμενη φασική τάση 130kV rms. Οι μονωτήρες αυτοί είχαν εγκατασταθεί από το 1995 και είχαν καλυφθεί από μούχλα που μπορούσε να παρατηρηθεί σε όλο το μήκος του μονωτήρα, αλλά μόνο στις περιοχές που δεν τις έβλεπε ο ήλιος, κι αυτό γιατί σ’ αυτές τις περιοχές η υγρασία παρέμενε για μεγαλύτερο διάστημα. Οι μονωτήρες αυτοί είχαν στηθεί σε ύψος περίπου ενός μέτρου πάνω απ’ το έδαφος. Ένα τέτοιο μονωτικό περίβλημα φαίνεται στο Σχήμα 2.7)



Σχήμα 2. 7: Μονωτικά περιβλήματα στα οποία έχει αναπτυχθεί μούχλα.

Οι διαστάσεις των μονωτήρων που χρησιμοποιήθηκαν στις δοκιμές δίνονται στον Πίνακα 2.1.

Πίνακας 2. 1: Διαστάσεις των μονωτήρων που χρησιμοποιήθηκαν στις δοκιμές

Μονωτήρας	Αριθμός πιάτων	Διάμετρος καλύμματος πυρήνα (sheath) / δίσκου (shed)	Μήκος ερπυσμού (mm)	Μήκος τόξου (mm)
#1 - #6	3	160/22	620	285
B1 – B3	6	112/26	684	210
H1 – H2	26	400/265	3600	1410
H3 – H4	30	285/216	3390	1430

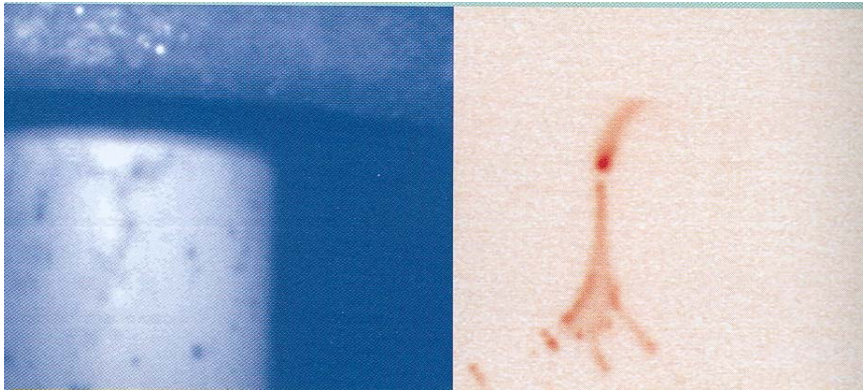
2.4.3 Ηλεκτρικές δοκιμές

Δοκιμές για το ρεύμα διαρροής έγιναν περιοδικά στους μονωτήρες #1-#6 της πρώτης ομάδας καθ' όλη τη διάρκεια της τεχνητής επιμόλυνσης που κράτησε περίπου 18 μήνες, ενώ στο τέλος της περιόδου αυτής οι μονωτήρες εκτέθηκαν και σε δοκιμές για την κρίσιμη τάση. Ανάλογη διαδικασία εφαρμόστηκε και στους μονωτήρες της δεύτερης ομάδας (B1-B3). Οι μελέτη για τους μονωτήρες της τρίτης κατηγορίας (H1-H4) περιορίστηκε σε δοκιμές για την κρίσιμη τάση.

Γενικά τα ρεύματα που καταγράφηκαν στις δοκιμές ήταν χαμηλά (της τάξης των δεκάδων μA), αλλά μεγαλύτερα κατά 3 ή 4 φορές απ' ότι σε καθαρούς μονωτήρες.

Ωστόσο περίπου έξι μήνες μετά την έναρξη των δοκιμών ξεκίνησαν εκκενώσεις στην επιφάνεια των μονωτήρων.

Συγκεκριμένα, οι περιοχές που εμφάνιζαν το πρόβλημα (ήταν μολυσμένες) λειτουργούσαν σαν πόλοι μεταξύ των οποίων εμφανίζονταν φωτεινές εκκενώσεις (Σχήμα 2.8).



Σχήμα 2.8: Κοντινή άποψη μικρής περιοχής του καλύμματος του πυρήνα (sheath) και του ακροδέκτη (end fitting) αριστερά και φωτεινές εκκενώσεις (δεξιά)

Για τις δοκιμές υπερπήδησης τηρήθηκε η διαδικασία που ορίζει το IEC 60060-1 με τη μεγαλύτερη δυνατή ακρίβεια. Για τους μονωτήρες τύπου διανομής οι δοκιμές πραγματοποιήθηκαν κατά IEC 60383-1, ενώ για τα μονωτικά περιβλήματα οι δοκιμές έγιναν σύμφωνα με το IEC 61462.

Όταν οι δοκιμές υπερπήδησης στα μονωτικά περιβλήματα (3^ο σύνολο μονωτήρων) ολοκληρώθηκαν, αυτά καθαρίστηκαν και αφαιρέθηκε από πάνω τους η βιολογική μόλυνση. Οι μονωτήρες αφέθηκαν στη συνέχεια για ένα μήνα στο εργαστήριο και μετά απ' αυτό το διάστημα οι δοκιμές επαναλήφθηκαν για τη δημιουργία δεδομένων αναφοράς.

Όσον αφορά το πρώτο σύνολο μονωτήρων, σύγκριση των μέσων τιμών και των τυπικών αποκλίσεων έδειξε ότι υπήρξε σημαντική μείωση στις τιμές των τάσεων υπερπήδησης για τους μονωτήρες που επιμολύνθηκαν τεχνητά στο εργαστήριο (#4-6) σε σχέση με όμοιους μονωτήρες που κρατήθηκαν καθαροί (#1-3). Θεωρώντας τις μέσες τιμές η μείωση είναι περίπου 33%. Οι μετρήσεις για το δεύτερο σύνολο μονωτήρων (B1-B3) που μολύνθηκαν κατά τη λειτουργία τους αποκάλυψαν ότι οι τιμές για την τάση υπερπήδησής τους ήταν ισοδύναμες με εκείνες των μονωτήρων που μολύνθηκαν τεχνητά

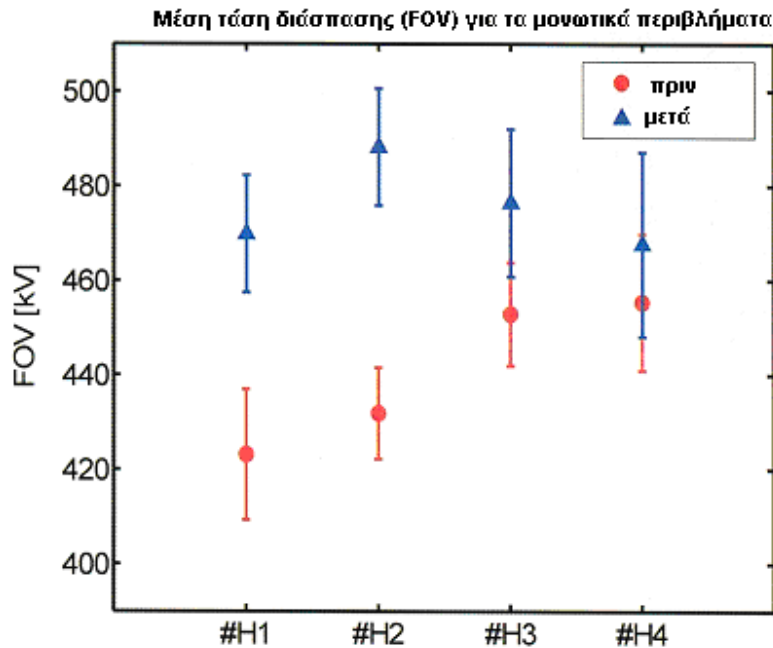
στο εργαστήριο. Δεν υπήρχαν καθαροί μονωτήρες όμοιοι με αυτούς της δεύτερης κατηγορίας και γι' αυτό ως αναφορά για τη σύγκριση χρησιμοποιήθηκαν οι καθαροί μονωτήρες #1-3 της πρώτης κατηγορίας που είχαν παρόμοια γεωμετρικά χαρακτηριστικά.

Παρότι οι μονωτήρες B1-B3 ήταν καλυμμένοι από μύκητες σε ποσοστό 60-80% της επιφάνειάς τους, ενώ οι μονωτήρες #4-6 μόλις σε ποσοστό 15-20% τα αποτελέσματα ήταν παρόμοια. Μια πιθανή εξήγηση γι' αυτήν την παρόμοια συμπεριφορά είναι ότι και στις δύο περιπτώσεις τα υλικά είχαν γίνει πλέον υδρόφιλα. Οι τιμές για τις τάσεις υπερπήδησης των παραπάνω μονωτήρων, όπως προέκυψαν απ' τις δοκιμές παρουσιάζονται στον Πίνακα 2.2

Πίνακας 2. 2: Μέσες τιμές για την τάση διάσπασης των μονωτήρων ανάρτησης

Μονωτήρας	Καθαροί μονωτήρες			Τεχνητά μολυσμένοι			Μολυσμένοι κατά τη λειτουργία τους		
	#1	#2	#3	#4	#5	#6	B1	B2	B3
Μέση Τάση Υπερπήδησης (kV)	114,6	119,2	118,4	78,6	79	79,2	77,8	80,6	77,4
Τυπική απόκλιση (kV)	6,3	7,2	10,4	1,7	3,5	1,8	6,1	4,6	4,2

Η επίδραση της βιολογικής μόλυνσης στις κρίσιμες τάσεις των μονωτικών περιβλημάτων (H1-H4) ήταν μικρότερη (βλ. Σχήμα 2.9). Η μείωση στην τιμή της κρίσιμης τάσης για τους μονωτήρες H1 και H2 πριν και μετά το πλύσιμο ήταν της τάξης του 11%, ενώ για τους μονωτήρες H3 και H4 ήταν μόλις 4,5%. Η αβεβαιότητα για την εκτίμηση αυτή είναι βέβαια αυξημένη, εξαιτίας των μεγάλων στατιστικών διακυμάνσεων.



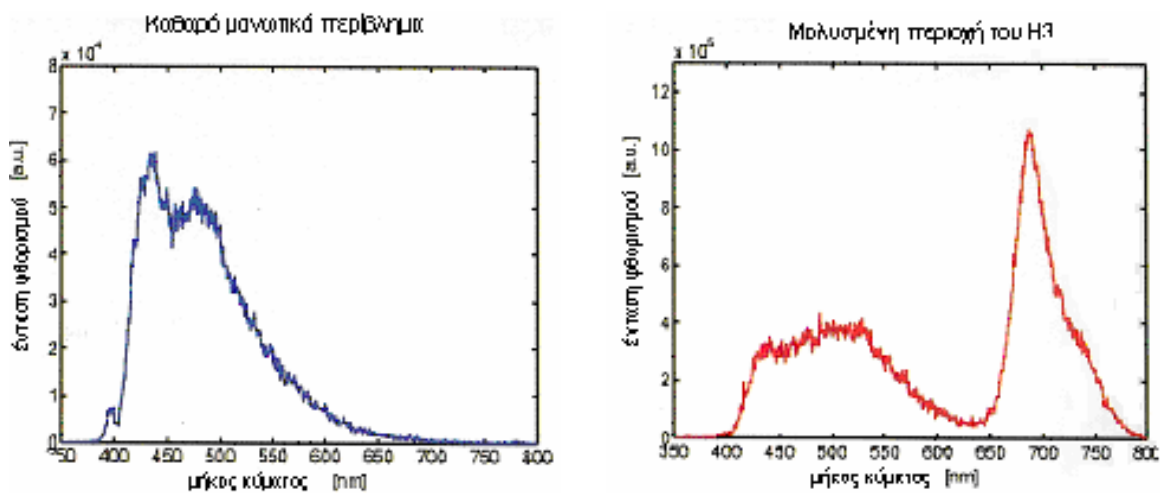
Σχήμα 2. 9: Μέση τάση διάσπασης (Flashover Voltage, FOV) για τα μονωτικά περιβλήματα πριν και μετά τον καθαρισμό (95% διαστήματα εμπιστοσύνης)

2.4.4 Φασματοσκοπία φθορισμού

Το φαινόμενο του φθορισμού προέρχεται από την αυθόρμητη εκπομπή φωτός όταν ένα άτομο ή μόριο χαλαρώνει και μεταπίπτει από ένα επίπεδο υψηλής ενέργειας σε ένα χαμηλότερης ενέργειας επίπεδο, ή στη βασική κατάσταση. Αν το μόριο έχει διεγερθεί με τη χρήση ακτίνων ή παλμών λέιζερ, τότε λέμε ότι ο προκύπτων φθορισμός προκλήθηκε από λέιζερ (laser-induced fluorescence, LIF). Στη φασματοσκοπία LIF το εκπεμπόμενο φως θα έχει εν μέρει το ίδιο μήκος κύματος με τη διεγείρουσα ακτινοβολία, αλλά θα έχει και φωτόνια με μεγαλύτερο μήκος κύματος. Με αυτόν τον τρόπο, το φάσμα του εκπεμπόμενου φωτός εξαρτάται σημαντικά από τα χαρακτηριστικά της επιφάνειας του υλικού το οποίο διεγείρουμε και είναι στην ουσία ένα αποτύπωμα της κατάστασης αυτής. Σ' αυτήν την έρευνα [7] χρησιμοποιήθηκε για τη διεγέρση ένα σύστημα που σχεδιάστηκε στο Lund Institute for Environmental Monitoring. Το λέιζερ που χρησιμοποιήθηκε ήταν Nd:YAG. Με το σύστημα αυτό απελευθερώνονται παλμοί φωτός μικρής διάρκειας (4-5ns) και μήκους κύματος 355nm με ρυθμό επανάληψης 20Hz. Μπορεί να στηθεί και σε φορητό για εξ αποστάσεως μετρήσεις. Μέσω ενός κάθετα τοποθετημένου τηλεσκοπίου η ακτίνα του λέιζερ επεκτείνεται σε μια διάμετρο 3cm προτού κατευθυνθεί προς τον

απομακρυσμένο στόχο με τη βοήθεια ενός πτυσσόμενου κατόπτρου κατευθυνόμενου από υπολογιστή. Ένα μέρος από το φως που εκπέμπει ο φθορισμός επιστρέφει στο ίδιο κάτοπτρο, συλλέγεται από ένα άλλο τηλεσκόπιο και οδηγείται σε έναν αναλυτή πολλών καναλιών (multi-channel analyser) μέσω μιας οπτικής ίνας. Στη συνέχεια το φάσμα αποθηκεύεται σε έναν υπολογιστή. Στο συγκεκριμένο πείραμα η απόσταση μεταξύ του φορητού (που μετέφερε το σύστημα διέγερσης) και των μονωτήρων ήταν περίπου 60 μέτρα.

Ένα τυπικό φάσμα που συλλέχθηκε από τμήμα μονωτικού περιβλήματος καλυμμένου από παχύ στρώμα βιολογικής μόλυνσης (χλωροφύκη) φαίνεται στο Σχήμα 2.10.

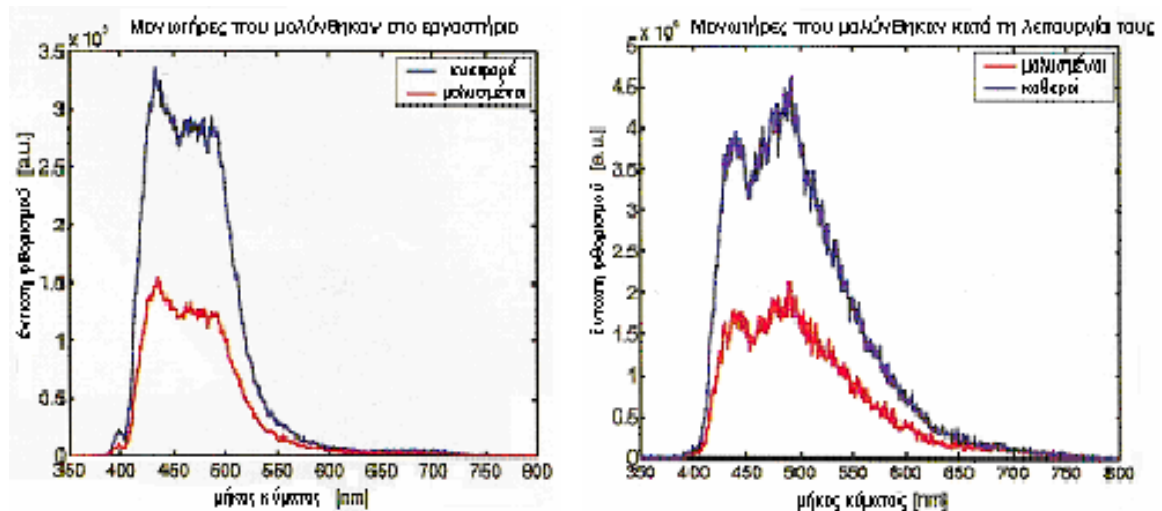


Σχήμα 2.10: Τυπικά φάσματα για καθαρή επιφάνεια (αριστερά) και για περιοχή αποικισμένη με φύκη του μονωτήρα Η3 (δεξιά)

Σ' αυτό υπάρχει μια έντονη κορυφή στα 680-690nm, που αντιστοιχεί στο φθορισμό της χλωροφύλλης. Αντίθετα, για τα τμήματα του μονωτήρα που δεν καλύπτονταν από φύκη, ο φθορισμός ήταν χαμηλός και μειούμενος για μήκη κύματος άνω των 600nm.

Τα φάσματα LIF για τους μονωτήρες που είχαν τεχνητά μολυνθεί στο εργαστήριο με μύκητες ήταν πιο δύσκολο να εξηγηθούν, για πολλούς λόγους. Πρώτον, διότι οι μύκητες σε αντίθεση με τα φύκη δεν φωτοσυνθέτουν και συνεπώς δεν εντοπίζεται χλωροφύλλη. Δεύτερον, γιατί οι μύκητες ήταν λίγο ως πολύ ομοιόμορφα κατανομημένοι στην επιφάνεια του μονωτήρα με τη μορφή μικρών και μεμονωμένων κηλίδων. Καθώς η ανάλυση αυτών των μετρήσεων ήταν της τάξης των 2-3cm, τα φάσματα δεν ήταν ούτε από περιοχές εντελώς καθαρές, ούτε από περιοχές εντελώς καλυμμένες με μύκητες. Εντούτοις, σύγκριση τυπικών φασμάτων από καθαρούς μονωτήρες αναφοράς και

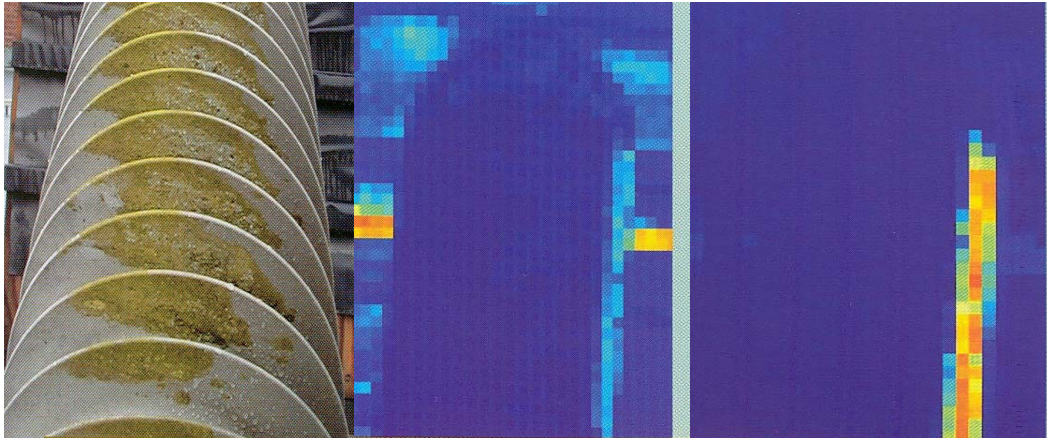
φασμάτων από μονωτήρες με ανάπτυξη μυκήτων έδειξαν ότι η ένταση φθορισμού ήταν χαμηλότερη για τους μολυσμένους μονωτήρες (βλ. Σχήμα 2.11). Επίσης, αν τα πλάτη κανονικοποιούνται, παρατηρούνταν ελαφρώς εντονότερος φθορισμός των μολυσμένων μονωτήρων για μήκη κύματος 520-600nm.



Σχήμα 2.11: Αριστερά: φάσματα από καθαρούς μονωτήρες αναφοράς και από μονωτήρες που μολύνθηκαν στο εργαστήριο. Δεξιά: φάσματα από μολυσμένες και καθαρές περιοχές μονωτήρων που μολύνθηκαν κατά τη λειτουργία τους.

Οι μετρήσεις που έγιναν για τους μονωτήρες που μολύνθηκαν κατά τη λειτουργία τους έδειξαν ότι, όπως και γι' αυτούς που είχαν μολυνθεί στο εργαστήριο, η ένταση φθορισμού στις προσβεβλημένες περιοχές ήταν πολύ μικρότερη απ' ότι στις καθαρές. Επίσης, οι μολυσμένες περιοχές εμφάνιζαν υψηλό φθορισμό στα 685nm περίπου.

Ένα από τα πλεονεκτήματα αυτού του συστήματος είναι ότι μπορεί να προγραμματιστεί για να κάνει μετρήσεις αυτόματα. Στην περίπτωση αυτή το σύστημα ρυθμίζεται να σαρώνει ένα μεγάλο μέρος του υπό εξέταση μονωτήρα, δίνοντας έτσι ένα μεγάλο σύνολο φασμάτων. Στη συνέχεια τα δεδομένα μπορούν να παρουσιαστούν σαν χρωματισμένες εικόνες, η κάθε μία απ' τις οποίες θα δείχνει την ένταση του φθορισμού εντός συγκεκριμένου εύρους μηκών κύματος. Το αποτέλεσμα μιας τέτοιας σάρωσης που έγινε σε μονωτικό περίβλημα με χλωροφύκη παρουσιάζεται στο Σχήμα 2.12.



Σχήμα 2.12: Φωτογραφία μονωτικού καλύμματος με φύκη (αριστερά), μέση ένταση φθορισμού για 400-800nm (κέντρο) και για 670-700nm (δεξιά)

Στη μεσαία εικόνα, που απεικονίζει τη μέση συνολική ένταση φθορισμού, μπορεί κανείς να παρατηρήσει ότι η ξύλινη κατασκευή υποστήριξης είχε τον εντονότερο φθορισμό απ' όλα τα υλικά που σαρώθηκαν (κίτρινο χρώμα). Σχετικά καθαρά φαίνονται και οι μολυσμένες άκρες του μονωτήρα (ανοιχτό μπλε). Ωστόσο, αναλύοντας την κορυφή για το φθορισμό της χλωροφύλλης μόνο, δηλαδή από 670-700nm, οι προσβεβλημένες περιοχές φαίνονται ακόμη πιο καθαρά (δεξιά εικόνα).

2.5 Μετασχηματισμός Κύματος (Wavelet Transformation) για την πρόβλεψη αστοχιών στους σύνθετους μονωτήρες

Με την εισαγωγή των σύνθετων (μη κεραμικών) μονωτήρων σε εφαρμογές υψηλών τάσεων, υπήρξε έντονο ενδιαφέρον για την ανάπτυξη μεθόδων παρακολούθησης που θα μπορούσαν να προβλέπουν έγκαιρα επικείμενες αστοχίες. Με αυτόν το σκοπό αναπτύχθηκαν ραγδαία πολλές μέθοδοι για την εκτίμηση των πολυμερών υλικών που χρησιμοποιούνταν στα καλύμματα των σύνθετων μονωτήρων. Η εισαγωγή των ΗΥ στις αρχές της δεκαετίας του '80 επέτρεψε την καταγραφή του ρεύματος διαρροής σε πραγματικό χρόνο, τουλάχιστον στο εργαστήριο. Δυστυχώς δεν έγινε εφικτό να χρησιμοποιηθεί η παράμετρος αυτή με αξιοπιστία για την πρόβλεψη αστοχιών ή ακόμα και για την κατάταξη της ηλεκτρικής απόδοσης διαφορετικών εναλλακτικών υλικών. Με το πέρασμα των χρόνων, αφιερώθηκε πολλή έρευνα για να μπορέσει αυτό να ξεπεραστεί – αλλά με μικρή ουσιαστικά μεταβολή του αποτελέσματος. Αυτό μπορεί να οφείλεται στο γεγονός ότι η αστοχία των υλικών είναι στην πραγματικότητα ένα ιδιαίτερα

περίπλοκο φαινόμενο που εξαρτάται από πολλούς ξεχωριστούς παράγοντες. Οι Rave Gogur και Sreeram Venkataraman [8] περιγράφουν την τεχνική της επεξεργασίας σημάτων με το Μετασχηματισμό Wavelet ως πρακτική και αξιόπιστη μέθοδο για την παρακολούθηση των σύνθετων μονωτήρων.

Η φθορά των πολυμερών υλικών που χρησιμοποιούνται ως περιβλήματα στους σύνθετους μονωτήρες κάτω από συνθήκες ρύπανσης είναι γενικά αποτέλεσμα των ξηρών ζωνών που προκαλούνται από το ρεύμα διαρροής. Το μέγεθος και η μορφή του ρεύματος διαρροής δεν είναι σταθερά, αλλά καθορίζονται από το πόσο βρεγμένο είναι το υλικό και από το ρυθμό εξάτμισης λόγω της θέρμανσης που προκαλεί το ρεύμα.

Ο Μετασχηματισμός Fourier (Fourier Transformation, FT) είναι μια πολύ γνωστή μέθοδος για την ανάλυση μη ημιτονοειδών σημάτων που επιτρέπει, για ένα δεδομένο χρονικό 'παράθυρο' να αποκτηθούν οι συνιστώσες του σήματος στο πεδίο της συχνότητας. Ωστόσο, ο Μετασχηματισμός Fourier δεν υποδεικνύει το χρόνο εμφάνισης των συνιστωσών συχνότητας σ' αυτό το 'παράθυρο'.

Μια βελτιωμένη παραλλαγή του συμβατικού Μετασχηματισμού Fourier είναι ο Σύντομος Μετασχηματισμός Fourier (Short Term FT, STFT), στον οποίο το χρονικό 'παράθυρο' συμπιέζεται. Για το σκοπό αυτό επιλέγεται μια συνάρτηση 'παραθύρου' (window function) w .

Δυστυχώς, το πρόβλημα με τον STFT είναι ότι, αν το 'παράθυρο' είναι πολύ στενό, επιτυγχάνεται καλή ανάλυση στο πεδίο του χρόνου, αλλά ανεπαρκής ανάλυση στο πεδίο της συχνότητας. Αντίθετα, με ένα μεγάλο 'παράθυρο', επιτυγχάνεται καλή ανάλυση συχνότητας, αλλά όχι χρόνου. Μια εναλλακτική προσέγγιση για να επιλυθεί το πρόβλημα της ανάλυσης που βασίζεται στο μέγεθος του 'παραθύρου', είναι ο Μετασχηματισμός Κύματος (Wavelet Transformation, WT).

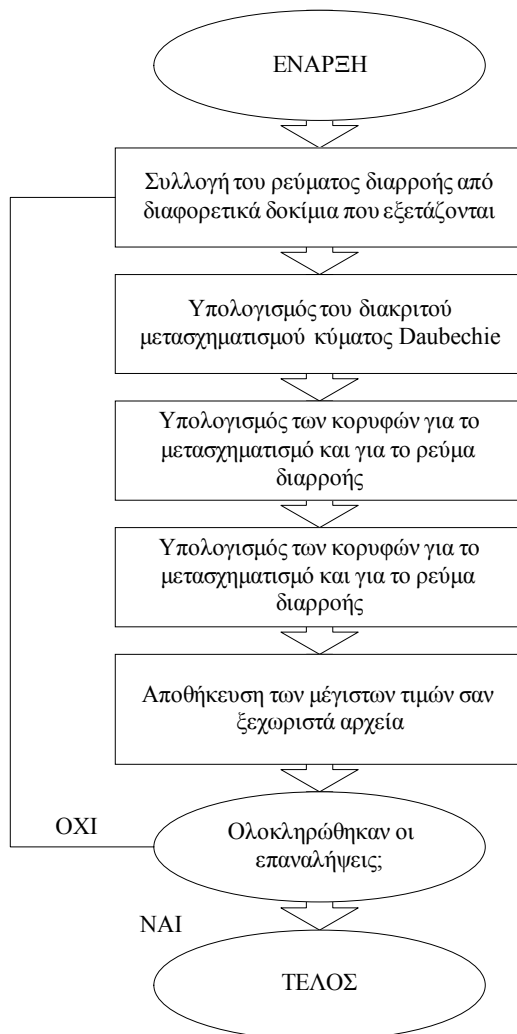
Ο όρος *κύματος* αναφέρεται σε ένα μικρό, τοπικά περιορισμένο κύμα που ενεργεί για ένα μικρό χρονικό διάστημα. Σε ένα Μετασχηματισμό Κύματος, τόσο η πληροφορία για το χρόνο, όσο και για τη συχνότητα του σήματος αποκτώνται ταυτόχρονα.

Τα δεδομένα αυτά μπορούν στη συνέχεια να σχεδιαστούν σε ένα τρισδιάστατο γράφημα, οι άξονες του οποίου θα είναι ο χρόνος, η συχνότητα και το πλάτος. Το 'παράθυρο' έχει την ιδιότητα να προσαρμόζεται με βάση το σήμα εισόδου. Ένα μητρικό κύμα (μια

μαθηματική συνάρτηση) επιλέγεται και με βάση αυτήν ολόκληρο το σήμα περιγράφεται με μια πεπερασμένη σειρά συντελεστών.

Οι μετασχηματισμοί κύματος έχουν χρησιμοποιηθεί σε πολυάριθμες εφαρμογές όπως η προστασία συστημάτων ισχύος, η εύρεση των σημείων σφαλμάτων και η μελέτη των μερικών εκκενώσεων.

Συγκεκριμένα στη μελέτη αυτή [8] εκτελέστηκαν δοκιμές σε δύο τύπους υλικών, κατά τις οποίες γινόταν παρακολούθηση και καταγραφή του ρεύματος διαρροής από σύστημα συλλογής δεδομένων και ταυτόχρονη εύρεση του Μετασχηματισμού Wavelet του ρεύματος. Στο παρακάτω σχήμα δίνεται το μπλοκ διάγραμμα του προγράμματος που χρησιμοποιήθηκε.



Σχήμα 2.13: Μπλοκ Διάγραμμα του προγράμματος

Βρέθηκε ότι με το Μετασχηματισμό Wavelet οι αστοχίες στα υλικά εντοπίζονταν με μεγαλύτερη ευαισθησία απ' ό,τι με τις πραγματικές τιμές του ρεύματος διαρροής.

Γενικά, οι συντελεστές Wavelet είναι ιδιαίτερα ευαίσθητοι σε μεταβολές του σήματος. Συνεπώς είναι αρκετά πιθανό ένα σύστημα που ενσωματώνει τη μεταβολή στα πλάτη των συντελεστών του μετασχηματισμού να είναι ιδιαίτερα αποτελεσματικό στην πρόβλεψη επικείμενων αστοχιών σε τέτοιου είδους υλικά.

Ένα τέτοιο σύστημα θα μπορούσε να συμπεριληφθεί σε περιοχές δοκιμών ή ακόμα για έλεγχο μονωτήρων κατά τη λειτουργία τους, όταν δεν υπάρχει άλλη προειδοποίηση, ώστε να λαμβάνονται διορθωτικές δράσεις πριν συμβεί κάποια αστοχία.

2.6 Τεχνητά Νευρωνικά Δίκτυα για τον έλεγχο των μονωτήρων

Ένα μεγάλο πεδίο στις εφαρμογές των Τεχνητών Νευρωνικών Δικτύων (ΤΝΔ) είναι η εκτίμηση συναρτήσεων, διότι οι χρήσιμες ιδιότητες των ΤΝΔ όπως είναι η προσαρμοστικότητα και η μη γραμμικότητα ταιριάζουν σε εργασίες εκτίμησης συναρτήσεων, όπου η εξίσωση που περιγράφει τη συνάρτηση είναι άγνωστη. Πολλές μελέτες έχουν γίνει για τον προσδιορισμό των παραμέτρων των ΤΝΔ που δίνουν τα καλύτερα αποτελέσματα. Έγιναν επίσης μελέτες για να καθοριστεί η επίδραση της παρουσίας ανεπαρκών δεδομένων του συνόλου εκπαίδευσης στην ακρίβεια του μοντέλου. Στον τομέα των συστημάτων ενέργειας, τα ΤΝΔ έχουν χρησιμοποιηθεί για ακριβή πρόβλεψη φορτίου, εκτίμηση ασφάλειας, έλεγχο πυκνωτών κ.ά.

Παρακάτω παρουσιάζονται αποτελέσματα εργασιών με θέμα την εφαρμογή των ΤΝΔ στις υψηλές τάσεις και συγκεκριμένα στους μονωτήρες.

2.6.1 Τεχνητό Νευρωνικό Δίκτυο για την εκτίμηση του επιπέδου ρύπανσης σε μονωτήρες υψηλών τάσεων.

Η έρευνα αυτή [9] κατευθύνεται στη μελέτη της συμπεριφοράς των μονωτήρων κάτω από θαλάσσια ρύπανση. Οι επιδράσεις των διαφόρων μετεωρολογικών παραγόντων στο επίπεδο της ρύπανσης έχουν μελετηθεί εκτενώς. Γίνεται μια προσπάθεια να εκτιμηθεί η σοβαρότητα της ρύπανσης κάτω από διάφορες καιρικές συνθήκες με τη χρήση Τεχνητών Νευρωνικών Δικτύων. Ένα ΤΝΔ μπορεί να μοντελοποιήσει μια συνάρτηση ακόμα κι αν η εξίσωση που την περιγράφει είναι άγνωστη. Το μόνο προαπαιτούμενο είναι ένα

αντιπροσωπευτικό δείγμα για τη συμπεριφορά της συνάρτησης, το οποίο προέρχεται από πειραματικά δεδομένα κι όχι από θεωρητική κατανόηση.

Οι Ahmad S. Ahmad et al. ανέπτυξαν ένα ΤΝΔ για την εκτίμηση συναρτήσεων και το χρησιμοποίησαν για την ακριβή μοντελοποίηση της σχέσης μεταξύ του ESDD (Equivalent Salt Deposit Density) και της θερμοκρασίας (T), της υγρασίας (H), της πίεσης (P), της βροχόπτωσης (R) και της ταχύτητας του ανέμου (WV). Η προσέγγιση που χρησιμοποιήθηκε για τη μεταφορά του επιπέδου ρύπανσης στην περιοχή των δοκιμών καλείται ESDD (Ισοδύναμη πυκνότητα εναπόθεσης άλατος) και μετριέται σε mg/cm^2 NaCl. Το ESDD είναι η ισοδύναμη ποσότητα NaCl που θα επέφερε την ίδια αγωγιμότητα σε πλήρη διάλυση. Αν η τιμή του ESDD είναι ίση ή μεγαλύτερη από $0,03 \text{ mg}/\text{cm}^2$ οι μονωτήρες θα πρέπει να πλυθούν. Στόχος της προσπάθειας αυτής ήταν η γνώση για τη συμπεριφορά και το επίπεδο της ρύπανσης, η οποία θα μπορούσε να βοηθήσει στην καθιέρωση μιας πολιτικής συντήρησης των μονωτήρων, αφού θα ήταν γνωστοί τόσο οι κρίσιμοι μήνες όσο και οι περίοδοι έκθεσης κατά τη διάρκεια του χρόνου.

2.6.1.1 Μεθοδολογία που χρησιμοποιήθηκε

Το νευρωνικό δίκτυο που χρησιμοποιήθηκε σ' αυτήν τη μελέτη ήταν ένα πολυστρωματικό δίκτυο εμπρόσθιας τροφοδότησης που χρησιμοποιούσε την τεχνική της προς τα πίσω διάδοσης του σφάλματος, προκειμένου να μοντελοποιήσει τη συνάρτηση $ESDD = f(T, H, R, P, WV)$, όπου το ESDD είναι η εξαρτημένη μεταβλητή και οι μετεωρολογικοί παράγοντες είναι οι ανεξάρτητες μεταβλητές. Το νευρωνικό δίκτυο εκπαιδεύτηκε με τη βοήθεια δεδομένων από μετρήσεις και η εκπαιδευτική ακρίβεια καθορίστηκε από το μέσο τετραγωνικό σφάλμα (RMSE).

2.6.1.2 Λεπτομέρειες για τον αλγόριθμο

Ο αλγόριθμος αυτός για το Τεχνητό Νευρωνικό Δίκτυο έχει χρησιμοποιηθεί επιτυχώς σε πολλές εφαρμογές. Είναι χρήσιμος γιατί λειτουργεί ως μοντέλο για ένα πραγματικό σύστημα ή συνάρτηση και η μελέτη για την πρόβλεψη ή τον έλεγχο του συστήματος γίνεται με βάση το μοντέλο αυτό. Σε ένα Τεχνητό Νευρωνικό Δίκτυο, οι νευρώνες μπορούν να ανήκουν σε κάποιο απ' τα εξής τρία επίπεδα (ή στρώματα): το επίπεδο εισόδου, το επίπεδο εξόδου και τα ενδιάμεσα (κρυμμένα) επίπεδα (hidden layers). Τα σήματα σε ένα ΤΝΔ εμπρόσθιας τροφοδότησης διαδίδονται από το επίπεδο εισόδου στο επίπεδο εξόδου μέσω των κρυμμένων επιπέδων. Κάθε νευρώνας του κρυμμένου

επιπέδου δέχεται ένα σήμα από όλες τις εισόδους μέσω των νευρώνων στο κρυμμένο επίπεδο μέσω συνδέσεων με μεταβλητά βάρη με αυτούς. Το ΤΝΔ μπορεί να αναγνωρίσει τα διανύσματα εισόδου εφόσον τα βάρη σύνδεσης προσαρμοστούν κατάλληλα με μια διαδικασία μάθησης. Ο αλγόριθμος ανάστροφης διάδοσης σφάλματος που είναι μια γενίκευση του κανόνα διόρθωσης σφάλματος Widrow-Hoff είναι η πιο γνωστή μέθοδος για την εκπαίδευση ΤΝΔ και έχει εφαρμοστεί και σ' αυτήν τη μελέτη [9]. Σε επόμενο κεφάλαιο θα δούμε αναλυτικότερα τον αλγόριθμο αυτό.

2.6.1.3 Περιγραφή του προβλήματος

Το προτεινόμενο μοντέλο πραγματοποιήθηκε με τη βοήθεια δεδομένων που αποκτήθηκαν έπειτα από μετρήσεις στο σταθμό παραγωγής Sultan Ismail Power Station. Οι μετεωρολογικές παράμετροι από τις οποίες εξαρτάται το ESDD είναι: η Θερμοκρασία (σε °C), η Υγρασία (%), η Πίεση του αέρα (σε mbar), η Βροχόπτωση (σε mm²) και η Ταχύτητα του ανέμου (σε m/s). Από τις 60 σειρές μετρήσεων που συλλέχθηκαν οι 46 σειρές προτύπων εισόδου-εξόδου χρησιμοποιήθηκαν για την εκπαίδευση και οι υπόλοιπες 14 για τον έλεγχο του ΤΝΔ.

2.6.2 Εκτίμηση του χρόνου έως την υπερπήδηση σε μονωτήρες που βρίσκονται σε συνθήκες ρύπανσης με χρήση ΤΝΔ.

Το άρθρο των Ghosh et al. [10] αναφέρεται στη χρήση ενός ΤΝΔ για την εκτίμηση της συνάρτησης $t = f(V, L, R_p)$, όπου t είναι ο χρόνος μέχρι να συμβεί υπερπήδηση σε ένα ρυπασμένο μονωτήρα, V η εφαρμοζόμενη τάση, L το μήκος του μονωτήρα και R_p η αντίσταση ανά μονάδα μήκους του μονωτήρα. τα προαπαιτούμενα δεδομένα για την εκπαίδευση πάρθηκαν από πειραματικές εργασίες που διενεργήθηκαν σε μοντέλο επίπεδης πλάκας για μονωτήρα με ρύπανση. Βρέθηκε ότι, όταν η εκπαίδευση ολοκληρώθηκε, το ΤΝΔ ήταν ικανό να εκτιμήσει τη συνάρτηση $t = f(V, L, R_p)$ πολύ ικανοποιητικά και αποτελεσματικά ακόμα κι όταν τα ανεπαρκή δεδομένα συμπεριλαμβάνονταν στο σύνολο εκπαίδευσης.

Όπως αναφέρθηκε και στο προηγούμενο κεφάλαιο, σε συνθήκες έντονης ατμοσφαιρικής ρύπανσης, ένα στρώμα ηλεκτρολυτικές ουσίες επικάθεται σταδιακά στην επιφάνεια του μονωτήρα και όταν προστεθεί υγρασία από ομίχλη ή βροχή σχηματίζεται ένα αγώγιμο στρώμα, στο οποίο ρέει το ρεύμα διαρροής. Σοβαρή επιφανειακή ρύπανση και

ανομοιογενής κατανομή του δυναμικού κατά μήκος της επιφάνειας του μονωτήρα προκαλούν την εμφάνιση φωτεινών εκκενώσεων (glow discharges) ή ασταθών τόξων. Μια εκκένωση είναι ικανή να γεφυρώσει εντελώς το μήκος L του μονωτήρα, όταν η εφαρμοζόμενη τάση είναι πάνω από μια κρίσιμη τιμή V_c . Για μια τέτοια τάση, οι εκκενώσεις ή τα ασταθή τόξα επιμηκύνονται μέσω διαδοχικών δενδροειδών σχηματισμών (root formation) πάνω στην επιφάνεια του βρώμικου μονωτήρα. Η ανάπτυξη αυτή συνεχίζεται μέχρι τα ασταθή τόξα να φτάσουν την κρίσιμη απόσταση X_c και είναι το X_c απ' όπου γίνεται η τελική υπερπήδηση για την επίτευξη της γεφύρωσης. Ωστόσο, η επιμήκυνση των τόξων κατά τη φάση που περιγράφηκε εξαρτάται από την ταχύτητα διάδοσης του τόξου. Η ταχύτητα του τόξου πάνω σε επιφάνεια με ρύπους δεν είναι μια σταθερή παράμετρος και εξαρτάται από διάφορους παράγοντες, όπως είναι η αντίσταση ανά μονάδα μήκους R_p , το μήκος L και η εφαρμοζόμενη ac τάση V . Συνεπώς η εξάρτηση του χρόνου t που μεσολαβεί μέχρι την υπερπήδηση από τα V , R_p και L είναι σημαντική στη μελέτη της ανάπτυξης των εκκενώσεων.

Στην εργασία αυτή [10] χρησιμοποιήθηκε το πολυεπίπεδο ΝΔ εμπρόσθιας τροφοδότησης. Τα δεδομένα εισόδου-εξόδου κανονικοποιήθηκαν πριν την έναρξη της διαδικασίας εκπαίδευσης του ΤΝΔ για την καλύτερη σύγκλιση και ακρίβεια της διαδικασίας εκμάθησης. Αρχικά χρησιμοποιήθηκε η τεχνική της προς τα πίσω διάδοσης του σφάλματος με σταθερό ρυθμό μάθησης και σταθερά ορμής και ένα κρυμμένο επίπεδο, προκειμένου να καθοριστούν τα βάρη σύνδεσης μεταξύ των νευρώνων. Στη συνέχεια δοκιμάστηκε ξανά η εκπαίδευση του ΤΝΔ χρησιμοποιώντας αυτή τη φορά δύο κρυμμένα επίπεδα, ενώ μελετήθηκε και το αποτέλεσμα διαφορετικών ρυθμών μάθησης και σταθερών ορμής στη σύγκλιση της διαδικασίας εκπαίδευσης.

Επιπλέον, δοκιμάστηκαν συνδυασμοί για τον αριθμό των νευρώνων στα κρυμμένα επίπεδα για να βρεθεί ο καλύτερος.

Βρέθηκε ότι η χρήση δύο κρυμμένων επιπέδων είχε καλύτερο αποτέλεσμα στη σύγκλιση απ' ότι η χρήση ενός κρυμμένου επιπέδου με τον ίδιο αριθμό κρυμμένων νευρώνων. Τέλος, για να επιταχυνθεί περισσότερο ο ρυθμός σύγκλισης της διαδικασίας μάθησης, παρουσιάστηκε ένας αλγόριθμος προσαρμοστικής μάθησης, στον οποίο τόσο ο ρυθμός μάθησης, όσο και η σταθερά ορμής προσαρμόζονται αυτόματα κατά τη διαδικασία της μάθησης.

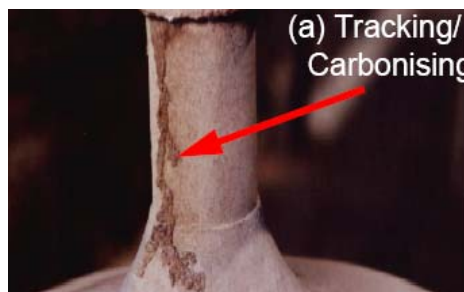
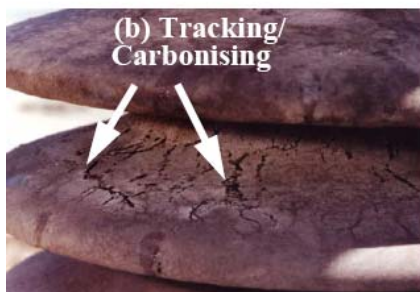
Απ' τη στιγμή που εκπαιδεύτηκε, το ΤΝΔ ήταν ικανό να προβλέπει το χρόνο μέχρι την υπερπήδηση, όταν του δίνονταν οι τιμές για τα V , R_p και L με ένα μέσο απόλυτο σφάλμα της τάξης του 6% σε σχέση με τα πειραματικά δεδομένα.

2.6.3 Χρήση συσκευών ελέγχου ρύπανσης σε συνδυασμό με ΤΝΔ για την πρόβλεψη υπερπήδησης σε μονωτήρες

Οι μονωτήρες των γραμμών μεταφοράς σε βιομηχανικές ζώνες ρυπαίνονται από σωματίδια στον αέρα, με συνέπεια τη διάσπασή τους. Οι μονωτήρες μπορούν να καθαρίζονται περιστασιακά, αλλά αυτή είναι μια δαπανηρή εργασία. Οι Cline et al. [11] εκπαίδευσαν ένα νευρωνικό δίκτυο να ερμηνεύει τα δεδομένα από δύο συσκευές ελέγχου ρύπανσης και να εκτιμά το αν επίκειται διάσπαση των μονωτήρων του υποσταθμού. Οι συσκευές ελέγχου είναι η UE-386 συσκευή υπερήχων και η CAT-ILD για την παρακολούθηση του ρεύματος διαρροής, που συνδυάζεται με μια τοπική ένδειξη της σχετικής υγρασίας. Τα δεδομένα πάρθηκαν από τεστ που διεξήχθησαν στο Ινστιτούτο Έρευνας Ηλεκτρικής Ισχύος (Electric Power Research Institute). Οι προβλέψεις του νευρωνικού δικτύου επικυρώνονται με ένα σύνολο πειραμάτων πάνω σε διασπάσεις.

2.6.4 ΤΝΔ για την ανάλυση επιφανειακών ρωγμών (surface tracking) σε συμπαγείς μονωτήρες

Οι επιφανειακές ρωγμές στους μονωτήρες είναι ένας απ' τους πιο σοβαρούς μηχανισμούς κατάρρευσης που σχετίζεται με τη συμπεριφορά των πολυμερών υλικών μετά από μακρόχρονη λειτουργία. Στα παρακάτω σχήματα φαίνονται τέτοιου είδους φθορές, όπως αυτές αποτυπώνονται στον οδηγό STRI [13].



Διάφοροι τύποι διακοπών μπορούν να ανιχνεύσουν σφάλματα σε μια γραμμή μεταφοράς και να αποτρέψουν την ολική κατάρρευση του συστήματος, αλλά σε ότι αφορά τους συμπαγείς μονωτήρες, λόγω των μη ιάσιμων χαρακτηριστικών τους, στις περισσότερες περιπτώσεις μετά την εμφάνιση και την ανάπτυξη τέτοιων ρωγμών είναι πολύ αργά για να “σωθεί” ένας μονωτήρας.

Οι M. Ugur et al. [12] ασχολήθηκαν με τη χρήση των ΤΝΔ για την ανίχνευση ρωγμών σε αρχικά στάδια εμφάνισης, η μέθοδος όμως που περιγράφεται στην εργασία τους μπορεί να εφαρμοστεί και για την ανίχνευση άλλων επικίνδυνων παραγόντων, όπως οι μερικές εκκενώσεις, το ρεύμα διαρροής και οι σοβαρές βλάβες σε ένα μονωτήρα.

Χρησιμοποιήθηκε ένα ΤΝΔ που χρησιμοποιούσε τον κανόνα ανάστροφης διάδοσης σφάλματος, το οποίο αρχικά εκπαιδεύτηκε με διαφορετικούς τύπους σημάτων. Εξαιτίας του ότι οι εισοδοί στο ΤΝΔ πρέπει να έχουν παρόμοιου τύπου, το σύστημα χρησιμοποιεί το μετασχηματισμό FFT (Fast Fourier Transform) του σήματος εισόδου, το οποίο μπορεί να έχει κι άλλες συνιστώσες συχνότητας πέρα απ’ τη βασική, ανάλογα με το είδος της επιφάνειας. (Λόγω της φύσης της μεθόδου ανάστροφης διάδοσης σφάλματος, διακυμάνσεις στο πλάτος των σημάτων εισόδου, ή στη χρονική στιγμή που αυτά ξεκινούν μπορούν να καταγραφούν σαν διαφορετικοί τύποι αν παρασταθούν απευθείας στο πεδίο του χρόνου. Το πρόβλημα αυτό λύνεται αν τα σήματα παρασταθούν στο χώρο της συχνότητας μέσω του FFT).

Το σύστημα λειτουργεί σε πραγματικό χρόνο, προειδοποιεί το χρήστη με την πρώτη ένδειξη βλάβης στην επιφάνεια του μονωτήρα και μπορεί να προστατεύσει έτσι το μονωτήρα από περαιτέρω φθορά. Για τη συλλογή των δεδομένων έγιναν πειράματα με βάση το πρότυπο IEC587 [12].

2.7 Σύνοψη

Στο κεφάλαιο αυτό παρουσιάστηκαν καινούριες τεχνικές για τον ποιοτικό έλεγχο των μονωτήρων. Μία απ’ αυτές είναι και τα Τεχνητά Νευρωνικά Δίκτυα, για τα οποία θα γίνει εκτενής αναφορά στα επόμενα κεφάλαια. Συγκεκριμένα, στα επόμενα τρία κεφάλαια θα παρουσιαστεί η θεωρία των νευρωνικών δικτύων και στη συνέχεια θα αναπτυχθεί αλγόριθμος για την κατασκευή ΤΝΔ για την εκτίμηση της κρίσιμης τάσης στους μονωτήρες.

Κεφάλαιο 3

Υπολογιστική Νοημοσύνη και Τεχνητά Νευρωνικά Δίκτυα

3.1 Υπολογιστική νοημοσύνη: Ένας νέος κλάδος

Τα τρία βασικά δομικά πεδία της Υπολογιστικής Νοημοσύνης (ΥΝ) είναι η «ασαφής» λογική, τα νευρωνικά δίκτυα και οι γενετικοί αλγόριθμοι.

Τα ασαφή σύνολα και η ασαφής λογική (fuzzy sets-fuzzy logic) παρέχουν τα μέσα για τη σχεδίαση και υλοποίηση συστημάτων τα οποία μπορούν να λειτουργούν και να βγάζουν συμπεράσματα με γλωσσικό τρόπο, μέσα σε συνθήκες αβεβαιότητας, όπως κάνουν οι εμπειρογνώμονες. Τα τεχνητά νευρωνικά δίκτυα (artificial neural networks) μιμούνται την κατανομημένη λειτουργία του ανθρώπινου εγκεφάλου και αποτελούνται από πολλά, απλά δομικά στοιχεία τα οποία ονομάζονται τεχνητοί νευρώνες, που ελέγχονται από προσαρμοζόμενες παραμέτρους και είναι ικανά να μαθαίνουν, να γενικεύουν και να αποκρίνονται με “εξυπνάδα” σε νέα ερεθίσματα. Τα νευρωνικά και ασαφή συστήματα συγκαταλέγονται ανάμεσα στα πιο ευρύτατα χρησιμοποιούμενα εργαλεία της σύγχρονης Τεχνητής Νοημοσύνης (ΤΝ) για τη μοντελοποίηση άγνωστων συστημάτων με μη γραμμική συμπεριφορά, χωρίς να χρειάζονται κάποιο μαθηματικό μοντέλο αυτών. Η ασαφής λογική είναι καταλληλότερη όταν υπάρχει κάποια μορφή γνώσης για το σύστημα, όπως π.χ. η γλωσσική πληροφορία που δίνει κάποιος εμπειρογνώμονας. Από το άλλο μέρος, τα ΤΝΔ έχουν μοναδικές ικανότητες μάθησης και γενίκευσης οι οποίες επιτρέπουν στο χρήστη να κατασκευάζει μοντέλα μη γραμμικών συστημάτων και να επιλύει πολύπλοκα προβλήματα ταξινόμησης με απλή χρήση δεδομένων εισόδου-εξόδου και μόνο. Οι γενετικοί αλγόριθμοι (genetic algorithms) προσφέρουν ένα χρήσιμο και ενδιαφέρον σύνολο γενικών εργαλείων για την επίλυση προβλημάτων βελτιστοποίησης μέσω συστηματικής τυχαίας αναζήτησης, ακολουθώντας τους μηχανισμούς της φυσικής γενετικής [14].

Οι εφαρμογές της υπολογιστικής νοημοσύνης εκτείνονται σε όλους τους τομείς της ανθρώπινης δραστηριότητας. Γενικά, οι τεχνικές της υπολογιστικής νοημοσύνης εφαρμόζονται σε συστήματα και διαδικασίες, οι οποίες απαιτούν μία ή περισσότερες από τις παρακάτω λειτουργίες:

- Προσέγγιση συναρτήσεων και απεικονίσεων
- Μοντελοποίηση στατικών και δυναμικών συστημάτων
- Αναγνώριση και ταξινόμηση προτύπων και γενικά διαδικασιών
- Αναγνώριση της δομής και των παραμέτρων φυσικών και τεχνολογικών συστημάτων
- Εκτίμηση και πρόβλεψη χρονοσειρών και ανελίξεων
- Βελτιστοποίηση διεργασιών και συστημάτων
- Επεξεργασία σημάτων και δεδομένων (1D, 2D)
- Εντοπισμός και αναγνώριση βλαβών και ανωμαλιών
- Επιλογή λύσεων και λήψη αποφάσεων
- Ρύθμιση και έλεγχος βιομηχανικών και μη βιομηχανικών συστημάτων

Στο χώρο της αναγνώρισης προτύπων οι εφαρμογές της υπολογιστικής νοημοσύνης περιλαμβάνουν μεταξύ άλλων τα εξής:

- ✓ Ταξινόμηση και αναγνώριση κειμένου
- ✓ Μάθηση χαοτικής ακολουθίας
- ✓ Επεξεργασία ψηφιακών εικόνων
- ✓ Αναγνώριση φωνής
- ✓ Αναγνώριση χειρόγραφων χαρακτήρων
- ✓ Συμπύεση εικόνας

Στο χώρο του αυτόματου ελέγχου η υπολογιστική νοημοσύνη βρίσκει εφαρμογή στην επίλυση των ακόλουθων προβλημάτων:

- ✓ Αναγνώριση συστήματος
- ✓ Μη γραμμικό φιλτράρισμα
- ✓ Μη γραμμική βελτιστοποίηση
- ✓ Έλεγχος μάθησης
- ✓ Βέλτιστος έλεγχος
- ✓ Προβλεπτικός έλεγχος
- ✓ Προσαρμοστικός έλεγχος

- ✓ Έλεγχος εσωτερικού μοντέλου
- ✓ Μη γραμμικός έλεγχος

Στο χώρο των επιχειρησιακών, διοικητικών και οικονομικών συστημάτων η υπολογιστική νοημοσύνη συμβάλλει στην επίλυση προβλημάτων όπως:

- ✓ Προβλήματα χρονοπρογραμματισμού και σχεδίαση εργασιών
- ✓ Προβλήματα επιλογής και οργάνωσης του προσωπικού επιχειρήσεων
- ✓ Προβλήματα καθορισμού τιμολογιακών πολιτικών
- ✓ Προβλήματα επιλογής τροφοδοτών και βελτιστοποίησης της παραγωγής βιομηχανικών προϊόντων
- ✓ Προβλήματα δικτύων διανομής προϊόντων
- ✓ Προβλήματα σχεδίασης συστημάτων υποστήριξης αποφάσεων
- ✓ Προβλήματα ανανέωσης εξοπλισμού / μοντερνοποίησης βιομηχανικών συστημάτων.

Η ιστορία της Υπολογιστικής Νοημοσύνης περιλαμβάνει από το ένα μέρος την ιστορία καθενός απ' τους κλάδους που τη συγκροτούν (Ασαφής Λογική, Νευρωνικά Δίκτυα, Γενετικοί Αλγόριθμοι) και απ' το άλλο την ιστορία της κοινής πορείας αυτών στο πλαίσιο της.

3.1.1 Ασαφή Σύνολα και Ασαφής Λογική

Η μαθηματική θεμελίωση της ασαφούς λογικής βασίζεται στη θεωρία των ασαφών συνόλων (fuzzy set theory), η οποία μπορεί να θεωρηθεί σαν μια γενίκευση της κλασικής θεωρίας συνόλων.

Απαραίτητη προϋπόθεση για την εκμάθηση των εφαρμογών και των αρχών που διέπουν την ασαφή λογική είναι η γνώση των αρχών των ασαφών συνόλων [15].

Η θεωρία ασαφών συνόλων θεμελιώθηκε στην παρούσα της μορφή από το μηχανικό αυτομάτου ελέγχου Lofti Zadeh το 1965. Η ιδέα όμως της χρήσης ασαφών μαθηματικών δομών έχει χρησιμοποιηθεί από τον Poincaré το 19^ο αιώνα, στο πλαίσιο της “οπτικής αντίληψης”. Εξάλλου, στις αρχές της δεκαετίας του '70 αναπτύχθηκε η λεγόμενη ασαφής γεωμετρία από τον Zeeman, η οποία όμως είχε πολύ περιορισμένη διάδοση σε σχέση με τη θεωρία του Zadeh. Η θεωρία του Zadeh βρήκε εφαρμογές στην τεχνολογία, στην

οργάνωση επιχειρήσεων, στην ψυχολογία, αλλά και σε περιοχές όπως η σημειολογία (semiotics).

Με τη θεωρία των ασαφών συνόλων και της ασαφούς λογικής βρίσκουμε χρήσιμες και αποδοτικές λύσεις σε δύσκολα πρακτικά προβλήματα, για τα οποία δεν έχουμε ακριβείς μαθηματικές περιγραφές και μοντέλα. Το γεγονός αυτής της αβεβαιότητας και ανακρίβειας διατυπώθηκε από τον Zadeh ως μια αρχή, την οποία ονόμασε “αρχή του ασυμβίβαστου”. Η αρχή του ασυμβίβαστου λέει ότι:

«Καθώς η πολυπλοκότητα ενός συστήματος αυξάνει, η ικανότητά μας να προβαίνουμε σε ακριβείς και σημαντικές δηλώσεις για τη συμπεριφορά του μειώνεται, μέχρις ότου φτάσουμε σε ένα όριο (κατώφλι) πέρα από το οποίο ακρίβεια και σημαντικότητα (ή σχετικότητα) καθίστανται σχεδόν αμοιβαίως αποκλειόμενα χαρακτηριστικά».

Ο τρόπος αντιμετώπισης του ασυμβίβαστου αυτού με τη θεωρία των ασαφών συνόλων είναι να παραστήσουμε τη συγκεκριμένη, αόριστη και μη ακριβή γνώση του ανθρώπου άμεσα, χωρίς τη μεσολάβηση κάποιας τεχνητής παράστασης, όπως είναι π.χ. ένας ακριβής μαθηματικός τύπος.

Μαθηματικά, ο όρος ασάφεια (fuzziness) σημαίνει πολλαπλότητα τιμών και πηγάζει από την αρχή αβεβαιότητας του Heisenberg [14]. Για παράδειγμα, έστω ένα σύνολο από καρέκλες σε ένα δωμάτιο. Στη θεωρία συνόλων, το σύνολο από καρέκλες μπορεί να προσδιοριστεί δείχνοντας κάθε αντικείμενο στο δωμάτιο και κάνοντας την ερώτηση “Είναι αυτό μια καρέκλα;”. Στην κλασική θεωρία συνόλων μας επιτρέπεται να χρησιμοποιήσουμε μόνο δυο απαντήσεις: *Ναι* ή *Όχι*. Αν κωδικοποιήσουμε το ναι με "1" και το όχι με "0", οι απαντήσεις μας θα βρίσκονται στο ζεύγος {0,1}. Εάν η απάντηση είναι "1" το στοιχείο ανήκει στο σύνολο, ενώ, αν η απάντηση είναι "0", δεν ανήκει. Στο τέλος, συλλέγουμε όλα τα αντικείμενα για τα οποία πήραμε την απάντηση "1" και βρίσκουμε το “σύνολο των καρεκλών σε ένα δωμάτιο”. Αν υποθέσουμε τώρα ότι γίνεται η ερώτηση “Ποια αντικείμενα σε ένα δωμάτιο μπορούν να χρησιμοποιηθούν σαν καρέκλα;”. Και πάλι δείχνοντας κάθε αντικείμενο γίνεται η ερώτηση "μπορεί αυτό να χρησιμοποιηθεί σαν καρέκλα;". Η απάντηση εδώ μπορεί επίσης, τεχνικά να περιοριστεί στο {0,1}. Ακόμα, το σύνολο από αντικείμενα σε ένα δωμάτιο τα οποία μπορούν να χρησιμοποιηθούν σαν καρέκλα μπορεί να περιλαμβάνει όχι μόνο καρέκλες αλλά επίσης τραπέζια, κουτιά, τμήματα του πατώματος κ.λ.π. δηλαδή είναι ένα σύνολο όχι

μονοσήμαντα ορισμένο. Όλα εξαρτώνται από το τι εννοούμε με τη λέξη χρήση. Λέξεις σαν τη «χρήση» έχουν διαφορετικές χροιές νοήματος και μπορούν να χρησιμοποιηθούν με πολλούς διαφορετικούς τρόπους. Το νόημα και η χρήση τους μπορεί να μεταβάλλεται από άνθρωπο σε άνθρωπο, ανά σκοπό και ανά περίπτωση. Αυτό εξαρτάται από τον ακριβή ορισμό μιας κατάστασης. Για το λόγο αυτό θεωρούμε ότι το "σύνολο των αντικειμένων τα οποία μπορούν να χρησιμοποιηθούν σαν καρέκλες", είναι ένα ασαφές σύνολο (fuzzy set), με την έννοια ότι μπορεί να μην υπάρχουν συγκεκριμένα καθορισμένα κριτήρια για να αποφασίσουμε τη συμμετοχή ενός αντικειμένου στο σύνολο. Αντικείμενα όπως τραπέζια, κιβώτια, τμήματα του πατώματος, μπορούν να χρησιμοποιηθούν σαν καρέκλες σε ένα βαθμό.

Θα πρέπει να σημειωθεί ότι δεν υπάρχει τίποτα το ασαφές (fuzzy) σχετικά με τα ίδια τα υλικά αντικείμενα: καρέκλες, κιβώτια και τραπέζια είναι πραγματικά αντικείμενα. Η ασάφεια είναι ένα χαρακτηριστικό της αναπαράστασης τους σε ένα περιβάλλον συμβόλων και γενικά μια ιδιότητα των μοντέλων και της γλώσσας [15]. Σύμφωνα με την κλασική θεωρία συνόλων υπάρχει μια μάλλον αυστηρή έννοια της συμμετοχής ενός αντικειμένου σε ένα σύνολο, δηλαδή ένα στοιχείο ή θα ανήκει ή δεν θα ανήκει σε ένα σύνολο. Στα ασαφή σύνολα περισσότεροι από δύο βαθμοί συμμετοχής είναι επιτρεπτοί. Ο βαθμός συμμετοχής (membership function) ενός στοιχείου σε ένα σύνολο περιγράφεται με έναν αριθμό μεταξύ του "0" και του "1", δηλαδή έναν αριθμό στο διάστημα $[0,1]$. Υπάρχει μια γενίκευση του συνόλου τιμών από το ζεύγος αριθμών $\{0,1\}$ σε όλους τους αριθμούς που υπάρχουν στο διάστημα $[0,1]$. Επεκτείνοντας το σύνολο τιμών αλλάζει η φύση της χαρακτηριστικής συνάρτησης, που τώρα λέγεται *συνάρτηση συμμετοχής* (membership function). Τώρα πλέον δεν έχουμε "crisp sets", αλλά έχουμε ασαφή σύνολα. Από τη στιγμή που το διάστημα $[0,1]$ περιέχει μία απειρία αριθμών, μία απειρία βαθμών συμμετοχής είναι δυνατή.

Οι συναρτήσεις συμμετοχής μπορούν να αναπαραστήσουν την υποκειμενική ιδέα μιας αόριστης έκφρασης. Για παράδειγμα αντικείμενα σε ένα δωμάτιο που μπορούν να λειτουργήσουν σαν καρέκλες, επιθυμητή απόδοση, ελάχιστη βελτίωση, μεγάλη βελτίωση είναι αόριστες εκφράσεις οι οποίες μπορούν να παρασταθούν με τη βοήθεια των συναρτήσεων συμμετοχής. Οι συναρτήσεις συμμετοχής μπορούν επίσης να υπολογιστούν με βάση τα στατιστικά δεδομένα.

Οι δύο πιο συχνά χρησιμοποιούμενοι τρόποι ορισμού των ασαφών συνόλων είναι οι εξής:

1. Εάν το X είναι ένας χώρος αναφοράς (universe of discourse) και το x είναι ένα στοιχείο του X , τότε το ασαφές σύνολο A που ορίζεται στον X μπορεί να παρασταθεί με τη βοήθεια της παρακάτω σχέσης :

$$A = \{(x, \mu_A(x))\}, x \in X$$

όπου το $\mu_A(x)$ είναι ο βαθμός συμμετοχής του στοιχείου x στο ασαφές σύνολο A . Κάθε ζευγάρι $(x, \mu_A(x))$ είναι το ζεύγος στοιχείου, τιμής συνάρτησης συμμετοχής και ονομάζεται “singleton”. Στα “crisp” σύνολα, ένα “singleton” είναι το στοιχείο x από μόνο του. Στα ασαφή σύνολα ένα “singleton” αποτελείται από το στοιχείο x και το βαθμό συμμετοχής του $\mu_A(x)$.

2. Ένα singleton μπορεί επίσης να παρασταθεί σαν $\mu_A(x)/x$, δηλαδή βάζοντας πρώτα τον βαθμό συμμετοχής στη συνέχεια το σύμβολο "/" και τέλος το στοιχείο x . Τα στοιχεία που έχουν βαθμό συμμετοχής μηδέν σε ένα ασαφές σύνολο μπορούν να παραληφθούν. Ο χώρος αναφοράς του ασαφούς συνόλου A είναι το σύνολο των στοιχείων του τα οποία έχουν βαθμό συμμετοχής διαφορετικό από το μηδέν. Με βάση τα παραπάνω ένα ασαφές σύνολο μπορεί να θεωρηθεί σαν η ένωση όλων των “singleton” $\mu_A(x)/x$:

$$A = \sum_{i=1}^n \mu_A(x_i) / x_i \quad (3.1)$$

Το σύμβολο του αθροίσματος στην παραπάνω σχέση παριστάνει την ένωση όλων των “singleton”. Η σχέση αυτή ισχύει στην περίπτωση που έχουμε ένα πεπερασμένο σύνολο αναφοράς για το A , ενώ για συνεχή χώρο αναφοράς, η σχέση αυτή γράφεται:

$$A = \int_x \mu_A(x) / x \quad (3.2)$$

Ο βαθμός συμμετοχής σε ένα ασαφές σύνολο μπορεί να είναι και αυτός ο ίδιος ένα ασαφές σύνολο [15].

3.1.2 Γενετικοί Αλγόριθμοι

Στις δεκαετίες του '50 και του '60 διάφοροι επιστήμονες υπολογιστών μελέτησαν ανεξάρτητα ο ένας από τον άλλο εξελικτικά συστήματα με τη σκέψη ότι η “εξέλιξη” θα μπορούσε να χρησιμοποιηθεί ως ένα εργαλείο βελτιστοποίησης σε τεχνολογικά

προβλήματα. Η ιδέα πίσω από τα συστήματα αυτά ήταν η ανάπτυξη ενός πληθυσμού από υποψήφιες λύσεις σε ένα δεδομένο πρόβλημα, χρησιμοποιώντας τελεστές εμπνεόμενους από τη φυσική γενετική μεταβολή και τη φυσική επιλογή [14]. Οι γενετικοί αλγόριθμοι εμφανίστηκαν για πρώτη φορά στα μέσα της δεκαετίας του 1970, όταν ο J. Holland δημιούργησε ένα υπολογιστικό μοντέλο που μπορούσε να λύνει προβλήματα βελτιστοποίησης. Σκοπός της έρευνας του Holland ήταν η ερμηνεία των προσαρμοστικών διαδικασιών των φυσικών συστημάτων και ο σχεδιασμός συστημάτων που έχουν ίδια χαρακτηριστικά με τα φυσικά συστήματα.

Στην περίπτωση των γενετικών αλγορίθμων, οι μεταβλητές αναπαριστώνται ως γονίδια σε ένα χρωμόσωμα. Οι γενετικοί αλγόριθμοι (ΓΑ) απεικονίζουν μία ομάδα υποψήφιων λύσεων (πληθυσμός) στην επιφάνεια ελέγχου (επιφάνεια ανίχνευσης της βέλτιστης λύσης). Με χρήση φυσικής επιλογής και τελεστών γενετικής, όπως μετάλλαξη και διασταύρωση γονιδίων, βρίσκονται τα χρωμοσώματα με την καλύτερη «καταλληλότητα» (fitness). Η φυσική επιλογή εγγυάται ότι τα χρωμοσώματα με την καλύτερη καταλληλότητα θα διαδίδονται στους μελλοντικούς πληθυσμούς. Χρησιμοποιώντας τον τελεστή διασταύρωσης γονιδίων, ο γενετικός αλγόριθμος συνδυάζει γονίδια από χρωμοσώματα δύο γονέων και παράγει δύο νέα χρωμοσώματα (απογόνους), τα οποία έχουν μεγάλη πιθανότητα να έχουν καλύτερη καταλληλότητα από τους γονείς τους. Η μετάλλαξη επιτρέπει να ελεγχθούν νέες περιοχές της επιφάνειας ελέγχου. Με αυτόν τον τρόπο η καταλληλότητα των χρωμοσωμάτων βελτιώνεται και μετά από πολλές γενιές θα δημιουργήσουν χρωμοσώματα που θα περιέχουν τις βέλτιστες τιμές των υπό μελέτη μεταβλητών (επίλυση του προβλήματος βελτιστοποίησης) [16].

Οι λόγοι που καθιστούν συμφέρουσα τη χρήση γενετικών αλγορίθμων είναι ότι είναι υπολογιστικά απλοί και ισχυροί στην αναζήτηση της βέλτιστης λύσης (ελάχιστο ή μέγιστο). Επιπλέον, δεν περιορίζονται ουσιαστικά από υποθέσεις για το χώρο αναζήτησης – υποθέσεις σχετικά με τη συνέχεια, την ύπαρξη παραγώγων και την ύπαρξη ενός μόνο μεγίστου. Λόγω της διαδικασίας της εξέλιξης δεν εγκλωβίζονται σε τοπικά ακρότατα, όπως συμβαίνει με άλλες τεχνικές βελτιστοποίησης. Οι γενετικοί αλγόριθμοι διαφέρουν αρκετά σε σχέση με τις υπόλοιπες τεχνικές, αφού αναπαριστούν τις παραμέτρους του προβλήματος με τη μορφή γονιδίων σε ένα χρωμόσωμα, όπως στη βιολογία.

Οι άλλες μέθοδοι βελτιστοποίησης (π.χ. μέθοδος σημείου-σημείου, στην οποία γίνεται μετακίνηση από ένα σημείο στο χώρο απόφασης σε ένα άλλο χρησιμοποιώντας κάποιον κανόνα μετάβασης) οδηγούνται πολύ συχνά σε λανθασμένο υπολογισμό του μεγίστου στις περιπτώσεις περιοχών με περισσότερα από ένα μέγιστα. Οι γενετικοί αλγόριθμοι, όμως, χρησιμοποιούν ταυτόχρονα μία ευρεία βάση δεδομένων σημείων ψάχνοντας πολλά ελάχιστα (ή μέγιστα ανάλογα με το πρόβλημα) ταυτόχρονα, με αποτέλεσμα να μειώνεται η πιθανότητα σφάλματος.

Σε αντίθεση με τις υπόλοιπες μεθόδους, που χρησιμοποιούν αιτιοκρατικούς κανόνες μετάβασης, οι γενετικοί αλγόριθμοι χρησιμοποιούν πιθανοτικούς κανόνες μετάβασης για την καθοδήγηση της αναζήτησης. Συνήθεις τεχνικές αναζήτησης, όπως είναι η τεχνική της κλίσης, απαιτούν τη γνώση των παραγώγων. Οι γενετικοί αλγόριθμοι πλεονεκτούν στο ότι δεν χρειάζονται βοηθητική πληροφορία. Δηλαδή μπορούν να επιτύχουν αποδοτική αναζήτηση με μοναδική απαίτηση τον υπολογισμό της τιμής της αντικειμενικής συνάρτησης που σχετίζεται με τις ατομικές συμβολοσειρές.

Τελικά η χρήση των γενετικών αλγορίθμων παρουσιάζει τα παρακάτω πλεονεκτήματα [16]:

- Δεν απαιτούν γνώση ή πληροφορία για την κλίση της επιφάνειας ελέγχου.
- Πιθανές ασυνέχειες στην επιφάνεια ελέγχου έχουν μικρή επίδραση στη συνολική απόδοση της βελτιστοποίησης.
- Συμπεριφέρονται πολύ καλά σε μεγάλης κλίμακας προβλήματα βελτιστοποίησης.
- Μπορούν να εφαρμοστούν σε μία ευρεία κλίμακα προβλημάτων βελτιστοποίησης.

Όμως η χρήση των γενετικών αλγορίθμων παρουσιάζει και κάποια μειονεκτήματα, όπως:

- Δεν καταφέρνουν πάντα να βρουν το ολικό ελάχιστο (ή μέγιστο ανάλογα με το πρόβλημα).
- Απαιτούν μεγάλο αριθμό υπολογισμών της συνάρτησης *καταλληλότητας*.
- Η *διαμόρφωση* (configuration) των παραμέτρων τους δεν είναι άμεση και απαιτεί δοκιμαστικές προσπάθειες.

3.1.2.1 Λειτουργία Γενετικών Αλγορίθμων

Ο πυρήνας του γενετικού αλγορίθμου αποτελείται από πέντε σημαντικά βήματα:

(α) *Αρχικοποίηση* (Initialization): Ο αρχικός πληθυσμός των χρωμοσωμάτων δημιουργείται είτε τυχαία, είτε διαταράσσοντας ένα χρωμόσωμα εισόδου. Ο τρόπος με τον οποίο γίνεται η αρχικοποίηση δεν θεωρείται κρίσιμος, καθώς ο αρχικός πληθυσμός εξελίσσεται και εκτείνεται σε ένα μεγάλο εύρος τιμών των υπό βελτιστοποίηση μεταβλητών (έχει ένα ποικιλόμορφο πληθυσμό). Όμως, αν υπάρχει σαφής γνώση για το σύστημα που πρόκειται να βελτιστοποιηθεί η πληροφορία αυτή μπορεί να συμπεριληφθεί στον αρχικό πληθυσμό. Για παράδειγμα, στο πρόβλημα επιλογής του μήκους κύματος στην υπέρυθη φασματοσκοπία, μπορεί αρχικά να επιλεγούν τα μήκη κύματος που αντιστοιχούν σε γνωστές ομάδες απορρόφησης στο πρώτο χρωμόσωμα του πληθυσμού. Τότε, όλα τα χρωμοσώματα στον αρχικό πληθυσμό μπορούν να παραχθούν με τυχαία διαταραχή των χρωμοσώματος εισόδου. Με αυτόν τον τρόπο ο αρχικός πληθυσμός είναι ακόμα ποικιλόμορφος, και επίσης ο αλγόριθμος βελτιστοποίησης δεν θα χάνει χρόνο εξερευνώντας περιοχές της επιφάνειας ελέγχου οι οποίες δεν παράγουν καλά μοντέλα.

(β) *Αξιολόγηση* (Evaluation): Στο δεύτερο βήμα υπολογίζεται η *καταλληλότητα*. Ο στόχος της συνάρτησης καταλληλότητας είναι να υπολογίσει αριθμητικά τη συμπεριφορά του χρωμοσώματος και να αξιολογήσει την ποιότητά του. Στις πραγματικές εφαρμογές μεθόδων βελτιστοποίησης, όπως οι γενετικοί αλγόριθμοι, το πιο κρίσιμο βήμα είναι η επιλογή της συνάρτησης καταλληλότητας.

(γ) *Αξιοποίηση* (Exploitation): Το τρίτο βήμα είναι η αξιοποίηση ή το βήμα της φυσικής επιλογής. Στο βήμα αυτό, τα χρωμοσώματα με τη *μεγαλύτερη βαθμολογία καταλληλότητας* (fitness score) τοποθετούνται μία ή περισσότερες φορές σε ένα *υποσύνολο ζευγαρώματος* (mating subset) με τρόπο περίπου τυχαίο. Τα χρωμοσώματα με χαμηλή βαθμολογία καταλληλότητας απομακρύνονται από τον πληθυσμό. Υπάρχουν διάφορες μέθοδοι για την υλοποίηση της αξιοποίησης. Μία από τις πιο κοινές μεθόδους είναι η *δυναδικού αγώνα μέθοδος επιλογής υποσυνόλου ζευγαρώματος* (binary tournament mating subset selection method). Σύμφωνα με τη μέθοδο αυτή, κάθε χρωμόσωμα στον πληθυσμό ανταγωνίζεται για μία θέση στο υποσύνολο ζευγαρώματος. Δύο χρωμοσώματα απομακρύνονται τυχαία από τον πληθυσμό, το χρωμόσωμα με τη μεγαλύτερη βαθμολογία καταλληλότητας τοποθετείται στο υποσύνολο ζευγαρώματος. Και τα δύο χρωμοσώματα επιστρέφουν στον πληθυσμό και ένας καινούριος αγώνας αρχίζει. Η διαδικασία αυτή συνεχίζεται μέχρι το υποσύνολο ζευγαρώματος να ολοκληρωθεί. Ένα

χαρακτηριστικό αυτού του σχήματος είναι ότι το χειρότερο χρωμόσωμα του πληθυσμού ποτέ δεν θα επιλεγεί στο υποσύνολο ζευγαρώματος.

(δ) *Διασταύρωση* (Ανασυνδυασμός - crossover): Στο τέταρτο βήμα αυτό οι λύσεις του ενδιαμέσου πληθυσμού ανασυνδυάζονται για την παραγωγή του επόμενου πληθυσμού με χρήση τελεστών που προσομοιώνουν αντίστοιχους γενετικούς μηχανισμούς. Συγκεκριμένα, από τον ενδιαμέσο πληθυσμό γίνεται τυχαία επιλογή των λύσεων ανά δύο, και έτσι δημιουργούνται οι λύσεις-γονείς. Έπειτα εφαρμόζεται στους γονείς ο τελεστής διασταύρωσης με μια πιθανότητα P_c . Η διασταύρωση ανασυνδυάζει τις συμβολοσειρές (γενετικό υλικό) των γονέων δημιουργώντας δύο απογόνους που κληρονομούν χαρακτηριστικά και των δύο γονέων.

Αναλυτικότερα δύο χρωμοσώματα-γονείς από το υποσύνολο ζευγαρώματος επιλέγονται τυχαία να ζευγαρωθούν. Συνήθως επιλέγεται μία υψηλή πιθανότητα (π.χ. 0,95) με την οποία τα χρωμοσώματα θα ανασυνδυάσουν τα γονίδια τους. Αν επιτραπεί στους γονείς να διασταυρωθούν, εφαρμόζεται ένας τελεστής ανασυνδυασμού των γονιδίων, ο οποίος ανταλλάσσει γονίδια ανάμεσα στους δύο γονείς και παράγει δύο απογόνους. Αν δεν επιτραπεί να διασταυρωθούν, οι γονείς τοποθετούνται στην επόμενη γενιά απaráλλακτοι. Οι δύο πιο συχνά χρησιμοποιούμενοι τελεστές ανασυνδυασμού των γονιδίων είναι οι μέθοδοι διασταύρωσης ενός και δύο σημείων. Στη μέθοδο διασταύρωσης ενός σημείου, επιλέγεται ένα σημείο διασταύρωσης στο χρωμόσωμα και τα γονίδια μέχρι εκείνο το σημείο ανταλλάσσονται ανάμεσα στους δύο γονείς. Στη μέθοδο διασταύρωσης δύο σημείων, επιλέγονται δύο σημεία διασταύρωσης και τα γονίδια ανταλλάσσονται ανάμεσα στους δύο γονείς. Οι απόγονοι αντικαθιστούν τους γονείς στην επόμενη γενιά. Ένας τρίτος τελεστής ανασυνδυασμού των γονιδίων, ο οποίος έχει γίνει αρκετά δημοφιλής τελευταία, είναι η μέθοδος της ομοιόμορφης διασταύρωσης. Σε αυτήν τη μέθοδο διασταύρωσης εφαρμόζεται ανασυνδυασμός των γονιδίων σε ατομικά γονίδια του χρωμοσώματος. Αν εκτελεστεί διασταύρωση, ανταλλάσσονται τα γονίδια ανάμεσα στους δύο γονείς, ενώ, αν δε γίνει διασταύρωση, τα γονίδια παραμένουν ανέπαφα. Αυτή η μέθοδος διασταύρωσης έχει μία υψηλότερη πιθανότητα να παράγει απογόνους οι οποίοι είναι πολύ διαφορετικοί από τους γονείς τους. Γι' αυτό η πιθανότητα ανασυνδυασμού των γονιδίων συνήθως τίθεται σε μία χαμηλή τιμή (π.χ. 0,1).

Η πιθανότητα να συμβεί η διασταύρωση είναι επίσης επιλέξιμη και συνήθως τίθεται σε μία χαμηλή τιμή (π.χ. 0,01) έτσι, ώστε να μην καταστρέφονται καλά χρωμοσώματα.

(ε) *Μετάλλαξη* (mutation): Στο πέμπτο βήμα πραγματοποιείται με τυχαία αλλαγή γονιδίων των απογόνων. Η διασταύρωση, αν και αποτελεί το βασικό μηχανισμό αναζήτησης νέων λύσεων, δεν είναι ωστόσο σε θέση να παράγει πληροφορία που δεν υπάρχει ήδη μέσα στον πληθυσμό. Ο τελεστής της μετάλλαξης καλύπτει αυτήν την ανάγκη εισάγοντας νέα πληροφορία στους απογόνους. Από πλευράς υλοποίησης η μετάλλαξη απλά αλλάζει την τιμή ενός συγκεκριμένου γονιδίου.

Κατά άλλους ερευνητές τα δύο τελευταία βήματα είναι ενοποιημένα σε ένα τέταρτο βήμα με όνομα «εξερεύνηση» (exploration) και αποτελείται από τους τελεστές ανασυνδυασμού των γονιδίων και μετάλλαξης.

Μετά το (ε) βήμα ο πληθυσμός είναι πλήρης από νέα χρωμοσώματα (απογόνους) και εκτελούνται ξανά τα βήματα (β) έως (ε). Αυτή η διαδικασία επαναλαμβάνεται όσο δεν ικανοποιούνται τα κριτήρια τερματισμού, δηλαδή είτε πρόκειται για ένα προκαθορισμένο αριθμό γενεών, είτε μέχρι ο βαθμός καταλληλότητας της καλύτερης λύσης να είναι μεγαλύτερος από κάποιο συγκεκριμένο όριο.

3.1.2.2 Πότε μπορούν να χρησιμοποιηθούν οι Γενετικοί Αλγόριθμοι

Η απάντηση στο ερώτημα αν μπορούν να χρησιμοποιηθούν οι γενετικοί αλγόριθμοι σε ένα πρόβλημα βελτιστοποίησης είναι αρκετά δύσκολη. Οι περισσότεροι ερευνητές ισχυρίζονται ότι δεν υπάρχει καλύτερη μέθοδος βελτιστοποίησης. Υπάρχουν παραπλήσιες δυνατότητες για κάθε τεχνική βελτιστοποίησης (*προσομοιωμένη απόπτωση-simulated annealing, γενετικοί αλγόριθμοι, ανύψωση κλίσης-gradient ascent, βελτιστοποίηση Simplex-Simplex optimization, Monte Carlo*, κτλ.) για να εκτελεστεί το ίδιο έργο. Το κλειδί είναι η χρήση της γνώσης του συστήματος που πρόκειται να βελτιστοποιηθεί, ώστε έπειτα να επιλεγεί η καλύτερη τεχνική, η οποία θα βρίσκει τη «βέλτιστη» λύση γρηγορότερα.

Όταν δεν υπάρχει αρκετή πληροφορία για την επιφάνεια ελέγχου-βελτιστοποίησης και ο υπολογισμός της κλίσης είναι είτε εντατικός, είτε αριθμητικά ασταθής, πολλοί ερευνητές προτιμούν τη χρήση μεθόδων βελτιστοποίησης, όπως *γενετικοί αλγόριθμοι* και *βελτιστοποίηση Simplex*, οι οποίες δεν απαιτούν πληροφορία για την κλίση. Ένας από

τους λόγους που μπορεί να προτιμηθούν οι γενετικοί αλγόριθμοι είναι η *προσαρμοστικότητα* τους (versatility). Χρησιμοποιώντας τη γνώση για ένα σύστημα μπορεί να προσαρμοστεί ο αλγόριθμος για τη συγκεκριμένη εφαρμογή. Αν ένα πρόβλημα παγιδευτεί σε ένα τοπικό ελάχιστο, μπορεί να αυξηθεί η μετάλλαξη. Έτσι, ενώ δεν υπάρχει εγγύηση ότι οι γενετικοί αλγόριθμοι μπορούν να συμπεριφερθούν καλύτερα σε μία εφαρμογή, συχνά υπάρχει η δυνατότητα μέσα από την τροποποίηση των παραμέτρων διαμόρφωσης του γενετικού αλγόριθμου ή με χρήση διαφορετικών γενετικών τελεστών να επιτευχθεί πολύ ικανοποιητική επίλυση του προβλήματος βελτιστοποίησης.

Ένα άλλο χαρακτηριστικό των γενετικών αλγορίθμων είναι ότι δεν βελτιστοποιούν απευθείας τις μεταβλητές, αλλά τις παραστάσεις τους. Για εφαρμογές, όπως επιλογή χαρακτηριστικών, όπου το χρωμόσωμα μπορεί να κωδικοποιηθεί με τέτοιο τρόπο, ώστε κάθε γονίδιο να παριστάνει ένα συγκεκριμένο χαρακτηριστικό, οι γενετικοί αλγόριθμοι είναι μία άριστη επιλογή.

Για εφαρμογές όπου οι υπό βελτιστοποίηση μεταβλητές διαφέρουν πολύ μεταξύ τους, π.χ. ακέραιες, δυαδικές, και πραγματικές μεταβλητές, οι γενετικοί αλγόριθμοι είναι επίσης μία άριστη επιλογή. Η διαμόρφωσή τους μπορεί να τροποποιηθεί, ώστε να συμπεριλάβει διαφορετικούς τελεστές μετάλλαξης για διαφορετικά τμήματα του χρωμοσώματος.

Για εφαρμογές όπου ο υπολογισμός της κλίσης του διανύσματος είναι αριθμητικά ακριβής και γρήγορος μπορεί να χρησιμοποιηθεί κάποια μορφή της μεθόδου της *καθοδικής κλίσης*. Οι γενετικοί αλγόριθμοι θα δουλέψουν και σε αυτού του τύπου τις εφαρμογές, αλλά θα φτάσουν τη βέλτιστη περιοχή πολύ πιο αργά από τις μεθόδους «*αναρρίχησης του λόφου*» (hill-climbing).

Εφαρμογές που απαιτούν την εύρεση του ακριβούς ολικού ελάχιστου ίσως είναι μία πρόκληση για τους γενετικούς αλγόριθμους. Οι τελευταίοι είναι καλύτεροι στην εύρεση της περιοχής του ολικού ελάχιστου, αλλά μερικές φορές έχουν πρόβλημα να βρουν τη θέση του ακριβούς ολικού ελάχιστου. Πολλοί ερευνητές χρησιμοποιούν τους γενετικούς αλγόριθμους για να φτάσουν κοντά στη βέλτιστη περιοχή και μετά χρησιμοποιούν μία άλλη μέθοδο για την τελική εξερεύνηση.

Μία από τις πιο συχνά παρατηρημένες δυσκολίες με τους γενετικούς αλγόριθμους είναι ότι συγκρινόμενοι με τις μεθόδους «*αναρρίχησης του λόφου*» γενικά απαιτούν

περισσότερους υπολογισμούς της συνάρτησης καταλληλότητας. Αν η επιφάνεια της συνάρτησης καταλληλότητας είναι εντελώς ομαλή, τότε μία μέθοδος αναρρίχησης του λόφου, όπως η βελτιστοποίηση *Simplex* θα συμπεριφέρεται καλύτερα από τον γενετικό αλγόριθμο για ένα δεδομένο αριθμό ανακυκλώσεων.

Δυστυχώς, υπάρχουν συγκεκριμένα προβλήματα βελτιστοποίησης, τα οποία αποτελούν μία δύσκολη πρόκληση για τους γενετικούς αλγόριθμους. Μία από τις κύριες περιοχές έρευνας είναι η μελέτη των εφαρμογών αυτού του τύπου και η ανάπτυξη μεθόδων που να προσδιορίζουν προκαταβολικά αν το πρόβλημα βελτιστοποίησης είναι δύσκολο να επιλυθεί με γενετικούς αλγόριθμους. Σχετικά πρόσφατα οι θεωρητικοί των γενετικών αλγορίθμων εντόπισαν μερικές από τις πιο συχνές αιτίες που δυσχεραίνουν την εφαρμογή τους. Μία από αυτές είναι όταν τα γονίδια του χρωμοσώματος δεν είναι κατάλληλα διευθετημένα και το μήκος των σχημάτων*, $L(H)$, είναι υπερβολικά μεγάλο για να επεξεργαστεί από τον γενετικό αλγόριθμο. Σύμφωνα με το θεώρημα του σχήματος, σχήματα μεγαλύτερου μήκους έχουν υψηλότερη πιθανότητα να καταστραφούν από μετάλλαξη και διασταύρωση των γονιδίων. Ένα σχήμα αλλαγής της διευθέτησης όπου τα «0» και τα «1» είναι πιο κοντά μεταξύ τους στο σχήμα του χρωμοσώματος, θα παρεμποδίζει την καταστροφή αυτή μειώνοντας το $L(H)$. Μία αυτόματη μέθοδος αλλαγής της διευθέτησης των χρωμοσωμάτων έχει αναπτυχθεί από τον D. Goldberg.

*Σημείωση: Ένα σχήμα (H) ορίζεται ως ένα πρότυπο για την περιγραφή ενός συνόλου χρωμοσωμάτων με παρόμοια τμήματα. Το πρότυπο αποτελείται από πολλαπλά «μηδέν» και «ένα», και από μετα-χαρακτήρες ή σύμβολα «μη ενδιαφέροντος», όπως #. Ο μετα-χαρακτήρας είναι απλά ένα συμβολικό εργαλείο που χρησιμοποιείται για να υποδηλώνει ότι σε εκείνο το πρότυπο ταιριάζει είτε ένα «0», είτε ένα «1». Σύμφωνα με τον Holland, η τάξη ενός σχήματος ($o(H)$) είναι ίση με τον αριθμό των σταθερών θέσεων, δηλαδή των μη μετα-χαρακτήρων, και το μήκος του σχήματος ($L(H)$) είναι ίσο με το συνολικό αριθμό των χαρακτήρων του. Έτσι το σχήμα #00#0 είναι σχήμα τάξης 3 ($o(H)=3$) και έχει μήκος 5 ($L(H)=5$) [17].

3.1.3 Τεχνητά Νευρωνικά δίκτυα

Τα Τεχνητά Νευρωνικά Δίκτυα (ΤΝΔ) μπορούν να μοντελοποιήσουν, με πολύ μεγάλη ακρίβεια, ένα πρόβλημα αξιοποιώντας τα δεδομένα ενός συνόλου μάθησης. Το μοντέλο αυτό χρησιμοποιείται για την πρόβλεψη των εξόδων για οποιεσδήποτε τιμές εισόδου [16]. Τα νευρωνικά δίκτυα εκτελούν υπό-συμβολική επεξεργασία πληροφορίας, η οποία

βασίζεται στη λειτουργία του ανθρώπινου εγκεφάλου και επικαλούνται την ιδέα της μοντελοποίησης “μαύρου κουτιού” (black box) χρησιμοποιώντας μοντέλα του ανθρώπινου εγκεφάλου, τα οποία εμπνέονται από τη βιολογία και τη νευροφυσιολογία. Για τη χρήση των μοντέλων αυτών δεν απαιτείται ρητή γνώση, σε αντίθεση με ότι ισχύει κατά την εφαρμογή συμβολικών μεθόδων της τεχνητής νοημοσύνης, οι οποίες στηρίζονται στη λογική. Στα συμβολικά συστήματα της τεχνητής νοημοσύνης η γνώση παριστάνεται ρητά, π.χ. με κανόνες παραγωγής. Στην υποσυμβολική προσέγγιση δεν δίνεται η υπό εξέταση σχέση ρητά, αλλά κωδικοποιείται στη δομή ενός νευρωνικού δικτύου [14].

Η εργασία αυτή ασχολείται με την εφαρμογή των ΤΝΔ για την εκτίμηση συναρτήσεων που αφορούν τη διάσπαση των μονωτήρων υπό συνθήκες ρύπανσης. Για το λόγο θα γίνει παρακάτω μια εκτενής αναφορά στις βασικές αρχές των ΤΝΔ, στις διάφορες αρχιτεκτονικές τους και στους αλγόριθμους μάθησης και διόρθωσης σφάλματος.

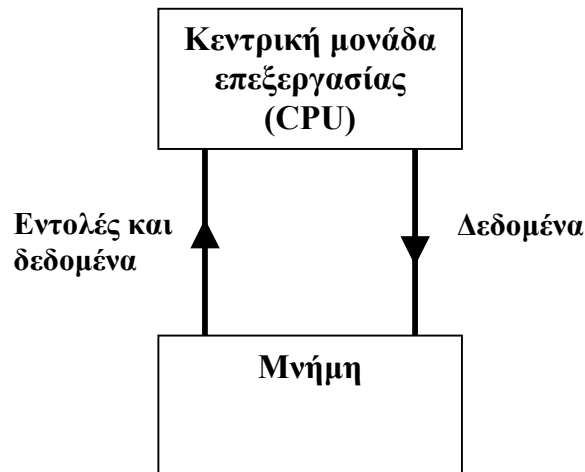
3.2 Θεωρία Τεχνητών Νευρωνικών Δικτύων

3.2.1 Ένας πρώτος ορισμός του Τεχνητού Νευρωνικού Δικτύου

Για να ξεκινήσουμε, θα ήταν χρήσιμο να δώσουμε έναν ορισμό για το τι εννοούμε με τον όρο “Τεχνητό Νευρωνικό Δίκτυο”. Ένα ΤΝΔ είναι ένα πλήθος απλών διασυνδεδεμένων μονάδων ή κόμβων, των οποίων η λειτουργία βασίζεται εν μέρει στους βιολογικούς νευρώνες. Η υπολογιστική ικανότητα του δικτύου είναι αποθηκευμένη στα βάρη που έχουν οι συνδέσεις των δομικών μονάδων και στα οποία (βάρη) καταλήγουμε μετά από μια διαδικασία προσαρμογής ή αλλιώς *μάθησης* από ένα σύνολο εκπαίδευσης. Για να δούμε πόσο διαφορετικό είναι κάτι τέτοιο από την επεξεργασία που γίνεται από τους συμβατικούς υπολογιστές αξίζει να εξετάσουμε τις θεμελιώδεις αρχές όλων αυτών των μηχανών.

Η μηχανή von Neumann και το συμβολικό παράδειγμα

Η λειτουργία όλων των συμβατικών υπολογιστών μπορεί να μοντελοποιηθεί με τον ακόλουθο τρόπο [18]:



Σχήμα 3.1: Η μηχανή von Neumann

Ο υπολογιστής εκτελεί επαναλαμβανόμενα τον ακόλουθο κύκλο γεγονότων:

1. Διαβάζει μια εντολή από τη μνήμη
2. Διαβάζει από τη μνήμη όσα δεδομένα χρειάζονται για την εντολή
3. Εκτελεί την εντολή
4. Αποθηκεύει τα αποτελέσματα στη μνήμη
5. Επιστρέφει στο βήμα 1

Ποια προβλήματα μπορεί αυτή η μέθοδος να λύσει εύκολα; Είναι δυνατό πολλά προβλήματα να τυποποιηθούν μέσω ενός αλγορίθμου, που είναι μια καλά ορισμένη διαδικασία που εγγυάται την ύπαρξη λύσης. Για παράδειγμα, ένας αλγόριθμος μπορεί να δίνει τη λύση σε ένα σύνολο εξισώσεων ή τον τρόπο αναζήτησης ενός αντικειμένου σε μια βάση δεδομένων. Αυτός ο αλγόριθμος μπορεί στη συνέχεια να αναλυθεί σε ένα σύνολο απλούστερων «προτάσεων», οι οποίες μπορούν διαδοχικά να απλοποιηθούν στις εντολές που τελικά θα εκτελέσει η Κεντρική Μονάδα Επεξεργασίας.

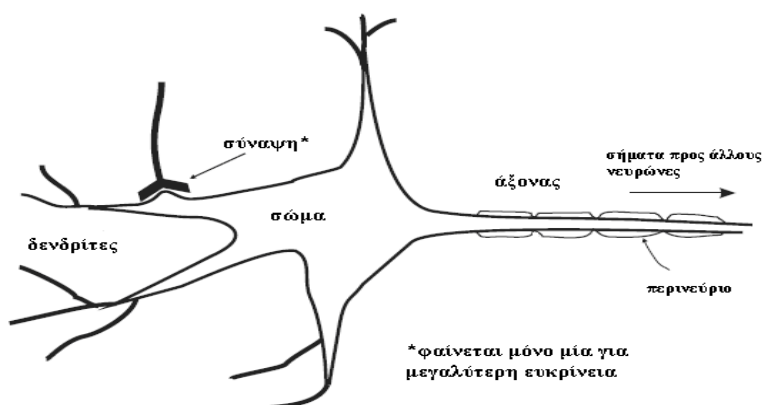
Παραθέτουμε κάποια από τα βασικά χαρακτηριστικά αυτών των μηχανών, ώστε να τα συγκρίνουμε στη συνέχεια με αυτά των δικτύων [18]:

- Ο υπολογιστής θα πρέπει να “γνωρίζει” εκ των προτέρων και με μεγάλη λεπτομέρεια την ακριβή ακολουθία βημάτων που χρειάζονται για να εκτελεστεί ο αλγόριθμος. Αυτή η ακολουθία βημάτων καλείται *πρόγραμμα*.
- Τα δεδομένα με τα οποία ασχολείται ο υπολογιστής πρέπει να έχουν συγκεκριμένη μορφή – αν κάτι τέτοιο δεν συμβαίνει, τα δεδομένα “μπερδεύουν” τη μηχανή.

- Το υλικό (hardware) του υπολογιστή φθείρεται εύκολα – αν καταστραφούν κάποιες θέσεις-κλειδιά στη μνήμη, η μηχανή θα σταματήσει να λειτουργεί.
- Υπάρχει μια καθαρή αντιστοιχία μεταξύ των σημασιολογικών αντικειμένων που χειρίζεται ο υπολογιστής (αριθμοί, λέξεις, κ.λ.π.) και του hardware του. Κάθε αντικείμενο μπορεί να αντιστοιχηθεί σε ένα μπλοκ της μνήμης του υπολογιστή.

Η επιτυχία της συμβολικής προσέγγισης στην Τεχνητή Νοημοσύνη εξαρτάται άμεσα από τις συνέπειες του πρώτου από τα παραπάνω σημεία, στο οποίο υποθέσαμε ότι μπορούμε να βρούμε έναν αλγόριθμο για να περιγράψουμε τη λύση του προβλήματος. Όμως αποδεικνύεται ότι πολλές από τις καθημερινές δραστηριότητες, που εμείς θεωρούμε δεδομένες, είναι δύσκολο να τυποποιηθούν με αυτόν τον τρόπο. Για παράδειγμα, η οπτική (ή ακουστική) αναγνώριση πραγμάτων στο περιβάλλον μας: πώς αναγνωρίζουμε χειρόγραφους χαρακτήρες, των οποίων δείγματα βλέπουμε για πρώτη φορά, ή το πρόσωπο κάποιου όταν το βλέπουμε από μια νέα οπτική γωνία; Πώς θυμόμαστε ολόκληρα περιστατικά όταν μας δίνονται απλώς κάποια ασαφή στοιχεία; Οι τεχνικές που χρησιμοποιούνται στις συμβατικές βάσεις δεδομένων είναι πολύ ελλιπείς για να αντιμετωπίσουν τη μεγάλη ποικιλία των συνειρμών που μπορεί να κάνει ο άνθρωπος. Ο τρόπος να αντιμετωπίσουμε αυτές τις δυσκολίες είναι να προσπαθήσουμε να αντιγράψουμε τη φυσική αρχιτεκτονική του εγκεφάλου, για να μιμηθούμε την εγκεφαλική λειτουργία με μεγαλύτερη ακρίβεια.

3.2.2 Ο Βιολογικός Νευρώνας



Σχήμα 3.2: Ο Βιολογικός Νευρώνας [18]

Η λειτουργία του βιολογικού συστήματος βασίζεται στη διασύνδεση εξειδικευμένων φυσικών κυττάρων, που ονομάζονται *νευρώνες*. Οι σημαντικές ιδιότητες των βιολογικών συστημάτων, κάποιες από τις οποίες αναφέραμε με παραδείγματα πιο πάνω, είναι η προσαρμοστικότητα, η ικανότητα αναγνώρισης από τα συμφραζόμενα, η ανοχή στα λάθη, η μεγάλη χωρητικότητα μνήμης και η ικανότητα επεξεργασίας βιολογικών πληροφοριών σε πραγματικό χρόνο [19]. Βέβαια ο τρόπος με τον οποίο λειτουργεί ο ανθρώπινος εγκέφαλος δεν είναι ακόμη επαρκώς γνωστός, πολλά πράγματα για το πώς ο εγκέφαλος εκπαιδεύεται για να επεξεργάζεται τις πληροφορίες είναι ακόμα άγνωστα. Επιπλέον, παρά το γεγονός ότι το βασικό στοιχείο υπολογισμού του ανθρώπινου συστήματος επεξεργασίας είναι σχετικά απλό (σε σχέση με τα ηλεκτρονικά στοιχεία), η συνολική επεξεργασία επιτυγχάνεται σε μερικές εκατοντάδες milliseconds. Αυτό μας οδηγεί στο συμπέρασμα ότι η βάση του βιολογικού υπολογισμού είναι ένας μικρός αριθμός ακολουθιακών βημάτων, καθένα από τα οποία εκτελείται σε μεγάλο παραλληλισμό. Επιπλέον στην έμφυτη αυτή παράλληλη αρχιτεκτονική κάθε μονάδα επεξεργασίας είναι σχετικά απλή και τοπικά συνδεδεμένη. Στο Σχήμα 3.2 παρουσιάζεται η βασική δομή ενός βιολογικού νευρώνα, ο οποίος αποτελείται από το *σώμα* (cell body), τον *άξονα* (axon), τους *δενδρίτες* (dendrites) και τις *συνάψεις* (synapses).

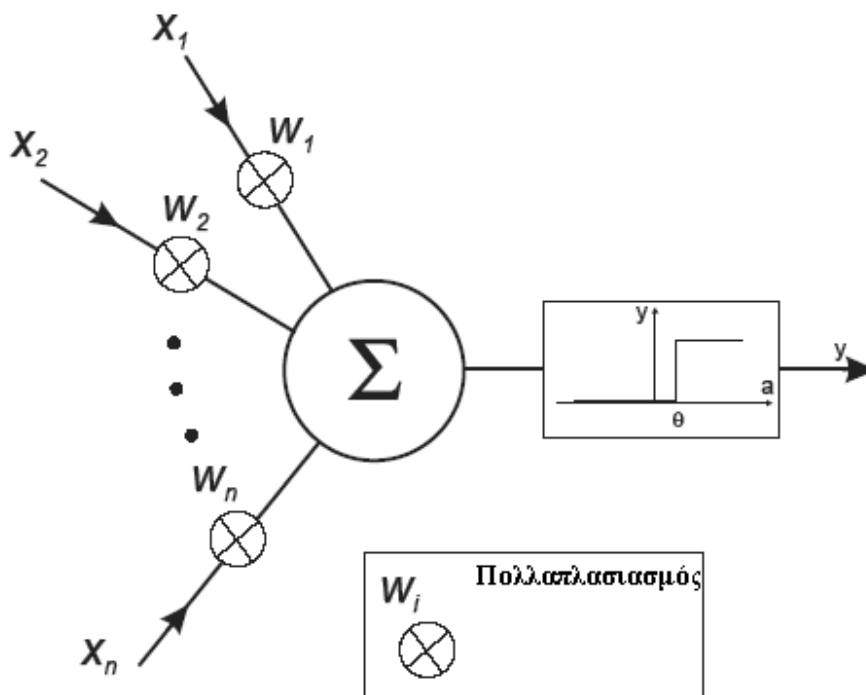
Η βασική λειτουργία που επιτελεί ένας βιολογικός νευρώνας είναι η συσσώρευση των σημάτων που δέχεται από τους νευρώνες με τους οποίους συνδέεται η είσοδός του, το φιλτράρισμα και η ενίσχυση αυτών των σημάτων και η παραγωγή ενός σήματος εξόδου, το οποίο στη συνέχεια μεταδίδεται μέσω των συνάψεων προς τους νευρώνες με τους οποίους συνδέεται η έξοδός του [19]. Αναλυτικότερα, στον ανθρώπινο εγκέφαλο ένας τυπικός νευρώνας συλλέγει σήματα από άλλους νευρώνες μέσω ενός πλήθους λεπτών δομών, τους *δενδρίτες*. Η ηλεκτρική δραστηριότητα, ή αλλιώς τα λεγόμενα δυναμικά δράσης, στέλνεται μέσω του *άξονα*, ο οποίος χωρίζεται σε χιλιάδες διακλαδώσεις. Στο τέλος κάθε διακλάδωσης, μια δομή που ονομάζεται *σύναψη* μετατρέπει το σήμα από τον *άξονα* σε ηλεκτρικό αποτέλεσμα (post synaptic potential – PSP). Κάποια απ' τα PSP αποτρέπουν τη μετάδοση του δυναμικού δράσης στους γειτονικούς νευρώνες που συνδέονται με τον αρχικό, ενώ άλλα την επιτρέπουν. Ο νευρώνας προσθέτει τα αποτελέσματα χιλιάδων τέτοιων PSP στις *δενδριτικές* του διακλαδώσεις σε κάποιο χρονικό διάστημα. Αν το άθροισμα των σημάτων στη βάση του *άξονα* υπερβεί ένα

κατώφλι, το κύτταρο εξαπολύει ένα καινούριο δυναμικό δράσης στο επόμενο κύτταρο και επαναλαμβάνεται η ίδια ακολουθία γεγονότων [18].

Η μάθηση λαμβάνει χώρα μέσω της αλλαγής στην αποτελεσματικότητα των συνάψεων, οπότε αλλάζει και η επίδρασή του ενός νευρώνα στον άλλο.

3.2.3 Τεχνητοί Νευρώνες

Η αντιστοιχία ενός τεχνητού νευρώνα με το απλό μοντέλο του βιολογικού νευρώνα που είδαμε παραπάνω είναι περιληπτικά η εξής: Τα σήματα (δυναμικά δράσης) εμφανίζονται στις εισόδους της μονάδας (συνάψεις). Το αποτέλεσμα κάθε σήματος (PSP) προκύπτει πολλαπλασιάζοντας το σήμα με κάποιον αριθμό (ή βάρος) που συμβολίζει το πόσο σημαντική είναι αυτή η σύναψη. Τα σταθμισμένα αυτά σήματα στη συνέχεια προστίθενται για να παράγουν τη συνολική ενεργοποίηση για τη μονάδα. Αν αυτή η ενεργοποίηση υπερβεί ένα συγκεκριμένο κατώφλι, η μονάδα παράγει μια έξοδο. Αυτή η λειτουργία αποτελεί και την αρχή λειτουργία του Τεχνητού Νευρώνα που προτάθηκε αρχικά από τους McCulloch και Pitts.



Σχήμα 3.3: Τεχνητός Νευρώνας

Υποθέτουμε ότι έχουμε n εισόδους, με σήματα x_1, x_2, \dots, x_n και βάρη w_1, w_2, \dots, w_n . Τα σήματα παίρνουν τις τιμές '1' ή '0' μόνο. Αυτό σημαίνει ότι τα σήματα παίρνουν τιμές Boolean. (Αυτό μας επιτρέπει να συζητήσουμε και τη σχέση τους με τα λογικά ψηφιακά κυκλώματα). Η ενεργοποίηση a δίνεται απ' τη σχέση:

$$a = w_1x_1 + w_2x_2 + \dots + w_nx_n \quad (3.3)$$

ή αλλιώς

$$a = \sum_{i=1}^n w_i x_i \quad (3.4)$$

Η έξοδος y στη συνέχεια υπολογίζεται ως εξής:

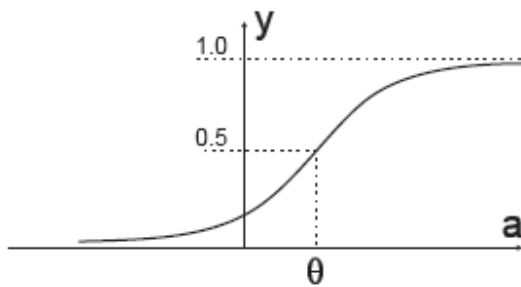
$$y = \begin{cases} 1, & \text{αν } a \geq \theta \\ 0, & \text{αν } a < \theta \end{cases} \quad (3.5)$$

Το κατώφλι θ συχνά είναι το μηδέν. Η συνάρτηση κατωφλίου είναι η λεγόμενη βηματική συνάρτηση. Αν θέλαμε να πάρουμε την αναλογία με τους πραγματικούς νευρώνες, θα λέγαμε ότι η παρουσία δυναμικού δράσης δηλώνεται με το δυαδικό '1', ενώ η απουσία δυναμικού δράσης με το δυαδικό '0'.

Μέχρι στιγμής, δεν έχουμε αναφέρει καθόλου τον παράγοντα χρόνο – η μονάδα αποκρίνεται ακαριαία στις εισόδους της, σε αντίθεση με τον πραγματικό νευρώνα. Αργότερα θα δούμε πώς αυτό μπορεί να βελτιωθεί.

3.2.4 Επικοινωνία σημάτων με μη δυαδικό τρόπο

Τα σήματα με τα οποία ασχοληθήκαμε ως τώρα (τόσο για τους πραγματικούς, όσο και για τους τεχνητούς νευρώνες) έπαιρναν μόνο δύο τιμές, ήταν δηλαδή δυαδικά σήματα. Όμως, είναι γενικά παραδεκτό ότι στους πραγματικούς νευρώνες, η πληροφορία κωδικοποιείται κυρίως σε σχέση με τη συχνότητα μετάδοσης του παλμού, παρά με την παρουσία ή απουσία αυτού. Αυτό στους τεχνητούς νευρώνες μπορεί να αναπαρασταθεί λαμβάνοντας σαν εύρος για τις τιμές εισόδους τους θετικούς πραγματικούς αριθμούς και «ομαλοποιώντας» τη βηματική συνάρτηση σε μια συνεχή συνάρτηση, τη λεγόμενη «σιγμοειδή».



Σχήμα 3.4: Σιγμοειδής συνάρτηση

Ένας εύχρηστος τύπος για τη σιγμοειδή συνάρτηση είναι ο:

$$y = f(a) \equiv \frac{1}{1 + e^{-a/\rho}} \quad (3.6)$$

Εδώ το ρ καθορίζει το σχήμα της καμπύλης: μια μεγάλη τιμή του ρ κάνει την καμπύλη πιο επίπεδη. Το κατώφλι αντιστοιχεί σ' εκείνη την τιμή του a για την οποία $y = 0,5$, άρα στη σχέση (3.6), το κατώφλι θ είναι το μηδέν. Αν δεν θέλουμε να έχει μηδενική τιμή, η συνάρτηση ενεργοποίησης γίνεται:

$$y = f(a) \equiv \frac{1}{1 + e^{-(a-\theta)/\rho}} \quad (3.7).$$

3.2.5 Χαρακτηριστικά ΤΝΔ – Σύγκριση με συμβατικούς υπολογιστές

Έχοντας δει τα συστατικά στοιχεία των Τεχνητών Νευρωνικών Δικτύων, μπορούμε να συγκρίνουμε τα χαρακτηριστικά τους με το παράδειγμα της μηχανής von Neumann.

- Προφανώς, ο τρόπος επεξεργασίας είναι εντελώς διαφορετικός – έχει ομοιότητες με την επεξεργασία σημάτων μάλλον, παρά με την επεξεργασία συμβόλων. Στα ΤΝΔ γίνεται συνδυασμός κάποιων σημάτων για την παραγωγή κάποιων άλλων, σε αντίθεση με τους υπολογιστές, όπου εκτελούνται εντολές αποθηκευμένες στη μνήμη.
- Η πληροφορία αποθηκεύεται σε ένα σύνολο βαρών κι όχι σε ένα πρόγραμμα. Τα βάρη αυτά προσαρμόζονται κατά την εκπαίδευση του ΤΝΔ με παραδείγματα.
- Τα ΤΝΔ παρουσιάζουν ευστάθεια: μικρές μεταβολές σε ένα σήμα εισόδου δεν επιδρούν δραστικά στην έξοδο.

- Μια μερική καταστροφή του δικτύου οδηγεί σε ανάλογη βλάβη της λειτουργίας του. Εντούτοις, κάποιες ικανότητες του δικτύου μπορούν να διατηρηθούν ακόμα κι αν το δίκτυο υποστεί σοβαρές βλάβες [20].
- Αυτό-οργάνωση: Ένα ΤΝΔ μπορεί να δημιουργήσει τη δική του οργάνωση ή αναπαράσταση της πληροφορίας που δέχεται κατά το χρόνο μάθησης [20].
- Λειτουργία πραγματικού χρόνου (real time operation): Οι υπολογισμοί του ΤΝΔ μπορούν να γίνονται παράλληλα ενώ σχεδιάζονται και κατασκευάζονται ειδικές συσκευές που εκμεταλλεύονται αυτήν την ικανότητα [20].
- Το ΤΝΔ μπορεί να χειριστεί και πρότυπα που βλέπει για πρώτη φορά, ενώ έχει και τη δυνατότητα της γενίκευσης από το σύνολο εκπαίδευσης [18].

Από την άλλη, τα ΤΝΔ εμφανίζουν και ανεπιθύμητα χαρακτηριστικά. Τα βασικότερα απ' αυτά είναι [19]:

- Δεν υπάρχουν σαφείς κανόνες για την ανάπτυξη ΤΝΔ για οποιαδήποτε εφαρμογή.
- Δεν υπάρχει γενικός τρόπος για την ερμηνεία της εσωτερικής λειτουργίας του δικτύου.
- Η εκπαίδευση μπορεί να είναι δύσκολη ή αδύνατη
- Η ικανότητα γενίκευσης είναι δύσκολα προβλέψιμη.

3.2.6 Διαδικασία Κατασκευής ΤΝΔ

Η ανάπτυξη συστηματικής μεθοδολογίας για την κατασκευή νευρωνικών δικτύων για την επίλυση κάποιου προβλήματος δεν έχει ακόμα επιτευχθεί. Βασικά ερωτήματα που προκύπτουν είναι [17]:

- ✓ Είναι οι τεχνικές των ΤΝΔ κατάλληλες ή εφαρμόσιμες στο πρόβλημα; Το πρόβλημα έχει μία ή περισσότερες λύσεις;
- ✓ Μπορούμε να τροποποιήσουμε γνωστά ΤΝΔ για να λύσουμε το πρόβλημα;
- ✓ Υπάρχουν τρόποι ανάλυσης του προβλήματος (π.χ. ποια η πολυπλοκότητά του);

Η εφαρμογή της τεχνολογίας νευρωνικών δικτύων για την αντιμετώπιση κάποιου προβλήματος απαιτεί τον καθορισμό των στοιχείων του ΤΝΔ, όπως της αρχιτεκτονικής, της τοπολογίας, των παραμέτρων των μονάδων και της διαδικασίας εκπαίδευσης. Αν και φαίνεται απλό, χρειάζεται εντούτοις αρκετή τεχνική κρίση. Υπάρχουν άπειροι

συνδυασμοί παραμέτρων που είναι αδύνατο να δοκιμασθούν όλοι. Επιπλέον πρέπει να εξεταστεί η καταλληλότητα της νευρωνικής λύσης.

Κατά τη διάρκεια της ανάπτυξης λύσεων βασισμένων σε ΤΝΔ προκύπτουν πολλά ερωτήματα, όπως:

- ✓ Είναι δυνατόν το ΤΝΔ να εκπαιδευτεί, ώστε να εκτελέσει την επιθυμητή λειτουργία;
- ✓ Υποθέτοντας ότι υπάρχει λύση, ποιες είναι οι παράμετροι του δικτύου;
- ✓ Ποιοι υπολογιστικοί πόροι είναι απαραίτητοι;

Σε πραγματικές εφαρμογές η σχεδίαση του τεχνητού νευρωνικού δικτύου είναι πολύπλοκη και συνήθως γίνεται αλληλεπιδραστικά και επαναληπτικά. Παρότι είναι αδύνατο να παρέχουμε έναν ολοκληρωμένο αλγόριθμο, στη συνέχεια παρουσιάζεται ένας βασικός σκελετός βημάτων που αντανakλά τα βασικά στάδια υλοποίησης. Η πληθώρα των παραμέτρων σχεδίασης περιλαμβάνει:

- α. τοπολογία δικτύου και στρατηγική διασύνδεσης των μονάδων,
- β. χαρακτηριστικά των μονάδων, τα οποία μπορούν να διαφέρουν ανάλογα με τη θέση τους,
- γ. διαδικασία εκπαίδευσης,
- δ. σύνολα εκπαίδευσης και ελέγχου,
- ε. αναπαράσταση εισόδου / εξόδου, προεπεξεργασία και μετεπεξεργασία.

3.2.7 Κατηγοριοποίηση των ΤΝΔ

Οποιαδήποτε περιγραφή ενός τεχνητού νευρωνικού δικτύου ξεκινάει με τον προσδιορισμό των εξής χαρακτηριστικών:

- Τοπολογία δικτύου
- Χαρακτηριστικά μονάδων
- Λειτουργικότητα του δικτύου

Μία προσέγγιση στο διαχωρισμό των ΤΝΔ προκύπτει από την επιθυμητή συμπεριφορά που θέλουμε αυτό να παρουσιάσει. Για παράδειγμα, η επιθυμητή λειτουργία ενός ΤΝΔ μπορεί να καθοριστεί με απαρίθμηση των καταστάσεων του δικτύου ή προσδιορίζοντας την επιθυμητή έξοδο βάσει των εισόδων και της τρέχουσας κατάστασης. Τα ΤΝΔ χωρίζονται στις παρακάτω κατηγορίες:

- *Συσχετιστές προτύπων*: Η λειτουργία αυτού του δικτύου είναι η συσχέτιση προτύπων και η υλοποίηση επιθυμητών απεικονίσεων εισόδου-εξόδου. Συνήθως υλοποιούνται με δίκτυα προσοτροφοδότησης.
- *Μοντέλο μνήμης συμφραζομένων*: Αντιπροσωπευτικό παράδειγμα είναι το δίκτυο *Hopfield*.
- *Αυτοοργανούμενα (self-organized) δίκτυα*: Στην κατηγορία αυτή ανήκουν δίκτυα με ικανότητα μάθησης χωρίς επίβλεψη, τα οποία δηλαδή κατηγοριοποιούν την είσοδο σύμφωνα με κάποια κριτήρια ομοιότητας.

Θεωρώντας την τοπολογία και τη δομή των ΤΝΔ μπορούμε να τα διαχωρίσουμε ως εξής:

- επαναληπτικά δίκτυα, δηλαδή με συνδέσεις ανάδρασης,
- δίκτυα πρόσθιας τροφοδότησης,
- δίκτυα με δομή επιπέδων ή ιεραρχική,
- δίκτυα με ανταγωνιστικές συνδέσεις.

Μέχρι τώρα θεωρήσαμε ότι ο σχεδιαστής του δικτύου πρέπει να επιλέξει μία τοπολογία δικτύου. Πρόσφατες έρευνες προτείνουν ένα ακόμη στάδιο εκπαίδευσης, το οποίο θα περιλαμβάνει συνεργασία και συνδυασμό διαφόρων τοπολογιών σε μία εφαρμογή. Παρότι η έρευνα είναι ακόμη σε αρχικό στάδιο, η ιδέα είναι η χρησιμοποίηση συνδυασμών εκπαιδευμένων επιμέρους δικτύων [17].

Στη συνέχεια θα αναλύσουμε λίγο περισσότερο το βασικό διαχωρισμό των τεχνητών νευρωνικών δικτύων: με επίβλεψη και χωρίς επίβλεψη.

Κεφάλαιο 4

Δίκτυα με επίβλεψη

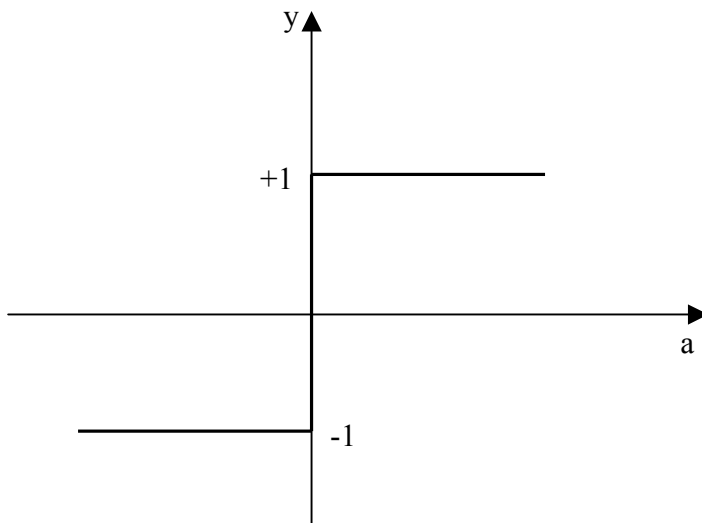
4.1 Γενικά

Τα ονομαζόμενα Τεχνητά Νευρωνικά Δίκτυα επιβλεπόμενης μάθησης ή αλλιώς απεικόνισης είναι γνωστά με αυτήν την ονομασία επειδή προσαρμόζουν (μορφοποιούν) τις απεικονίσεις εισόδου – εξόδου σύμφωνα με ένα δεδομένο σύνολο ζευγών εκπαίδευσης εισόδου – εξόδου [14]. Τέτοιου είδους δίκτυα είναι τα *δίκτυα προσοτροφοδότησης*, τα *δίκτυα ακτινικών συναρτήσεων βάσης* και τα *δίκτυα συναρτήσεων B-spline*. Βάση για τα ΤΝΔ προσοτροφοδότησης αποτελεί το μονοστρωματικό δίκτυο προσοτροφοδότησης (γνωστό και ως μονοστρωματικό Perceptron).

4.2 Το Perceptron

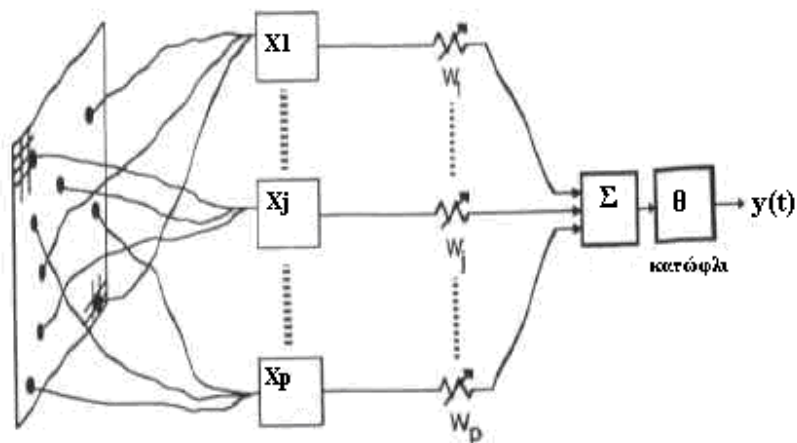
Το ΤΝΔ Perceptron αναπτύχθηκε από τον Roseblatt (1958) και αποτελεί μια από τις πρώτες προσπάθειες σχεδίασης και κατασκευής ευφύων συστημάτων με δυνατότητες αυτομάθησης. Το Perceptron είναι το απλούστερο ΤΝΔ που χρησιμοποιείται για την ταξινόμηση γραμμικά διαχωρίσιμων προτύπων, δηλαδή προτύπων τα οποία διαχωρίζονται από ένα υπερεπίπεδο. Αποτελείται από ένα απλό νευρώνα, σαν κι αυτόν που είδαμε προηγουμένως, με προσαρμόσιμα βάρη, ο οποίος ακολουθείται από μια διπολική συνάρτηση ενεργοποίησης της μορφής:

$$y = f(a) \equiv \begin{cases} -1, & a < 0 \\ 0, & a = 0 \\ +1, & a > 0 \end{cases} \quad (4.1)$$



Σχήμα 4.1: Συνάρτηση ενεργοποίησης του Perceptron

Η δομή του Perceptron (ενός νευρώνα) φαίνεται στο Σχήμα 4.2. Αυτό το δίκτυο μπορεί να πραγματοποιήσει ταξινόμηση στην περίπτωση που έχουμε μόνο δύο κατηγορίες. Για να είναι δυνατή η ταξινόμηση περισσότερων γραμμικά διαχωρίσιμων κατηγοριών (κλάσεων), το στρώμα εξόδου χρειάζεται να έχει περισσότερους από έναν νευρώνες.



Σχήμα 4.2: Το Perceptron

Η έξοδος του Perceptron δίνεται από τη σχέση:

$$y(t) = f_h(a(t)) \quad (4.2)$$

με

$$a(t) = \sum_{i=0}^p w_i x_i = \mathbf{w}^T(t) \mathbf{x}(t) \quad (4.3)$$

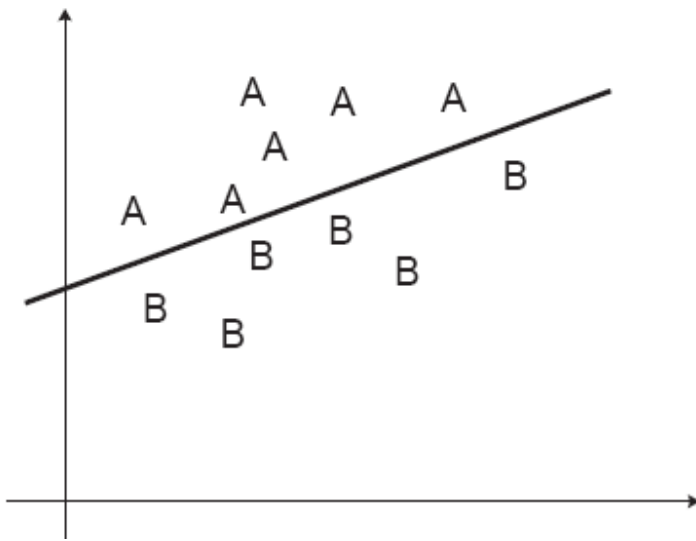
και

$$\mathbf{x}(t) = [-1, x_1(t), x_2(t), \dots, x_p(t)]^T \quad (4.4)$$

$$\mathbf{w}(t) = [\theta(t), w_1(t), w_2(t), \dots, w_p(t)]^T \quad (4.5)$$

(το κατώφλι θ μεταθέτει το σύνορο διαχωρισμού)

Ο σκοπός του Perceptron είναι να ταξινομήσει τα εξωτερικά πρότυπα x_1, x_2, \dots, x_n σε μία από δύο κατηγορίες A ή B. Αυτό γίνεται εκχωρώντας ένα σημείο $[x_1, x_2, \dots, x_p]^T$ στην κατηγορία A, αν η έξοδος y είναι +1 και στην κατηγορία B, αν η έξοδος y είναι -1.



Σχήμα 4.3: Γραμμικός διαχωρισμός δύο κλάσεων A και B στην περίπτωση δύο διαστάσεων

Η διδασκαλία του Perceptron πρέπει να γίνει με καλά παραδείγματα κάθε μιας κατηγορίας. Εάν οι προς διαχωρισμό κατηγορίες μοιάζουν πολύ μεταξύ τους, τότε είναι δύσκολο να εκπαιδευτεί το perceptron να ταξινομεί σωστά τα πρότυπα που του παρουσιάζονται στη μία ή στην άλλη κατηγορία.

Για να κατανοήσουμε πώς λειτουργεί ένας ταξινομητής προτύπων, θεωρούμε την περίπτωση δύο μεταβλητών x_1 και x_2 , όπου το σύνορο απόφασης (διαχωρισμού ανάμεσα στις κατηγορίες A και B) είναι μια ευθεία γραμμή (βλ. Σχήμα 4.3), ο τύπος της οποίας είναι: $w_1x_1 + w_2x_2 - \theta = 0$. Οποιοδήποτε σημείο βρίσκεται πάνω από τη διαχωριστική γραμμή ταξινομείται στην κατηγορία A, διαφορετικά αποδίδεται στην κατηγορία B. Ο αλγόριθμος επιλογής (ανανέωσης) των βαρών (αλγόριθμος perceptron) στηρίζεται στο

γεγονός ότι, εάν οι κλάσεις A και B είναι γραμμικά διαχωρίσιμες, τότε υπάρχει ένα διάνυσμα βαρών w τέτοιο ώστε:

$$w^T x \geq 0 \text{ όταν το } x \text{ ανήκει στην κατηγορία A}$$

$$w^T x < 0 \text{ όταν το } x \text{ ανήκει στην κατηγορία B.}$$

Συνεπώς, όταν το perceptron λαμβάνει ένα ζευγάρι εκπαίδευσης (x_1, x_2) πρέπει να προσδιορίσει ένα διάνυσμα βάρους w τέτοιο, ώστε να ικανοποιούνται οι δύο παραπάνω ανισότητες. Ο αλγόριθμος perceptron περιλαμβάνει τα ακόλουθα βήματα [14]:

Βήμα 1: Διάλεξε μία αρχική τιμή $w(0)$ του διανύσματος βάρους $w(t)$.

Βήμα 2: Παρουσίασε στο perceptron το διάνυσμα εισόδου x και την επιθυμητή έξοδο $d(t)$.

Βήμα 4: Υπολόγισε την προκύπτουσα έξοδο $y(t)$ με βάση τις (4.2) και (4.3):

$$y(t) = f_h(w^T(t)x(t)) = \text{sgn}\{w^T(t)x(t)\}, \text{ όπου } \text{sgn}\{\cdot\} \text{ είναι η συνάρτηση προσήμου:}$$

$$\text{sgn}(a) = \begin{cases} -1, & a < 0 \\ +1, & a \geq 0 \end{cases}$$

Βήμα 4: Εάν η πραγματική έξοδος είναι σωστή, δηλαδή εάν $w^T(t)x(t) \geq 0$ και το $x(t)$ ανήκει στην κατηγορία A ή $w^T(t)x(t) < 0$ και το $x(t)$ ανήκει στην κατηγορία B, τότε δεν γίνεται καμία διόρθωση του βάρους, δηλαδή $w(t+1) = w(t)$. Διαφορετικά γίνεται διόρθωση (ανανέωση) του $w(t)$ ως εξής:

$$w(t+1) = w(t) - \gamma x(t), \text{ εάν } w^T(t)x(t) \geq 0 \text{ και } x(t) \in B$$

ή

$$w(t+1) = w(t) + \gamma x(t) \text{ εάν } w^T(t)x(t) < 0 \text{ και } x(t) \in A.$$

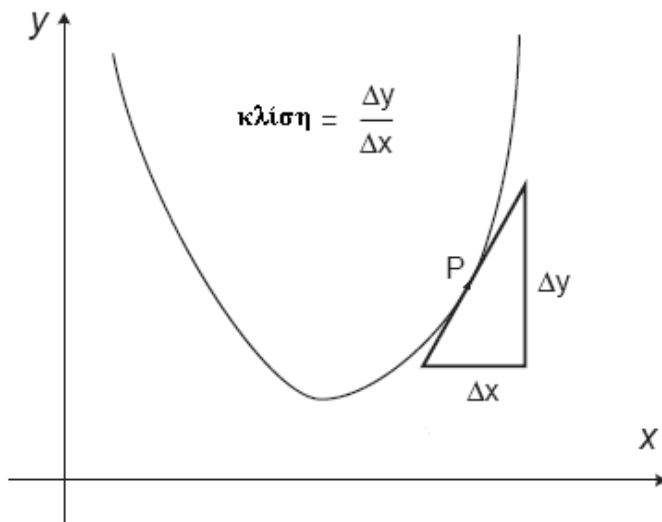
Όπου $\gamma > 0$ είναι η παράμετρος που καθορίζει το ρυθμό μάθησης (ανανέωσης). Ο Roseblatt απέδειξε ότι ο αλγόριθμος αυτός συγκλίνει [14].

Ένας άλλος αλγόριθμος δόθηκε από τους Widrow και Hoff και είναι γνωστός ως αλγόριθμος (ή κανόνας) δέλτα, από το συμβολισμό του λάθους $\delta(t) = d(t) - y(t)$ που χρησιμοποίησαν οι Widrow και Hoff.

4.2.1 Ελαχιστοποιώντας το σφάλμα: ο κανόνας δέλτα

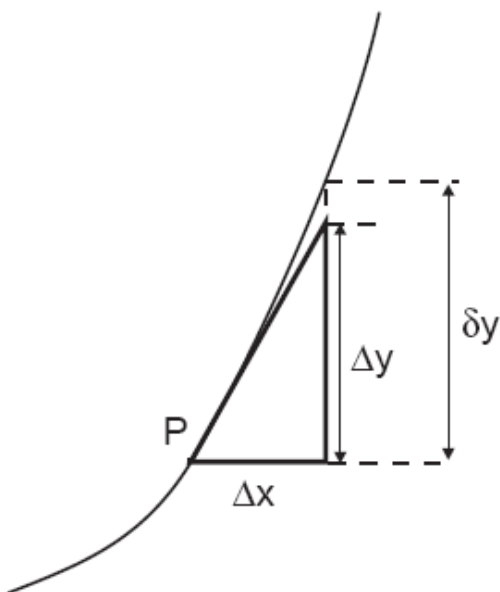
Ας υποθέσουμε ότι το y είναι κάποια συνάρτηση του x ($y = y(x)$), αλλά δεν ξέρουμε την

ακριβή μορφή αυτής της συνάρτησης. Επιπλέον, ας υποθέσουμε ότι θέλουμε να βρούμε τη θέση (συντεταγμένη x) της ελάχιστης τιμής της συνάρτησης και ότι μπορούμε να βρούμε την κλίση (το ρυθμό μεταβολής του y) σε κάθε σημείο. Η κλίση αυτή είναι ο λόγος $\Delta y / \Delta x$, όπως φαίνεται και στο διάγραμμα του Σχήματος 2.4.



Σχήμα 4.4

Όταν το Δx είναι μικρό, τότε το Δy είναι περίπου το ίδιο με τη μεταβολή δy στη συνάρτηση y για την ίδια μεταβολή Δx , όπως φαίνεται στο Σχήμα 4.5.



Σχήμα 4.5: Μικρές μεταβολές Δx

$$\text{Έτσι θα είναι: } \delta y \approx \Delta y = \frac{\Delta y}{\Delta x} \Delta x \quad (4.6)$$

$$\text{ώστε } \delta y \approx \text{slope} \times \Delta x \quad (4.7)$$

$$\text{Θέτουμε τώρα } \Delta x = -\alpha \times \text{κλίση} \quad (4.8)$$

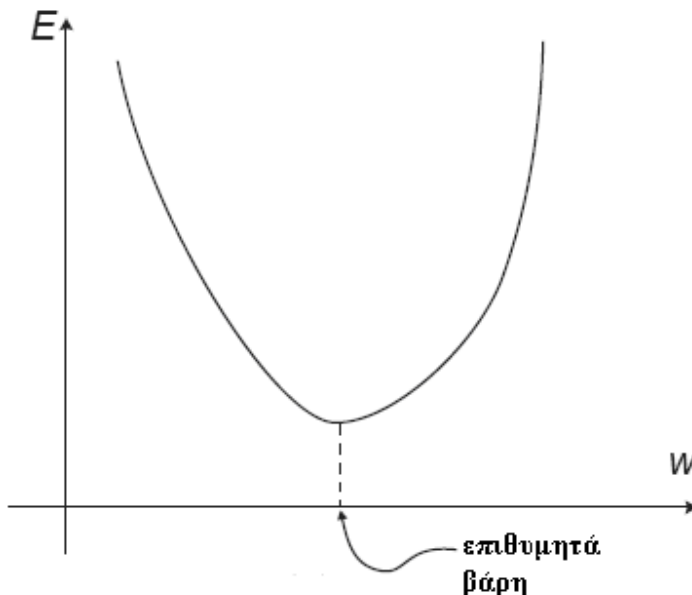
όπου $\alpha > 0$ και αρκετά μικρό, ώστε να εξασφαλίζει ότι $\delta y \approx \Delta y$. Άρα,

$$\delta y \approx -\alpha (\text{slope})^2 \quad (4.9).$$

Αυτό σημαίνει ότι $\delta y < 0$ και ότι έχουμε “κατηφορίσει” στην καμπύλη, προς το ελάχιστο σημείο της. Αν συνεχίσουμε να επαναλαμβάνουμε το βήμα (4.9), θα προσεγγίσουμε την τιμή του x που σχετίζεται με το ελάχιστο της συνάρτησης. Αυτή η τεχνική ονομάζεται *κανόνας του διανύσματος κλίσης (gradient descent)*. Το ερώτημα είναι πώς κάτι τέτοιο θα μπορούσε να χρησιμοποιηθεί στην εκπαίδευση ΤΝΔ [18].

Ο κανόνας του διανύσματος κλίσης για το σφάλμα

Η ιδέα είναι να μπορέσουμε να υπολογίσουμε το σφάλμα, κάθε φορά που στο δίκτυο παρουσιάζεται ένα διάνυσμα εκπαίδευσης (δεδομένου ότι έχουμε επιβλεπόμενη μάθηση, στην οποία δίνεται η έξοδος) και να εφαρμόσουμε τον κανόνα διανύσματος κλίσης στο σφάλμα αυτό θεωρώντας το συνάρτηση των βαρών. Για κάθε βάρος θα υπάρχει η αντίστοιχη εφαπτομένη στην καμπύλη του σφάλματος, όπως φαίνεται και στο Σχήμα 4.6.



Σχήμα 4.6: Ο κανόνας του διανύσματος κλίσης – Το σφάλμα (E) συναρτίζει των βαρών (w)

Τυπικά, για κάθε πρότυπο p , προσδιορίζουμε ένα σφάλμα E_p , το οποίο είναι συνάρτηση των βαρών, δηλαδή:

$$E_p = E_p(w_1, w_2, \dots, w_n) \quad (4.10).$$

Το σφάλμα αυτό ορίζεται ως το τετράγωνο της διαφοράς μεταξύ της πραγματικής εξόδου t και της εξόδου του δικτύου y . Έτσι (για ένα νευρώνα):

$$E_p = \frac{1}{2}(t - y)^2 \quad (4.11)$$

όπου θεωρούμε το y ως συνάρτηση των βαρών. Το συνολικό σφάλμα E είναι απλώς το άθροισμα όλων των E_p .

$$E = \sum_p E_p \quad (4.12)$$

Για να μπορέσει να εφαρμοστεί ο κανόνας του διανύσματος κλίσης, το σφάλμα θα πρέπει να είναι μια συνεχής και παραγωγίσιμη συνάρτηση των βαρών σε κάθε σημείο. Πολλές φορές βέβαια κάτι τέτοιο δεν συμβαίνει. Αν και η ενεργοποίηση είναι συνεχής συνάρτηση των βαρών, η έξοδος αλλάζει απότομα, καθώς η συνάρτηση ενεργοποίησης υπερβαίνει την τιμή κατωφλίου.

Μπορεί ναδειχθεί ότι η κλίση του E_p ως προς το w_j είναι $-(t-a)x_j$. Ο κανόνας μάθησης (κανόνας δέλτα) ορίζεται πλέον κάνοντας μια μεταβολή Δw_j στο βάρος σύμφωνα με την εξίσωση (4.8), δηλαδή:

$$\Delta w_j = -\alpha \times (\text{κλίση του } E_p \text{ ως προς το } w_j) = \alpha (t - a) x_j \quad (4.13)$$

Ο Widrow ονόμασε τους νευρώνες που χρησιμοποιούν τον κανόνα δέλτα ADALINES (ADaptive LINear Elements) και συνδέοντας πολλά απ' αυτά σχημάτισε τα Νευρωνικά Δίκτυα MADALINES (Many- ADALINES). Ο κανόνας μάθησης που βασίζεται στο διάνυσμα κλίσης με αυτούς τους τύπους νευρώνων είναι, επομένως, γνωστός ως κανόνας Widrow Hoff, αλλά συνηθέστερα πλέον ως κανόνας δέλτα.

Το μονοστρωματικό perceptron δεν μπορεί να ταξινομήσει πρότυπα που δεν είναι γραμμικά διαχωρίσιμα, δηλαδή πρότυπα για τα οποία δεν μπορούμε να βρούμε ένα γραμμικό σύνορο απόφασης (δηλ. δεν μπορούμε να φέρουμε μία ευθεία γραμμή) που να διαχωρίζει τις δύο κατηγορίες [14], γι' αυτό και αμφισβητήθηκε. Η πρώτη σημαντική κριτική του perceptron του Rosenblatt, παρουσιάστηκε από τους Minsky και Selfridge (1961), οι οποίοι υπογράμμισαν ότι το perceptron δεν μπορεί ούτε καν να υλοποιήσει τη

συνάρτηση της ισοτιμίας δύο bits, δηλαδή το πρόβλημα XOR κατά το οποίο η έξοδος είναι μονάδα, μόνο όταν μία και μόνο μία από τις δύο εισόδους είναι μονάδα, διαφορετικά ως έξοδο δίνει μηδέν. Τα υπολογιστικά όρια του perceptron του Rosenblatt καθορίστηκαν με αυστηρά μαθηματική διατύπωση στο διάσημο βιβλίο των Minsky και Papert “Perceptrons” (1969), όπου απέδειξαν με αυστηρή μαθηματική ανάλυση ότι το perceptron του Rosenblatt είναι έμφυτα ανίκανο να αντιμετωπίσει δύσκολα προβλήματα ταξινόμησης. Στο τελευταίο κεφάλαιο του βιβλίου τους υποστήριξαν ότι οι περιορισμοί της που είχαν ανακαλύψει για το perceptron θα έπρεπε λογικά να ισχύουν και για τις πολυεπίπεδες επεκτάσεις του. Ισχυρίστηκαν δε ότι: *«Το perceptron μας έχει προσφέρει πολύτιμη γνώση εξαιτίας των πολλών περιορισμών του. Έχει πολλά χαρακτηριστικά που είναι άξια προσοχής: τη γραμμικότητά του, το ενδιαφέρον θεώρημα μάθησης, την απλότητα του ως ένα μοντέλο παράλληλου υπολογισμού. Ωστόσο δεν υπάρχει κανένας λόγος να πιστέψουμε ότι κάποιο από αυτά τα προτερήματα του οδηγεί σε μια πολυεπίπεδη έκδοση. Ωστόσο, θεωρούμε ότι είναι σημαντικό ερευνητικό πρόβλημα η διαφώτιση (ή η απόρριψη) της διαισθητικής μας χρήσης, ότι η επέκταση σε πολυεπίπεδα συστήματα θα είναι άγονη.»*

Η παραπάνω εικασία δημιούργησε πολλές απορίες για τις υπολογιστικές ικανότητες όχι μόνο του perceptron, αλλά και των νευρωνικών δικτύων.

Εν τούτοις, η ιστορία έδειξε ότι η εικασία των Minsky και Papert δεν δικαιώνεται, αφού σήμερα έχουμε διάφορους τύπους νευρωνικών δικτύων, οι οποίοι είναι υπολογιστικά πολύ πιο ισχυροί από το perceptron του Rosenblatt. Χαρακτηριστικά παραδείγματα, το πολυεπίπεδο τεχνητό νευρωνικό δίκτυο που εκπαιδεύεται, με τον αλγόριθμο της προς τα πίσω διάδοσης σφάλματος και τα δίκτυα ακτινικών συναρτήσεων βάσης [19].

4.2.2 Πολυεπίπεδα δίκτυα

Τα πολυεπίπεδα perceptron περιλαμβάνουν, εκτός από το στρώμα κόμβων εισόδου και το στρώμα εξόδου, ένα ή περισσότερα στρώματα κρυμμένων κόμβων. Επίσης, είναι πρόσθιας τροφοδότησης, δηλαδή τροφοδοτούνται από μπροστινές εισόδους μόνο. Το πολυεπίπεδο νευρωνικό δίκτυο πρόσθιας τροφοδότησης χρησιμοποιούνται σε ποικίλες, αλλά και δύσκολες εφαρμογές, στις οποίες εκπαιδεύονται βάσει του αλγόριθμου της προς τα πίσω διάδοσης σφάλματος (Error Back-Propagation). Ο παραπάνω αλγόριθμος σχετίζεται με το γνωστό κανόνα της διόρθωσης σφάλματος. Η διαδικασία εκπαίδευσης

συνίσταται από δύο περάσματα σε όλα τα επίπεδα (στρώματα) του δικτύου που καλούνται ευθύ και ανάστροφο πέραςμα και για τα οποία θα αναφερθούμε πιο αναλυτικά παρακάτω.

Ένα πολυεπίπεδο Perceptron έχει κυρίως δύο χαρακτηριστικά :

1. Οι νευρώνες των κρυμμένων επιπέδων περιέχουν μία μη-γραμμική συνάρτηση ενεργοποίησης (activate function), η οποία είναι συνεχής και παραγωγίσιμη σε κάθε σημείο. Η εμφάνιση της μη γραμμικότητας είναι καθοριστική για τις μεγάλες υπολογιστικές δυνατότητες που μας παρέχουν τα πολυεπίπεδα νευρωνικά δίκτυα.
2. Δεν υπάρχει σύνδεση μεταξύ νευρώνων του ίδιου επιπέδου ή συνήθως μεταξύ νευρώνων που ανήκουν σε μη διαδοχικά επίπεδα. Αντίθετα στις πιο πολλές περιπτώσεις υπάρχει πλήρης διασύνδεση μεταξύ των νευρώνων διαδοχικών επιπέδων.

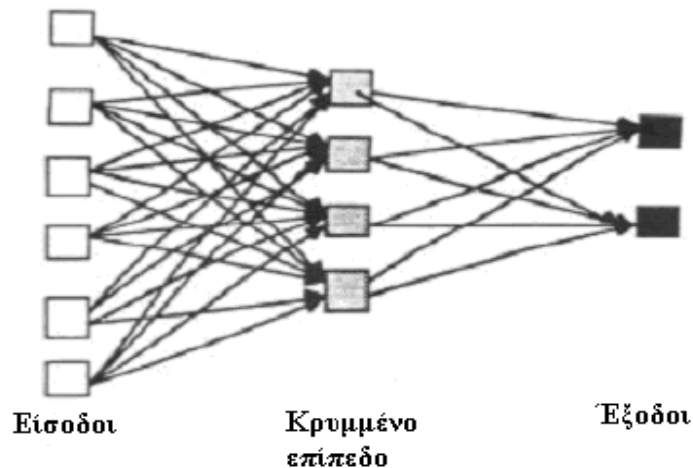
Η υπολογιστική ισχύς του πολυεπίπεδου perceptron προέρχεται από τον συνδυασμό αυτών των χαρακτηριστικών και από την ικανότητα μάθησης μέσω εκπαίδευσης. Αυτά τα χαρακτηριστικά είναι επίσης υπεύθυνα και για τις ελλείψεις που παρουσιάζουν οι γνώσεις μας γύρω από τη συμπεριφορά αυτών των δικτύων. Η παρουσία κατανεμημένης μορφής μη-γραμμικότητας και η μεγάλη συνδεσιμότητα καθιστούν υπερβολικά δύσκολη τη θεωρητική ανάλυση τους. Επίσης η χρήση κρυμμένων νευρώνων δυσκολεύει την κατανόηση και την εύκολη απεικόνιση της διαδικασίας μάθησης.

Το ερευνητικό ενδιαφέρον γι' αυτά τα δίκτυα ξεκίνησε με τη δουλειά του Rosenblatt πάνω στο perceptron και του Widrow πάνω στο δίκτυο Madaline το 1962. Αυτό που έλειπε από τα πρόσθια τροφοδοτούμενα πολυεπίπεδα δίκτυα αυτής της εποχής είναι αυτό που σήμερα ονομάζουμε μάθηση με τον *αλγόριθμο ανάστροφης διάδοσης σφάλματος* ή αλλιώς *error back-propagation*.

Ο όρος back-propagation εμφανίστηκε μετά το 1985, αλλά η ιδέα περιγράφηκε για πρώτη φορά από τον Werbos στη διδακτορική διατριβή του το 1974. Την περίοδο 1985-1986 χρησιμοποιήθηκε από τους Rumelhart, Hinton, Williams, McClelland, Parker και LeCun.

Η ανάπτυξη του αλγορίθμου back-propagation αποτελεί ορόσημο στην εξέλιξη των νευρωνικών δικτύων, διότι παρέχει μία υπολογιστικά αποτελεσματική μέθοδο εκπαίδευσης των πολυεπίπεδων perceptrons. Αν και δεν μπορούμε να ισχυριστούμε ότι ο αλγόριθμος αυτός μπορεί να επιλύσει όλα τα προβλήματα, πρέπει να διευκρινίσουμε ότι

εξάλειψε κάθε απαισιοδοξία που είχαν εισαγάγει για τα πολυεπίεδα perceptrons οι Minsky και Papert το 1969. Παρακάτω, φαίνεται η δομή ενός απλού ΝΔ πρόσθιας τροφοδότησης, με ένα κρυμμένο στρώμα.



Σχήμα 4.7: Νευρωνικό δίκτυο ενός κρυφού επιπέδου πλήρως διασυνδεδεμένο [20].

Στο Σχήμα 4.7 παριστάνεται η αρχιτεκτονική ενός πολυεπιπέδου τεχνητού νευρωνικού δικτύου ενός κρυφού επιπέδου έξι εισόδων και δύο εξόδων. Το δίκτυο είναι πλήρως διασυνδεδεμένο, δηλαδή κάθε νευρώνας συνδέεται με όλους τους νευρώνες του επόμενου επιπέδου. Η ροή όλων των σημάτων έχει κατεύθυνση από την είσοδο προς την έξοδο μέσω των επιπέδων του δικτύου. Τα σήματα διακρίνονται σε δύο ειδών :

- Τα σήματα λειτουργίας, τα οποία εμφανίζονται καταρχήν στην είσοδο του δικτύου ως ερέθισμα και προωθούνται κατά την ορθή φορά ανά επίπεδο, μέχρι να φτάσουν στο επίπεδο εξόδου, όπου παράγονται τα σήματα εξόδου. Ονομάζονται έτσι, καθώς κατά το πέρασμα τους από κάθε νευρώνα υλοποιείται ένας υπολογισμός συναρτήσεως των αντίστοιχων βαρών.
- Τα σήματα σφάλματος που δημιουργούνται μονάχα κατά τη διάρκεια της εκπαίδευσης και προωθούνται κατά την αντίστροφη φορά, από την έξοδο προς την είσοδο, διαμέσου του νευρωνικού δικτύου.

4.3 Ο αλγόριθμος Backpropagation

Παρουσιάσαμε προηγουμένως τον κανόνα δέλτα για την περίπτωση ενός νευρώνα. Για περισσότερους από ένα νευρώνες στο στρώμα εξόδου (έστω N), το σφάλμα θα πρέπει να

προστίθεται για όλους τους κόμβους:

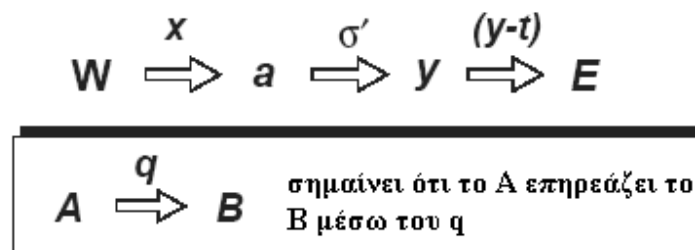
$$E_p = \frac{1}{2} \sum_{j=1}^N (t^j - y^j)^2 \quad (4.14)$$

Η ιδέα παραμένει η ίδια, η εφαρμογή δηλαδή του κανόνα διανύσματος κλίσης στο σφάλμα θεωρώντας το συνάρτηση των βαρών. Αυτή τη φορά ωστόσο θα πρέπει να λάβουμε υπόψη όλα τα βάρη, τόσο για τους κόμβους του *κρυμμένου* επιπέδου, όσο και για τους κόμβους εξόδου. Οι πρώτοι είναι κόμβοι (νευρώνες) του ενδιάμεσου επιπέδου (ή επιπέδων) στους οποίους δεν έχουμε άμεση πρόσβαση κατά τη διαδικασία της εκπαίδευσης (δεν μπορούμε να ξέρουμε ποια θα είναι η έξοδός τους). Οι κόμβοι εξόδου είναι αυτοί που μας δείχνουν την απόκριση του δικτύου σε μία είσοδο. Η ανάλυση για τους κόμβους εξόδου είναι η ίδια με αυτή του κανόνα δέλτα, όπως διατυπώθηκε στην εξίσωση (4.13) παραπάνω:

$$\Delta w_i^j = \alpha \sigma'(a^j) (t^j - y^j) x_i^j \quad (4.15)$$

όπου ο άνω δείκτης εισάγεται για να δηλώσουμε ποιος κόμβος (νευρώνας) περιγράφεται. Για μεγαλύτερη κατανόηση, είναι χρήσιμο να σπάσουμε το δεξιό μέλος της εξίσωσης (4.15) με τον ακόλουθο τρόπο: Ο όρος $(t^j - y^j)$ αντιπροσωπεύει ένα μέτρο του σφάλματος στο j -οστό νευρώνα. Ο όρος $\sigma'(a^j)$ σχετίζεται με το πόσο γρήγορα η ενεργοποίηση μπορεί να μεταβάλλει την έξοδο (και συνεπώς το σφάλμα). Αν αυτό είναι μικρό όταν βρισκόμαστε σε μία απ' τις «ουρές» της σιγμοειδούς, μια αλλαγή της ενεργοποίησης δεν θα αλλάξει πολύ την έξοδο. Αν, ωστόσο, είναι μεγάλο, τότε θα πρέπει να περιμένουμε απότομη μεταβολή για δεδομένη μεταβολή στην ενεργοποίηση. Ο παράγοντας x_i^j σχετίζεται, με τη σειρά του, με το ποσοστό κατά το οποίο η i -οστή είσοδος του υπό μελέτη επιπέδου έχει επηρεάσει την ενεργοποίηση. Σημειώνουμε ότι ο παράγοντας x_i^j δεν είναι απαραίτητα η είσοδος του συστήματος, αλλά η είσοδος που προέρχεται κάθε φορά απ' το προηγούμενο επίπεδο. Αν είναι μηδέν, τότε η συγκεκριμένη είσοδος δεν «ενοχοποιείται» για το σφάλμα και συνεπώς η μεταβολή στο βάρος της θα πρέπει επίσης να είναι μηδέν. Αν απ' την άλλη έχει μεγάλη τιμή (π.χ. 1), τότε η i -οστή είσοδος είχε μεγάλη συνεισφορά στην ενεργοποίηση που έδωσε το σφάλμα και έτσι το βάρος πρέπει να μεταβληθεί κατά ένα ανάλογα μεγάλο ποσοστό.

Ανακεφαλαιώνοντας: Το x_i^j μας λέει πόσο η i -οστή είσοδος του υπό μελέτη επιπέδου είναι υπεύθυνη για την ενεργοποίηση, το $\sigma'(a^j)$ μας λέει πόσο γρήγορα μεταβάλλεται η έξοδος μετά από μεταβολές στην ενεργοποίηση και το $(t^j - y^j)$ είναι το σφάλμα στο j -οστό νευρώνα. Είναι επομένως λογικό, το γινόμενο αυτών των παραγόντων να μας δίνει κάτι το οποίο είναι μέτρο του ρυθμού μεταβολής του σφάλματος ως προς το βάρος w_i^j . Το πώς επιδρούν τα παραπάνω στον προσδιορισμό του σφάλματος φαίνεται διαγραμματικά παρακάτω:



Τυπική αλυσίδα για τον προσδιορισμό του σφάλματος

Χρησιμοποιώντας τους προηγούμενους συμβολισμούς, μπορούμε να συνδυάσουμε δύο απ' τους παράγοντες αυτούς ως εξής:

$$\delta^j = \sigma'(a^j)(t^j - y^j) \quad (4.16)$$

Ο κανόνας δέλτα για τους νευρώνες εξόδου γράφεται τώρα:

$$\Delta w_i^j = \alpha \delta^j x_i \quad (4.17)$$

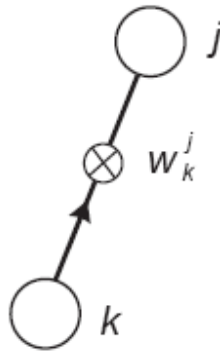
Ας θεωρήσουμε τώρα το ΝΔ με ένα κρυμμένο επίπεδο, που είδαμε και στο Σχήμα 2.7 και συγκεκριμένα τον k -οστό κρυμμένο νευρώνα. Το πρόβλημα στο να προσδιορίσουμε ένα σύνολο μεταβολών βάρους σ' αυτόν τον τύπο νευρώνα έγκειται στο πόσο επηρεάστηκε το σφάλμα από αυτόν το νευρώνα (credit assignment problem).

Για αρχή παρατηρούμε ότι για την i -οστή είσοδο στον κρυμμένο νευρώνα, η τιμή της εισόδου θα έχει παρόμοιο ρόλο με πριν, οπότε μπορούμε να γράψουμε:

$$\Delta w_i^k = \alpha \delta^k x_i \quad (4.18)$$

και του ζητούμενο τώρα είναι να βρούμε τι περιλαμβάνεται στον παράγοντα δ^k .

Ας θεωρήσουμε μία και μόνο έξοδο από τον κρυμμένο νευρώνα στο νευρώνα εξόδου.



Διάγραμμα του k-οστού κρυμμένου νευρώνα και j-οστού νευρώνα εξόδου

Η επίδραση που έχει αυτός ο νευρώνας στο σφάλμα εξαρτάται από δύο πράγματα: Πρώτον, πόσο μπορεί να επηρεάσει την έξοδο του νευρώνα j και, μέσω αυτού, πόσο η έξοδος του νευρώνα j επηρεάζει το σφάλμα. Όσο πιο πολύ ο k επηρεάζει τον j, τόσο μεγαλύτερη περιμένουμε να είναι η επίδρασή του στο σφάλμα, αλλά αυτή θα είναι σημαντική μόνο όταν ο j έχει κάποια επίδραση στο σφάλμα στην έξοδό του. Η συνεισφορά του j στο σφάλμα εκφράζεται με το 'δέλτα', γι' αυτόν τον κόμβο – δ^j . Η επιρροή που έχει ο k στον j δίνεται από το βάρος w_k^j . Συνεπώς, στην έκφραση για το δ^k περιμένουμε να βρούμε το γινόμενο $w_k^j \delta^j$. Όμως η έξοδος από τον k-οστό κόμβο μπορεί να καταλήγει και σε άλλους κόμβους, οπότε οι συνεισφορές στο σφάλμα απ' όλους αυτούς θα πρέπει να ληφθούν υπόψη, δηλαδή θα πρέπει να αθροίσουμε αυτά τα γινόμενα για όλα τα j.

Αυτό έχει σαν αποτέλεσμα η έκφραση για το δ^k να είναι τελικά

$$\delta^k = \sigma'(a^k) \sum_{j \in I_k} \delta^j w_k^j \quad (4.19)$$

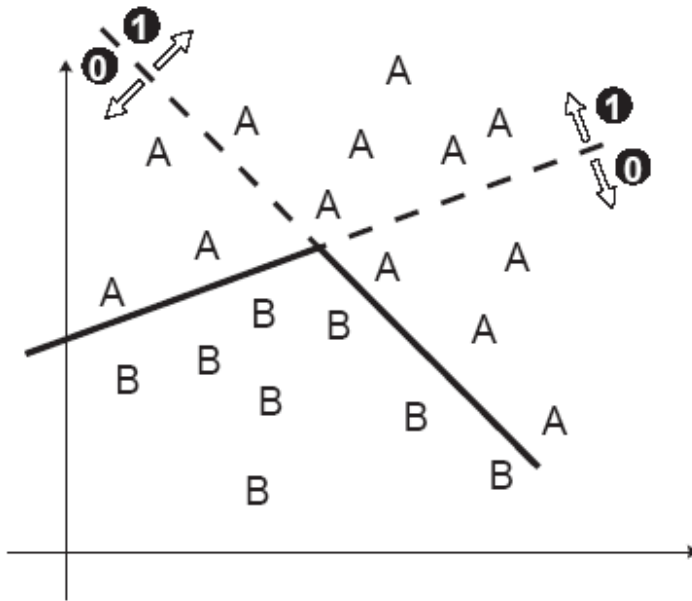
όπου I_k είναι το σύνολο των νευρώνων που παίρνουν είσοδο απ' τον κρυμμένο νευρώνα k (fan-out set of k).

Ο παράγοντας $\sigma'(a^k)$ εμφανίζεται στην έκφραση για το δ^k για τους ίδιους λόγους που εμφανιζόταν και στις περιπτώσεις νευρώνων εξόδου. Χρησιμοποιώντας την εξίσωση (4.17), μπορούμε μέσω της (4.16) να υπολογίσουμε τις μεταβολές στα βάρη για τους κρυμμένους νευρώνες. Η εκπαίδευση ενός δικτύου μπορεί τώρα να γίνει με τα ακόλουθα βήματα:

1. Παρουσιάζουμε το πρότυπο στο στρώμα εισόδου
2. Αφήνουμε τις κρυμμένες μονάδες να υπολογίσουν την έξοδό τους χρησιμοποιώντας το πρότυπο.
3. Αφήνουμε τις μονάδες εξόδου να υπολογίσουν την έξοδό τους, χρησιμοποιώντας ως είσοδο το αποτέλεσμα του βήματος 2 από τις κρυμμένες μονάδες.
Τα βήματα 1-3 είναι γνωστά ως *ευθύ πέρασμα (forward pass)*, καθώς η πληροφορία ρέει προς τα εμπρός, με τη φυσική έννοια, μέσω του δικτύου.
4. Εφαρμόζουμε το πρότυπο-στόχο στο στρώμα εξόδου.
5. Υπολογίζουμε τα δ στους νευρώνες εξόδου, σύμφωνα με την εξίσωση (4.16).
6. Εκπαιδεύουμε κάθε νευρώνα εξόδου χρησιμοποιώντας τον κανόνα διανύσματος κλίσης (gradient descent).
7. Για κάθε κρυμμένο νευρώνα υπολογίζουμε το δ σύμφωνα με την εξίσωση (4.19).
8. Για κάθε κρυμμένο νευρώνα χρησιμοποιούμε το δ που βρήκαμε στο βήμα 7, για να τον εκπαιδεύσουμε με βάση τον κανόνα gradient descent.
Τα βήματα 4-8 είναι γνωστά ως *ανάστροφο πέρασμα (backward pass)*.
Το βήμα 7 περιλαμβάνει *διάδοση* των δ από τους νευρώνες εξόδου, *προς τα πίσω*, στους νευρώνες του κρυμμένου επιπέδου με τους οποίους συνδέονται. Απ' αυτό προέρχεται και το όνομα του αλγορίθμου Error Backpropagation (προς τα πίσω διάδοση του σφάλματος). Ο αλγόριθμος αυτός αναφέρεται και ως *γενικευμένος κανόνας δέλτα*.

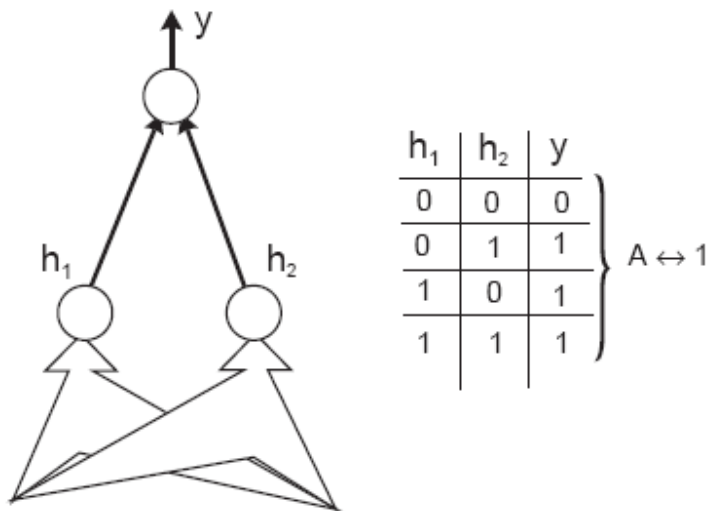
4.3.1 Μη γραμμικώς διαχωρίσιμα προβλήματα

Το δίκτυο δύο επιπέδων παρουσιάστηκε πριν ως ταξινομητής δύο και παραπάνω κλάσεων. Ας θεωρήσουμε τώρα την ακόλουθη περίπτωση:



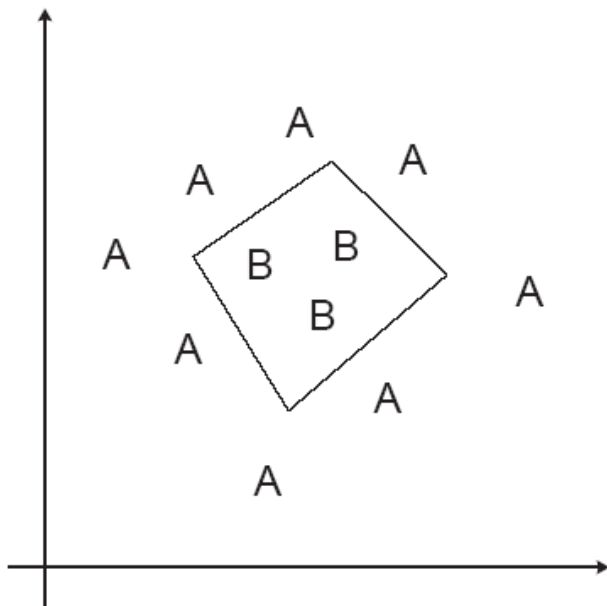
Σχήμα 4.8: Δύο μη γραμμικώς διαχωρίσιμες κλάσεις

Οι δύο κλάσεις A και B δεν μπορούν να διαχωριστούν από ένα και μόνο υπερεπίπεδο. Γενικά απαιτείται μια *επιφάνεια απόφασης* (*decision surface*) αυθαίρετα σχηματισμένη. Στο παραπάνω σχήμα, η επιφάνεια αυτή προσδιορίζεται από δύο επίπεδα. Μπορούμε τώρα να κατασκευάσουμε ένα δίκτυο δύο επιπέδων για να λύσουμε το πρόβλημα αυτό. Κάθε επίπεδο καθορίζεται από ένα ζεύγος κρυμμένων νευρώνων. Ας υποθέσουμε τώρα ότι καθένας απ' αυτούς τους νευρώνες μαθαίνει να δίνει '1' (ή τουλάχιστον μια τιμή κοντά σ' αυτή, με τη σιγμοειδή συνάρτηση), όταν πρόκειται για την κλάση A. Αν ο νευρώνας εξόδου μπορεί να ταξινομήσει το A σαν '1', τότε θα έχει μάθει τη λογική συνάρτηση OR.



Σχήμα 4.9: Πίνακας αληθείας για το νευρώνα εξόδου – πύλη OR

Για πιο σύνθετες επιφάνειες απόφασης χρειαζόμαστε περισσότερες κρυμμένες μονάδες.



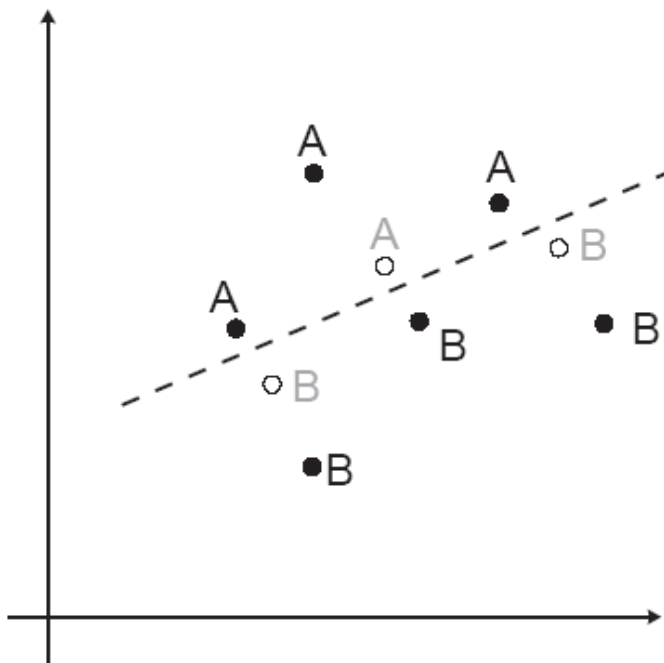
Σχήμα 4.10: Σύνθετη επιφάνεια απόφασης

Για να συνοψίσουμε την ισχύ των εργαλείων που έχουν αναπτυχθεί μέχρι τώρα: μπορούμε να εκπαιδεύσουμε ένα πολυεπίπεδο δίκτυο, ώστε να εκτελεί κατηγοριοποίηση αυθαίρετου αριθμού κλάσεων με μια αυθαίρετη επιφάνεια απόφασης. Αυτό που απαιτείται είναι να έχουμε το σύνολο των εισόδων και των επιθυμητών εξόδων, να

καθορίσουμε τον αριθμό των κρυμμένων μονάδων που θα χρησιμοποιηθούν και να εφαρμόσουμε τον κανόνα διανύσματος κλίσης στο σφάλμα με τον αλγόριθμο της προς τα πίσω διάδοσης. Υπάρχουν βέβαια και λεπτομέρειες που δυσκολεύουν τη διαδικασία. Η πρώτη απ' αυτές αφορά τον αριθμό των κρυμμένων μονάδων που χρησιμοποιούνται και σχετίζεται με την ανεπαρκή γενίκευση του συνόλου εκπαίδευσης (*training set generalisation*). Η δεύτερη αφορά τη φύση του κανόνα διανύσματος κλίσης (*gradient descent*).

4.3.1.1 Γενίκευση

Ας θεωρήσουμε την παρακάτω περίπτωση για το χώρο των προτύπων:

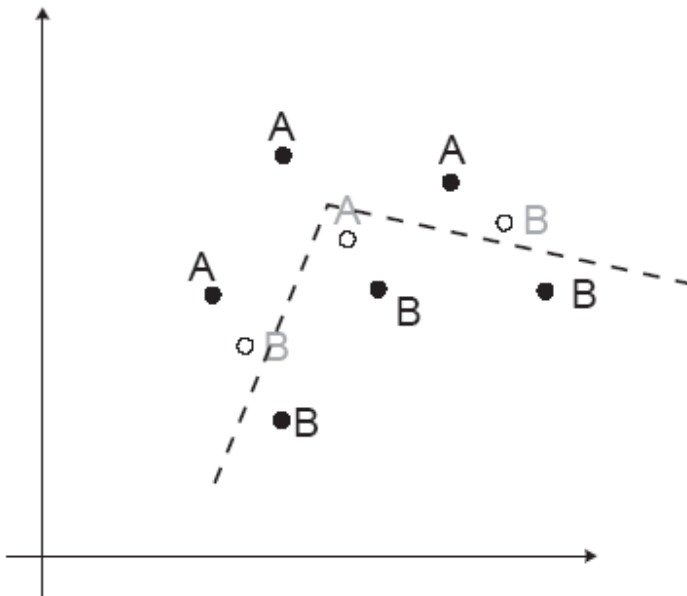


Σχήμα 4.11: Διάγραμμα εκπαίδευσης για δύο κλάσεις.

Τα πρότυπα εκπαίδευσης φαίνονται με τα έντονα σημεία και υπάρχουν δύο κλάσεις A και B. Μόνο ένας νευρώνας χρησιμοποιείται για να τις ταξινομήσει. Οι κύκλοι που φαίνονται σε κάθε κλάση αποτελούν διανύσματα που δεν χρησιμοποιήθηκαν κατά τη διάρκεια της εκπαίδευσης. Αυτά είναι τα πρότυπα επαλήθευσης (*test patterns*). Είναι αντιπροσωπευτικά της κάθε κλάσης, πάρθηκαν από το σύνολο επαλήθευσης και ταξινομήθηκαν σωστά, αν και το δίκτυο δεν τα «είδε» κατά την εκπαίδευση. Αυτή είναι

η δύναμη του δικτύου και ένας απ' τους βασικούς λόγους που το χρησιμοποιούμε. Λέμε τότε ότι το δίκτυο έχει γενικεύσει από το σύνολο εκπαίδευσης.

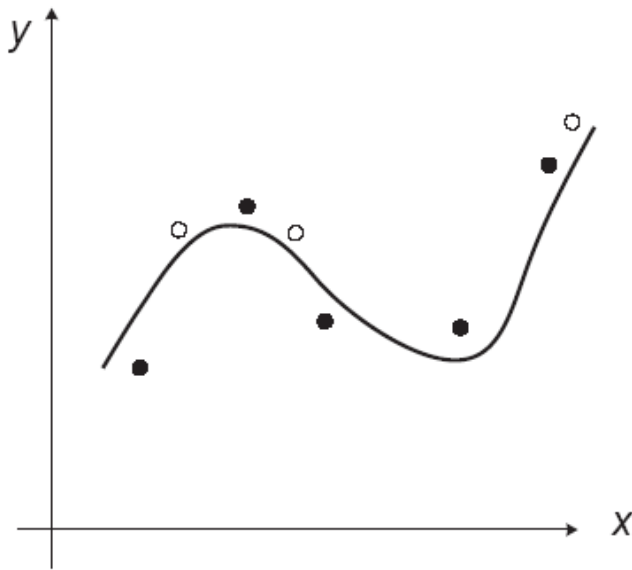
Η σωστή ταξινόμηση των προτύπων επαλήθευσης που φαίνεται στο Σχήμα 4.11 φαίνεται να δικαιώνει ενός μόνο υπερεπίπεδου ως επιφάνεια απόφασης. Ας υποθέσουμε ότι χρησιμοποιούσαμε δύο υπερεπίπεδα (δύο κρυμμένες μονάδες σε ένα διεπίπεδο δίκτυο). Για την ελαχιστοποίηση του σφάλματος, τα επίπεδα θα μπορούσαν να ευθυγραμμιστούν όσο το δυνατό πιο κοντά στο σύνολο εκπαίδευσης.



Σχήμα 4.12: Χρησιμοποίηση 2^{ου} επιπέδου για την ταξινόμηση των κλάσεων A και B

Κάποια από τα δεδομένα επαλήθευσης είναι τώρα λάθος ταξινομημένα. Το πρόβλημα είναι ότι το δίκτυο, με δύο κρυμμένες μονάδες, έχει πλέον πάρα πολλή ελευθερία και ταίριαξε ένα σύνορο απόφασης στα δεδομένα εκπαίδευσης, χωρίς να υπολογίζει τις υποκείμενες τάσεις.

Υπάρχει κι άλλος τρόπος να το δούμε αυτό, μέσω της συνάρτησης εισόδου-εξόδου του δικτύου. Το παρακάτω διάγραμμα δείχνει, σχηματικά, την έξοδο y ενός δυαδικού (δύο-κλάσεων) ταξινομητή σαν συνάρτηση της εισόδου του.

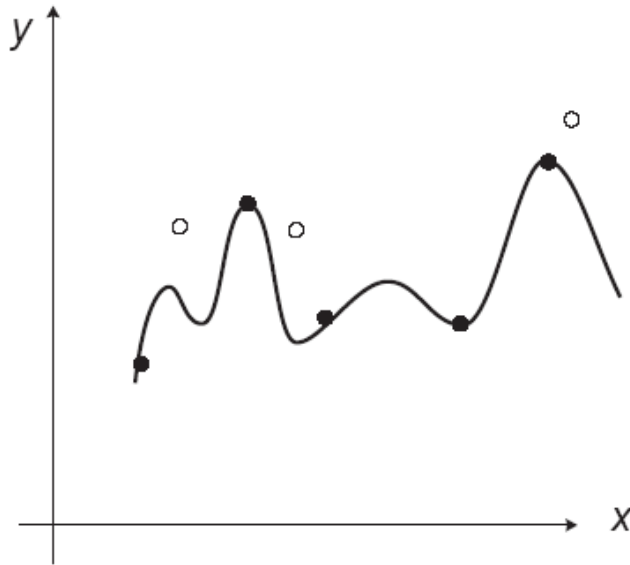


Σχήμα 4.13: y συναρτήσει του x για ένα δυαδικό ταξινομητή

Η καμπύλη που φαίνεται στο Σχήμα 4.13 είναι η πραγματική έξοδος, για την είσοδο x , ενώ τα σημεία αντιπροσωπεύουν τα δεδομένα εκπαίδευσης. Αν η καμπύλη περνούσε ακριβώς από το σύνολο εκπαίδευσης, το σφάλμα θα ήταν ίσο με το μηδέν. Κάτι τέτοιο δεν συμβαίνει στο προηγούμενο παράδειγμα, αλλά η έξοδος έχει συλλάβει την τάση για μεταβολή των δεδομένων. Αν χρησιμοποιούσαμε περισσότερες κρυμμένες μονάδες, το δίκτυο θα είχε μεγαλύτερη ελευθερία, η έξοδός του θα μεταβαλλόταν πιο γρήγορα, αποκρινόμενη σε μια μεταβολή της εισόδου και μπορεί να είχαμε μια καμπύλη όπως αυτή του επόμενου σχήματος.

Τώρα, αν και η καμπύλη έχει προσαρμοστεί σχεδόν πλήρως στα δεδομένα εκπαίδευσης, δίνοντας σχεδόν μηδενικό σφάλμα, τα δεδομένα επαλήθευσης δεν είναι σωστά ταξινομημένα. Το δίκτυο έχει κάνει ανεπαρκή γενίκευση. Το ερώτημα που παραμένει αναπάντητο είναι πώς θα καθορίσουμε τον αριθμό των υπερεπιπέδων ή αλλιώς των κρυμμένων μονάδων που θα χρησιμοποιήσουμε. Αρχικά αυτό δεν φαινόταν να είναι πρόβλημα, αφού υπήρχε η εντύπωση ότι μπορούμε πάντοτε να χρησιμοποιούμε περισσότερες κρυμμένες μονάδες απ' όσες είναι αυστηρά απαραίτητες. Θα υπήρχε απλώς ένα πλεόνασμα, με περισσότερους από έναν κρυμμένους νευρώνες ανά υπερεπίπεδο. Δυστυχώς όμως αυτό δεν ισχύει. Οι πάρα πολλοί κρυμμένοι νευρώνες μπορούν να κάνουν το δίκτυό μας να ταξινομεί άριστα το σύνολο εκπαίδευσης σε βάρος

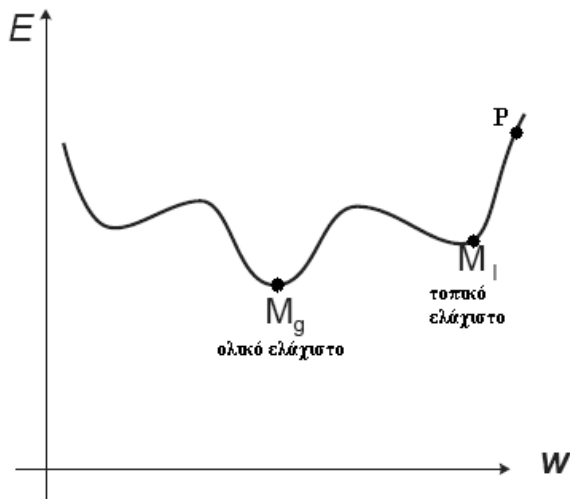
οποιασδήποτε χρήσιμης γενίκευσης. Ο υπολογισμός του βέλτιστου αριθμού των κρυμμένων νευρώνων είναι ένα πρόβλημα που ερευνάται.



Σχήμα 4.14: Ταξινόμηση με περισσότερες κρυμμένες μονάδες

4.3.2 Τοπικά Ακρότατα

Ας θεωρήσουμε τη συνάρτηση σφάλματος που φαίνεται στο παρακάτω σχήμα:



Σχήμα 4.15: Ολικό και τοπικό ελάχιστο της συνάρτησης σφάλματος

Ας υποθέσουμε ότι ξεκινάμε με ένα σύνολο βαρών για το δίκτυο, που αντιστοιχούν στο σημείο P. Εάν εφαρμόσουμε τον κανόνα διανύσματος κλίσης, το ελάχιστο που θα συναντήσουμε είναι αυτό στο M_l , κι όχι αυτό στο M_g . Το M_l ονομάζεται τοπικό ελάχιστο και αντιστοιχεί σε μία μερική λύση για το δίκτυο σε σχέση με τα δεδομένα εκπαίδευσης.

Το M_g είναι το ολικό ελάχιστο που αναζητούμε και, αν δεν έχουν ληφθεί μέτρα για να ξεφύγουμε από το M_l , δεν θα φτάσουμε ποτέ στο M_g . Ουσιαστικά, θα πρέπει να επιλέξουμε μια τέτοια κατάσταση, στην οποία κάθε μεταβολή στα βάρη να ελέγχεται όχι μόνο από την κλίση του σφάλματος, αλλά να περιλαμβάνει και ένα τυχαίο στοιχείο, ώστε κάποιες φορές να “ανεβαίνουμε” την καμπύλη, αντί να την “κατεβαίνουμε”.

4.3.3 Επιταχύνοντας τη μάθηση: ο ρόλος της ορμής (*momentum*)

Η ταχύτητα της μάθησης εξαρτάται από το ρυθμό μάθησης α . Αν αυτός αυξηθεί πάρα πολύ, η μάθηση γίνεται ασταθής και το δίκτυο ταλαντεύεται γύρω από το ελάχιστο σφάλμα. Άρα, για να έχουμε υψηλό επίπεδο μάθησης, ο ρυθμός μάθησης θα πρέπει να είναι μικρός. Ένας τρόπος να ξεπεράσουμε τους περιορισμούς που επιβάλλει κάτι τέτοιο είναι να μεταβάλλουμε τον κανόνα μάθησης, ώστε να περιλάβει έναν όρο που περιέχει ένα ποσοστό της τελευταίας μεταβολής στο βάρος. Ο νέος αυτός κανόνας εκφράζεται από τη σχέση:

$$\Delta w_i^j(n) = \alpha \delta^j x_i^j + \lambda \Delta w_i^j(n-1) \quad (4.20)$$

Έτσι, αν η προηγούμενη μεταβολή βάρους $\Delta w_i^j(n-1)$ ήταν μεγάλη, αντίστοιχα μεγάλη θα είναι και η τρέχουσα μεταβολή $\Delta w_i^j(n)$. Αυτό σημαίνει ότι η μεταβολή βάρους μεταφέρει κάποια ορμή στην επόμενη επανάληψη. Κάτι τέτοιο έχει την τάση να εξομαλύνει μικρές διακυμάνσεις στο διάστημα σφάλματος-βάρους. Η παράμετρος λ ελέγχει τη συνεισφορά της ορμής (*momentum*). Η σταθερά ορμής (λ) είναι συνήθως θετικός αριθμός.

4.3.4 Κριτήρια τερματισμού του αλγόριθμου *Backpropagation*

Όσον αφορά τον τερματισμό του αλγόριθμου της ανάστροφης διάδοσης σφάλματος θα πρέπει αναφερθεί ότι δεν συγκλίνει γενικά και ότι δεν υφίστανται απόλυτα ορισμένα κριτήρια τερματισμού της λειτουργίας του. Στην πράξη όμως υπάρχουν κάποια λογικά κριτήρια που οδηγούν στον τερματισμό της διαδικασίας εκπαίδευσης. Αυτά τα κριτήρια έχουν να κάνουν κυρίως με ιδιότητες των τοπικών και ολικών ελαχίστων της επιφάνειας σφάλματος. Αν θεωρήσουμε ότι το διάνυσμα βαρών w_0 είναι τοπικό ή ολικό ελάχιστο, τότε, για να ισχύει αυτό, θα πρέπει το διάνυσμα κλίσης $\vec{g}(\vec{w})$, δηλαδή η μερική

παράγωγος του σφάλματος ως προς το διάνυσμα βαρών \bar{w} , να ισούται με το μηδέν, για $\bar{w} = \bar{w}_0$. Πιθανά κριτήρια σύγκλισης του αλγόριθμου είναι τα εξής [17]:

1. Ο αλγόριθμος οπίσθιας τροφοδότησης θεωρείται ότι έχει συγκλίνει, όταν η ευκλείδεια νόρμα του διανύσματος κλίσης ξεπεράσει ένα πολύ μικρό κατώφλι κλίσης. Το μειονέκτημα αυτής της μεθόδου είναι ότι ο χρόνος εκπαίδευσης μπορεί να είναι μεγάλος, ενώ πρέπει να υπολογίζεται και το διάνυσμα κλίσης. Επιπρόσθετα μια ακόμα ιδιότητα του ελαχίστου \bar{w}_0 είναι ότι η συνάρτηση του μέσου τετραγωνικού σφάλματος είναι στατική, κάτι που μας οδηγεί στο παρακάτω κριτήριο τερματισμού:

2. Ο αλγόριθμος ανάστροφης διάδοσης θεωρείται ότι έχει συγκλίνει, όταν η απόλυτη τιμή του ρυθμού μεταβολής του σφάλματος ανά εποχή είναι ικανοποιητικά μικρή.

3. Ο αλγόριθμος ανάστροφης διάδοσης σφάλματος τερματίζει σε ένα διάνυσμα βαρών \bar{w}_{final} , όταν $\|\bar{g}(\bar{w}_{\text{final}})\| \leq \varepsilon$, όπου ε ένα αρκετά ικανοποιητικά μικρό κατώφλι κλίσης ή αν $E_{\text{av}}(\bar{w}_{\text{final}}) \leq \tau$, όπου τ ένα ικανοποιητικά μικρό κατώφλι για τη μέση τιμή του σφάλματος (E_{av}).

4.3.5 Τρόποι εκπαίδευσης

Η διαδικασία εκπαίδευσης κατά τον αλγόριθμο Back-Propagation λαμβάνει χώρα με την παρουσίαση και εφαρμογή στο νευρωνικό δίκτυο ενός συνόλου παραδειγμάτων εκπαίδευσης. Η παρουσίαση όλων των προτύπων εκπαίδευσης μια φορά το καθένα στο δίκτυο ονομάζεται εποχή (epoch). Κατά τη συνολική διαδικασία εκπαίδευσης εκτελούνται επαναλήψεις των εποχών, ώσπου τα βάρη του δικτύου να σταθεροποιηθούν σε συγκεκριμένες τιμές που θα προκαλούν σύγκλιση της μέσης τιμής των σφαλμάτων, για όλα τα πρότυπα εκπαίδευσης, στην ελάχιστη δυνατή τιμή της.

Εδώ αξίζει να αναφερθεί ότι σε κάθε εποχή υπάρχει ο τυχαίος και ο σειριακός τρόπος παρουσίασης των προτύπων εκπαίδευσης. Ο πρώτος τρόπος έχει το πλεονέκτημα της εξασφάλισης του στοχαστικού χαρακτήρα της μάθησης και το μειονέκτημα των συνεχών και μη αμελητέων ταλαντώσεων των σφαλμάτων εκπαίδευσης και αξιολόγησης, ώσπου να φτάσουμε στο τελικό αποτέλεσμα. Από την άλλη ο δεύτερος τρόπος, της εν σειρά παρουσίασης προτύπων, δίνει συνήθως χαμηλότερο σφάλμα εκπαίδευσης και μεγαλύτερη δυνατότητα ελέγχου αυτού του σφάλματος, αλλά εμπεριέχει τον κίνδυνο της

λεγόμενης «αποστήθισης» του δικτύου της συγκεκριμένης σειράς παρουσίασης των προτύπων, με αποτέλεσμα το τελικό σφάλμα αξιολόγησης να είναι κάποιες φορές υψηλότερο σε σχέση με εκείνο του τυχαίου τρόπου παρουσίασης.

Για ένα δεδομένο σύνολο εκπαίδευσης η διαδικασία μπορεί να εκτελεστεί με δύο τρόπους: ανά πρότυπο και ανά εποχή.

Στην περίπτωση της εκπαίδευσης ανά πρότυπο τα βάρη αναπροσαρμόζονται έπειτα από την παρουσίαση κάθε πρότυπου εκπαίδευσης στο δίκτυο. Για να γίνει αυτό πιο κατανοητό έστω ότι έχουμε μια εποχή που αποτελείται από N πρότυπα εκπαίδευσης $\{\bar{x}(1), \bar{t}(1)\}, \{\bar{x}(2), \bar{t}(2)\}, \dots, \{\bar{x}(n), \bar{t}(n)\}$, όπου τα $\{\bar{x}(i), \bar{t}(i)\}$ είναι στη γενική περίπτωση διανύσματα. Το πρώτο πρότυπο εφαρμόζεται στο δίκτυο και εκτελούνται τόσο το ευθύ, όσο και το ανάστροφο πέρασμα που περιγράψαμε προηγουμένως, με αποτέλεσμα να μεταβάλλονται τα βάρη. Στη συνέχεια παρουσιάζεται το δεύτερο στη σειρά πρότυπο και εκτελείται η ίδια διαδικασία που οδηγεί σε νέα μεταβολή των βαρών. Η εποχή τελειώνει με την παρουσίαση του N -οστού προτύπου. Στην περίπτωση της εκπαίδευσης ανά εποχή τα βάρη ενημερώνονται έπειτα από την παρουσίαση στο δίκτυο όλου του συνόλου προτύπων της εποχής. Στην πράξη η μέθοδος που προτιμάται είναι η εκπαίδευση ανά πρότυπο, καθώς ενισχύει το στοχαστικό χαρακτήρα και είναι πιο ικανή στην αποτροπή παγίδευσης του αλγορίθμου σε τοπικά ελάχιστα. Βέβαια η εκπαίδευση ανά εποχή, υπολογίζει με πιο ακριβή τρόπο το διάνυσμα κλίσης. Σε κάθε πρόβλημα λοιπόν σταθμίζονται τα πλεονεκτήματα και μειονεκτήματα και επιλέγεται μια εκ των δύο μεθόδων εκπαίδευσης [19].

4.4 Σύνοψη δυνατοτήτων του Πολυστρωματικού Perceptron

Το πολυστρωματικό (πολυεπίπεδο) perceptron έχει τη δυνατότητα να εκτελεί ικανοποιητικά τις παρακάτω πρακτικές εργασίες [14]:

- Ταξινόμηση προτύπων, στην οποία αναφερθήκαμε εκτενώς παραπάνω.
- Προσέγγιση συναρτήσεων
- Γενίκευση
- Ανεκτικότητα σε βλάβες

4.4.1 Ταξινόμηση προτύπων

Όπως είδαμε, το μονοστρωματικό perceptron δεν μπορεί να λύσει το πρόβλημα XOR ή άλλα πολύπλοκα (μη γραμμικά) προβλήματα ταξινόμησης, γιατί παράγει ένα απλό επίπεδο διαχωρισμού μεταξύ των κατηγοριών. Το πρόβλημα XOR όμως λύνεται επιτυχώς, αν προστεθεί ένα κρυμμένο στρώμα. Ένα ΤΝΔ με δύο στρώματα μπορεί να διακρίνει (ταξινομήσει) περιοχές (βλ. Σχήμα 4.8). Το πρώτο στρώμα μπορεί να παράγει γραμμές (επίπεδα) ταξινόμησης στο χώρο προτύπων, οι οποίες δεν μπορούν να υπερβούν το πλήθος των κόμβων του. Οι γραμμές αυτές συνδυάζονται από τους νευρώνες εξόδου και παράγουν κυρτές περιοχές, σαν κι αυτή του Σχήματος 4.12. Προσθέτοντας ένα ακόμη κρυμμένο στρώμα κόμβων υπολογισμού, αυτό συνδυάζει τις γραμμές (επίπεδα) που δέχεται από τους κόμβους του πρώτου κρυμμένου στρώματος και στέλνει κυρτές περιοχές στους νευρώνες του στρώματος εξόδου, οι οποίοι παράγουν περιοχές οποιουδήποτε σχήματος στο χώρο προτύπων. Η πολυπλοκότητα του σχήματος των περιοχών αυτών περιορίζεται μόνο από τον αριθμό των κόμβων του δικτύου.

4.4.2 Προσέγγιση συναρτήσεων

Ένα πολυστρωματικό ΝΔ perceptron εκπαιδευόμενο με τον αλγόριθμο ανάστροφης διάδοσης σφάλματος (BP) μπορεί να χρησιμοποιηθεί για την προσέγγιση μη γραμμικών συναρτήσεων εισόδου-εξόδου. Το θεώρημα Kolmogorov, που θα διατυπώσουμε παρακάτω, εξασφαλίζει ότι με ένα κρυμμένο στρώμα η προσέγγιση αυτή μπορεί να έχει οποιαδήποτε ακρίβεια επιθυμούμε.

Θεώρημα Kolmogorov:

Για κάθε $\varepsilon > 0$ και κάθε συνεχή συνάρτηση $F(\mathbf{x}): [0,1]^n \subset \mathcal{R}^n \rightarrow \mathcal{R}$, υπάρχει ένα ΝΔ με ένα μοναδικό κρυμμένο στρώμα, διάνυσμα εισόδου \mathbf{x} και έξοδο $y(\mathbf{x})$:

$$y(\mathbf{x}) = \sum_{i=1}^N \lambda_i f \left(\sum_{j=1}^n w_{ij} x_j - \theta_i \right) \quad (4.21)$$

τέτοιο ώστε για όλα τα $\mathbf{x} \in [0,1]^n$ να ισχύει:

$$|F(\mathbf{x}) - y(\mathbf{x})| < \varepsilon \quad (4.22)$$

όπου $f(\cdot)$ είναι μια συνεχής σιγμοειδής (δηλαδή πεπερασμένη και μονότονα αύξουσα) συνάρτηση ενεργοποίησης και λ_i, θ_i ($i=1,2,\dots,N$) και w_{ij} ($i=1,2,\dots,N, j=1,2,\dots,n$) είναι κατάλληλες πραγματικές σταθερές.

Αυτό σημαίνει ότι η έξοδος $y(\mathbf{x})$ του ΝΔ, οι κρυμμένοι νευρώνες του οποίου έχουν βάρη w_{ij} και κατώφλια θ_i είναι μια ομοιόμορφη προσέγγιση της δοσμένης μη-γραμμικής συνάρτησης $F(\mathbf{x})$ για όλες τις τιμές του διανύσματος εισόδου \mathbf{x} στο n -διάστατο μοναδιαίο υπερκύβο $[0,1]^n$. Σημειώνεται ότι ο νευρώνας εξόδου είναι γραμμικός με βάρη $\lambda_1, \lambda_2, \dots, \lambda_N$ (δεν πρέπει να συγχέονται με τη σταθερά ορμής λ). Πρέπει να τονισθεί ότι το θεώρημα Kolmogorov εξασφαλίζει μόνο την ύπαρξη ενός ΤΝΔ με ένα κρυμμένο στρώμα που να προσεγγίζει τη μη γραμμική συνάρτηση με οποιαδήποτε επιθυμητή ακρίβεια, αλλά δεν μας λέει αν το δίκτυο αυτό είναι βέλτιστο σε σχέση με το χρόνο μάθησης ή την ευκολία υλοποίησης του. Τέλος, οι (4.21) και (4.22) δεν εγγυώνται ότι με ένα δεδομένο αλγόριθμο μάθησης το ΤΝΔ θα φτάσει κάποτε σε κατάλληλα βάρη w_{ij} και λ_i που να ικανοποιούν τη συνθήκη ακριβείας (4.22).

4.4.3 Γενίκευση

Γενίκευση (generalization) είναι η ικανότητα ενός ΤΝΔ να ταξινομεί πρότυπα, τα οποία δεν γνώρισε ποτέ. Η γενίκευση είναι μία από τις πιο σημαντικές ιδιότητες των ΤΝΔ. Υποτίθεται βεβαίως ότι οι προς ταξινόμηση εισοδοί (πρότυπα) λαμβάνονται από τον ίδιο πληθυσμό προτύπων, ο οποίος χρησιμοποιήθηκε για την εκπαίδευση του δικτύου. Η γενίκευση επιτυγχάνεται ανιχνεύοντας ιδιότητες του προτύπου εισόδου που είναι σημαντικές και έχουν εμφυτευθεί (κωδικοποιηθεί) στους εσωτερικούς κόμβους. Ένα γνωστό πρότυπο ταξινομείται τότε μαζί με τα άλλα πρότυπα που έχουν τις ίδιες χαρακτηριστικές ιδιότητες. Έτσι, μάθηση μέσω παραδειγμάτων είναι μια καλή μέθοδος εκπαίδευσης του ΤΝΔ. Το ΤΝΔ έχει επίσης την ικανότητα να διακρίνει και να ταξινομεί (σχετικά καλά) ελλιπή, παραμορφωμένα και διαταραγμένα με θόρυβο πρότυπα. Πρέπει να σημειωθεί ότι ένα ΤΝΔ μπορεί να προβεί καλύτερα σε παρεμβολή δεδομένων εισόδου (interpolation), παρά σε πρόβλεψη (επέκταση, extrapolation), γιατί μπορεί να ανιχνεύσει καλά ενδιάμεσες καταστάσεις που δεν έχει δει, παρά καταστάσεις που βρίσκονται έξω από την περιοχή στην οποία έχει εκπαιδευτεί. Τούτο γίνεται κατανοητό αν θεωρήσουμε τη διαδικασία εκπαίδευσης του ΤΝΔ ως μια διαδικασία προσαρμογής καμπύλης (curve

fitting). Το TND εκπαιδεύεται με ένα δεδομένο σύνολο ζευγών εισόδου-εξόδου και καλείται να ανιχνεύσει ενδιάμεσα σημεία με παρεμβολή. Εάν το δίκτυο έχει εκπαιδευτεί σωστά (βλ. Σχήμα 4.6), τότε είναι ικανό να δώσει ένα σωστό σημείο (ζεύγος τιμών εισόδου-εξόδου), έστω κι αν είναι ελαφρά διαφορετικό από τα παραδείγματα με τα οποία εκπαιδεύτηκε. Εάν όμως το δίκτυο έχει “υπερεκπαιδευτεί” (βλ. Σχήμα 4.7), τότε συνήθως η ικανότητά του να γενικεύει σωστά μειώνεται.

Η γενίκευση ενός TND επηρεάζεται από τα ακόλουθα:

- α) Το μέγεθος και την καταλληλότητα του συνόλου εκπαίδευσης.
- β) Την αρχιτεκτονική του δικτύου.
- γ) Την πολυπλοκότητα του θεωρούμενου προβλήματος.

Στην πράξη, προσπαθούμε να βελτιώσουμε την ποιότητα της γενίκευσης επιλέγοντας το πιο κατάλληλο σύνολο εκπαίδευσης. Η καλύτερη αρχιτεκτονική του δικτύου επιλέγεται συνήθως με προσεκτική μελέτη του υπό εξέταση προβλήματος.

4.4.4 Ανεκτικότητα σε βλάβες

Το πολυστρωματικό perceptron είναι ενδογενώς ανεκτικό σε βλάβες, γιατί είναι ένα σύστημα παράλληλης επεξεργασίας που κάθε κόμβος επεξεργασίας συμβάλλει στο τελικό αποτέλεσμα της εξόδου. Εάν χαθεί ή χαλάσει ένας κόμβος, ή υποστούν βλάβη τα συναπτικά του βάρη, η “ανάκτηση” πέφτει σε ποιότητα, αλλά η κατανεμημένη φύση της πληροφορίας εξασφαλίζει ότι η βλάβη θα πρέπει να είναι πολύ μεγάλη, για να υποβιβασθεί η ποιότητα της απόκρισης σε μεγάλο βαθμό. Συνεπώς το δίκτυο παρουσιάζει ελαφρά πτώση της συμπεριφοράς του και όχι καταστροφική. Επίσης, το πολυστρωματικό perceptron αντέχει σε θορύβους, λόγω της ικανότητας της γενίκευσης που έχει. Η βλάβη στο δίκτυο αντιμετωπίζεται με επανεκπαίδευση και η “ανάρρωση” αυτή του δικτύου είναι συνήθως πολύ γρήγορη.

Κεφάλαιο 5

Δίκτυα χωρίς επίβλεψη

5.1 Γενικά

Εκτός από τα προβλήματα μάθησης με επίβλεψη (π.χ. ταξινόμηση), που συναντήσαμε στο προηγούμενο κεφάλαιο, έχουν αναπτυχθεί σημαντικά μοντέλα ΝΔ κατάλληλα για μάθηση χωρίς επίβλεψη.

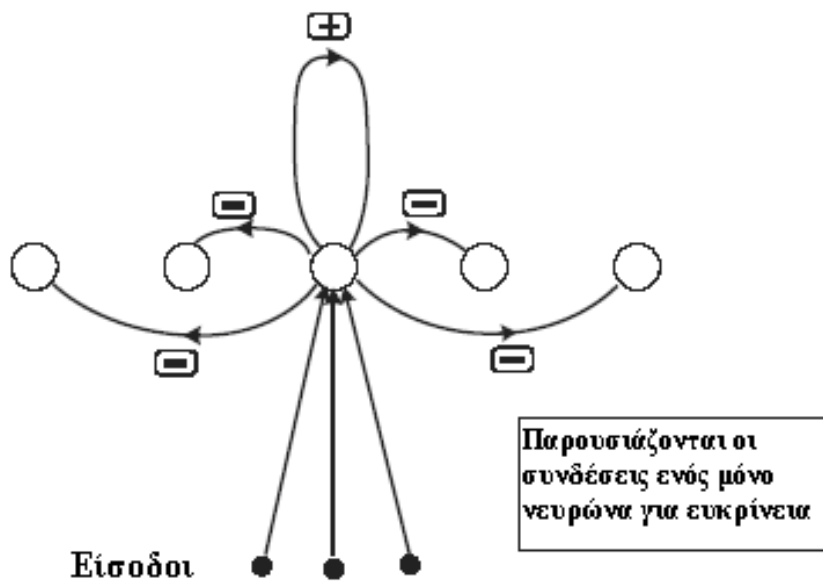
Ένα σύστημα μη επιβλεπόμενης μάθησης λειτουργεί εξάγοντας ιδιότητες ή χαρακτηριστικά των προτύπων τα οποία του παρουσιάζονται, χωρίς να του έχει “ειπωθεί” ποιες έξοδοι ή κατηγορίες εισόδων είναι επιθυμητές. Δηλαδή το σύστημα ανιχνεύει ή κατηγοριοποιεί μόνιμες ιδιότητες, χωρίς καμία ανατροφοδότηση από το περιβάλλον. Τα ΝΔ μη επιβλεπόμενης μάθησης προσπαθούν να μάθουν να αποκρίνονται σε διάφορα πρότυπα εισόδου με διαφορετικά τμήματα του δικτύου. Το ΝΔ συνήθως εκπαιδεύεται να δυναμώνει την πυροδότηση σε απόκριση συχνά εμφανιζόμενων προτύπων. Με τον τρόπο αυτό το ΝΔ αναπτύσσει ορισμένες εσωτερικές παραστάσεις για την κωδικοποίηση προτύπων εισόδου. Παραδείγματα ΝΔ μη επιβλεπόμενης μάθησης είναι: i) τα ΝΔ ανταγωνιστικής μάθησης, ii) το αυτο-οργανούμενο ΝΔ Kohonen, iii) τα ΝΔ ανάλυσης σε κύριες συνιστώσες και iv) το αυτο-οργανούμενο ΝΔ βασισμένο στην αρχή ‘διατήρησης μέγιστης πληροφορίας’ (maximum information preservation – MIP) [14].

5.2 Ανταγωνιστική μάθηση

Η ανταγωνιστική μάθηση υλοποιείται με τη χρήση ενός επιπέδου από ‘ανταγωνιστικούς’ νευρώνες, το οποίο ονομάζεται ανταγωνιστικό επίπεδο (competitive layer). Κάθε φορά που εμφανίζεται κάποιο διάνυσμα εισόδου στο ανταγωνιστικό επίπεδο, έχουμε ‘ανταγωνισμό’ (με βάση κάποιο κριτήριο επίδοσης) μεταξύ των νευρώνων του ανταγωνιστικού επιπέδου για το ποιος θα βγει ‘νικητής’ για το συγκεκριμένο πρότυπο. Ο νικητής νευρώνας είναι αυτός ο οποίος συνήθως μεταβάλλει τις τιμές των βαρών του

περισσότερο σε σχέση με τους υπόλοιπους νευρώνες. Με τον τρόπο αυτό έχει βρεθεί ότι επιτυγχάνεται αυτό-οργάνωση (self-organisation) και συγκεκριμένοι νευρώνες μαθαίνουν να αποκρίνονται σε συγκεκριμένα ερεθίσματα (διανύσματα εισόδου). Συνήθως το κριτήριο που χρησιμοποιείται για την εύρεση του νικητή νευρώνα είναι η απόσταση (στις περισσότερες περιπτώσεις ευκλείδεια) του διανύσματος εισόδου $\bar{x} = (x_1, \dots, x_d)^T$ από το διάνυσμα των βαρών $w_i = (w_{i1}, \dots, w_{id})^T$ κάθε νευρώνα i . Ο νικητής νευρώνας είναι αυτός με τη μικρότερη απόσταση $|\bar{x} - \bar{w}_j|$. Η αυτό-οργάνωση επιτυγχάνεται στη συνέχεια μετακινώντας το διάνυσμα βαρών του νικητή νευρώνα j προς το διάνυσμα εισόδου \bar{x} . Η ίδια διαδικασία επαναλαμβάνεται κάθε φορά για κάθε νέο διάνυσμα εισόδου που εμφανίζεται στο ανταγωνιστικό επίπεδο [19]. Στη συνέχεια θα δούμε αναλυτικά όσα είπαμε παραπάνω για την ανταγωνιστική μάθηση και τα ΝΔ που βασίζονται σ' αυτήν.

Ας θεωρήσουμε ένα ανταγωνιστικό επίπεδο, δηλαδή μια ομάδα νευρώνων, όπως φαίνεται στο παρακάτω σχήμα:



Σχήμα 5. 1: Ανταγωνιστικό επίπεδο [18]

Κάθε κελί (νευρώνας) λαμβάνει το ίδιο σύνολο εισόδων από ένα στρώμα εισόδου και υπάρχουν ενδοστρωματικές ή *παράλληλες* συνδέσεις έτσι, ώστε κάθε κόμβος να

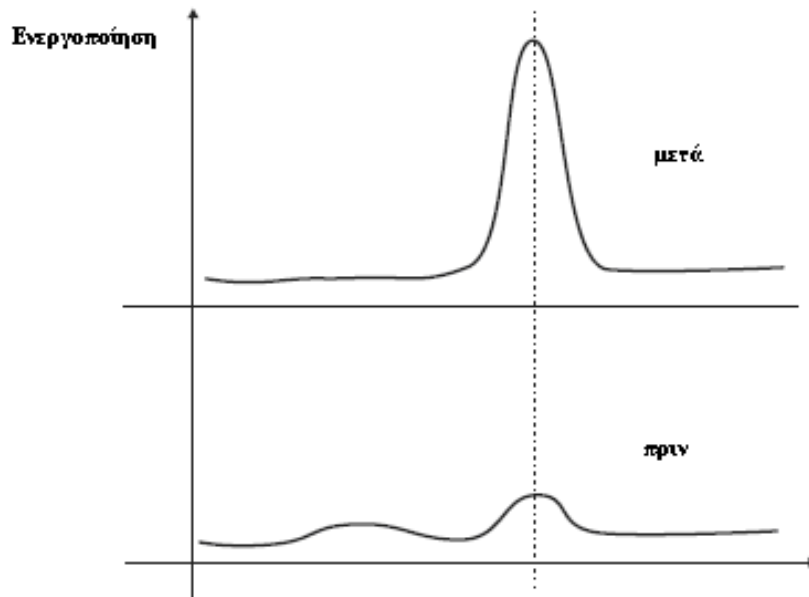
συνδέεται με τον εαυτό του μέσω ενός διεγερτικού (θετικού) βάρους και να εμποδίζει όλους τους άλλους κόμβους στο στρώμα με αρνητικά βάρη. Ας υποθέσουμε τώρα ότι ένα διάνυσμα \bar{x} παρουσιάζεται στην είσοδο. Κάθε μονάδα τώρα υπολογίζει ένα σταθμισμένο άθροισμα s των εισόδων που προέρχονται από το διάνυσμα \bar{x} . Δηλαδή:

$$s = \sum_i w_i x_i \quad (5.1)$$

Με διανυσματικό συμβολισμό, αυτό δεν είναι άλλο από το εσωτερικό γινόμενο $\bar{w} \cdot \bar{x}$. Έστω τώρα ότι κάποιος k νευρώνας θα έχει μια τιμή για το s μεγαλύτερη από όλους τους άλλους νευρώνες (κόμβους) στο στρώμα.

Αυτό που συμβαίνει είναι ότι η ενεργοποίηση του νευρώνα με τη μεγαλύτερη διέγερση s από την είσοδο επηρεάζεται άμεσα (αυξάνεται) απ' αυτό και έμμεσα από την αυτό-διεγερόμενη (παράλληλη) σύνδεση. Αυτός στη συνέχεια παρεμποδίζει τους γειτονικούς νευρώνες κι έτσι η επίδραση τους στον k μειώνεται ακόμη περισσότερο. Αυτή η διαδικασία συνεχίζεται μέχρι να επιτευχθεί κάποια σταθερότητα. Υπάρχει συνεπώς ένας 'ανταγωνισμός' για την ενεργοποίηση μέσα στο επίπεδο και το δίκτυο λέμε ότι εξελίχθηκε μέσω *ανταγωνιστικών δυναμικών (competitive dynamics)*. Κάτω από κατάλληλες συνθήκες, η δραστηριότητα των νευρώνων, των οποίων η είσοδος s ήταν μικρότερη από του νευρώνα-νικητή k , θα μηδενιστεί. Λέμε τότε ότι πρόκειται για ένα 'winner-takes-all' δίκτυο, καθώς ο νευρώνας με τη μεγαλύτερη είσοδο 'κερδίζει' όλη τη διαθέσιμη δραστηριότητα.

Αν η δραστηριότητα του δικτύου αναπαρασταθεί με βάση τους νευρώνες, τότε μια αρχική κατάσταση στην οποία ξεχωρίζει κάποιος νευρώνας θα εξελιχθεί όπως φαίνεται στο Σχήμα 5.2.



Σχήμα 5. 2: Χρονική εξέλιξη της ενεργοποίησης νευρώνων υπό την επίδραση ανταγωνιστικών δυναμικών

Ας θεωρήσουμε ένα σύνολο διανυσμάτων εισόδου, του οποίου τα διανύσματα έχουν όλα το ίδιο μέτρο και ας υποθέσουμε, χωρίς βλάβη της γενικότητας, ότι αυτό είναι μονάδα.

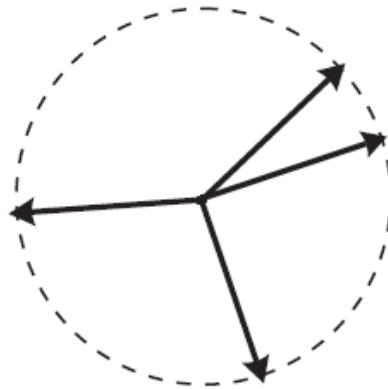
Γνωρίζουμε ότι το μέτρο $\|\mathbf{x}\|$ ενός διανύσματος \mathbf{x} δίνεται από τη σχέση:

$$\|\mathbf{x}\| = \sum_i x_i^2 \quad (5.2)$$

Ένα σύνολο διανυσμάτων, για τα οποία ισχύει $\|\mathbf{x}\|=1 \forall \mathbf{x}$, λέμε ότι είναι κανονικοποιημένο. Δηλαδή

$$\sum_i x_i^2 = 1 \quad (5.3)$$

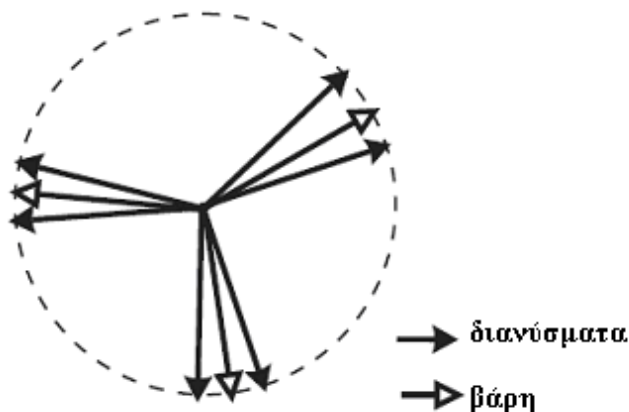
Εφόσον όλα τα διανύσματα έχουν μοναδιαίο μέτρο, μπορούν να αναπαρασταθούν σαν ακτίνες μιας μοναδιαίας σφαίρας.



Σχήμα 5.3: Μοναδιαία Σφαίρα

Ας υποθέσουμε τώρα ότι ένα ανταγωνιστικό επίπεδο έχει και τα διανύσματα των βαρών του κανονικοποιημένα με βάση την (5.3). Τότε και αυτά τα διανύσματα μπορούν να αναπαρασταθούν στην ίδια σφαίρα.

Αυτό που χρειάζεται το δίκτυο για να κωδικοποιήσει το σύνολο εκπαίδευσης είναι να ευθυγραμμιστούν τα διανύσματα βαρών με όλες τις ομάδες που βρίσκονται σ' αυτό το σύνολο και κάθε ομάδα να αντιπροσωπεύεται από τουλάχιστον ένα νευρώνα. Τότε, όταν ένα διάνυσμα παρουσιάζεται στο δίκτυο ως είσοδος θα υπάρχει ένας νευρώνας, ή μια ομάδα νευρώνων που θα έχουν τη μέγιστη απόκριση στην είσοδο και που θα αποκρίνονται έτσι μόνο όταν το συγκεκριμένο διάνυσμα εμφανίζεται ως είσοδος.



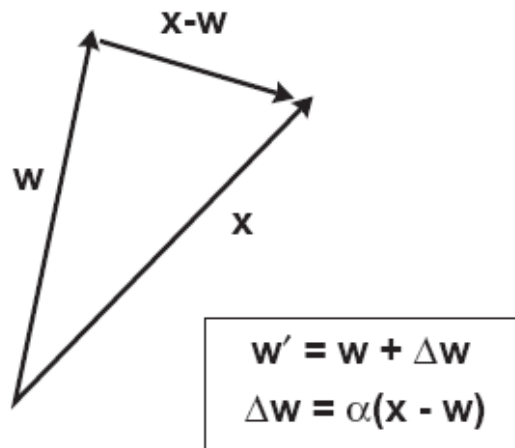
Σχήμα 5.4: Ευθυγράμμιση βαρών και διανυσμάτων εισόδου

Αν το δίκτυο μπορέσει να μάθει να διαμορφώνει τα διανύσματα βαρών με αυτόν τον τρόπο, χωρίς να του αναφέρεται ρητά η ύπαρξη ομάδων από ομοειδείς εισόδους, τότε

λέγεται ότι υποβάλλεται σε διαδικασία *αυτό-οργανούμενης* ή αλλιώς *μη επιβλεπόμενης* μάθησης. Αυτό είναι κάτι τελείως διαφορετικό από τα δίκτυα που εκπαιδεύονται με τον κανόνα δέλτα ή τον αλγόριθμο Backpropagation, όπου έπρεπε να δίνεται και η επιθυμητή έξοδος.

Για να μπορέσει να επιτευχθεί αυτός ο στόχος, τα διανύσματα των βαρών θα πρέπει να περιστρέφονται στη σφαίρα, ώστε να ευθυγραμμιστούν με το σύνολο εκπαίδευσης. Αυτό μπορεί γίνει με σταδιακό και αποτελεσματικό τρόπο, μετακινώντας ελαφρά το διάνυσμα βάρους που είναι πιο κοντά (κατά γωνία) στο τρέχον διάνυσμα εισόδου προς το διάνυσμα αυτό. Ο νευρώνας k , με το κοντινότερο διάνυσμα, είναι αυτός που δίνει τη μεγαλύτερη διέγερση εισόδου s , δηλαδή το μεγαλύτερο εσωτερικό γινόμενο των διανυσμάτων βαρών και εισόδων. Όπως φαίνεται παρακάτω, το διάνυσμα βαρών του νευρώνα k μπορεί να πλησιάσει την είσοδο, αν μεταβληθεί κατά

$$\Delta \bar{w} = \alpha(\bar{x} - \bar{w}) \quad (5.4)$$



Θα μπορούσαμε εποπτικά με υπολογιστή να αποφασίσουμε ποιος νευρώνας έχει τη μεγαλύτερη διέγερση s , αλλά είναι πιο ικανοποιητικό το δίκτυο να μπορεί να αποφασίζει μόνο του. Στο σημείο αυτό εισάγονται τα ανταγωνιστικά δυναμικά. Ας υποθέσουμε ότι πρόκειται για ένα ‘winner-takes-all’ δίκτυο, οπότε ο νευρώνας-νικητής έχει τιμή 1 και όλοι οι άλλοι νευρώνες έχουν τιμές κοντά στο μηδέν. Αφού αφήσουμε το δίκτυο να φτάσει σε ισορροπία, επιβάλλουμε τον κανόνα:

$$\Delta \bar{w} = \alpha(\bar{x} - \bar{w})y \quad (5.5)$$

σε όλο το δίκτυο. Έτσι, θα υπάρχει ένας και μόνο νευρώνας (αυτός του οποίου το εσωτερικό γινόμενο s ήταν μεγαλύτερο) ο οποίος θα έχει $y=1$ και για τον οποίο η

μεταβολή στο βάρος θα δίνεται από τη σχέση (5.5). Όλοι οι άλλοι νευρώνες θα έχουν $y = 0$ κι έτσι και οι μεταβολές στα βάρη τους θα είναι επίσης μηδέν. Τα στάδια της μάθησης θα είναι:

1. Παρουσίαση των διανυσμάτων εισόδου στο δίκτυο και υπολογισμός του s για κάθε νευρώνα.
2. Ενημέρωση του δικτύου (στην πράξη αυτό γίνεται σε διακριτά βήματα) με βάση τη σχέση (5.4) για πεπερασμένο χρόνο, ή μέχρι να φτάσει σε ισορροπία.
3. Εκπαίδευση όλων των νευρώνων με βάση την (5.5).

Υπάρχουν ορισμένα σημεία στον παραπάνω κανόνα μάθησης που αξίζει να αναλυθούν. Πρώτον, αν τα βάρη είναι αρχικά κανονικοποιημένα με βάση την (5.3) και τα διανύσματα εισόδου είναι κι αυτά κανονικοποιημένα με τον ίδιο τρόπο, τότε η κανονικοποίηση των βαρών διατηρείται στον κανόνα μάθησης. Δηλαδή η μεταβολή στο διάνυσμα \bar{w} δίνεται από το άθροισμα των μεταβολών των συνιστωσών του

$$\sum_i \Delta w_i = \alpha y \left(\sum_i x_i - \sum_i w_i \right) \quad (5.6)$$

και καθένα απ' τα αθροίσματα στην παρένθεση είναι 1 (λόγω κανονικοποίησης), οπότε το αριστερό μέλος είναι μηδέν. Τα αποτελέσματα της κανονικοποίησης αγνοήθηκαν μέχρι στιγμής, για να διευκρινιστούν τα βασικά. Υποθέσαμε ότι το εσωτερικό γινόμενο s παρέχει μια ένδειξη για τη γωνιακή απόσταση μεταξύ των βαρών και των διανυσμάτων εισόδου. Αυτό είναι αλήθεια μέχρι ένα σημείο, αλλά ας θυμηθούμε ότι το εσωτερικό γινόμενο περικλείει και τα γινόμενα των μέτρων των διανυσμάτων. Αν είτε τα διανύσματα εισόδου, είτε τα διανύσματα βαρών είναι μεγάλα, τότε το s μπορεί επίσης να είναι μεγάλο, όχι όμως σαν αποτέλεσμα γωνιακής εγγύτητας (δηλαδή ευθυγράμμισης των διανυσμάτων), αλλά σαν αποτέλεσμα του μεγέθους τους. Αυτό που χρειάζεται είναι ένα μέτρο για τη συγγραμμικότητα των διανυσμάτων, που να μην εξαρτάται από τα μήκη των διανυσμάτων και η κανονικοποίηση είναι ένας τρόπος να επιτευχθεί αυτό.

Δεύτερον, ο κανόνας μάθησης μπορεί να επεκταθεί και για το άθροισμα των δύο όρων:

$$\Delta \bar{w} = \alpha \bar{x}y - \alpha \bar{w}y \quad (5.7)$$

Ο πρώτος από τους όρους μοιάζει με παράγοντα Hebb, ενώ ο δεύτερος επιδρά στο βάρος (weight decay). Έτσι η ανταγωνιστική μάθηση μπορεί να θεωρηθεί ως συνδυασμός της

μάθησης Hebb και ενός κανόνα προσαρμογής βαρών (weight decay) που εγγυάται την κανονικοποίηση [18]. Χαρακτηριστικά παραδείγματα δικτύων που βασίζονται στην ανταγωνιστική μάθηση είναι το δίκτυο *LVQ* και το δίκτυο *SOFM*

5.3 Μάθηση Hebb

Η μάθηση Hebb είναι ο δεύτερος τύπος αυτό-οργανούμενης μάθησης πέρα από την ανταγωνιστική μάθηση. Ένα αυτό-οργανούμενο ΤΝΔ μπορεί να έχει διάφορες δομές. Οι δύο πιο διαδεδομένες απ' αυτές είναι οι εξής:

- Το ΤΝΔ αποτελείται από ένα στρώμα εισόδου (πηγής) και ένα στρώμα εξόδου (αναπαράστασης) με συνδέσεις προσοτροφοδότησης από την είσοδο στην έξοδο και πλάγιες συνδέσεις μεταξύ των νευρώνων του στρώματος εξόδου.
- Ένα πολυστρωματικό ΤΝΔ στο οποίο η αυτό-οργάνωση προχωράει στρώμα με στρώμα.

Η μάθηση και στις δύο περιπτώσεις συνίσταται στην επαναληπτική προσαρμογή (ανανέωση) των συναπτικών βαρών όλων των συνδέσεων του δικτύου σε απάντηση προτύπων εισόδου, σύμφωνα με δεδομένους κανόνες, μέχρις ότου αναπτυχθεί ένα τελικό ΤΝΔ. Η ανατροφοδότηση που λαμβάνει χώρα (μεταξύ των μεταβολών στα συναπτικά βάρη και των μεταβολών στα πρότυπα εισόδου) για να επιτευχθεί αυτό-οργάνωση (και όχι ευστάθεια) πρέπει να είναι θετική. Οι τρεις θεμελιακές αρχές τις οποίες ακολουθούν τα αυτό-οργανούμενα ΤΝΔ είναι οι εξής:

1. Οι μεταβολές ανανέωσης των συναπτικών βαρών τείνουν να αυτό-ενισχύονται.
2. Οι περιορισμοί στους διατιθέμενους πόρους οδηγεί σε ανταγωνισμό ανάμεσα στις συνάψεις (επιλέγονται οι συνάψεις που αναπτύσσονται πιο ισχυρά σε βάρος των υπολοίπων).
3. Οι τροποποιήσεις (μεταβολές ανανέωσης) των συναπτικών βαρών τείνουν να συνεργάζονται.

Μετά τα παραπάνω για την αυτό-οργανούμενη μάθηση, εξετάζεται η μάθηση Hebb στην οποία όταν δύο νευρώνες πυροδοτούνται ταυτόχρονα (δηλαδή έχουν ισχυρές αποκρίσεις), το βάρος σύνδεσής τους αυξάνει. Η αύξηση αυτή του συναπτικού βάρους μεταξύ των νευρώνων είναι ανάλογη προς τη συχνότητα με την οποία αυτοί

πυροδοτούνται ταυτόχρονα. Ανάμεσα στους διάφορους μαθηματικούς τύπους έκφρασης της παραπάνω αρχής (υπόθεσης Hebb) ο πιο απλός είναι ο ακόλουθος:

$$\Delta \bar{w} = \alpha \bar{x}y .$$

Τα πιο συχνά εμφανιζόμενα πρότυπα έχουν τη μεγαλύτερη επίδραση στα συναπτικά βάρη και τελικά οδηγούν το ΤΝΔ στην παραγωγή των μεγαλύτερων εξόδων. Εφαρμόζοντας τον κανόνα Hebb προκαλείται αδιάκοπη αύξηση των βαρών, οπότε στην πράξη χρειάζεται να περιορίσουμε την αύξηση αυτή (βλ. παράγραφο 5.2) [14].

5.4 Αυτο-οργανούμενοι χάρτες Kohonen – Το δίκτυο SOM

Ένας τύπος ΤΝΔ που βασίζεται στην ανταγωνιστική μάθηση είναι το δίκτυο SOM (self-organising map) ή δίκτυο SOFM (self-organising feature map). Το δίκτυο αυτό προτάθηκε απ' τον Kohonen στα μέσα της δεκαετίας του 1980, αποτελεί έναν *αυτο-οργανούμενο χάρτη χαρακτηριστικών* και βασίζεται στην έννοια της ανταγωνιστικής μάθησης. Δηλαδή οι νευρώνες του ανταγωνιστικού επιπέδου 'μάχονται' μεταξύ τους για το ποιος θα επικρατήσει για κάποιο συγκεκριμένο ερέθισμα (πρότυπο εισόδου) [19].

Ειδικότερα οι νευρώνες τοποθετούνται στη θέση των κόμβων ενός πλέγματος μίας ή δύο διαστάσεων. Μπορεί να δημιουργηθούν και με περισσότερες διαστάσεις, αλλά δεν συνηθίζεται. Στα πλαίσια της ανταγωνιστικής μάθησης οι νευρώνες επιλέγουν ορισμένα πρότυπα εισόδου και ρυθμίζουν τα αντίστοιχα βάρη τους, με κύριο στόχο να διαταχθούν με τέτοιο τρόπο, ώστε η θέση τους στο πλέγμα να συσχετίζεται με συγκεκριμένα χαρακτηριστικά των προτύπων εισόδου δημιουργώντας ένα είδος τοπογραφικού χάρτη των τελευταίων. Εξαιτίας τούτου προέκυψε και ο όρος του αυτο-οργανούμενου χάρτη - Self Organizing Map.

Ο αυτο-οργανούμενος χάρτης προήλθε από το νευρο-βιολογικό τρόπο λειτουργίας του φλοιού του εγκεφάλου. Συγκεκριμένα στηρίζεται στην αρχή ότι η χωρική θέση ενός νευρώνα εξόδου σε ένα τοπογραφικό χάρτη αντιστοιχεί σε έναν τομέα ή σε ένα χαρακτηριστικό των δεδομένων εισόδου. Οπότε μπορούν να σχηματιστούν δύο διαφορετικά μοντέλα απεικόνισης χαρακτηριστικών. Στο πρώτο μοντέλο - κατά το *Willshaw* - διατάσσονται οι νευρώνες εξόδου σ' ένα πλέγμα δύο διαστάσεων εξασφαλίζοντας ένα σύνολο γειτόνων. Ειδικότερα υπάρχουν δύο ξεχωριστά δισδιάστατα πλέγματα νευρώνων που συνδέονται μεταξύ τους και το ένα προβάλλεται πάνω στο

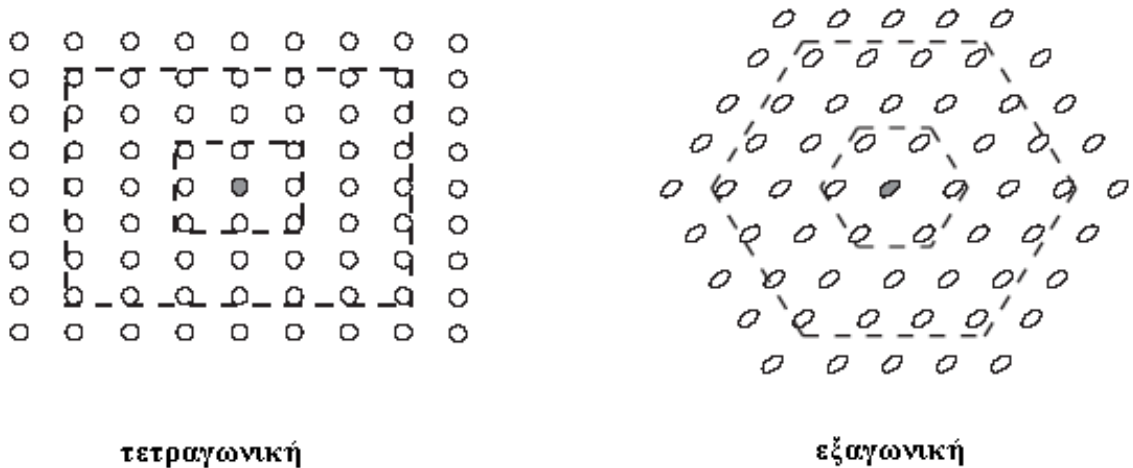
άλλο. Το ένα αποτελεί το πλέγμα εισόδου που έχει ένα διεγερτικό μηχανισμό μικρής εμβέλειας και ένα μηχανισμό αναστολής μεγάλης εμβέλειας και το άλλο είναι το πλέγμα εξόδου. Οι μηχανισμοί αυτοί είναι τοπικής εμβέλειας και επιτυγχάνουν οι νευρώνες του πλέγματος εξόδου να μη λειτουργούν ως «νικητής που τα παίρνει όλα», αλλά να ενεργοποιούνται λίγοι σε κάθε χρονική στιγμή. Παράλληλα για να μην υπάρξει ασύμμετρη αύξηση κάποιων βαρών που θα οδηγήσει σε αστάθεια το δίκτυο, υπάρχουν αντίστοιχοι περιορισμοί των βαρών μέσω ανώτατων ορίων. Επομένως κάποια βάρη του ίδιου νευρώνα αυξάνονται και κάποια άλλα μειώνονται. Με αυτόν τον τρόπο παράγεται μία τοπολογικά διατεταγμένη απεικόνιση ανάμεσα στην είσοδο και στην έξοδο των ίδιων διαστάσεων, εξηγώντας αρκετά ικανοποιητικά τη σύνδεση του αμφιβληστροειδή με το οπτικό τμήμα του φλοιού του εγκεφάλου· αυτό αποτελεί μία από τις σημαντικότερες εφαρμογές του.

Το δεύτερο μοντέλο -που προτάθηκε από τον *Kohonen*- δεν εξηγεί νευρο-βιολογικές λειτουργίες, αλλά ενσωματώνει τα βασικά χαρακτηριστικά των υπολογιστικών χαρτών του εγκεφάλου έχοντας τη δυνατότητα να διενεργεί συμπίεση των δεδομένων. Ουσιαστικά ανήκει στους αλγόριθμους του διανυσματικού κβαντιστή και παρέχει μία τοπολογική απεικόνιση που τοποθετεί με τον καλύτερο τρόπο ένα σταθερό αριθμό από διανύσματα σ' ένα χώρο μικρότερης διάστασης πραγματοποιώντας συμπίεση.

Στην παρούσα περίπτωση θα χρησιμοποιηθεί το μοντέλο του *Kohonen*, οπότε ο βασικός σκοπός του αυτό-οργανωμένου δικτύου είναι η απεικόνιση ενός προτύπου εισόδου διάστασης d σ' ένα διακριτό χάρτη μίας ή δύο διαστάσεων. Κάθε νευρώνας του πλέγματος συνδέεται πλήρως με όλους τους κόμβους του επιπέδου εισόδου.

Ο αλγόριθμος ξεκινά αρχικοποιώντας κατάλληλα τα βάρη του δικτύου που συνδέουν τους νευρώνες εισόδου με τους νευρώνες του χάρτη - είτε με τυχαίους αριθμούς, είτε με συγκεκριμένη διάταξη. Έπειτα ακολουθεί η φάση του *ανταγωνισμού*, όπου για κάθε πρότυπο εισόδου οι νευρώνες του πλέγματος υπολογίζουν την αντίστοιχη τιμή της συνάρτησης ανταγωνισμού των νευρώνων κερδίζοντας εκείνος ο νευρώνας με τη μεγαλύτερη τιμή. Στη συνέχεια υλοποιείται η φάση της *συνεργασίας*, όπου ο νευρώνας νικητής καθορίζει τη χωρική θέση μίας τοπολογικής γειτονιάς παρέχοντας τη βάση για τη συνεργασία μεταξύ των γειτονικών νευρώνων. Μετά ακολουθεί η φάση της *προσαρμογής των βαρών*, όπου ο μηχανισμός αυτός επιτρέπει στους νευρώνες που ανήκουν στη

νικήτρια γειτονιά να προσαρμόσουν τις τιμές των βαρών τους, ώστε να ενισχυθεί η απόκριση του νευρώνα-νικητή σε μία επόμενη εφαρμογή ενός παρόμοιου προτύπου εκπαίδευσης.



Σχήμα 5. 3: Οι “γειτονιές” του Kohonen

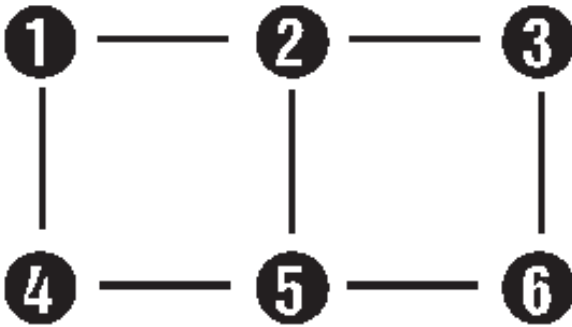
5.4.1 Ο αλγόριθμος

Στην παρουσίαση κάθε νέου διανύσματος εισόδου λαμβάνει χώρα η παρακάτω ακολουθία βημάτων [18]:

- Εύρεση του νευρώνα k του οποίου το διάνυσμα βαρών είναι πιο κοντά στο τρέχον διάνυσμα εισόδων.
- Εκπαίδευση του νευρώνα k και όλων των νευρώνων της γειτονιάς του.
- Ελαφριά μείωση του ρυθμού μάθησης.
- Μετά από M κύκλους, μείωση του μεγέθους της γειτονιάς.

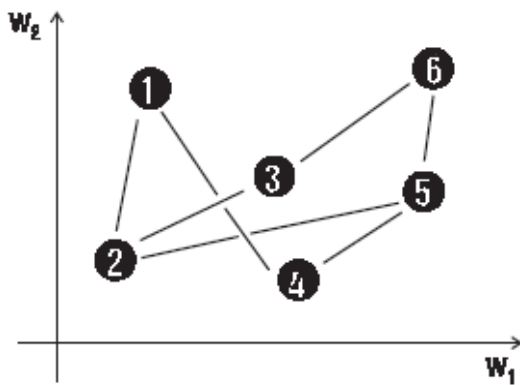
5.4.2 Ένα παράδειγμα

Η αυτο-οργάνωση ενός δικτύου Kohonen είναι δυνατό να εξηγηθεί γραφικά, χρησιμοποιώντας ένα δίκτυο όπου το διάστημα εισόδου έχει μόνο δύο στοιχεία. Θεωρώντας ένα δίκτυο με έξι νευρώνες σε ορθογώνιο πλέγμα.



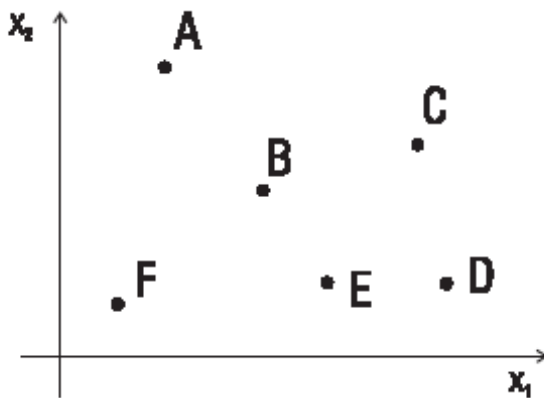
Σχήμα 5. 4: Δίκτυο έξι κόμβων σε πλέγμα

Μια άλλη αναπαράσταση του δικτύου αυτού είναι στο χώρο των βαρών. Εφόσον υπάρχουν μόνο δύο βάρη, μπορεί να σχεδιαστεί όπως φαίνεται στο Σχήμα 5.5.



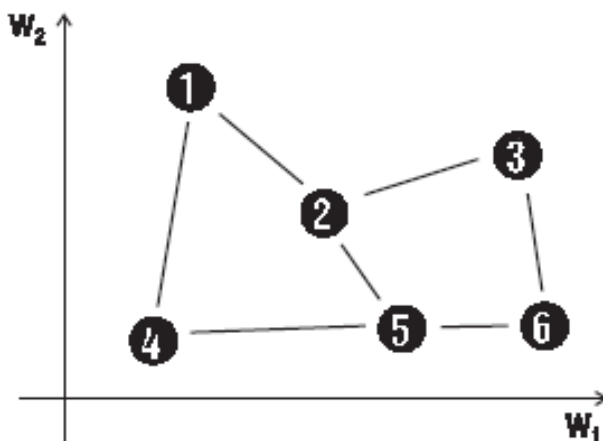
Σχήμα 5. 5: Τυχαία βάρη

Οι γραμμές στο Σχήμα 5.5 σχεδιάστηκαν για να συνδέουν νευρώνες που γειτνιάζουν (όπως φαίνεται απ' το Σχήμα 5.4). Έστω ότι υπάρχουν έξι διανύσματα εισόδου που μπορούν να αναπαρασταθούν στο χώρο των προτύπων, όπως φαίνεται παρακάτω:



Σχήμα 5. 6: Διανύσματα εισόδου για το δίκτυο Kohonen 6 νευρώνων

Σε ένα καλά εκπαιδευμένο δίκτυο που έχει αναπτύξει έναν τοπογραφικό χάρτη, το διάγραμμα για τα βάρη θα έπρεπε να έχει την ίδια τοπολογία με αυτό για τα διανύσματα και να αντανakλά τις ιδιότητες του συνόλου εκπαίδευσης, όπως φαίνεται στο Σχήμα 5.7.



Σχήμα 5. 7: Εκπαιδευμένο δίκτυο

5.5 Αυτοσυσχετιστικές μνήμες – Δίκτυα Hopfield

Η ανάκληση ενός γεγονότος σε μία τυχαία χρονική στιγμή προκαλείται από τη συσχέτιση αυτού του γεγονότος με κάποιο ερέθισμα. Για παράδειγμα κάποιος αναφέρει ένα γεγονός και εμείς αμέσως ανακαλούμε στο μυαλό μας μία αλυσίδα από στιγμές ή εμπειρίες σχετικές με το συγκεκριμένο γεγονός. Μπορεί επίσης να δούμε μία εικόνα από ένα σπίτι που επισκεφτήκαμε σε μικρή ηλικία, και αυτή η εικόνα ανακαλεί στη μνήμη μας στιγμές εκείνης της περιόδου.

Σε μια πιο συνηθισμένη περίπτωση, αλλά της ίδιας κατηγορίας, τυχαίνει κάποιες φορές να καλούμαστε να αναγνωρίσουμε μερικώς κατεστραμμένα γράμματα ή να τα δούμε μέσα από ένα παράθυρο ενώ βρέχει -παρουσία θορύβου. Τα κοινό χαρακτηριστικό αυτών των παραδειγμάτων είναι το εξής: Υπάρχει μία συλλογή πληροφοριών η οποία είναι με κάποιο τρόπο διατεταγμένη, αλληλοσυσχετιζόμενη και αποθηκευμένη στη μνήμη. Μπορεί να θεωρηθεί ότι τα δεδομένα αυτά σχηματίζουν ένα αποθηκευμένο πρότυπο. Στην περίπτωση της αναγνώρισης χαρακτήρων, τα διάφορα τμήματα τους καθορίζονται από την κανονική παράσταση του γράμματος. Όταν παρουσιάζεται τμήμα του γράμματος, τότε ανακαλείται ή συσχετίζεται το υπόλοιπο αυτού. Δεν έχει σημασία ποιο τμήμα παρουσιάζεται για να ανακτηθεί ολόκληρο το πρότυπο.

Οι συμβατικοί υπολογιστές μπορούν να υλοποιήσουν μια τέτοια εργασία με πολύ περιορισμένο τρόπο. Το λογισμικό που συνήθως χρησιμοποιείται για αποθήκευση, αναζήτηση και ανάκτηση είναι οι βάσεις δεδομένων. Σε τέτοια προγράμματα η αναζήτηση γίνεται με τη χρήση ενός κλειδιού. Για παράδειγμα έχουμε έναν κατάλογο από βιβλία σε μία βιβλιοθήκη, όπου αποθηκεύουμε τους τίτλους, τα ονόματα των συγγραφέων και διάφορες άλλες παρόμοιες πληροφορίες. Η αναζήτηση σε έναν τέτοιο κατάλογο θα γίνει για ένα συγκεκριμένο πεδίο από αυτά που αναφέραμε και χρησιμοποιώντας ολόκληρο το όνομα-κλειδί που αναζητούμε. Αν υποθέσουμε όμως ότι γνωρίζουμε ένα τμήμα μονάχα του πραγματικού πεδίου της βάσης, π.χ. «Προσομο» αντί για «Προσομοίωση», τότε δεν υπάρχει τρόπος για μία τέτοια βάση δεδομένων να εκτελέσει αυτήν την αναζήτηση, διότι όχι μόνο δεν γνωρίζει σε ποιο πεδίο αντιστοιχεί, αλλά ούτε και σε ποια θέση του πεδίου να ψάξει. Οι πληροφορίες που η βάση απαιτεί είναι πολύ συγκεκριμένες σε αντίθεση με τα παραδείγματα μας, όπου ανακαλούμε πρότυπα με ελλείψεις πληροφορίες.

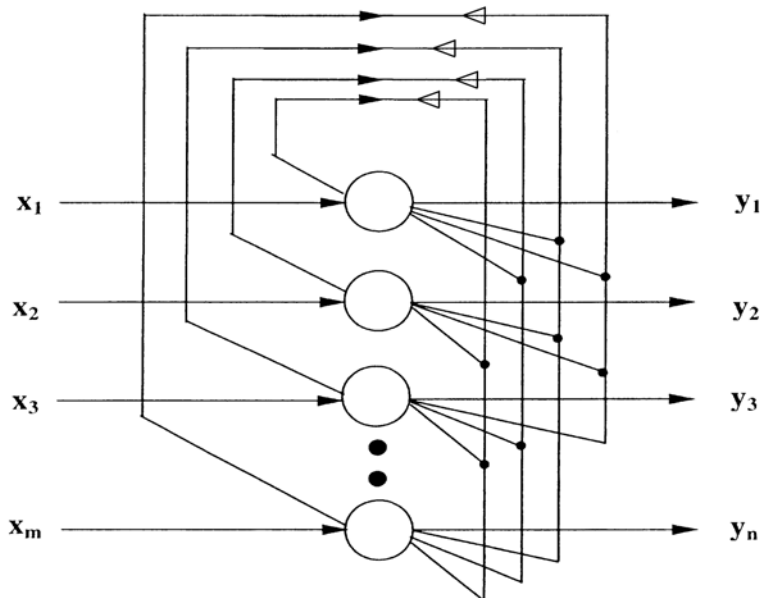
Το φυσικό ανάλογο της αυτό-συσχετιστικής μνήμης βρίσκεται στο παράδειγμα μίας σφαίρας εντός μίας κοιλότητας, όπου η σφαίρα θα ισορροπήσει τελικά στον πυθμένα, που αποτελεί το σημείο ευσταθούς ισορροπίας. Διαφορετικά θα μπορούσαμε να ισχυριστούμε ότι η σφαίρα καταλήγει σε εκείνο το σημείο, επειδή «θυμάται» που είναι ο πυθμένας. Αν χρησιμοποιήσουμε μια κυματοειδή επιφάνεια αντί για μία κοιλότητα, μπορούμε να αποθηκεύσουμε πολλές «μνήμες». Αν η σφαίρα ξεκινήσει από κάποιο τυχαίο σημείο, τότε θα καταλήξει στο πλησιέστερο τοπικό κοίλο -τοπικό ελάχιστο.

Επομένως ανακαλεί το πλησιέστερο αποθηκευμένο πρότυπο. Αυτό το σημείο είναι και ένα τοπικό ελάχιστο της ενέργειας του συστήματος.

Υπάρχουν λοιπόν δύο αλληλένδετοι τρόποι αντιμετώπισης του θέματος. Σύμφωνα με τον πρώτο, λέμε ότι το σύστημα οδηγείται σε ένα ενεργειακό ελάχιστο, ενώ σύμφωνα με το δεύτερο, αποθηκεύουμε ένα σύνολο προτύπων και ανακαλούμε αυτό που βρίσκεται πλησιέστερα στην αρχική θέση. Αν θελήσουμε να κατασκευάσουμε ένα δίκτυο που θα μιμείται αυτήν τη συμπεριφορά, τότε θα πρέπει να συμπεριλάβουμε τα παρακάτω σημεία-κλειδιά.

1. Το δίκτυο περιγράφεται πλήρως από ένα διάνυσμα κατάστασης $\vec{x} = (x_1, x_2, \dots, x_d)^T$.
2. Υπάρχει ένα σύνολο ευσταθών καταστάσεων $\vec{x}^1, \vec{x}^2, \dots, \vec{x}^m$. Αυτές αντιστοιχούν στα αποθηκευμένα πρότυπα, τα οποία αποτελούν τα τοπικά ελάχιστα της κυματοειδούς επιφάνειας.
3. Το σύστημα ξεκινάει από μία τυχαία αρχική κατάσταση -ερέθισμα και καταλήγει σε μία από τις ευσταθείς καταστάσεις. Η διαδικασία αυτή συνοδεύεται από μείωση της ενέργειας E του συστήματος και αντιστοιχεί στην ανάκληση της μνήμης.

Η αυτοσυσχετιστική μνήμη υλοποιείται συνήθως μέσω ανατροφοδοτούμενων δικτύων, όπως είναι τα δίκτυα Hopfield, που μία μορφή του φαίνεται στο ακόλουθο σχήμα.



Σχήμα 5. 8: Δίκτυο Hopfield

Κάθε κόμβος συνδέεται με όλους τους άλλους εκτός του εαυτού του με αμφίδρομες συνδέσεις και τα βάρη των συνδέσεων είναι *συμμετρικά*. Δηλαδή για δύο οποιουσδήποτε κόμβους i και j τα βάρη είναι ίσα και προς τις δύο κατευθύνσεις ($w_{ij} = w_{ji}$ & $w_{ii} = 0$). Επιπλέον οι πολώσεις θ_i θεωρούμε ότι είναι μηδενικές. Μπορεί να παρατηρηθεί ότι η ροή της πληροφορίας δεν έχει μία μόνο κατεύθυνση, όπως συμβαίνει με τα πολυεπίπεδα νευρωνικά δίκτυα πρόσθιας τροφοδότησης. Είναι πιθανό μία πληροφορία να επιστρέψει στον αρχικό κόμβο μέσω κάποιου άλλου. Αυτό το φαινόμενο ονομάζεται *ανάδραση* και τα δίκτυα *επαναληπτικά* ή *ανατροφοδοτούμενα* (recurrent).

Η κατάσταση του δικτύου οποιαδήποτε χρονική στιγμή δίνεται από το διάνυσμα των εξόδων των κόμβων. Έστω τώρα ότι ξεκινάμε με κάποια αρχική κατάσταση και διαλέγουμε τυχαία με ίση πιθανότητα έναν οποιοδήποτε κόμβο του δικτύου τον οποίο και ενημερώνουμε. Δηλαδή υπολογίζουμε την ενεργοποίηση του κόμβου - εσωτερικό γινόμενο εισόδων επί βάρη συν πόλωση - και δίνουμε ως έξοδο «1», αν αυτό είναι μεγαλύτερο από μηδέν και «-1» διαφορετικά. Το δίκτυο τώρα βρίσκεται είτε στην ίδια κατάσταση από την οποία ξεκίνησε, είτε σε μία άλλη με απόσταση *Hamming 1* από την αρχική. Στη συνέχεια επαναλαμβάνουμε την ίδια διαδικασία για πολλά βήματα. Με τον τρόπο αυτό, για κάθε κατάσταση, δεδομένου του δικτύου -τιμές βαρών-, μπορούμε να υπολογίζουμε τις επόμενες. Σε ένα τέτοιο δίκτυο μπορούν να αποθηκευτούν πρότυπα $\bar{x} = (x_1, x_2, \dots, x_d)^T$ όπου $x_i \in \{-1, 1\}$. Τα πρότυπα αυτά αποθηκεύονται ως *καταστάσεις ευσταθείας-ισορροπίας* του δικτύου, με κατάλληλη επιλογή των τιμών των βαρών w_{ij} . Υπάρχουν δύο τρόποι λειτουργίας του δικτύου *Hopfield*: *σύγχρονη λειτουργία* και *ασύγχρονη λειτουργία*. Και στους δύο τρόπους λειτουργίας η ενημέρωση ενός κόμβου i τη χρονική στιγμή t γίνεται σύμφωνα με τη σχέση:

$$y_i(t+1) = \begin{cases} 1 & , u_i(t) \geq 0 \\ -1 & , u_i(t) < 0 \end{cases}$$

$$u_i(t) = \sum_{j=1}^n w_{ij} \cdot y_j(t) + \theta_i$$

Μια κατάσταση $\bar{y}(t) = (y_1(t), y_2(t), \dots, y_n(t))^T$ αποτελεί *κατάσταση ισορροπίας* (ευστάθειας) ενός δικτύου *Hopfield*, όταν $y_i(t+1) = y_i(t)$ για κάθε $i = 1, \dots, n$. Η διαφορά

της ασύγχρονης και της σύγχρονης λειτουργίας βρίσκεται στο γεγονός ότι στην πρώτη ενημερώνονται διαδοχικά όλοι οι κόμβοι με τυχαία επιλογή, ενώ στη δεύτερη ενημερώνονται ταυτόχρονα όλοι οι κόμβοι του δικτύου υπολογίζοντας συγχρόνως όλα τα $u_i(t)$. Με την πρώτη μέθοδο καταλήγουμε πάντα σε ευσταθή κατάσταση ισορροπίας, καθώς ελαχιστοποιείται η συνάρτηση ενέργειας, η οποία δεν μπορεί να μειώνεται διαρκώς. Από την άλλη η τεχνική της σύγχρονης ενημέρωσης δεν μπορεί να εγγυηθεί τη σύγκλιση του δικτύου σε κάποια κατάσταση ισορροπίας. Συγκεκριμένα αποδεικνύεται ότι το δίκτυο είτε θα καταλήξει σε κατάσταση ισορροπίας, είτε θα εμπλακεί σε κύκλο μήκους δύο, δηλαδή θα παλινδρομεί συνεχώς μεταξύ δύο καταστάσεων. Η εύρεση των τιμών των βαρών στα δίκτυα Hopfield γίνεται με βάση τον κανόνα του Hebb [19].

5.6 Σύνοψη

Στο παρόν κεφάλαιο έγινε αναφορά στα Τεχνητά Νευρωνικά Δίκτυα των οποίων η εκπαίδευση γίνεται χωρίς επίβλεψη. Περιγράφηκαν οι βασικές αρχές της ανταγωνιστικής μάθησης και της μάθησης Hebb, ενώ παρουσιάστηκαν και τα δίκτυα Kohonen και Hopfield. Παρότι στη συγκεκριμένη εργασία δεν θα χρησιμοποιηθούν δίκτυα χωρίς επίβλεψη, η αναφορά έγινε για λόγους πληρότητας.

Στο επόμενο κεφάλαιο θα γίνει περιγραφή του προγράμματος που αναπτύχθηκε για την κατασκευή ενός ΤΝΔ με επίβλεψη, για την εκτίμηση της κρίσιμης τάσης σε μονωτήρες υψηλών τάσεων.

Κεφάλαιο 6

Ανάπτυξη Αλγορίθμου για την εκτίμηση της κρίσιμης τάσης σε μονωτήρες ΥΤ

6.1 Περιγραφή Προβλήματος

Σε προηγούμενα κεφάλαια παρουσιάστηκε ένα απ' τα σημαντικότερα προβλήματα που παρουσιάζονται σε μονωτήρες ΥΤ όταν αυτοί λειτουργούν υπό συνθήκες ρύπανσης (βιομηχανικής ή θαλάσσιας): η εμφάνιση του φαινομένου της υπερπήδησης, δηλαδή η γεφύρωση τόξου και η διέλευση ρεύματος από τμήμα του μονωτήρα ή σε όλο το μήκος του μονωτήρα. Το φαινόμενο αυτό μπορεί να οδηγήσει το μονωτήρα σε υπερπήδηση, δηλαδή στην πλήρη καταστροφή του. Επομένως είναι πολύ σημαντικό να γνωρίζουμε ποιες είναι εκείνες οι παράμετροι που επηρεάζουν το φαινόμενο της υπερπήδησης και για ποιες τιμές των παραμέτρων αυτών είναι αυξημένος ο κίνδυνος να συμβεί κάτι τέτοιο.

Ενδεχομένως ο τρόπος με τον οποίο καθεμιά απ' τις παραμέτρους αυτές συμβάλλει στην εμφάνιση του φαινομένου να μην είναι γνωστός, ή η πολυπλοκότητα του φαινομένου να εισάγει αβεβαιότητα για την οποία θα πρέπει να πάρουμε κατάλληλες προσεγγίσεις. Στο σημείο αυτό τα Τεχνητά Νευρωνικά Δίκτυα αποδεικνύονται ένα πολύ χρήσιμο εργαλείο για την εκτίμηση και πρόβλεψη της τάσης διάσπασης ενός μονωτήρα, όταν αυτός βρίσκεται υπό την επίδραση των παραπάνω παραγόντων.

Όπως έχει ήδη αναφερθεί, τα ΤΝΔ έχουν την ικανότητα να “μαθαίνουν” αυτόματα προσεγγιστικές σχέσεις μεταξύ εισόδων και εξόδων, χωρίς να υποτάσσονται από το μέγεθος και την πολυπλοκότητα του προβλήματος. Οι προσεγγιστικές αυτές σχέσεις βασίζονται σε πραγματικά δεδομένα που χρησιμοποιούνται για την εκπαίδευση του ΤΝΔ (π.χ. δεδομένα από πειραματικές μετρήσεις). Αναφέρθηκαν επίσης οι δύο βασικές κατηγορίες ΤΝΔ με βάση τον κανόνα μάθησης: τα δίκτυα με επίβλεψη και τα δίκτυα χωρίς επίβλεψη. Στην παρούσα διπλωματική θα χρησιμοποιηθεί δίκτυο επιβλεπόμενης μάθησης με τον κανόνα αναστροφής διάδοσης για τη διόρθωση του σφάλματος.

Η εκμάθηση με επίβλεψη είναι μία προσέγγιση μάθησης μίας απεικόνισης του χώρου εισόδου n διαστάσεων στο χώρο εξόδου m διαστάσεων. Η απεικόνιση είναι μία συναρτησιακή προσέγγιση ανάμεσα στο χώρο εισόδου και εξόδου και δίνει τη δυνατότητα στο δίκτυο να γενικεύσει, δηλαδή να δίνει σωστές εξόδους-απαντήσεις για πρότυπα εισόδου που δεν έχουν χρησιμοποιηθεί πριν για την εκπαίδευσή του.

Η ικανότητα γενίκευσης μετράται απ' το πόσο καλά το δίκτυο ανταποκρίνεται στα δεδομένα επιβεβαίωσης. Τα δεδομένα επιβεβαίωσης έχουν την ίδια δομή με τα δεδομένα εκπαίδευσης, όμως αντιστοιχούν σε διαφορετικά πρότυπα εισόδου/εξόδου. Τα δεδομένα επιβεβαίωσης, όπως φυσικά και τα δεδομένα εκπαίδευσης πρέπει να είναι αντιπροσωπευτικά της λειτουργίας που προσπαθεί να μάθει το δίκτυο.

Γενικά η εκπαίδευση του δικτύου γίνεται με χρήση περίπου του 70% των διαθέσιμων δεδομένων (σύνολο εκπαίδευσης). Τα υπόλοιπα δεδομένα χρησιμοποιούνται για επιβεβαίωση της απόδοσης του δικτύου (σύνολο επιβεβαίωσης).

Στο πρόβλημα που μελετήθηκε στη διπλωματική αυτή τα πρότυπα (συνδυασμοί) εισόδου/εξόδου που χρησιμοποιήθηκαν ήταν συνολικά 118 [βλ. Παράρτημα]. Κάποια από αυτά είχαν αποκτηθεί από παλαιότερα πειράματα που είχαν γίνει στο Εργαστήριο Υψηλών Τάσεων, ενώ για να αυξηθεί ο αριθμός των προτύπων τα αρχικά αυτά δεδομένα εμπλουτίστηκαν με κάποια ακόμα που προέκυψαν από εφαρμογή μαθηματικού τύπου βασισμένου σε απλοποιημένο μοντέλο για την εύρεση της κρίσιμης τάσης. Απ' τα δεδομένα αυτά τα σύνολα εκπαίδευσης και επιβεβαίωσης διαμορφώθηκαν με τυχαίο τρόπο με χρήση βοηθητικού προγράμματος.

Ως προς τη δημιουργία και την εκπαίδευση του ΤΝΔ χρησιμοποιήθηκε το πρόγραμμα MATLAB, ενώ στον αλγόριθμο έχουν ληφθεί υπόψη δυνατότητες εκτέλεσης διαφορετικών σεναρίων που αφορούν την αρχιτεκτονική, το ρυθμό μάθησης και τη συνάρτηση εκπαίδευσης.

Παρακάτω θα παρουσιαστούν τα δεδομένα που χρησιμοποιήθηκαν ως είσοδοι και έξοδοι για την εκπαίδευση του δικτύου, η πειραματική μέθοδος για τη συλλογή των δεδομένων αυτών και η προσεγγιστική σχέση εύρεσης της τάσης διάσπασης για τα δεδομένα που δεν προέκυψαν από πειράματα.

6.2 Δεδομένα εκπαίδευσης

Για την εκπαίδευση του ΤΝΔ χρησιμοποιήθηκαν δεδομένα που αφορούν μονωτήρες τύπου cap & pin. Συγκεκριμένα, ως μεταβλητές εισόδου χρησιμοποιήθηκαν τα εξής χαρακτηριστικά:

- Μέγιστη Διάμετρος D_m (σε cm) του δίσκου του μονωτήρα.
- Ύψος H (σε cm) του μονωτήρα: Η απόσταση από το ένα μεταλλικό άκρο στο άλλο.
- Μήκος ερπυσμού L (σε cm): Είναι η συντομότερη διαδρομή επί της επιφανείας του μονωτήρα ανάμεσα σε δύο μεταλλικά άκρα του (μήκος της εξωτερικής ημιπεριμέτρου του μονωτικού υλικού σε επίπεδο τομής διερχόμενο από τον άξονα συμμετρίας του μονωτήρα) [21].
- Συντελεστής μορφής F : Καθορίζεται από τις διαστάσεις του μονωτήρα. Προκειμένου να προσδιοριστεί γραφικά, σχεδιάζεται η αντίστροφη τιμή της περιφέρειας του μονωτήρα ($1/p$) συναρτήσει του μερικού μήκους ερπυσμού (l) υπολογιζόμενο από το τέλος του μονωτήρα μέχρι το μετρούμενο σημείο. Ο συντελεστής μορφής δίνεται από το εμβαδό κάτω από τη σχηματιζόμενη καμπύλη και υπολογίζεται από τη σχέση [22]:

$$F = \int_0^L \frac{dl}{p(l)}$$
 Αν $D(l)$ είναι η διάμετρος του μονωτήρα, τότε η προηγούμενη σχέση τροποποιείται ως εξής [23]:
$$F = \int_0^L \frac{dl}{\pi \cdot D(l)}$$
- Αγωγιμότητα στρώματος σ_s ή επιφανειακή αγωγιμότητα (σε μS): Η αγωγιμότητα του στρώματος ρύπανσης πολλαπλασιασμένη με το συντελεστή μορφής. Συνήθως εκφράζεται σε μS [22].

Ενώ η εξαρτημένη μεταβλητή εξόδου του νευρωνικού δικτύου θεωρήθηκε η κρίσιμη τάση U_c (σε kV).

6.3 Πειραματική Διάταξη

Η πειραματική διάταξη για τη συλλογή των δεδομένων εγκαταστάθηκε αρχικά στο Κέντρο Δοκιμών Ερευνών και Προτύπων (ΚΔΕΠ) της ΔΕΗ. Ο απαιτούμενος εξοπλισμός για τη διεξαγωγή των πειραμάτων περιλάμβανε ένα σταθμό δοκιμών, ένα

μετασηματιστή τροφοδότησης και διατάξεις μέτρησης. Απ' αυτά, το ΚΔΕΠ ήδη διέθετε το Μ/Σ δοκιμών και τα όργανα μέτρησης, ενώ ο σταθμός δοκιμών κατασκευάστηκε στα πλαίσια διπλωματικής εργασίας που έγινε στο ΕΜΠ [24]. Ο σταθμός δοκιμών αποτελείται κυρίως από δύο θαλάμους: το θάλαμο ρύπανσης και το θάλαμο ομίχλης.

6.3.1 Θάλαμος ρύπανσης

Ο θάλαμος ρύπανσης είναι ένα μεταλλικό (DEXION 140) κουτί διαστάσεων $1,5\text{m} \times 1,5\text{m} \times 1,5\text{m}$, που καλύπτεται από φύλλα πολυαιθυλενίου, πάχους 0,2mm. Τα δοκίμια αναρτήθηκαν από την οροφή και περιστρέφονταν σε οριζόντια θέση, με ταχύτητα περίπου 140 r.p.m. (η ταχύτητα μπορούσε να μεταβάλλεται συνεχώς μεταξύ 0 και 260 r.p.m).

Πριν οι μονωτήρες κρεμαστούν στο θάλαμο ρύπανσης καθαρίστηκαν προσεκτικά, ώστε να αφαιρεθεί κάθε ίχνος γράσου. Ο καθαρισμός έγινε με βύθιση των μονωτήρων σε διάλυμα φωσφορικού νατρίου. Στη συνέχεια η επιφάνεια του μονωτήρα ξεπλύθηκε επιμελώς με νερό βρύσης. Η δημιουργία στερεάς επίστρωσης ρύπανσης στην επιφάνεια του μονωτήρα γίνεται με ψεκάσμο από διάλυμα ρύπανσης, καθώς ο μονωτήρας περιστρέφεται γύρω απ' τον κατακόρυφο άξονα, για να επιτευχθεί ομοιόμορφο στρώμα ρύπανσης. Οι ρυπαντές που χρησιμοποιήθηκαν ήταν: 75g/l πηλός πορσελάνης, 675g/l πυριτικό άλευρο (silica flour) και NaCl όπως απαιτείται, διαλυμένα σε ισοπροπυλική αλκοόλη. Η ισοπροπυλική αλκοόλη χρησιμοποιήθηκε, διότι ο μικρός χρόνος που χρειαζόταν για να στεγνώσει απέτρεπε το στάξιμο του διαλύματος και παρείχε ένα ομοιόμορφο στρώμα επικάλυψης. Ο χρόνος ρύπανσης ανήλθε σε περίπου 30 λεπτά. Το μίγμα αυτό αναδεύτηκε για αρκετή ώρα πριν τη μετάγγισή του στο δοχείο πίεσης. Ο ψεκάσμος έγινε από 3 μπεκ διαμέτρου στομίου 1mm, τα δύο απ' τα οποία ψέκαζαν την κάτω επιφάνεια του μονωτήρα, ενώ το άλλο την πάνω, από αποστάσεις περίπου 35cm, όπως ορίζει ο κανονισμός IEC-507/1975. Η πίεση του αέρα στην είσοδο του δοχείου ελέγχεται με ρυθμιστή πίεσης. Αφού ρυπάνθηκαν οι μονωτήρες, αφέθηκαν να στεγνώσουν για περίπου μία ώρα. Επιπλέον τοποθετήθηκαν σε χώρο θερμοκρασίας $8\text{ }^{\circ}\text{C}$ προκειμένου να στερεοποιηθεί η ρύπανση στην επιφάνειά τους και να μην ξεπλένεται με την έκθεσή τους σε ομίχλη.

Στη συνέχεια έγινε μέτρηση του επιπέδου της ρύπανσης στους μονωτήρες, σαν δείκτης

του οποίου χρησιμοποιήθηκε η πυκνότητα εναπόθεσης άλατος (salt deposit density). Τρεις μονωτήρες από κάθε σύνολο μονωτήρων όμοια ρυπασμένων, πλύθηκαν πολύ προσεκτικά με απιονισμένο νερό και μετρήθηκε η αγωγιμότητα του διαλύματος που προέκυψε απ' την πλύση.

6.3.2 Θάλαμος ομίχλης

Τα υπόλοιπα δοκίμια οδηγήθηκαν στη συνέχεια στο θάλαμο ομίχλης, ένα μεταλλικό κουτί διαστάσεων 4,6m x 4,6m x 4,6m, καλυμμένο με πολυαιθυλένιο πάχους 0,2mm. Για να γίνουν οι δοκιμές απαιτήθηκε οι μονωτήρες να αναρτηθούν σε κάποιο ύψος και να γειωθεί το σημείο ανάρτησης. Στο στέλεχος του μονωτήρα-δοκιμίου συνδέθηκε τμήμα αγωγού των γραμμών μεταφοράς μήκους 3m (1,5m εκατέρωθεν του μονωτήρα) με τα πραγματικά εξαρτήματα σύνδεσης, ώστε να γίνεται προσομοίωση του ηλεκτρικού πεδίου στην περιοχή του μονωτήρα. Για την προσαγωγή υψηλής τάσεως από το Μ/Σ δοκιμών χρησιμοποιήθηκε αγωγός σπирάλ διαμέτρου 24mm για την αντιμετώπιση του φαινομένου Corona. Η είσοδος του αγωγού στο θάλαμο ομίχλης έγινε από τρύπα διαμέτρου 150cm. Η διάμετρος αυτή καλύπτει λειτουργία για τιμές τάσης έως 150kV, εφόσον η ελάχιστη απόσταση μεταξύ κάθε γειωμένου σημείου του θαλάμου και του αγωγού ΥΤ ισούται με 0,5m ανά 100kV (IEC-507/1975).

Οι εκτοξευτήρες ομίχλης κατασκευάστηκαν στο μηχανουργείο του ΚΔΕΠ σύμφωνα με τους κανονισμούς IEC-507/1975 του Κεφ. III. Η λειτουργία του συστήματος απαιτούσε την ταυτόχρονη τροφοδότηση των ζευγών-ακροφυσίων α) με πεπιεσμένο αέρα μέχρι 7bars και β) με αλατοδιάλυμα ή νερό.

Η ομίχλη που δημιουργήθηκε από το σύστημα αυτό ήταν σχεδόν πάντα πιο πυκνή απ' ότι παρατηρείται στη φύση. Η πυκνότητα της ομίχλης δεν φάνηκε να επηρεάζει τα αποτελέσματα των δοκιμών, αλλά μόνο το ρυθμό ύγρανσης. Εντούτοις, όταν η ομίχλη γινόταν πολύ πυκνή, τα σταγονίδια νερού ξέπλεναν την τεχνητή ρύπανση από την επιφάνεια των μονωτήρων, με αποτέλεσμα να αυξάνεται η αντοχή τους κατά τη διάρκεια της δοκιμής.

6.3.3 Δοκιμές αντοχής

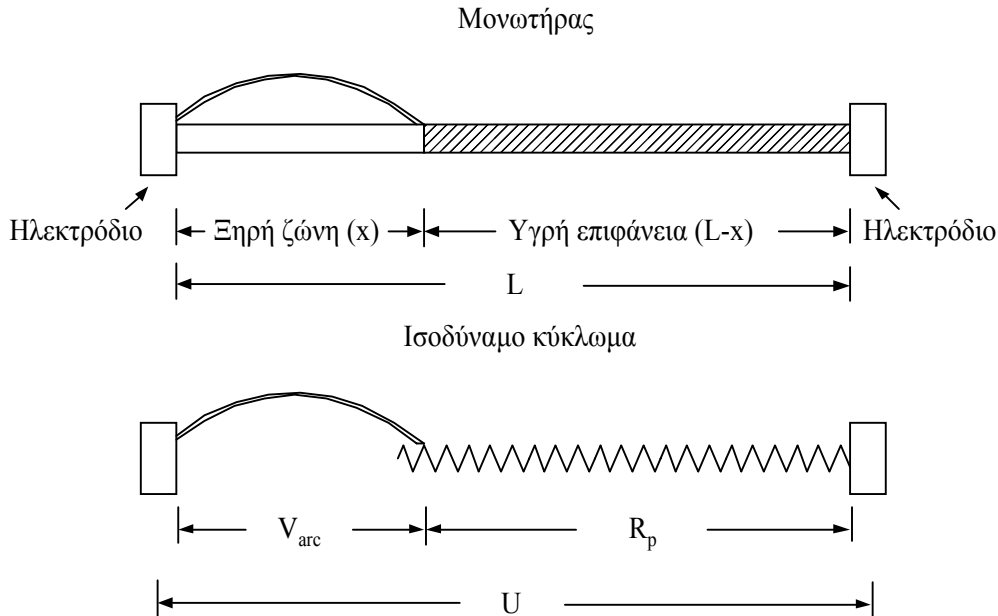
Η μικρότερη τιμή αντίστασης του μονωτήρα (ή της αλυσίδας μονωτήρων) αντιστοιχεί στη μέγιστη τιμή αγωγιμότητας στρώματος. Μετά την εμφάνιση αυτής της μικρότερης τιμής αντίστασης, η τάση δοκιμής εφαρμόζεται στο μονωτήρα μέσα στα επόμενα 5s και διατηρείται μέχρις ότου συμβεί υπερπήδηση, ή για 15min αν δεν συμβεί υπερπήδηση. Στοιχειώδεις παρατηρήσεις έδειξαν ότι οι περισσότερες υπερπηδήσεις συνέβησαν μέσα στο πρώτο λεπτό μετά την εφαρμογή της τάσης (περίπου το 80%), ενώ ένα 5% των υπερπηδήσεων συνέβησαν κατά το δεύτερο λεπτό. Στη συνέχεια ο μονωτήρας αφαιρείται από το θάλαμο ομίχλης και αφήνεται να στεγνώσει. Τοποθετείται και δεύτερη φορά στο θάλαμο ομίχλης, μέχρι η αγωγιμότητα στρώματος να φτάσει μια μέγιστη τιμή. Η τάση εφαρμόζεται και πάλι και η παραπάνω διαδικασία επαναλαμβάνεται. Η διαδικασία αυτή επαναλαμβάνεται και τρίτη φορά, εκτός αν και τις δύο προηγούμενες φορές είχε συμβεί υπερπήδηση. Αν και τις τρεις φορές δεν συμβεί υπερπήδηση, η τάση αυξάνεται με βήμα 5% και οι δοκιμές επαναλαμβάνονται. Αν συμβεί μία μόνο υπερπήδηση στις τρεις πρώτες δοκιμές, τότε εκτελείται και τέταρτη δοκιμή στο ίδιο επίπεδο τάσης. Γενικά, αν ο αριθμός των υπερπηδήσεων φτάσει τις δύο, οι δοκιμές δεν συνεχίζονται στο ίδιο ή σε υψηλότερο επίπεδο τάσης. Όταν ο συνολικός αριθμός των δοκιμών αντοχής σε κάθε επίπεδο τάσης φτάσει τις τρεις, δεν γίνονται περαιτέρω δοκιμές στο ίδιο ή σε χαμηλότερο επίπεδο τάσης. Η μέγιστη τάση αντοχής (V_c) είναι η υψηλότερη τάση στην οποία μπορούν να πραγματοποιηθούν δοκιμές αντοχής. Η ελάχιστη τάση υπερπήδησης (V_f) είναι η μικρότερη τάση στην οποία πραγματοποιούνται δύο δοκιμές υπερπήδησης.

Οι δοκιμές που έγιναν ήταν για διάφορους τύπους μονωτήρων ανάρτησης που χρησιμοποιούνται από τη ΔΕΗ στα συστήματα μεταφοράς και διανομής. Σημειώνεται επίσης ότι η πειραματική διαδικασία επαναλήφθηκε για τιμές ρύπανσης που φαίνεται να είναι αντιπροσωπευτικές της εκτιμώμενης ρύπανσης σε αρκετές περιοχές της Ελλάδας με έντονη ρύπανση.

6.4 Μαθηματικό μοντέλο για την εύρεση της κρίσιμης τάσης

Όπως αναφέρθηκε παραπάνω, κάποια απ' τα δεδομένα για την εκπαίδευση του ΤΝΔ προέκυψαν από την εφαρμογή του μαθηματικού μοντέλου για την εύρεση της κρίσιμης

τάσης, όπως αυτό παρουσιάστηκε από το Εργαστήριο Υψηλών Τάσεων του Ε.Μ.Π. [23, 26, 27]. Το μοντέλο αυτό είναι απλοποιημένο, σκοπό έχει τον υπολογισμό της τάσης υπερπήδησης ενός ρυπασμένου μονωτήρα και στηρίζεται κυρίως στο κυκλωματικό μοντέλο του Obenhaus [28] και του Alston [29]. Το μοντέλο του Obenhaus σχηματικά παρουσιάζεται στο Σχήμα 6.1.



Σχήμα 6. 1: Το μοντέλο του Obenhaus.

Στο μοντέλο του Obenhaus υπάρχει ένα τόξο μήκους x , το οποίο παριστάνει το τόξο που δημιουργείται κατά τη γεφύρωση των ξηρών ζωνών, εν σειρά με μία αντίσταση R_p , η οποία παριστάνει το στρώμα υγρής ρύπανσης. Το κύκλωμα τροφοδοτείται από πηγή σταθερής τάσης U .

Στο μοντέλο που προτάθηκε απ' το Εργαστήριο Υψηλών Τάσεων η κρίσιμη τάση U_c δίνεται από τον ακόλουθο τύπο:

$$U_c = \frac{A}{n+1} \cdot (L + \pi \cdot n \cdot D_m \cdot F \cdot K) \cdot (\pi \cdot A \cdot D_m \cdot \sigma_s)^{-\left(\frac{n}{n+1}\right)} \quad (6.1)$$

όπου L είναι το μήκος ερπυσμού του μονωτήρα, D_m είναι η μέγιστη διάμετρος του δίσκου του μονωτήρα, F είναι ο συντελεστής μορφής και A , n είναι οι σταθερές του τόξου. Η επιφανειακή αγωγιμότητα σ_s (σε Ω^{-1}) υπολογίζεται από τη σχέση [30]:

$$\sigma_s = (369.05 \cdot C + 0.42) \cdot 10^{-6} \quad (6.2)$$

όπου C είναι ισοδύναμη πυκνότητα εναποθέματος αλατιού (ESDD) σε mg/cm².

Ο συντελεστής της αντίστασης του στρώματος ρύπανσης K, στην περίπτωση που ο μονωτήρας είναι cap-and-pin, εκφράζεται από τη σχέση:

$$K = 1 + \frac{n+1}{2 \cdot \pi \cdot F \cdot n} \cdot \ln\left(\frac{L}{2 \cdot \pi \cdot R \cdot F}\right) \quad (6.3)$$

όπου R είναι η ακτίνα του ίχνους του τόξου, η οποία δίνεται από τη σχέση [31]:

$$R = \sqrt{\frac{I}{\pi \cdot 1.45}} = 0.469 \cdot (\pi \cdot A \cdot D_m \cdot \sigma_s)^{\frac{1}{2 \cdot (n+1)}} \quad (6.4)$$

Όπου I στην παραπάνω σχέση είναι το ρεύμα διαρροής.

Στην περίπτωση μονωτήρων stab-type ο συντελεστής K δίνεται από τη σχέση:

$$K = \frac{N \cdot (n+1)}{2 \cdot \pi \cdot F \cdot n} \cdot \left(\ln\left(\frac{4 \cdot L}{\pi \cdot N \cdot R}\right) - \ln\left(\tan \frac{\pi}{2 \cdot (n+1)}\right) \right) \quad (6.5)$$

όπου N είναι ο αριθμός των “δίσκων”.

Στη συνέχεια το παραπάνω μαθηματικό μοντέλο επεκτάθηκε, ώστε να καλύπτει περιπτώσεις ανομοιόμορφης ρύπανσης στην επιφάνεια του μονωτήρα [32, 33]. Με βάση το συγκεκριμένο μοντέλο το μήκος ερπυσμού χωρίζεται σε τμήματα και σε κάθε τμήμα θεωρείται σταθερή ποσότητα ρύπανσης. Τελικά, υπολογίζεται μία μέση τιμή της ισοδύναμης πυκνότητας εναποθέματος αλατιού (ESDD), η οποία υπεισέρχεται στη σχέση (6.1), προκειμένου να υπολογιστεί η κρίσιμη τάση σε μονωτήρα ρυπασμένο ανομοιόμορφα.

6.5 Επιλογή των συνόλων εκπαίδευσης και επιβεβαίωσης

Η επιλογή των δύο συνόλων για την εκπαίδευση του νευρωνικού δικτύου και την επιβεβαίωση για το ότι εκπαιδεύτηκε σωστά έγινε με τη βοήθεια ενός προγράμματος που υλοποιήθηκε στο MATLAB. Συνολικά τα δεδομένα μας αποτελούνταν από 118 διανύσματα (βλ. Πίνακα Data του Παραρτήματος). Αρχικά το σύνολο εκπαίδευσης

αποφασίστηκε να είναι το 80% του συνόλου των δεδομένων (δηλαδή 94 διανύσματα), ενώ το σύνολο επιβεβαίωσης το υπόλοιπο 20% (δηλαδή 24 διανύσματα).

Ο αλγόριθμος αυτός ήταν ο εξής:

```
a=rand(24,1);
b=118*a;
c=round(b);
for i=1:24
    testdata(i,1)=Data(c(i),1);
    testdata(i,2)=Data(c(i),2);
    testdata(i,3)=Data(c(i),3);
    testdata(i,4)=Data(c(i),4);
    testdata(i,5)=Data(c(i),5);
    testdata(i,6)=Data(c(i),6);
    traindata=Data;
    traindata(c(i),:)=[];
```

Η λειτουργία του παραπάνω προγράμματος είναι η εξής:

Ο πίνακας a είναι ένας πίνακας 24×1 , τα στοιχεία του οποίου είναι τυχαίοι αριθμοί στο διάστημα $(0,1)$ που παράγονται από τη συνάρτηση `rand` του MATLAB. Ο πίνακας a πολλαπλασιάζεται με 118 και στρογγυλοποιείται, ώστε τελικά να προκύψει ένας πίνακας c , διάστασης 24×1 , του οποίου τα στοιχεία έχουν πλέον τιμές από 1 έως 118. Τα στοιχεία του πίνακα c αντιπροσωπεύουν εκείνες τις γραμμές που θα αφαιρεθούν από τον πίνακα `Data` που περιέχει όλα τα δεδομένα και θα αποτελέσουν τις γραμμές του πίνακα `testdata`, ο οποίος θα είναι το σύνολο επιβεβαίωσης. Το σύνολο εκπαίδευσης (πίνακας `traindata`) θα είναι ο αρχικός πίνακας `Data`, απ' τον οποίο όμως θα έχουν αφαιρεθεί πλέον οι γραμμές που ανήκουν στον `testdata`.

Σημειώνεται ότι στο σύνολο εκπαίδευσης πρέπει οπωσδήποτε να περιλαμβάνονται οι μέγιστες και ελάχιστες τιμές για κάθε μία από τις μεταβλητές εισόδου. Γι' αυτό, αν η συνάρτηση `rand` δώσει κάποια τυχαία γραμμή που περιλαμβάνει μέγιστο ή ελάχιστο κάποιας εκ των μεταβλητών, αυτή θα πρέπει να εξαιρεθεί από το σύνολο `testdata`.

Στην προκειμένη περίπτωση όλες οι γραμμές του πίνακα `testdata` που προέκυψαν με τον τρόπο που περιγράφηκε ήταν αποδεκτές και οι πίνακες `testdata` και `traindata`, καθώς και ο πίνακας `c` διαμορφώθηκαν όπως φαίνεται στο Παράρτημα.

6.6 Κατασκευή του νευρωνικού δικτύου

Στην εργασία αυτή έγινε η κατασκευή και μελέτη ενός ΤΝΔ, το οποίο εκπαιδεύτηκε έτσι, ώστε να κάνει πρόβλεψη για την κρίσιμη τάση ενός μονωτήρα, αν του δίνονται τα υπόλοιπα χαρακτηριστικά.

Για την κατασκευή του ΤΝΔ έγινε χρήση των έτοιμων εργαλείων του MATLAB. Ο κώδικας παρατίθεται στη συνέχεια.

Αρχικά παρουσιάζεται το κυρίως πρόγραμμα (`main`) μέσω του οποίου γίνεται κλήση της συνάρτησης `Net_1` για την κατασκευή του νευρωνικού και στο οποίο μπορεί κανείς να μεταβάλλει τις παραμέτρους που ορίζουν τον αριθμό των επιπέδων, τον αριθμό των νευρώνων σε κάθε επίπεδο, τον αριθμό των εποχών και το ρυθμό μάθησης.

Το πρόγραμμα `main` είναι το εξής:

```
%Εντολές για να κλείσει όλα τα παράθυρα που έχουν ανοίξει από
προηγούμενη εφαρμογή και όλες τις μεταβλητές που χρησιμοποίησε η
προηγούμενη εφαρμογή
clear all;
close all;

%Για την εύρεση του ελάχιστου σφάλματος στον πίνακα apothiki (βλ.
παρακάτω)
Minimum = 1000000;

%Κλήση των αρχείων TestData_file και TrainData_file που περιέχουν
τους πίνακες TestData και TrainData αντίστοιχα.
TestData_file;
TrainData_file;

%Επιλογή αριθμού κρυμμένων επιπέδων (1 ή 2)
NumOfLayers = ...;
```

```
%Επιλογή αριθμού εποχών
NumOfEpochs = ...;

%Επιλογή του στόχου εκπαίδευσης, δηλαδή του μέσου τετραγωνικού
σφάλματος μεταξύ πραγματικών εξόδων και εξόδων από την εκπαίδευση
του νευρωνικού, το οποίο αν επιτευχθεί σταματά η εκπαίδευση
Goal = ...;

%Επιλογή του ρυθμού μάθησης
LearningRate = ...;

%Επιλογή αριθμού νευρώνων t
t = ...;

j = 1;
i = 1;
iprox=0;
%Αν το κρυμμένο επίπεδο είναι 1, καλείται η συνάρτηση Net_1 με τις
παραμέτρους που φαίνονται παρακάτω
if NumOfLayers == 1;
    for NumOfNeuronsL1 = t;
        iprox=iprox+1;
        Error=
Net_1 (NumOfLayers,NumOfNeuronsL1,0,TrainData,TestData,
        'traingd',NumOfEpochs,Goal,LearningRate);
%Σχηματίζεται ο πίνακας apothiki ο οποίος στην πρώτη στήλη εμφανίζει
τον αύξοντα αριθμό, στη δεύτερη στήλη των αριθμό νευρώνων στο πρώτο
επίπεδο, στην τρίτη στήλη τον αριθμό νευρώνων στο δεύτερο επίπεδο
(μηδέν) και στην τέταρτη στήλη το μέσο τετραγωνικό σφάλμα (Error)
        apothiki (iprox,4)=Error;
        apothiki (iprox,1)=iprox;
        apothiki (iprox,2)=NumOfNeuronsL1;
    end
end
```

Αν τα κρυμμένα επίπεδα είναι 2, καλείται η συνάρτηση Net_1 με τις παραμέτρους που φαίνονται παρακάτω

```

if NumOfLayers == 2;
    for NumOfNeuronsL1 = t;
        for NumOfNeuronsL2 = t;
            if NumOfNeuronsL2<=NumOfNeuronsL1;
                iprox=iprox+1;

Error=Net_1 (NumOfLayers,NumOfNeuronsL1,NumOfNeuronsL2,
              TrainData,TestData, 'traingd', NumOfEpochs, Goal,
              LearningRate);
%Σχηματίζεται ο πίνακας apothiki ο οποίος στην πρώτη στήλη εμφανίζει
τον αύξοντα αριθμό, στη δεύτερη στήλη των αριθμό νευρώνων στο πρώτο
επίπεδο, στην τρίτη στήλη τον αριθμό νευρώνων στο δεύτερο επίπεδο
και στην τέταρτη στήλη το μέσο τετραγωνικό σφάλμα (Error)
                apothiki (iprox,4)=Error;
                apothiki (iprox,1)=iprox;
                apothiki (iprox,2)=NumOfNeuronsL1;
                apothiki (iprox,3)=NumOfNeuronsL2;
            end
            j=j+1;
        end
    end
    i=i+1;
    j=1;
end
end
[row column]=size(apothiki);

%Εύρεση του ελάχιστου στοιχείου της τέταρτης γραμμής του πίνακα
apothiki, δηλαδή του ελάχιστου μέσου τετραγωνικού σφάλματος
for i=1:row;
    if ((apothiki (i,4)<Minimum)&&(apothiki (i,4)~=0));
        Minimum=apothiki (i,4);
        NumOfNeuronsL1_min=apothiki (i,2);
        NumOfNeuronsL2_min=apothiki (i,3);
    end
end

```

```
end
```

```
%Εμφάνιση του πίνακα arothiki, του ελάχιστου σφάλματος και του  
συνδυασμού νευρώνων για τον οποίο αυτό εμφανίστηκε
```

```
arothiki
```

```
Minimum
```

```
NumOfNeuronsL1_min
```

```
NumOfNeuronsL2_min
```

```
%Κλήση του προγράμματος correlation για την εύρεση της συσχέτισης  
[βλ. Παράρτημα
```

```
correlation
```

Η συνάρτηση Net_1 που καλείται από το κυρίως πρόγραμμα επιστρέφει κάθε φορά το μέσο τετραγωνικό σφάλμα και είναι η εξής:

```
Function
```

```
Error=Net_1 (NumOfLayers, NumOfNeuronsL1, NumOfNeuronsL2, TrainData,  
             TestData, TrainingMethod, NumOfEpochs, Goal, LearningRate)
```

```
% Επιλογή διανυσμάτων εισόδου από τον πίνακα TrainData που περιέχει  
τα δεδομένα εκπαίδευσης
```

```
Input = TrainData(:,1:5);
```

```
% Ανάστροφος του πίνακα εισόδων
```

```
RevInput = Input';
```

```
% Επιλογή του διανύσματος εξόδου από τον πίνακα TrainData
```

```
Target = TrainData(:,6);
```

```
% Ανάστροφος του διανύσματος εξόδου
```

```
RevTarget = Target';
```

```
% Κανονικοποίηση των δεδομένων στο διάστημα [-1 1] με χρήση της
συνάρτησης premnmx, η οποία δέχεται σαν ορίσματα τους ανεστραμμένους
πίνακες RevInput και RevTarget
[NormalizedRevInput,mininput,maxinput,NormalizedRevTarget,mintarget,
maxtarget] = premnmx(RevInput,RevTarget);

%Εύρεση της μέγιστης και ελάχιστης τιμής των κανονικοποιημένων
δεδομένων με χρήση της συνάρτησης minmax
Limit = minmax(NormalizedRevInput);

% Κατασκευή του νευρωνικού με τη συνάρτηση newff. Αν έχουμε δύο
κρυφά επίπεδα στο νευρωνικό δίκτυο, τότε στα ορίσματα της newff
προστίθεται και η παράμετρος NumOfNeuronsL2. Ως συνάρτηση μεταφοράς
χρησιμοποιείται η tansig
if NumOfLayers == 1
    NeuralNet=newff(Limit,[NumOfNeuronsL1 1],
                    {'tansig','tansig'},TrainingMethod);
else
    NeuralNet=newff(Limit,[NumOfNeuronsL1 NumOfNeuronsL2 1],
                    {'tansig','tansig','tansig'},TrainingMethod);
end

% Ορισμός των παραμέτρων του δικτύου, των οποίων η μεταβολή γίνεται
μέσω του κυρίως προγράμματος
NeuralNet.trainParam.epochs=NumOfEpochs;
NeuralNet.trainParam.goal=Goal;
NeuralNet.trainParam.lr=LearningRate;

% Αρχικοποίηση των βαρών του κρυμμένου επιπέδου με την initnw και
απόδοση τυχαίων τιμών στα βάρη εισόδου με τη rands
if NumOfLayers == 1
    NeuralNet.layers{1}.initFcn = 'initnw';
    NeuralNet.layers{2}.initFcn = 'initnw';
    NeuralNet.biases{1}.initFcn = 'rands';
    NeuralNet.biases{2}.initFcn = 'rands';
else
```

```
NeuralNet.layers{1}.initFcn = 'initnw';
NeuralNet.layers{2}.initFcn = 'initnw';
NeuralNet.layers{3}.initFcn = 'initnw';
NeuralNet.biases{1}.initFcn = 'rands';
NeuralNet.biases{2}.initFcn = 'rands';
NeuralNet.biases{3}.initFcn = 'rands';
end

% Εφαρμογή των αρχικοποιήσεων με την init
NeuralNet = init(NeuralNet);

% Εκπαίδευση του νευρωνικού
NeuralNet = train(NeuralNet,NormalizedRevInput,NormalizedRevTarget);

% Εξομοίωση νευρωνικού
SimNet = sim(NeuralNet,NormalizedRevInput);
% Επιλογή διανυσμάτων εισόδου από τον πίνακα TestData που περιέχει
τα δεδομένα επαλήθευσης
TestInput = TestData(:,1:5);
RevTestInput = TestInput';
% Επιλογή διανύσματος εξόδου από το αρχείο εισόδου TestData
TestTarget = TestData(:,6);
Num=length(TestTarget)
RevTestTarget = TestTarget';

% Κανονικοποίηση των δεδομένων στο διάστημα [-1 1]
[NormalizedRevTestInput,testdmininput,testdmaxinput,
NormalizedRevTestTarget,testdmintarget,testdmaxtarget]=premnmx(RevTestInput,RevTestTarget);

% Εξομοίωση δικτύου με βάση τις κανονικοποιημένες εισόδους
NorSimTestdNet = sim(NeuralNet,NormalizedRevTestInput);

% Αποκανονικοποίηση
```



```

[SimTestdNet] =
postmmx(NorSimTestdNet, testdmintarget, testdmaxtarget);

%Ανοιγμα αρχείων στα οποία αποθηκεύεται η απόλυτη διαφορά μεταξύ
εκτιμώμενων εξόδων SimTestdNet και πραγματικών εξόδων RevTestTarget
fid1=fopen('senario1_sfalma.txt', 'a');
fid2=fopen('senario1_error.txt', 'a');
fid3=fopen('senario1_target.txt', 'a');

Sfalma=(SimTestdNet-RevTestTarget)';

%Υπολογισμός μέσου τετραγωνικού σφάλματος με χρήση της συνάρτησης
mse
Error = mse(Sfalma)

Estimated_Target = SimTestdNet'

fprintf(fid2, ' %12.8f\n', Error);
for i=1:Num
    fprintf(fid1, ' %12.8f ', Sfalma(i));
    fprintf(fid3, ' %12.8f ', Estimated_Target(i));
end
    fprintf(fid1, ' \n');
    fprintf(fid3, ' \n');

%Κλείσιμο των αρχείων
fclose(fid1);
fclose(fid2);
fclose(fid3);

```

6.7 Σύνοψη

Παραπάνω περιγράφηκε ο αλγόριθμος για τη κατασκευή, την εκπαίδευση και τον έλεγχο του ΤΝΔ (Net_1), καθώς και το πρόγραμμα main, απ' το οποίο γίνεται κλήση της συνάρτησης Net_1 και μεταβολή των παραμέτρων που ενδιαφέρουν για τις διάφορες

δοκιμές. Στο επόμενο κεφάλαιο θα παρουσιαστούν τα αποτελέσματα των δοκιμών αυτών και η επεξεργασία τους.

Κεφάλαιο 7

Αποτελέσματα

7.1 Εισαγωγή

Παρακάτω παρατίθενται τα αποτελέσματα από διαδοχικές δοκιμές που έγιναν με το πρόγραμμα που περιγράφηκε στο Κεφάλαιο 6. Οι δοκιμές έγιναν για διάφορες έτοιμες μεθόδους εκπαίδευσης του MATLAB (`traingd`, `traingda`, `traingdx` και `trainlm`), οι οποίες είναι ουσιαστικά παραλλαγές του αλγόριθμου για τη διάδοση σφάλματος (`backpropagation`). Για κάθε μία απ' τις παραπάνω μεθόδους πάρθηκε ένα σύνολο σεναρίων στα οποία η παράμετρος που μεταβαλλόταν ήταν ο αριθμός των εποχών. Έτσι, για κάθε μέθοδο προέκυψε ένα σύνολο 10 σεναρίων για μεταβολή εποχών από 500-5000 με βήμα 500. Σε κάθε σενάριο γινόταν εσωτερική μεταβολή στον αριθμό των νευρώνων (από 2 έως 25) προκειμένου να βρεθεί για πόσους νευρώνες παρουσιαζόταν ελάχιστο σφάλμα. Για τον αριθμό νευρώνων που δίνουν κάθε φορά το καλύτερο σφάλμα έγινε η γραφική παράσταση της συσχέτισης μεταξύ πραγματικών και εκτιμώμενων τιμών, κοινή γραφική για τα στοιχεία των πινάκων που περιέχουν τις πραγματικές και εκτιμώμενες τιμές, καθώς και η γραφική παράσταση της απόλυτης διαφοράς των στοιχείων αυτών. Τέλος, για κάθε περίπτωση συνάρτησης εκπαίδευσης έγινε μια τρισδιάστατη παρουσίαση του μέσου τετραγωνικού σφάλματος συναρτήσει του αριθμού εποχών και νευρώνων, προκειμένου να καθοριστεί η βέλτιστη αρχιτεκτονική του δικτύου και να γίνουν περαιτέρω δοκιμές για μεταβολή του ρυθμού μάθησης, ο οποίος αρχικά θεωρήθηκε σταθερός και ίσος με 0,5. Παρακάτω περιγράφεται η λειτουργία των μεθόδων εκπαίδευσης που χρησιμοποιήθηκαν.

7.1.1 Λειτουργία της `traingd`

Υπάρχουν πολλές παραλλαγές του αλγορίθμου `backpropagation`. Η πιο απλή υλοποίηση της μάθησης με ανάστροφη διάδοση σφάλματος είναι αυτή στην οποία τα βάρη ανανεώνονται προς την κατεύθυνση της αρνητικής κλίσης (`negative gradient`). Υπάρχουν

δύο διαφορετικοί τρόποι με τους οποίους μπορεί να υλοποιηθεί ο κανόνας κλίσης: με τον πρώτο τρόπο ο υπολογισμός της κλίσης και η ανανέωση των βαρών γίνεται μετά από την παρουσίαση της κάθε εισόδου στο δίκτυο, ενώ με το δεύτερο τρόπο πρώτα παρουσιάζονται όλες οι εισοδοί στο δίκτυο και στη συνέχεια ανανεώνονται τα βάρη (λειτουργία ‘παρτίδας’ προτύπων – batch mode).

Η *traingd* είναι μέθοδος εκπαίδευσης που λειτουργεί με βάση το δεύτερο τρόπο. Τα βάρη ανανεώνονται κατά τη φορά αρνητικής κλίσης της συνάρτησης λειτουργίας. Οι παράμετροι που σχετίζονται με την *traingd* είναι οι εποχές, ο στόχος (goal) και ο ρυθμός μάθησης (*lr*). Ο ρυθμός μάθησης πολλαπλασιάζεται με την αντίθετη τιμή της κλίσης και καθορίζονται έτσι οι μεταβολές στα βάρη. Όσο μεγαλύτερος είναι ο ρυθμός μάθησης, τόσο μεγαλύτερο είναι το βήμα της μεταβολής. Αν ο ρυθμός μάθησης είναι πολύ μεγάλος, ο αλγόριθμος γίνεται ασταθής. Αν ο ρυθμός μάθησης ορισθεί πολύ μικρός, ο αλγόριθμος χρειάζεται πολύ χρόνο για να συγκλίνει. Οι υπόλοιπες παράμετροι καθορίζουν το πότε θα σταματήσει η εκπαίδευση. Η εκπαίδευση σταματά όταν ο αριθμός των επαναλήψεων ξεπεράσει τον αριθμό των εποχών που έχει ορισθεί ή αν η συνάρτηση λειτουργίας (μέσο τετραγωνικό σφάλμα εκπαίδευσης) πάρει τιμή μικρότερη του στόχου. Μια άλλη συνάρτηση που λειτουργεί με ‘παρτίδες’ προτύπων είναι η *traingdm* που εκμεταλλεύεται και τον όρο ορμής για ταχύτερη σύγκλιση. Η *traingdm* δεν δοκιμάστηκε στην παρούσα εφαρμογή.

7.1.2 Λειτουργία των *traingda* και *traingdx*

Και οι δύο αυτές μέθοδοι λειτουργούν όπως η *traingd*, δηλαδή σε batch mode, αλλά παρουσιάζουν το πλεονέκτημα ότι είναι πολύ πιο γρήγορες στη σύγκλιση απ’ την *traingd*. Γενικά οι γρήγοροι αλγόριθμοι κατατάσσονται σε δύο κατηγορίες. Η πρώτη κατηγορία χρησιμοποιεί ευριστικές τεχνικές [34], που σχηματίζονται μετά από ανάλυση της λειτουργίας της πρότυπης ‘μεθόδου κλίσης’. Μία ευριστική τροποποίηση είναι η τεχνική με προσθήκη όρου ορμής, η οποία χρησιμοποιείται και απ’ την *traingdm*, που αναφέρθηκε παραπάνω. Οι *traingda* και *traingdx* χρησιμοποιούν μία άλλη ευριστική τεχνική, αυτή του μεταβαλλόμενου ρυθμού μάθησης.

Με τον κλασικό κανόνα κλίσης, ο ρυθμός μάθησης διατηρείται σταθερός κατά τη διάρκεια της εκπαίδευσης. Η λειτουργία του αλγορίθμου είναι πολύ ευαίσθητη στη

σωστή ρύθμιση του ρυθμού μάθησης. Αν ο ρυθμός μάθησης ορισθεί πολύ μεγάλος, ο αλγόριθμος παρουσιάζει ταλαντώσεις και γίνεται ασταθής, ενώ σε αντίθετη περίπτωση αργεί να συγκλίνει. Η λειτουργία του κανόνα κλίσης μπορεί να βελτιωθεί αν επιτραπεί στο ρυθμό μάθησης να αλλάζει κατά τη διάρκεια της εκπαίδευσης. Ένας προσαρμοζόμενος ρυθμός μάθησης θα βοηθούσε, ώστε και το βήμα μάθησης να είναι όσο το δυνατό μεγαλύτερο (μεγαλύτερη ταχύτητα), αλλά και η μάθηση να διατηρείται ευσταθής. Για να υλοποιηθεί ο προσαρμοζόμενος ρυθμός μάθησης χρειάζεται να γίνουν κάποιες αλλαγές στον αλγόριθμο που χρησιμοποιεί η `traingd`. Πρώτον, υπολογίζονται η έξοδος και το σφάλμα για το αρχικό δίκτυο. Σε κάθε εποχή υπολογίζονται καινούρια βάρη, χρησιμοποιώντας τον τρέχοντα ρυθμό μάθησης. Στη συνέχεια υπολογίζονται καινούριες εξοδοί και σφάλματα.

Όπως και με τον όρο ορμής, έτσι κι εδώ, αν το καινούριο σφάλμα υπερβαίνει το παλιό κατά μία προκαθορισμένη τιμή, τα καινούρια βάρη απορρίπτονται. Επιπλέον ο ρυθμός μάθησης μειώνεται. Αλλιώς τα καινούρια βάρη γίνονται αποδεκτά. (Τόσο ο λόγος καινούριου/παλιού σφάλματος, όσο και η μείωση στο ρυθμό μάθησης είναι παράμετροι που ορίζονται εντός της έτοιμης συνάρτησης εκπαίδευσης του MATLAB και στην παρούσα εργασία χρησιμοποιήθηκαν οι default τιμές. Γενικά πάντως οι τιμές αυτές μπορούν να μεταβληθούν με κατάλληλη τροποποίηση του αλγορίθμου).

Εάν το καινούριο σφάλμα είναι μικρότερο απ' το παλιό, ο ρυθμός μάθησης αυξάνεται. Εκπαίδευση με χρήση του αλγορίθμου ανάστροφης διάδοσης σφάλματος και προσαρμοζόμενο ρυθμό μάθησης επιτυγχάνεται με τη χρήση της συνάρτησης `traingda`, ενώ η συνάρτηση `traingd` συνδυάζει τον προσαρμοζόμενο ρυθμό μάθησης με προσθήκη όρου ορμής.

7.1.3 Λειτουργία της `trainlm`

Ο αλγόριθμος με τον οποίο λειτουργεί η `trainlm` είναι ο αλγόριθμος ανάστροφης διάδοσης σφάλματος κατά Levenberg-Marquardt, ο οποίος με τη σειρά του είναι προσέγγιση των μεθόδων quasi-Newton [34]. Ο αλγόριθμος αυτός φαίνεται ότι δίνει καλά αποτελέσματα για δίκτυα μεσαίου και μεγάλου μεγέθους (έως και αρκετών εκατοντάδων βαρών).

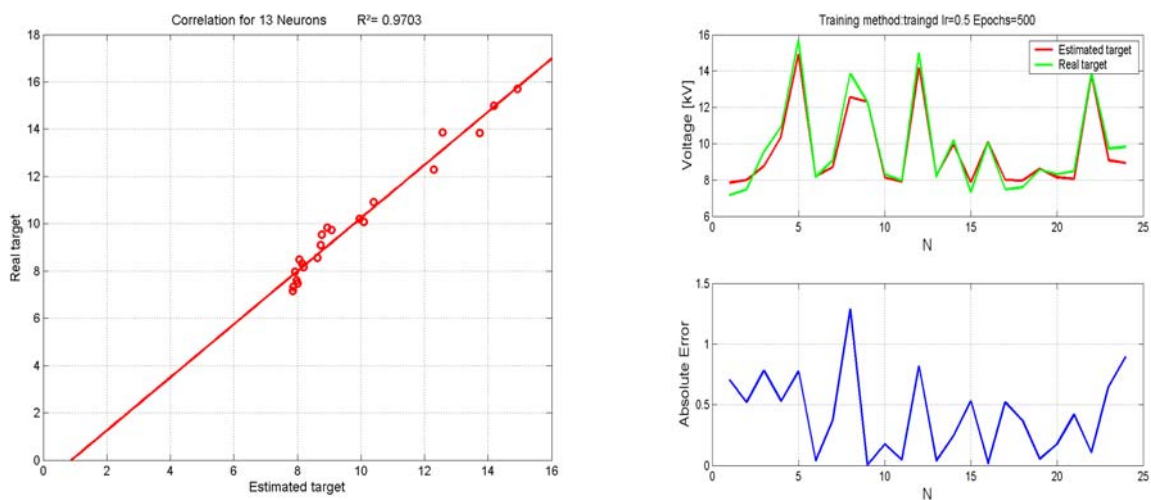
7.2 Σενάρια 1-10 για μέθοδο εκπαίδευσης την traingd

Η μέθοδος εκπαίδευσης traingd εκπαιδεύει το νευρωνικό δίκτυο με βάση τον κανόνα ανάστροφης διάδοσης σφάλματος (gradient descent backpropagation). Τα στοιχεία του δικτύου είναι:

- 1 επίπεδο
- Μεταβολή νευρώνων: από 2 έως 25
- Ρυθμός μάθησης: 0,5
- Μεταβολή εποχών: από 500 έως 5000 με βήμα 500.

ΣΕΝΑΡΙΟ 1

Για 500 Εποχές: Το μικρότερο σφάλμα προέκυψε για 13 νευρώνες και είχε τιμή 0,2903. Η συσχέτιση μεταξύ πραγματικών και εκτιμώμενων τιμών στην περίπτωση αυτή είναι 0,9703.

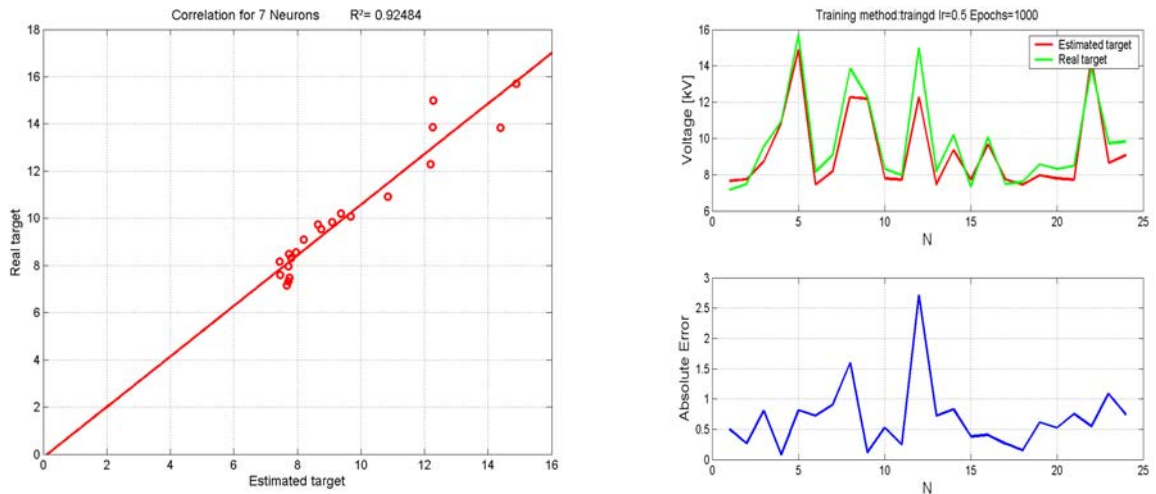


Σχήμα 7.1

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμώμενης τιμής για το σύνολο ελέγχου για 13 νευρώνες.
Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμώμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

ΣΕΝΑΡΙΟ 2

Για 1000 Εποχές: Μικρότερο σφάλμα 0,7559 για 7 νευρώνες. Η συσχέτιση μεταξύ πραγματικών και εκτιμώμενων τιμών στην περίπτωση αυτή είναι 0,92484.

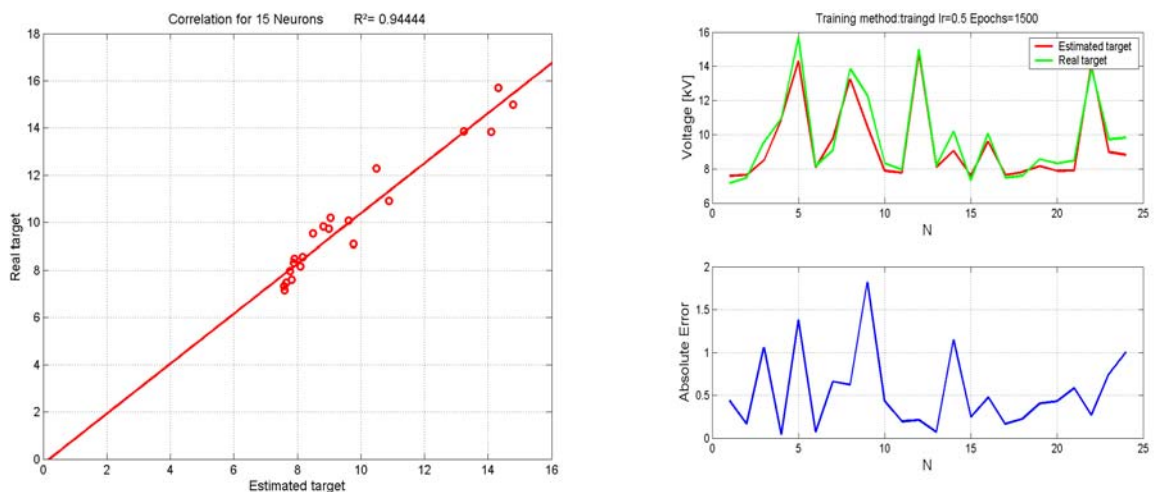


Σχήμα 7.2

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμώμενης τιμής για το σύνολο ελέγχου για 7 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμώμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

ΣΕΝΑΡΙΟ 3

Για 1500 Εποχές: Μικρότερο σφάλμα 0,4868 για 15 νευρώνες. Η συσχέτιση μεταξύ πραγματικών και εκτιμώμενων τιμών στην περίπτωση αυτή είναι 0,9444.

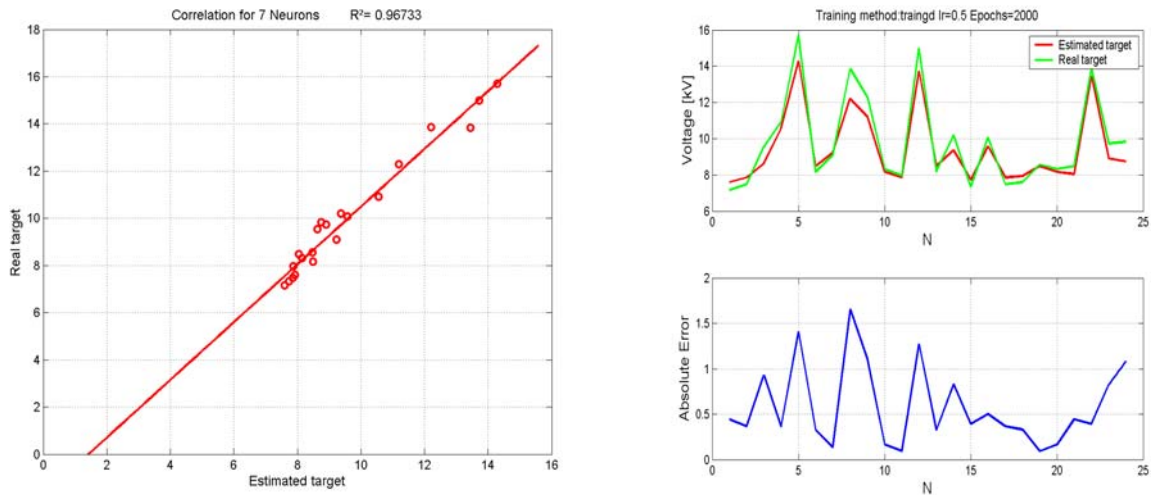


Σχήμα 7.3

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμώμενης τιμής για το σύνολο ελέγχου για 15 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμώμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

ΣΕΝΑΡΙΟ 4**Για 2000 Εποχές**

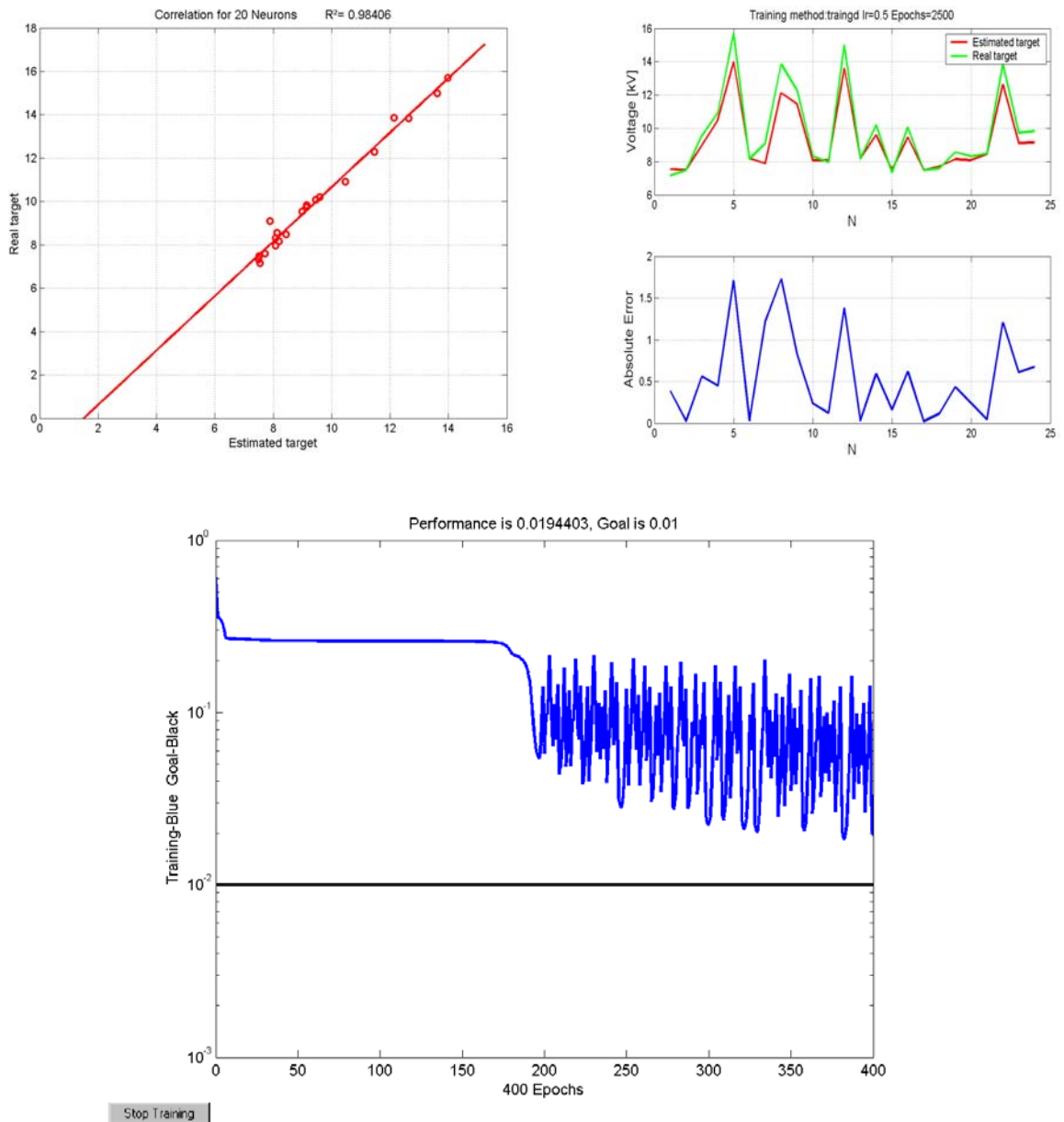
Μικρότερο σφάλμα 0,5301 για 7 νευρώνες. Η συσχέτιση μεταξύ πραγματικών και εκτιμούμενων τιμών στην περίπτωση αυτή είναι 0,96733.

**Σχήμα 7.4**

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμούμενης τιμής για το σύνολο ελέγχου για 7 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμούμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

ΣΕΝΑΡΙΟ 5**Για 2500 Εποχές**

Μικρότερο σφάλμα 0,5859 για 20 νευρώνες. Η συσχέτιση μεταξύ πραγματικών και εκτιμούμενων τιμών στην περίπτωση αυτή είναι 0,98406. Πρέπει να σημειωθεί ότι η καμπύλη εκπαίδευσης για την traingd εμφάνιζε μεγάλη ταλάντωση για αριθμό νευρώνων πάνω από 15. Αυτό σημαίνει ότι για πολλούς νευρώνες η μέθοδος δεν συγκλίνει και συνεπώς το ότι δίνει μικρό σφάλμα είναι τυχαίο και δεν μπορεί να θεωρηθεί σωστό. Παρακάτω φαίνεται η καμπύλη εκπαίδευσης για την traingd και οι ταλαντώσεις αυτής για μεγάλο αριθμό νευρώνων.



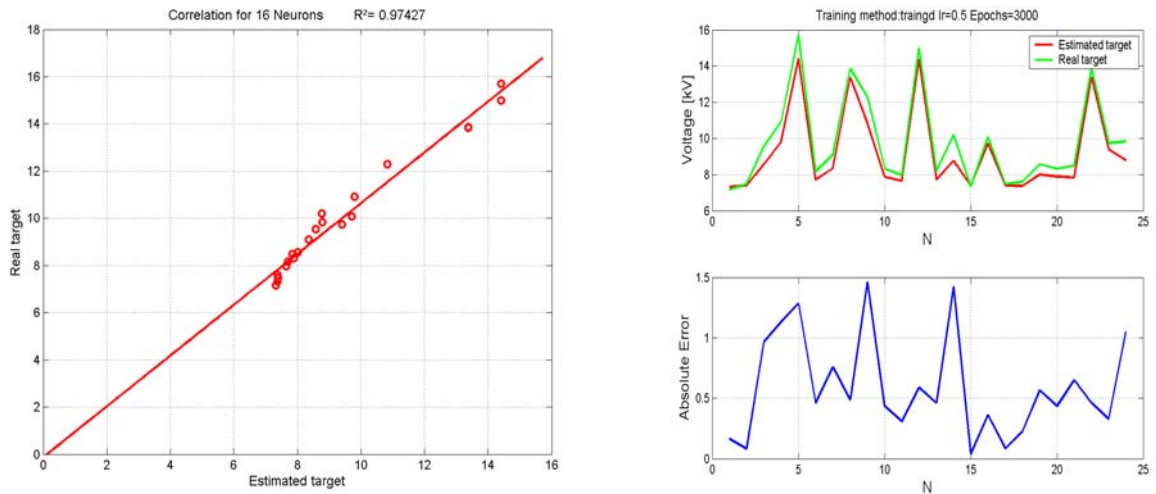
Σχήμα 7.5

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμούμενης τιμής για το σύνολο ελέγχου για 20 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμούμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

Κάτω: η ταλάντωση που παρατηρείται στην καμπύλη εκπαίδευσης για μεγάλο αριθμό νευρώνων.

ΣΕΝΑΡΙΟ 6:

Για 3000 Εποχές: Μικρότερο σφάλμα 0,5197 για 16 νευρώνες. Η συσχέτιση μεταξύ πραγματικών και εκτιμώμενων τιμών στην περίπτωση αυτή είναι 0,97427.

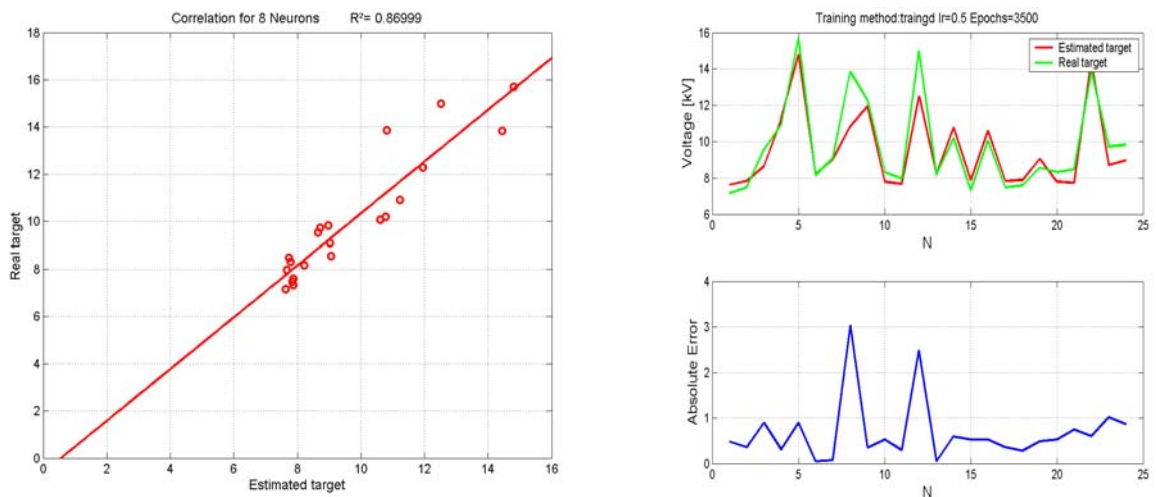


Σχήμα 7.6

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμώμενης τιμής για το σύνολο ελέγχου για 16 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμώμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

ΣΕΝΑΡΙΟ 7

Για 3500 Εποχές: Καλύτερο σφάλμα 0,9301 για 8 νευρώνες. Η συσχέτιση μεταξύ πραγματικών και εκτιμώμενων τιμών στην περίπτωση αυτή είναι 0,86999.

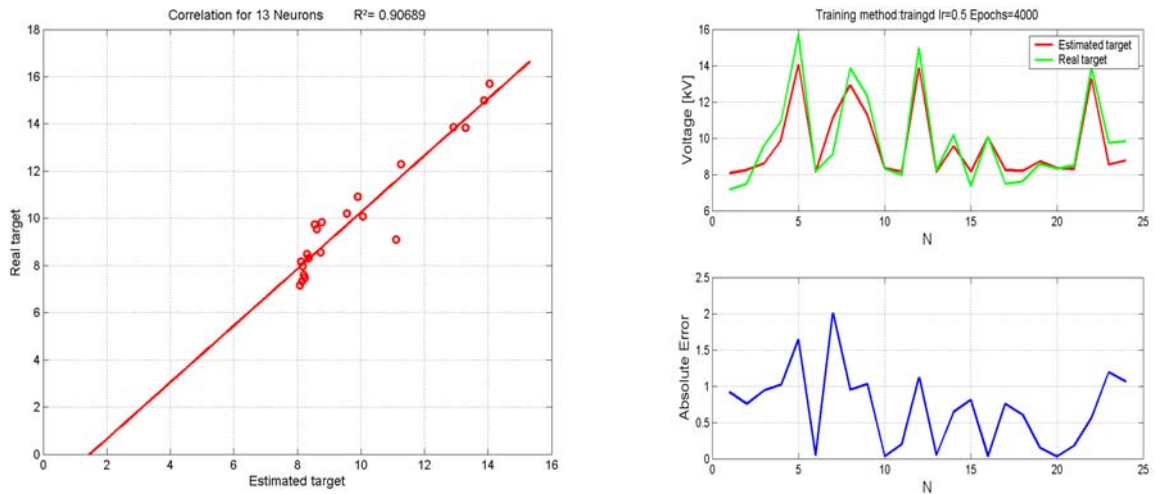


Σχήμα 7.7

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμώμενης τιμής για το σύνολο ελέγχου για 8 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμώμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

ΣΕΝΑΡΙΟ 8

Για 4000 Εποχές: Μικρότερο σφάλμα 0,7618 για 13 νευρώνες. Η συσχέτιση μεταξύ πραγματικών και εκτιμώμενων τιμών στην περίπτωση αυτή είναι 0,90689.

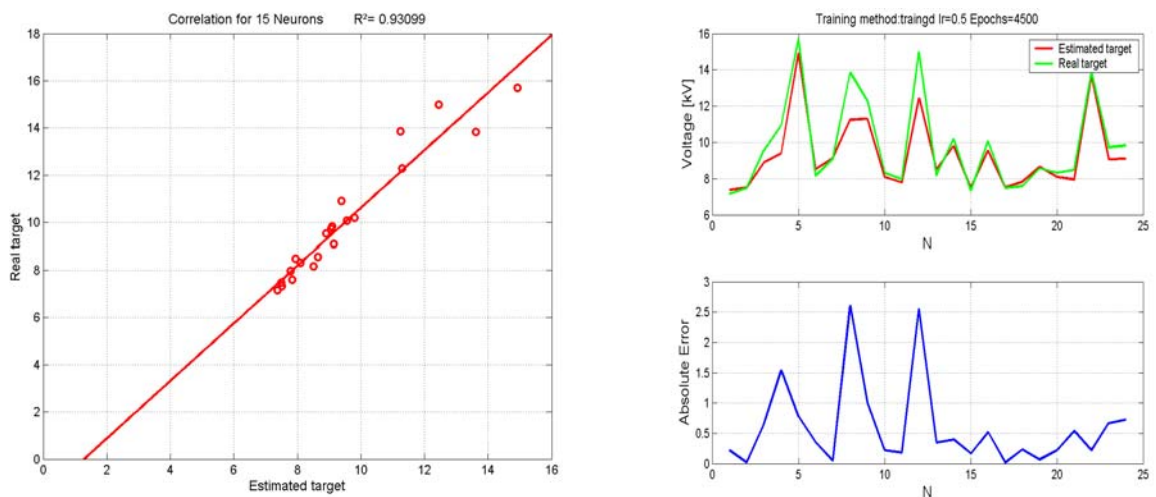


Σχήμα 7.8

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμώμενης τιμής για το σύνολο ελέγχου για 13 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμώμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

ΣΕΝΑΡΙΟ 9

Για 4500 Εποχές: Μικρότερο σφάλμα 0,8345 για 15 νευρώνες. Η συσχέτιση μεταξύ πραγματικών και εκτιμώμενων τιμών στην περίπτωση αυτή είναι 0,93099.

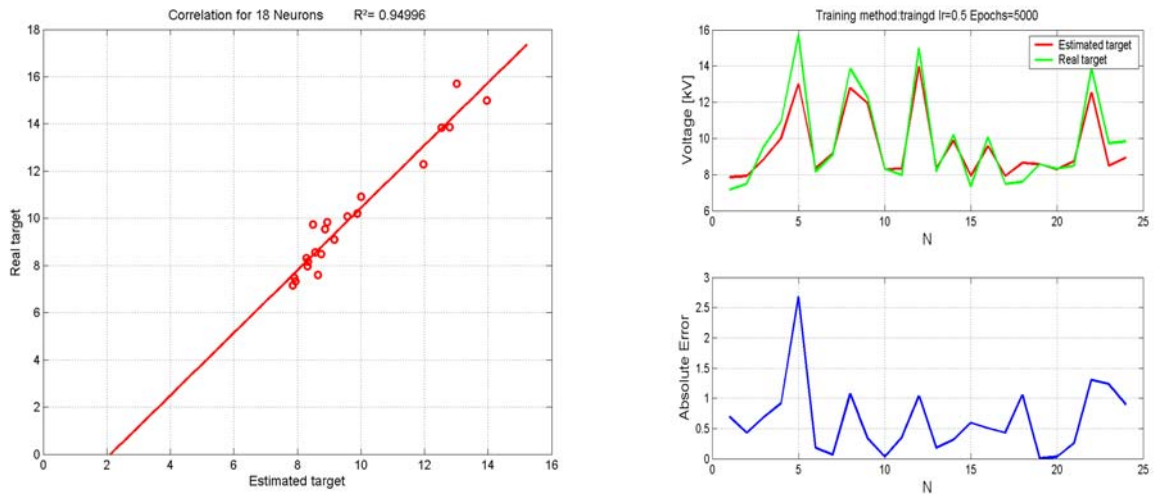


Σχήμα 7.9

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμώμενης τιμής για το σύνολο ελέγχου για 15 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμώμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

ΣΕΝΑΡΙΟ 10

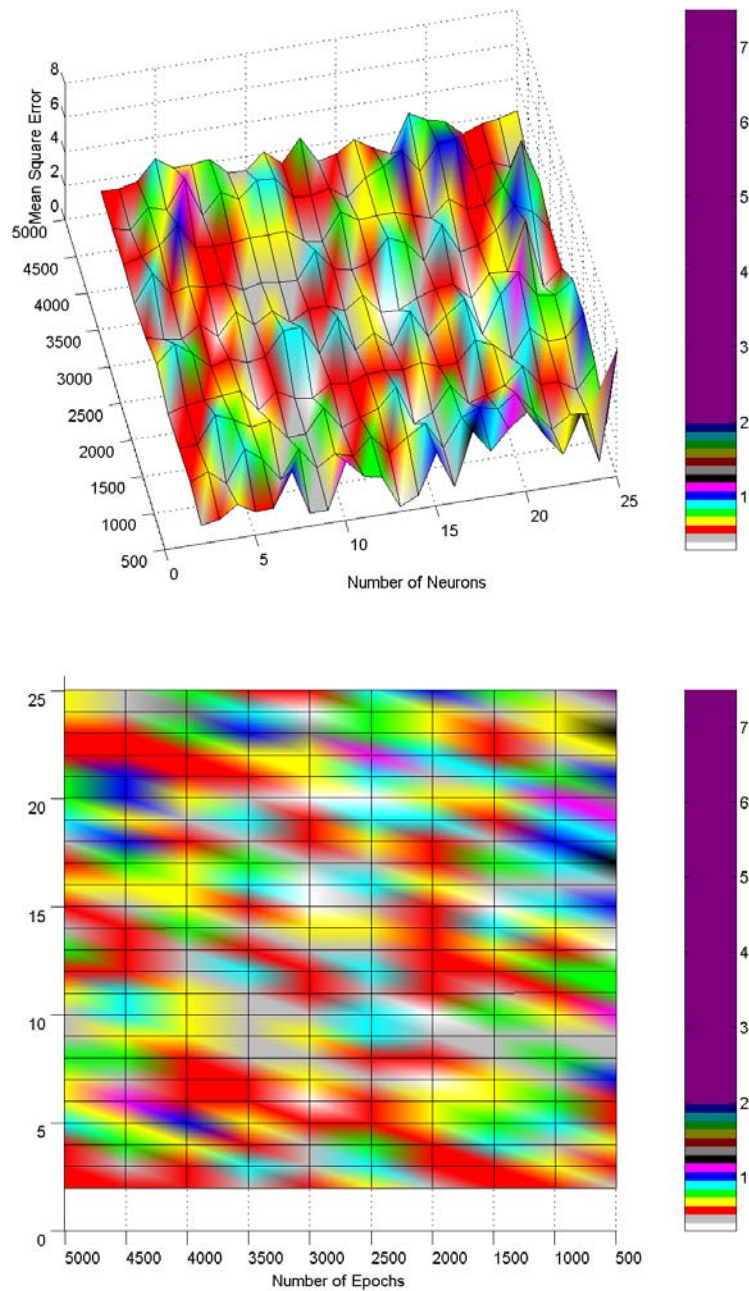
Για 5000 Εποχές: Μικρότερο σφάλμα 0,7425 για 18 νευρώνες. Η συσχέτιση μεταξύ πραγματικών και εκτιμώμενων τιμών στην περίπτωση αυτή είναι 0,94996.



Σχήμα 7.10

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμώμενης τιμής για το σύνολο ελέγχου για 18 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμώμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

Στο σημείο αυτό, αφού τελείωσαν τα σενάρια για την traingd, παρουσιάζεται η τρισδιάστατη καμπύλη για το μέσο τετραγωνικό σφάλμα σαν συνάρτηση των νευρώνων και των εποχών.



Σχήμα 7.11 Τρισδιάστατη καμπύλη για το μέσο τετραγωνικό σφάλμα σαν συνάρτηση των νευρώνων και των εποχών. Γράφημα με προοπτική (πάνω) και κάτοψη γραφήματος (κάτω).

Οι συνδυασμοί που έδωσαν καλύτερα σφάλματα φαίνονται στα παραπάνω γραφήματα σαν περιοχές άσπρου ή ανοιχτού γκρι χρώματος. Τέτοιες περιοχές εντοπίζονται για αριθμό νευρώνων από 8-10 και για 2000-3500 εποχές.

7.3 Σενάρια 1-10 για μέθοδο εκπαίδευσης την traingda

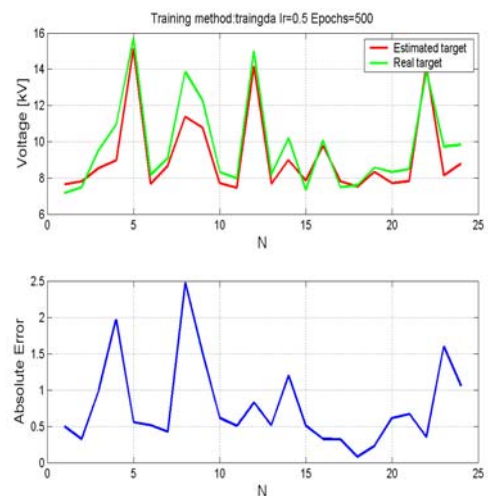
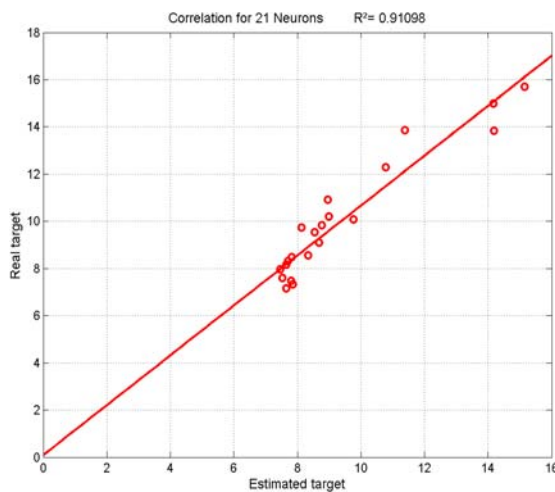
Η μέθοδος εκπαίδευσης traingda εκπαιδεύει το νευρωνικό δίκτυο με βάση τον κανόνα ανάστροφης διάδοσης σφάλματος, κάνοντας παράλληλα προσαρμογή στο ρυθμό μάθησης (gradient descent with adaptive learning rate backpropagation). Τα στοιχεία του δικτύου είναι:

- 1 επίπεδο
- Μεταβολή νευρώνων: από 2 έως 25
- Ρυθμός μάθησης: 0,5
- Μεταβολή εποχών: από 500 έως 5000 με βήμα 500.

ΣΕΝΑΡΙΟ 1

Για 500 Εποχές: Το μικρότερο σφάλμα έχει τιμή 0,9414 και εμφανίζεται για 21 νευρώνες.

Η συσχέτιση μεταξύ πραγματικών και εκτιμώμενων τιμών στην περίπτωση αυτή είναι 0,91098.

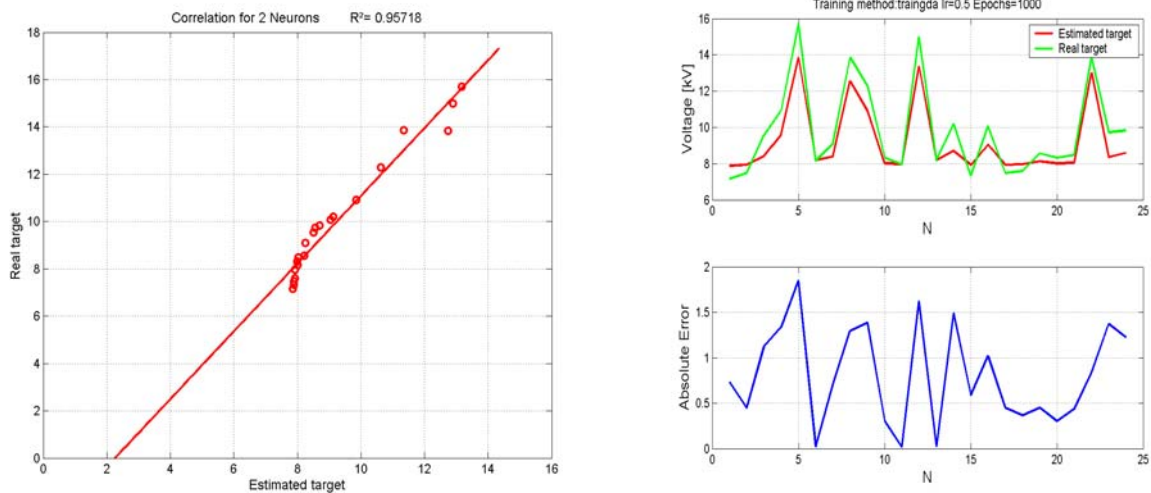


Σχήμα 7.12

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμώμενης τιμής για το σύνολο ελέγχου για 21 νευρώνες.
 Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμώμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

ΣΕΝΑΡΙΟ 2

Για 1000 Εποχές: Μικρότερο σφάλμα 0,9396 για 2 νευρώνες. Η συσχέτιση μεταξύ πραγματικών και εκτιμώμενων τιμών στην περίπτωση αυτή είναι 0,95718.

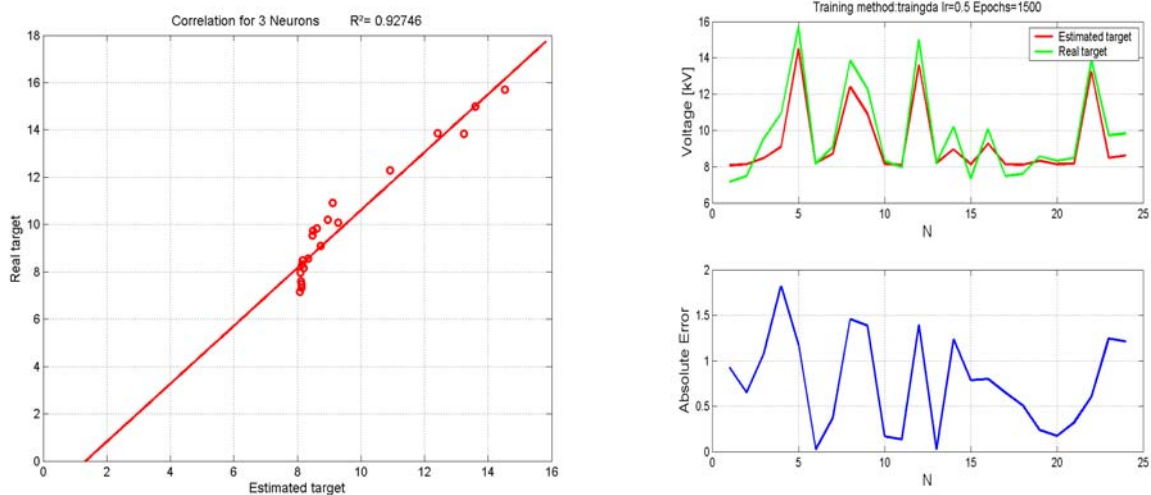


Σχήμα 7.13

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμώμενης τιμής για το σύνολο ελέγχου για 2 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμώμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

ΣΕΝΑΡΙΟ 3:

Για 1500 Εποχές: Μικρότερο σφάλμα 0,8473 για 3 νευρώνες. Η συσχέτιση μεταξύ πραγματικών και εκτιμώμενων τιμών στην περίπτωση αυτή είναι 0,92746.

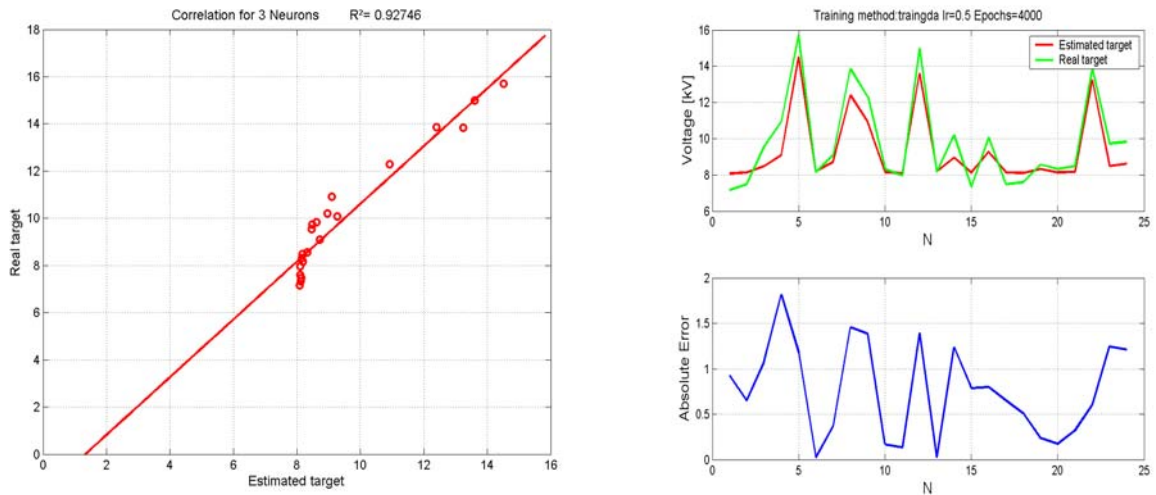


Σχήμα 7.14

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμώμενης τιμής για το σύνολο ελέγχου για 3 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμώμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

ΣΕΝΑΡΙΟ 4:

Για 2000 Εποχές: Μικρότερο σφάλμα 0,8473 για 3 νευρώνες. Η συσχέτιση μεταξύ πραγματικών και εκτιμώμενων τιμών στην περίπτωση αυτή είναι 0,92746.

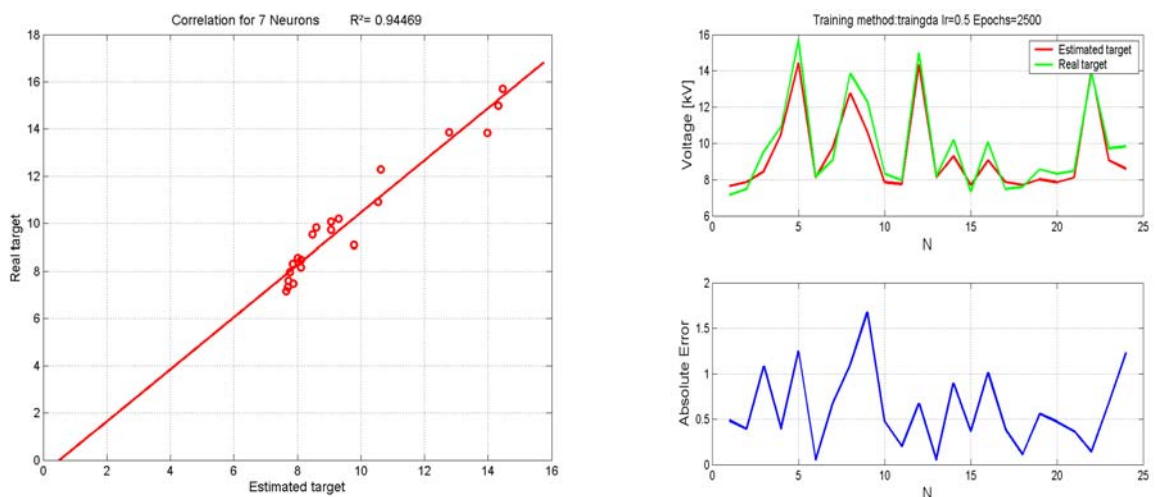


Σχήμα 7.15

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμώμενης τιμής για το σύνολο ελέγχου για 3 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμώμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

ΣΕΝΑΡΙΟ 5

Για 2500 Εποχές: Μικρότερο σφάλμα 0,5523 για 7 νευρώνες. Η συσχέτιση μεταξύ πραγματικών και εκτιμώμενων τιμών στην περίπτωση αυτή είναι 0,94469

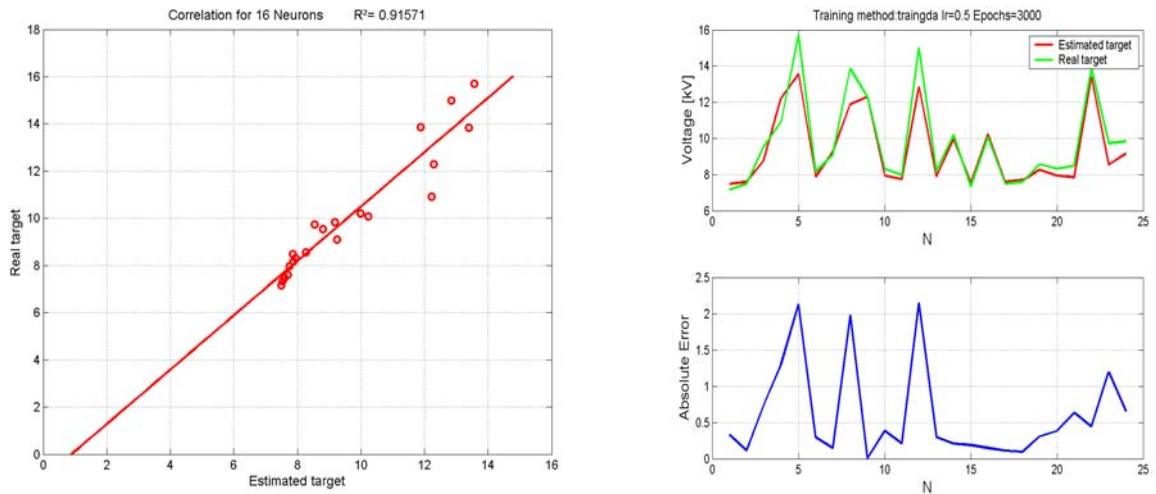


Σχήμα 7.16

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμώμενης τιμής για το σύνολο ελέγχου για 7 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμώμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

ΣΕΝΑΡΙΟ 6

Για 3000 Εποχές: Μικρότερο σφάλμα 0,7784 για 16 νευρώνες. Η συσχέτιση μεταξύ πραγματικών και εκτιμώμενων τιμών στην περίπτωση αυτή είναι 0,91571.

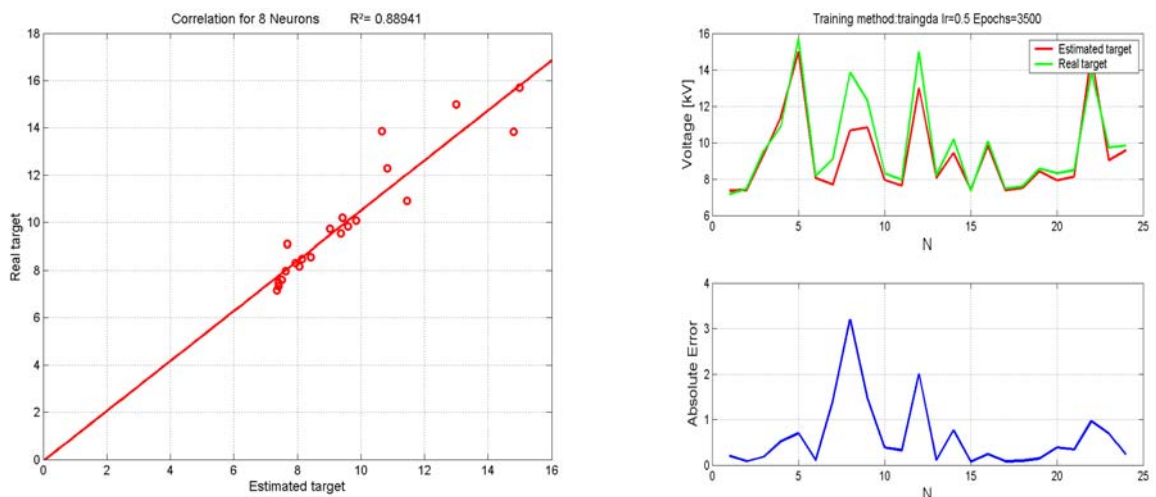


Σχήμα 7.17

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμώμενης τιμής για το σύνολο ελέγχου για 16 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμώμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

ΣΕΝΑΡΙΟ 7

Για 3500 Εποχές: Μικρότερο σφάλμα 0.9147 για 8 νευρώνες. Η συσχέτιση μεταξύ πραγματικών και εκτιμώμενων τιμών στην περίπτωση αυτή είναι 0,88941.

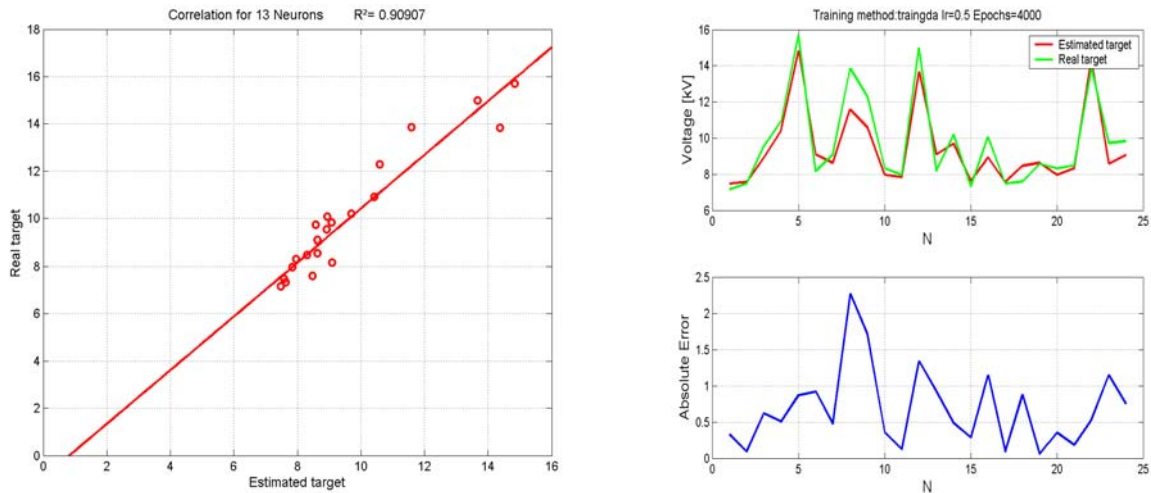


Σχήμα 7.18

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμώμενης τιμής για το σύνολο ελέγχου για 8 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμώμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

ΣΕΝΑΡΙΟ 8

Για 4000 Εποχές: Μικρότερο σφάλμα 0,7587 για 13 νευρώνες. Η συσχέτιση μεταξύ πραγματικών και εκτιμώμενων τιμών στην περίπτωση αυτή είναι 0,90907.

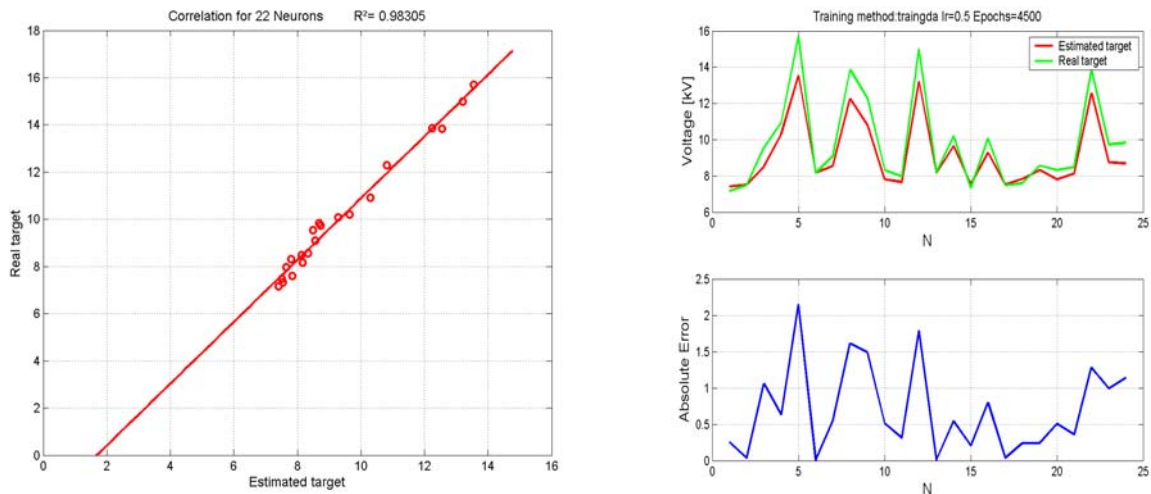


Σχήμα 7.19

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμώμενης τιμής για το σύνολο ελέγχου για 13 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμώμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

ΣΕΝΑΡΙΟ 9

Για 4500 Εποχές: Μικρότερο σφάλμα 0,8453 για 22 νευρώνες. Η συσχέτιση μεταξύ πραγματικών και εκτιμώμενων τιμών στην περίπτωση αυτή είναι 0,98305.

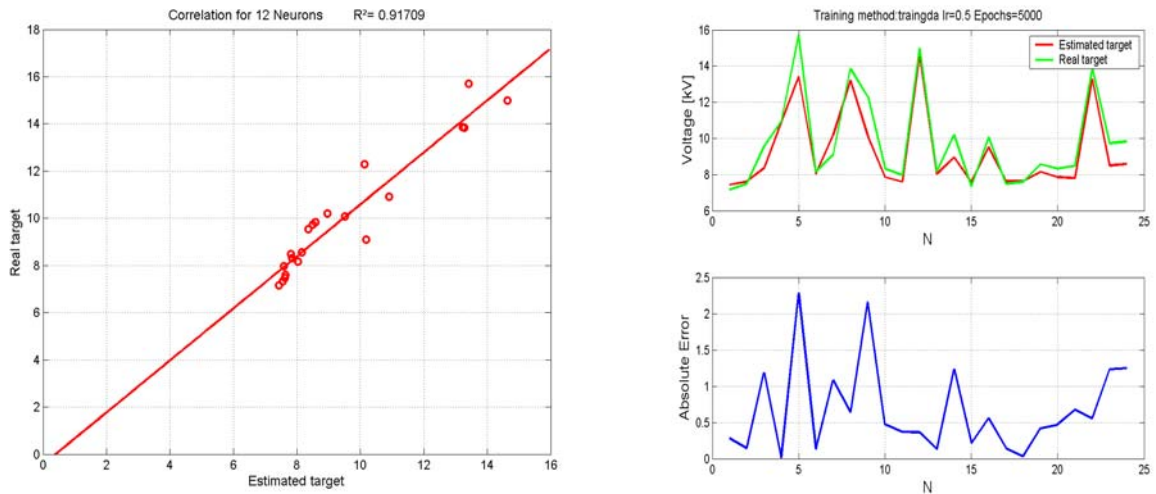


Σχήμα 7.20

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμώμενης τιμής για το σύνολο ελέγχου για 22 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμώμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

ΣΕΝΑΡΙΟ 10

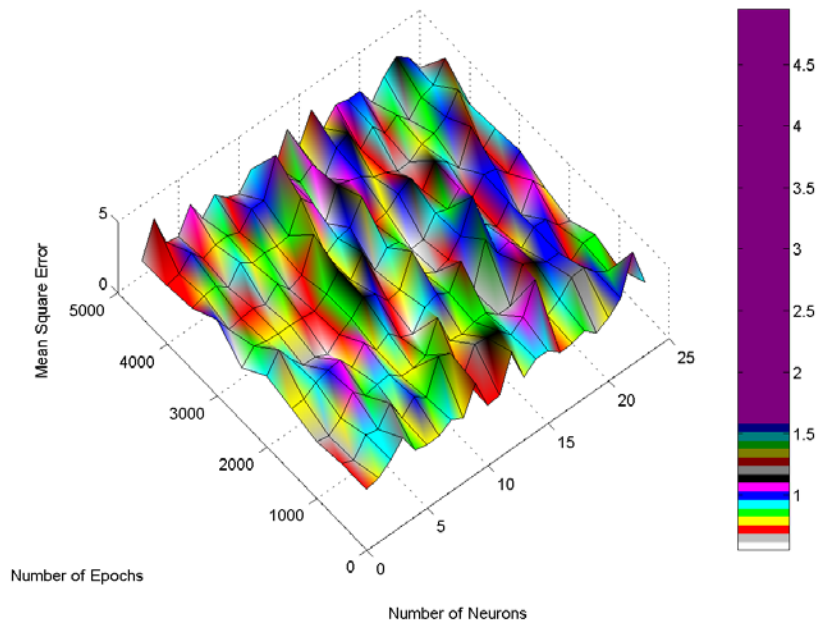
Για 5000 Εποχές: Μικρότερο σφάλμα 0,8199 για 12 νευρώνες. Η συσχέτιση μεταξύ πραγματικών και εκτιμώμενων τιμών στην περίπτωση αυτή είναι 0,91709.

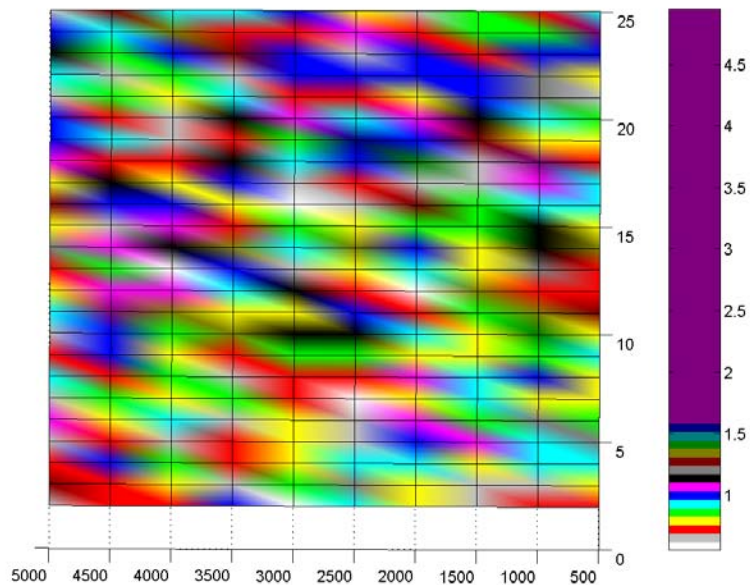


Σχήμα 7.21

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμώμενης τιμής για το σύνολο ελέγχου για 12 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμώμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

Το τρισδιάστατο γράφημα για την περίπτωση της traingda φαίνεται παρακάτω:





7.22 Τρισδιάστατη καμπύλη για το μέσο τετραγωνικό σφάλμα σαν συνάρτηση των νευρώνων και των εποχών. Γράφημα με προοπτική (πάνω) και κάτοψη γραφήματος (κάτω).

Γενικά τα σφάλματα είναι μικρότερα απ' ότι στην περίπτωση της `traingd`, ενώ αρκετά ικανοποιητικά αποτελέσματα λαμβάνονται για μικρό αριθμό νευρώνων και αριθμό εποχών από 1500-3000.

7.4 Σενάρια 1-10 για μέθοδο εκπαίδευσης την `traingdx`

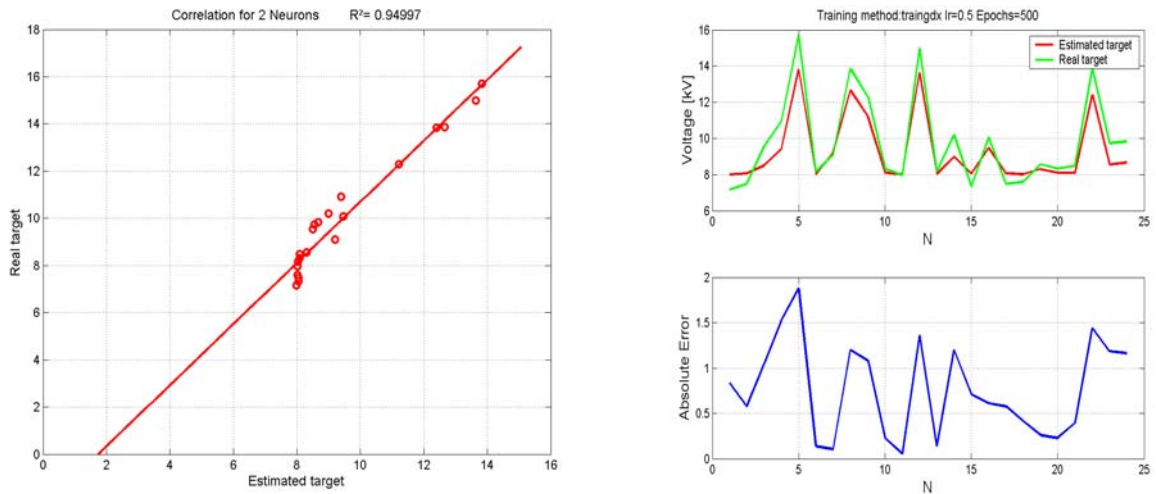
Η `traingdx` κάνει ό,τι ακριβώς και η `traingda`, εισάγοντας επιπλέον και όρο ορμής (gradient descent with momentum and adaptive learning rate backpropagation).

Τα στοιχεία του δικτύου είναι και πάλι:

- 1 επίπεδο
- Μεταβολή νευρώνων: από 2 έως 25
- Ρυθμός μάθησης: 0,5
- Μεταβολή εποχών: από 500 έως 5000 με βήμα 500.

ΣΕΝΑΡΙΟ 1

Για 500 Εποχές: Μικρότερο σφάλμα 0,8543 για 2 νευρώνες. Η συσχέτιση μεταξύ πραγματικών και εκτιμώμενων τιμών στην περίπτωση αυτή είναι 0,94997.

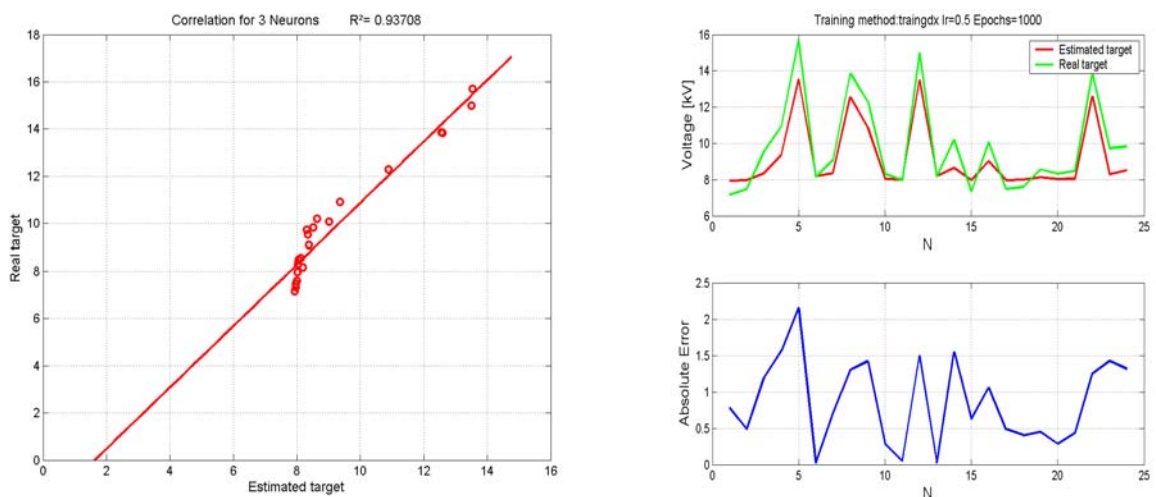


Σχήμα 7.23

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμώμενης τιμής για το σύνολο ελέγχου για 2 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμώμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

ΣΕΝΑΡΙΟ 2

Για 1000 Εποχές: Μικρότερο σφάλμα 1,0874 για 3 νευρώνες. Η συσχέτιση μεταξύ πραγματικών και εκτιμώμενων τιμών στην περίπτωση αυτή είναι 0,93708.

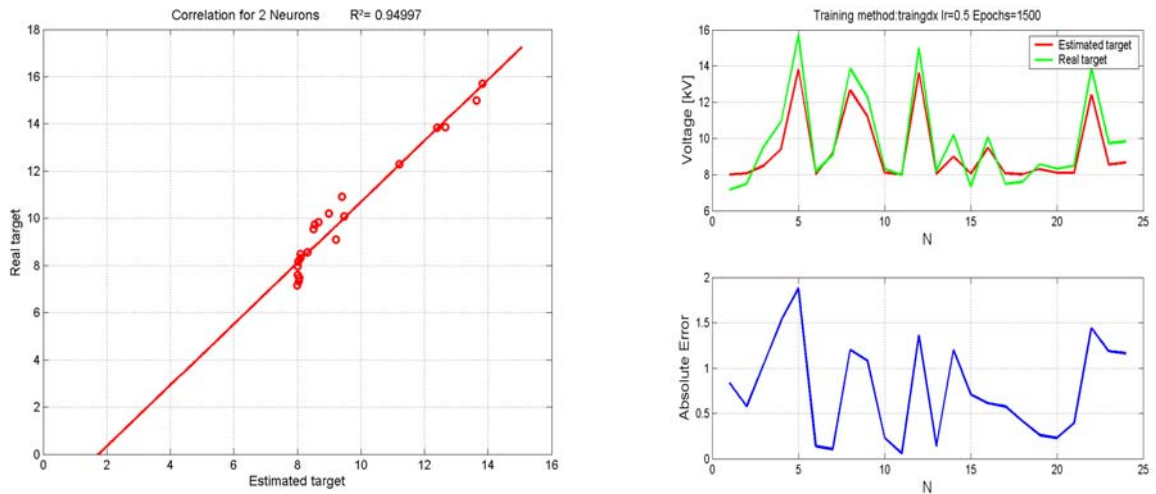


Σχήμα 7.24

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμώμενης τιμής για το σύνολο ελέγχου για 3 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμώμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

ΣΕΝΑΡΙΟ 3

Για 1500 Εποχές: Μικρότερο σφάλμα 0,8543 για 2 νευρώνες. Η συσχέτιση μεταξύ πραγματικών και εκτιμώμενων τιμών στην περίπτωση αυτή είναι 0,94997.

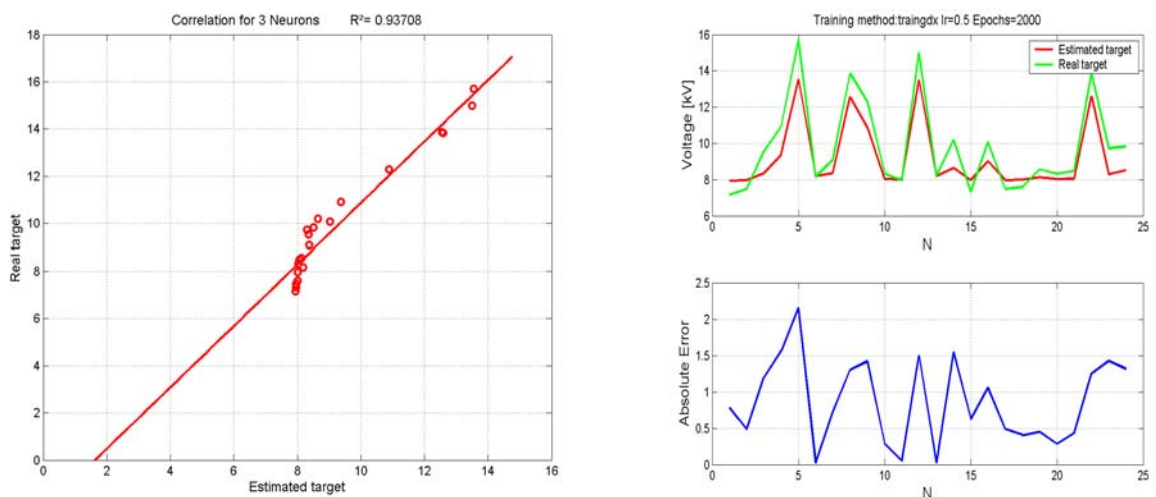


Σχήμα 7.25

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμώμενης τιμής για το σύνολο ελέγχου για 2 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμώμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

ΣΕΝΑΡΙΟ 4

Για 2000 Εποχές: Μικρότερο σφάλμα 1,0874 για 3 νευρώνες. Η συσχέτιση μεταξύ πραγματικών και εκτιμώμενων τιμών στην περίπτωση αυτή είναι 0,93708.

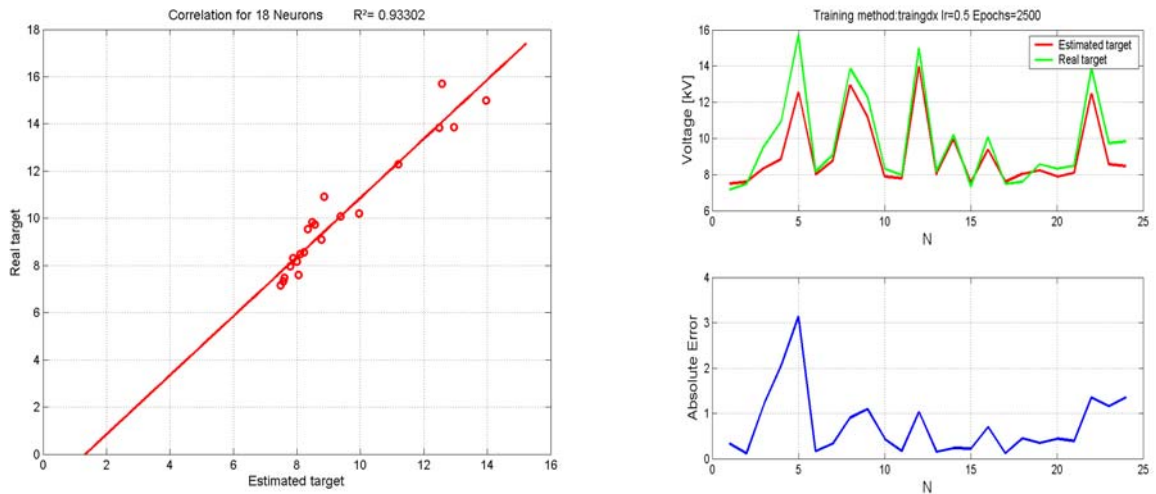


Σχήμα 7.26

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμώμενης τιμής για το σύνολο ελέγχου για 3 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμώμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

ΣΕΝΑΡΙΟ 5

Για 2500 Εποχές: Μικρότερο σφάλμα 1,0599 για 18 νευρώνες. Η συσχέτιση μεταξύ πραγματικών και εκτιμώμενων τιμών στην περίπτωση αυτή είναι 0,93302.

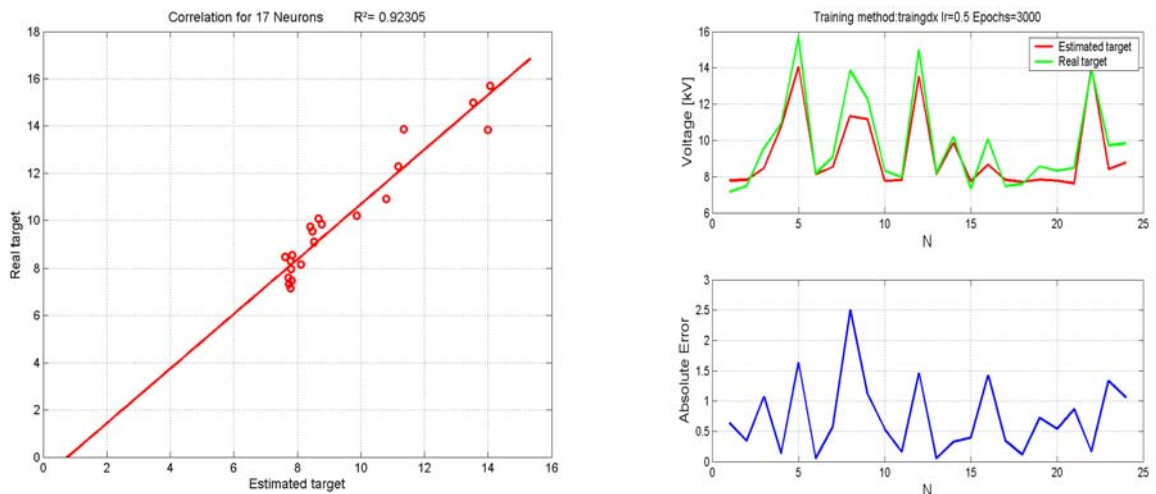


Σχήμα 7.27

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμώμενης τιμής για το σύνολο ελέγχου για 18 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμώμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

ΣΕΝΑΡΙΟ 6

Για 3000 Εποχές: Μικρότερο σφάλμα 0,9 για 17 νευρώνες. Η συσχέτιση μεταξύ πραγματικών και εκτιμώμενων τιμών στην περίπτωση αυτή είναι 0,92305

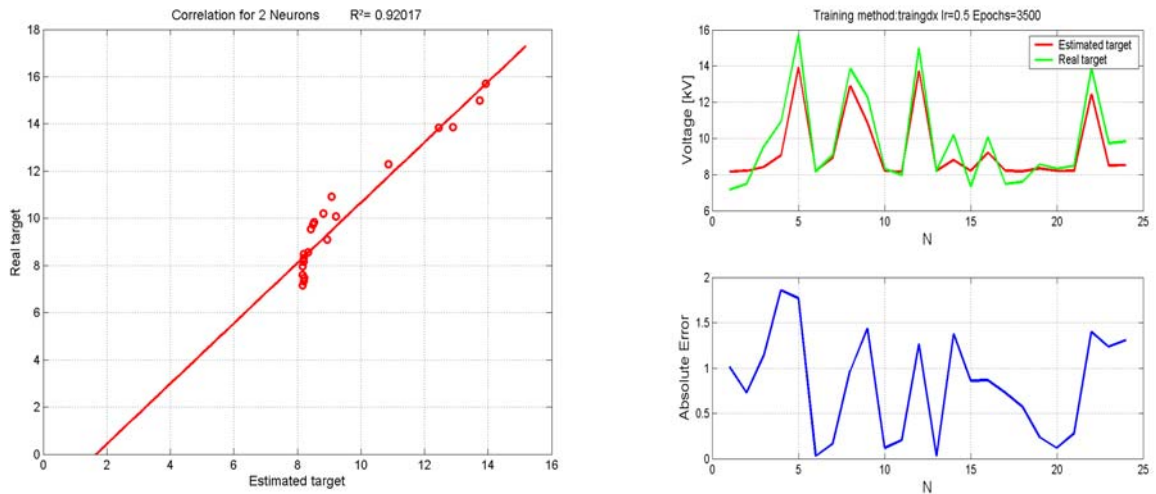


Σχήμα 7.28

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμώμενης τιμής για το σύνολο ελέγχου για 17 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμώμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

ΣΕΝΑΡΙΟ 7

Για 3500 Εποχές: Μικρότερο σφάλμα 0,9866 για 2 νευρώνες. Η συσχέτιση μεταξύ πραγματικών και εκτιμούμενων τιμών στην περίπτωση αυτή είναι 0,92017.

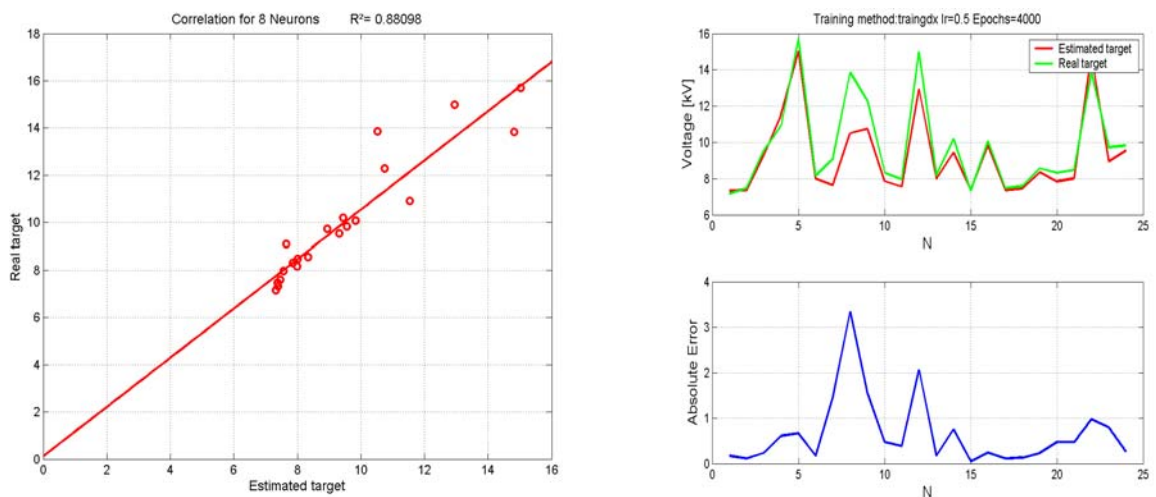


Σχήμα 7.29

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμούμενης τιμής για το σύνολο ελέγχου για 2 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμούμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

ΣΕΝΑΡΙΟ 8

Για 4000 Εποχές: Μικρότερο σφάλμα 1,0057 για 8 νευρώνες. Η συσχέτιση μεταξύ πραγματικών και εκτιμούμενων τιμών στην περίπτωση αυτή είναι 0,88098.

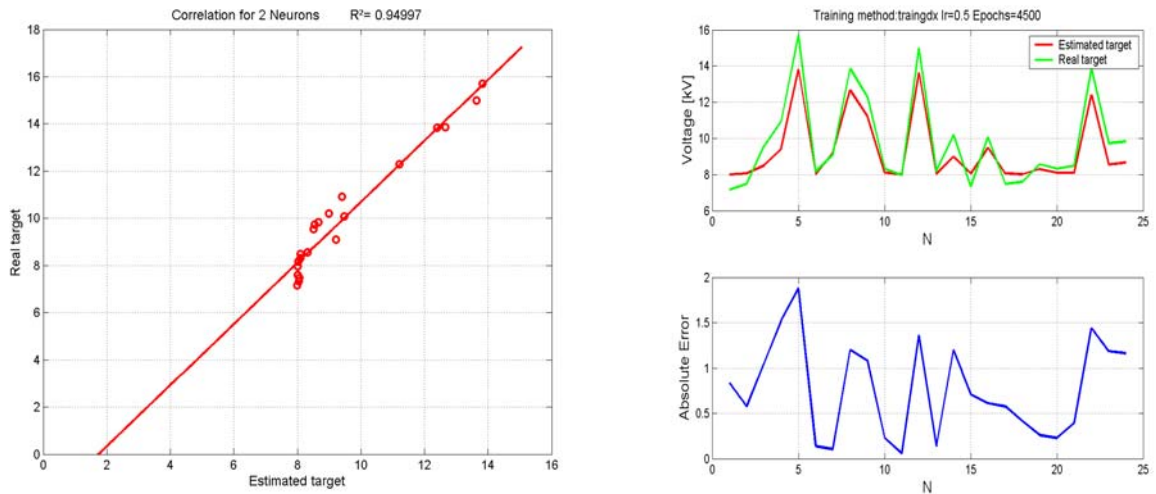


Σχήμα 7.30

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμούμενης τιμής για το σύνολο ελέγχου για 8 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμούμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

ΣΕΝΑΡΙΟ 9

Για 4500 Εποχές: Μικρότερο σφάλμα 0,8543 για 2 νευρώνες. Η συσχέτιση μεταξύ πραγματικών και εκτιμούμενων τιμών στην περίπτωση αυτή είναι 0,94997.

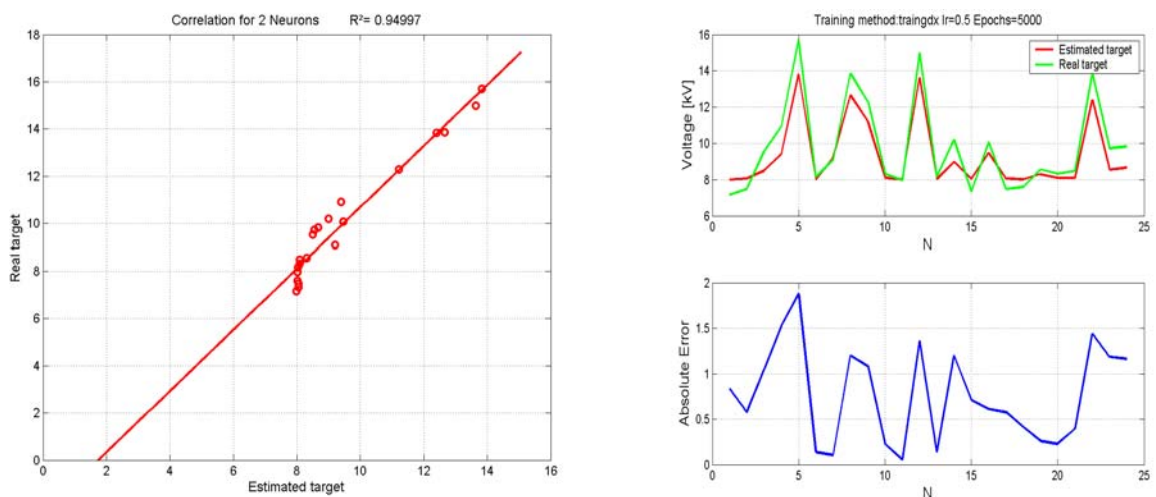


Σχήμα 7.31

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμούμενης τιμής για το σύνολο ελέγχου για 2 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμούμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

ΣΕΝΑΡΙΟ 10

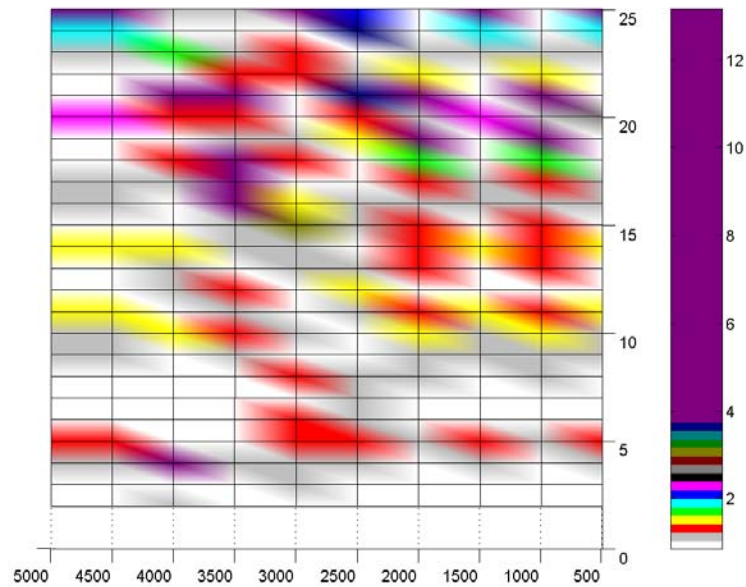
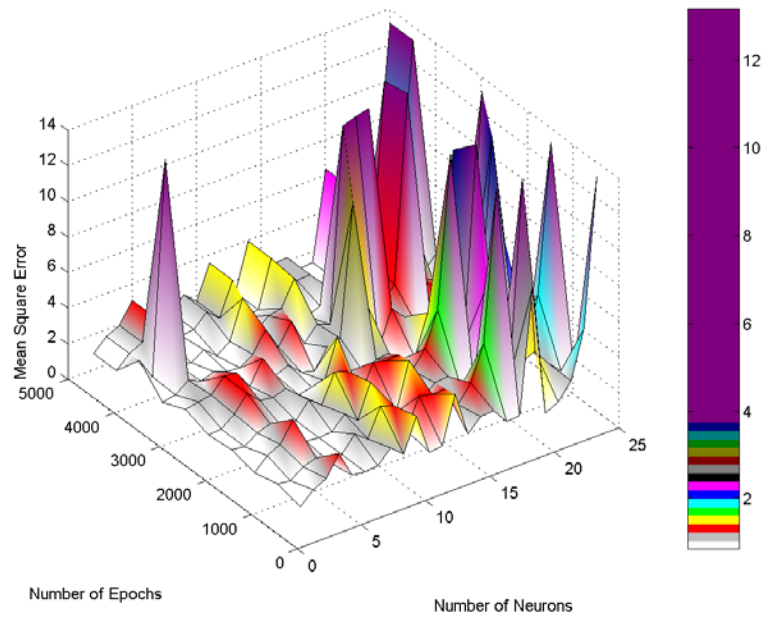
Για 5000 Εποχές: Μικρότερο σφάλμα 0,8543 για 2 νευρώνες. Η συσχέτιση μεταξύ πραγματικών και εκτιμούμενων τιμών στην περίπτωση αυτή είναι 0,94997



Σχήμα 7.32

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμούμενης τιμής για το σύνολο ελέγχου για 2 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμούμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

Το τρισδιάστατο γράφημα για την περίπτωση της `trainngdx` φαίνεται παρακάτω:



7.33 Τρισδιάστατη καμπύλη για το μέσο τετραγωνικό σφάλμα σαν συνάρτηση των νευρώνων και των εποχών. Γράφημα με προοπτική (πάνω) και κάτοψη γραφήματος (κάτω).

Παρατηρείται ότι υπάρχουν περιοχές που η `trainngdx` δίνει πολύ μεγαλύτερα σφάλματα απ' τις δύο προηγούμενες συναρτήσεις, αλλά για συγκεκριμένη αρχιτεκτονική δικτύου, η

trainIdx δίνει πολύ καλά αποτελέσματα. Συγκεκριμένα, για αριθμό νευρώνων μικρότερο από 10 και μέχρι τις 3000 εποχές, τα σφάλματα είναι πολύ μικρά.

7.5 Σενάρια 1-10 για μέθοδο εκπαίδευσης την trainlm

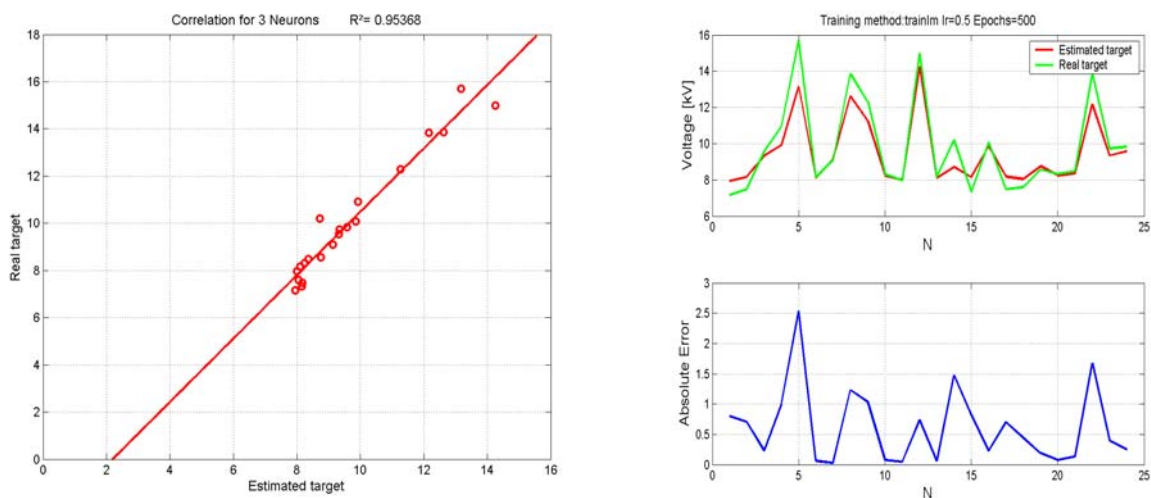
Η trainlm εκπαιδεύει το νευρωνικό δίκτυο με βάση τον κανόνα ανάστροφης διάδοσης σφάλματος κατά Levenberg-Marquardt (Levenberg-Marquardt backpropagation).

Τα στοιχεία δικτύου είναι:

- 1 επίπεδο
- Μεταβολή νευρώνων: από 2 έως 25
- Ρυθμός μάθησης: 0,5
- Μεταβολή εποχών: από 500 έως 5000 με βήμα 500.

ΣΕΝΑΡΙΟ 1

Για 500 Εποχές: Μικρότερο σφάλμα 0,7669 για 3 νευρώνες. Η συσχέτιση μεταξύ πραγματικών και εκτιμώμενων τιμών στην περίπτωση αυτή είναι 0,95368.

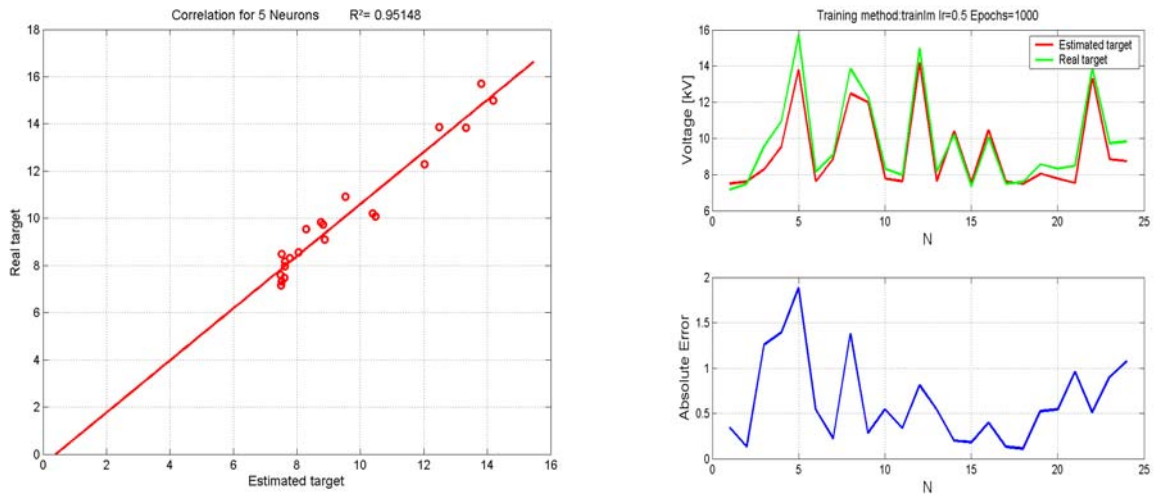


Σχήμα 7.34

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμώμενης τιμής για το σύνολο ελέγχου για 3 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμώμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

ΣΕΝΑΡΙΟ 2

Για 1000 Εποχές: Μικρότερο σφάλμα 0,6198 για 5 νευρώνες. Η συσχέτιση μεταξύ πραγματικών και εκτιμούμενων τιμών στην περίπτωση αυτή είναι 0,95148.

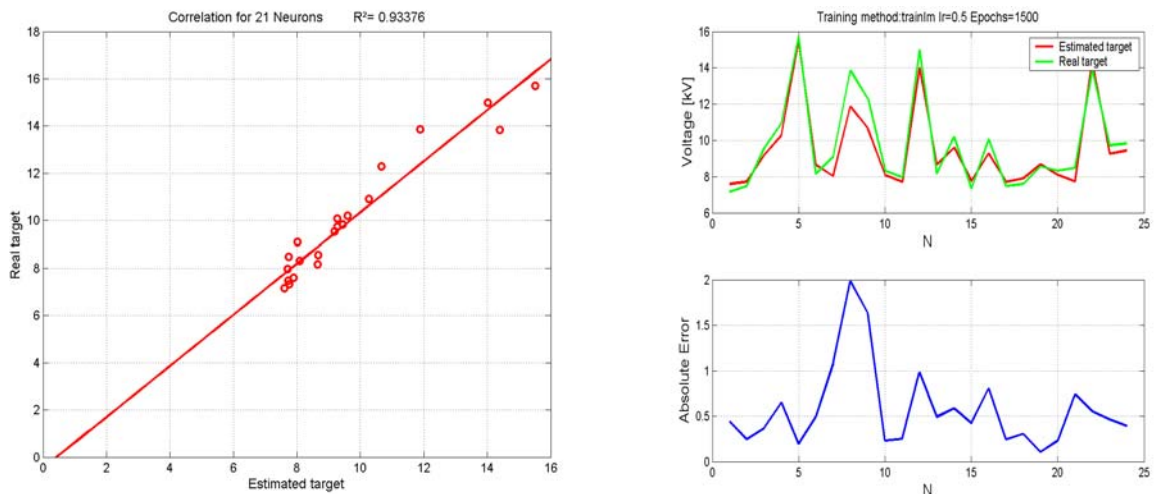


Σχήμα 7.35

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμούμενης τιμής για το σύνολο ελέγχου για 5 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμούμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

ΣΕΝΑΡΙΟ 3

Για 1500 Εποχές: Μικρότερο σφάλμα 0,5327 για 21 νευρώνες. Η συσχέτιση μεταξύ πραγματικών και εκτιμούμενων τιμών στην περίπτωση αυτή είναι 0,93376.

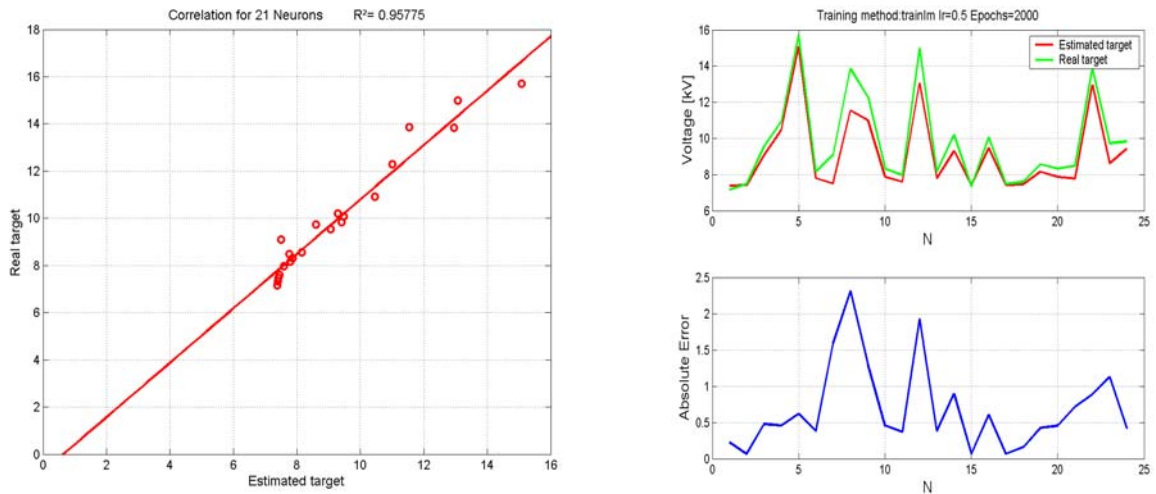


Σχήμα 7.36

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμούμενης τιμής για το σύνολο ελέγχου για 21 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμούμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

ΣΕΝΑΡΙΟ 4

Για 2000 Εποχές: Μικρότερο σφάλμα 0,7980 για 21 νευρώνες. Η συσχέτιση μεταξύ πραγματικών και εκτιμώμενων τιμών στην περίπτωση αυτή είναι 0,95775.

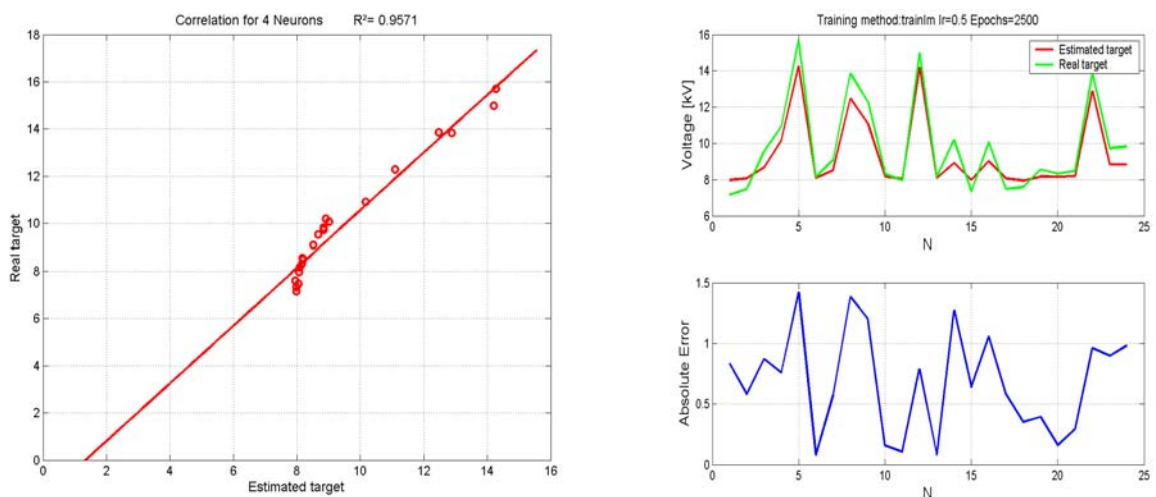


Σχήμα 7.37

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμώμενης τιμής για το σύνολο ελέγχου για 21 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμώμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

ΣΕΝΑΡΙΟ 5

Για 2500 Εποχές: Μικρότερο σφάλμα 0.6385 για 4 νευρώνες. Η συσχέτιση μεταξύ πραγματικών και εκτιμώμενων τιμών στην περίπτωση αυτή είναι 0.9571.

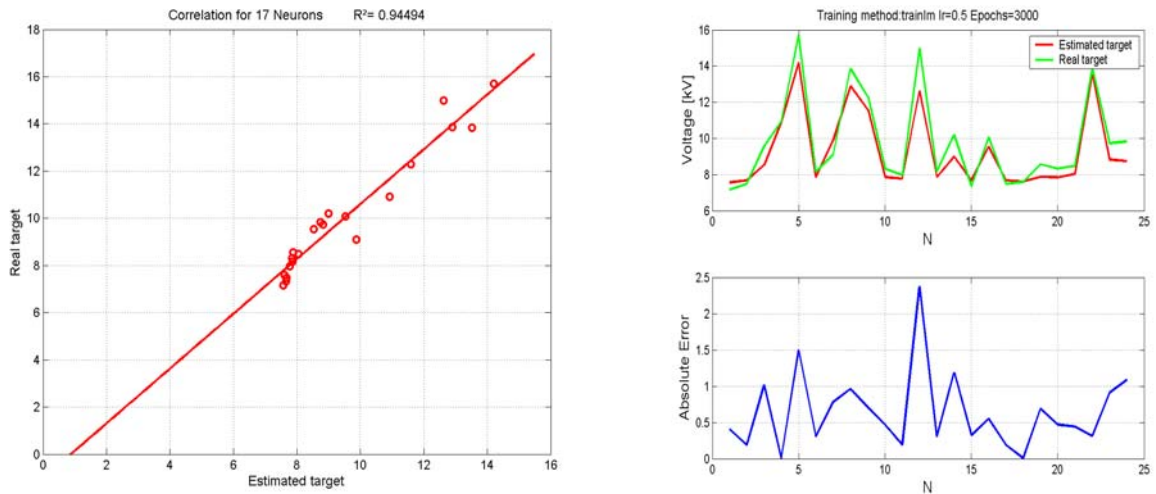


Σχήμα 7.38

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμώμενης τιμής για το σύνολο ελέγχου για 4 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμώμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

ΣΕΝΑΡΙΟ 6

Για 3000 Εποχές: Μικρότερο σφάλμα 0,6858 για 17 νευρώνες. Η συσχέτιση μεταξύ πραγματικών και εκτιμώμενων τιμών στην περίπτωση αυτή είναι 0,94494.

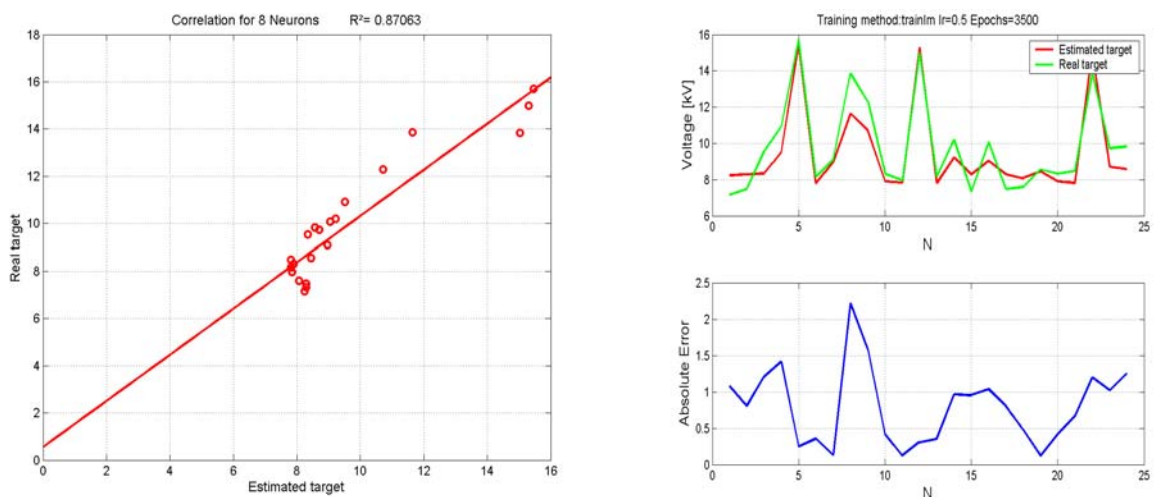


Σχήμα 7.39

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμώμενης τιμής για το σύνολο ελέγχου για 17 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμώμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

ΣΕΝΑΡΙΟ 7

Για 3500 Εποχές: Μικρότερο σφάλμα 0,9095 για 8 νευρώνες. Η συσχέτιση μεταξύ πραγματικών και εκτιμώμενων τιμών στην περίπτωση αυτή είναι 0,87063.

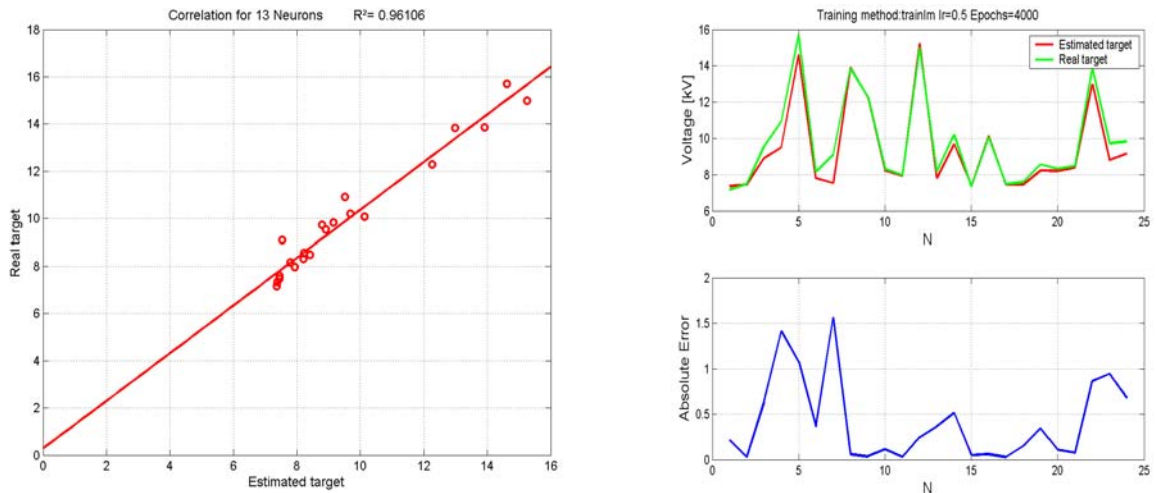


Σχήμα 7.40

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμώμενης τιμής για το σύνολο ελέγχου για 8 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμώμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

ΣΕΝΑΡΙΟ 8

Για 4000 Εποχές: Μικρότερο σφάλμα 0,3719 για 13 νευρώνες. Η συσχέτιση μεταξύ πραγματικών και εκτιμούμενων τιμών στην περίπτωση αυτή είναι 0,96106.

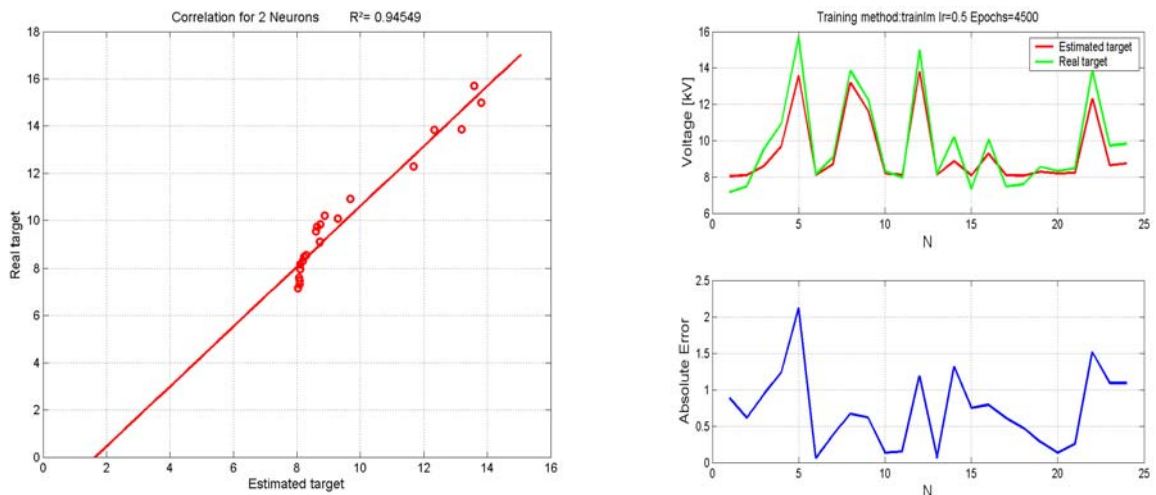


Σχήμα 7.41

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμούμενης τιμής για το σύνολο ελέγχου για 13 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμούμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

ΣΕΝΑΡΙΟ 9

Για 4500 Εποχές: Μικρότερο σφάλμα 0,7870 για 2 νευρώνες. Η συσχέτιση μεταξύ πραγματικών και εκτιμούμενων τιμών στην περίπτωση αυτή είναι 0,94549.

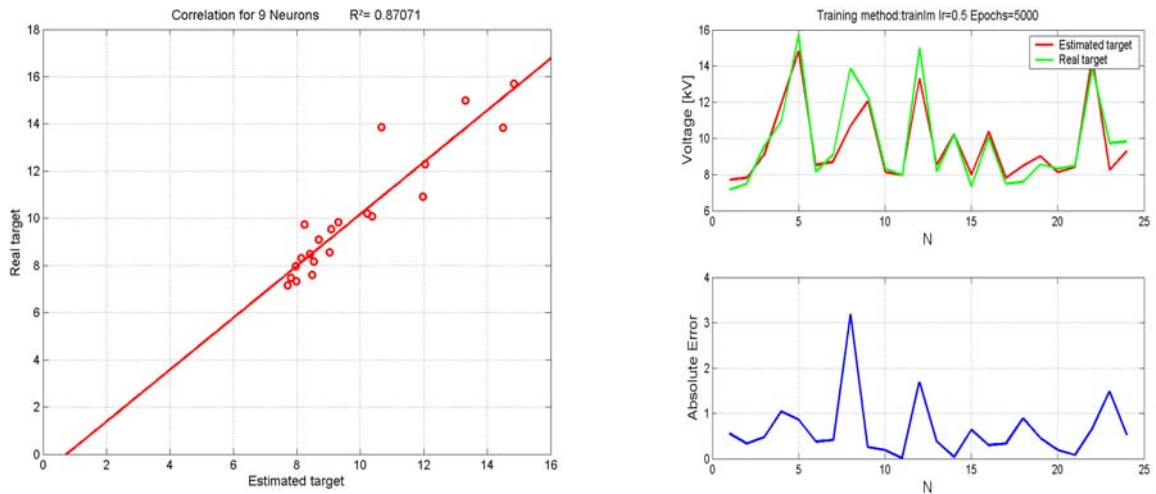


Σχήμα 7.42

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμούμενης τιμής για το σύνολο ελέγχου για 2 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμούμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

ΣΕΝΑΡΙΟ 10

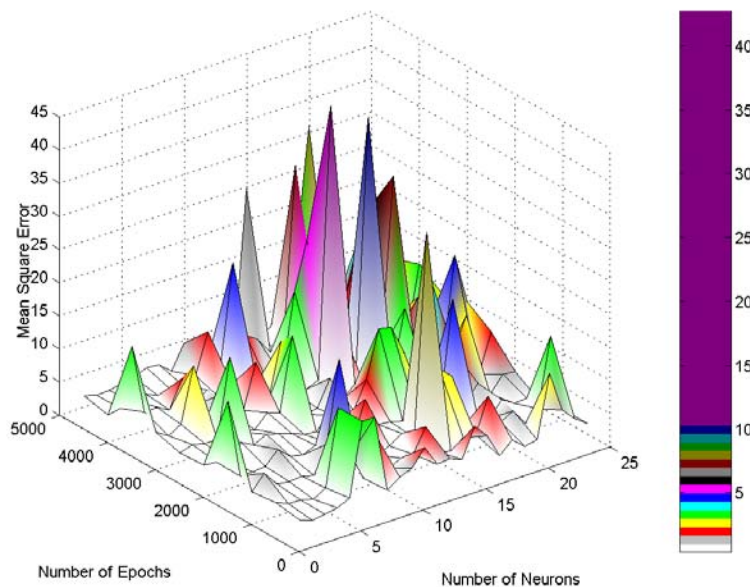
Για 5000 Εποχές: Μικρότερο σφάλμα 0,8601 για 9 νευρώνες. Η συσχέτιση μεταξύ πραγματικών και εκτιμώμενων τιμών στην περίπτωση αυτή είναι 0,87071.

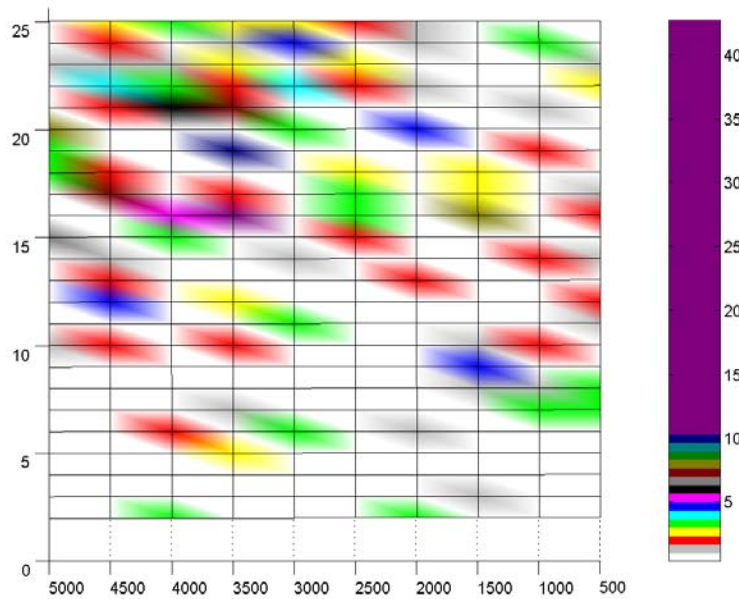


Σχήμα 7.43

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμώμενης τιμής για το σύνολο ελέγχου για 9 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμώμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

Το τρισδιάστατο γράφημα για την trainlm είναι το εξής:





7.44 Τρισδιάστατη καμπύλη για το μέσο τετραγωνικό σφάλμα σαν συνάρτηση των νευρώνων και των εποχών. Γράφημα με προοπτική (πάνω) και κάτοψη γραφήματος (κάτω).

Είναι προφανές ότι η `trainlm` για μεγάλο πλήθος νευρώνων ή/και για πολλές εποχές δεν συγκλίνει και δίνει πολύ μεγάλα σφάλματα, τα οποία βέβαια δεν μπορούν να γίνουν αποδεκτά. Όμως, η `trainlm` για μικρό αριθμό νευρώνων και κυρίως για λίγες εποχές φαίνεται να δίνει καλά αποτελέσματα, ενώ παρουσιάζει και το πλεονέκτημα ότι είναι πιο γρήγορη μέθοδος από τις προηγούμενες.

7.6 Σχολιασμός των αποτελεσμάτων

Αρχικά πρέπει να αναφερθεί ότι οι μέθοδοι που χρησιμοποιήθηκαν παραπάνω είναι κάποιες απ' τις μεθόδους που περιέχονται στο πακέτο των νευρωνικών του MATLAB. Ενδεικτικά χρησιμοποιήθηκαν τέσσερις μέθοδοι που χρησιμοποιούν τον αλγόριθμο ανάστροφης διάδοσης σφάλματος (`traingd`) ή παραλλαγές αυτού (`traingda`, `traingdx`, `trainlm`). Φαίνεται ότι ορισμένες απ' αυτές δεν συγκλίνουν (π.χ. `trainlm`), ενώ άλλες θα μπορούσαν να δώσουν καλύτερα αποτελέσματα, αν μεταβαλλόταν ο ρυθμός μάθησης. Το επόμενο βήμα είναι λοιπόν να γίνουν δοκιμές στις οποίες ο αριθμός των νευρώνων θα περιοριστεί στα πλαίσια που φαίνεται να δίνουν καλά αποτελέσματα, ο αριθμός τον

εποχών θα κρατηθεί σταθερός και θα μεταβάλλεται ο ρυθμός μάθησης. Στις δοκιμές που θα γίνουν, ο ρυθμός μάθησης θα κυμανθεί από 0,1-0,9 με βήμα 0,05.

7.7 Μεταβάλλοντας το ρυθμό μάθησης

Σε κάθε μία απ' τις παρακάτω δοκιμές ο αριθμός εποχών κρατήθηκε σταθερός και οι νευρώνες κυμάνθηκαν στο εύρος που έδινε τα καλύτερα αποτελέσματα με βάση τα τρισδιάστατα γραφήματα.

Η μέθοδος που δοκιμάστηκε ήταν η `trainngdx`, διότι έδωσε τα καλύτερα αποτελέσματα.

ΔΟΚΙΜΕΣ ΜΕ `trainngdx`

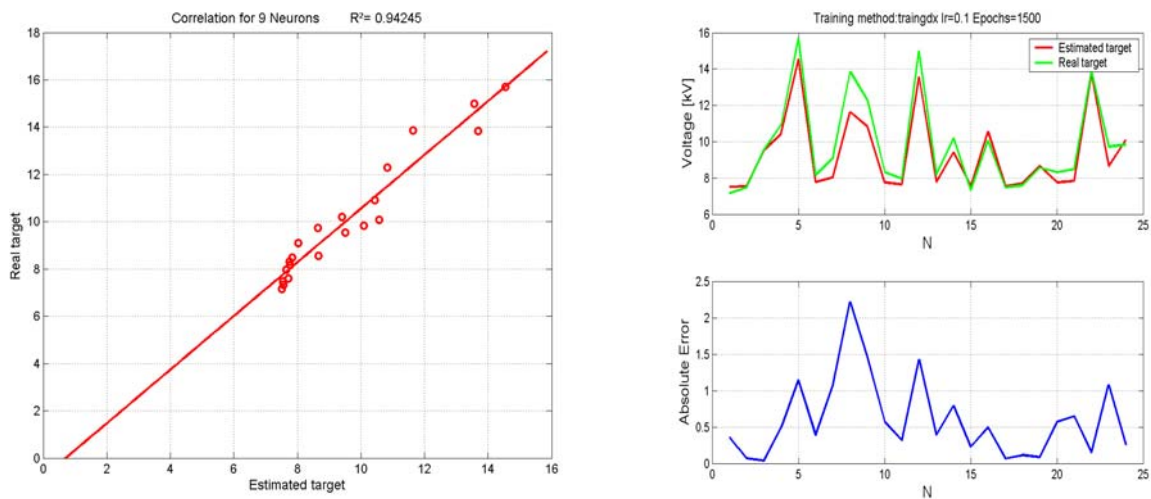
ΔΟΚΙΜΗ 1:

Αριθμός νευρώνων: μεταβολή από 6-12

Αριθμός εποχών: σταθερά 1500 εποχές

Ρυθμός μάθησης: 0,1

Ελάχιστο σφάλμα 0,6512 για 9 νευρώνες. Συσχέτιση για την περίπτωση αυτή: 0,94245.



Σχήμα 7.45

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμούμενης τιμής για το σύνολο ελέγχου για 9 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμούμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

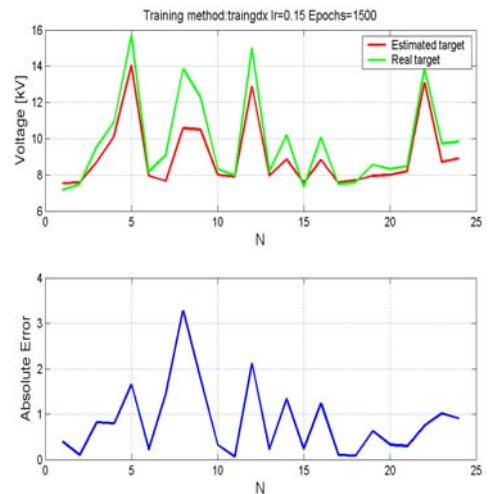
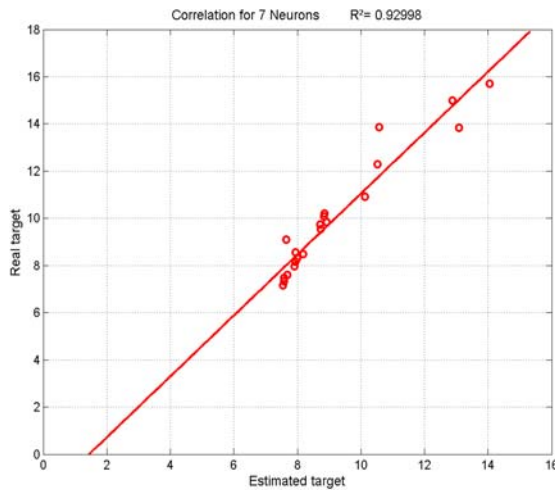
ΔΟΚΙΜΗ 2:

Αριθμός νευρώνων: μεταβολή από 6-12

Αριθμός εποχών: σταθερά 1500 εποχές

Ρυθμός μάθησης: 0,15

Ελάχιστο σφάλμα 1,3063 για 7 νευρώνες. Συσχέτιση για την περίπτωση αυτή: 0,93



Σχήμα 7.46

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμούμενης τιμής για το σύνολο ελέγχου για 7 νευρώνες.

Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμούμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

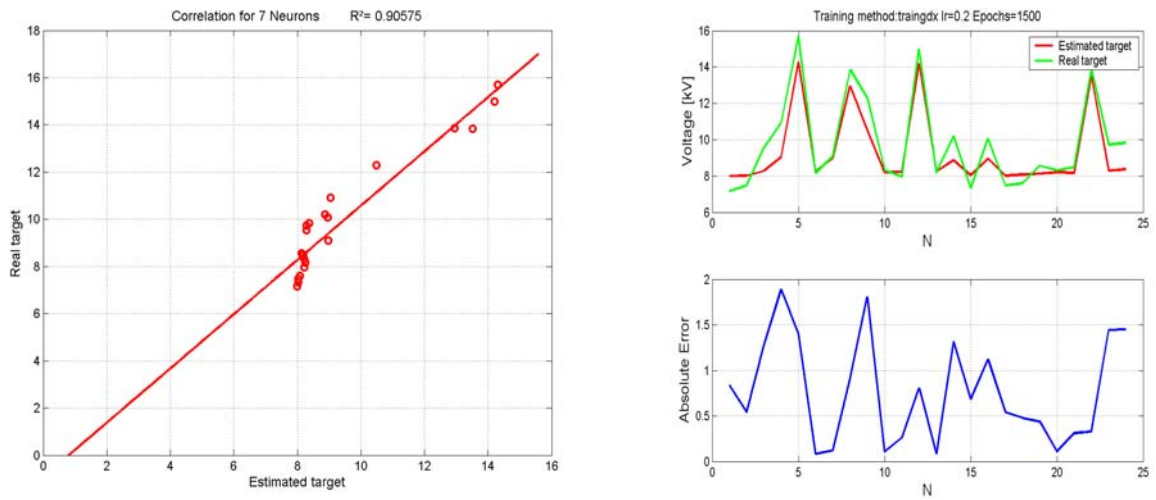
ΔΟΚΙΜΗ 3:

Αριθμός νευρώνων: μεταβολή από 6-12

Αριθμός εποχών: σταθερά 1500 εποχές

Ρυθμός μάθησης: 0,2

Ελάχιστο σφάλμα 0,8987 για 7 νευρώνες. Συσχέτιση για την περίπτωση αυτή: 0,90575



Σχήμα 7.47

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμούμενης τιμής για το σύνολο ελέγχου για 7 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμούμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

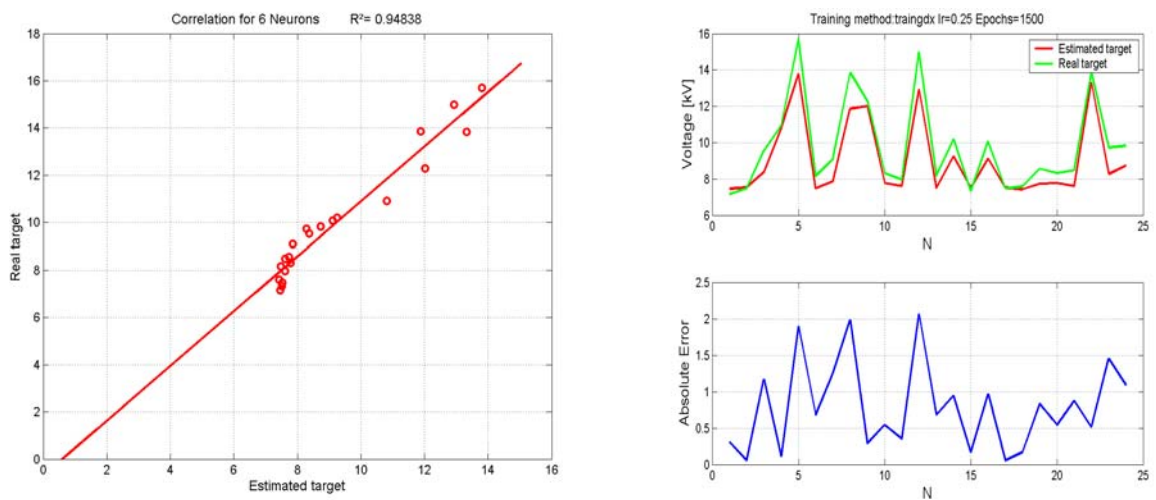
ΔΟΚΙΜΗ 4:

Αριθμός νευρώνων: μεταβολή από 6-12

Αριθμός εποχών: σταθερά 1500 εποχές

Ρυθμός μάθησης: 0,25

Ελάχιστο σφάλμα 0,9803 για 6 νευρώνες. Συσχέτιση για την περίπτωση αυτή: 0,9484



Σχήμα 7.48

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμούμενης τιμής για το σύνολο ελέγχου για 6 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμούμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

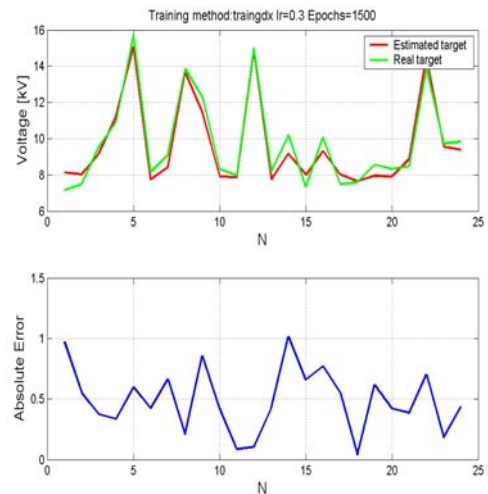
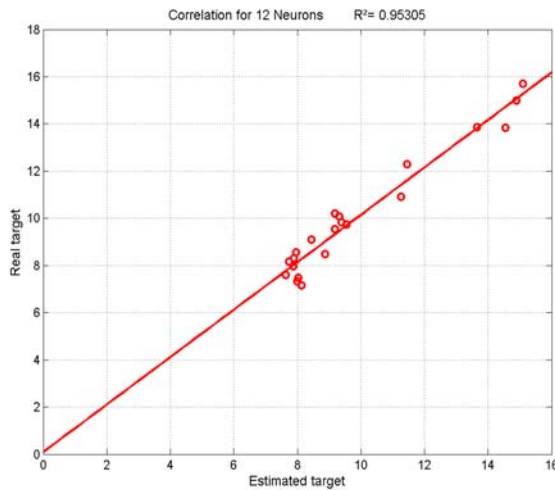
ΔΟΚΙΜΗ 5:

Αριθμός νευρώνων: μεταβολή από 6-12

Αριθμός εποχών: σταθερά 1500 εποχές

Ρυθμός μάθησης: 0,3

Ελάχιστο σφάλμα 0,3099 για 12 νευρώνες. Συσχέτιση για την περίπτωση αυτή: 0,95305



Σχήμα 7.49

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμούμενης τιμής για το σύνολο ελέγχου για 12 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμούμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

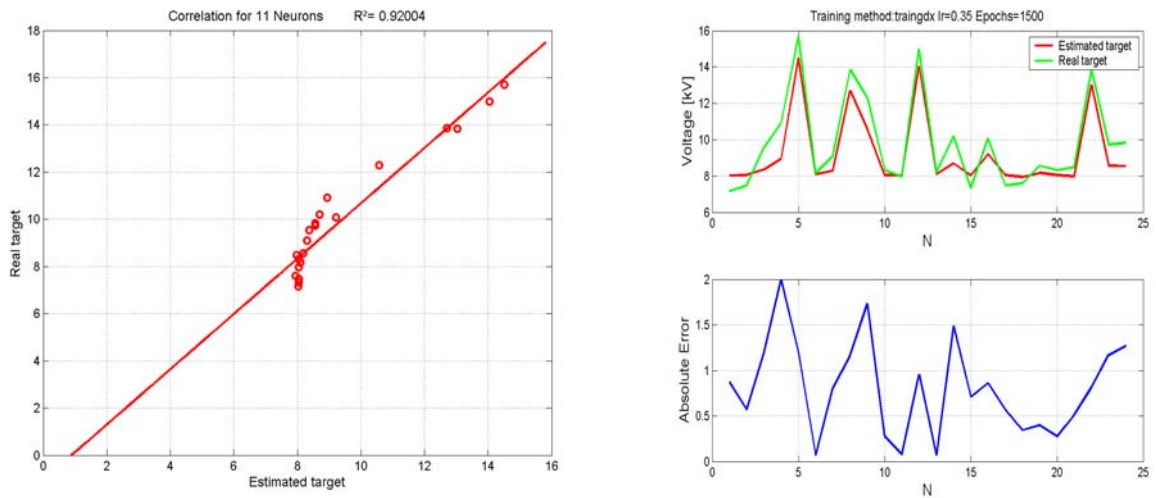
ΔΟΚΙΜΗ 6:

Αριθμός νευρώνων: μεταβολή από 6-12

Αριθμός εποχών: σταθερά 1500 εποχές

Ρυθμός μάθησης: 0,35

Ελάχιστο σφάλμα 0,9132 για 11 νευρώνες. Συσχέτιση για την περίπτωση αυτή: 0,92004



Σχήμα 7.50

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμούμενης τιμής για το σύνολο ελέγχου για 11 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμούμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

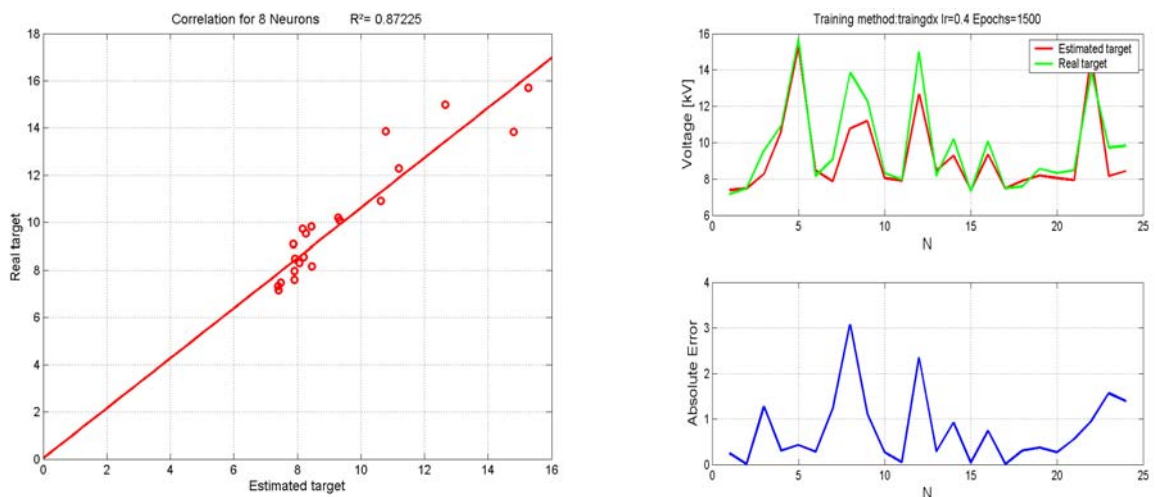
ΔΟΚΙΜΗ 7:

Αριθμός νευρώνων: μεταβολή από 6-12

Αριθμός εποχών: σταθερά 1500 εποχές

Ρυθμός μάθησης: 0,4

Ελάχιστο σφάλμα 1,13 για 8 νευρώνες. Συσχέτιση για την περίπτωση αυτή: 0,87225



Σχήμα 7.51

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμούμενης τιμής για το σύνολο ελέγχου για 8 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμούμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

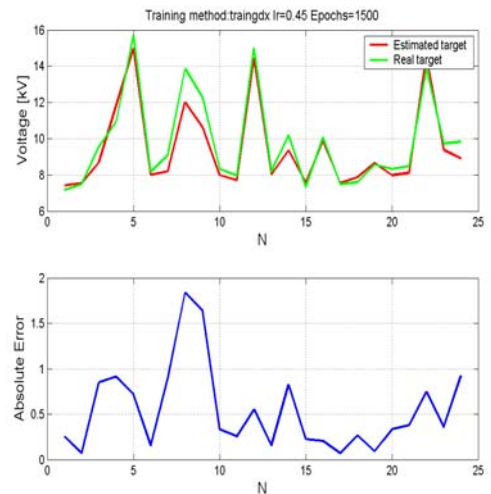
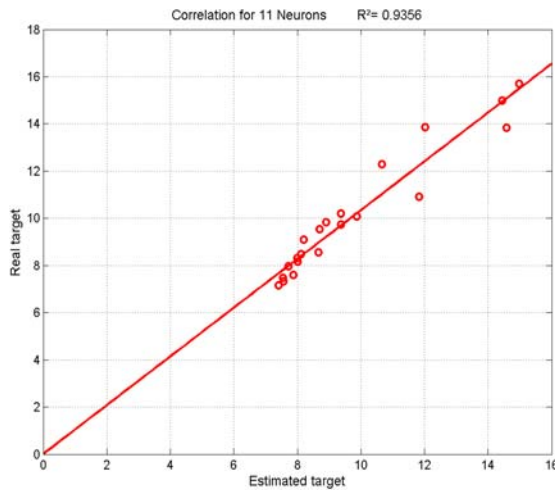
ΔΟΚΙΜΗ 8:

Αριθμός νευρώνων: μεταβολή από 6-12

Αριθμός εποχών: σταθερά 1500 εποχές

Ρυθμός μάθησης: 0,45

Ελάχιστο σφάλμα 0,5085 για 11 νευρώνες. Συσχέτιση για την περίπτωση αυτή: 0,9356



Σχήμα 7.52

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμούμενης τιμής για το σύνολο ελέγχου για 11 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμούμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

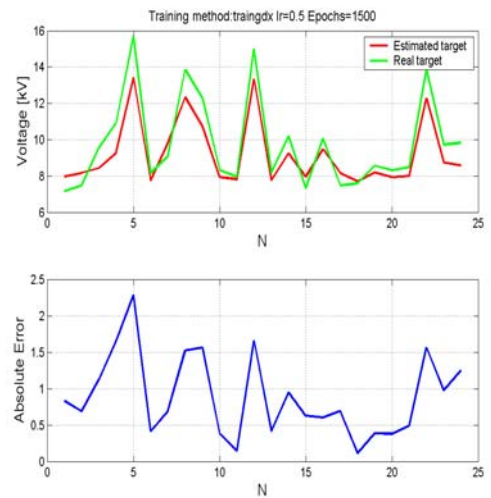
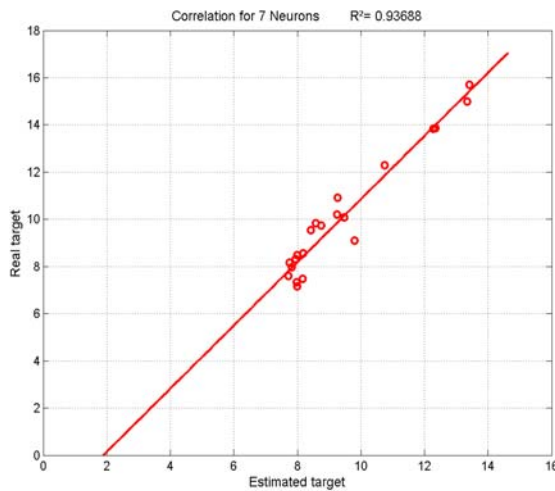
ΔΟΚΙΜΗ 9:

Αριθμός νευρώνων: μεταβολή από 6-12

Αριθμός εποχών: σταθερά 1500 εποχές

Ρυθμός μάθησης: 0,5

Ελάχιστο σφάλμα 1,1054 για 7 νευρώνες. Συσχέτιση για την περίπτωση αυτή: 0,93688



Σχήμα 7.53

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμούμενης τιμής για το σύνολο ελέγχου για 7 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμούμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

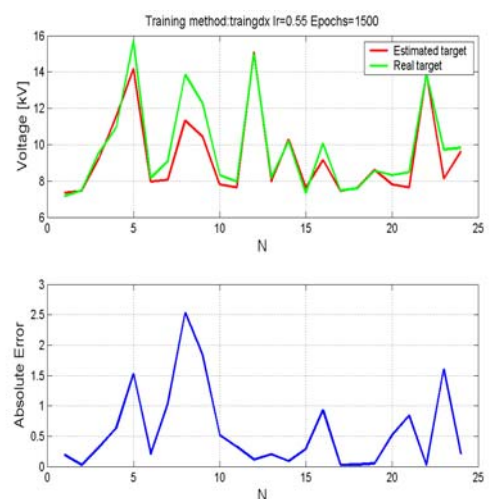
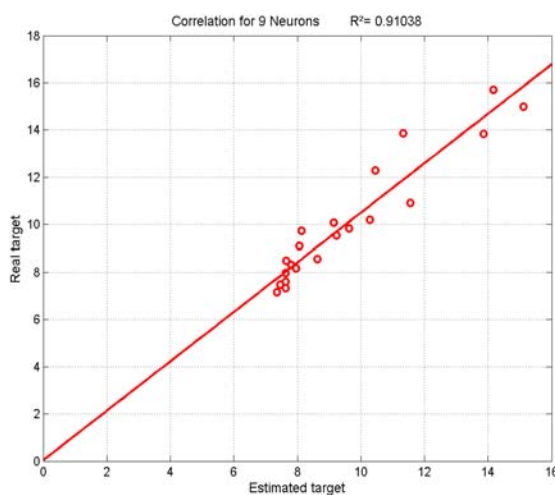
ΔΟΚΙΜΗ 10:

Αριθμός νευρώνων: μεταβολή από 6-12

Αριθμός εποχών: σταθερά 1500 εποχές

Ρυθμός μάθησης: 0,55

Ελάχιστο σφάλμα 0,7829 για 9 νευρώνες. Συσχέτιση για την περίπτωση αυτή: 0,91038



Σχήμα 7.54

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμούμενης τιμής για το σύνολο ελέγχου για 9 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμούμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

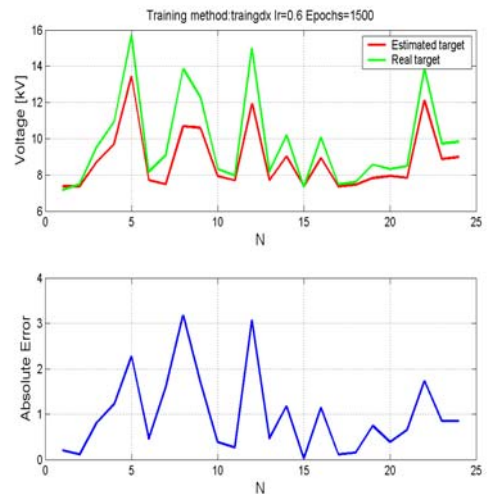
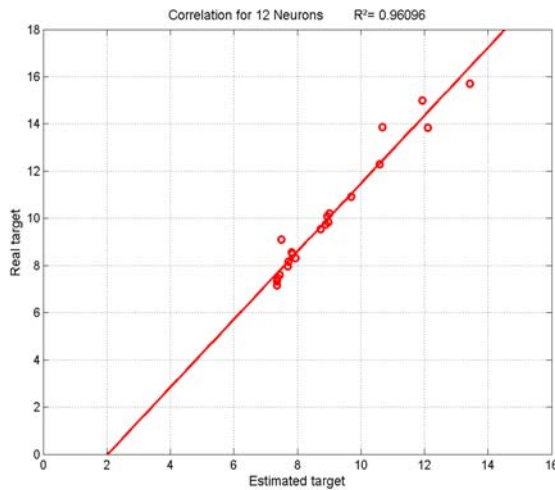
ΔΟΚΙΜΗ 11:

Αριθμός νευρώνων: μεταβολή από 6-12

Αριθμός εποχών: σταθερά 1500 εποχές

Ρυθμός μάθησης: 0,6

Ελάχιστο σφάλμα 1,7219 για 12 νευρώνες. Συσχέτιση για την περίπτωση αυτή: 0,96096



Σχήμα 7.55

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμούμενης τιμής για το σύνολο ελέγχου για 12 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμούμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

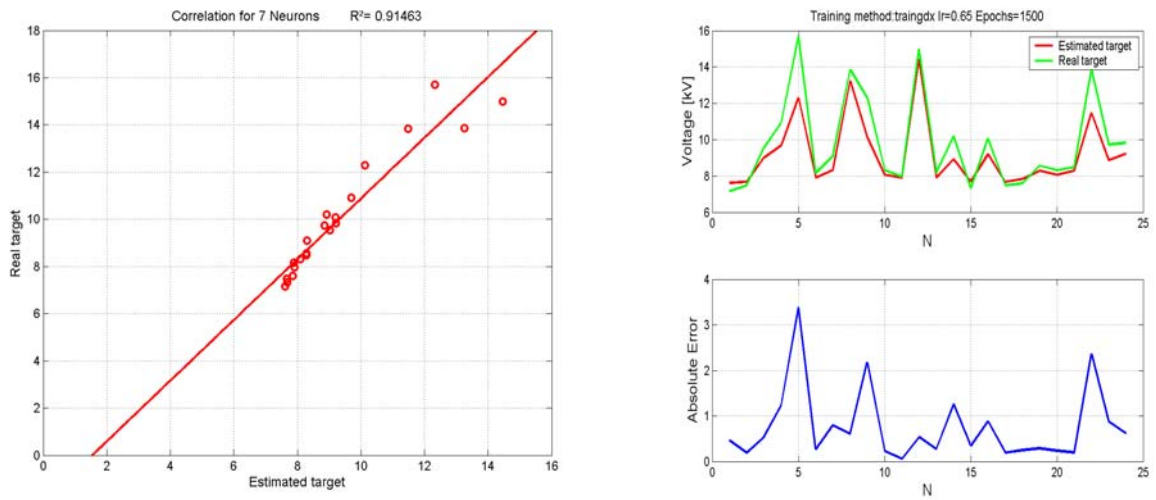
ΔΟΚΙΜΗ 12:

Αριθμός νευρώνων: μεταβολή από 6-12

Αριθμός εποχών: σταθερά 1500 εποχές

Ρυθμός μάθησης: 0,65

Ελάχιστο σφάλμα 1,2156 για 7 νευρώνες. Συσχέτιση για την περίπτωση αυτή: 0,91463



Σχήμα 7.56

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμούμενης τιμής για το σύνολο ελέγχου για 7 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμούμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

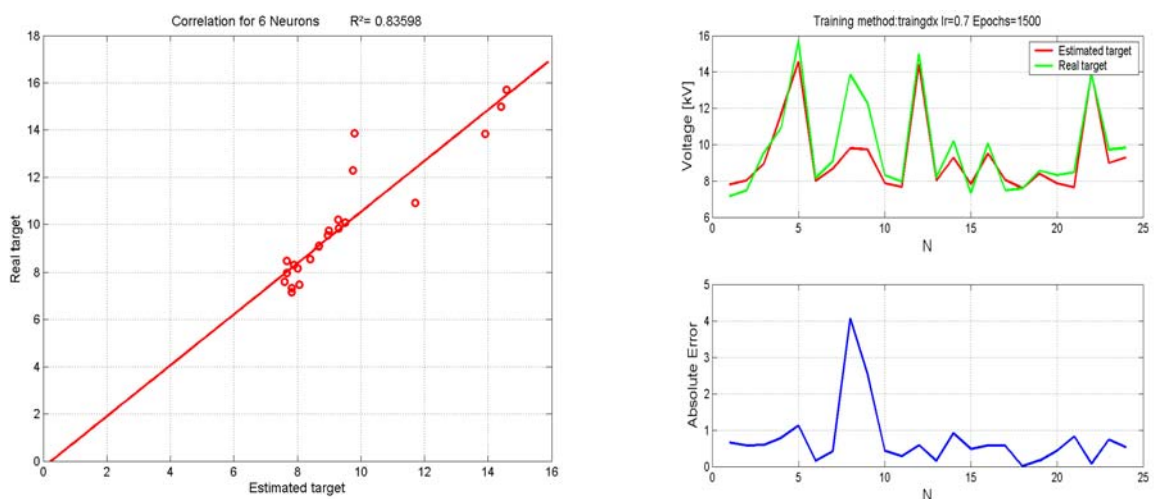
ΔΟΚΙΜΗ 13:

Αριθμός νευρώνων: μεταβολή από 6-12

Αριθμός εποχών: σταθερά 1500 εποχές

Ρυθμός μάθησης: 0,7

Ελάχιστο σφάλμα 1,2658 για 6 νευρώνες. Συσχέτιση για την περίπτωση αυτή: 0,83598



Σχήμα 7.57

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμούμενης τιμής για το σύνολο ελέγχου για 6 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμούμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

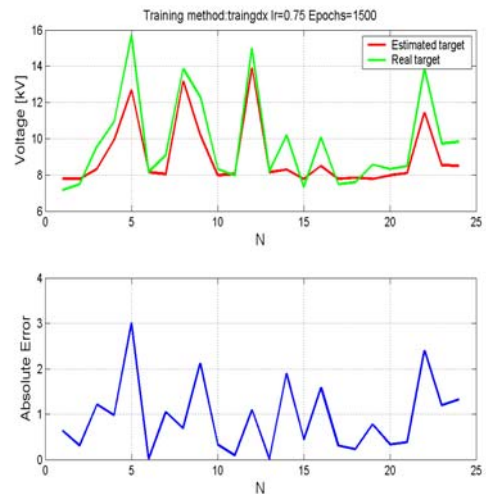
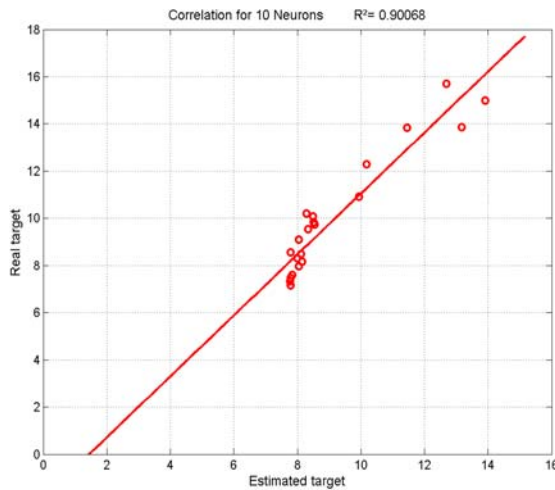
ΔΟΚΙΜΗ 14:

Αριθμός νευρώνων: μεταβολή από 6-12

Αριθμός εποχών: σταθερά 1500 εποχές

Ρυθμός μάθησης: 0,75

Ελάχιστο σφάλμα 1,4836 για 10 νευρώνες. Συσχέτιση για την περίπτωση αυτή: 0,90068



Σχήμα 7.58

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμούμενης τιμής για το σύνολο ελέγχου για 10 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμούμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

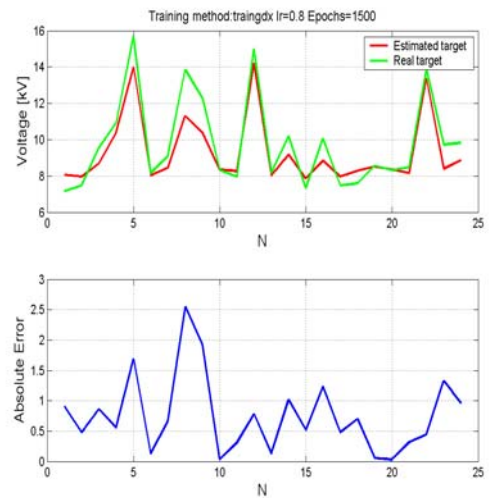
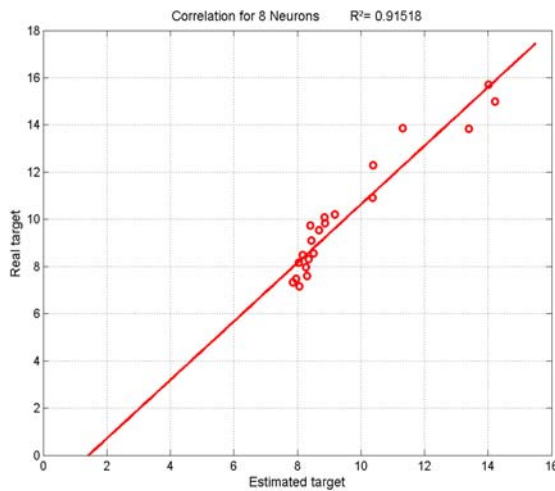
ΔΟΚΙΜΗ 15:

Αριθμός νευρώνων: μεταβολή από 6-12

Αριθμός εποχών: σταθερά 1500 εποχές

Ρυθμός μάθησης: 0,8

Ελάχιστο σφάλμα 0,9527 για 8 νευρώνες. Συσχέτιση για την περίπτωση αυτή: 0,91518



Σχήμα 7.59

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμούμενης τιμής για το σύνολο ελέγχου για 8 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμούμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

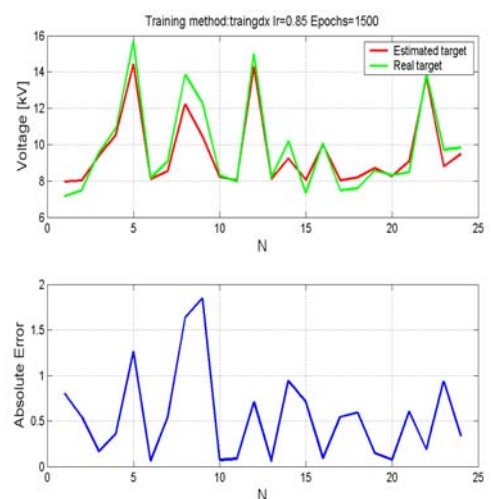
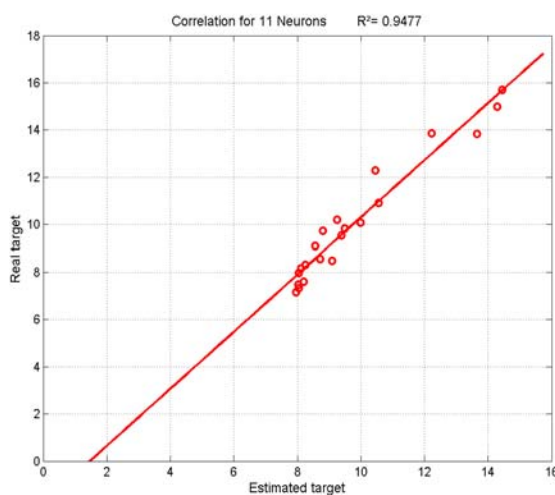
ΔΟΚΙΜΗ 16:

Αριθμός νευρώνων: μεταβολή από 6-12

Αριθμός εποχών: σταθερά 1500 εποχές

Ρυθμός μάθησης: 0,85

Ελάχιστο σφάλμα 0,5438 για 11 νευρώνες. Συσχέτιση για την περίπτωση αυτή: 0,9477



Σχήμα 7.60

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμούμενης τιμής για το σύνολο ελέγχου για 11 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμούμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

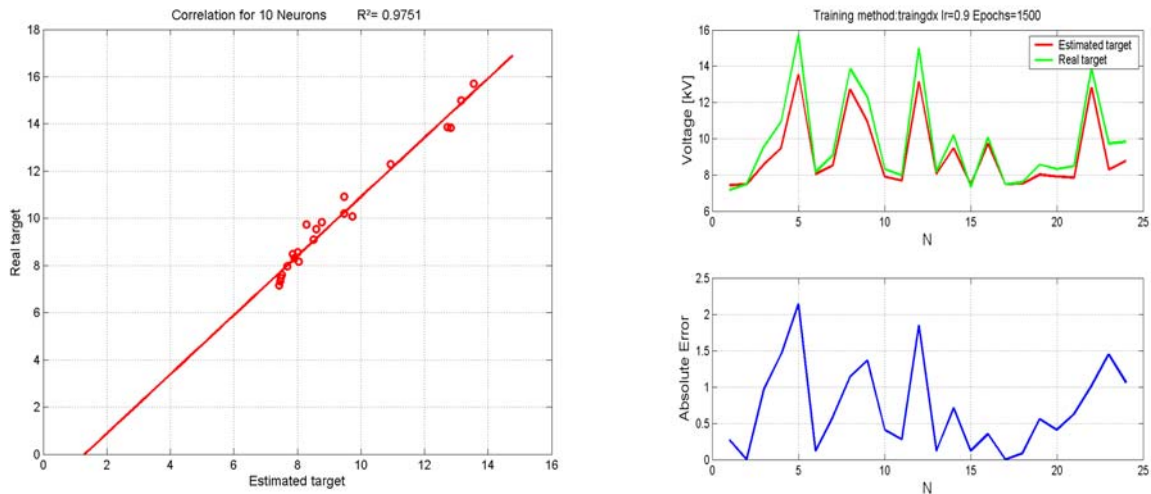
ΔΟΚΙΜΗ 17:

Αριθμός νευρώνων: μεταβολή από 6-12

Αριθμός εποχών: σταθερά 1500 εποχές

Ρυθμός μάθησης: 0,9

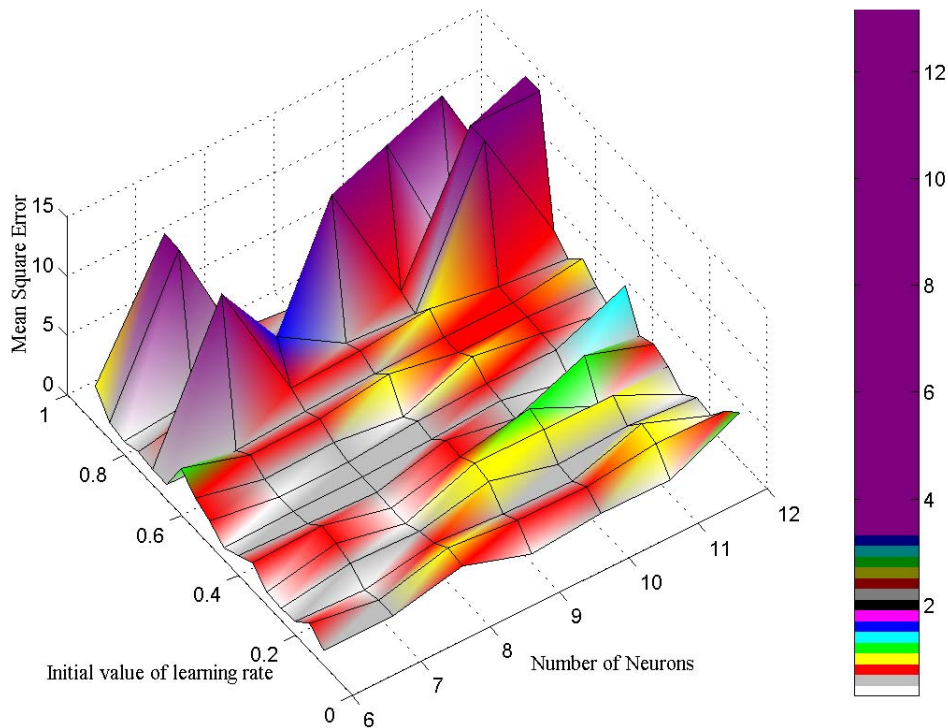
Ελάχιστο σφάλμα 0,8622 για 10 νευρώνες. Συσχέτιση για την περίπτωση αυτή: 0,9751



Σχήμα 7.61

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμούμενης τιμής για το σύνολο ελέγχου για 10 νευρώνες. Δεξιά: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμούμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

Η μεταβολή του μέσου τετραγωνικού σφάλματος συναρτήσει των νευρώνων και του ρυθμού μάθησης φαίνεται σε τρισδιάστατο γράφημα παρακάτω:



Σχήμα 7.62: Μεταβολή του μέσου τετραγωνικού σφάλματος συναρτήσει των νευρώνων και του ρυθμού μάθησης για το σύνολο ελέγχου

Παρατηρείται ότι για ρυθμό μάθησης άνω του 0,5 τα σφάλματα είναι πάρα πολύ μεγάλα, ενώ αντίθετα για ρυθμό μάθησης 0,3-0,4 και αριθμό νευρώνων περίπου 8 η μέθοδος `trainIdx` δίνει πολύ ικανοποιητικά αποτελέσματα

7.8 ΔΟΚΙΜΗ ΓΙΑ ΔΥΟ ΕΠΙΠΕΔΑ

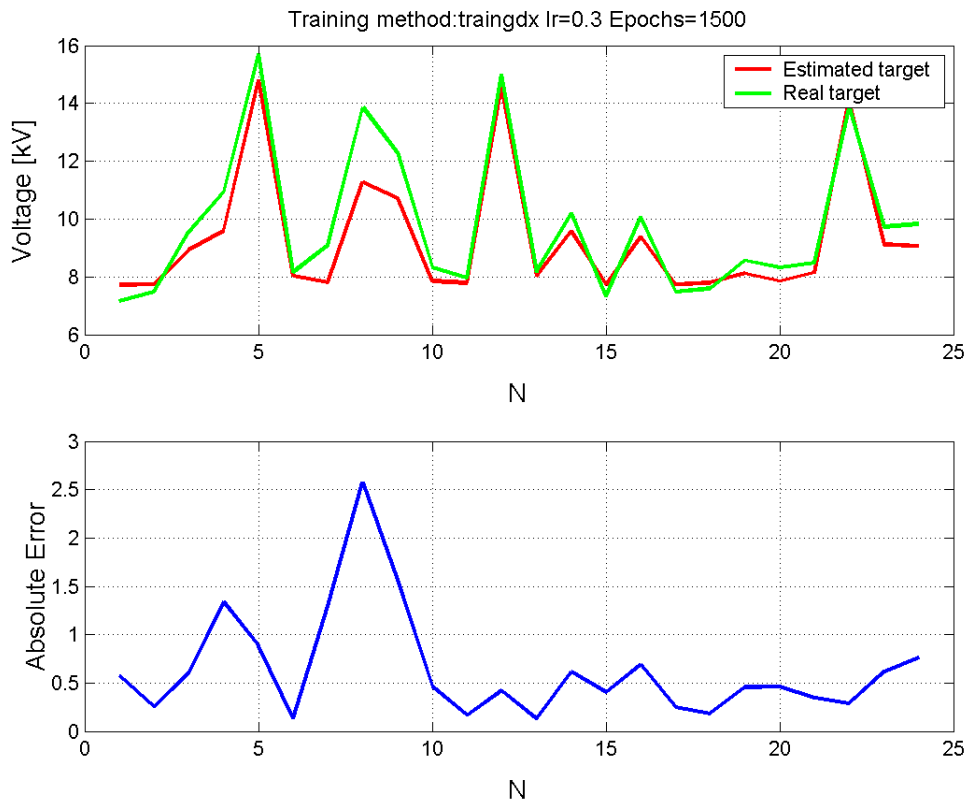
Αριθμός νευρώνων: μεταβολή από 6-12

Αριθμός εποχών: σταθερά 1500 εποχές

Ρυθμός μάθησης: 0,3

Ελάχιστο σφάλμα 0,6493 για αριθμό νευρώνων στο πρώτο επίπεδο: 12 και αριθμό νευρώνων στο δεύτερο επίπεδο: 6.

Στο Σχήμα 7.63 φαίνονται οι διαφορές στις τιμές μεταξύ των πραγματικών και εκτιμώμενων τιμών τάσης (σε kV) για το σύνολο ελέγχου, στην περίπτωση που το νευρωνικό δίκτυο έχει δύο κρυμμένα επίπεδα.



Σχήμα 7.63: Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμούμενες τιμές που δίνει το νευρωνικό με δύο κρυμμένα επίπεδα, στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

7.9 Παρατηρήσεις

Μια γενική παρατήρηση που μπορεί να γίνει είναι ότι σε κάποιες από τις δοκιμές (π.χ. δοκιμές 4, 5, 8, 10, 13) η συσχέτιση μεταξύ εκτιμούμενων και πραγματικών τιμών εκτός του ότι παρουσιάζει καλή τιμή, περνάει και απ' την αρχή των αξόνων. Αυτό είναι μια πολύ καλή ένδειξη ότι το δίκτυο δίνει καλά αποτελέσματα. Στις υπόλοιπες περιπτώσεις παρατηρείται μετατόπιση της ευθείας της συσχέτισης προς τα δεξιά (παρότι οι τιμές της είναι καλές), γεγονός που υποδηλώνει ότι το δίκτυο κάνει ένα συστηματικό σφάλμα.

Τέλος θα πρέπει να σημειωθεί ότι το δίκτυο μπορεί να λειτουργήσει και με δύο επίπεδα, γεγονός όμως που δεν συνηθίζεται, αφού με βάση το θεώρημα του Kolmogorov η ίδια λειτουργία μπορεί να επιτευχθεί και με ένα επίπεδο. Πάντως, αυτό είναι ένα θέμα που θα μπορούσε να εξετασθεί στο μέλλον, για να φανεί αν το δίκτυο με δύο κρυμμένα επίπεδα μπορεί να δώσει και καλύτερα αποτελέσματα.

Κεφάλαιο 8

Πρόγραμμα σε Fortran

8.1 Εισαγωγή

Στο κεφάλαιο αυτό γίνεται η παρουσίαση των αποτελεσμάτων που έδωσε για το ίδιο πρόβλημα ένα προσαρμοστικό νευρωνικό δίκτυο (adaptive neural network) που βρίσκει αυτόματα το βέλτιστο συνδυασμό πλήθους νευρώνων, παραμέτρων ρυθμού μάθησης και όρου ορμής, ώστε να επιλεγεί το καταλληλότερο μοντέλο [17].

Εφόσον επιλεγούν το πλήθος και η μορφή των μεταβλητών εισόδου και εξόδου, εκπαιδεύεται το τεχνητό νευρωνικό δίκτυο μέσω της μεθόδου της ανάστροφης διάδοσης σφάλματος, τα στάδια της οποίας είναι:

1. Αρχικοποίηση: Ορίζεται η αρχιτεκτονική του δικτύου -ουσιαστικά δίνονται συγκεκριμένες τιμές του πλήθους νευρώνων, εκκίνησης και χρονικής παραμέτρου όρου ορμής, εκκίνησης και χρονικής παραμέτρου ρυθμού μάθησης, το είδος της συνάρτησης ενεργοποίησης ανά επίπεδο και των αντίστοιχων χαρακτηριστικών τους τιμών και αρχικοποιούνται τα βάρη των συνδέσεων σε κατάλληλες τυχαίες μικρές τιμές.

2. Παρουσίαση των προτύπων εκπαίδευσης: Παρουσιάζονται στο νευρωνικό δίκτυο μία εποχή προτύπων εκπαίδευσης και για κάθε πρότυπο επαναλαμβάνονται τα βήματα 3 & 4, αφού ρυθμιστούν οι παράμετροι του ρυθμού μάθησης και του όρου ορμής. Τονίζεται ότι η σειρά παρουσίασης των προτύπων σε κάθε εποχή ep είναι τυχαία.

3. Υπολογισμοί ευθείας φοράς: Αν θεωρηθεί το n -οστό πρότυπο εκπαίδευσης $\{\vec{x}(n), \vec{t}(n)\}$ με $\vec{x}(n)$ το διάνυσμα εισόδου που εφαρμόζεται στο δίκτυο και $\vec{t}(n)$ το επιθυμητό διάνυσμα εξόδου του δικτύου, τότε υπολογίζονται τα σήματα που προκύπτουν σε κάθε επίπεδο. Το σήμα ενεργοποίησης του νευρώνα k του l επιπέδου -είτε πρόκειται για το επίπεδο εισόδου, είτε για το κρυμμένο- είναι :

$$u_k^{(l)}(n) = \sum_{v=0}^p w_{kv}^{(l)}(n) y_v^{(l-1)}(n) \quad (8.1)$$

όπου $w_{kv}^{(l)}$ το βάρος που συνδέει τον νευρώνα k του l επιπέδου με το νευρώνα v του $(l-1)$ επιπέδου, $y_v^{(l-1)}$ το σήμα εξόδου του νευρώνα v του προηγούμενου επιπέδου και p ο συνολικός αριθμός των νευρώνων του $(l-1)$ επιπέδου. Για $v = 0$ προκύπτει $y_0 = -1$ και $w_{k0} = \theta_k$, όπου θ_k η τιμή πόλωσης. Οι συναρτήσεις μεταξύ των δύο επιπέδων μπορεί να είναι είτε η τροποποιημένη υπερβολική εφαπτομένη, είτε η λογιστική, οπότε το σήμα εξόδου του παραπάνω νευρώνα είναι το εξής :

$$y_k^{(l)}(n) = f(u_k^{(l)}(n)) \quad (8.2)$$

όπου

$$f(x) = \tanh(h_1 \cdot x + h_2) \text{ -υπερβολική εφαπτομένη-} \quad (8.3)$$

$$\text{ή } f(x) = 1 / (1 + \exp(-(h_1 \cdot x + h_2))) \text{ -σιγμοειδής-}$$

με h_1, h_2 νέες παραμέτρους προς προσδιορισμό, όπως επίσης και το είδος των συναρτήσεων.

Αν ο νευρώνας k ανήκει στο πρώτο κρυμμένο επίπεδο, δηλαδή $l = 1$, τότε τίθεται:

$$y_v^{(0)}(n) = x_v(n), \quad \forall v \quad (8.4)$$

όπου $x_v(n)$ το v -οστό στοιχείο του διανύσματος εισόδου $\bar{x}(n)$.

Αν όμως ο νευρώνας k ανήκει στο επίπεδο εξόδου $-L = l = 2$ για την παρούσα περίπτωση- τότε ισχύει:

$$y_k^{(L)}(n) = o_k(n), \quad \forall k \quad (8.5)$$

όπου $o_k(n)$ το k -οστό στοιχείο του πραγματικού διανύσματος εξόδου $\bar{o}(n)$. Οπότε το σφάλμα εξόδου του νευρώνα k είναι ίσο με:

$$e_k(n) = t_k(n) - o_k(n) \quad (8.6)$$

4. Υπολογισμοί αντίστροφης φοράς: Υπολογίζεται η τοπική κλίση του κάθε νευρώνα θεωρώντας τις εξής περιπτώσεις:

- ◆ Αν ο νευρώνας k είναι κόμβος εξόδου ($l=L=2$), τότε ισχύει:

$$\delta_k^{(L)}(n) = e_k^{(L)}(n) \cdot f'(u_k^{(L)}(n)) \quad (8.7)$$

- ◆ Αν ο νευρώνας k ανήκει στο κρυμμένο επίπεδο l - που εδώ μπορεί να είναι μόνο 1-, τότε ισχύει :

$$\delta_k^{(l)}(n) = f' \left(u_k^{(l)}(n) \right) \cdot \sum_i \delta_i^{(l+1)}(n) \cdot w_{ik}^{(l+1)}(n) \quad (8.8)$$

Στη συνέχεια προσαρμόζονται τα βάρη του επιπέδου l σύμφωνα με τον γνωστό κανόνα δ :

$$w_{kv}^{(l)}(n+1) = w_{kv}^{(l)}(n) + \alpha \cdot [w_{kv}^{(l)}(n) - w_{kv}^{(l-1)}(n-1)] + \eta \cdot \delta_k^{(l)}(n) \cdot y_v^{(l-1)}(n) \quad (8.9)$$

όπου η ο ρυθμός μάθησης και α ο όρος ορμής, όπως προκύπτουν από τις σχέσεις (8.13) και (8.14).

5. Εύρεση Σφαλμάτων: Σ' αυτό το στάδιο πραγματοποιούνται οι υπολογισμοί σφαλμάτων για την τρέχουσα εποχή ep , όπως της μέσης τιμής αθροίσματος τετραγωνικών σφαλμάτων των νευρώνων εξόδου G_{av-va} για το σύνολο αξιολόγησης έχοντας q_{out} το πλήθος νευρώνων εξόδου και N_{va} το πλήθος των διανυσμάτων αξιολόγησης:

$$G_{av-va} = \frac{1}{N_{va}} \sum_{m=1}^{N_{tr}} \left(\frac{1}{2} \sum_{k=1}^{q_{out}} e_k^2(m) \right) \quad (8.10)$$

Εναλλακτικά μπορεί να χρησιμοποιηθεί η μέση τετραγωνική ρίζα σφάλματος εξόδου για το σύνολο αξιολόγησης, η οποία δίνεται από τη σχέση:

$$MSE_{va} = \sqrt{\frac{1}{N_{va} \cdot q_{out}} \sum_{m=1}^{N_{va}} \sum_{k=1}^{q_{out}} e_k^2(m)} = \sqrt{\frac{2}{q_{out}} G_{av-va}} \quad (8.11)$$

Άλλα σφάλματα που μπορούν να χρησιμοποιηθούν είναι η μέση τιμή αθροίσματος σφαλμάτων των νευρώνων εξόδου ή η μέση τετραγωνική ρίζα σφάλματος για τα σύνολα εκπαίδευσης και τελικής πρόβλεψης, το μέγιστο κατ' απόλυτη τιμή σφάλμα, το μέσο απόλυτο εκατοστιαίο σφάλμα (MAPE).

για τα σύνολα εκπαίδευσης, ελέγχου και τελικής πρόβλεψης.

6. Επανάληψη: Επαναλαμβάνονται οι παραπάνω υπολογισμοί για νέες εποχές προτύπων, ώσπου οι ελεύθεροι παράμετροι του δικτύου, δηλαδή τα βάρη, να σταθεροποιηθούν και να μην ξεπεραστεί ο μέγιστος αριθμός εποχών. Συγκεκριμένα για να τερματιστεί η εκπαίδευση, πρέπει να ισχύουν ένα από τα εξής κριτήρια: το πλήθος των εποχών να είναι μεγαλύτερο από ένα μέγιστο αριθμό εποχών \max_epochs , όλες οι μεταβολές των βαρών και των πολώσεων μεταξύ δύο διαδοχικών εποχών να είναι μικρότερες από μία μικρή οριακή τιμή $limit_1$, δηλαδή:

$$ep \geq \max_epochs \quad \text{ή} \quad \left| w_{kv}^{(l)}(ep) - w_{kv}^{(l)}(ep-1) \right| < limit_1, \forall k, v, l \quad (8.12)$$

Αν δεν ισχύουν τα κριτήρια, κατά την επιστροφή στο 2^ο βήμα γίνεται νέος υπολογισμός του ρυθμού μάθησης και του όρου ορμής:

$$\eta(ep) = \begin{cases} \eta(ep-1), & G_{av-tr}(ep) > G_{av-tr}(ep-1) \\ \eta(ep-1) \cdot \exp(-1/T_\eta), & G_{av-tr}(ep) \leq G_{av-tr}(ep-1) \end{cases} \quad (8.13)$$

$$a(ep) = \begin{cases} a(ep-1), & G_{av-tr}(ep) \leq G_{av-tr}(ep-1) \\ a(ep-1) \cdot \exp(-1/T_a), & G_{av-tr}(ep) > G_{av-tr}(ep-1) \end{cases} \quad (8.14)$$

όπου T_η και $\eta_0 = \eta(0)$ είναι η χρονική παράμετρος και η αρχική τιμή του ρυθμού μάθησης αντίστοιχα, T_a και $a_0 = a(0)$ είναι η χρονική παράμετρος και η αρχική τιμή του όρου ορμής. Ουσιαστικά το τεχνητό νευρωνικό δίκτυο προσαρμόζει τις παραμέτρους του ανάλογα με την εξέλιξη του σφάλματος, γι' αυτό ανήκει στα προσαρμοστικά νευρωνικά δίκτυα. Σημειώνεται ότι κατά τη μεταβολή του όρου ορμής δεν χρησιμοποιήθηκε στην περίπτωση βελτίωσης της μέσης τιμής αθροίσματος τετραγωνικών σφαλμάτων των νευρώνων εξόδου ως νέα τιμή το $a(n-1) \cdot \exp(1/T_a)$, αλλά προτιμήθηκε η σταθεροποίηση της τιμής του όρου ορμής για λόγους ευστάθειας κατά τη σύγκλιση του αλγορίθμου εκπαίδευσης.

Αν δεν ισχύουν τα κριτήρια μάθησης (8.12), τότε γίνεται η μετάβαση στο επόμενο βήμα.

7. Υπολογισμός τελικών σφαλμάτων. Υπολογίζονται τα αντίστοιχα σφάλματα όπως το μέσο τετραγωνικό σφάλμα ή το MAPE για το τελικό μέγεθος εκτίμησης με έμφαση στο μέσο απόλυτο εκατοστιαίο σφάλμα.

Σημειώνεται ότι, προκειμένου να αποφευχθεί ο κορεσμός των συναρτήσεων ενεργοποίησης, πραγματοποιείται η αντίστοιχη κανονικοποίηση των μεταβλητών εισόδου και εξόδου μέσω της σχέσης:

$$\hat{z} = c + \frac{b-c}{r_{\max} - r_{\min}} \cdot (z - r_{\min}) \quad (8.15)$$

όπου r_{\max} και r_{\min} τα άνω και κάτω όρια του υπό κανονικοποίηση μεγέθους z , b και c τα αντίστοιχα όρια του κανονικοποιημένου μεγέθους \hat{z} .

Η παραπάνω μεθοδολογία εφαρμόζεται, εφόσον είναι δεδομένα τα στοιχεία διαμόρφωσης του νευρωνικού δικτύου, όπως το πλήθος των νευρώνων, οι αρχικές τιμές και οι χρονικές παράμετροι του ρυθμού εκπαίδευσης και του όρου ορμής - T_η , η_0 , T_a και

α_0 -, το είδος των συναρτήσεων ενεργοποίησης και οι αντίστοιχες παράμετροι τους. Για την εύρεση των καταλληλότερων τιμών των αντίστοιχων παραμέτρων μπορούν να γίνουν οι αντίστοιχοι συνδυασμοί. Αν για την i -οστή παράμετρο υπό βελτιστοποίηση υπάρχουν s_i τιμές, τότε οι αντίστοιχοι συνδυασμοί είναι ίσοι με $s_i = \prod s_i$.

Προκειμένου να μειωθούν οι αντίστοιχοι συνδυασμοί για τον προσδιορισμό των βέλτιστων τιμών των παραμέτρων του τεχνητού νευρωνικού δικτύου, πραγματοποιούνται δύο στάδια. Κατά το πρώτο ο βασικός αλγόριθμος εκτελείται ξεχωριστά για την περιοχή τιμών της κάθε μεταβλητής υπό βελτιστοποίηση και καταγράφονται εκείνες οι τιμές που δίνουν ικανοποιητικά αποτελέσματα. Στο δεύτερο στάδιο οι παράμετροι μεταβάλλουν τις τιμές τους συγχρόνως εντός των περιοχών, που προέκυψαν κατά το πρώτο βήμα, και η κύρια διαδικασία επαναλαμβάνεται για όλους τους δυνατούς συνδυασμούς. Με αυτόν τον τρόπο γίνεται η βελτιστοποίηση των παραμέτρων του ΤΝΔ.

Συνοψίζοντας τα βασικά βήματα της προτεινόμενης μεθοδολογίας βελτιστοποίησης είναι:

α. Ορισμός των συνδυασμών των n μεταβλητών εισόδου, για καθένα από τους οποίους πραγματοποιούνται τα εξής:

- Για κάθε μεταβλητή του νευρωνικού δικτύου εφαρμόζεται επαναληπτικά ο προσαρμοστικός αλγόριθμος ανάστροφης τροφοδότησης ξεχωριστά για κατάλληλο εύρος τιμών και προσδιορίζονται οι περιοχές με τα ικανοποιητικά αποτελέσματα.
- Ο ίδιος αλγόριθμος επαναλαμβάνεται, ενώ όλες οι μεταβλητές μεταβάλλουν ταυτόχρονα τις τιμές τους, οι οποίες κυμαίνονται στις αντίστοιχες περιοχές της προηγούμενης φάσης, έτσι, ώστε ο συνδυασμός που παράγει το μικρότερο σφάλμα για το σύνολο αξιολόγησης να επιλεγεί.

β. Τελικά εκτιμάται το επιθυμητό μέγεθος, που σε αυτήν την περίπτωση είναι η κρίσιμη τάση για τους μονωτήρες.

Τα βασικά μειονεκτήματα της μεθοδολογίας με τη χρήση των τεχνητών νευρωνικών δικτύων είναι η μη εύρεση τυπικής απόκλισης και η μη αναγνώσιμη μορφή του μοντέλου που προέκυψε από το χρήστη. Δηλαδή από τη μία μεριά πρέπει να γίνει μία στοιχειώδης ανάλυση ευαισθησίας του μοντέλου, για να μελετηθεί η ανταπόκριση του σε μη πλήρως

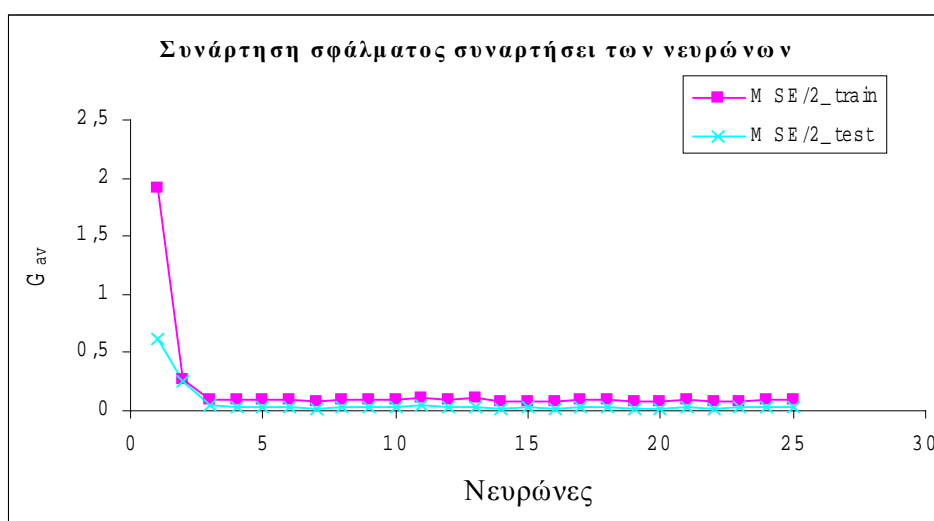
γνωστά δεδομένα. Από την άλλη δεν είναι εμφανής λόγω της δομής του νευρωνικού δικτύου πόσο συμμετέχει η κάθε μεταβλητή εισόδου.

Επίσης στην προκειμένη περίπτωση συναντιέται μία επιπρόσθετη δυσκολία κατά την εκπαίδευση, καθώς το πλήθος των διαθέσιμων δεδομένων διανυσμάτων δεν είναι μεγάλο, γεγονός που οδηγεί στην ταύτιση των συνόλων εκπαίδευσης και αξιολόγησης του ΤΝΔ και πρακτικά στην αύξηση του κινδύνου της υπερεκπαίδευσης. Το τελευταίο αντιμετωπίζεται εν μέρει από την επιλογή κατάλληλου μέγιστου αριθμού επαναλήψεων, που πρέπει να μην είναι ούτε μικρός μην επιτρέποντας τη σύγκλιση, ούτε μεγάλος οδηγώντας σε αποστήθιση το ΤΝΔ [17].

8.2 Δοκιμές

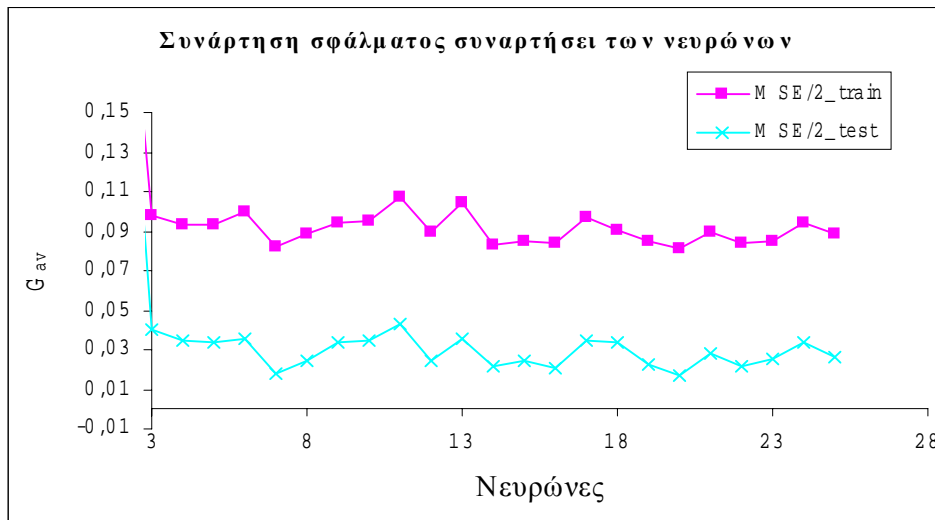
Σημειώνουμε αρχικά ότι για το επίπεδο εισόδου και το κρυμμένο επίπεδο ως συνάρτηση ενεργοποίησης των νευρώνων επελέγη η υπερβολική εφαπτομένη.

Προκειμένου να προσδιοριστεί ποιος αριθμός νευρώνων δίνει τα καλύτερα αποτελέσματα έγινε το εξής: Για σταθερές τιμές στις παραμέτρους του όρου ορμής a και του ρυθμού μάθησης η και με σταθερό αριθμό εποχών (5000) μεταβάλλονται οι νευρώνες από 1 έως 25 με βήμα 1 και παρατηρείται το σφάλμα. Θα πρέπει να αναφερθεί ότι σαν δείκτης για τη βελτιστοποίηση δεν χρησιμοποιείται το μέσο τετραγωνικό σφάλμα απλώς, όπως στα προηγούμενα κεφάλαια, αλλά η συνάρτηση σφάλματος G_{av} [19]. Τα αποτελέσματα φαίνονται στην παρακάτω γραφική παράσταση:



Σχήμα 8. 1: Δοκιμή για την εύρεση του κατάλληλου πλήθους νευρώνων

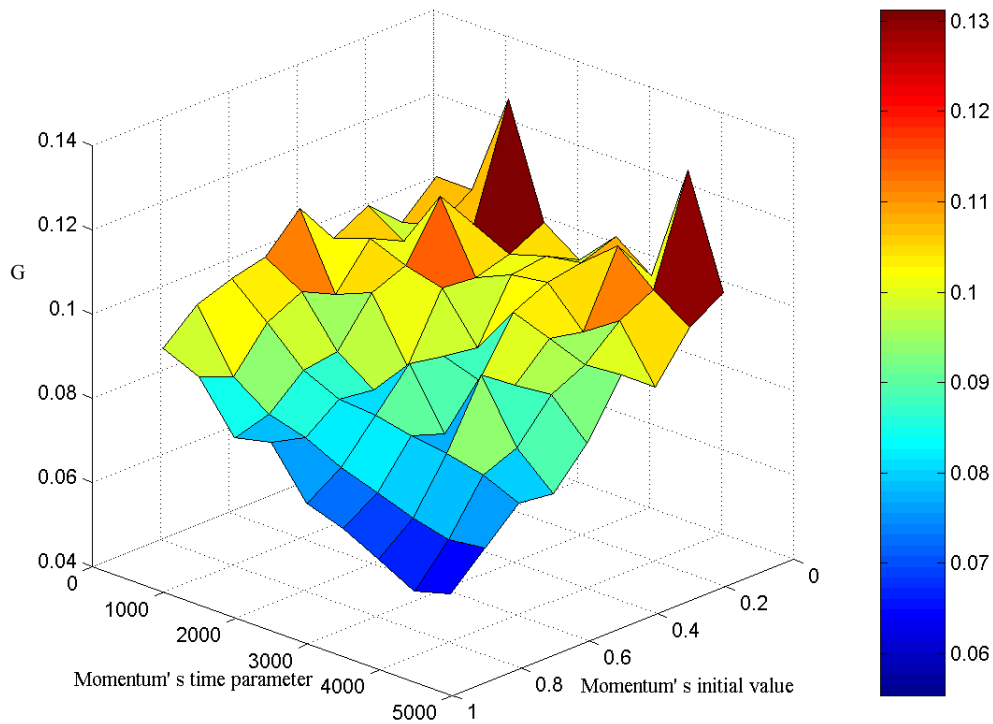
Για μεγαλύτερη ευκρίνεια το παραπάνω σχήμα τροποποιείται ως εξής:



Σχήμα 8. 2: Μετατοπισμένοι άξονες για μεγαλύτερη ευκρίνεια

Όπως φαίνεται, το μικρότερο σφάλμα εντοπίζεται για 7 νευρώνες.

Στη συνέχεια διατηρείται σταθερός ο ρυθμός μάθησης, ο αριθμός νευρώνων και ο αριθμός εποχών, ενώ μεταβάλλεται ο όρος ορμής (ο σταθερός όρος από 0,1-0,9 και η χρονική παράμετρος από 500-4500), προκειμένου να βρεθεί για ποια τιμή του όρου ορμής εμφανίζονται καλά αποτελέσματα. Το τρισδιάστατο γράφημα για τη δοκιμή αυτή φαίνεται παρακάτω:



Σχήμα 8.3 Μεταβολή της συνάρτησης σφάλματος G συναρτήσει του όρου ορμής (σταθερού όρου $\eta_0=\{0.1, 0.2, \dots, 0.9\}$ και χρονικής παραμέτρου $T_0=\{500, 1000, \dots, 4500\}$) για σταθερό ρυθμό μάθησης ($\alpha_0=0.5$, $T_{\alpha_0}=3000$), αριθμό νευρώνων και αριθμό εποχών (5000).

Παρατηρήθηκε λοιπόν ότι το ελάχιστο μέσο τετραγωνικό σφάλμα $G_{av-training}$ για το σύνολο εκπαίδευσης προέκυψε για σταθερό όρο ορμής ίσο με 0,9 και για χρονική παράμετρο ορμής ίση με 4000.

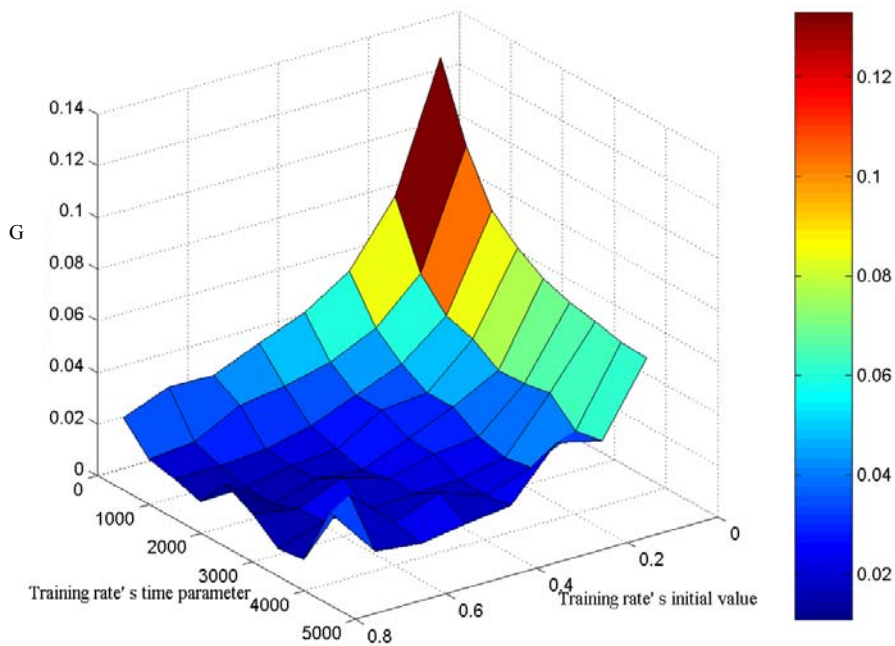
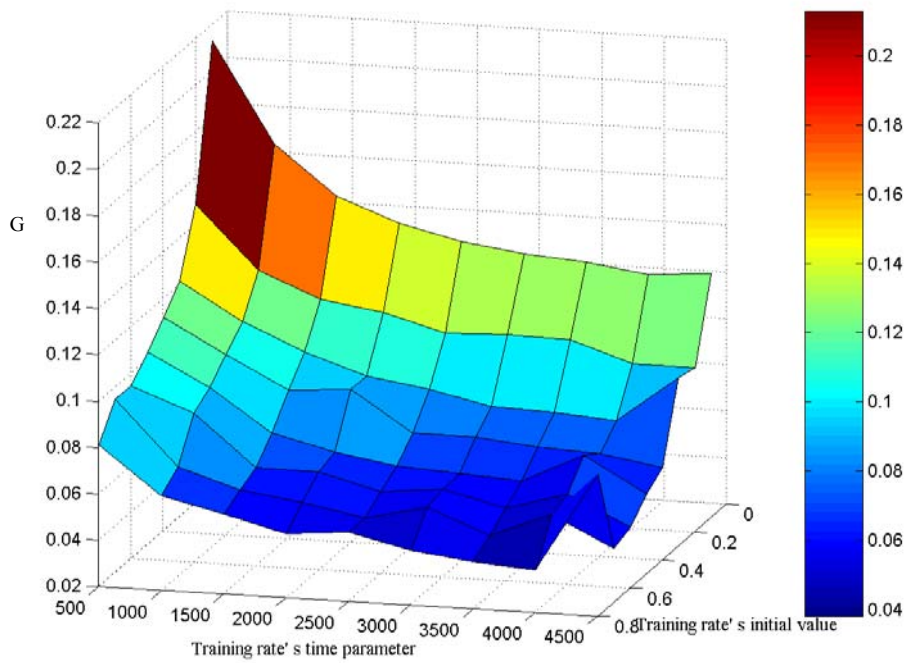
Το επόμενο στάδιο ήταν να καθοριστεί για ποια τιμή του ρυθμού μάθησης λαμβάνεται το καλύτερο $G_{av-training}$. Για το λόγο αυτό κρατήθηκαν σταθερά τα εξής

Σταθερός όρος ορμής: $\alpha_0=0,9$

Χρονική παράμετρος ορμής: $T_{\alpha_0}=4000$

Αριθμός νευρώνων: 7

Και μεταβλήθηκε ο σταθερός όρος του ρυθμού μάθησης (από 0,1-0,9) και η χρονική παράμετρος του ρυθμού μάθησης (από 500-4500).



Σχήμα 8.4 Μεταβολή της συνάρτησης σφάλματος G συναρτήσει του ρυθμού μάθησης (σταθερού όρου και χρονικής παραμέτρου) $n=7$, $\eta_0=0.9$, $T_0=4000$

Πάνω: Για το σύνολο εκπαίδευσης, Κάτω: Για το σύνολο ελέγχου

Παρατηρείται ότι δεν είναι απαραίτητο το μικρότερο σφάλμα να παρουσιάζεται στα ίδια σημεία και για το σύνολο εκπαίδευσης και για το σύνολο ελέγχου, αλλά γενικά τα δύο γραφήματα έχουν την παρόμοια μορφή.

Στη συνέχεια επαναλήφθηκε η διαδικασία για 5-9 νευρώνες, $\alpha_0=0,8-0,9$, $T_{\alpha_0}=3500-4500$ με βήμα 500, $\eta_0=0,7-0,9$ και $T_{\eta_0}=3500-4500$ με βήμα 500 και καταλήξαμε στα εξής τελικά αποτελέσματα:

Το μικρότερο σφάλμα για το σύνολο εκπαίδευσης επετεύχθη για σταθερό όρου ρυθμού μάθησης ίσο με 0,8 και χρονική παράμετρο ρυθμού μάθησης ίση με 4000. Γι' αυτές τις τιμές παραμέτρων παρατηρείται μικρό σφάλμα και στο σύνολο ελέγχου.

Εφόσον σταθεροποιήθηκαν οι παραπάνω παράμετροι μπορεί να γίνει εκτέλεση του προγράμματος για τα εξής:

Αριθμός νευρώνων: 7

Σταθερός όρος ορμής α_0 : 0,9

Χρονική παράμετρος ορμής T_{α_0} : 4000

Σταθερός όρος ρυθμού μάθησης η_0 : 0,8

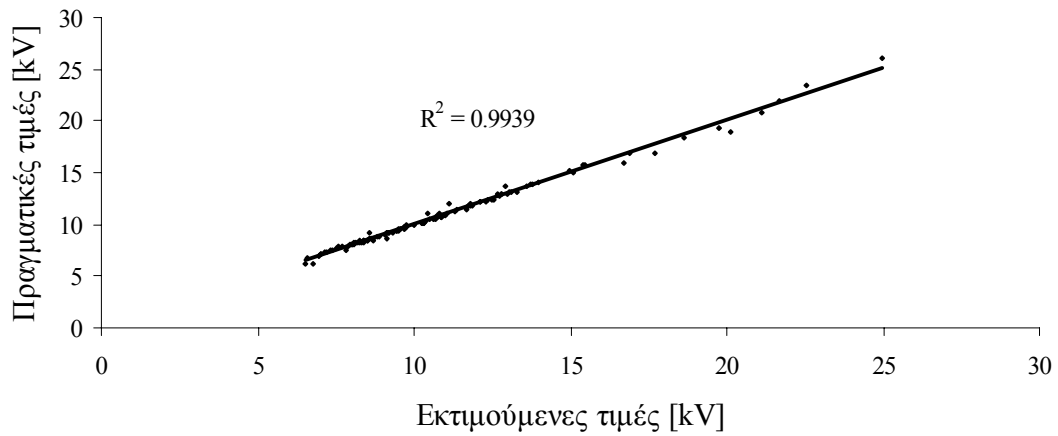
Χρονική παράμετρος ρυθμού μάθησης T_{η_0} : 4000.

Τα σφάλματα εκπαίδευσης και ελέγχου προέκυψαν αντίστοιχα:

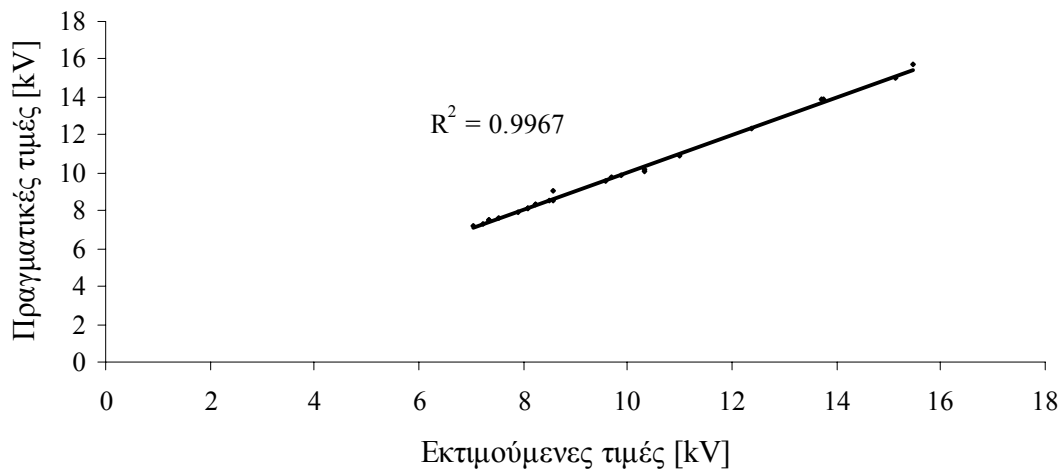
$G_{\text{av-training}}=0,03820679$ και $G_{\text{av-test}}=0,01212812$.

Μετά την εκτέλεση του προγράμματος, βρέθηκε η συσχέτιση μεταξύ πραγματικών και εκτιμώμενων τιμών του ΤΝΔ. Στη συνέχεια φαίνονται τα γραφήματα για τη συσχέτιση μεταξύ εκτιμώμενων και εκτιμώμενων και πραγματικών, τόσο για το σύνολο εκπαίδευσης, όσο και για το σύνολο ελέγχου.

Συσχέτιση μεταξύ πραγματικών και εκτιμώμενων τιμών για την κρίσιμη τάση για το σύνολο εκπαίδευσης



Συσχέτιση μεταξύ πραγματικών και εκτιμώμενων τιμών για την κρίσιμη τάση για το σύνολο ελέγχου



Κεφάλαιο 9

Συμπεράσματα – Η επόμενη μέρα

Η παρούσα διπλωματική πραγματεύεται την εφαρμογή της μεθοδολογίας των Τεχνητών Νευρωνικών Δικτύων (ΤΝΔ) για την εκτίμηση της κρίσιμης τάσης σε μονωτήρες Υψηλών Τάσεων. Για το σκοπό αυτό δοκιμάστηκαν δύο διαφορετικά νευρωνικά δίκτυα, τα οποία ως μεταβλητές εισόδου χρησιμοποιούσαν τα εξής χαρακτηριστικά του μονωτήρα: μέγιστη διάμετρο D_m (σε cm) του δίσκου του μονωτήρα, ύψος H (σε cm), μήκος ερπυσμού L (σε cm), συντελεστή μορφής F και αγωγιμότητα στρώματος σ_s (σε μS), ενώ ως έξοδο έδιναν την κρίσιμη τάση σε kV.

Το πρώτο απ' τα δύο νευρωνικά δίκτυα κατασκευάστηκε με χρήση έτοιμων εργαλείων της πλατφόρμας του MATLAB για τα νευρωνικά. Ο όρος 'έτοιμα εργαλεία' περιλαμβάνει τις συναρτήσεις εκπαίδευσης, ενεργοποίησης, αρχικοποίησης κ.ά., οι οποίες σε αυτήν την περίπτωση δεν υλοποιήθηκαν εξ αρχής, αλλά χρησιμοποιήθηκαν ως είχαν, χωρίς να μεταβληθεί ο κώδικας υλοποίησής τους. Ο δεύτερος τρόπος ήταν με χρήση αυτόνομου προγράμματος σε digital Fortran, το οποίο κατασκευάζει ένα προσαρμοστικό ΤΝΔ, παρέχοντας ταυτόχρονα τη δυνατότητα αυτόματης εύρεσης του βέλτιστου συνδυασμού νευρώνων, ρυθμού μάθησης και όρου ορμής.

Και στις δύο περιπτώσεις έγιναν δοκιμές μεταβάλλοντας τον αριθμό εποχών, το πλήθος των νευρώνων και το ρυθμό μάθησης, προκειμένου να βρεθεί η βέλτιστη αρχιτεκτονική του δικτύου, δηλαδή ο συνδυασμός που δίνει το ελάχιστο σφάλμα. Επιπλέον, στην πρώτη περίπτωση δοκιμάστηκαν και διάφορες μέθοδοι εκπαίδευσης που χρησιμοποιούν τη μέθοδο ανάστροφης διάδοσης σφάλματος, για να βρεθεί ποια έχει την καλύτερη σύγκλιση. Ως κριτήριο βελτιστοποίησης (σφάλμα) στην πρώτη περίπτωση χρησιμοποιήθηκε το μέσο τετραγωνικό σφάλμα (συνάρτηση MSE της MATLAB με

αναλυτική έκφραση $MSE = \frac{1}{N_{va}} \left(\sum_{m=1}^{N_{va}} \sqrt{\sum_{k=1}^{q_{out}} e_k^2(m)} \right)$), ενώ στη δεύτερη περίπτωση

χρησιμοποιήθηκε η συνάρτηση σφάλματος $G_{av-va} = \frac{1}{N_{va}} \sum_{m=1}^{N_m} \left(\frac{1}{2} \sum_{k=1}^{q_{out}} e_k^2(m) \right)$ [19]. Το πλήθος

των νευρώνων εξόδου q_{out} είναι στην προκειμένη περίπτωση 1.

Τα αποτελέσματα και των δύο εναλλακτικών κατασκευών του ΤΝΔ συμφωνούν ως προς τη μέθοδο εκπαίδευσης και την αρχιτεκτονική του δικτύου, με τη διαφορά ότι το δίκτυο που κατασκευάστηκε στο MATLAB δίνει γενικά μεγαλύτερα σφάλματα απ' αυτό που κατασκευάστηκε σε digital Fortran. Η βελτιστοποίηση έγινε και στις δύο περιπτώσεις με βάση το σύνολο εκπαίδευσης.

Συγκεκριμένα, οι δοκιμές που έγιναν είναι οι εξής:

Για το πρώτο δίκτυο δοκιμάστηκαν οι εξής μέθοδοι εκπαίδευσης: εκπαίδευση με τον κανόνα απότομης κλίσης (traingd), εκπαίδευση με κανόνα απότομης κλίσης και ταυτόχρονη προσαρμογή του ρυθμού μάθησης (traingda), εκπαίδευση με κανόνα απότομης κλίσης, ταυτόχρονη προσαρμογή του ρυθμού μάθησης και προσθήκη όρου ορμής (traingdx) και εκπαίδευση με τον κανόνα Levenberg-Marquardt (trainlm). Για κάθε μία απ' τις παραπάνω μεθόδους πάρθηκε ένα σύνολο σεναρίων στα οποία η παράμετρος που μεταβαλλόταν ήταν ο αριθμός των εποχών. Έτσι, για κάθε μέθοδο προέκυψε ένα σύνολο 10 σεναρίων για μεταβολή εποχών από 500-5000 (εξωτερικά από το πρόγραμμα main), ενώ σε κάθε σενάριο γινόταν μεταβολή στον αριθμό των νευρώνων (από 2 έως 25) αυτόματα, προκειμένου να βρεθεί για πόσους νευρώνες παρουσιαζόταν ελάχιστο σφάλμα. Η μέθοδος εκπαίδευσης που έδωσε τα καλύτερα αποτελέσματα απ' τις παραπάνω ήταν η traingdx (εκπαίδευση με κανόνα απότομης κλίσης, ταυτόχρονη προσαρμογή του ρυθμού μάθησης και προσθήκη όρου ορμής), διότι παρότι και άλλες μέθοδοι έδωσαν μικρά σφάλματα, η traingdx παρουσίασε επιπλέον και καλύτερη σύγκλιση. Σε άλλες μεθόδους (π.χ. traingd) παρουσιάστηκε παλινδρόμηση (αστάθεια), που οφείλεται στο μεγάλο αριθμό νευρώνων από ένα σημείο κι ύστερα. Αυτό σημαίνει ότι τα αποτελέσματα που δίνει η συγκεκριμένη μέθοδος για μεγάλο πλήθος νευρώνων δεν μπορούν να γίνουν αποδεκτά, ακόμη κι αν το σφάλμα προκύπτει μικρό.

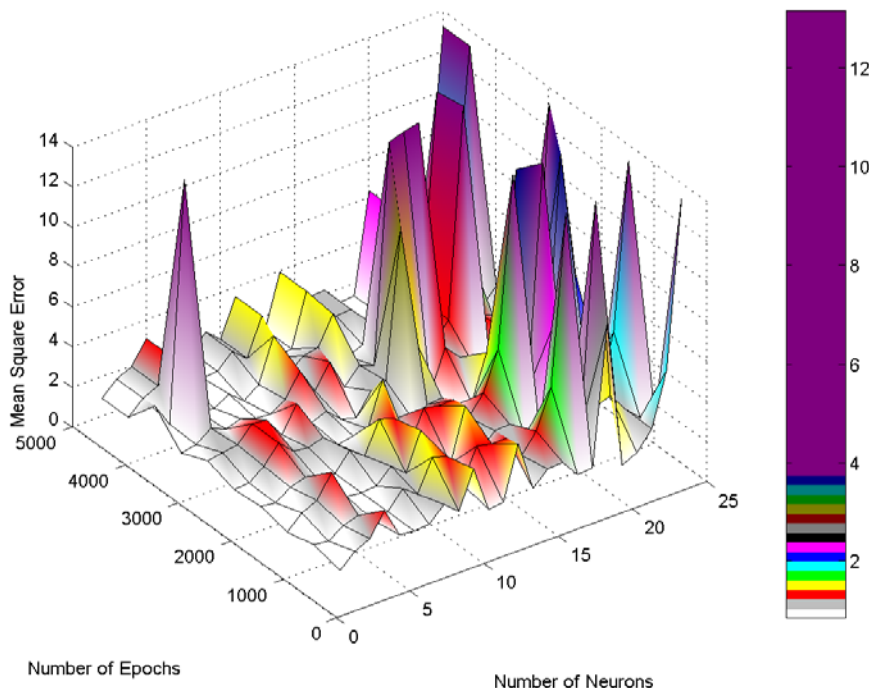
Η μέθοδος που δίνει τα καλύτερα αποτελέσματα (traingdx) βασίζει τη λειτουργία της στην παρουσίαση όλων των προτύπων εισόδου στο δίκτυο και στη μετέπειτα ανανέωση των βαρών. Με τη μέθοδο αυτή το δίκτυο εκπαιδεύεται προσαρμοστικά, δηλαδή με μεταβαλλόμενο ρυθμό μάθησης κατά τη διάρκεια της εκπαίδευσης. Στον κλασικό

κανόνα κλίσης, αν ο ρυθμός οριστεί πολύ μεγάλος, ο αλγόριθμος παρουσιάζει ταλαντώσεις και γίνεται ασταθής, ενώ αν οριστεί πολύ μικρός ο αλγόριθμος αργεί να συγκλίνει. Ένας προσαρμοζόμενος ρυθμός μάθησης βοηθά, ώστε και το βήμα μάθησης να είναι όσο το δυνατό μεγαλύτερο (μεγαλύτερη ταχύτητα), αλλά και η μάθηση να διατηρείται ευσταθής. Η προσαρμογή του ρυθμού μάθησης στη συγκεκριμένη περίπτωση γίνεται ως εξής:

Αρχικά υπολογίζονται η έξοδος και το σφάλμα για το δίκτυο. Σε κάθε εποχή υπολογίζονται καινούρια βάρη, χρησιμοποιώντας τον τρέχοντα ρυθμό μάθησης. Στη συνέχεια υπολογίζονται καινούριες εξοδοί και σφάλματα. Αν το καινούριο σφάλμα υπερβαίνει το παλιό κατά μία προκαθορισμένη τιμή, τα καινούρια βάρη απορρίπτονται. Επιπλέον ο ρυθμός μάθησης μειώνεται. Αλλιώς τα καινούρια βάρη γίνονται αποδεκτά. Εάν το καινούριο σφάλμα είναι μικρότερο απ' το παλιό, ο ρυθμός μάθησης αυξάνεται.

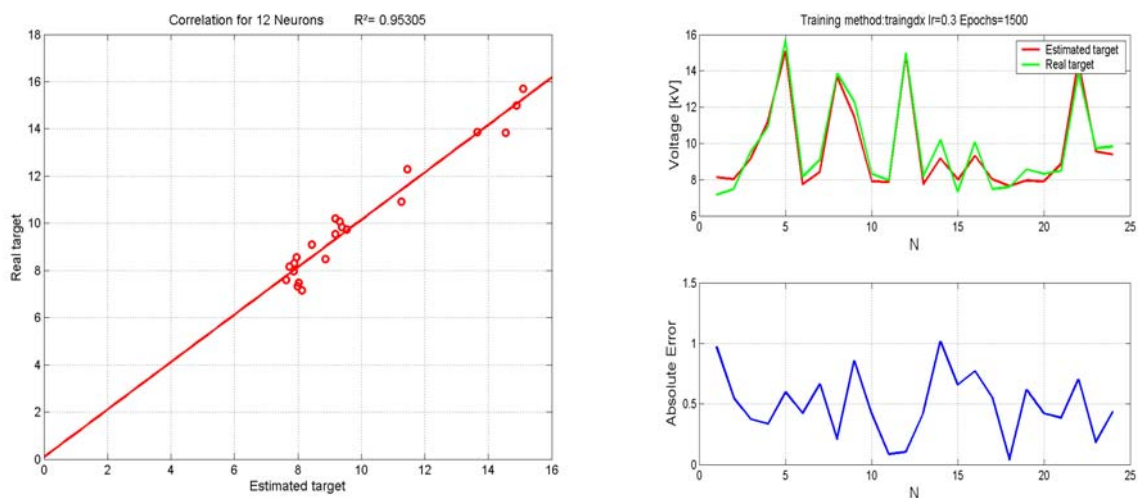
Η `trainIdx` εκπαιδεύει το δίκτυο με βάση τον κανόνα ανάστροφης διάδοσης σφάλματος, με προσαρμοζόμενο ρυθμό μάθησης και παράλληλη προσθήκη όρου ορμής.

Μετά το τέλος των δοκιμών, το τρισδιάστατο γράφημα που παρουσιάζει την εξέλιξη του MSE συναρτήσει του αριθμού εποχών και του πλήθους νευρώνων είναι το εξής:



Σχήμα 9. 1: Τρισδιάστατη καμπύλη για το μέσο τετραγωνικό σφάλμα σαν συνάρτηση των νευρώνων και των εποχών για την περίπτωση της `trainIdx`.

Από το γράφημα αυτό μπορεί κανείς να εντοπίσει τους συνδυασμούς εποχών και νευρώνων για τους οποίους το μέσο τετραγωνικό σφάλμα ελαχιστοποιείται. Αυτό συμβαίνει στις περιοχές όπου ο αριθμός νευρώνων είναι μικρότερος από 12 και ο αριθμός εποχών μικρότερος από 3500. Για μεγαλύτερο αριθμό νευρώνων είναι εμφανές ότι το σύστημα δεν συγκλίνει. Ένα τυπικό γράφημα των δοκιμών αυτών είναι το παρακάτω. Για την περίπτωση αυτή η αρχική τιμή του ρυθμού μάθησης ήταν 0,3, καλύτερο σφάλμα προέκυψε για 12 νευρώνες και ήταν το $MSE=0,3099$ και η συσχέτιση για τους νευρώνες αυτούς ήταν 0,95305.

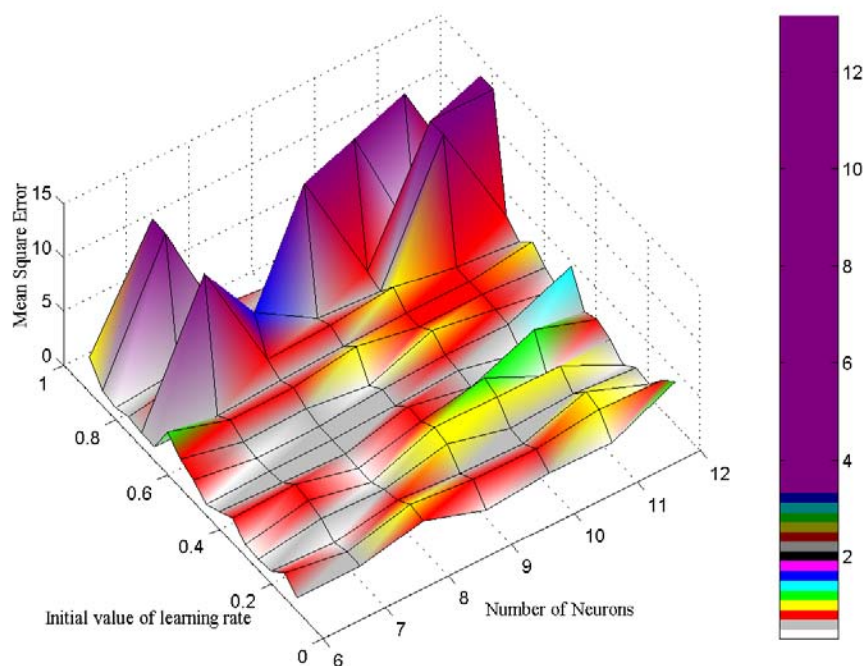


Σχήμα 9.2

Αριστερά: η συσχέτιση μεταξύ πραγματικής και εκτιμούμενης τιμής για το σύνολο ελέγχου για 12 νευρώνες. **Δεξιά:** Κοινή γραφική για τις 24 τιμές του συνόλου ελέγχου και τις αντίστοιχες εκτιμούμενες τιμές που δίνει το νευρωνικό στην περίπτωση ελάχιστου μέσου τετραγωνικού σφάλματος.

Η δεύτερη φάση δοκιμών είχε σαν στόχο τον προσδιορισμό εκείνης της αρχικής τιμής ρυθμού μάθησης (αρχική γιατί στη συνέχεια μεταβάλλεται λόγω προσαρμογής) για την οποία το δίκτυο παρουσιάζει καλύτερη σύγκλιση. Αυτό έγινε κρατώντας σταθερό τον αριθμό των εποχών (1500) και μεταβάλλοντας τους νευρώνες σε μικρότερο πλέον εύρος (από 6 έως 12), αφού γι' αυτήν την περιοχή, όπως φάνηκε και πριν, λαμβάνεται μικρό σφάλμα.

Από αυτό το σύνολο δοκιμών προέκυψε ένα δεύτερο τρισδιάστατο γράφημα που αποτυπώνει συνολικά τα αποτελέσματα και το οποίο είναι το εξής:



Σχήμα 9. 2: Μεταβολή του μέσου τετραγωνικού σφάλματος συναρτήσει των νευρώνων και του ρυθμού μάθησης για μέθοδο εκπαίδευσης την `traingdx`.

Παρατηρείται λοιπόν ότι για 1500 εποχές, καλύτερα αποτελέσματα λαμβάνονται όταν ο αριθμός νευρώνων είναι περίπου 8 και η αρχική τιμή ρυθμού μάθησης όχι μεγαλύτερη από 0,5. Ως προς τον αριθμό των νευρώνων το αποτέλεσμα που προέκυψε είναι αναμενόμενο, αφού συνήθως το πλήθος των νευρώνων δεν πρέπει να υπερβαίνει κατά πολύ το διπλάσιο του αριθμού των μεταβλητών εισόδου (δηλαδή το 10 στην προκειμένη περίπτωση, αφού οι μεταβλητές εισόδου είναι 5).

Το δεύτερο δίκτυο είναι ένα προσαρμοστικό δίκτυο, δηλαδή εκπαιδεύεται με προσαρμοζόμενο ρυθμό μάθησης. Όπως φάνηκε κι απ' τα προηγούμενα, ο προσαρμοζόμενος ρυθμός μάθησης έδωσε τα καλύτερα αποτελέσματα και στο πρώτο δίκτυο. Το προσαρμοστικό αυτό ΤΝΔ δίνει τη δυνατότητα εύρεσης του βέλτιστου συνδυασμού για το πλήθος νευρώνων και για τις παραμέτρους του ρυθμού μάθησης και του όρου ορμής (σταθερός όρος και χρονική παράμετρος). Όπως αναφέρθηκε και πριν το δίκτυο αυτό δίνει καλύτερα αποτελέσματα απ' αυτό που κατασκευάστηκε στο MATLAB, γεγονός που οφείλεται σε δύο κυρίως λόγους:

- Στο πρώτο δίκτυο οι αρχικοποιήσεις των βαρών γίνονται μέσω της συνάρτησης `init` του MATLAB και χωρίς να υπάρχει η δυνατότητα ελέγχου τους και κατάλληλης προσαρμογής τους. Αντίθετα, στο δεύτερο δίκτυο υπάρχει η δυνατότητα αυτή.
- Με την `trainidx` γίνεται μεν προσαρμογή στο ρυθμό μάθησης, αλλά χωρίς να μπορεί κανείς να επέμβει στις παραμέτρους αυτού, κυρίως στο πώς αυτός θα μεταβάλλεται χρονικά. Επιπλέον η ορμή διατηρείται σταθερή. Αντίθετα, στο δεύτερο δίκτυο ορίζεται η χρονική παράμετρος και στο ρυθμό μάθησης αλλά και στην ορμή και μπορεί να δοκιμαστούν διάφορες τιμές για τις παραμέτρους αυτές, ώστε να προκύψει βέλτιστο αποτέλεσμα.

Στον παρακάτω πίνακα παρατίθενται τα βέλτιστα αποτελέσματα για καθένα από τα δύο δίκτυα, σε συγκρίσιμη μορφή, δηλαδή έχει γίνει η μετατροπή του G_{av} σε MSE μέσω της σχέσης $MSE = \sqrt{2 \cdot G_{av}}$.

Το βέλτιστο MSE για το πρώτο δίκτυο είναι, όπως αναφέρθηκε και στο Κεφάλαιο 7, το $MSE_1=0,3099$, ενώ το βέλτιστο G_{av} , όπως αναφέρθηκε στο Κεφάλαιο 8, είναι το $G_{av-test}=0,01212812$ για το σύνολο επιβεβαίωσης (ελέγχου), το οποίο με χρήση του παραπάνω τύπου μπορεί να γίνει $MSE_2=0,16$.

Η συσχέτιση και το σφάλμα για το πρώτο δίκτυο αντιστοιχούν σε πλήθος νευρώνων 12, αριθμό εποχών 1500, και αρχική τιμή ρυθμού μάθησης 0,3. Η συσχέτιση και το σφάλμα για το δεύτερο δίκτυο αντιστοιχούν σε πλήθος νευρώνων 7, σταθερό όρο ρυθμού μάθησης ίσο με $\eta_0=0,8$, χρονική παράμετρο ρυθμού μάθησης ίση με $T_{\eta_0}=4000$, σταθερό όρο ορμής $\alpha_0=0,9$ και χρονική παράμετρο ορμής $T_{\alpha_0}=4000$.

Νευρωνικό δίκτυο	Συσχέτιση για το σύνολο ελέγχου	Μέσο Τετραγωνικό Σφάλμα [kV]
A' (MATLAB)	0,9531	0,31
B' (FORTRAN)	0,9967	0,16

Παρατηρείται ότι το δίκτυο B' δίνει εμφανώς μικρότερο σφάλμα, αλλά κυρίως καλύτερη συσχέτιση, συνεπώς σε αυτό το σημείο υπερτερεί του δικτύου A'. Στα πλεονεκτήματα του πρώτου δικτύου (A') θα πρέπει ωστόσο να συμπεριληφθεί το γεγονός ότι κατασκευάστηκε στο MATLAB, στο οποίο είναι πιο εύκολη και προσιτή η ανάπτυξη

λογισμικού σε χρήστη που δεν είναι εξοικειωμένος με τον προγραμματισμό. Βασικό χαρακτηριστικό του προγραμματισμού σε MATLAB είναι το ότι δίνει τη δυνατότητα εύκολου χειρισμού των σφαλμάτων (δεν είναι δηλαδή δυνατό να κάνει κάποιος ένα λογικό σφάλμα, χωρίς να το αντιληφθεί), είναι φιλικό προς το χρήστη και η κατασκευή του ΤΝΔ γίνεται με αλγόριθμο μικρό σε μέγεθος σε σχέση με αυτόν της Fortran.

Αυτό που θα μπορούσε να εξεταστεί στο μέλλον είναι αν στο πρώτο δίκτυο, το υλοποιημένο με εργαλεία του MATLAB, θα μπορούσε να γίνει κάποια αλλαγή, είτε με εσωτερική τροποποίηση του κώδικα των έτοιμων συναρτήσεων, είτε με την προσθήκη κάποιων επιπλέον παραμέτρων (π.χ. ορισμός βήματος για την προσαρμογή του ρυθμού μάθησης), ώστε να βελτιωθεί η λειτουργία του δικτύου. Επίσης, θα μπορούσαν να γίνουν δοκιμές για διαφοροποιημένα σύνολα εκπαίδευσης και ελέγχου. Στη συγκεκριμένη εφαρμογή ως σύνολο εκπαίδευσης λήφθηκε το 80% του συνόλου των δεδομένων, ενώ ως σύνολο ελέγχου το υπόλοιπο 20%. Θα μπορούσαν λοιπόν να μεταβληθούν οι πληθυσμοί (π.χ. 90% και 10% αντίστοιχα για το σύνολο εκπαίδευσης και ελέγχου ή 70% και 30%) και να ξαναεκτελεστεί το πρόγραμμα για να παρατηρηθούν τα νέα αποτελέσματα. Επίσης, θα μπορούσαν να συνεχιστούν οι δοκιμές για δύο κρυμμένα επίπεδα, αφού μια τέτοια αρχιτεκτονική δείχνει ότι μπορεί να δώσει καλά αποτελέσματα και έχει εφαρμοστεί και σε άλλες περιπτώσεις [10].

Συνοψίζοντας, τα νευρωνικά δίκτυα είναι ένα πολύ χρήσιμο εργαλείο με μεγάλο εύρος εφαρμογών. Μία απ' τις εφαρμογές αυτές θα μπορούσε να είναι και η εκτίμηση της κρίσιμης τάσης σε μονωτήρες. Βέβαια, όπως και κάθε μαθηματικό μοντέλο, έτσι και τα ΤΝΔ δεν δίνουν πάντα σωστά αποτελέσματα (π.χ. περιπτώσεις αστάθειας σαν κι αυτές που εμφανίστηκαν στις εκτελέσεις του προγράμματος). Υπάρχει όμως περιθώριο βελτίωσης αν και τα ήδη υπάρχοντα αποτελέσματα μπορούν να χαρακτηριστούν ικανοποιητικά.

Βιβλιογραφία

- [1] Κ. Θ. Δέρβος: «Μονωτικά Υλικά Υψηλών Τάσεων», Αθήνα 2001
- [2] Tomas Gustavsson: “Outdoor Aging of Silicone Rubber Formulations in Coastal Environment”, Licentiate thesis. Göteborg, Sweden, Chalmers University of Technology, School of Electrical and Computer Engineering, Technical report No. 353L, June 2000
- [3] Eugenio Concha: “ILMOS, Insulator Leakage Current Monitor”. Άρθρο διαθέσιμο στην ιστοσελίδα:
http://www.psptechnologies.com/Assets/nonhtml/articulo_1_ingles.pdf
- [4] IEC 815, “Guide for the selection of insulators in respect of polluted conditions”, 1986
- [5] J.S.T. Looms, “Insulators for high voltages”, Peter Peregrinus Ltd., London, Unighted Kingdom, 1990
- [6] Ravi S. Gorur: “Using Corona Cameras to Inspect Composite Insulators”, INMR-Quarterly Review 2004, Issue 63, Vol. 12
- [7] Andreas Dernfalk: “New Laser Technique for Diagnostics of Composite Insulators”, INMR-Quarterly Review 2004, Issue 65, Vol. 12
- [8] Ravi S. Gorur, Sreeram Venkataraman: “Wavelet Transformation as a Tool for Predicting Failure of Polymeric Materials”, INMR-Quarterly Review 2004, Issue 65, Vol. 12
- [9] Ahmad S. Ahmad, P. S. Ghosh, Syed Abdul Kader Aljunid, Hussein Ahmad Ismail Said, Halil Hussain: “Artificial Neural Network for Contamination Severity Assessment of High Voltage Insulators Under Various Meteorological Conditions”, AUPEC, Perth, 23-26 September 2001
- [10] P. S. Ghosh, S. Chakravorti, N. Chatterjee: “Estimation of Time-to-flashover Characteristics of Contaminated Electrolytic Surfaces using a Neural Network”,

- IEEE Transactions on Dielectrics and Electrical Insulation, Vol. 2, No. 6, December 1995
- [11] Paul Cline, Will Lannes, Gill Richards: "Use of pollution monitors with a neural network to predict insulator flashover", ELSEVIER Electric Power Systems Research 42 (1997)
- [12] M. Ugur, D. W. Auckland, B. R. Varlow, Z. Emin: "Neural Networks to Analyse Surface Tracking on Solid Insulators", IEEE Transactions on Dielectrics and Electrical Insulation, Vol. 4, No. 6, December 1997
- [13] STRI Guide, Guide for Visual Identification of Deterioration & Damages on Suspension Composite Insulators, 2003
- [14] Σπύρος Γ. Τζαφέστας: «Υπολογιστική Νοημοσύνη, Τόμος Α: Μεθοδολογίες», Αθήνα 2002
- [15] Χαράλαμπος Ν. Ηλίας: «Ανάπτυξη μοντέλου βραχυπρόθεσμης πρόβλεψης φορτίου με χρήση ασαφούς λογικής και κατηγοριοποίηση τυπικών ημερών», Μεταπτυχιακή Διπλωματική Εργασία, Διατμηματικό Πρόγραμμα Μεταπτυχιακών Σπουδών «Παραγωγή και Διαχείριση Ενέργειας», Ε.Μ.Π., Αθήνα, Οκτώβριος 2002
- [16] Γ. Κονταξής, Ν. Χατζηαργυρίου: «Κέντρα Ελέγχου Ενέργειας», Αθήνα 2003
- [17] Γ. Τσεκούρας: «Εφαρμογή της μεθόδου των Τεχνητών Νευρωνικών Δικτύων σε θέματα σηράγγων», Διπλωματική Εργασία, Αθήνα 2004
- [18] Kevin Gurney: "Neural Nets", Brunel University Middx, διαθέσιμο στην ιστοσελίδα: <http://www.shef.ac.uk>
- [19] Αριστείδης Λύκας: «Υπολογιστική Νοημοσύνη», Ιωάννινα 1999
- [20] Christos Stergiou, Dimitrios Siganos:
http://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol4/cs11/report.html
- [21] Ι. Α. Σταθόπουλος: «Υψηλές Τάσεις Ι», Εκδόσεις Συμεών, Αθήνα
- [22] IEC 507, "Artificial pollution tests on high-voltage insulators to be used on a.c. systems", 1991
- [23] F. V. Topalis, I. F. Gonos, I. A. Stathopoulos, "Dielectric behaviour of polluted porcelain insulators", IEE Proceedings – Generation, Transmission and Distribution, Vol. 148, No. 4, July 2001, pp. 269-274

- [24] Ξευγένης Αλέξανδρος: «Μελέτη της λειτουργίας μονωτήρων αναρτήσεως σε τεχνητά ρυπασμένο περιβάλλον με τη μέθοδο δημιουργίας αλατούχου ομίχλης», Διπλωματική Εργασία, Αθήνα 1986
- [25] F.V. Topalis, I.F. Gonos, I.A. Stathopoulos, “Dielectric behaviour of polluted porcelain insulators”, IEE Proceedings – Generation, Transmission and Distribution, Vol. 148, No. 4, July 2001, pp. 269-274
- [26] G.I. Theodorakis, F.V. Topalis, I.A. Stathopoulos, “Parameter identification of the polluted insulator model”, Proceedings of international symposium on simulation and modelling, Lugano, Switzerland 1989, pp. 103-106
- [27] I.F. Gonos, S.A. Sufliis, F.V. Topalis, I.A. Stathopoulos, “Computer methods for the determination of the critical parameters of polluted insulators”, Proceedings of the 1st Mediterranean Conference on Power Generation, Transmission and Distribution (MEDPOWER '98), Session 7, Insulation Co-ordination and Optimisation, Lefkosia, (Nicosia), Cyprus, 16 – 18 November 1998
- [28] F. Obenhaus, “Fremdschichtüberschlag und Kriechweglänge”, Deutsche Electrotechnik, Vol. 4, 1958, pp. 135-136
- [29] L. L. Alston, S. Zoledziowski, “Growth of discharges on polluted insulation”, Proc. IEE, Vol. 110, No. 7, July 1963, pp. 1260-1266
- [30] R.J. Xavier, Y.N. Rao, “Study of surface conductivity and E.S.D.D. on contaminated porcelain insulating surfaces”, 5th International Symposium on High Voltage Engineering, Braunschweig, August 1987, Vol. II, paper 51.12
- [31] R. Wilkins, “Flashover voltage of high-voltage insulators with uniform surface-pollution films”, Proc. IEE, Vol. 116, No. 3, March 1969, pp. 457-465
- [32] I.F. Gonos, F.V. Topalis, I.A. Stathopoulos, “A model for the flashover process on non-uniformly polluted insulators”, International Journal of Modelling and Simulation, Vol. 22, No. 3, 2002, pp. 137-141
- [33] S.A. Sufliis, I.F. Gonos, F.V. Topalis, I.A. Stathopoulos, “Study of the dielectric behaviour of non-uniformly polluted insulators”, XIIIth International Symposium on High Voltage Engineering, Netherlands, August 2003
- [34] Matlab Help, Version 6.5

Παράρτημα

Π.1 Σύνολο Δεδομένων

Στον Πίνακα Data παρουσιάζονται όλα τα δεδομένα και αυτά που προέκυψαν από πειραματικές μετρήσεις, αλλά και αυτά που προέκυψαν από εφαρμογή του μαθηματικού μοντέλου. Στα αριστερά κελιά δίνονται τα μοντέλα των μονωτήρων (No1-No14), οι οποίοι είναι όλοι τύπου cap & pin.

Πίνακας Data: Πίνακας που περιλαμβάνει όλα τα δεδομένα

	ΕΙΣΟΔΟΙ					ΕΞΟΔΟΣ
	D _m [cm]	H [cm]	L [cm]	F	σ _s (μS)	U _c (kV)
No 1	26,8	15,9	33	0,79	49	12,32
	26,8	15,9	33	0,79	60	11,479
	26,8	15,9	33	0,79	87	10,079
	26,8	15,9	33	0,79	103	9,499
	26,8	15,9	33	0,79	127	8,825
	26,8	15,9	33	0,79	138	8,57
	26,8	15,9	33	0,79	180	7,803
	26,8	15,9	33	0,79	194	7,599
	26,8	15,9	33	0,79	204	7,465
No 2	26,8	15,9	40,6	0,86	49	13,643
	26,8	15,9	40,6	0,86	60	12,727
	26,8	15,9	40,6	0,86	87	11,199
	26,8	15,9	40,6	0,86	103	10,566
	26,8	15,9	40,6	0,86	127	9,828
	26,8	15,9	40,6	0,86	138	9,55
	26,8	15,9	40,6	0,86	180	8,71
	26,8	15,9	40,6	0,86	194	8,486
	26,8	15,9	40,6	0,86	204	8,34
No 3	25,4	16,5	43,2	0,9	49	13,86
	25,4	16,5	43,2	0,9	60	12,937
	25,4	16,5	43,2	0,9	87	11,396
	25,4	16,5	43,2	0,9	103	10,757
	25,4	16,5	43,2	0,9	127	10,012
	25,4	16,5	43,2	0,9	138	9,731
	25,4	16,5	43,2	0,9	180	8,883
	25,4	16,5	43,2	0,9	194	8,657
	25,4	16,5	43,2	0,9	204	8,508
No 4	25,4	14,6	27,9	0,68	49	12
	25,4	14,6	27,9	0,68	60	11,1

	25,4	14,6	27,9	0,68	87	8,7
	25,4	14,6	27,9	0,68	103	9,1
	25,4	14,6	27,9	0,68	127	7,5
	25,4	14,6	27,9	0,68	138	7,8
	25,4	14,6	27,9	0,68	180	6,2
	25,4	14,6	27,9	0,68	194	6,8
	25,4	14,6	27,9	0,68	204	6,1
No 5	25,4	14,6	43,2	0,92	7,801	26
	25,4	14,6	43,2	0,92	18,8725	19
	25,4	14,6	43,2	0,92	37,325	15
	25,4	14,6	43,2	0,92	59,468	13
	25,4	14,6	43,2	0,92	81,611	12
	25,4	14,6	43,2	0,92	111,135	10,5
No 6	25,4	14,6	31,8	0,72	49	11,814
	25,4	14,6	31,8	0,72	60	11,007
	25,4	14,6	31,8	0,72	87	9,664
	25,4	14,6	31,8	0,72	103	9,107
	25,4	14,6	31,8	0,72	127	8,46
	25,4	14,6	31,8	0,72	138	8,216
	25,4	14,6	31,8	0,72	180	7,48
	25,4	14,6	31,8	0,72	194	7,284
	25,4	14,6	31,8	0,72	204	7,156
No 7	29,2	15,9	47	0,92	49	15,129
	29,2	15,9	47	0,92	60	14,119
	29,2	15,9	47	0,92	87	12,435
	29,2	15,9	47	0,92	103	11,737
	29,2	15,9	47	0,92	127	10,923
	29,2	15,9	47	0,92	138	10,616
	29,2	15,9	47	0,92	180	9,688
	29,2	15,9	47	0,92	194	9,441
	29,2	15,9	47	0,92	204	9,279
No 8	27,9	15,6	36,8	0,76	49	13,101
	27,9	15,6	36,8	0,76	60	12,211
	27,9	15,6	36,8	0,76	87	10,73
	27,9	15,6	36,8	0,76	103	10,116
	27,9	15,6	36,8	0,76	127	9,402
	27,9	15,6	36,8	0,76	138	9,132
	27,9	15,6	36,8	0,76	180	8,319
	27,9	15,6	36,8	0,76	194	8,103
	27,9	15,6	36,8	0,76	204	7,961
No 9	32,1	17,8	54,6	0,96	49	16,818
	32,1	17,8	54,6	0,96	60	15,701
	32,1	17,8	54,6	0,96	87	13,839
	32,1	17,8	54,6	0,96	103	13,066
	32,1	17,8	54,6	0,96	127	12,165

	32,1	17,8	54,6	0,96	138	11,825
	32,1	17,8	54,6	0,96	180	10,798
	32,1	17,8	54,6	0,96	194	10,525
	32,1	17,8	54,6	0,96	204	10,345
No 10	25,4	14,6	30,5	0,7	7,801	22
	25,4	14,6	30,5	0,7	18,8725	16
	25,4	14,6	30,5	0,7	37,325	13
	25,4	14,6	30,5	0,7	59,468	11
	25,4	14,6	30,5	0,7	81,611	10
	25,4	14,6	30,5	0,7	111,135	8,5
No 11	28	17	37	0,8	49	13,195
	28	17	37	0,8	60	12,3
	28	17	37	0,8	87	10,809
	28	17	37	0,8	103	10,191
	28	17	37	0,8	127	9,472
	28	17	37	0,8	138	9,201
	28	17	37	0,8	180	8,382
	28	17	37	0,8	194	8,165
	28	17	37	0,8	204	8,022
No 12	25,4	14,5	30,5	0,74	49	11,608
	25,4	14,5	30,5	0,74	60	10,813
	25,4	14,5	30,5	0,74	87	9,49
	25,4	14,5	30,5	0,74	103	8,941
	25,4	14,5	30,5	0,74	127	8,304
	25,4	14,5	30,5	0,74	138	8,064
	25,4	14,5	30,5	0,74	180	7,339
	25,4	14,5	30,5	0,74	194	7,146
	25,4	14,5	30,5	0,74	204	7,02
No 13	20	16,5	40	1,29	49	12,989
	20	16,5	40	1,29	60	12,136
	20	16,5	40	1,29	87	10,711
	20	16,5	40	1,29	103	10,119
	20	16,5	40	1,29	127	9,43
	20	16,5	40	1,29	138	9,169
	20	16,5	40	1,29	180	8,382
	20	16,5	40	1,29	194	8,172
	20	16,5	40	1,29	204	8,034
No 14	22,9	16,6	43,2	1,38	7,801	23,5
	22,9	16,6	43,2	1,38	11,4915	20,9
	22,9	16,6	43,2	1,38	15,182	19,4
	22,9	16,6	43,2	1,38	18,8725	18,3
	22,9	16,6	43,2	1,38	22,563	16,9
	22,9	16,6	43,2	1,38	37,325	15,8
	22,9	16,6	43,2	1,38	74,23	13,6

Μετά την εκτέλεση του βοηθητικού προγράμματος, τα σύνολα εκπαίδευσης και επιβεβαίωσης προέκυψαν ως εξής:

TrainData:

D_m (cm)	H (cm)	L (cm)	F	σ_s (μ S)	U_c (kV)
26,8000	15,9000	33,0000	0,7900	49,0000	12,3200
26,8000	15,9000	33,0000	0,7900	60,0000	11,4790
26,8000	15,9000	33,0000	0,7900	87,0000	10,0790
26,8000	15,9000	33,0000	0,7900	103,0000	9,4990
26,8000	15,9000	33,0000	0,7900	127,0000	8,8250
26,8000	15,9000	33,0000	0,7900	138,0000	8,5700
26,8000	15,9000	33,0000	0,7900	180,0000	7,8030
26,8000	15,9000	33,0000	0,7900	194,0000	7,5990
26,8000	15,9000	33,0000	0,7900	204,0000	7,4650
26,8000	15,9000	40,6000	0,8600	49,0000	13,6430
26,8000	15,9000	40,6000	0,8600	60,0000	12,7270
26,8000	15,9000	40,6000	0,8600	87,0000	11,1990
26,8000	15,9000	40,6000	0,8600	103,0000	10,5660
26,8000	15,9000	40,6000	0,8600	138,0000	9,5500
26,8000	15,9000	40,6000	0,8600	180,0000	8,7100
26,8000	15,9000	40,6000	0,8600	194,0000	8,4860
26,8000	15,9000	40,6000	0,8600	204,0000	8,3400
25,4000	16,5000	43,2000	0,9000	49,0000	13,8600
25,4000	16,5000	43,2000	0,9000	60,0000	12,9370
25,4000	16,5000	43,2000	0,9000	87,0000	11,3960
25,4000	16,5000	43,2000	0,9000	103,0000	10,7570
25,4000	16,5000	43,2000	0,9000	127,0000	10,0120
25,4000	16,5000	43,2000	0,9000	138,0000	9,7310
25,4000	16,5000	43,2000	0,9000	180,0000	8,8830
25,4000	16,5000	43,2000	0,9000	194,0000	8,6570
25,4000	16,5000	43,2000	0,9000	204,0000	8,5080
25,4000	14,6000	27,9000	0,6800	49,0000	12,0000

25,4000	14,6000	27,9000	0,6800	60,0000	11,1000
25,4000	14,6000	27,9000	0,6800	87,0000	8,7000
25,4000	14,6000	27,9000	0,6800	103,0000	9,1000
25,4000	14,6000	27,9000	0,6800	127,0000	7,5000
25,4000	14,6000	27,9000	0,6800	138,0000	7,8000
25,4000	14,6000	27,9000	0,6800	180,0000	6,2000
25,4000	14,6000	27,9000	0,6800	194,0000	6,8000
25,4000	14,6000	27,9000	0,6800	204,0000	6,1000
25,4000	14,6000	43,2000	0,9200	7,8010	26,0000
25,4000	14,6000	43,2000	0,9200	18,8725	19,0000
25,4000	14,6000	43,2000	0,9200	37,3250	15,0000
25,4000	14,6000	43,2000	0,9200	59,4680	13,0000
25,4000	14,6000	43,2000	0,9200	81,6110	12,0000
25,4000	14,6000	43,2000	0,9200	111,1350	10,5000
25,4000	14,6000	31,8000	0,7200	49,0000	11,8140
25,4000	14,6000	31,8000	0,7200	60,0000	11,0070
25,4000	14,6000	31,8000	0,7200	87,0000	9,6640
25,4000	14,6000	31,8000	0,7200	103,0000	9,1070
25,4000	14,6000	31,8000	0,7200	127,0000	8,4600
25,4000	14,6000	31,8000	0,7200	138,0000	8,2160
25,4000	14,6000	31,8000	0,7200	180,0000	7,4800
25,4000	14,6000	31,8000	0,7200	194,0000	7,2840
25,4000	14,6000	31,8000	0,7200	204,0000	7,1560
29,2000	15,9000	47,0000	0,9200	49,0000	15,1290
29,2000	15,9000	47,0000	0,9200	60,0000	14,1190
29,2000	15,9000	47,0000	0,9200	87,0000	12,4350
29,2000	15,9000	47,0000	0,9200	103,0000	11,7370
29,2000	15,9000	47,0000	0,9200	127,0000	10,9230
29,2000	15,9000	47,0000	0,9200	138,0000	10,6160
29,2000	15,9000	47,0000	0,9200	180,0000	9,6880
29,2000	15,9000	47,0000	0,9200	194,0000	9,4410

29,2000	15,9000	47,0000	0,9200	204,0000	9,2790
27,9000	15,6000	36,8000	0,7600	49,0000	13,1010
27,9000	15,6000	36,8000	0,7600	60,0000	12,2110
27,9000	15,6000	36,8000	0,7600	87,0000	10,7300
27,9000	15,6000	36,8000	0,7600	103,0000	10,1160
27,9000	15,6000	36,8000	0,7600	127,0000	9,4020
27,9000	15,6000	36,8000	0,7600	138,0000	9,1320
27,9000	15,6000	36,8000	0,7600	180,0000	8,3190
27,9000	15,6000	36,8000	0,7600	194,0000	8,1030
27,9000	15,6000	36,8000	0,7600	204,0000	7,9610
32,1000	17,8000	54,6000	0,9600	49,0000	16,8180
32,1000	17,8000	54,6000	0,9600	60,0000	15,7010
32,1000	17,8000	54,6000	0,9600	87,0000	13,8390
32,1000	17,8000	54,6000	0,9600	103,0000	13,0660
32,1000	17,8000	54,6000	0,9600	127,0000	12,1650
32,1000	17,8000	54,6000	0,9600	138,0000	11,8250
32,1000	17,8000	54,6000	0,9600	180,0000	10,7980
32,1000	17,8000	54,6000	0,9600	194,0000	10,5250
32,1000	17,8000	54,6000	0,9600	204,0000	10,3450
25,4000	14,6000	30,5000	0,7000	7,8010	22,0000
25,4000	14,6000	30,5000	0,7000	18,8725	16,0000
25,4000	14,6000	30,5000	0,7000	37,3250	13,0000
25,4000	14,6000	30,5000	0,7000	59,4680	11,0000
25,4000	14,6000	30,5000	0,7000	81,6110	10,0000
25,4000	14,6000	30,5000	0,7000	111,1350	8,5000
28,0000	17,0000	37,0000	0,8000	49,0000	13,1950
28,0000	17,0000	37,0000	0,8000	60,0000	12,3000
28,0000	17,0000	37,0000	0,8000	87,0000	10,8090
28,0000	17,0000	37,0000	0,8000	103,0000	10,1910
28,0000	17,0000	37,0000	0,8000	127,0000	9,4720
28,0000	17,0000	37,0000	0,8000	138,0000	9,2010

28,0000	17,0000	37,0000	0,8000	180,0000	8,3820
28,0000	17,0000	37,0000	0,8000	194,0000	8,1650
28,0000	17,0000	37,0000	0,8000	204,0000	8,0220
25,4000	14,5000	30,5000	0,7400	49,0000	11,6080
25,4000	14,5000	30,5000	0,7400	60,0000	10,8130
25,4000	14,5000	30,5000	0,7400	87,0000	9,4900
25,4000	14,5000	30,5000	0,7400	103,0000	8,9410
25,4000	14,5000	30,5000	0,7400	127,0000	8,3040
25,4000	14,5000	30,5000	0,7400	138,0000	8,0640
25,4000	14,5000	30,5000	0,7400	180,0000	7,3390
25,4000	14,5000	30,5000	0,7400	194,0000	7,1460
25,4000	14,5000	30,5000	0,7400	204,0000	7,0200
20,0000	16,5000	40,0000	1,2900	49,0000	12,9890
20,0000	16,5000	40,0000	1,2900	60,0000	12,1360
20,0000	16,5000	40,0000	1,2900	87,0000	10,7110
20,0000	16,5000	40,0000	1,2900	103,0000	10,1190
20,0000	16,5000	40,0000	1,2900	127,0000	9,4300
20,0000	16,5000	40,0000	1,2900	138,0000	9,1690
20,0000	16,5000	40,0000	1,2900	180,0000	8,3820
20,0000	16,5000	40,0000	1,2900	194,0000	8,1720
20,0000	16,5000	40,0000	1,2900	204,0000	8,0340
22,9000	16,6000	43,2000	1,3800	7,8010	23,5000
22,9000	16,6000	43,2000	1,3800	11,4915	20,9000
22,9000	16,6000	43,2000	1,3800	15,1820	19,4000
22,9000	16,6000	43,2000	1,3800	18,8725	18,3000
22,9000	16,6000	43,2000	1,3800	22,5630	16,9000
22,9000	16,6000	43,2000	1,3800	37,3250	15,8000
22,9000	16,6000	43,2000	1,3800	74,2300	13,6000

TestData:

D _m (cm)	H (cm)	L (cm)	F	σ _s (μS)	U _c (kV)
25,4000	14,6000	31,8000	0,7200	204,0000	7,1560
25,4000	14,6000	31,8000	0,7200	180,0000	7,4800
26,8000	15,9000	40,6000	0,8600	138,0000	9,5500
29,2000	15,9000	47,0000	0,9200	127,0000	10,9230
32,1000	17,8000	54,6000	0,9600	60,0000	15,7010
28,0000	17,0000	37,0000	0,8000	194,0000	8,1650
25,4000	14,6000	27,9000	0,6800	103,0000	9,1000
25,4000	16,5000	43,2000	0,9000	49,0000	13,8600
28,0000	17,0000	37,0000	0,8000	60,0000	12,3000
27,9000	15,6000	36,8000	0,7600	180,0000	8,3190
27,9000	15,6000	36,8000	0,7600	204,0000	7,9610
25,4000	14,6000	43,2000	0,9200	37,3250	15,0000
28,0000	17,0000	37,0000	0,8000	194,0000	8,1650
28,0000	17,0000	37,0000	0,8000	103,0000	10,1910
25,4000	14,5000	30,5000	0,7400	180,0000	7,3390
26,8000	15,9000	33,0000	0,7900	87,0000	10,0790
25,4000	14,6000	31,8000	0,7200	180,0000	7,4800
26,8000	15,9000	33,0000	0,7900	194,0000	7,5990
26,8000	15,9000	33,0000	0,7900	138,0000	8,5700
27,9000	15,6000	36,8000	0,7600	180,0000	8,3190
26,8000	15,9000	40,6000	0,8600	194,0000	8,4860
32,1000	17,8000	54,6000	0,9600	87,0000	13,8390
25,4000	16,5000	43,2000	0,9000	138,0000	9,7310
26,8000	15,9000	40,6000	0,8600	127,0000	9,8280

Π.2 Βοηθητικά προγράμματα

R2.m

```
function F = R2(X,Y)

Cxx=0;
Cyy=0;
Cxy=0;
N=length(X);
x=mean(X);
y=mean(Y);
for i=1:N
    Cxx=Cxx+(X(i)-x)^2;
    Cyy=Cyy+(Y(i)-y)^2;
    Cxy=Cxy+(X(i)-x)*(Y(i)-y);
end
b=Cxy/Cxx;
a=y-b*x;

F(2)=a;
F(1)=b;
F(3)=Cxy*Cxy/Cxx/Cyy;
```

MakeLine.m

```
function MakeLine = R2(a,b, Xmin, Xmax)

dx=Xmax/100;
for i=1:110
    x=(i-1)*dx;
    y=a*x+b;
    X(1,i)=x;
    X(2,i)=y;
end
MakeLine=X;
```

Correlation.m

```
clear all
close all

load senario1_target.txt
load senario1_error.txt
load senario1_sfalma.txt
load TestData_File.txt

target=senario111_target;
error=senario111_error;
sfalma=senario111_sfalma;
M=length(error);

for N=1:M
    Xax=target(N,:);
    Yax=TestData_File(:,6);
    A=R2(Xax, Yax);
    a=A(1);    b=A(2);    r2=A(3)
    Xmn=min(Xax);    Xmx=max(Xax);

    Dat=MakeLine(a,b,Xmn,Xmx);
    XX=Dat(1,:);
    YY=Dat(2,:);

    Neurones=N+1
    figure(N+1)
    p2=plot(Xax, Yax,'or');
    hold on
    grid
    p1=plot(XX, YY,'r');
    x1=xlabel('Estimated target');
    y1=ylabel('Real target');
    set(p1,'linewidth',2)
    set(p2,'linewidth',2)
    set(x1,'FontName','TimesNewRoman','FontSize',12);
    set(y1,'FontName','TimesNewRoman','FontSize',12);
```

```
axis([0 16 0 18])
str1 = num2str(N+1);
str2 = num2str(r2);
out = ['Correlation for ' str1 ' Neurons          R2= ' str2];
t1=title(out);
set(t1,'FontName', 'TimesNewRoman', 'FontSize', 12);
end
```

```
Minerror=min(error)
ii=1;
while Minerror ~error(ii)
    ii=ii+1;
end
```

```
figure (111)
subplot(2,1,1)
p1=plot(target(ii,:), 'r');
hold on
p2=plot(TestData_File(:,6), 'g');
grid
x1=xlabel('N');
y1=ylabel('Voltage [kV]');
set(x1,'FontName', 'TimesNewRoman', 'FontSize', 14);
set(y1,'FontName', 'TimesNewRoman', 'FontSize', 14);
set(p1,'linewidth',2)
set(p2,'linewidth',2)
legend('Estimated target', 'Real target');
t1=title('Training method:traingd lr=0.5 Epochs=500');
set(t1,'FontName', 'TimesNewRoman', 'FontSize', 12);
```

```
subplot(2,1,2)
ww=target(ii,:)-TestData_File(:,6)';
p3=plot(abs(ww), 'b');
```

```
grid
x1=xlabel('N');
y1=ylabel('Absolute Error');
set(x1,'FontName','TimesNewRoman','FontSize',14);
set(y1,'FontName','TimesNewRoman','FontSize',14);
set(p3,'linewidth',2)
```