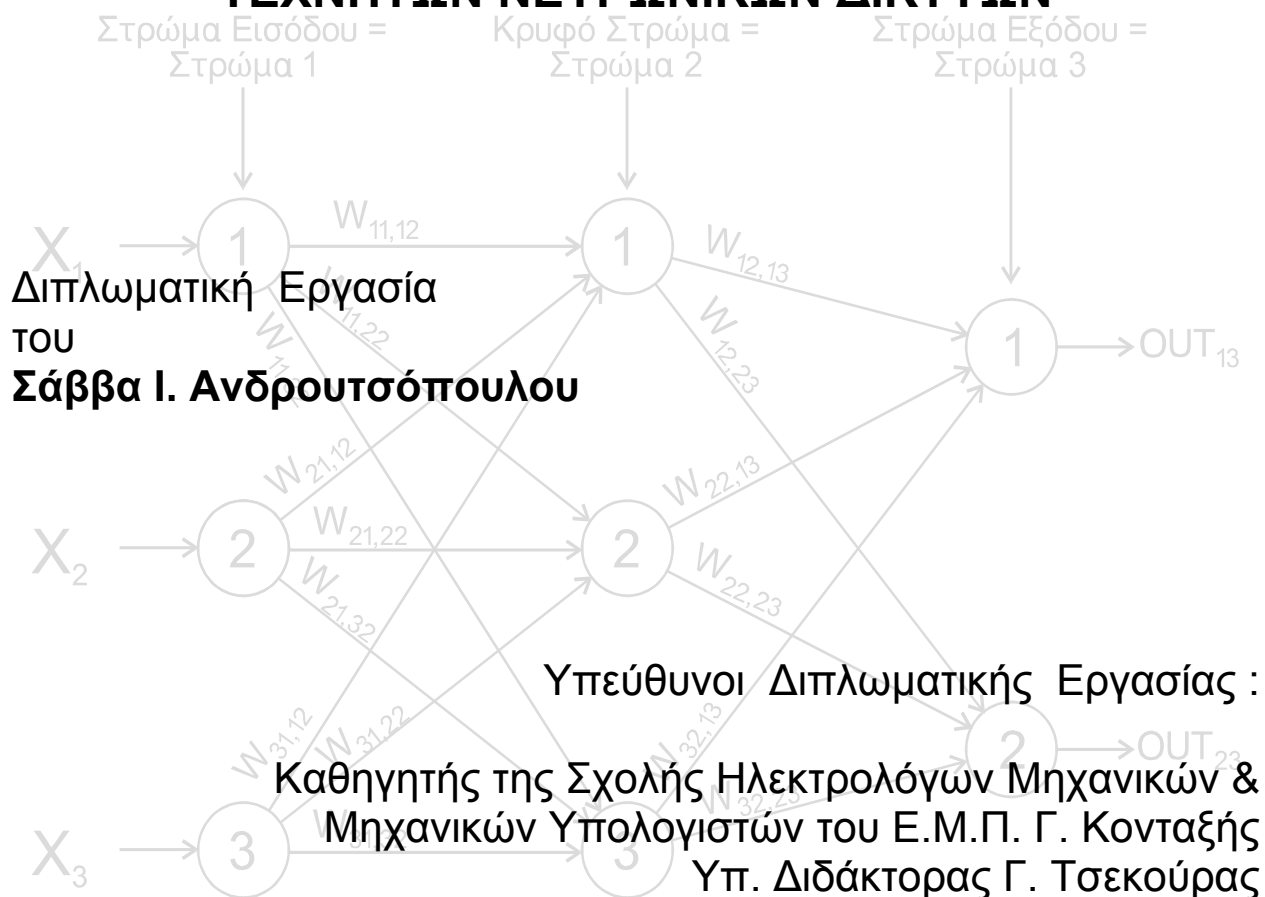




ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

Σχολή Ηλεκτρολόγων Μηχανικών και
Μηχανικών Υπολογιστών
Τομέας Ηλεκτρικής Ισχύος

ΒΡΑΧΥΠΡΟΘΕΣΜΗ ΠΡΟΒΛΕΨΗ ΦΟΡΤΙΟΥ ΜΕ ΤΗ ΧΡΗΣΗ ΠΑΡΑΛΛΑΓΩΝ ΤΟΥ ΑΛΓΟΡΙΘΜΟΥ ΤΗΣ ΑΝΑΣΤΡΟΦΗΣ ΔΙΑΔΟΣΗΣ ΣΦΑΛΜΑΤΟΣ ΤΩΝ ΤΕΧΝΗΤΩΝ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ



*Στην οικογένειά μου
και στους φίλους μου*

Πρόλογος

Από την αρχή της δημιουργίας των συστημάτων ηλεκτρικής ενέργειας, κρίθηκε απαραίτητη η προσέγγιση των μελλοντικών ζητούμενων φορτίων. Αυτή η ανάγκη πήγαζε κυρίως από το γεγονός ότι, όταν παράγεται η ηλεκτρική ενέργεια από κάποιον σταθμό παραγωγής, θα πρέπει να καταναλώνεται επί τόπου καθώς υπάρχει ελάχιστη δυνατότητα αποθήκευσης. Συνεπώς το κάθε σύστημα παραγωγής ηλεκτρικής ενέργειας θέλει να γνωρίζει με ικανοποιητική ακρίβεια τα μελλοντικά φορτία, ώστε η παραγόμενη ισχύς του να καλύπτει τα φορτία μαζί με τις απώλειες του δικτύου.

Μια πρώτη, αλλά σημαντική προσέγγιση αυτού του προβλήματος έγινε μέσω των κλασικών μεθόδων που περιλαμβάνουν τη χρήση πινάκων και την κατασκευή μαθηματικών συναρτήσεων, οι οποίες επεξεργάζονται ιστορικά δεδομένα και επιστρέφουν ως αποτέλεσμα την πρόβλεψη για τα μελλοντικά φορτία.

Την τελευταία δεκαετία κυρίως, έχει γίνει μια πολύ σοβαρή προσπάθεια για την αντιμετώπιση του προβλήματος με τη χρήση νέων εξελιγμένων μεθόδων, όπως η ασαφής λογική και τα τεχνητά νευρωνικά δίκτυα, μέθοδοι που βέβαια απαιτούν σημαντική υπολογιστική ισχύ.

Σκοπός της συγκεκριμένης διπλωματικής εργασίας είναι κυρίως η μελέτη του αλγορίθμου εκπαίδευσης νευρωνικών δικτύων της ανάστροφης διάδοσης σφάλματος - Back Propagation, για τη βραχυπρόθεσμη πρόβλεψη φορτίου που αφορά τις επόμενες 24 ώρες.

Στο πρώτο κεφάλαιο, αναφερόμαστε γενικά στην πρόβλεψη φορτίου και τη σημασία της, στα είδη αυτής και κυρίως στη βραχυπρόθεσμη πρόβλεψη φορτίου.

Στο δεύτερο κεφάλαιο περιγράφουμε με αναλυτικό τρόπο τις κλασικές και τις νέες μεθόδους πρόβλεψης φορτίου. Οι τελευταίες αφορούν τη χρήση τεχνητών νευρωνικών δικτύων και ασαφούς λογικής, καθώς και συνδυασμών αυτών.

Στο τρίτο κεφάλαιο περιγράφουμε αναλυτικά το θεωρητικό μέρος του αλγορίθμου της ανάστροφης διάδοσης σφάλματος και κατόπιν παραθέτουμε ένα αριθμητικό παράδειγμα εφαρμογής του.

Στο τέταρτο κεφάλαιο περιγράφουμε με διεξοδικό τρόπο το πρόγραμμα που υλοποιεί στην πράξη τον αλγόριθμο, τα σενάρια εκπαίδευσης του δικτύου που κατασκευάσαμε και τις εκτελέσεις που έλαβαν χώρα. Επίσης σημειώνουμε τις παρατηρήσεις και τα σχόλιά μας σχετικά με τα αποτελέσματα που προέκυψαν.

Στο πέμπτο κεφάλαιο αναφέρουμε κάποια γενικά συμπεράσματα, καθώς και ορισμένες, χρήσιμες κατά τη γνώμη μας, προτάσεις για περαιτέρω έρευνα και εφαρμογές.

Στο παράρτημα I περιέχονται όλες οι γραφικές παραστάσεις που προέκυψαν από τις εκτελέσεις των προγραμμάτων.

Τέλος το Παράρτημα I περιλαμβάνει τη βιβλιογραφία, από την οποία αντλήσαμε στοιχεία για την παρούσα διπλωματική εργασία.

ΕΥΧΑΡΙΣΤΙΕΣ

Στο σημείο αυτό θα ήθελα να ευχαριστήσω θερμά και να εκφράσω την ειλικρινή μου ευγνωμοσύνη, σε όσους στάθηκαν δίπλα μου με κάθε τρόπο και με βοήθησαν στην ολοκλήρωση της παρούσας διπλωματικής εργασίας.

Ξεκινώντας θα ήθελα να ευχαριστήσω ιδιαίτερος τον καθηγητή και Πρόεδρο της σχολής Ηλεκτρολόγων Μηχανικών και Μηχανικών Η/Υ του Ε.Μ.Π. κύριο **Γεώργιο Κονταξή**, επειδή μου ανέθεσε αυτήν τη διπλωματική εργασία, καθώς και για την πολύτιμή βοήθεια και καθοδήγηση του κατά την διάρκεια εκπόνησης αυτής.

Επίσης θα ήθελα να ευχαριστήσω πολύ τον Διπλωματούχο Ηλεκτρολόγο Μηχανικό και Μηχανικό Η/Υ του Ε.Μ.Π. και υποψήφιο διδάκτορα της σχολής κύριο **Γεώργιο Τσεκούρα** για τον χρόνο που μου αφιέρωσε, τις πολύτιμες συμβουλές του σε όλα τα θέματα που προέκυψαν καθώς και για την ηθική του συμπαράσταση.

Επίσης θα ήθελα να ευχαριστήσω την Υποτομεάρχη του τμήματος πληροφορικής στη Διεύθυνση Διαχείρισης Ενέργειας και υπεύθυνη σχεδιασμού προγραμμάτων, κυρία **Ειρήνη Σκανδάλη** για την παροχή δεδομένων από τα αρχεία της Δ.Ε.Η., που συνέβαλαν ουσιαστικά στην περάτωση της παρούσας διπλωματικής εργασίας.

Ακόμα θα ήθελα να ευχαριστήσω την φίλη και συμφοιτήτρια μου **Μαρίνα Καμηλάκη** για τη βοήθεια που μου προσέφερε με την παροχή δεδομένων για την κατηγοριοποίηση των ημερών, καθώς και για την πολύτιμη συμπαράσταση που μου παρείχε κατά τη διάρκεια εκπόνησης αυτής της εργασίας.

Επιπρόσθετα θα ήθελα να ευχαριστήσω τους γονείς μου **Γιάννη** και **Γεωργία** και τον αδερφό μου **Βασίλη**, οι οποίοι από τη στιγμή έναρξης των σπουδών μου στο Ε.Μ.Π. μέχρι και τώρα, μου συμπαράσθηκαν το μέγιστο και με κάθε δυνατό τρόπο.

Τέλος θα ήθελα να ευχαριστήσω όλους τους **φίλους** μου που στάθηκαν δίπλα μου, καθόλη αυτήν την περίοδο εκπόνησης της παρούσας εργασίας.

Αθήνα – Ιούλιος 2003
Ανδρουτσόπουλος Σάββας

ΠΕΡΙΕΧΟΜΕΝΑ

ΠΡΟΛΟΓΟΣ

ΚΕΦΑΛΑΙΟ 1	Εισαγωγή – Γενικά για την Πρόβλεψη Φορτίου	1
1.1	Πρόβλεψη Φορτίου και Χρησιμότητά της στα Συστήματα Ηλεκτρικής Ενέργειας	1
1.2	<i>Παράγοντες μεταβολής του φορτίου</i>	3
1.3	<i>Βραχυπρόθεσμη Πρόβλεψη Φορτίου</i>	4
1.4	Απαραίτητα στοιχεία για την πρόβλεψη φορτίου	5
1.5	Συνιστώσες του φορτίου	7
ΚΕΦΑΛΑΙΟ 2	Μέθοδοι πρόβλεψης φορτίου	9
2.1	Κλασικές μέθοδοι πρόβλεψης φορτίου	9
2.2	Μέθοδοι εξομάλυνσης	9
2.2.1	Η μέθοδος των κινητών μέσων όρων (<i>MA</i>)	9
2.2.2	Η μέθοδος των γενικών γραμμικών κινητών μέσων όρων	10
2.2.3	Η μέθοδος εκθετικής εξομάλυνσης	10
2.3	Μέθοδος αποσύνθεσης	12
2.4	Μέθοδος πολλαπλής γραμμικής παλινδρομής	12
2.5	Οικονομετρική μέθοδος	13
2.6	Αυτοπαλινδρομικά μοντέλα κινουμένου μέσου όρου	14
2.7	Ολοκληρωμένα αυτοπαλινδρομικά μοντέλα κινουμένου μέσου όρου	15
2.8	Λοιπές παλινδρομικές και μη μέθοδοι	15
2.9	Μέθοδοι βραχυπρόθεσμης πρόβλεψης φορτίου με χρήση Ασαφούς Λογικής	16
2.9.1	Εισαγωγή στην ασαφή λογική	16
2.9.2	Εφαρμογή της ασαφούς λογικής στη βραχυπρόθεσμη πρόβλεψη φορτίου και ο συνδυασμός της με τα τεχνητά νευρωνικά δίκτυα	18
2.9.3	Μέθοδοι βραχυπρόθεσμης πρόβλεψης φορτίου με χρήση υβριδικών μοντέλων	19
2.10	Μέθοδοι βραχυπρόθεσμης πρόβλεψης φορτίου στηριζόμενες στα Τεχνητά Νευρωνικά Δίκτυα	21
2.11	Γενικά περί πολυεπίπεδων τεχνητών νευρωνικών δικτύων για τη βραχυπρόθεσμη πρόβλεψη φορτίου	22
2.12	Εφαρμογές πολυεπίπεδων τεχνητών νευρωνικών δικτύων βραχυπρόθεσμης πρόβλεψης φορτίου	24
2.13	Εφαρμογές ανατροφοδοτούμενων πολυεπίπεδων τεχνητών νευρωνικών δικτύων βραχυπρόθεσμης πρόβλεψης φορτίου	31
2.14	Λοιπές μορφές τεχνητών νευρωνικών δικτύων βραχυπρόθεσμης πρόβλεψης φορτίου	

		32
2.15	Σύγκριση τεχνητών νευρωνικών δικτύων για το πρόβλημα της βραχυπρόθεσμης πρόβλεψης φορτίου και συμπεράσματα	35
2.16	Επίδραση της θερμοκρασίας στη βραχυπρόθεσμη πρόβλεψη φορτίου με έμφαση στα τεχνητά νευρωνικά δίκτυα	36
2.17	Επίδραση του ανέμου στη βραχυπρόθεσμη πρόβλεψη φορτίου με έμφαση στα τεχνητά νευρωνικά δίκτυα	37
2.18	Επίδραση της διαχείρισης φορτίου στη βραχυπρόθεσμη πρόβλεψή του	38
2.19	Η σημασία του κόστους στη βραχυπρόθεσμη πρόβλεψη φορτίου	40
2.20	Οι τελευταίες δυνατότητες βελτιστοποίησης των μεθοδολογιών για τη βραχυπρόθεσμη πρόβλεψη φορτίου και γενικότερα συμπεράσματα	41
2.21	Πολύ Βραχυπρόθεσμη Πρόβλεψη Φορτίου	42
2.21	Μεσοπρόθεσμη και Μακροπρόθεσμη Πρόβλεψη Φορτίου	43
ΚΕΦΑΛΑΙΟ 3	Η μέθοδος της ανάστροφης διάδοσης-Back Propagation	45
3.1	Εισαγωγή - Το πολυεπίπεδο Perceptron	45
3.2	Ροή σημάτων	46
3.3	Ο αλγόριθμος Back-Propagation	47
3.4	Αντίστροφη προώθηση σημάτων σφάλματος διαμέσου δικτύου	50
3.5	Τα δύο στάδια του υπολογισμού του αλγορίθμου ανάστροφης διάδοσης σφάλματος (B-P)	52
3.6	Σιγμοειδής μη-γραμμικότητα	53
3.7	Προσθήκη του όρου ορμής { momentum }	54
3.8	Τρόποι εκπαίδευσης	55
3.8.1	Εκπαίδευση ανά πρότυπο	56
3.8.2	Εκπαίδευση ανά εποχή	56
3.9	Κριτήρια τερματισμού	57
3.10	Αριθμητικό παράδειγμα – Επικοινωνία σε μια γραμμή κόμβων	58
3.11	Περίληψη της διαδικασίας της ανάστροφης διάδοσης	59
ΚΕΦΑΛΑΙΟ 4	Εφαρμογές & σενάρια του αλγορίθμου της ανάστροφης διάδοσης	63
4.1	Γενικά για το πρόγραμμα εκπαίδευσης T.N.Δ	63
4.2	Διάβασμα αρχείων εισόδου	66
4.3	Κανονικοποίηση τιμών	73
4.4	Πραγματοποίηση διανυσμάτων εκπαίδευσης	74
4.5	Μετάβαση δεδομένων στο κυρίως πρόγραμμα	75
4.6	Παρουσίαση προτύπων εκπαίδευσης	75
4.7	Ο αλγόριθμος της ανάστροφης διάδοσης- Back Propagation	76
4.8	Σφάλματα	79
4.9	Μεταβολές των ρυθμών μάθησης και ορμής	80
4.9.1	Μεταβολές των παραμέτρων μάθησης	80
4.9.2	Αποτελέσματα για μεταβολή των παραμέτρων μάθησης - Γραφήματα	81
4.9.3	Παρατηρήσεις σχετικά με τη μεταβολή των παραμέτρων μάθησης	84

4.9.4	Μεταβολή των παραμέτρων ορμής	85
4.9.5	Αποτελέσματα για μεταβολή των παραμέτρων ορμής - Γραφήματα	85
4.9.6	Παρατηρήσεις σχετικά με τη μεταβολή των παραμέτρων ορμής	88
4.10	Σενάριο εκπαίδευσης 1	89
4.10.1	Αποτελέσματα σεναρίου 1 - Γραφήματα	93
4.10.2	Παρατηρήσεις – Σχόλια για το σενάριο 1	95
4.11	Σενάριο εκπαίδευσης 2	96
4.11.1	Αποτελέσματα σεναρίου 2 - Γραφήματα	97
4.11.2	Παρατηρήσεις – Σχόλια για το σενάριο 2	98
4.12	Σενάριο εκπαίδευσης 3	98
4.12.1	Αποτελέσματα σεναρίου 3 - Γραφήματα	101
4.12.2	Παρατηρήσεις – Σχόλια για το σενάριο 3	103
4.13.	Κατηγοριοποίηση των ημερών	103
4.13.1	Περιγραφή του προγράμματος ενεργοποίησης της κατηγοριοποίησης	105
4.14	Σενάριο εκπαίδευσης 4	108
4.14.1	Αποτελέσματα σεναρίου 4 - Γραφήματα	111
4.14.2	Παρατηρήσεις – Σχόλια για το σενάριο 4	113
4.15	Σενάριο εκπαίδευσης 5	114
4.15.1	Αποτελέσματα σεναρίου 5 - Γραφήματα	115
4.15.2	Παρατηρήσεις – Σχόλια για το σενάριο 5	117
4.16	Σενάριο εκπαίδευσης 6	118
4.16.1	Αποτελέσματα σεναρίου 6 - Γραφήματα	119
4.16.2	Παρατηρήσεις – Σχόλια για το σενάριο 6	125
4.17	Τρισδιάστατο Γράφημα	118
		127
ΚΕΦΑΛΑΙΟ 5	Τελικά συμπεράσματα και προτάσεις για περαιτέρω έρευνα	129
5.1	Περιγραφή της διπλωματικής εργασίας	129
5.2	Τελικά συμπεράσματα και προτάσεις	130
ΠΑΡΑΡΤΗΜΑ I	Γραφήματα	133
ΠΙ.1	Γραφήματα εκτελέσεων του προγράμματος εκπαίδευσης Τ.Ν.Δ. για μεταβολές των παραμέτρων μάθησης και ορμής	135
ΠΙ.2	Γραφήματα εκτελέσεων του προγράμματος εκπαίδευσης Τ.Ν.Δ. για τα σενάρια εκπαίδευσης 1 - 6	252
ΠΑΡΑΡΤΗΜΑ II	Βιβλιογραφία	167

Κεφάλαιο 1

Εισαγωγή - Γενικά για την Πρόβλεψη Φορτίου

➤ 1.1 Πρόβλεψη Φορτίου και χρησιμότητά της στα Συστήματα Ηλεκτρικής Ενέργειας

Ένα σύστημα παραγωγής , μεταφοράς και διανομής ηλεκτρικής ενέργειας οφείλει να παρέχει στον καταναλωτή ηλεκτρική ενέργεια ανά πάσα στιγμή, σε οποιοδήποτε σημείο ζήτησης και με χαρακτηριστικά τέτοια, που θα καλύπτουν τις ζητούμενες απαιτήσεις του όσον αφορά την ποιότητα (επίπεδα τάσης, συχνότητας, συντελεστή παραμόρφωσης αρμονικών κτλ.), αλλά και την ποσότητα, δηλαδή πρέπει πάντα να υπάρχει επάρκεια ισχύος και ενέργειας.

Για να είναι δυνατή η πλήρης κάλυψη της ζήτησης στο άμεσο ή στο απώτερο μέλλον, προφανώς απαιτείται να αναπτυχθούν οι κατάλληλες μεθοδολογίες πρόβλεψης του φορτίου. Οι τελευταίες διακρίνονται στις παρακάτω 4 κατηγορίες :

- ❑ πολύ βραχυπρόθεσμη, που αφορά την πρόβλεψη της ζήτησης του φορτίου για τα επόμενα 30 λεπτά ως μία ώρα, με βήματα μερικών δευτερολέπτων ως λεπτού,
- ❑ βραχυπρόθεσμη, όπου γίνεται η πρόβλεψη του φορτίου για τα επόμενο εικοσιτετράωρο ως μία εβδομάδα, με χρονικό βήμα της μισής ή της μίας ώρας,

- ❑ μεσοπρόθεσμη, που σχετίζεται με την πρόβλεψη που γίνεται για ένα έτος, με χρονικό βήμα της μιας εβδομάδας,
- ❑ μακροπρόθεσμη, όπου η αντίστοιχη πρόβλεψη γίνεται με έναν χρονικό ορίζοντα των 10-20 ετών, με χρονικό βήμα ενός έτους.

Η πολύ βραχυπρόθεσμη πρόβλεψη φορτίου έχει ως στόχο την κάλυψη των αναγκών των συστημάτων αυτομάτου ελέγχου των γεννητριών, τα οποία χρειάζονται όσο το δυνατό καλύτερη πληροφόρηση για τις μελλοντικές αλλαγές φόρτισής τους, ώστε να προχωρήσουν στις αναγκαίες μεταβολές της παροχής καυσίμου και ρύθμισης των επιπέδων τάσης εξόδου τους. Με αυτόν τον τρόπο καθίσταται δυνατός ο έλεγχος της σχέσης συχνότητας-φορτίου που συνιστά καθοριστικό παράγοντα για την ασφάλεια του συστήματος. Επίσης σήμερα μπορεί να διαδραματίσει σημαντικό ρόλο στη διαμόρφωση των τιμών ηλεκτρικής ενέργειας στα αντίστοιχα χρηματιστήρια ηλεκτρικής ενέργειας στην Καλιφόρνια.

Η βραχυπρόθεσμη πρόβλεψη έχει ως στόχο τη ρύθμιση των βασικών επιπέδων λειτουργίας των μονάδων, όπως επίσης παίζει πολύ σημαντικό ρόλο στη διαμόρφωση της σειράς ένταξης τους βάσει των κριτηρίων της οικονομικής κατανομής, αλλά και λαμβάνοντας υπόψη τα αντίστοιχα προγράμματα συντήρησης. Ουσιαστικά είναι υπεύθυνη για την ενεργειακή διαχείριση του συστήματος.

Η μεσοπρόθεσμη πρόβλεψη αφορά κυρίως τη ρύθμιση των προγραμμάτων συντήρησης των μονάδων παραγωγής και αξιοποίησης των διαθέσιμων υδάτινων και υπόλοιπων πόρων στις καταλληλότερες χρονικές περιόδους του έτους. Για παράδειγμα στο ηπειρωτικό ελληνικό σύστημα οι θερμικές μονάδες υφίστανται κατά κύριο λόγο τη μία ή τις δύο ετήσιες συντηρήσεις τους συνήθως σε μη καλοκαιρινούς μήνες, ώστε να καλυφθεί η ανάγκη της ετήσιας αιχμής που εμφανίζεται κατά το καλοκαίρι και οφείλεται κατά κύριο λόγο στη χρήση κλιματιστικών, όπως μας δίνεται από τις αντίστοιχες προβλέψεις.

Η μακροπρόθεσμη πρόβλεψη φορτίου αφορά κυρίως την ετήσια αιχμή φορτίου, ή και τις εποχιακές, καθώς και τη συνολική ετήσια ενέργεια για τα επόμενα 10 με 20 χρόνια, ώστε να είναι δυνατός ο κατάλληλος σχεδιασμός και η κατασκευή του συστήματος ηλεκτρικής ενέργειας. Για παράδειγμα αναφέρουμε την δημιουργία νέων μονάδων παραγωγής, την κατασκευή νέων γραμμών μεταφοράς ή την αναδιαμόρφωση του δικτύου διανομής μίας μεγάλης περιοχής, π.χ. η ανύψωση του δικτύου μέσης τάσης στα 20 kV από τα 15 kV για την περιοχή της Θεσσαλονίκης και η κατασκευή του υδροηλεκτρικού σταθμού στο Θησαυρό. Όλα τα παραπάνω είναι προφανώς αρκετά χρονοβόρα και δαπανηρά και για αυτό κρίνεται απαραίτητη μια σωστή μακροπρόθεσμη πρόβλεψη, ώστε να αξιοποιηθούν στο έπακρο.

Στη συνέχεια θα μιλήσουμε αναλυτικότερα για τη βραχυπρόθεσμη η οποία μαζί με τη μακροπρόθεσμη πρόβλεψη φορτίου, έχουν εξελιχθεί περισσότερο εν συγκρίσει με τα άλλα είδη και για τις οποίες έχουν αναπτυχθεί ιδιαίτερα ενδιαφέρουσες εφαρμογές με νευρωνικά δίκτυα. Πριν από αυτό όμως θα γίνει μία σύντομη αναφορά στους παράγοντες που επηρεάζουν τα φορτία.

➤ 1.2 Παράγοντες μεταβολής του φορτίου

Η κατανάλωση ενέργειας μιας ορισμένης περιοχής επηρεάζεται από ένα σημαντικό πλήθος διαφορετικών και σε αρκετές περιπτώσεις απροσδιόριστων παραγόντων, οι οποίοι είναι κυρίως αυτοί που αναφέρονται παρακάτω :

1. Οικονομικοί, δηλαδή η οικονομική κατάσταση, το οικονομικό περιβάλλον και οι ρυθμοί ανάπτυξης κάποιας περιοχής επηρεάζουν καθοριστικά τη ζήτηση ηλεκτρικής ενέργειας. Αυτό συμβαίνει, διότι αλλάζει το βιοτικό επίπεδο των καταναλωτών, εξελίσσεται δημογραφικά ο πληθυσμός της χώρας, γίνεται η αυτοματοματοποίηση των δραστηριοτήτων, και η βελτιστοποίηση της λειτουργίας των συσκευών και των μηχανών επιδρώντας σε μεσοπρόθεσμη και μακροπρόθεσμη βάση. Η οικονομική ανάπτυξη μιας περιοχής έχει ως άμεσο αποτέλεσμα να μεταβάλλονται :

- ❑ η σύνθεση των καταναλωτών, καθώς έχουμε αύξηση του αστικού πληθυσμού και μείωση του γεωργικού,
- ❑ η ποσότητα της καταναλισκόμενης ηλεκτρικής ενέργειας, καθώς από τη μία μεριά βελτιώνεται το βιοτικό επίπεδο των καταναλωτών με αποτέλεσμα να σημειώνεται αύξηση και από την άλλη υπάρχουν τα φαινόμενα κορεσμού και η χρησιμοποίηση λιγότερο ενεργοβόρων συσκευών,
- ❑ οι μέθοδοι καλλιέργειας, άρδευσης και αυτοματοποίησης της βιομηχανίας.

2. Εποχιακοί παράγοντες , καθώς τα φορτία παρουσιάζουν έντονες διακυμάνσεις κατά τη διάρκεια τόσο του εικοσιτετραώρου, όσο και του έτους. Οι αλλαγές αυτές έχουν να κάνουν κυρίως με τις μεταβολές των καιρικών συνθηκών και των ανθρώπινων δραστηριοτήτων. Ο τελευταίος αυτός παράγοντας επηρεάζεται σε σημαντικό βαθμό από τη γεωγραφική περιοχή, καθώς παραδείγματος χάρη οι αστικές περιοχές κατά τη διάρκεια των καλοκαιρινών διακοπών παρουσιάζουν μείωση της ζήτησης, ενώ αντίστοιχα οι παραθαλάσσιες αύξηση.

3. Μετεωρολογικοί, διότι η θερμοκρασία κατά κύριο λόγο και , σε αρκετά μικρότερο ποσοστό , η υγρασία, η ανεμόπτωση και η ηλιοφάνεια επιδρούν στην κατανάλωση ηλεκτρικής ενέργειας του πληθυσμού μιας περιοχής.

4. Τυχαίοι παράγοντες , όπως είναι ορισμένα τυχαία γεγονότα -παραδείγματος χάρη μεγάλες απεργίες, εκλογές, μετάδοση κάποιων τηλεοπτικών προγραμμάτων κτλ.

5. Αβέβαιοι, δηλαδή παράγοντες που είναι πολύ δύσκολο να προβλεφθούν και επηρεάζουν μακροπρόθεσμα την κατανάλωση, όπως οι προοπτικές ανάπτυξης μίας περιοχής, ο ρυθμός αύξησης του πληθυσμού και η πορεία της οικονομίας.

Ανάλογα με το είδος της πρόβλεψης η δράση ορισμένων από αυτούς περιορίζεται, όπως δεν είναι δυνατό να επιδράσει ο ετήσιος ρυθμός οικονομικής ανάπτυξης της χώρας στην πρόβλεψη του φορτίου για την επόμενη μισή ώρα (πολύ βραχυπρόθεσμη πρόβλεψη) . Αντίστοιχα η επίδραση των καιρικών φαινομένων στη μακροπρόθεσμη πρόβλεψη φορτίου δεν διαδραματίζει σημαντικό ρόλο, αν και στο τελευταίο μπορεί να αναφερθεί, ως απόδειξη του αντιθέτου, η μετατόπιση της αιχμής του φορτίου στο καλοκαίρι για το ελληνικό σύστημα λόγω των κλιματιστικών. Βέβαια σ' αυτήν την περίπτωση πρέπει να τονισθεί ότι έγινε η διείσδυση μίας νέας συσκευής τελικής χρήσης, οπότε θα πρέπει ο μελετητής, πριν την εφαρμογή του οποιοδήποτε μοντέλου, να προδιαγράψει την αντίστοιχη επίδραση στο φορτίο, αλλιώς θα υπάρξει σε κάθε είδους πρόβλεψη σημαντική απόκλιση.

➤ 1.3 Βραχυπρόθεσμη Πρόβλεψη Φορτίου

Η βραχυπρόθεσμη πρόβλεψη φορτίου αφορά τον προσδιορισμό της ζήτησης ηλεκτρικής ισχύος για τις επόμενες 24 ως 168 ώρες , ή αλλιώς για την επόμενη μέρα ως μία εβδομάδα , ανά μία ώρα ή πιο σπάνια ανά μισάωρο. Επηρεάζεται σε σημαντικό βαθμό, όπως θα αποδειχθεί και από την ανάλυση της αντίστοιχης αρθρογραφίας, από την ετήσια αύξηση του φορτίου, από το φορτίο των προηγούμενων ημερών και από τα αντίστοιχα επίπεδα θερμοκρασίας και λοιπών καιρικών στοιχείων, όπως υγρασίας, ανέμου, ηλιοφάνειας κτλ.. Συνήθως όμως λαμβάνεται υπόψη μόνο η επίδραση της θερμοκρασίας , της οποίας είναι απαραίτητη η αντίστοιχη πρόβλεψη για τις μέρες που επιθυμούμε να προσδιορίσουμε το φορτίο -αν όχι σε ωριαίο επίπεδο, τουλάχιστον σε μέγιστη και ελάχιστη ημερήσια θερμοκρασία με τις αντίστοιχες διασπορές τους.

Η βραχυπρόθεσμη πρόβλεψη φορτίου διαδραματίζει καθοριστικό ρόλο στην ομαλή λειτουργία του συστήματος και στον προσδιορισμό του επιπέδου ασφαλείας του. Έχουν προσδιορισθεί, μέσω κατάλληλων μοντέλων, οι επιδράσεις της μη ακριβούς πρόβλεψης της θερμοκρασίας [A1] ή της απότομης βηματικής αλλαγής του φορτίου -χρήση μεθόδου διχοτόμου-[A2] που μπορεί να προσομοιώνει μία απότομη απόρριψη φορτίου από μία μεγάλη βιομηχανική μονάδα λόγω σφάλματος. Επίσης επηρεάζει σημαντικά τη διαμόρφωση της θέσης και του μεγέθους των γεννητριών εκτός των εταιρειών -NUG ή Non utility generation-, οι οποίες ανήκουν σε ιδιοπαραγωγούς, νοσοκομεία, κτλ., σε συνδυασμό πάντοτε με τη διαστασιολόγηση του εκάστοτε δικτύου διανομής και του ρυθμού μη προγραμματισμένων διακοπών αυτών των γεννητριών [A3].

Ακόμη είναι απαραίτητη για τη ρύθμιση και τον έλεγχο συσκευών αποθήκευσης ενέργειας -θερμοσυσσωρευτών, θερμαντήρων νερού, κλιματιστικών διατάξεων με δεξαμενές πάγου, όταν συνδυάζεται με τα αντίστοιχα ηλεκτρικά τιμολόγια προς μείωση του τελικού κόστους [A4], [A5].

Δηλαδή για να είναι δυνατή η διαχείριση ζήτησης φορτίου και ο έλεγχος της εφαρμογής των διαφόρων μέτρων εξοικονόμησης ηλεκτρικής ενέργειας , ιδιαίτερα στα πλαίσια της μελλοντικής απελευθέρωσης της αγοράς ενέργειας , είναι αναγκαία η ύπαρξη ενός πλήθους δεδομένων που θα

αφορούν τις διάφορες κατηγορίες καταναλωτών -βαθμός κορεσμού ηλεκτρικών συσκευών, εισόδημα, σύστημα θέρμανσης, κλιματισμός, απόδοση κινητήρων κτλ.-, ώστε να είναι δυνατός ο προσδιορισμός της χρονολογικής καμπύλης φορτίου και μέσω τεχνικοοικονομικής μελέτης να βρίσκονται τα εφαρμόσιμα μέτρα [E1].

Η ατυχής πρόβλεψη φορτίου έχει δυσμενείς επιπτώσεις, τόσο ως προς την ορθή αξιοποίηση των δυνατοτήτων του δικτύου μεταφοράς ισχύος, όσο και ως προς την ευστάθεια του συστήματος ηλεκτρικής ενέργειας [A6]. Στην περίπτωση που η πρόβλεψη φορτίου υπολείπεται του πραγματικού, οδηγούμαστε σε αγορά ακριβών MW αιχμής από άλλους παραγωγούς, από το εξωτερικό ή το χρηματιστήριο ηλεκτρικής ενέργειας. Αν είναι πιο μεγάλη της πραγματικής, τότε η επιχείρηση έχει μεγαλύτερο κόστος εκκίνησης και σταθερό κόστος λειτουργίας λόγω της ένταξης μεγαλύτερων μονάδων. Αν το μέσο απόλυτο σφάλμα -που κυμαίνεται ανάμεσα σε 3% και 5%- μειωθεί κατά 1%, τότε υπάρχει μείωση του κόστους παραγωγής κατά 0,1% με 0,3% τουλάχιστον. Ιδιαίτερα δυσμενέστερα αποτελέσματα έχουμε στην περίπτωση της υποεκτίμησης του φορτίου. Γι' αυτό απαιτείται ακριβέστερη βραχυπρόθεσμη πρόβλεψη φορτίου [A7]. Παράλληλα η αβεβαιότητα πρόβλεψης φορτίου επηρεάζει τα αποτελέσματα των μοντέλων βραχυπρόθεσμης κατανομής της παραγωγής των γεννητριών και της απαιτούμενης εφεδρείας, μέσω πιθανοτικών μοντέλων [A8] ή μέσω του κριτηρίου της επιλογής εφεδρείας ισχύος, ίσης ή μεγαλύτερης από τη μεγαλύτερη γεννήτρια ή από ένα όριο φορτίου [Δ1].

Εξαιτίας των παραπάνω, πέρα από τις διάφορες μεθοδολογίες πρόβλεψης έχουν αναπτυχθεί κατάλληλοι αλγόριθμοι με μαθηματικά μοντέλα για τη βελτίωση της απόδοσής τους, ώστε με τη λήψη περισσότερων μετρήσεων σε κατάλληλα προεπιλεγμένα σημεία του δικτύου, να γίνεται διόρθωση των μετρητικών σφαλμάτων των ωριαίων φορτίων [A9], [Δ2].

Συνεπώς λόγω της μεγάλης σπουδαιότητας της βραχυπρόθεσμης πρόβλεψης φορτίου έχουν αναπτυχθεί δύο βασικές κατηγορίες μεθοδολογιών. Η μια κατηγορία αποτελείται από τα κλασικά μοντέλα, όπως είναι οι μέθοδοι εξομάλυνσης, αποσύνθεσης, γραμμικών παλινδρομήσεων και αυτοπαλινδρομήσεων κινητού μέσου όρου. Η δεύτερη κατηγορία συνίσταται από νεότερα μοντέλα, όπως τα τεχνητά νευρωνικά δίκτυα, την ασαφή λογική καθώς και τους υβριδικούς σχηματισμούς τους.

Σε κάθε περίπτωση όμως τα αρχικά δεδομένα μας είναι τα ωριαία φορτία προηγούμενων ημερών, τα ημερολογιακά χαρακτηριστικά και τα διάφορα στοιχεία που αντιπροσωπεύουν τις καιρικές συνθήκες, με ιδιαίτερη έμφαση κυρίως στη θερμοκρασία σε ποσοστό πάνω από το 90% της υπάρχουσας αρθρογραφίας.

Στις επόμενες παραγράφους, καθώς και στο επόμενο κεφάλαιο θα αναφερθούμε αναλυτικότερα σ' αυτά τα δεδομένα, στην κατηγοριοποίηση των επιμέρους συνιστωσών του φορτίου, στις κλασικές και στις νεότερες μεθοδολογίες και στην τελική επίδραση της ακρίβειάς τους στο κόστος των ηλεκτρικών επιχειρήσεων.

➤ 1.4 Απαραίτητα στοιχεία για την πρόβλεψη φορτίου

Τα βασικότερα στοιχεία εισόδου στη βραχυπρόθεσμη πρόβλεψη φορτίου είναι φυσικά τα ωριαία φορτία των προηγούμενων ωρών ή ημερών, φθάνοντας σε χρονικό βάθος μίας μέρας ως μίας εβδομάδας. Συνηθέστερα λαμβάνονται υπόψη για την πρόβλεψη του ωριαίου φορτίου h τα

ωριαία φορτία των προηγούμενων τριών ωρών, τα αντίστοιχα φορτία της προηγούμενης μέρας και της αντίστοιχης μέρας της περασμένης εβδομάδας.

Βέβαια σε κάθε περίπτωση σημαντικό ρόλο διαδραματίζει και το είδος της μέρας, καθώς π.χ. άλλη συμπεριφορά παρουσιάζει η Τρίτη σε σχέση με την Κυριακή. Αυτό μπορεί να αντιμετωπιστεί είτε με διαφορετικές εφαρμογές του βασικού μοντέλου, οπότε για το κάθε είδος μέρας έχουμε και διαφορετικές τιμές συντελεστών, είτε χρησιμοποιώντας αντίστοιχες εισόδους, όπως με κατάλληλο ακέραιο αριθμό ή με κατάλληλο δυαδικό αριθμό. Για παράδειγμα η μέρα Τρίτη αντιστοιχεί στο 2 και η Κυριακή στο 7, ή αντίστοιχα έχουμε για Τρίτη και Κυριακή 0100000 και 0000001 [Δ2].

Όσον αφορά το πρόβλημα της εποχικότητας των φορτίων, δηλαδή τη διαφορετική συμπεριφορά που παρουσιάζουν τα φορτία, π.χ. την άνοιξη από ότι το καλοκαίρι, αντιμετωπίζεται πάλι είτε με διαφορετικές εφαρμογές μοντέλων, είτε με τη χρήση περιοδικών συναρτήσεων, όπως του ζεύγους $(\cos(2\pi/365), \sin(2\pi/365))$ [Δ2], όπου με μονοσήμαντο τρόπο χαρακτηρίζεται μία μέρα ως προς τη θέση της μέσα στο έτος.

Τελευταίο σημαντικό στοιχείο είναι τα καιρικά στοιχεία, τονίζοντας όμως σ' αυτήν την περίπτωση την έντονη επίδραση του κλίματος σε μία περιοχή. Δηλαδή για τον ελλαδικό χώρο έχει αποδειχθεί από τις εφαρμογές [Δ2] ότι η θερμοκρασία καλύπτει σχεδόν πλήρως τις ανάγκες μας για την πρόβλεψη της επίδρασης του καιρού στο φορτίο. Αντίθετα στο Χονγκ-Κονγκ λόγω του υψηλού ποσοστού βροχοπτώσεων και των υψηλών τιμών υγρασίας δεν περιορίζονται στη θερμοκρασία, αλλά ως δεδομένο χρειάζεται και η σχετική υγρασία. Αντίστοιχες παρατηρήσεις ισχύουν και για τον άνεμο, την ηλιοφάνεια κτλ.[Α13].

Βέβαια ο κάθε μελετητής λαμβάνει υπόψη του τα διάφορα στοιχεία με διαφορετικό τρόπο. Αν δεν χρησιμοποιηθεί απλά η θερμοκρασία, μπορεί να ορισθεί ένα μοντέλο καιρού, όπου παραδείγματος χάρη στο [Α10] προσδιορίζονται οι αντίστοιχοι παράγοντες $F_d^c(T)$, $F_d^h(T)$, όπου $T_{lc}(d)$, $T_{2c}(d)$ τα κατώφλια θερμοκρασίας για τις ψυχρές ώρες και $T_{lh}(d)$, $T_{2h}(d)$ τα αντίστοιχα για τις ζεστές ώρες ανάλογα με το είδος της d μέρας της εβδομάδας:

$$F_d^c(T) = \begin{cases} 0 & , \quad T < T_{lc}(d) \\ (T - T_{lc}(d))^2 / 2(T_{2c}(d) - T_{lc}(d)), & T_{lc}(d) \leq T < T_{2c}(d) \\ T - (T_{lc}(d) + T_{2c}(d)) / 2 & , \quad T \geq T_{2c}(d) \end{cases} \quad (1.4-1)$$

$$F_d^h(T) = \begin{cases} 0 & , \quad T > T_{lh}(d) \\ (T - T_{lh}(d))^2 / 2(T_{lh}(d) - T_{2h}(d)), & T_{2h}(d) < T \leq T_{lh}(d) \\ (T_{lh}(d) + T_{2h}(d)) / 2 - T & , \quad T \leq T_{2h}(d) \end{cases} \quad (1.4-2)$$

Μία ενδιαφέρουσα τακτική, όταν θέλουμε να λάβουμε και άλλα στοιχεία πέρα από τη θερμοκρασία, είναι ο ορισμός της «ισοδύναμης» ή αλλιώς της «ενεργού» θερμοκρασίας, την οποία και προσδιορίζουμε μέσω εμπειρικών σχέσεων. Ενδεικτικά αναφέρουμε ότι η ισοδύναμη θερμοκρασία T_{eq} συνδέεται με τη θερμοκρασία T , τη σχετική υγρασία RH και την ταχύτητα του ανέμου V κατά τις προδιαγραφές που σχετίζονται με την ανθρώπινη άνεση, σύμφωνα με την ακόλουθη σχέση [Α11]:

$$T_{eq} = \left\{ \begin{array}{ll} \frac{T - 0,55 \cdot (1,0 - RH) \cdot (T - 58) - 15,95}{0,725} & , \quad T \geq 66^{\circ}F \\ T & , \quad 65^{\circ}F < T < 66^{\circ}F \\ 91,4 - \frac{(10,45 - 10,0\sqrt{V} - V) \cdot (91,4 - T)}{22,952} & , \quad T \leq 65^{\circ}F \end{array} \right\} \quad (1.4-3)$$

όπου T και T_{eq} σε $^{\circ}F$, RH σε δεκαδικό, V σε m/sec.

Άλλοι μελετητές [A12] έχουν ορίσει και απλούστερες εκφράσεις για την «ενεργό» θερμοκρασία T_{eff} που χρησιμοποιείται ως είσοδος στο δίκτυο, η οποία προκύπτει -με χρήση των άνω συμβόλων, όπου a εμπειρικός συντελεστής- από τις ακόλουθες σχέσεις:

$$T_{eff} = \left\{ \begin{array}{ll} T + a \cdot RH & , \quad T > 75^{\circ}F \\ T & , \quad 65^{\circ}F \leq T \leq 75^{\circ}F \\ T - \frac{V \cdot (65^{\circ} - T)}{100} & , \quad T < 65^{\circ}F \end{array} \right\} \quad (1.4-4)$$

Αδιαφιλονίκητο συμπέρασμα από όλα τα προαναφερθέντα, είναι ότι η κάθε κατηγορία δεδομένων μπορεί να αναπαρασταθεί με πολλούς και διαφορετικούς μεταξύ τους τρόπους. Οι τελικές επιλογές βέβαια γίνονται από τον εκάστοτε μελετητή και επηρεάζονται σε σημαντικό βαθμό από τη μορφή του μοντέλου, ενώ με τη σειρά τους επιδρούν στη διαμόρφωση των τελικών αποτελεσμάτων.

➤ 1.5 Συνιστώσες του φορτίου

Πριν από τον προσδιορισμό του μοντέλου σημαντικό ρόλο διαδραματίζει και ο διαχωρισμός του φορτίου σε επιμέρους συνιστώσες. Αυτό εξαρτάται φυσικά από την επιλογή του μοντέλου, καθώς και από τους παράγοντες που επηρεάζουν τα αποτελέσματά που δίνει. Γι' αυτό και υπάρχουν διάφοροι τρόποι επιμερισμού του φορτίου, οι κυριότεροι εκ των οποίων αναφέρονται παρακάτω :

- i. Σε τέσσερις συνιστώσες : 1) το ντετερμινιστικό γνωστό φορτίο , 2) το ανεξάρτητο από τον καιρό , 3) το εξαρτημένο από τον καιρό και 4) το υπόλοιπο που αποτελεί τον «θόρυβο» [A11]. Αντίστοιχα για το [A13] τα τέσσερα μέρη είναι η συνιστώσα καιρού, η συνιστώσα κανονικού φορτίου, το φορτίο ειδικών περιστάσεων και το τυχαίο φορτίο.
- ii. Σε τρεις συνιστώσες : στην ετήσια, στην εποχιακή και στην ημερήσια συνιστώσα του [A14], λόγω χρήσης παλινδρομικού μοντέλου. Αντίστοιχα στο [A15] τα τρία μέρη είναι το ετήσιο, το ημερήσιο και το στοχαστικό ενώ στο [A16] είναι το ημερήσιο, το εβδομαδιαίο και το κυκλικό.

- iii. Σε δύο συνιστώσες : τη βασική μορφή του φορτίου και την αντίστοιχη εξαρτώμενη από τον καιρό συνιστώσα, για κάθε συσκευή ηλεκτρικής ενέργειας χωριστά [A17] ή για τη συνολική ζήτηση του συστήματος [A18].

Ουσιαστικά η ανάλυση του φορτίου σε συνιστώσες γίνεται βάσει, είτε των κριτηρίων του χρόνου, σύμφωνα με την περίπτωση ii , είτε των παραγόντων εξάρτησης του φορτίου, όπως στις περιπτώσεις i και iii.

Στο επόμενο κεφάλαιο θα αναφερθούμε πιο αναλυτικά, τόσο στις κλασικές, όσο και στις νεότερες μεθόδους βραχυπρόθεσμης πρόβλεψης φορτίου.

Κεφάλαιο 2

Μέθοδοι πρόβλεψης φορτίου

➤ 2.1 Κλασικές μέθοδοι πρόβλεψης φορτίου

Στην αρχή αυτού του κεφαλαίου και πριν αναφερθούμε στις νεότερες μεθόδους πρόβλεψης φορτίου, κρίνουμε σκόπιμο να παραθέσουμε, εν συντομία, παλαιότερες κλασικές μεθόδους, εστιάζοντας παράλληλα στα πλεονεκτήματα και τα μειονεκτήματά τους. Οι διάφορες μεθοδολογίες διακρίνονται σε δύο κυρίως κατηγορίες, οι οποίες δεν είναι άλλες από τις μεθόδους εξομάλυνσης και από τις μεθόδους αποσύνθεσης. Παρακάτω θα αναφερθούμε και στις δύο αυτές κατηγορίες.

➤ 2.2 Μέθοδοι εξομάλυνσης

Για τη βραχυπρόθεσμη πρόβλεψη φορτίου, οι μέθοδοι εξομάλυνσης αποτελούν μια από τις απλούστερες κατηγορίες μεθόδων. Σε αυτές ανήκουν οι εξής :

➤ 2.2.1 Η μέθοδος των κινητών μέσων όρων (MA)

Το φορτίο σε αυτή την περίπτωση $\hat{P}_m(t+k|t)$ που αναφέρεται στην ώρα $t+k$, ενώ βρίσκεται στη χρονική στιγμή t , προσδιορίζεται από την παρακάτω σχέση :

$$\hat{P}_m(t+k|t) = \frac{1}{m} \cdot \sum_{i=0}^{m-1} P(t-i) \text{ με } k = 1, \dots, m \quad (2.2.1-1)$$

Σε ορισμένες περιπτώσεις [A19], μελετώνται τα ωριαία φορτία, είτε του προηγούμενου εικοσιτετραώρου, είτε του εικοσιτετραώρου προ μίας εβδομάδας, οπότε διαμορφώνονται αντίστοιχα τα μοντέλα «Ημέρα 1» και «Ημέρα 7». Συγκρίνοντας [A19] τα δύο παραπάνω αυτά μοντέλα, έχει βρεθεί ότι η μέση τιμή του απόλυτου σφάλματος είναι μικρότερη για το μοντέλο «Ημέρα 7» για τη Δευτέρα και την Κυριακή, ενώ για τις υπόλοιπες ημέρες είναι μικρότερη για το μοντέλο «Ημέρα 1». Επιπρόσθετα αν χρησιμοποιηθεί ένας συνδυασμός των δύο μοντέλων, τότε υπεισέρχεται ένα σφάλμα περίπου της τάξης του 5%, το οποίο είναι αρκετά υψηλό.

➤ 2.2.2 Η μέθοδος των γενικών γραμμικών κινητών μέσων όρων

Η μέθοδος αυτή περιγράφεται από τις ακόλουθες σχέσεις [B1] :

$$S'_t = \frac{X_t + X_{t-1} + \dots + X_{t-N+1}}{N} \quad (2.2.2-1)$$

$$S''_t = \frac{S'_t + S'_{t-1} + \dots + S'_{t-N+1}}{N} \quad (2.2.2-2)$$

$$\alpha_t = S'_t + (S'_t - S''_t) \quad (2.2.2-3)$$

$$b_t = \frac{2}{N-1} (S'_t - S''_t) \quad (2.2.2-4)$$

$$\hat{X}_t(m) = \alpha_t + b_t \cdot m \quad (2.2.2-5)$$

δηλαδή ουσιαστικά οι μέσοι όροι αποκτούν κάποια βάρη, τα οποία παίζουν σημαντικό ρόλο στο τελικό αποτέλεσμα .

➤ 2.2.3 Η μέθοδος εκθετικής εξομάλυνσης

Η μέθοδος αυτή χωρίζεται σε διάφορες επιμέρους κατηγορίες, όπως τα μη εποχιακά μοντέλα, τα προσθετικά εποχιακά και τα πολλαπλασιαστικά με υπομέρους κλάσεις του σταθερού επιπέδου, της γραμμικής τάσης και της μη γραμμικής τάσης.

Σαν ενδεικτικό παράδειγμα της μεθόδου αναφέρουμε το άρθρο [A16], όπου ο εκάστοτε καταναλωτής με τη βραχυπρόθεσμη πρόβλεψη φορτίου, ρυθμίζει τις ώρες ζήτησης ισχύος για θερμικούς ή ψυκτικούς σκοπούς, χρησιμοποιώντας αντίστοιχα θερμοσυσσωρευτές ή κεντρικές κλιματιστικές μονάδες με παγοκολώνες. Συγκεκριμένα, η κλιματιστική συσκευή μοντελοποιείται κατάλληλα, ώστε κατά τη διάρκεια ενός εικοσιτετραώρου να κλείνει ο κύκλος ψύξης. Η απαραίτητη, λοιπόν, για την επιτυχή λειτουργία της συσκευής, πρόβλεψη φορτίου πραγματοποιείται μέσω της ακόλουθης πολλαπλασιαστικής εποχιακής τεχνικής εξομάλυνσης, λαμβάνοντας υπόψη την καθημερινή κυκλικότητα της θερμοκρασίας και την άνοδο ή την κάθοδο της μέσης ημερήσιας θερμοκρασίας.

$$e(t) = X(t) - T_1(t-1) \quad (2.2.3-1) \quad S(t) = S(t-1) + \frac{h_1 \cdot e(t)}{I(t-24+m)} \quad (2.2.3-2)$$

$$I(t) = I(t-d) + \frac{h_3 \cdot e(t)}{S(t)} \quad (2.2.3-4) \quad T_m(t) = S(t) \cdot I(t-24+m) \quad (2.2.3-4)$$

όπου X_t είναι η πραγματική τιμή της χρονοσειράς για τη χρονική περίοδο t , e_t το σφάλμα πρόβλεψης για την αντίστοιχη περίοδο, S_t το επίπεδο της χρονοσειράς στο τέλος της περιόδου t , T_t η τάση της χρονοσειράς στο τέλος της αντίστοιχης περιόδου t , I_t ο εποχιακός παράγοντας για τη χρονική περίοδο t , h_1 ο συντελεστής εξομάλυνσης για το επίπεδο της χρονοσειράς και h_3 ο συντελεστής εξομάλυνσης για τον εποχιακό παράγοντα.

Επίσης με βάση τις σχέσεις :

$$T_{low}(t) = S'(t) \cdot [1.0 - \beta \cdot [1.0 - I(t-24+m_{low})]] \quad (2.2.3-5)$$

$$T_{high}(t) = S'(t) \cdot [1.0 - \beta \cdot [1.0 - I(t-24+m_{high})]] \quad (2.2.3-6)$$

προσδιορίζουμε μέσω των δεικτών m_{low} και m_{high} την ελάχιστη και τη μέγιστη θερμοκρασία $T_{low}(t)$ και $T_{high}(t)$, λαμβάνοντας υπόψη την παράμετρο κλίμακας β και το προσαρμοσμένο επίπεδο της κλίμακας $S'(t)$.

Στη συνέχεια, προσδιορίζεται το φορτίο L_{total} της t ώρας της d ημέρας που αποτελείται από τρεις συνιστώσες : την καθημερινή, την εβδομαδιαία και την κυκλική συνιστώσα, $L_{daily}(t,d)$, $L_{weekly}(t,d)$, $L_{trend}(t,d)$ αντίστοιχα. Εφαρμόζοντας τη μέθοδο του κινούμενου μέσου όρου $M(t,d)$ και υπολογίζοντας τον κινούμενο λόγο των κέντρων $CMR(t,d)$ έχουμε τα ακόλουθα:

$$M(t,d) = \frac{\left[\frac{L_{total}(t-12,d) + L_{total}(t+12,d)}{2} + \sum_{i=-11}^{11} L_{total}(t+i,d) \right]}{24} \quad (2.2.3-7)$$

$$CMR(t,d) = L_{total}(t,d) - M(t,d) \quad (2.2.3-8)$$

$$L_{daily}(t,d) = \frac{1}{N} \sum_{i=1}^N CMR(t,i) \quad (2.2.3-9)$$

όπου $N = 7 \cdot W - 1$, με W ο αριθμός των εβδομάδων που χρησιμοποιούνται ως ιστορικά δεδομένα.

Με τα πιο πάνω προσδιορίζουμε την καθημερινή συνιστώσα φορτίου. Παρομοίως, προσδιορίζεται και η εβδομαδιαία συνιστώσα με 168 ώρες, έναντι των 24 ωρών της καθημερινής και από το υπόλοιπο προσδιορίζεται η κυκλική συνιστώσα. Συσχετίζοντας την καθημερινή συνιστώσα φορτίου με τη θερμοκρασία, μέσω της μεθόδου των ελαχίστων τετραγώνων, επαναπροσδιορίζεται η καθημερινή συνιστώσα και κατόπιν από το άθροισμα των τριών μερών,

προκύπτει το τελικό φορτίο της επόμενης μέρας, ώστε να μπορεί να γίνει ο προγραμματισμός αποθήκευσης ενέργειας για το επόμενο εικοσιτετράωρο.

➤ 2.3 Μέθοδοι αποσύνθεσης

Κατά τις μεθόδους αποσύνθεσης εφαρμόζουμε κάποιες απλές μαθηματικές σχέσεις, με σκοπό την απομόνωση των τεσσάρων βασικών συνιστωσών των χρονοσειρών:

- 1) την εποχιακότητα S_t ,
- 2) την κυκλικότητα C_t ,
- 3) την τάση T_t και
- 4) την τυχαιότητα R_t .

Αυτές συνδέονται με μία συνάρτηση f και με την παρατήρηση X_t , σύμφωνα με την ακόλουθη σχέση:

$$X_t = f(S_t, C_t, T_t, R_t) \quad (2.3-1)$$

Η μεθοδολογία που ακολουθείται, όταν η συνάρτηση είναι πολλαπλασιαστική [A2], είναι η εξής:

- α) υπολογισμός των κεντρικών κινητών μέσων όρων,
- β) υπολογισμός των λόγων εποχιακότητας,
- γ) απαλοιφή της τυχαιότητας,
- δ) εύρεση της αποεποχικοποιημένης σειράς,
- ε) εύρεση της σειράς τάσης-κύκλου,
- στ) διαχωρισμός των συνιστωσών τάσης και κυκλικότητας.

Στη συνέχεια με βάση αυτά τα στοιχεία σε και σε αντίστοιχη εφαρμογή για την εύρεση της αιχμής του τριμηνιαίου φορτίου, προκύπτει ακρίβεια της τάξης, περίπου του 5% για το επόμενο έτος.

➤ 2.4 Μέθοδος πολλαπλής γραμμικής παλινδρόμησης

Η μέθοδος γραμμικής πολλαπλής παλινδρόμησης στηρίζεται στην παρακάτω σχέση της μορφής [B1], [A20] :

$$y(t) = a_0 + a_1 \cdot x_1(t) + \dots + a_n \cdot x_n(t) + b(t) \quad (2.4-1)$$

όπου $y(t)$ είναι το ηλεκτρικό φορτίο, $x_i(t)$ η i -οστή παράμετρος που συσχετίζεται με την έξοδο, $a_i(t)$ ο i -οστός συντελεστής συσχέτισης και $b(t)$ η τυχαία μεταβλητή του σφάλματος με μηδενική μέση τιμή. Λόγω της απλής μορφής του μοντέλου, αυτό δίνει μεγαλύτερα σφάλματα από τις υπόλοιπες μεθοδολογίες από 2 ως 10 φορές [A20].

Ένα απλό παράδειγμα αναφέρεται στο [A13], όπου το φορτίο αποτελείται από τέσσερις συνιστώσες:

α) από τη **συνιστώσα καιρού**, που προσδιορίζεται μέσω της σχέσης:

$$W(t) = a_1 \cdot T(t) + a_2 \cdot T(t-24) + a_3 \cdot S(t) + a_4 \cdot V(t) + a_5 \cdot H(t) \quad (2.4-2)$$

λαμβάνοντας υπόψη τη θερμοκρασία της αντίστοιχης ώρας $T(t)$ και προ εικοσιτεσσάρων ωρών $T(t-24)$, τη διάρκεια της ηλιοφάνειας $S(t)$, την ταχύτητα ανέμου $V(t)$ και τη σχετική υγρασία της τρέχουσας ώρας $H(t)$,

β) από τη **συνιστώσα κανονικού φορτίου**, που είναι ανεξάρτητη του καιρού, λαμβάνοντας συγκεκριμένες μορφές ανάλογα με το είδος της μέρας,

γ) από το **φορτίο ειδικών περιστάσεων**, όπως αργίες, απεργίες κτλ.- σε αυτό το σημείο μπαίνει η εμπειρία του χειριστή,

δ) από το **τυχαίο φορτίο**.

Εδώ δίνεται ιδιαίτερη προσοχή στην επιλογή μεταβλητών καιρού και μέσω των συντελεστών συσχέτισης, προσδιορίζονται τα απαραίτητα στοιχεία -πιθανοτική ανάλυση. Επίσης τα ανώμαλα στοιχεία απομακρύνονται από το σύνολο δεδομένων και τελικά υπάρχει ικανοποιητική απόδοση της μεθόδου, με ποσοστό απόλυτου σφάλματος της τάξης του 2%.

➤ 2.5 Οικονομετρική μέθοδος

Τα κύρια βήματα που ακολουθούνται σε αυτήν τη μέθοδο είναι : 1) η ανάπτυξη αρχικά ενός συγκεκριμένου μοντέλου, 2) η επιλογή των δεδομένων του εκάστοτε προβλήματος, 3) η μεθοδική ανάλυσή τους και 4) η εκ νέου προσαρμογή του μοντέλου μας βάσει των συμπερασμάτων μας. Τα στοιχεία που προβληματίζουν είναι η εκτίμηση της τρέχουσας κατάστασης και η πρόβλεψη των επικείμενων μεταβολών. Οι λύσεις σε αυτά τα δύο προβλήματα, είναι δυνατό να διαφέρουν σημαντικά, παρόλη την ομοιότητα των διαδικασιών που εφαρμόζουμε για τις παραπάνω λύσεις.

Για περαιτέρω κατανόηση των παραπάνω, θα αναφέρουμε το ακόλουθο παράδειγμα [A21]. Σε αυτό προσεγγίζουμε τη ζήτηση φορτίου μέσω των τελικών χρήσεων. Οι βασικές αρχές που χρησιμοποιούμε εδώ, είναι η κατηγοριοποίηση των πελατών σε οικιακούς, βιομηχανικούς και εμπορικούς, και η διακριτοποίηση των χρήσεων ηλεκτρικής ενέργειας σε δύο ομάδες : των εξαρτώμενων από τη θερμοκρασία και των ανεξάρτητων. Για παράδειγμα στην πιο απλή περίπτωση η συνολική ενέργεια E_i ενός στοιχείου τελικής χρήσης i για μία συγκεκριμένη περίοδο, συνδέεται με τη ζήτηση D_{ij} στην ώρα j για το στοιχείο τελικής χρήσης i , που προσδιορίζεται μέσω του αριθμού των στοιχείων τελικής χρήσης N_i φορτίου C_i ανά στοιχείο χρήσης i , τα οποία συμμετέχουν κατά την ώρα j κατά το μέρος F_{ij} :

$$E_i = \sum_j D_{ij} \quad (2.5-1)$$

$$D_{ij} = N_i \cdot C_i \cdot F_{ij} \quad (2.5-2)$$

Με την επανάληψη της παραπάνω μεθοδολογίας για το σύνολο των τελικών χρήσεων, καθίσταται τελικά δυνατό να προσδιορίσουμε το ωριαίο φορτίο και τη συνολική ενέργεια μίας περιόδου.

➤ 2.6 Αυτοπαλινδρομικά μοντέλα κινούμενου μέσου όρου

Σ' αυτήν τη μέθοδο περιλαμβάνονται τα αυτοπαλινδρομικά μοντέλα κινούμενου μέσου όρου ARMA (p,q), όπου ορίζονται μέσω των ακόλουθων σχέσεων:

$$\phi(B) \cdot y(t) = \theta(B) \cdot a(t) \quad (2.6-1)$$

όπου

$$\theta(B) = 1 - \theta_1 \cdot B - \dots - \theta_q \cdot B^q \quad (2.6-2)$$

$$\phi(B) = 1 - \phi_1 \cdot B - \dots - \phi_p \cdot B^p \quad (2.6-3)$$

όπου $y(t)$ είναι το ηλεκτρικό φορτίο, $a(t)$ ο λευκός θόρυβος, θ_i ο i -οστός συντελεστής βάρους του θορύβου πριν i χρονικά βήματα, ϕ_i ο i -οστός συντελεστής βάρους του φορτίου πριν i χρονικά βήματα. Το σημαντικό μειονέκτημα αυτής της μεθόδου είναι ότι προσεγγίζει στάσιμες χρονοσειρές φορτίου, με αποτέλεσμα να περιορίζεται αρκετά το πλήθος των εφαρμογών της.

Στη συνέχεια θα γίνει μια απλή αναφορά μίας βελτιωμένης παραλλαγής αυτής της μεθόδου. Σε αυτήν την περίπτωση [A11] λαμβάνει χώρα σε πραγματικό χρόνο η βραχυπρόθεσμη πρόβλεψη φορτίου, που συγκεκριμένα αφορά δίκτυα διανομής και στηρίζεται κατά κύριο λόγο σε μοντέλο ARMA, με χρήση μεταβλητού παράγοντα απώλειας μνήμης. Η πρόβλεψη γίνεται με σφάλμα 2% για το επόμενο 24-ώρο και 2,5% για τις επόμενες 168 ώρες. Οι παράγοντες που λαμβάνονται υπόψη είναι η θερμοκρασία, η υγρασία, και η ταχύτητα του ανέμου, όχι όμως η συννεφιά ή οι καταιγίδες. Αυτό που δημιουργείται τελικά, είναι μία ισοδύναμη θερμοκρασία T_{eq} που δίνεται από τη σχέση (1.4-3) του πρώτου κεφαλαίου.

Το ζητούμενο φορτίο με τη σειρά του αποτελείται από τέσσερις σημαντικές συνιστώσες :

1) το ντετερμινιστικό γνωστό φορτίο, 2) το ανεξάρτητο από τον καιρό φορτίο, 3) το εξαρτημένο από τον καιρό φορτίο και το 4) το υπόλοιπο του φορτίου που συνιστά για μας το «θόρυβο». Η πρώτη συνιστώσα είναι γνωστή και καθ' όλα προγραμματισμένη, γι' αυτό και διαχωρίζεται από το υπόλοιπο φορτίο. Η ανεξάρτητη από τον καιρό συνιστώσα επηρεάζεται από κοινωνικούς και γεωγραφικούς παράγοντες και προσεγγίζεται μέσω ανάλυσης Fourier, έχοντας ως περίοδο της βασικής συχνότητας τις 168 ώρες. Η συνιστώσα του φορτίου που εξαρτάται από τον καιρό προσεγγίζεται με ένα κυβικό πολυώνυμο, όπου υπάρχει και το κατάλληλο χρονικό παράθυρο των 6 προηγούμενων ωρών. Για το υπόλοιπο φορτίο, που θεωρείται για μας ως στοχαστικός –τυχαίος– θόρυβος, χρησιμοποιείται ένα μοντέλο της λεγόμενης μορφής ARMA που λαμβάνει υπόψη του την πρώτη, τη δεύτερη και την εικοστή τέταρτη προηγούμενη ώρα. Το σύστημα που προκύπτει τελικά μας δίνει αποδόσεις λύσεων βελτιωμένες κατά 50% περίπου, σε σχέση με την αρχική μέθοδο. Επιπρόσθετα θα πρέπει να τονιστεί ότι δε λαμβάνονται υπόψη κατά τον προσδιορισμό των παραμέτρων της μεθόδου οι ειδικές μέρες –αργίες, ημιαργίες κ.λ.π.– καθαυτές, αλλά για την πρόβλεψή τους γίνονται διάφορες παραδοχές, όπως για παράδειγμα η ημέρα των Χριστουγέννων θεωρείται ότι παρουσιάζει παρόμοια συμπεριφορά με αυτή των Κυριακών και η Μεγάλη

Παρασκευή αναλογεί σε συμπεριφορά Σαββάτου. Αν όμως η αργία εμφανιστεί το σαββατοκύριακο, τότε η Παρασκευή και η Δευτέρα θα προκύψουν από τη μέση τιμή των αντίστοιχων εργάσιμων ημερών με το Σάββατο. Τέλος η παραπάνω μέθοδος δεν είναι σε θέση να αντιμετωπίσει μεγάλες και απότομες μεταβολές των καιρικών συνθηκών.

➤ 2.7 Ολοκληρωμένα αυτοπαλινδρομικά μοντέλα κινούμενου μέσου όρου

Σ' αυτή τη μέθοδο περιλαμβάνονται τα ολοκληρωμένα αυτοπαλινδρομικά μοντέλα κινούμενου μέσου όρου ARIMA (p,q), όπου ορίζονται μέσω των σχέσεων (2.6-2) και (2.6-3) και των ακόλουθων :

$$\phi(B) \cdot \nabla^d y(t) = \theta(B) \cdot a(t) \quad (2.7-1)$$

$$\text{όπου} \quad \nabla^d y(t) = (1 - B)^d \cdot y(t) \quad (2.7-2)$$

και $y(t)$, $a(t)$, θ_i , ϕ_i , ορίστηκαν στην προηγούμενη παράγραφο. Με αυτήν τη μέθοδο προσεγγίζουμε μη στάσιμες χρονοσειρές φορτίου, οπότε είναι σε θέση να αντιμετωπίσει πιο εύκολα το πρόβλημα της πρόβλεψης και άρα χρησιμοποιείται σε ένα ευρύτερο πλήθος εφαρμογών.

Επίσης εδώ, αξίζει να αναφέρουμε ότι υπάρχει μία βασική παραλλαγή των ολοκληρωμένων αυτοπαλινδρομικών μοντέλων κινούμενου μέσου όρου, όπου μπορούμε να εκφράσουμε δύο ή και περισσότερες διαφορετικές εποχικότητες [A20].

Γενικότερα αυτό το μοντέλο αναφέρεται σε ένα σημαντικό πλήθος άρθρων και, ιδιαίτερα από τη δεκαετία του '90, είναι το κύριο και άμεσο στοιχείο σύγκρισης με τα νεότερα μοντέλα των νευρωνικών δικτύων, λόγω της πολύ καλής συμπεριφοράς του σε σχέση με τις υπόλοιπες υπάρχουσες κλασικές μεθοδολογίες, αλλά και λόγω της κατανοητής μορφής του [A22].

➤ 2.8 Λοιπές παλινδρομικές και μη μέθοδοι

Σε αυτό το σημείο θα πρέπει να πούμε ότι υπάρχει ένα μεγάλο πλήθος παραλλαγών όλων των παραπάνω μεθόδων, για παράδειγμα με εφαρμογή των διαφόρων τεχνικών μετασχηματισμού. Επίσης μπορούν να υπάρξουν και οι μη παραμετρικές μέθοδοι που χρησιμοποιούν μορφές και εκτιμητές που έχουν παρθεί από τον κλάδο των πιθανοτήτων, καθώς και ντετερμινιστικές προσεγγίσεις του προβλήματος ή υβριδικές μορφές κλασικών μεθόδων.

Για παράδειγμα αναφέρουμε μια μορφή συνεχούς αυτοπαλινδρομικού μοντέλου που είναι η τύπου *Hilbert*, για μη στάσιμες στοχαστικές χρονοσειρές βραχυπρόθεσμης πρόβλεψης φορτίου [A23]. Οι ισότητες στις οποίες στηρίζεται αυτή η μέθοδος δεν είναι άλλες από αυτές της απλής παλινδρόμησης, κατάλληλα βέβαια τροποποιημένες βάσει του μετασχηματισμού *Hilbert*. Η αντίστοιχη εφαρμογή της μεθόδου με 59 εβδομάδες ιστορικών δεδομένων σε δίκτυο μέσης τάσης, παρέχει ικανοποιητικά αποτελέσματα με σφάλμα κάτω του 2%, λόγω του μεγάλου πλήθους εβδομαδιαίων δεδομένων, παρά την έλλειψη μετεωρολογικών δεδομένων.

Πάντως σε κάθε περίπτωση οι κλασικές μέθοδοι έχουν τεράστια αξία καθώς κάλυψαν, συνεχίζουν να καλύπτουν και θα αποτελούν πάντοτε τη βάση και τον κορμό για την ανάπτυξη οποιωνδήποτε νεότερων μεθόδων, αφού άλλωστε μέσω αυτών έγιναν οι αναλύσεις και προσδιορίστηκαν ποιοτικά αλλά και ποσοτικά οι μεταβλητές εκείνες που επηρεάζουν τα ωριαία φορτία ζήτησης.

➤ 2.9 Μέθοδοι βραχυπρόθεσμης πρόβλεψης φορτίου με χρήση Ασαφούς Λογικής

Μια πολύ σημαντική κατηγορία νεότερων μεθόδων πρόβλεψης φορτίου, είναι και αυτή που σχετίζεται με την χρήση και εφαρμογή της ασαφούς λογικής- *Fuzzy Logic*. Παρά το ότι τα τεχνητά νευρωνικά δίκτυα βρήκαν πλούσιο έδαφος εφαρμογών στο πρόβλημα της πρόβλεψης φορτίου σε συνδυασμό με την ασαφή λογική, η τελευταία είχε αρχίσει ανεξάρτητα από τα νευρωνικά δίκτυα να βρίσκει εφαρμογή στα προβλήματα των συστημάτων ηλεκτρικής ενέργειας. Γι' αυτόν το λόγο παρακάτω θα πραγματοποιήσουμε μία μικρή εισαγωγή στις έννοιες της ασαφούς λογικής και στη συνέχεια θα αναφερθούμε στα υβριδικά μοντέλα που περιλαμβάνουν ασαφή λογική μαζί με τεχνητά νευρωνικά δίκτυα.

➤ 2.9.1 Εισαγωγή στην ασαφή λογική

Οι προσπάθειες για αυτοματοποίηση διαφόρων εκφράσεων της ανθρώπινης δραστηριότητας, από τη διάγνωση ενός ασθενούς μέχρι και τις διαδικασίες χειρισμού μίας μονάδας παραγωγής ενέργειας, χαρακτηρίζονται έντονα από το διαφορετικό τρόπο με τον οποίο σκέφτονται και κρίνουν οι άνθρωποι, σε σχέση με τον τρόπο με τον οποίο είναι προγραμματισμένοι να δρουν οι υπολογιστές. Για παράδειγμα αναφέρουμε τον τρόπο με τον οποίο αποθηκεύονται οι θερμοκρασίες μίας μέρας στον υπολογιστή, όπου κρατά μία αριθμητική τιμή, χωρίς να είναι σε θέση να τη χαρακτηρίσει. Αντίθετα ο άνθρωπος, χωρίς να κρατά μία συγκεκριμένη θερμοκρασία χαρακτηρίζει τη θερμοκρασία ως υψηλή για την τρέχουσα ημέρα, αν είναι πάνω από τους 30 °C το καλοκαίρι και 20 °C το χειμώνα, δεδομένο που δε θα μπορούσε ο υπολογιστής να χρησιμοποιήσει.

Η ασαφής λογική θα λέγαμε ότι είναι ουσιαστικά μια προσπάθεια αντιγραφής και αφομοίωσης του ανθρώπινου τρόπου σκέψης, ο οποίος δεν καθορίζεται με ένα απλό «ναι» ή «όχι» που θα προσομοιωνόταν εύκολα με μία δυαδική μεταβλητή, αλλά εμπεριέχει σε μεγάλο βαθμό την έννοια της ασάφειας και της αοριστίας. Με λίγα λόγια η ασαφής λογική είναι ένα χαρακτηριστικό της αναπαράστασης των διαφόρων «αντικειμένων» σ' ένα περιβάλλον συμβόλων και γενικά μία ιδιότητα των μοντέλων και της γλώσσας.

Στη συνέχεια θα ορίσουμε ένα σύνολο βασικών εννοιών, όπως είναι οι εξής :

- ❑ Ασαφή σύνολα, δηλαδή εκείνα τα σύνολα που δεν έχουν ένα ορισμένο τρόπο συμμετοχής σε μία ομάδα, αλλά επιτρέπουν τα στοιχεία-αντικείμενα x που ανήκουν σε αυτά να έχουν βαθμούς συμμετοχής $\mu_F(x)$ με την ιδιότητα $0 < \mu_F(x) < 1$. Ένα ασαφές σύνολο με πεδίο ορισμού το σύνολο U , ορίζεται από τη συνάρτηση συμμετοχής $\mu_F: U \rightarrow [0,1]$ και χαρακτηρίζεται από μία λεκτική μεταβλητή F .

- Λεκτικές μεταβλητές, δηλαδή οι συνηθισμένες λεκτικές εκφράσεις που χρησιμοποιούνται, για να συμβολίσουν ένα συγκεκριμένο ασαφές σύνολο σ' ένα δοσμένο πρόβλημα και παίρνουν τιμές, όπως για παράδειγμα «μεγάλο», «μεσαίο» και «μικρό».
- Λογικό συμπέρασμα, δηλαδή η διαδικασία εξαγωγής συμπεράσματος με την ακόλουθη δομή : Έστω ότι το δεδομένο 1 είναι «Το x είναι A » και η υπόθεση 2 είναι «Αν x είναι A , τότε το y είναι B », τότε προκύπτει το συμπέρασμα ότι «Το y είναι B »
- Δομή ασαφούς δικτύου, δηλαδή εκείνο το σύστημα, το οποίο χρησιμοποιεί την ασαφή λογική για την επίλυση ενός προβλήματος, ενώ πρακτικά του δίνεται μία μη-ασαφής είσοδος και λαμβάνεται πάλι μια μη-ασαφής έξοδος. Ουσιαστικά αποτελείται από τέσσερα στοιχεία :

α) τη θύρα ασάφειας, που είναι μία απεικόνιση του πραγματικού μη ασαφούς χώρου κατάστασης των εισόδων $U \subseteq R^n$, στα ασαφή σύνολα που ορίζονται στο U . Ουσιαστικά είναι ο σύνδεσμος ανάμεσα στον εξωτερικό κόσμο και στο ασαφές σύστημα.

β) τη βάση ασαφών κανόνων, το οποίο είναι ένα σύνολο λεκτικών κανόνων της μορφής αν-τότε. Το αν-τμήμα είναι το τμήμα της υπόθεσης, ενώ το τότε-τμήμα είναι το τμήμα συμπεράσματος. Δηλαδή έχει την παρακάτω μορφή:

$$R^i: \text{Αν } (x_1 \text{ είναι } A_1^i), (x_2 \text{ είναι } A_2^i) \dots \text{ και } (x_n \text{ είναι } A_n^i), \text{ τότε } y^i \text{ είναι ίσο με } W^i$$

όπου R^i είναι ο ασαφής κανόνας i , x_i οι μεταβλητές εισόδου του ασαφούς συστήματος, y_i η έξοδος του i -οστού κανόνα και A_n^i οι λεκτικοί όροι του κανόνα που χαρακτηρίζονται από τις συναρτήσεις συμμετοχής.

γ) το μηχανισμό εξαγωγής ασαφών συμπερασμάτων, ο οποίος είναι ο μηχανισμός μέσω του οποίου εξάγονται συμπεράσματα, βασιζόμενα στους ασαφείς κανόνες της μορφής αν-τότε και στους συνθετικούς κανόνες εξαγωγής συμπερασμάτων. Με αυτόν τον τρόπο διαμορφώνονται όλα τα σύνολα εξόδου του συστήματος.

δ) τη θύρα αποασαφοποίησης, όπου λαμβάνει τα αποτελέσματα του μηχανισμού εξαγωγής ασαφών συμπερασμάτων και παράγει μία σαφή έξοδο, η οποία αντιπροσωπεύει και την πραγματική έξοδο του συστήματος. Η αποασαφοποίηση γίνεται συνήθως με τη μέθοδο του σταθμισμένου μέσου όρου.



Στη συνέχεια θα παρουσιάσουμε τις εφαρμογές τις ασαφούς λογικής στο πρόβλημα της βραχυπρόθεσμης πρόβλεψης φορτίου.

➤ 2.9.2 Εφαρμογή της ασαφούς λογικής στη βραχυπρόθεσμη πρόβλεψη φορτίου και ο συνδυασμός της με τα τεχνητά νευρωνικά δίκτυα

Το πρόβλημα της βραχυπρόθεσμης πρόβλεψης φορτίου γίνεται ιδιαίτερα δύσκολο στην περίπτωση μελέτης ενός συστήματος διανομής και στην εύρεση, είτε της ωριαίας φόρτισης, είτε της μέγιστης ζήτησης, που αποτελεί και το στοιχείο διαστασιολόγησης του δικτύου. Οι μετρήσεις που διαθέτει κανείς γι' αυτά τα δίκτυα είναι πολύ περιορισμένες, περιλαμβάνοντας ορισμένες μετρήσεις σε τροφοδότες και την κατανάλωση ενέργειας των καταναλωτών, μέσω τιμολογίων. Οι παράγοντες τυχαιότητας που συμμετέχουν είναι πολλοί, όπως η τυχαία συμπεριφορά της ζήτησης των καταναλωτών σε ενέργεια και ισχύ, η επίδραση του παράγοντα τυχαιότητας στο σύστημα, π.χ. βλάβες κτλ., με συνέπεια να οδηγούμαστε στην εφαρμογή της ασαφούς λογικής.

Συγκεκριμένα στο [A24] προτείνεται η χρήση της ασαφούς λογικής για τον προσδιορισμό της μέγιστης ζήτησης ισχύος σε βραχυπρόθεσμη πρόβλεψη. Με αυτόν τον τρόπο είναι δυνατό να εκφραστούν ο ασαφής χαρακτήρας των συναρτήσεων ελέγχου, λόγω έλλειψης πλήρους μοντελοποίησης του συστήματος και χρήσης ευριστικών μεθόδων απόφασης, ο ασαφής χαρακτήρας της κατάστασης του δικτύου, του ελέγχου του και της σχετικής πληροφορίας γύρω από αυτό. Προηγουμένως είχαν προταθεί δύο μεθοδολογίες, είτε η χρήση των τυπικών καμπυλών φορτίου και πρόσθετης πληροφορίας για την κατανάλωση σε διαφορετικές περιόδους κτλ., είτε η διακριτοποίηση του φορτίου σε σχέση με το είδος των καταναλωτών που συναντά κανείς στο σύστημα. Οι αδυναμίες τους βρίσκονται στις διάφορες εκτιμήσεις των στοιχείων που λείπουν και στην εξάλειψη του ασαφούς χαρακτήρα ορισμένων παραγόντων. Γι' αυτό μία μεθοδολογία με χρήση ασαφούς λογικής είναι πολλά υποσχόμενη.

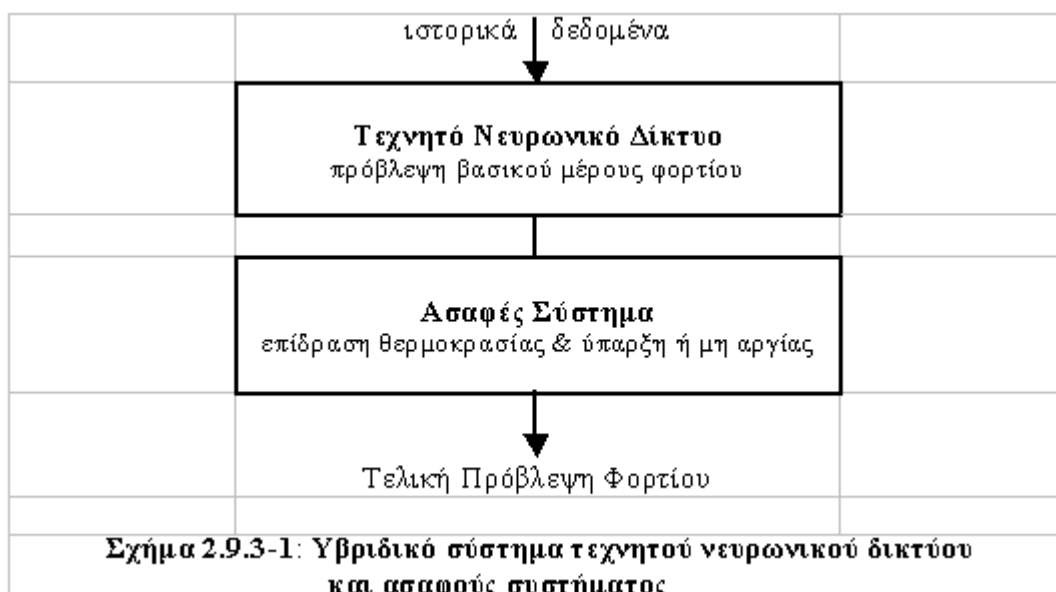
Πιο πρωτοποριακή μέθοδος είναι η χρήση ασαφών νευρωνικών δικτύων που εκμεταλλεύεται κανείς τόσο τα ποιοτικά, όσο και τα ποσοτικά στοιχεία που διαθέτουν οι ειδικοί οργανισμοί και στατιστικές υπηρεσίες [A25], [Δ2]. Επίσης μπορούμε να αξιοποιήσουμε υπάρχοντες εμπειρικούς κανόνες, που με άλλο τρόπο θα ήταν αδύνατο. Η όλη διαδικασία είναι πιο κοντά στην πραγματικότητα, μέσω των μεταβλητών -μέση τιμή και διασπορά τους.

Με την εκτέλεση προγράμματος για το ελληνικό ηπειρωτικό σύστημα του 1997, χωρίς να περιλαμβάνονται οι μέρες διακοπών και ανώμαλων καιρικών συνθηκών, ελήφθησαν αρκετά ικανοποιητικά αποτελέσματα. Συγκρινόμενο με ένα τεχνητό πολυεπίπεδο νευρωνικό δίκτυο των ίδιων εισόδων, έχει μικρή διαφοροποίηση ως προς την ακρίβεια (μέση τιμή απόλυτου σφάλματος 2,9% και μέση απόκλιση 1,1%, έναντι 2,88% και 1,15% αντίστοιχα για το τεχνητό). Όμως το σημαντικό πλεονέκτημα της χρήσης ασαφούς λογικής, είναι ο 5 φορές λιγότερος υπολογιστικός χρόνος που απαιτείται σε σχέση με τα τεχνητά Ν.Δ.-18 sec σε Η/Υ 66 MHz για την πρώτη, έναντι 92 sec των δεύτερων.

➤ 2.9.3 Μέθοδοι βραχυπρόθεσμης πρόβλεψης φορτίου με χρήση υβριδικών μοντέλων

Τα τεχνητά νευρωνικά δίκτυα δεν περιορίστηκαν μόνο στη δημιουργία των ασαφών δικτύων, αλλά προχώρησαν και στο σχηματισμό μίας σειράς υβριδικών μοντέλων, τόσο με τα ασαφή συστήματα, όσο και με τις κλασικές μεθόδους.

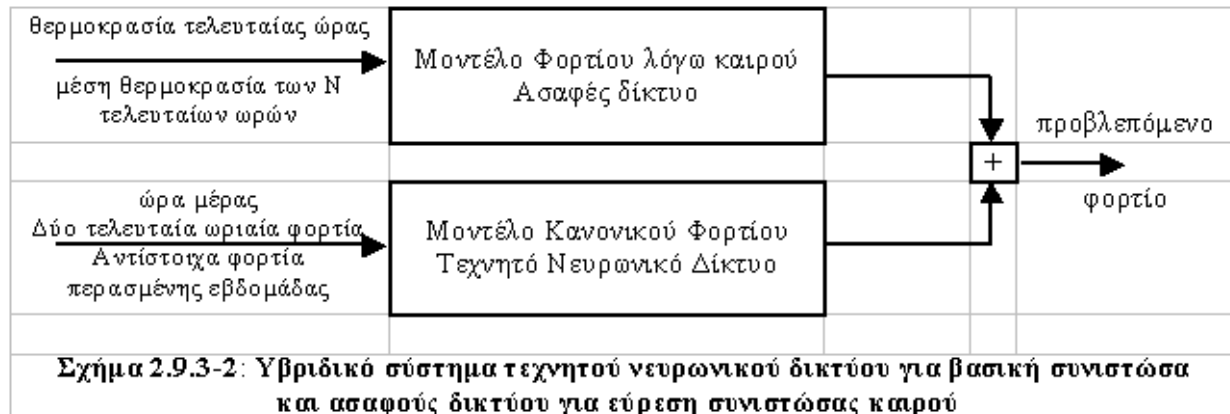
Πιο συγκεκριμένα στο [Α26] παρουσιάζεται ένα υβριδικό μοντέλο που συνδυάζει τα τεχνητά νευρωνικά δίκτυα με τα ασαφή συστήματα. Συγκεκριμένα, υπάρχει το τεχνητό πολυεπίπεδο νευρωνικό δίκτυο δύο κρυφών επιπέδων με 5 και 3 νευρώνες αντίστοιχα, που χρησιμοποιείται για την πρόβλεψη της συνιστώσας του βασικού μέρους του φορτίου, στηριζόμενο στα ωριαία φορτία των δύο προηγούμενων ωρών, της αντίστοιχης ώρας και των δύο προηγούμενων ωρών για μία, δύο, επτά και δεκατέσσερις μέρες πιο πριν. Με τη βοήθεια των ασαφών δικτύων και με κατάλληλη επιμέρους κατηγοριοποίηση, γίνεται ο προσδιορισμός της επίδρασης της μέσης θερμοκρασίας εκείνης της μέρας και της διαφοράς της με των αντίστοιχων ημερών για τις τρεις προηγούμενες εβδομάδες, όπως και των διακοπών λαμβάνοντας υπόψη τη συμπεριφορά των καταναλωτών κατά τα προηγούμενα έτη. Η βασική μορφή του τελικού μοντέλου που προκύπτει παριστάνεται με το παρακάτω διάγραμμα :



Από όλη τη διαδικασία-αλληλεπίδραση των δύο συστημάτων, προκύπτει η τελική πρόβλεψη φορτίου με μέσο όρο απόλυτου σφάλματος για τις κανονικές ημέρες γύρω στο 1,3% και για τις αργίες περίπου στο 2,5%, όντας πολύ καλύτερο από εκείνο της μεθόδου με την εκθετική εξομάλυνση που δίνει μεγαλύτερα σφάλματα κατά 1% τουλάχιστον.

Επίσης πολύ ενδιαφέρουσα περίπτωση είναι το υβριδικό μοντέλο των ασαφών και των τεχνητών νευρωνικών δικτύων, που βασίζεται στο διαχωρισμό του φορτίου σε δύο συνιστώσες: στην εξαρτώμενη και στην ανεξάρτητη από τη θερμοκρασία συνιστώσα [A18].

Η συνολική δομή του δικτύου βρίσκεται παρακάτω :



Η παραπάνω διαδικασία, όσον αφορά κομμάτι του νευρωνικού δικτύου πραγματοποιείται με σύγκριση των μηνιαίων μέσων ωριαίων φορτίων ημερήσιας διάρκειας, με βάση τρία είδη ημερών, εργάσιμες, Σάββατα και Κυριακές. Από αυτήν τη διαδικασία εντοπίζεται το ελάχιστο φορτίο που σε κάθε περίπτωση αποτελεί τη βάση. Στη συνέχεια με κατάλληλες πράξεις προκύπτουν τα βασικά φορτία που εκπαιδεύουν τα τεχνητά νευρωνικά δίκτυα, χρησιμοποιώντας την ώρα της μέρας, το ωριαίο φορτίο των τελευταίων δύο ωρών και των αντίστοιχων ωρών πριν μία εβδομάδα και ένα μήνα. Τα ασαφή δίκτυα, επιπρόσθετα, χρησιμοποιούνται για την πρόβλεψη της συνιστώσας του φορτίου, η οποία εξαρτάται από την θερμοκρασία. Ουσιαστικά βρίσκουμε τη μέση θερμοκρασία κάθε ώρας μέχρι προ 72 ωρών και μέσω κατάλληλων κανόνων συσχετίζεται η μέση θερμοκρασία με την ωριαία θερμοκρασία. Στη συνέχεια χρησιμοποιώντας συναρτήσεις τύπου Gauss και κατάλληλες διαδικασίες ασαφοποίησης και αποασαφοποίησης, προκύπτει τελικά το επιθυμητό φορτίο μέσω άθροισης.

Η προαναφερθείσα μέθοδος έχει μέση τιμή απόλυτου σφάλματος από 1,12% ως 2,38%, έναντι του σχεδόν διπλάσιου που προκύπτει με κλασικό πολυεπίπεδο νευρωνικό δίκτυο με τα ίδια δεδομένα.

Σε αυτό το σημείο θα πρέπει να πούμε ότι τα τεχνητά νευρωνικά δίκτυα δε συνδυάζονται μόνο με τα ασαφή δίκτυα, αλλά και με κλασικές μεθόδους. Για παράδειγμα μία ενδιαφέρουσα περίπτωση έχουμε στην περίπτωση χρήσης μοντέλων επεξεργασίας μη μόνιμων χρονοσειρών, συνδυαζόμενα με τεχνητά νευρωνικά δίκτυα και λαμβάνοντας υπόψη τις επιπτώσεις των καιρικών συνθηκών –θερμοκρασία, υγρασία, βροχόπτωση, καθώς και λιακάδα (προηγούμενης μέρας).

Επιπρόσθετα τα υβριδικά μοντέλα δε συνεπάγονται κατ' ανάγκη την ύπαρξη τεχνητών νευρωνικών δικτύων. Υπάρχουν και αρκετές περιπτώσεις, τόσο συνδυασμού κλασικών μεθοδολογιών, όσο και ασαφούς λογικής με κλασική μέθοδο, χωρίς την ύπαρξη των Τ.Ν.Δ.

Σε κάθε περίπτωση όμως, η πληθώρα αυτών των υβριδικών δικτύων υποδηλώνει την προσπάθεια για μείωση του σφάλματος στη βραχυπρόθεσμη πρόβλεψη φορτίου. Επακόλουθο αυτής της προσπάθειας είναι και η ανάπτυξη μεθόδων αξιολόγησης και κοστολόγησης των διαφόρων τρόπων πρόβλεψης. Μάλιστα έχει δοθεί ιδιαίτερο βάρος στη βελτίωση των εισερχόμενων παραγόντων στα οποιαδήποτε μορφής δίκτυα, όπως είναι η θερμοκρασία και ο άνεμος, καθώς και στη βελτίωση της πρόβλεψης κατά τη διάρκεια των ωρών υψηλότερου κόστους, ώστε να υπάρξει τελικά μεγαλύτερη μείωση του κόστους και άρα αύξηση του κέρδους.

Ακολουθώντας, θα αναφερθούμε αναλυτικά στις εφαρμογές των τεχνητών νευρωνικών δικτύων, όσον αφορά τη βραχυπρόθεσμη πρόβλεψη φορτίου.

➤ 2.10 Μέθοδοι βραχυπρόθεσμης πρόβλεψης φορτίου, στηριζόμενες στα Τεχνητά Νευρωνικά Δίκτυα

Λόγω της υφής του προβλήματος της βραχυπρόθεσμης πρόβλεψης φορτίου, δηλαδή λόγω έλλειψης καθορισμένων μαθηματικών σχέσεων που προσδιορίζουν το ωριαίο φορτίο των επόμενων ωρών μέσω παραμέτρων που πολλές φορές δεν είναι πλήρως γνωστές, όπως για παράδειγμα οι καιρικές συνθήκες, οδηγούμαστε στην έρευνα νέων μεθοδολογιών, όπως είναι τα τεχνητά νευρωνικά δίκτυα διαφόρων μορφών – *A.N.N.*

Όταν αρχικά είχε μελετηθεί η παραπάνω μεθοδολογία, λόγω μη ύπαρξης αυστηρών μαθηματικών μοντέλων και κριτηρίων, π.χ. το πλήθος των επιπέδων και των νευρώνων είναι μεταβλητό, οδήγησε στην αποφυγή εφαρμογής της μεθόδου, λόγω της μη ελεγχόμενης αξιοπιστίας της [A15], [A27]. Επίσης τα τεχνητά νευρωνικά δίκτυα παρουσιάζουν μεγαλύτερη ευαισθησία με το σύστημα της ηλεκτρικής εταιρείας, για παράδειγμα έχουμε άλλη συμπεριφορά για τα ελληνικά δεδομένα, όσον αφορά το κλίμα και άλλη για τις βορειοανατολικές πολιτείες της Αμερικής, με συνέπεια ανάλογα με την περίπτωση να χρειάζεται μερική τροποποίηση των εισόδων [A27].

Με την πειραματική εφαρμογή τους όμως και με τη γνωστοποίηση των αποτελεσμάτων, ξεπεράστηκαν αυτοί οι δισταγμοί, με συνέπεια σήμερα πολλές εταιρίες ενέργειας να χρησιμοποιούν τέτοιους αλγόριθμους στα κέντρα ελέγχου ενέργειας.

Γενικότερα τα τεχνητά νευρωνικά δίκτυα δεν προγραμματίζονται, αλλά μαθαίνουν. Φέρουν κάποιες ελεύθερες παραμέτρους που καθορίζουν τη σχέση εισόδου-εξόδου και ανάλογα με παλαιότερα δεδομένα εισόδου-εξόδου (πρότυπα εκπαίδευσης) αυτές προσδιορίζονται, ούτως ώστε για τις συγκεκριμένες εισόδους να δίνουν τις συγκεκριμένες εξόδους . Το σημαντικό εδώ είναι ότι δε χρειάζεται να έχουμε πλήρως κατανοήσει ένα πρόβλημα, ούτως ώστε να αναπτύξουμε τα νευρωνικά δίκτυα. Στην επίλυση ενός προβλήματος, ένα σημαντικό πλεονέκτημα των Τ.Ν.Δ. σε σχέση με τους ανθρώπους, είναι ότι μπορούν να «δουν» πολυδιάστατα πρότυπα εισόδου, τα οποία μπορεί να μην έχουμε τη δυνατότητα να αναγνωρίσουμε και να κατανοήσουμε, λόγω του

γεγονότος ότι έχουμε τρισδιάστατη άποψη του κόσμου. Η ικανότητα των δικτύων αυτών, να μαθαίνουν σχέσεις μεταξύ των δεδομένων δίνει αυξημένες δυνατότητες για την επίλυση προβλημάτων. Το βασικό τους μειονέκτημα είναι όμως ότι δεν μπορούν όμως να ερμηνεύσουν στο πώς έφτασαν στο συμπέρασμά τους, δηλαδή να εξηγήσουν πώς έχουν μάθει να αναγνωρίζουν ένα πρότυπο. Απλώς «κυοφορούν» μέσα τους τη σχέση μεταξύ της πολυδιάστατης εισόδου και εξόδου.

Επίσης ένα άλλο σημαντικό θέμα είναι η δημιουργία αποτελεσματικών συνόλων εκπαίδευσης. Όσον αφορά αυτό το θέμα, θα πρέπει να έχουμε υπόψη μας ότι τα νευρωνικά δίκτυα είναι κατάλληλα για παρεμβολή και όχι για εξαγωγή συμπερασμάτων. Αυτό σημαίνει ότι τα σύνολα εκπαίδευσης πρέπει να αποτελούνται από τον πλήρη χώρο προτύπων που βρίσκεται στη διάθεση μας και τα οποία χρειάζεται να ταξινομούνται και να αναγνωρίζονται σωστά, κατά τη διάρκεια της εκπαίδευσης του δικτύου.

Επιπρόσθετα, πρέπει να τονισθεί ότι ιδιαίτερη βαρύτητα θα πρέπει να δίνεται στην επιλογή του αριθμού των επιπέδων (κρυφών στρωμάτων) και των νευρώνων ανά επίπεδο. Για παράδειγμα, υπερβολικά πολλοί νευρώνες στα κρυφά στρώματα σε ένα δίκτυο εμπρόσθιας τροφοδότησης πολλών στρωμάτων, μπορεί να οδηγήσει σε απομνημόνευση των συνόλων εκπαίδευσης με τον κίνδυνο απώλειας της ικανότητας γενίκευσης. Από την άλλη μεριά, για ένα πολύπλοκο και πολυσύνθετο πρόβλημα, η έλλειψη ικανού αριθμού νευρώνων ενδέχεται να παρεμποδίζει την κατάλληλη ταξινόμηση των προτύπων. Βέβαια, πάντοτε η επιλογή της τοπολογίας του δικτύου και των συνόλων εκπαίδευσης εξαρτάται σε μεγάλο βαθμό από το ίδιο το πρόβλημα.

Τέλος υπάρχει και το ζήτημα του χρόνου εκπαίδευσης, ιδιαίτερα σε περιπτώσεις εφαρμογών πραγματικού χρόνου, όπως είναι και η πρόβλεψη φορτίου. Αυτός ο χρόνος είναι πιθανό να είναι μεγάλος, είτε λόγω της σύνθετης τοπολογίας του δικτύου, είτε λόγω του μεγάλου αριθμού των προτύπων εκπαίδευσης που συνιστούν το σύνολο μάθησης, αλλά όταν το δίκτυο εκπαιδευτεί, η ανταπόκρισή του είναι πολύ γρήγορη.

Στη συνέχεια θα παρουσιάσουμε εκείνες τις μορφές των τεχνητών νευρωνικών δικτύων που βρίσκουν εφαρμογή στη βραχυπρόθεσμη πρόβλεψη φορτίου, όπου έχουν σχεδόν απόλυτη κυριαρχία τα πολυεπίπεδα νευρωνικά δίκτυα με τις διάφορες μορφές του αλγορίθμου της ανάστροφης διάδοσης σφάλματος -Back Propagation algorithm. Συγχρόνως θα μιλήσουμε και για τα προαναφερθέντα βασικά χαρακτηριστικά των νευρωνικών δικτύων.

➤ 2.11 Γενικά περί πολυεπίπεδων τεχνητών νευρωνικών δικτύων για τη βραχυπρόθεσμη πρόβλεψη φορτίου

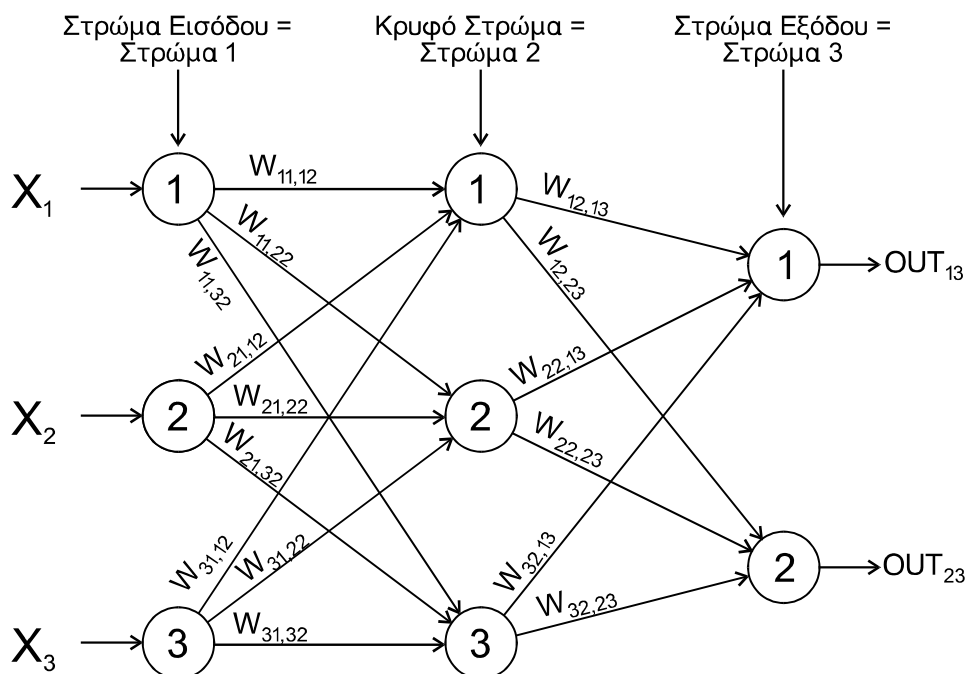
Όπως αναφέραμε και προηγουμένως, τα πλέον ευρέως χρησιμοποιούμενα νευρωνικά δίκτυα είναι τα πολυεπίπεδα ή αλλιώς πολλών στρωμάτων -multi layer perceptrons -MLPs. Τα συγκεκριμένα είναι νευρωνικά δίκτυα εμπρόσθιας τροφοδότησης, που αποτελούνται από ένα στρώμα εισόδου -Input Layer-, ένα ή περισσότερα κρυφά στρώματα -Hidden Layers- και ένα στρώμα εξόδου -Output Layer. Κάθε στρώμα αποτελείται από κόμβους που καλούνται νευρώνες και κάθε νευρώνας συνδέεται με τους νευρώνες του επόμενου στρώματος, μέσω διαφορετικών ελεύθερων παραμέτρων που καλούνται βάρη.

Οι νευρώνες στο στρώμα εισόδου είναι παθητικοί, δηλαδή ο καθένας απλά μεταδίδει μία τιμή, μέσω των διασυνδεδετικών βαρών στους κρυφούς νευρώνες. Οι κρυφοί νευρώνες και οι νευρώνες εξόδου επεξεργάζονται τις εισόδους τους σε δύο βήματα. Κάθε νευρώνας πολλαπλασιάζει κάθε είσοδο με το βάρος που της αντιστοιχεί, προσθέτει όλα αυτά τα γινόμενα και μετά περνάει το άθροισμα που προκύπτει από μια συνάρτηση μεταφοράς, ή αλλιώς ενεργοποίησης, για να παράγει το αποτέλεσμα που δεν είναι άλλο από την έξοδο αυτού. Κάποιοι τύποι συναρτήσεων ενεργοποίησης που χρησιμοποιούνται ευρέως είναι και η σιγμοειδής, η βηματική, η γραμμική, η αναρρίχησης και η γκαουσιανή οι οποίες και παρουσιάζονται στον παρακάτω πίνακα:

$f(x) = 1/(1 + e^{-ax})$ σιγμοειδής με πεδίο τιμών $[0,1]$	$f(x) = \tanh(x)$ σιγμοειδής με πεδίο τιμών $[-1,1]$	$f(x) = a \cdot x$ γραμμική
$f(x) = \begin{cases} a & , \quad x \geq \gamma \\ b & , \quad x < \gamma \end{cases}$ βηματική	$f(x) = \exp\left(\frac{-x^2}{v}\right)$ Γκαουσιανή	$f(x) = \begin{cases} \gamma & , \quad x \geq \gamma \\ x & , \quad x < \gamma \\ -\gamma & , \quad x \leq -\gamma \end{cases}$ αναρρίχησης

Πίνακας 2.11-1 - Συναρτήσεις Ενεργοποίησης

Παρακάτω παρουσιάζουμε ένα πολυεπίπεδο τεχνητό νευρωνικό δίκτυο δομής 3-3-2.



Σχήμα 2.11-2: Πολυεπίπεδο τεχνητό νευρωνικό δίκτυο 3-3-2

Στο προηγούμενο σχήμα το νευρωνικό αυτό δίκτυο των 3 στρωμάτων διαθέτει 3 νευρώνες στο στρώμα εισόδου, 3 νευρώνες στο κρυφό στρώμα και 2 νευρώνες στο στρώμα εξόδου. Αυτό το νευρωνικό δίκτυο συμβολίζεται και ως δίκτυο αρχιτεκτονικής 3-3-2 -από τους αντίστοιχους αριθμούς των νευρώνων στο στρώμα εισόδου, στο κρυφό και στο στρώμα εξόδου. Έχει $3 \times 3 = 9$ συνδέσεις μεταξύ των 3 νευρώνων εισόδου και των 3 κρυφών νευρώνων και $3 \times 2 = 6$ συνδέσεις μεταξύ των 3 κρυφών νευρώνων και των 2 νευρώνων εξόδου. Συνολικά λοιπόν, το δίκτυο έχει 15 συνδέσεις. Σε κάθε σύνδεση αντιστοιχεί και μία τιμή βάρους $w_{pj,ki}$ που ενώνει το νευρώνα p του στρώματος j με το νευρώνα k του στρώματος i .

Υπάρχουν πολλές μέθοδοι που μπορούν να χρησιμοποιηθούν για την εκπαίδευση, δηλαδή τη ρύθμιση των βαρών ενός πολυεπίπεδου νευρωνικού δικτύου. Όμως δεν υπάρχει τρόπος για να προσδιοριστεί εκ των προτέρων, ποια μέθοδος μάθησης λειτουργεί καλύτερα σε μία συγκεκριμένη εφαρμογή. Ένας από τους βασικότερους τρόπους είναι ο αλγόριθμος εκπαίδευσης της ανάστροφης διάδοσης του σφάλματος -Back Propagation-, ο οποίος παρέχει αρκετά ικανοποιητικά αποτελέσματα για ένα μεγάλο πλήθος διαφορετικών προβλημάτων. Ο παραπάνω αλγόριθμος δεν είναι τίποτα άλλο από μια μέθοδο επαναληπτικής ρύθμισης των βαρών, μέχρι να επιτευχθεί η επιθυμητή ακρίβεια στην έξοδο. Κατά την εφαρμογή του στηρίζεται σε μία μέθοδο βελτιστοποίησης και συγκεκριμένα τη μέθοδο ελαχιστοποίησης με χρήση του γενικευμένου κανόνα δέλτα, η οποία εφαρμόζεται σε μία συνάρτηση σφάλματος, όπως είναι το άθροισμα των τετραγώνων των σφαλμάτων.

Άλλες παράμετροι που συμμετέχουν σε πολλές περιπτώσεις στον τελικό προσδιορισμό των βαρών είναι : 1) ο ρυθμός μάθησης , μια σταθερά που ελέγχει το ρυθμό αναπροσαρμογής των βαρών του δικτύου από τα υπολογιζόμενα σφάλματα και 2) ο συντελεστής ορμής , μια σταθερά που επιτρέπει την αξιοποίηση της πληροφορίας από προηγούμενες τιμές σφάλματος, κατά τη διαδικασία ενημέρωσης των βαρών στη διάρκεια μίας ανακύκλωσης, βοηθώντας πρακτικά στην αποφυγή εγκλωβισμού σε κάποιο ενδεχόμενο τοπικό ελάχιστο και όντας σχεδόν πάντοτε μικρότερη από το ρυθμό μάθησης. Η αναδρομική διαδικασία εκπαίδευσης εκτελείται ώσπου να επιτευχθεί ικανοποιητική πρόβλεψη των τιμών εξόδου, ελέγχοντας το άθροισμα των τετραγώνων των σφαλμάτων ή το μέγιστο σφάλμα ή και τα δύο. Η διαδικασία εκπαίδευσης στην πράξη τερματίζεται όταν η τιμή του σφάλματος είναι μικρότερη από μία ορισμένη ανοχή, η οποία έχει καθοριστεί εξ αρχής και η τιμή της εξαρτάται συνήθως από το πρόβλημα. Βέβαια δεν καθίσταται πάντοτε δυνατό η επίτευξη της δεδομένης τιμής ανοχής και γι' αυτό κάποιες φορές αναπροσαρμόζεται και αυτή η τιμή.

➤ 2.12 Εφαρμογές πολυεπίπεδων τεχνητών νευρωνικών δικτύων βραχυπρόθεσμης πρόβλεψης φορτίου

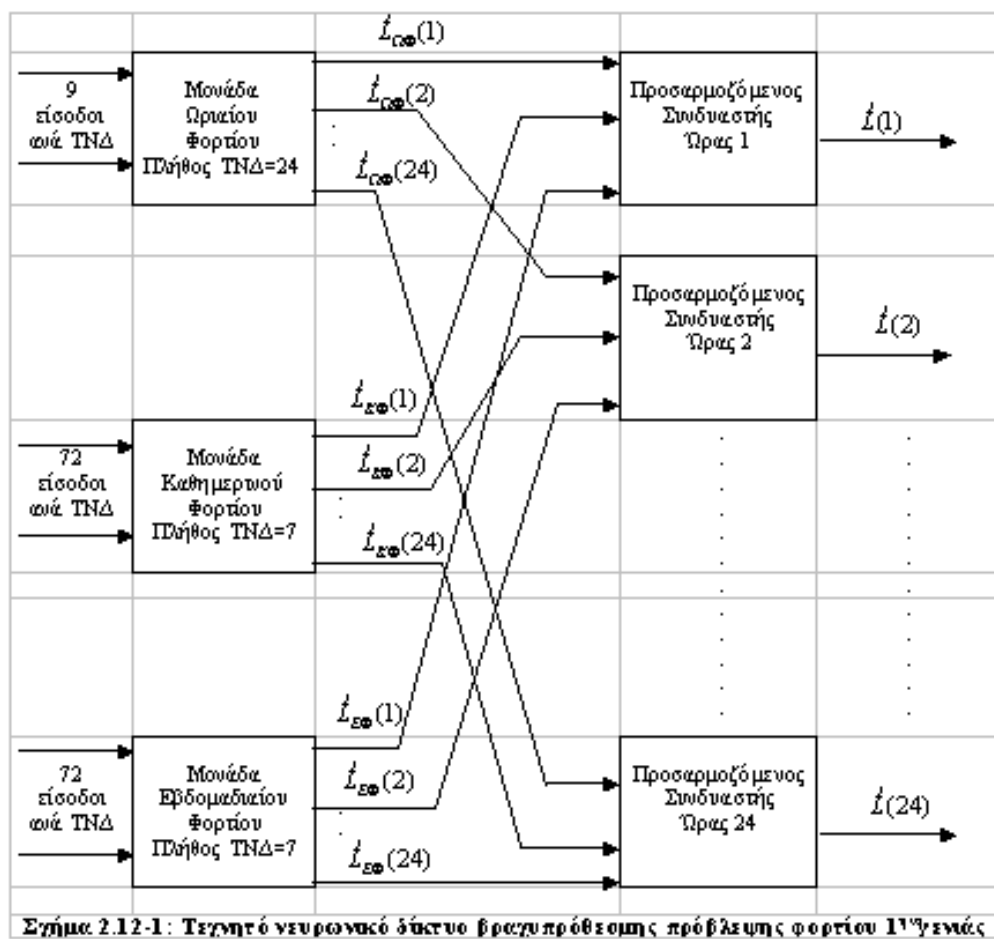
Σε αυτήν την παράγραφο θα παρουσιάσουμε ορισμένες χαρακτηριστικές εφαρμογές των πολυεπίπεδων τεχνητών νευρωνικών δικτύων στη βραχυπρόθεσμη πρόβλεψη φορτίου, κατά τη διάρκεια της τελευταίας δεκαετίας. Μία πολύ ενδιαφέρουσα περίπτωση είναι η περίπτωση της ανάπτυξης νευρωνικού δικτύου και της διαρκούς ανανέωσής του από πανεπιστημιακούς φορείς,

υπό την αιγίδα του αμερικάνικου οργανισμού έρευνας ηλεκτρικής ισχύος -EPRI- [A12], [A28], [A29].

Από το 1992, τόσο στον Καναδά, όσο και στις Η.Π.Α. έχουν αρχίσει να χρησιμοποιούνται στην πράξη συγκεκριμένα μοντέλα νευρωνικών δικτύων βραχυπρόθεσμης πρόβλεψης φορτίου [A28]. Η αρχική δομή αυτών των μοντέλων ήταν ένα πολυεπίπεδο τεχνητό νευρωνικό δίκτυο, με αλγόριθμο εκπαίδευσης αυτόν της ανάστροφης διάδοσης σφάλματος, αποτελούμενο από τρία επίπεδα : το εβδομαδιαίο, το ημερήσιο και το ωριαίο επίπεδο.

Το πρώτο επίπεδο λάμβανε ως δεδομένα τα ωριαία φορτία και τις αντίστοιχες θερμοκρασίες της ίδιας μέρας της τελευταίας εβδομάδας, όπως και τις θερμοκρασίες της προβλεπόμενης μέρας. Το δεύτερο επίπεδο είχε ως δεδομένα τα ωριαία φορτία και τις αντίστοιχες θερμοκρασίες της προηγούμενης μέρας και τις θερμοκρασίες της μέρας υπό πρόβλεψη, ενώ το τρίτο επίπεδο είχε το φορτίο, τη θερμοκρασία και τη σχετική υγρασία της ώρας των τελευταίων τριών περασμένων ημερών, την πρόβλεψη της θερμοκρασίας και της σχετικής υγρασίας για τη συγκεκριμένη μέρα και για τον τύπο της μέρας.

Η βασική μορφή του παρουσιάζεται στο παρακάτω σχήμα :



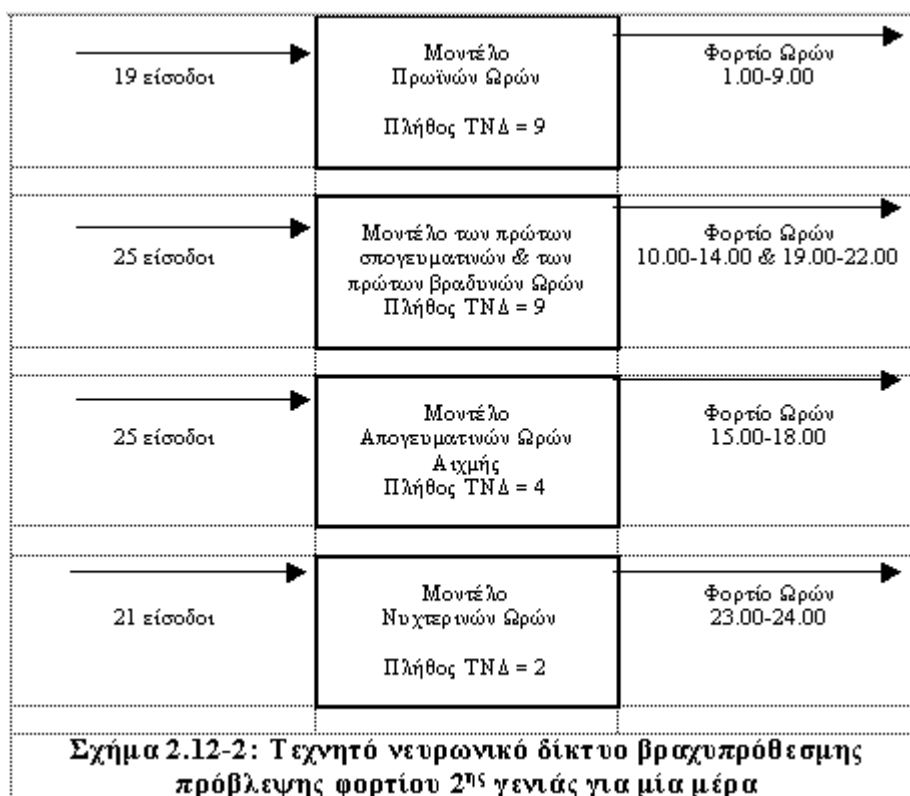
Το πρόβλημα που είχε προκύψει με τη χρήση αυτού του μοντέλου, ήταν κυρίως η αλληλοεπικάλυψη μερικών δεδομένων και γι' αυτό προτάθηκε κατηγοριοποίηση των ωρών της ημέρας. Για παράδειγμα για τις Η.Π.Α. προέκυψαν πέντε κατηγορίες :

α) 1.00-9.00, β) 10.00-14.00, γ) 19.00-22.00, δ) 15.00-18.00 και ε) 23.00-24.00.

Για την κάθε κατηγορία ελήφθησαν χωριστά το φορτίο, η θερμοκρασία, η σχετική υγρασία των δύο τελευταίων ημερών, προ 7 ημερών και οι προβλέψεις της θερμοκρασίας και της υγρασίας για την επόμενη μέρα, καθώς και το είδος της μέρας. Σε κάθε κατηγορία προστέθηκαν επιμέρους στοιχεία ανάλογα με τα βάρη τους, που μπορεί να είναι ένα μεταβαλλόμενο πλήθος ωριαίων θερμοκρασιών ή φορτίων προηγούμενων ωρών ή των αντίστοιχων ωρών προηγούμενων ημερών.

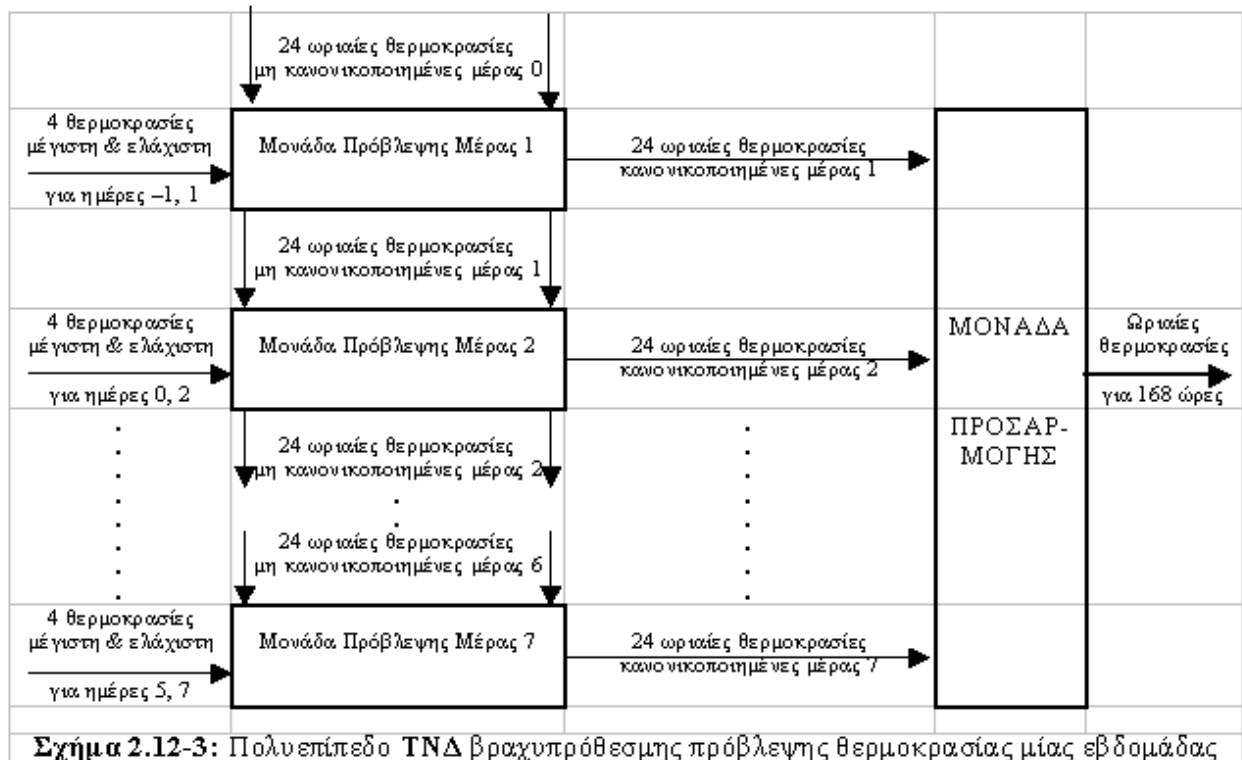
Το αποτέλεσμα που προέκυψε ήταν η σημαντική βελτίωση της ακρίβειας, όσον αφορά την πρόβλεψη που είναι άλλωστε και το ζητούμενο, αλλά και η αύξηση της ταχύτητας της όλης διαδικασίας εκπαίδευσης του νευρωνικού δικτύου. Επίσης οι θερμοκρασίες που πρέπει να προβλέψουμε για την ημέρα που ενδιαφερόμαστε, προσδιορίζονται πάλι μέσω ενός πολυεπίπεδου τεχνητού νευρωνικού δικτύου 28 εισόδων και 24 εξόδων, με την μέθοδο της ανάστροφης διάδοσης σφάλματος- Back Propagation. Με ανάλογο τρόπο φυσικά προκύπτουν και οι σχετικές υγρασίες. Κατά τη διάρκεια λειτουργίας του δικτύου υπάρχει βελτίωση των βαρών μέσω κατάλληλου μηχανισμού του νευρωνικού δικτύου, πραγματοποιώντας μία μικρή εκπαίδευση με τα πιο πρόσφατα δεδομένα.

Παρακάτω παρουσιάζουμε το αντίστοιχο μοντέλο δεύτερης γενιάς :



Όμως, όπως και προαναφέραμε πέραν από την πρόβλεψη του φορτίου, γίνεται και η αντίστοιχη πρόβλεψη της θερμοκρασίας με τη βοήθεια ενός τεχνητού νευρωνικού δικτύου, όπου οι 28 εισοδοί κάθε μονάδας για μία μέρα είναι οι 24 ωριαίες θερμοκρασίες, η μέγιστη και η ελάχιστη

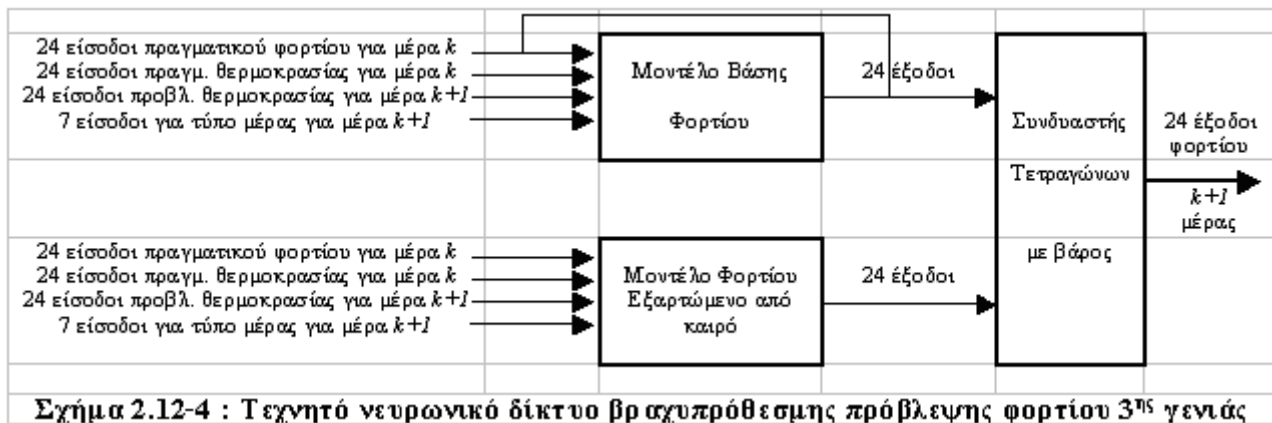
θερμοκρασία της μέρας υπό πρόβλεψη, καθώς και η μέγιστη και η ελάχιστη της μέρας προ 48 ωρών. Ως εξόδους λαμβάνονται 24 ωριαίες προβλέψεις θερμοκρασίας με δύο μορφές, δηλαδή με και χωρίς κανονικοποίηση. Οι κανονικοποιημένες μεταβλητές θερμοκρασίας χρησιμοποιούνται στο προσαρμοστικό μοντέλο για την εύρεση των ωριαίων θερμοκρασιών της συνολικής εβδομάδας, ενώ οι μη κανονικοποιημένες ως είσοδοι στην άλλη βαθμίδα του νευρωνικού δικτύου για την πρόβλεψη της θερμοκρασίας του επόμενου εικοσιτετράωρου.



Κατά την εφαρμογή της παραπάνω μεθόδου, το μέσο απόλυτο σφάλμα προβλεπόμενου ωριαίου φορτίου κυμαίνεται ανάμεσα σε 2%-4%, το αντίστοιχο σφάλμα για την θερμοκρασία από 1,5% ως 2% και της σχετικής υγρασίας από 5,5% μέχρι 6%.

Στη συνέχεια ακολούθησε η τρίτη γενιά τεχνητών νευρωνικών δικτύων βραχυπρόθεσμης πρόβλεψης φορτίου [A12], η οποία αντί για 38 ή 24 νευρωνικά δίκτυα διαθέτει μόλις 2, μέσω των οποίων, το πρώτο προσδιορίζει το φορτίο βάσης και το δεύτερο την αλλαγή του φορτίου βάσης, λαμβάνοντας υπόψη τα εικοσιτέσσερα ωριαία φορτία της τρέχουσας μέρας, τις αντίστοιχες ωριαίες θερμοκρασίες, όπως επίσης τις αντίστοιχες προβλέψεις για τις ωριαίες θερμοκρασίες της επόμενης μέρας, αλλά και το είδος αυτής. Από τα δύο πολυεπίπεδα νευρωνικά δίκτυα που εκπαιδεύονται με την προς τα πίσω διάδοση σφάλματος, προκύπτει ο γραμμικός συνδυασμός αυτών, όπου εφαρμόζεται η μέθοδος των ελαχίστων τετραγώνων με χρήση βαρών, δίνοντας μεγαλύτερη σπουδαιότητα στα σφάλματα των πιο πρόσφατων φορτίων.

Η βασική δομή του παρουσιάζεται ευθύς αμέσως :



Η θερμοκρασία που χρησιμοποιείται ως είσοδος στο δίκτυο είναι η «ενεργός», που συναντήσαμε και προηγούμενα, και η οποία προκύπτει από κατάλληλους γραμμικούς συνδυασμούς της πραγματικής θερμοκρασίας με τη σχετική υγρασία και την ταχύτητα του ανέμου -βλέπε και σχέση (1.4-3) του προηγούμενου κεφαλαίου. Πέραν τούτου οι διακοπές και οι ειδικές μέρες αντιμετωπίζονται με τη χρήση ειδικών αλγορίθμων, όπου μέσω γραμμικών παρεμβολών και προσδιορίζοντας τις τιμές βαρών προκύπτει η τελική τιμή πρόβλεψης του φορτίου, που καθορίζεται είτε από το διαχειριστή του συστήματος ηλεκτρικής ενέργειας, είτε από τη μέρα διακοπών αναπροσαρμόζοντας τη μορφή του φορτίου.

Η βελτίωση που προκύπτει σε σχέση με την προηγούμενη έκδοση-γενιά έχει διαπιστωθεί ότι κυμαίνεται από 9% ως 24%. Εδώ αξίζει απλά να αναφερθεί, ότι έχει βρει πρακτική εφαρμογή σε 35 ηλεκτρικές αμερικάνικες και καναδικές εταιρείες, μέχρι τα τέλη του έτους 1998.

Όπως έχουμε προαναφέρει [A22] η βραχυπρόθεσμη πρόβλεψη φορτίου σε ένα σύστημα ηλεκτρικής ενέργειας, καθορίζει άμεσα και σε πολύ μεγάλο βαθμό την πολιτική της ενεργειακής διαχείρισης. Η μη γραμμικότητα μεταξύ εισόδων και εξόδων, η συσχέτιση ανάμεσα στις μεταβλητές εισόδου περιορίζει τον αριθμό τους και κυρίως αυτά τα μοντέλα δεν είναι ιδιαίτερα εύκαμπτα στις γρήγορες μεταβολές των φορτίων. Εξαιτίας των προαναφερθέντων στην εταιρεία PG&E, που είναι μια γνωστή εταιρεία ηλεκτρισμού και φυσικού αερίου του Ειρηνικού, αναπτύχθηκε ένα τεχνητό πολυεπίπεδο νευρωνικό δίκτυο με ένα κρυμμένο επίπεδο, το οποίο εκπαιδεύεται με τον αλγόριθμο της ανάστροφης διάδοσης σφάλματος, έχοντας 77 εισόδους, 24 κρυμμένους νευρώνες και 24 εξόδους, ώστε να είναι δυνατή η αντιμετώπιση ειδικών γεγονότων, βελτιώνοντας συγχρόνως τις προβλέψεις για τις υπόλοιπες ώρες της μέρας. Οι εισοδοί είναι : δύο μεταβλητές προσδιορισμού της εποχικότητας της υπό πρόβλεψη μέρας, με τη χρήση ζεύγους περιοδικών συναρτήσεων ($\cos(2\pi t/365)$, $\sin(2\pi t/365)$), οι μέγιστες και ελάχιστες θερμοκρασίες έξι περιοχών, οι μέσες μέγιστες και ελάχιστες θερμοκρασίες των αντίστοιχων περιοχών για την τρέχουσα μέρα και προ μίας εβδομάδας, οι αντίστοιχες μέγιστες μεταβολές σε μία μέρα, τα ισοδύναμα πλήθη ψυχρών και θερμών θερμοκρασιακών ημερών, τα ωριαία φορτία των δύο τελευταίων ημερών και η μέση τιμή του φορτίου της περασμένης εβδομάδας. Επιπρόσθετα, δοκιμάστηκαν και άλλες παράμετροι, όπως η υγρασία και η ηλιοφάνεια, αλλά είχαν μικρή επίδραση στο συγκεκριμένο σύστημα, γι' αυτό δεν χρησιμοποιήθηκαν, όπως επίσης δε λήφθηκε υπόψη η ανάπτυξη του φορτίου λόγω της ιδιόμορφης εξέλιξης της ανάπτυξης της εταιρείας.

Ωστόσο αντιμετωπίζεται το πρόβλημα της διαφοροποίησης ανάμεσα στις διακοπές και στις κανονικές μέρες της εβδομάδας.

Η χρονική διάρκεια της πρόβλεψης εκτείνεται σε ορίζοντα μίας εβδομάδας. Τα δεδομένα που χρησιμοποιήθηκαν εκτείνονταν σε επίπεδο ενός έτους και όχι σε λίγες εβδομάδες, διότι τα τελευταία δίνουν ικανοποιητικά αποτελέσματα για εκτελέσεις μεγάλων χρονικών διαστημάτων. Από υλοποίηση του δικτύου και εκτέλεση του αντίστοιχου προγράμματος, λαμβάνουμε την ημερήσια εβδομαδιαία πρόβλεψη με ακρίβεια γύρω στο $2,0 \pm 0,2\%$.

Άλλες εταιρείες, όπως η Florida Power & Light Company [A30], χρησιμοποίησαν πολυεπίπεδα τεχνητά νευρωνικά δίκτυα για την πρόβλεψη φορτίου, λαμβάνοντας υπόψη την ανάπτυξη του φορτίου μέσω της μέσης τιμής των ωριαίων φορτίων ομαλών συνθηκών ανάμεσα στους 55 και στους 85°F ανά 2°F, προσδιορίζοντας ανά περιοχή την τιμή και βλέποντας την αντίστοιχη εξέλιξη. Το φορτίο της επόμενης μέρας προσδιορίζεται από τεχνητό πολυεπίπεδο νευρωνικό δίκτυο με 3 κρυμμένους νευρώνες, 5 ως 15 εισόδους -ανάλογα με το είδος της ημέρας που θεωρείται σταθερό ανά εποχή. Οι βασικές μεταβλητές είναι οι προβλέψεις των θερμοκρασιών, το αντίστοιχο ωριαίο φορτίο της προηγούμενης μέρας, η μέγιστη και η ελάχιστη θερμοκρασία της παρούσας και της προηγούμενης μέρας.

Η προσαρμογή των δεδομένων γίνεται σε ημερήσιο, εβδομαδιαίο και μηνιαίο επίπεδο και στηρίζεται στο νόμο των ελαχίστων τετραγώνων, χρησιμοποιώντας το σύνολο της επανεκπαίδευσης που ορίζεται από τα νέα δεδομένα που συγκεντρώνονται με την πάροδο του χρόνου. Παραδείγματος χάρη σε ημερήσιο επίπεδο για τη δεύτερη μέρα πρόβλεψης, αφού έχει ολοκληρωθεί και έχει εξαντληθεί χρονικά, επιτρέπονται μικροδιορθώσεις των βαρών με βάση την απόσταση των πραγματικών και των προβλεπόμενων φορτίων της πρώτης μέρας που προσδιορίστηκε αρχικά.

Επίσης υπάρχει και ειδική μεταχείριση των αργιών. Συγκεκριμένα γι' αυτές λαμβάνονται δεδομένα για πέντε έτη, έναντι των δύο που χρειάζονται για τις κανονικές μέρες. Αντίστοιχα συμβαίνει και για τις μέρες μεγάλου ψύχους ή ζέστης, δηλαδή ακραίων καιρικών συνθηκών. Η απόδοση του νευρωνικού δικτύου είναι καλύτερη από την αντίστοιχη των παλαιότερων μοντέλων, που βασίζονταν στις χρονοσειρές και σε ένα μεγαλύτερο πλήθος μεταβλητών, ιδιαίτερα στις ειδικές μέρες των αργιών.

Τέλος, ιδιαίτερο ενδιαφέρον παρουσιάζει η εφαρμογή των πολυεπίπεδων νευρωνικών δικτύων με την ανάστροφη διάδοση σφάλματος στη βραχυπρόθεσμη πρόβλεψη φορτίου, για το ελληνικό ηπειρωτικό σύστημα [A31], [Δ2], όπου ως εισόδους έχουμε τα ωριαία φορτία της σημερινής και της χθεσινής μέρας, τη μέγιστη και την ελάχιστη θερμοκρασία για την τρέχουσα μέρα και την πρόβλεψη για την επόμενη μέρα από δύο μετεωρολογικούς σταθμούς, ένας στη βόρεια και ένας στη νότια Ελλάδα. Ακόμη, ως δεδομένα δίνονται για τη μέρα της εβδομάδας ένας δυαδικός επτανήφιος αριθμός, παραδείγματος χάρη για την Τρίτη είναι 0100000, ενώ για την εποχικότητα της μέρας, χρησιμοποιείται το ζεύγος περιοδικών συναρτήσεων ($\cos(2\pi/365), \sin(2\pi/365)$). Τα δεδομένα δε, που χρησιμοποιούμε είναι 6 ετών. Επίσης, η συνάρτηση ενεργοποίησης είναι η υπερβολική εφαπτομένη και το δίκτυο έχει ένα κρυμμένο επίπεδο με πλήθος νευρώνων ίσο με 48.

Όσον αφορά τις γιορτές λαμβάνεται ιδιαίτερη μέριμνα για αυτές και ειδικά αν είναι διαδοχικές. Αυτό ήταν ένα πολύ σοβαρό πρόβλημα στις προβλέψεις φορτίου, που έπρεπε να

αντιμετωπίζεται κατάλληλα και το οποίο λύθηκε τελικά με κατάλληλο, αρχικά, προσδιορισμό του φορτίου μίας κανονικής μέρας και έπειτα με τη μείωση του φορτίου λόγω των διακοπών, λαμβάνοντας υπόψη και το είδος των ημερών που προηγούνται και έπονται της μέρας πρόβλεψης. Συγκεκριμένα κατά την πρόβλεψη φορτίου εορτών $\bar{L}_{holiday}$ θεωρούμε ότι αποτελείται από δύο συνιστώσες : μία συνιστώσα κανονικής μέρας $\bar{L}_{normal}$ και μία διόρθωση της επίδρασης της εορτής $\Delta\bar{L}_{holiday}$, οπότε ισχύει ότι:

$$\bar{L}_{holiday} = \bar{L}_{normal} - \Delta\bar{L}_{holiday} \quad (2.12-5)$$

Όταν στο δiάνυσμα εισόδου του νευρωνικού δικτύου υπάρχουν φορτία εορτών, τότε αυξάνονται κατά ένα διορθωτικό όρο { όρος επίδρασης εορτής }, πριν τροφοδοτηθούν στο δίκτυο. Συνεπώς η απεικόνιση εισόδου-εξόδου του δικτύου, το οποίο είναι εκπαιδευμένο για πρόβλεψη με δεδομένα κανονικών ημερών, θα είναι της μορφής:

$$\bar{L}_{normal} = N(\bar{X} + \Delta\bar{X}_{holiday}) \quad (2.12-6)$$

όπου $\bar{X}$ είναι το δiάνυσμα εισόδου του δικτύου με ιστορικά δεδομένα φορτίου και θερμοκρασίας και $\Delta\bar{X}_{holiday}$ η διόρθωση επίδρασης εορτής για το δiάνυσμα εισόδου, που επηρεάζεται μόνο από τους όρους που σχετίζονται με τα φορτία των εορτών. Σε αυτό το σημείο θα πρέπει να τονιστεί ότι η διόρθωση επίδρασης εορτής υπολογίζεται ξανά, χρησιμοποιώντας ιστορικά δεδομένα προηγούμενων ετών, για την υπό εξέταση εορτή. Αν $\bar{L}_1$ και $\bar{L}_2$ είναι τα σημερινά και τα χθεσινά φορτία, $\bar{T}$ είναι οι θερμοκρασιακές μεταβλητές εισόδου, τότε το δiάνυσμα εισόδου $\bar{X}$ είναι ίσο με:

$$\bar{X} = [\bar{L}_1, \bar{L}_2, \bar{T}] \quad (2.12-7)$$

Αν σήμερα είναι η μέρα γιορτής, ενώ χθες ήταν εργάσιμη, τότε η διόρθωση επίδρασης εορτής για το δiάνυσμα εισόδου θα είναι η εξής:

$$\Delta\bar{X}_{holiday} = [\Delta\bar{L}_{1,holiday}, 0, 0] \quad (2.12-8)$$

$$\text{όπου} \quad \Delta\bar{L}_{1,holiday} = \frac{1}{m} \cdot \sum_{i=1}^m [N(X_1^i) - L_1^i] \quad (2.12-9)$$

και όπου $\{(X_1^i / L_1^i), i = 1 \dots m\}$ είναι τα ιστορικά δεδομένα εισόδου-εξόδου των m περασμένων ετών που αφορούν την συγκεκριμένη γιορτή.

Αν υπάρξουν πολλές συνεχόμενες εορτές, τότε η διαδικασία πρέπει να επαναληφθεί μέχρι να προσδιοριστεί ένα πρότυπο εισόδου-εξόδου (X_i, L_i) με όλα τα φορτία εισόδου από κανονικές μέρες. Με τη νέα μέθοδο τα σφάλματα κατά τη διάρκεια των εορτών μειώνονται κατά 5% για τις

απλές γιορτές και κατά 25% στην περίπτωση των συνεχόμενων γιορτών. Η μέση τιμή του απόλυτου σφάλματος αυξάνεται από 2,35% για την πρώτη μέρα μέχρι και 4% για την έβδομη μέρα της εβδομάδας, αλλά θεωρείται αρκετά ικανοποιητική η συμπεριφορά της σε σχέση με τα γραμμικά κλασικά μοντέλα, μειώνοντας το σφάλμα άνω του 30%.

Από το παραπάνω δίκτυο προέκυψε μία ενδιαφέρουσα παραλλαγή για αυτόνομα συστήματα, όπως της Κρήτης [A32], όπου το τεχνητό νευρωνικό δίκτυο έχει 64 εισόδους -ωριαία φορτία, μέγιστη και ελάχιστη θερμοκρασία τρέχουσας και προηγούμενης μέρας, αντίστοιχες θερμοκρασιακές προβλέψεις επόμενης μέρας, η διασπορά της μέγιστης θερμοκρασίας της τρέχουσας μέρας και της επόμενης σε σχέση με τις θερμοκρασίες κατωφλιού ψύξης και θέρμανσης -17°C και 26°C, μεταβολή της μέγιστης θερμοκρασίας, έκφραση της μέρας του έτους μέσω του ζεύγους συναρτήσεων ($\cos(2\pi t/365)$, $\sin(2\pi t/365)$) και η μέρα της εβδομάδας με επτανήφιο δυαδικό αριθμό. Η συνάρτηση ενεργοποίησης είναι η υπερβολική εφαπτομένη, το πλήθος των κρυφών επιπέδων είναι ένα και ο αριθμός των νευρώνων είναι 48, δηλαδή η υπόλοιπη δομή του δικτύου παραμένει ίδια με του [A31].

Επιπλέον, αξίζει να πούμε ότι δίνεται ιδιαίτερη έμφαση στην ανάπτυξη της μεθοδολογίας πρόβλεψης φορτίου, αν δεν έχουν ολοκληρωθεί οι μετρήσεις της τρέχουσας μέρας. Το πιο σημαντικό στοιχείο εδώ είναι ο έλεγχος των στοιχείων που χρησιμοποιούνται μέσω διαδικασίας κανονικοποίησης των φορτίων και σύγκρισης με τις τυπικές μέρες, ώστε να αντιμετωπιστούν οι πιθανές εσφαλμένες μετρήσεις. Ιδιαίτερα για το δίκτυο της Κρήτης οι διορθώσεις έφθασαν το 2,5% των δεδομένων. Αν χρησιμοποιηθούν τόσο στην εκπαίδευση, όσο και στην πρόβλεψη προεπεξεργασμένα δεδομένα, τότε το σφάλμα της πρόβλεψης φθάνει στο 1,86%. Ως προς τις αργίες εφαρμόζεται η μεθοδολογία του [A31].

Συγκεντρωτικά θα μπορούσαμε να συμπεράνουμε ότι η βραχυπρόθεσμη πρόβλεψη φορτίου με τα κλασικά πολυεπίπεδα τεχνητά νευρωνικά δίκτυα και με τη χρήση του αλγόριθμου εκπαίδευσης της ανάστροφης διάδοσης σφάλματος- Back Propagation-, έχει μέση τιμή απόλυτου σφάλματος η οποία κυμαίνεται από 1,5% μέχρι 2,5%, σε συνεργασία πολλές φορές με απλούς αλγόριθμους διόρθωσης ειδικών περιπτώσεων, είτε πρόκειται για μη ομαλές μέρες, για παράδειγμα αργίες, είτε πρόκειται για εσφαλμένες μετρήσεις.

➤ 2.13 Εφαρμογές ανατροφοδοτούμενων πολυεπίπεδων τεχνητών νευρωνικών δικτύων βραχυπρόθεσμης πρόβλεψης φορτίου

Επειδή το υπό πρόβλεψη ωριαίο φορτίο εξαρτάται σε σημαντικό βαθμό από τα φορτία των προηγούμενων ωρών, το νευρωνικό δίκτυο καθίσταται αναγκαίο να θυμάται εν μέρει τα πρόσφατα ιστορικά στοιχεία. Γι' αυτόν τον λόγο προτάθηκε και η δημιουργία ανατροφοδοτούμενων πολυεπίπεδων νευρωνικών δικτύων.

Σε αυτού του είδους τα δίκτυα ουσιαστικά η ανατροφοδότηση γίνεται από το στρώμα εξόδου προς το στρώμα εισόδου. Δημιουργούνται για κάθε μέρα της εβδομάδας ένα νευρωνικό δίκτυο, με την Τετάρτη, Πέμπτη και Παρασκευή να ταυτίζονται, οπότε ο συνολικός αριθμός τους είναι 5 [A33]. Επίσης η απόδοσή του σε σχέση με το απλό πολυεπίπεδο τεχνητό νευρωνικό δίκτυο δεν είναι σταθερή, καθώς άλλες χρονιές υπερτερεί το ένα και άλλες χρονιές το άλλο μοντέλο. Το γενικό συμπέρασμα είναι ότι το εμπρόσθιο T.N.Δ. αποδίδει καλύτερα, όσο οι χρησιμοποιούμενες

χρονοσειρές πλησιάζουν πολύ τις αντίστοιχες τις ομάδες εκπαίδευσης. Το πότε θα παρουσιασθεί η ανωτερότητα του ανατροφοδοτούμενου Τ.Ν.Δ., εξαρτάται κυρίως από τις χρονοσειρές που εφαρμόζονται. Σε κάθε περίπτωση όμως, η απόδοσή του εξαρτάται από την τελική επιλογή του δικτύου.

Το ανατροφοδοτούμενο πολυεπίπεδο νευρωνικό δίκτυο με την προς τα πίσω διάδοση του σφάλματος στο [A34] επιτρέπει την πρόβλεψη φορτίου για μία εβδομάδα, λαμβάνοντας υπόψη ωριαία φορτία, μέγιστη, ελάχιστη, μέση και ωριαίες θερμοκρασίες, υγρασίες, ταχύτητες ανέμου και συνθήκες ηλιοφάνειας, μέσω ενός κρυμμένου επιπέδου με κατάλληλο αριθμό νευρώνων και με χρήση παντού σιγμοειδών συναρτήσεων ενεργοποίησης. Χρησιμοποιούνται δεδομένα των δύο τελευταίων εβδομάδων με μειονέκτημα να μη λαμβάνει υπόψη του τα κρύα ή τα θερμά κύματα και τις μεταβολές των εποχών.

Πάντως σε κάθε περίπτωση έχουμε ότι με ανατροφοδοτούμενο πολυεπίπεδο Τ.Ν.Δ. βελτιώνεται το μέσο τετραγωνικό σφάλμα, καθώς και η διασπορά του από 2,56% και 2,07% σε 2,02% και 1,86% αντίστοιχα, έχοντας ως αρχικά δεδομένα τα ωριαία φορτία του προηγούμενου εικοσιτετραώρου, το δείκτη της ώρας, τον τύπο της μέρας, τη μέση τιμή της μέγιστης και της ελάχιστης θερμοκρασίας της προβλεπόμενης και της τρέχουσας μέρας για το δίκτυο της Νοτίου Αφρικής. Για άλλα κράτη μπορεί να επιλεγεί άλλο πλήθος μεταβλητών.

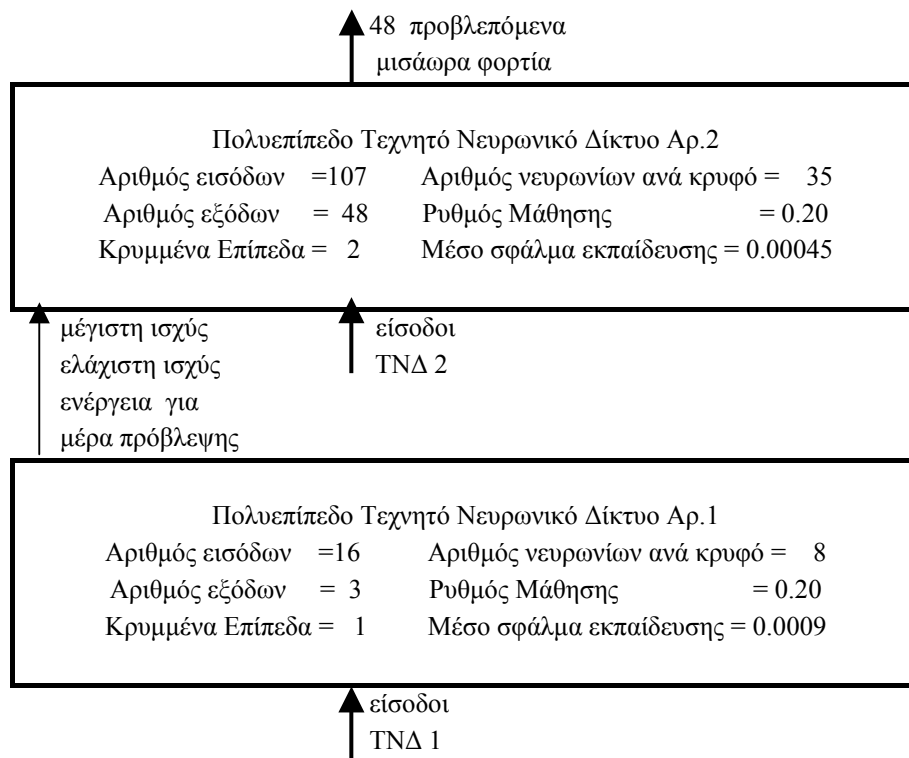
➤ 2.14 Λοιπές μορφές τεχνητών νευρωνικών δικτύων βραχυπρόθεσμης πρόβλεψης φορτίου

Πέραν από τα πλήρη πολυεπίπεδα και ανατροφοδοτούμενα τεχνητά νευρωνικά δίκτυα, έχουν δοκιμασθεί και μη πλήρη δίκτυα, μέσω των οποίων επιδιώκεται κατά κύριο λόγο η μείωση του χρόνου εκπαίδευσης, αλλά και η εξάλειψη του λεγόμενου κινδύνου της αποστήθισης.

Συγκεκριμένα υπάρχει στο [A35] το μη πλήρες πολυεπίπεδο τεχνητό νευρωνικό δίκτυο με αλγόριθμο εκπαίδευσης της ανάστροφης διάδοσης σφάλματος. Σε αυτό το δίκτυο εφαρμόζονται 30 είσοδοι, οι οποίοι περιλαμβάνουν τη μέση θερμοκρασία των τριών τελευταίων ημερών, την ωριαία θερμοκρασία των τριών τελευταίων ωρών της τρέχουσας μέρας, τα ωριαία φορτία των τριών τελευταίων ωρών της τρέχουσας, της προηγούμενης μέρας και τα αντίστοιχα της προηγούμενης εβδομάδας, τα είδη της μέρας και της ώρας με τις αντίστοιχες δυαδικές μεταβλητές. Σαν έξοδο λαμβάνεται το φορτίο της επόμενης ώρας.

Για να προσαρμόζεται εύκολα το μοντέλο και να μειωθεί ο χρόνος εκτέλεσης προτάθηκε η μερική σύνδεση των κρυμμένων κόμβων με τις εισόδους, εφαρμόζοντας τη σιγμοειδή συνάρτηση. Με αυτόν τον τρόπο συσχετίζονται σχεδόν χωριστά οι ομάδες πληροφόρησης -θερμοκρασίες, μέρα εβδομάδας, ώρα μέρας και ωριαία φορτία. Μόνο οι νευρώνες που αντιστοιχούν στα ωριαία φορτία τροφοδοτούνται από όλες τις εισόδους.

Ακολουθεί το σχήμα το οποίο απεικονίζει το παραπάνω μερικώς συνδεδεμένο πολυεπίπεδο Τ.Ν.Δ. :



Σχήμα 2.14-2: Δύο εν σειρά πολυεπίπεδα τεχνητά νευρωνικά δίκτυα

Συγκεκριμένα, για να προβλέψουμε τα ωριαία φορτία του επόμενου εικοσιτετραώρου χρησιμοποιούμε το πολυεπίπεδο τεχνητό νευρωνικό δίκτυο 2, το οποίο έχει ως εισόδους τα φορτία ανά μισή ώρα του τελευταίου εικοσιτετραώρου, τις ωριαίες θερμοκρασίες της τελευταίας μέρας, τις αντίστοιχες υγρασίες, τη μέγιστη και ελάχιστη προβλεπόμενη θερμοκρασία της επόμενης μέρας, την αντίστοιχη μέγιστη υγρασία, όπως επίσης και το είδος της μέρας και του μήνα με τη βοήθεια πέντε ψηφίων. Ακόμα εισέρχονται ως είσοδοι οι έξοδοι του δικτύου 1, που είναι η προβλεπόμενη καταναλισκόμενη ενέργεια της επόμενης μέρας, η μέγιστη και η ελάχιστη ζήτηση φορτίου. Το βοηθητικό νευρωνικό δίκτυο έχει ως εισόδους του την καταναλισκόμενη ενέργεια της μέρας που πέρασε, το συντελεστή ισχύος, τη μέγιστη και την ελάχιστη ζήτηση φορτίου, τα αντίστοιχα μισάωρα του εικοσιτετραώρου που πέρασε, τα όρια διακύμανσης του φορτίου, τις αντίστοιχες θερμοκρασίες εκείνες τις χρονικές στιγμές, τη μέγιστη υγρασία της μέρας και τον τύπο της μέρας με τέσσερα ψηφία. Και στα δύο νευρωνικά δίκτυα εφαρμόζεται ο αλγόριθμος της ανάστροφης διάδοσης σφάλματος.

Με αυτόν τον τρόπο βελτιώνεται η συμπεριφορά του αρχικού δικτύου, που έχει ως έξοδο την πρόβλεψη της επόμενης μέρας, ανά τριάντα λεπτά. Με τη βελτίωση, η μέση τιμή απόλυτου σφάλματος πέφτει στο 2,707% από το 3,367%.

➤ 2.15 Σύγκριση των τεχνητών νευρωνικών δικτύων για το πρόβλημα της βραχυπρόθεσμης πρόβλεψης φορτίου και συμπεράσματα

Βάσει, όλων των προαναφερθέντων θα μπορούσαμε γενικά να συμπεράνουμε από τη σύγκριση των κλασικών μεθόδων με τα τεχνητά νευρωνικά δίκτυα, ότι από τη μία μεριά τα αποτελέσματα των τελευταίων ως προς την πρόβλεψη του φορτίου είναι αρκετά καλύτερα και ακριβέστερα. Σαν παράδειγμα αναφέρουμε το [Δ2] όπου η σύγκριση των T.N.Δ. με μοντέλα ARIMA έδειξε βελτίωση άνω του 50% ως προς τη μέση τιμή του απόλυτου σφάλματος. Από την άλλη πλευρά ένα μειονέκτημα που παρουσιάζει το T.N.Δ. είναι η ευαισθησία, η οποία σχετίζεται με το σύστημα της ηλεκτρικής ενέργειας, δηλαδή άλλη συμπεριφορά έχει ένα νευρωνικό δίκτυο π.χ. για τα ελληνικά κλιματικά δεδομένα και άλλη για τις βορειοανατολικές πολιτείες της Αμερικής, με συνέπεια βέβαια ανά περίπτωση, να χρειάζεται κατάλληλη μερική τροποποίηση των εισόδων του δικτύου.

Όσον αφορά τις εισόδους των τεχνητών νευρωνικών δικτύων για βραχυπρόθεσμη πρόβλεψη φορτίου, χρησιμοποιώντας κατάλληλες μεθόδους και με σύγκριση των στοιχείων ορισμένων ηλεκτρικών εταιρειών, προσδιορίστηκαν στο [A37] οι ανεξάρτητες μεταβλητές που διαμορφώνουν τις χρονοσειρές των φορτίων. Το βασικό συμπέρασμα που προκύπτει από την ανάλυση, είναι ότι επηρεάζονται κατά κύριο λόγο από το φορτίο της προηγούμενης μέρας, κυρίως προ 24 ωρών, και έπειτα από τις θερμοκρασιακές συνθήκες. Γι' αυτό καταλήγουν σε 15 μεταβλητές που είναι αρκετά λιγότερες, τόσο από αυτές που χρησιμοποιούνται στο [A30] - άνω των 500, όσο και από τις 50 με 100 που χρησιμοποιούν συνήθως. Πιο συγκεκριμένα στο [A37] προτείνονται τα ωριαία φορτία της τρέχουσας ημέρας, προ 2, 4, 6 και 8 ωρών και τέλος το είδος του κύκλου της μέρας για την προβλεπόμενη ώρα μέσω των συναρτήσεων $\cos(2\pi t/24)$, $\sin(2\pi t/24)$, $\cos(2\pi t/12)$, $\sin(2\pi t/12)$, αποσιωπώντας βέβαια ότι προβλέπεται μόνο την επόμενη ώρα, έστω με τα ίδια επίπεδα σφάλματος. Πάντως σε αυτήν την περίπτωση γίνεται χρήση πολύ λίγων μεταβλητών με μικρότερο σύνολο στοιχείων, με αποτέλεσμα φυσικά την ταχύτερη εκτέλεση της διαδικασίας.

Τέλος υπάρχουν ορισμένοι [A38] που ισχυρίζονται ότι με βάση την ήδη πλούσια αρθρογραφία που υπάρχει πάνω στο πρόβλημα της βραχυπρόθεσμης πρόβλεψης φορτίου, καθίσταται πλέον δυνατό να διατυπωθούν μία σειρά κανόνων που να προσδιορίζουν τον αριθμό των κρυμμένων επιπέδων του T.N.Δ. – από το θεώρημα του Kromologon προκύπτει ότι είναι ένα-, το είδος της συνάρτησης στο επίπεδο της εισόδου που είναι σιγμοειδής αντί της γραμμικής, το είδος της συνάρτησης στο κρυμμένο επίπεδο που είναι ημιτονοειδής, το είδος της μάθησης που είναι της ανάστροφης διάδοσης σφάλματος με μειούμενο ρυθμό μάθησης και με συντελεστή ορμής κατάλληλων τιμών. Επίσης η εφαρμογή υλοποιείται με προσθήκη λευκού θορύβου Gauss.

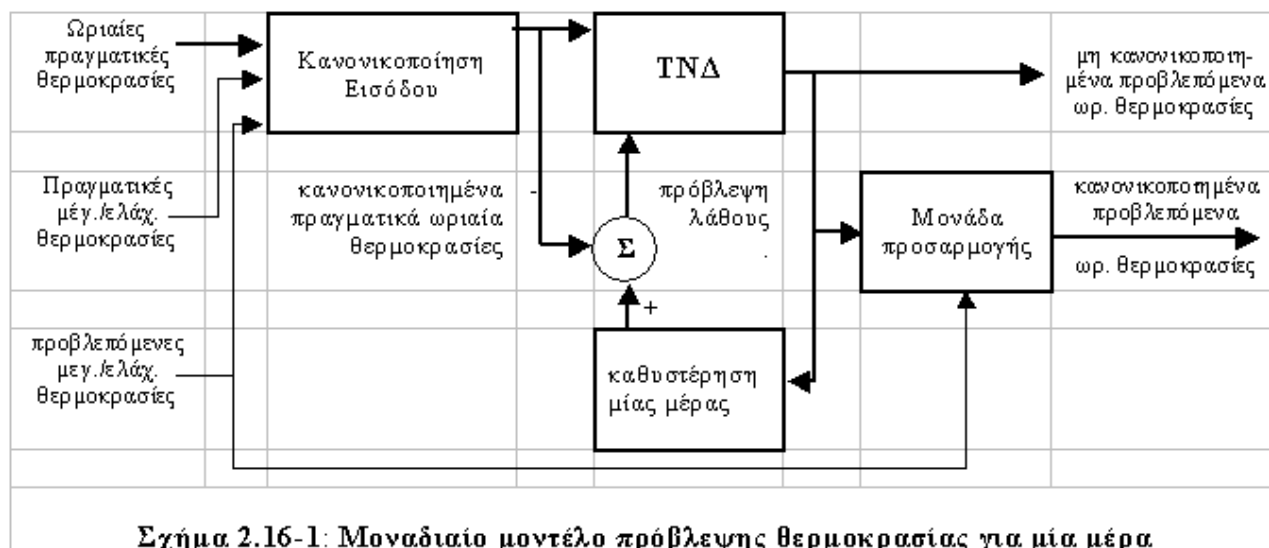
Επιπλέον, πρέπει να υπάρξει η κατάλληλη ανατροφοδότηση του κρυμμένου επιπέδου ως προς την είσοδο, να περιληφθούν και οι ανώμαλες μέρες, ενώ όλο το δίκτυο να επανεκπαιδεύεται τουλάχιστον σε μηνιαίο επίπεδο. Βέβαια οι περιπτώσεις που μελετώνται, δίνουν μέση τιμή απόλυτου σφάλματος κοντά στο 2%, αρκετά μικρότερο από το αντίστοιχο των κλασικών μεθόδων, αλλά σίγουρα σημαντικό ρόλο διαδραματίζει και το πλήθος και η ακρίβεια των δεδομένων, οπότε δεν μπορούν να ληφθούν αυτοί οι κανόνες ως επίλυση του προβλήματος, αλλά ως μία γενική οδηγία.

Στην επόμενη παράγραφο θα αναπτύξουμε την επίδραση της θερμοκρασίας στη βραχυπρόθεσμη πρόβλεψη φορτίου.

➤ 2.16 Επίδραση της θερμοκρασίας στη βραχυπρόθεσμη πρόβλεψη φορτίου με έμφαση στα τεχνητά νευρωνικά δίκτυα

Η θερμοκρασία συνιστά ίσως το βασικότερο μέγεθος μετά τα ωριαία φορτία των περασμένων ωρών, όσον αφορά τη βραχυπρόθεσμη πρόβλεψη φορτίου, ιδιαίτερα η θερμοκρασία της υπό πρόβλεψη ώρας. Η ενδιαφερόμενη ηλεκτρική εταιρεία, στις πιο πολλές περιπτώσεις, ενημερώνεται για τις προβλεπόμενες μελλοντικές θερμοκρασίες από την αντίστοιχη εθνική μετεωρολογική υπηρεσία. Υπάρχουν όμως και άλλες περιπτώσεις, που προχωρά κανείς σε αυτόνομη πρόβλεψη της θερμοκρασίας, όπως στην περίπτωση ενός μεμονωμένου καταναλωτή για τη ρύθμιση των συσκευών του [A16], αλλά και στην περίπτωση που, είτε τα προβλεπόμενα στοιχεία δε διαθέτουν την απαραίτητη ακρίβεια, είτε είναι αρκετά αντιοικονομικό να τα λαμβάνει κανείς σε καθημερινή βάση. Γι' αυτό έχουμε οδηγηθεί αρκετές φορές στην ανάπτυξη ξεχωριστών μοντέλων πρόβλεψης θερμοκρασίας, ακόμη και με τη χρήση νευρωνικών τεχνητών δικτύων.

Ένα παράδειγμα μοντέλου πρόβλεψης της ωριαίας θερμοκρασίας με τη βοήθεια τεχνητών νευρωνικών δικτύων είναι και το παρακάτω. Η πρόβλεψη είναι για μία εβδομάδα, χρησιμοποιώντας επτά ημερήσιες δομές πρόβλεψης ωριαίας θερμοκρασίας, έχοντας ως εισόδους τη μέγιστη και την ελάχιστη θερμοκρασία, τις ωριαίες θερμοκρασίες της προηγούμενης μέρας και την πρόβλεψη της μέγιστης και της ελάχιστης θερμοκρασίας για την υπό πρόβλεψη μέρα, μέσω στοιχείων από μία μετεωρολογική υπηρεσία.



Στο παραπάνω σχήμα ως εξόδους λαμβάνονται οι 24 ωριαίες θερμοκρασίες κανονικοποιημένες, καθώς και οι πραγματικές τους τιμές με βάση τις μέγιστες και τις ελάχιστες προβλεπόμενες θερμοκρασίες κάθε ημέρας.

Με τη βοήθεια αυτού του μοντέλου μία ηλεκτρική εταιρεία δύναται να προβλέψει την ωριαία θερμοκρασία μίας εβδομάδας και να προσδιορίσει με ικανοποιητική ακρίβεια ακόμη και τον ερχομό κυμάτων ψύχους. Τα παλαιότερα απλούστερα μοντέλα πρόβλεψης θερμοκρασίας που χρησιμοποιούσαν τις μέσες τιμές των τελευταίων τριάντα ημερών, δεν είχαν τέτοιες δυνατότητες. Συνεπώς το παραπάνω μοντέλο καθίσταται πολύ χρήσιμο για την πρόβλεψη φορτίου.

Ακόμα θα πρέπει να αναφερθεί ότι έχει μελετηθεί, μέσω των πιθανοτικών μεθοδολογιών Bayes, η επίδραση της θερμοκρασίας στην αβεβαιότητα της πρόβλεψης του φορτίου [A39]. Η προσφορότερη προσέγγιση είναι ο προσδιορισμός της μέγιστης ημερήσιας αιχμής και σε συνδυασμό με ένα τυπικό πρότυπο φορτίου η γραμμική προσαρμογή, για να δημιουργηθεί η τελική μορφή του φορτίου, αρκεί να μην υπάρχουν ακραία καιρικά φαινόμενα. Στη συνέχεια μπορεί να εφαρμοσθεί ένα δυναμικό γραμμικό μοντέλο, το οποίο χρησιμοποιεί ως μεταβλητές προσαρμογής την αιχμή, τη μέση τιμή της θερμοκρασίας και το φορτίο της προηγούμενης μέρας, όπως και τις προβλέψεις για θερμοκρασία και αιχμή για τις επόμενες μέρες. Το βασικό συμπέρασμα που προκύπτει είναι ότι υπάρχει διαφορά στην επίδραση της εσφαλμένης προβλεπόμενης θερμοκρασίας, ανάλογα με την κάθε εποχή του χρόνου, όπως π.χ. το καλοκαίρι είναι δυνατό από τις εκτιμήσεις να οδηγηθούμε σε σφάλματα ακόμα και της τάξης του 30% για ένα μεσαίου μεγέθους σύστημα ηλεκτρικής ενέργειας, εν αντιθέσει με την άνοιξη ή το φθινόπωρο που η επίδραση είναι σημαντικά μικρότερη. Σε κάθε περίπτωση η τελική συμπεριφορά της πρόβλεψης, εξαρτάται και από το μέγεθος του εξαρτώμενου από τον καιρό φορτίου.

➤ 2.17 Επίδραση του ανέμου στη βραχυπρόθεσμη πρόβλεψη φορτίου με έμφαση στα τεχνητά νευρωνικά δίκτυα

Αναμφισβήτητα στη σύγχρονη εποχή η ανάπτυξη μοντέλων πρόβλεψης φορτίου, αλλά και ανέμου διαδραματίζουν σημαντικό ρόλο στο λειτουργικό σχεδιασμό των συστημάτων ηλεκτρικής ενέργειας και ιδιαίτερα σε αυτά που βρίσκονται σε απομονωμένες περιοχές.

Η πρόβλεψη της συμπεριφοράς του ανέμου γίνεται σε ορίζοντα 4 μέχρι και 48 ωρών με χρονικό βήμα των είκοσι λεπτών ως μία ώρα, καθορίζοντας αντίστοιχα και την πρόβλεψη φορτίου. Η πρόβλεψη του ανέμου είναι ιδιαίτερα σημαντικός παράγοντας για τη λειτουργία του συστήματος ηλεκτρικής ενέργειας, ιδιαίτερα στην περίπτωση υψηλής διείσδυσης απομονωμένων ανεμογεννητριών ή αιολικών πάρκων, καθώς μπορεί να μειωθεί το μέγεθος της εφεδρείας, να υπάρξουν πρόσθετες μειώσεις καυσίμων, να ρυθμιστούν σε πιο αποδοτικά επίπεδα οι γεννήτριες, να υπάρξει η μείωση των συμβατικών ορίων λειτουργίας, να μεταβληθεί ο αριθμός ένταξης ή λήξης λειτουργίας των μονάδων και να αυξηθεί η αξιοπιστία του συστήματος .

Είναι απαραίτητο να διακριθούν οι διάφορες μέθοδοι, τόσο ως προς τη μεθοδολογία, όσο και ως προς τη χρονική κλίμακα -που είναι από μερικά δευτερόλεπτα ως και μία ώρα για προβλέψεις μερικών λεπτών ως 48 ωρών- [A19]. Για βραχυπρόθεσμη πρόβλεψη της ταχύτητας του ανέμου, δηλαδή μέχρι 10 λεπτά, μπορούν να χρησιμοποιηθούν τεχνικές φίλτρων Kalman [A40] δίνοντας ιδιαίτερα ικανοποιητικά αποτελέσματα μέχρι χρόνους ενός λεπτού, αλλά από 10 λεπτά και άνω κρίνονται ακατάλληλα [A41]. Επίσης αναπτύχθηκαν πολυεπίπεδα νευρωνικά δίκτυα με προς τα πίσω διάδοση σφάλματος, με βήμα 1 sec για χρόνους ενός λεπτού και με ικανοποιητικά

αποτελέσματα [A42], ενώ δημιουργήθηκαν ανατροφοδοτούμενα και ασαφή νευρωνικά δίκτυα με ικανοποιητικά αποτελέσματα [A43]. Παράλληλα υπήρξαν και μοντέλα για τη λειτουργία συστημάτων ηλεκτρικής ενέργειας με πετρελαιοκινητήρες και ανεμογεννήτριες, για βραχυπρόθεσμη πρόβλεψη [A44].

Θα πρέπει να τονιστεί ότι σε απομονωμένα συστήματα, όπως της Κρήτης ή της Λήμνου [A47], η πρόβλεψη του φορτίου είναι αρκετά δύσκολη και δεν επιτυγχάνεται ο στόχος της μέσης τιμής του απόλυτου σφάλματος 3%. Αυτό γίνεται ακόμη δυσκολότερο από τη μεγάλη διείσδυση αιολικών πάρκων, που περιορίζουν το χρονικό διάστημα πρόβλεψης και αυξάνουν το αντίστοιχο σφάλμα λόγω του σφάλματος στον προσδιορισμό της ταχύτητας του ανέμου, που κινείται στα επίπεδα του 10% με 20%, περιορίζοντας σημαντικά τις δυνατότητες πρόβλεψης της συνολικής ζήτησης.

Παράλληλα για να βελτιωθεί η πρόβλεψη, μπορεί να απαιτείται η αγορά οργάνων ή η πραγματοποίηση συνεχών μετρήσεων, που είναι αρκετά δαπανηρές, με συνέπεια να τίθεται το πρόβλημα σε οικονομικοτεχνικό επίπεδο. Εξαιτίας τούτου, είναι αναγκαία η αξιολόγηση ως προς το κόστος των διαφόρων μεθόδων πρόβλεψης. Στο [A46] αναπτύσσεται μεθοδολογία αξιολόγησης των κερδών από τη χρήση βελτιωμένων μοντέλων πρόβλεψης φορτίου και ταχύτητας του ανέμου, δίνοντας ιδιαίτερη έμφαση στην περίπτωση των αυτόνομων συστημάτων. Υπάρχει ένα ενιαίο πρόγραμμα για τη λειτουργία του συστήματος ηλεκτρικής ενέργειας και την κοστολόγηση των σχετικών δαπανών που εκτελείται σε πραγματικό χρόνο, λαμβάνοντας υπόψη την αβεβαιότητα των προβλέψεων.

Σημαντικό ρόλο παίζει η μορφή και η δομή του δικτύου, καθορίζοντας σε μεγάλο βαθμό την ένταξη των βασικών μονάδων, το είδος και το πλήθος των εφεδρικών μονάδων. Γι' αυτό στην Κρήτη από τη μετάβαση της απλής πρόβλεψης στην τέλεια, με διείσδυση των ανεμογεννητριών κατά 20%, υπάρχει βελτίωση του κόστους μόλις κατά 1,5%, χωρίς να λάβουμε υπόψη μας το παραπάνω κόστος της τέλει πρόβλεψης [A48]. Πλέον το πρόβλημα που τίθεται είναι, αν συμφέρει οικονομικά την ηλεκτρική επιχείρηση να βελτιώσει την πρόβλεψή της, εφαρμόζοντας ακόμη και προγράμματα διαχείρισης φορτίου -που θα αναφερθούμε ακολούθως- και ξεφεύγοντας από τα στενά όρια της πρόβλεψης φορτίου.

➤ 2.18 Επίδραση της διαχείρισης φορτίου στη βραχυπρόθεσμη πρόβλεψή του

Οι περισσότερες ηλεκτρικές εταιρείες εφαρμόζουν προγράμματα διαχείρισης φορτίου, κατά την προσπάθειά τους να ελέγξουν αποτελεσματικότερα τη ζήτηση. Εδώ περιλαμβάνονται κυρίως η αποκοπή της ημερήσιας αιχμής, το γέμισμα των κοιλάδων και οι διάφορων ειδών μεταφορές φορτίου, είτε μόνιμες -αλλαγή χρήσης ενέργειας-, είτε περιοδικές -μετατόπιση ημερήσιου φορτίου κατά τη διάρκεια της νύχτας. Ο έλεγχος των συσκευών κάθε κατηγορίας καταναλωτών πραγματοποιείται, δίνοντας κυρίως τιμολογιακά κίνητρα.

Σ' αυτό το σημείο η βραχυπρόθεσμη πρόβλεψη φορτίου είναι απαραίτητη για τη ρύθμιση και τον έλεγχο συσκευών, όταν συνδυάζεται με τα ηλεκτρικά τιμολόγια [A5]. Η σχέση που δημιουργείται είναι αμφίδρομη, διότι και επηρεάζεται μέσω των «χαμηλών» ωρών κόστους τη ρύθμιση των συσκευών, π.χ. ο θερμοσίφωνας θα λειτουργήσει τη νύχτα, και επηρεάζεται, αφού

μεταβάλλεται η συμπεριφορά των καταναλωτών, π.χ. αύξηση του νυχτερινού φορτίου και μείωση του ημερήσιου. Γι' αυτό το λόγο έχει γίνει μία συστηματική προσπάθεια μοντελοποίησης της συμπεριφοράς ορισμένων οικιακών συσκευών, όπως του θερμοσίφωνα [A49] και των κλιματιστικών [A50].

Αρχικά το πρόβλημα της διαχείρισης αντιμετωπίστηκε με απλές κλασικές μεθοδολογίες, όπως στο [A51] εφαρμόστηκε η μέθοδος των ελαχίστων τετραγώνων για τον προσδιορισμό της συμπεριφοράς των αντλιών θερμότητας -ηλεκτρικών και φυσικού αερίου-, μέσω κατάλληλου δείγματος σε σχέση με τον αριθμό των θερμών και των ψυχρών ημερών, την οικονομική κατάσταση, το πλήθος των ανθρώπων στο σπίτι, την ηλικία του οικήματος, τις θερμικές απώλειες του, το κόστος της μονάδας φυσικού αερίου και της kWh κτλ.. Σε άλλη περίπτωση [A50] η πρόβλεψη γίνεται μέσω πολλαπλασιαστικών εποχιακών μεθόδων εξομάλυνσης σταθερού επιπέδου, διαχειριζόμενοι τα κλιματιστικά με βάση τη θερμοκρασία, ώστε σε χρόνο ενός εικοσιτετραώρου να κλείνει είτε η διαδικασία ψύξης των μεγάλων δεξαμενών αποθήκευσης παγοκολώνων κατά το καλοκαίρι, είτε η διαδικασία θέρμανσης των αντίστοιχου μέσου, κατά τη διάρκεια του χειμώνα. Αυτό γίνεται σε συνδυασμό με τις τιμές της ενέργειας και της ισχύος, ώστε ο πελάτης να μεταφέρει την αιχμή του σε κατάλληλη περίοδο της μέρας. Για την εξασφάλιση των ενεργειακών θερμικών ή ψυκτικών αναγκών του καταναλωτή χρησιμοποιείται συντελεστής ασφαλείας.

Ένας άλλος τρόπος πρόβλεψης υλοποιείται, μέσω μίας απλοποιημένης οικονομετρικής μεθόδου που παρουσιάζεται στο [A52]. Μία από τους βασικότερες επιθυμίες μίας ηλεκτρικής εταιρείας σε σχέση με τη λειτουργία του ηλεκτρικού συστήματος, είναι η ομαλοποίηση -αν όχι η επιπεδοποίηση- της χρονολογικής καμπύλης φορτίου, ώστε να γίνει πιο εύκολη η αντιμετώπιση της ζήτησης. Η συμπεριφορά επηρεάζεται από τα είδη των καταναλωτών. Γι' αυτό είναι απαραίτητα σύμφωνα με την οικονομετρική μέθοδο η ανάλυση σε ανεξάρτητα στοιχεία του προγράμματος διαχείρισης ηλεκτρικού φορτίου, η πρόταση αρκετών εναλλακτικών λύσεων και η ανάλυση των φορτίων σε ωριαία βάση, λαμβάνοντας υπόψη την ετήσια και την εποχιακή συμπεριφορά τους.

Συγχρόνως πραγματοποιείται και η οικονομική αξιολόγηση του προγράμματος διαχείρισης φορτίου, ώστε να διαπιστωθεί κατά πόσο οικονομικά συμφέρουσα είναι μία τέτοια επένδυση. Χρειάζονται πληροφορίες για το σύστημα του φορτίου, τις κατηγορίες των πελατών, τις ενεργειακές καταναλώσεις, την αναγνώριση των ειδών των τελικών χρήσεων. Σημαντική προσπάθεια καταναλίσκεται στην εύρεση των χαμένων δεδομένων. Συνήθως διατίθενται η συνολική περίοδος μέτρησης ανά κατηγορία, οι καθημερινές καμπύλες φορτίου στα βασικά σημεία του συστήματος διανομής και περιορισμένα ποσοτικά δεδομένα, σχετικά με τη χρήση συσκευών και κατηγοριών τελικής χρήσης φορτίου. Στη συνέχεια ανάλογα με την τελική χρήση και το είδος της καμπύλης διαμορφώνεται η τελική μορφή, ρίχνοντας ιδιαίτερο βάρος στα εποχιακά ευαίσθητα φορτία. Ο αλγόριθμος της οικονομετρικής μεθόδου -που αναλύεται στο [A52]- αναπτύχθηκε με βάση τα ωριαία φορτία μίας περιόδου από 24 ώρες ως ένα έτος, τους μηνιαίους συντελεστές διόρθωσης και τα υπόλοιπα στοιχεία, σε σχέση με το σημείο αναφοράς -υποσταθμός, τροφοδότης- και τελικά με ένα πλήθος εναλλακτικών σεναρίων για κάθε χρήση, καταλήγει κανείς στη διαμόρφωση του προγράμματος διαχείρισης φορτίου και στις επιπτώσεις του σε ωριαίο επίπεδο.

➤ 2.19 Η σημασία του κόστους στη βραχυπρόθεσμη πρόβλεψη φορτίου

Σύμφωνα με τα όσα έχουμε προαναφέρει, η βραχυπρόθεσμη πρόβλεψη φορτίου επηρεάζεται σε σημαντικό βαθμό από τις αντίστοιχες προβλέψεις θερμοκρασίας, ταχύτητας και διεύθυνσης ανέμου -που μέσω των ανεμογεννητριών επιδρά στην παραγωγή-, όπως επίσης και από τα εφαρμοζόμενα προγράμματα διαχείρισης φορτίου σε συνδυασμό με την αντίστοιχη τιμολογιακή πολιτική [A16]. Όμως οι βασικοί λόγοι υλοποίησης της πρόβλεψης της ζήτησης, είναι από τη μία πλευρά η κάλυψη των αναγκών των πελατών, επιδιώκοντας την ομαλή λειτουργία του ηλεκτρικού συστήματος και από την άλλη η μεγιστοποίηση των καθαρών κερδών της ηλεκτρικής επιχείρησης, υπακούοντας βέβαια στους εθνικούς νόμους και εξασφαλίζοντας την απαραίτητη κοινωνική πολιτική, πχ. την ηλεκτροδότηση απομονωμένων ακριτικών νησιών κτλ.. Εξαιτίας τούτων είναι αναγκαίο να προσδιορίσουμε τις οικονομικές επιπτώσεις της ανακριβούς πρόβλεψης φορτίου.

Η βραχυπρόθεσμη πρόβλεψη φορτίου επηρεάζει το πρόγραμμα λειτουργίας των γεννητριών, τις αγορές καυσίμων, το πρόγραμμα συντήρησης και την ανάλυση ασφάλειας του συστήματος. Η ακρίβεια που πετυχαίνει κανείς με τις διάφορες μεθοδολογίες είναι 2% με 5%, αλλά η πραγματική απαίτηση καθορίζεται από την οικονομική επίδραση. Ο υπολογισμός που παρουσιάζεται στο [A53] βασίζεται στην εφαρμογή της μεθόδου Monte-Carlo. Τα βασικά βήματα είναι τα εξής :

α) Ο ένας χρόνος διαιρείται σε 12 μήνες.

β) Για κάθε μήνα γίνεται ο προσδιορισμός των διαθέσιμων γεννητριών λαμβάνοντας υπόψη τους διάφορους περιορισμούς των γεννητριών και του συστήματος

γ) Τα ωριαία πραγματικά φορτία είναι διαθέσιμα.

δ) Χρησιμοποιείται το πρόγραμμα ένταξης μονάδων για το βέλτιστο ανά μέρα προγραμματισμό.

ε) Υποτίθεται ότι το ωριαίο προβλεπόμενο φορτίο είναι ίσο με το αντίστοιχο πραγματικό και έναν τυχαίο αριθμό που προβλέπεται με γνωστή κατανομή πιθανότητας, όπως η ομοιόμορφη ή κανονική.

στ) Το πρόγραμμα ένταξης και ρύθμισης μονάδων γίνεται σε μοναδιαίο επίπεδο χρησιμοποιώντας τα προβλεπόμενα ωριαία φορτία. Με αυτόν τον τρόπο προσομοιώνεται η πραγματική λειτουργία που επιτελούν οι χειριστές του συστήματος. Συγχρόνως γίνεται και ο υπολογισμός του συνολικού κόστους, λαμβάνοντας υπόψη τα λειτουργικά κόστη, τα κόστη ένταξης ή σταματήματος μονάδων, ποινές για μη τήρηση συμφωνιών και αγορές για τη συνάντηση της ζήτησης φορτίου.

ζ) Υπολογίζεται η οικονομική ζημιά από τη διαφορά του προκύπτοντος κόστους και του κόστους με ιδανική πρόβλεψη.

η) Τα τρία τελευταία βήματα επαναλαμβάνονται πολλές φορές για τη στατιστική ανάλυση του προβλήματος. Όσο πιο μεγάλο είναι το πλήθος, τόσο πιο ακριβής θα είναι η απάντηση.

θ) Με την εύρεση των απαραίτητων δειγμάτων και την επανάληψη της διαδικασίας όσες φορές είναι απαραίτητη, προκύπτουν τόσο τα μηνιαία στοιχεία, όσο και τα ετήσια, όπου η μέση ετήσια

τιμή είναι το άθροισμα των 12 μηνιαίων μέσων τιμών και η αντίστοιχη τυπική απόκλιση η ρίζα του αθροίσματος των τετραγώνων των αντίστοιχων μηνιαίων τυπικών αποκλίσεων.

Στη συνέχεια πραγματοποιήθηκε μία εφαρμογή του μοντέλου για ένα υποθετικό σύστημα μεσαίου μεγέθους 20 θερμικών μονάδων, θεωρώντας τες όλες διαθέσιμες, ενώ η εφεδρεία του συστήματος είναι ίση με το 10% του φορτίου. Επίσης περιλαμβάνεται μία διαδικασία «αγορών» που καλείται είτε να καλύψει τις ανάγκες ενέργειας με αγορές, όταν υπάρχει έλλειψη ισχύος, είτε να πουλήσει ισχύ, όταν το άθροισμα των πιθανών γεννητριών του συστήματος είναι μεγαλύτερο από τη ζήτηση του συστήματος.

Το γενικό συμπέρασμα είναι ότι όσο αυξάνεται το πραγματικό φορτίο και όσο αυξάνεται το σφάλμα πρόβλεψης φορτίου, τόσο αυξάνουν αντίστοιχα η μέση τιμή και η τυπική απόκλιση του επιπλέον κόστους, λόγω της ανακρίβειας της πρόβλεψης. Βέβαια σημαντικό ρόλο παίζουν και οι παράμετροι της εφεδρείας και της διαθεσιμότητας των διαφόρων μονάδων, που στην παρούσα φάση αγνοούνται. Ομοίως αποδεικνύεται ότι όσο αυξάνεται ο συντελεστής χρησιμοποίησης του συστήματος ή όσο αυξάνεται η ανακρίβεια γνώσης της αιχμής του φορτίου -που παίζει πιο σημαντικό ρόλο από τα άλλα ωριαία φορτία-, τόσο αυξάνεται η μέση τιμή και η τυπική απόκλιση της διαφοράς κόστους, λόγω της έλλειψης ακρίβειας στην πρόβλεψη του φορτίου.

Επίσης αναπτύχθηκαν και μεθοδολογίες αξιολόγησης των κερδών από τη χρήση βελτιωμένων μοντέλων πρόβλεψης φορτίου και ταχύτητας του ανέμου, όπως αυτής του [A46], την οποία αναλύσαμε στην παράγραφο 2.17. Βέβαια στην αντίστοιχη εφαρμογή για το νησί της Κρήτης δεν υπήρξε σημαντική μείωση του κόστους, γεγονός που αποδεικνύει τη σημασία που έχει σε κάθε περίπτωση η δομή και η ανάπτυξη του κάθε ηλεκτρικού συστήματος, όπως και τα αντίστοιχα κλιματικά στοιχεία μίας περιοχής, ως προς τη συμπεριφορά των μεθοδολογιών πρόβλεψης φορτίου.

➤ 2.20 Οι τελευταίες δυνατότητες βελτιστοποίησης των μεθοδολογιών για τη βραχυπρόθεσμη πρόβλεψη φορτίου και συμπεράσματα γενικότερης υφής

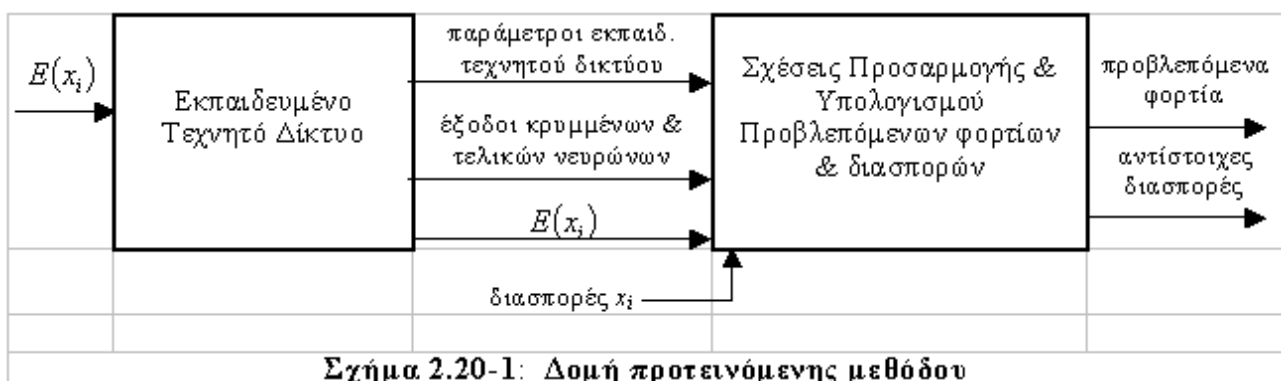
Κατά τα τελευταία δύο έτη έχουν προταθεί ορισμένες βελτιώσεις, ώστε να αυξήσουν την ακρίβειά τους οι μέθοδοι βραχυπρόθεσμης πρόβλεψης φορτίου, επιβεβαιώνοντας τη στροφή προς τα νευρωνικά δίκτυα.

Συγκεκριμένα στο [A54] προτείνεται η κατάλληλη επεξεργασία και ομαδοποίηση των δεδομένων για τη βελτίωση της συμπεριφοράς των τεχνητών νευρωνικών δικτύων, στη βραχυπρόθεσμη πρόβλεψη φορτίου. Αναλυτικότερα για κάθε μέρα προσδιορίζεται η μέση, η ελάχιστη, η μέγιστη θερμοκρασία και για κάθε ώρα, μέσω της διασποράς και της μέσης τιμής υπολογίζεται η ισοδύναμη θερμοκρασία. Στη συνέχεια για κάθε μέρα θεωρείται ένα εξαδιάστατο διάνυσμα -σημερινής και προηγούμενης μέρας- και διαμορφώνονται κατάλληλες γειτονιές με βάση αυτά τα ιστορικά στοιχεία, χωρίς να λαμβάνονται υπόψη οι ιδιαίτερες μέρες και προσδιορίζεται για την ημέρα «στόχο» η απόστασή της από την κάθε γειτονιά. Για κάθε τύπο μέρας υπάρχει το αντίστοιχο νευρωνικό δίκτυο και μέσω κατάλληλης εκπαίδευσης του δικτύου προσδιορίζεται ο αριθμός των κρυμμένων νευρώνων, ώστε να ελαχιστοποιηθεί το σφάλμα, έχοντας χωρίσει τα στοιχεία της κάθε ομάδας σε δύο υποομάδες -για εκπαίδευση και δοκιμή. Τελικά μπορούμε να προχωρήσουμε σε προβλέψεις ως 120 ώρες μπροστά με σφάλματα, των οποίων η μέση τιμή δεν

ξεπερνά το 2,5%. Βέβαια σε κάθε περίπτωση το αποτέλεσμα εξαρτάται από τα δεδομένα της κάθε εταιρείας. Η μέθοδος στηρίζεται στην εύρεση του κ-οστού γείτονα μέσω κατάλληλων διανυσμάτων και αποστάσεων, λαμβάνοντας υπόψη τα αντίστοιχα δεδομένα έτσι, ώστε η μέρα που μοιάζει να μπαίνει στην αντίστοιχη υποομάδα εκπαίδευσης.

Τέλος έχει συστηματοποιηθεί η μελέτη των επιπτώσεων των πιθανοτικών εισόδων στα τεχνητά νευρωνικά δίκτυα για την πρόβλεψη ηλεκτρικού φορτίου [A55]. Συγκεκριμένα συνήθως λαμβάνονται υπόψη τα φορτία της προηγούμενης μέρας, οι αντίστοιχες θερμοκρασίες και οι προβλεπόμενες θερμοκρασίες της μέρας που θέλουμε την πρόβλεψη.

Στην παρούσα περίπτωση γίνεται λόγος για την πρόβλεψη της ημερήσιας αιχμής. Συγκεκριμένα η μέθοδος χωρίζεται σε δύο τμήματα : α) στο αρχικό νευρωνικό δίκτυο που εκπαιδεύεται από πραγματικά ιστορικά στοιχεία και β) στο προηγούμενο δίκτυο, όπου χρησιμοποιώντας μία ομάδα εξισώσεων προσδιορίζονται τα προβλεπόμενα φορτία και η διασπορά τους.



Από την εκτέλεση του αντίστοιχου προγράμματος προέκυψε με τη χρήση στοιχείων ενός έτους για εκπαίδευση και ενός έτους για έλεγχο, ότι η μέση τιμή του απόλυτου σφάλματος είναι ίση με 2,21%, ενώ αν αγνοηθούν οι θερμοκρασίες γίνεται ίσο με 4,87%. Σε κάθε περίπτωση λήφθηκαν 100 διαφορετικές ψευδοτιμές, οπότε με βάση αυτά τα στοιχεία προσδιορίζει κανείς τα αντίστοιχα ζητούμενα μεγέθη της τροποποιημένης μεθόδου, παρουσιάζοντας βελτίωση έναντι της κλασικής και με τις τρεις πιθανοτικές συναρτήσεις, δίνοντας σαφές προβάδισμα στην κανονική μέθοδο.

Σε κάθε περίπτωση όμως διαπιστώνουμε ότι τα πολυεπίπεδα τεχνητά νευρωνικά δίκτυα έχουν καθιερωθεί στη βραχυπρόθεσμη πρόβλεψη φορτίου, λόγω των καλύτερων αποτελεσμάτων που δίνουν σε σχέση με τις κλασικές μεθοδολογίες και λόγω της απλούστερης -εννοιολογικά- δομής τους, σε σχέση με τις υπόλοιπες μορφές τεχνητής νοημοσύνης.

Τα σφάλματα, όπως μέση τιμή απόλυτου σφάλματος και η ρίζα της μέσης τιμής των τετραγώνων των σφαλμάτων, είναι της τάξης του 2% και μικρότερα. Όμως η τελική τιμή δεν εξαρτάται μόνο από τη δομή του νευρωνικού δικτύου -αρχιτεκτονική δικτύου, πλήθος εισόδων, κρυμμένα επίπεδα, μέθοδος εκπαίδευσης κτλ.-, αλλά και από το ίδιο το ηλεκτρικό σύστημα και τις κλιματολογικές συνθήκες της περιοχής που καλύπτει.

Παρακάτω θα αναφερθούμε συνοπτικά στην πολύ βραχυπρόθεσμη, μεσοπρόθεσμη, αλλά και μακροπρόθεσμη πρόβλεψη φορτίου.

➤ 2.21 Πολύ Βραχυπρόθεσμη Πρόβλεψη Φορτίου

Η πολύ βραχυπρόθεσμη πρόβλεψη φορτίου -very short-term load forecasting- είναι η πρόβλεψη της ζήτησης του φορτίου για τα επόμενα 30 λεπτά ως μία ώρα, με βήματα μερικών δευτερολέπτων ως λεπτού. Έχει ως στόχο τη ρύθμιση των βασικών επιπέδων λειτουργίας των μονάδων, όπως επίσης διαδραματίζει σημαντικό ρόλο στη διαμόρφωση της σειράς ένταξής τους με βάση τα κριτήρια της οικονομικής κατανομής και λαμβάνοντας υπόψη τα αντίστοιχα προγράμματα συντήρησης. Βασικά διαδραματίζει σημαντικό ρόλο στην ενεργειακή διαχείριση του συστήματος, ενώ δίνει συγχρόνως χρήσιμα στοιχεία στα προγράμματα ασφάλειας του συστήματος.

Για την αντιμετώπιση του προβλήματος έχουν εφαρμοστεί ασαφή λογική και τεχνητά νευρωνικά δίκτυα, καθώς και αυτοπαλινδρομικά μοντέλα συνολικού χρόνου 30 λεπτών ανά διαστήματα 1 με 2 λεπτών.

➤ 2.22 Μεσοπρόθεσμη και Μακροπρόθεσμη Πρόβλεψη Φορτίου

Η μεσοπρόθεσμη πρόβλεψη φορτίου γίνεται για ένα έτος με χρονικό βήμα εβδομάδας, ενώ η μακροπρόθεσμη για χρονικό ορίζοντα 10-20 ετών με χρονικό βήμα ενός έτους. Η πρώτη αφορά κυρίως τη ρύθμιση των προγραμμάτων συντήρησης των μονάδων παραγωγής και αξιοποίησης των διαθέσιμων υδάτινων και λοιπών πόρων, στις καταλληλότερες χρονικές περιόδους του έτους. Η δεύτερη ασχολείται κυρίως με την ετήσια αιχμή φορτίου -ή και τις εποχιακές- και τη συνολική ετήσια ενέργεια για τα επόμενα 10 με 20 χρόνια, ώστε να είναι δυνατός ο σχεδιασμός και η κατασκευή του συστήματος ηλεκτρικής ενέργειας. Συνήθως αντιμετωπίζονται μαζί, διότι ξεκινώντας κανείς από τη μακροπρόθεσμη πρόβλεψη, εύκολα μπορεί να μεταβεί στο επόμενο επίπεδο, αν λάβει υπόψη του την εποχιακή κυκλικότητα του κλίματος κτλ.. Γι' αυτό άλλωστε συνδυάζονται κατάλληλα με οικονομοτεχνικές μελέτες μεσοπρόθεσμου και μακροπρόθεσμου σχεδιασμού [A56].

Στο κεφάλαιο που ακολουθεί θα αναφερθούμε αναλυτικά στα πολυεπίπεδα τεχνητά νευρωνικά δίκτυα, καθώς και στον αλγόριθμο της ανάστροφης διάδοσης σφάλματος -Back Propagation-, ο οποίος εφαρμόζεται στα συγκεκριμένα δίκτυα.

Κεφάλαιο 3

Η μέθοδος της ανάστροφης διάδοσης-Back Propagation

➤ 3.1 Εισαγωγή - Το πολυεπίπεδο Perceptron

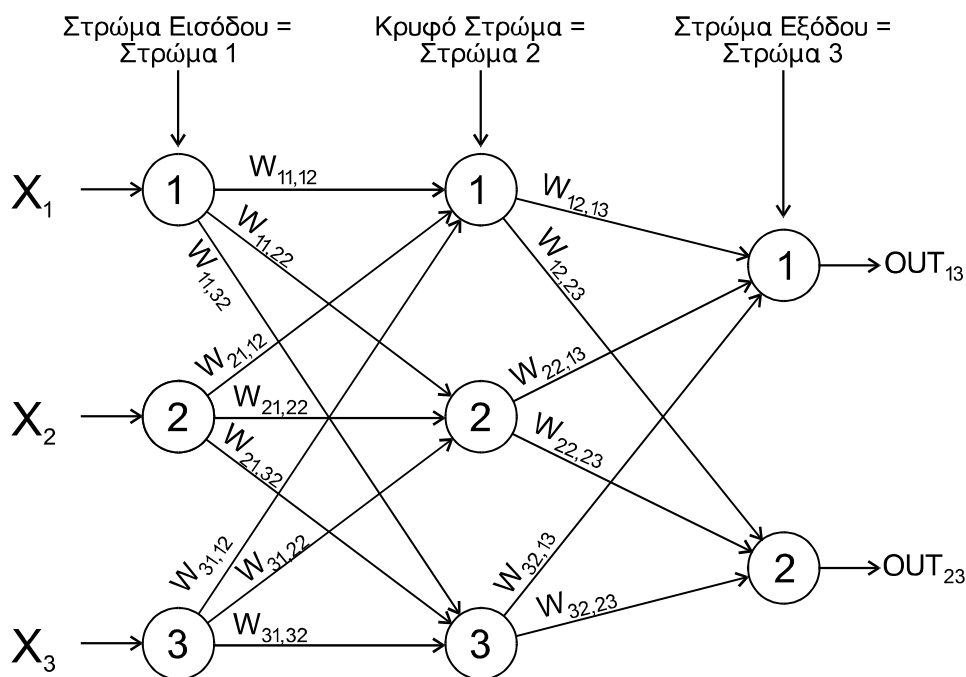
Μια ευρέως χρησιμοποιούμενη κατηγορία νευρωνικών δικτύων είναι τα πολυεπίπεδα Perceptron (multilayer perceptron ή MLP) . Τα δίκτυα αυτά, ένα παράδειγμα των οποίων φαίνεται στην επόμενη παράγραφο, αποτελούνται από κόμβους οι οποίοι καλούνται νευρώνες και είναι διασυνδεδεμένοι μεταξύ τους με ελεύθερες παραμέτρους, που καλούνται βάρη. Κάθε στήλη νευρώνων ορίζει και ένα επίπεδο, δηλαδή αυτά τα δίκτυα περιλαμβάνουν τα επίπεδα εισόδου και εξόδου καθώς και ένα ή περισσότερα εσωτερικά κρυμμένα επίπεδα. Επίσης είναι πρόσθιας τροφοδότησης , δηλαδή τροφοδοτούνται από μπροστινές εισόδους μόνο .

Το πολυεπίπεδο νευρωνικό δίκτυο Perceptron χρησιμοποιούνται σε ποικίλες, αλλά και δύσκολες εφαρμογές, στις οποίες εκπαιδεύονται βάσει του αλγόριθμου της ανάστροφης διάδοσης του σφάλματος (Error Back-Propagation). Ο παραπάνω αυτός αλγόριθμος έχει να κάνει με τον 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 δίκτυα MLPs.
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.

➤ 3.2 Ροή Σημάτων



Σχήμα 3.2-1 - Τεχνητό νευρωνικό δίκτυο δύο επιπέδων πλήρως διασυνδεδεμένο

Στο παραπάνω σχήμα παριστάνεται η αρχιτεκτονική ενός πολυεπίπεδου Τ.Ν.Δ. Perceptron δύο επιπέδων, τριών 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 συναρτήσεων των αντίστοιχων βαρών.

- Τα σήματα σφάλματος που δημιουργούνται μονάχα κατά την διάρκεια της εκπαίδευσης και προωθούνται κατά την αντίστροφη φορά, από την έξοδο προς την είσοδο, διαμέσου του νευρωνικού δικτύου.

➤ 3.3 Ο αλγόριθμος Back-Propagation

Έστω ότι έχουμε ένα σύνολο εκπαίδευσης που αποτελείται από N ζεύγη της μορφής $[\bar{x}, \bar{d}]$, όπου $\bar{x}$ διάνυσμα εισόδου (πρότυπο) και $\bar{d}$ το διάνυσμα των επιθυμητών εξόδων όταν εφαρμοσθεί σαν είσοδος το πρότυπο $\bar{x}$.

Το σφάλμα στην έξοδο του τυχαίου νευρώνα j , ο οποίος είναι κόμβος εξόδου, για την n -οστή επανάληψη όπου και εφαρμόζεται η είσοδος $\bar{x}(n)$ $\{n = 1, 2, \dots, N\}$ δίδεται από τον τύπο :

$$e_j(n) = d_j(n) - y_j(n) \quad (3.3-1)$$

όπου d_j και y_j η επιθυμητή και η πραγματική έξοδος αντίστοιχα του νευρώνα j όταν εφαρμόζεται η $\bar{x}(n)$.

Η στιγμιαία τιμή του τετραγωνικού σφάλματος για τον νευρώνα j ορίζεται ως $\frac{1}{2} e_j^2(n)$ και το άθροισμα των τετραγωνικών σφαλμάτων όλων των νευρώνων εξόδου για την επανάληψη n δίδεται από τη σχέση :

$$G(n) = \frac{1}{2} \sum_{j \in \text{ex}} e_j^2(n) \quad (3.3-2)$$

Επιπρόσθετα ορίζουμε τη μέση τιμή των σφαλμάτων για όλα τα N πρότυπα ως εξής :

$$G_{av} = \frac{1}{N} \sum_{n=1}^N G(n) \quad (3.3-3)$$

Προφανώς τόσο το G , όσο και το G_{av} , εξαρτώνται από τις ελεύθερες παραμέτρους του δικτύου, που δεν είναι άλλες από τα βάρη μεταξύ των συνδέσεων, αλλά και τις πολώσεις. Ο σκοπός της διαδικασίας εκπαίδευσης είναι φυσικά η ελαχιστοποίηση της μέσης τιμής των σφαλμάτων G_{av} , με την ανάλογη προσαρμογή των ελευθέρων παραμέτρων του δικτύου. Για να επιτευχθεί αυτό χρησιμοποιούμε μια συγκεκριμένη λογική όπου θεωρούμε ότι τα βάρη ενημερώνονται μετά το πέρασμα κάθε προτύπου, σύμφωνα με τα σφάλματα που παρουσιάζονται κατά την παρουσίαση του αντίστοιχου προτύπου στο δίκτυο. Ουσιαστικά θα λέγαμε ότι η μέση τιμή των μεταβολών όλων των βαρών είναι μια προσεκτική εκτίμηση της πραγματικής μεταβολής των βαρών, που προκύπτει από τη διαδικασία ελαχιστοποίησης του G_{av} .

Όπως παρατηρούμε από το σχήμα (3.2-1) ο νευρώνας εξόδου j τροφοδοτείται από τα σήματα εξόδου όλων των νευρώνων του προηγούμενου επιπέδου. Οπότε το σήμα $u_j(n)$ δίνεται από τη σχέση :

$$u_j(n) = \sum_{i=0}^p w_{ji}(n) y_i(n) \quad (3.3-4)$$

όπου p είναι το σύνολο όλων των νευρώνων του προηγούμενου επιπέδου. Θα πρέπει να σημειωθεί ότι το βάρος w_{j0} (για $i = 0$) είναι η λεγόμενη πόλωση και αντιστοιχεί σε εξωτερική είσοδο $y_0 = -1$.

Συνεπώς το σήμα εξόδου $y_j(n)$ του νευρώνα εξόδου j κατά την επανάληψη n θα δοθεί από τη σχέση :

$$y_j(n) = \phi(u_j(n)) \quad (3.3-5)$$

Ο αλγόριθμος προβαίνει σε κάθε επανάληψη σε μια διόρθωση $\Delta w_{ji}(n)$ στο βάρος $w_{ji}(n)$ και η οποία είναι ανάλογη της κλίσης $\frac{\partial G(n)}{\partial w_{ji}(n)}$. Βάσει του γνωστού κανόνα της αλυσίδας η μερική αυτή παράγωγος παίρνει την μορφή :

$$\frac{\partial G(n)}{\partial w_{ji}(n)} = \frac{\partial G(n)}{\partial e_j(n)} \frac{\partial e_j(n)}{\partial y_j(n)} \frac{\partial y_j(n)}{\partial u_j(n)} \frac{\partial u_j(n)}{\partial w_{ji}(n)} \quad (3.3-6)$$

Η παραπάνω αυτή κλίση καθορίζει την κατεύθυνση ελαχιστοποίησης στον χώρο των βαρών. Αν διαφορίσουμε και τις δύο πλευρές της εξίσωσης (3.3-2) ως προς $e_j(n)$ παίρνουμε :

$$\frac{\partial G(n)}{\partial e_j(n)} = e_j(n) \quad (3.3-7)$$

Διαφορίζοντας ως προς $y_j(n)$ και τα δύο μέλη της εξίσωσης (3.3-1) παίρνουμε :

$$\frac{\partial e_j(n)}{\partial y_j(n)} = -1 \quad (3.3-8)$$

Ακολουθώντας διαφορίζοντας ως προς $u_j(n)$ και τα δύο μέλη της εξίσωσης (3.3-5) παίρνουμε :

$$\frac{\partial y_j(n)}{\partial u_j(n)} = \phi'_j(u_j(n)) \quad (3.3-9)$$

Τέλος διαφορίζοντας ως προς $w_{ji}(n)$ και τα δύο μέλη της εξίσωσης (3.3-4) παίρνουμε :

$$\frac{\partial u_j(n)}{\partial w_{ji}(n)} = y_i(n) \quad (3.3-10)$$

Αντικαθιστώντας στην (3.3-6) τις εξισώσεις (3.3-7)-(3.3-10) προκύπτει τελικά :

$$\frac{\partial G(n)}{\partial w_{ji}(n)} = -e_j(n)\phi'_j(u_j(n))y_i(n) \quad (3.3-11)$$

Η διόρθωση $\Delta w_{ji}(n)$, που είναι άλλωστε και αυτό που μας ενδιαφέρει , προσδιορίζεται από τον κανόνα δέλτα :

$$\Delta w_{ji}(n) = -\eta \frac{\partial G(n)}{\partial w_{ji}(n)} \quad (3.3-12)$$

Όπου η μια σταθερά που καλείται ρυθμός μάθησης (learning rate).

Αν ορίσουμε ως τοπική κλίση $\delta_j(n)$ για τον νευρώνα j την παράσταση :

$$\delta_j(n) = \frac{\partial G(n)}{\partial e_j(n)} \frac{\partial e_j(n)}{\partial y_j(n)} \frac{\partial y_j(n)}{\partial u_j(n)} = e_j(n)\phi'_j(u_j(n)) \quad (3.3-13)$$

τότε οι (3.3-11) και (3.3-12) μας δίνουν τελικά :

$$\Delta w_{ji}(n) = \eta * \delta_j(n) * y_i(n) \quad (3.3-14)$$

Όπως παρατηρούμε από την παραπάνω αποδεικτική διαδικασία ο πιο σημαντικός παράγοντας από τον οποίο εξαρτάται τελικά η αναπροσαρμογή των βαρών $\Delta w_{ji}(n)$ δεν είναι άλλος, όπως είναι και λογικό άλλωστε, από το σφάλμα $e_j(n)$ στον νευρώνα j . Συνεπώς θα πρέπει να θεωρήσουμε δύο περιπτώσεις, από τις οποίες η πρώτη θα είναι ο νευρώνας j να συνιστά κόμβο εξόδου και η δεύτερη περίπτωση ο νευρώνας j να αποτελεί εσωτερικό κόμβο του νευρωνικού δικτύου.

Στην πρώτη περίπτωση και εφόσον γνωρίζουμε την επιθυμητή έξοδο του μπορούμε απλά να υπολογίσουμε το σήμα σφάλματος. Στη δεύτερη περίπτωση μπορούμε να σημειώσουμε ότι παρόλο

που αυτοί οι νευρώνες δεν είναι άμεσα προσπελάσιμοι, έχουν κι αυτοί μερίδιο ευθύνης όσον αφορά τα σφάλματα εξόδου του δικτύου. Ο τρόπος λοιπόν με τον οποίο συνυπολογίζουμε τη συμμετοχή των εσωτερικών κόμβων ενός δικτύου στο προκύπτον σφάλμα εξόδου, περιγράφεται παρακάτω και λέγεται αντίστροφη προώθηση των σημάτων σφάλματος, διαμέσου του δικτύου.

➤ 3.4 Αντίστροφη προώθηση σημάτων σφάλματος διαμέσου δικτύου

Όταν ο νευρώνας βρίσκεται σε κρυμμένο επίπεδο δεν υφίσταται κάποια επιθυμητή έξοδος για αυτόν. Το σήμα σφάλματος λοιπόν θα πρέπει να καθορισθεί αναδρομικά σε σχέση με τα σφάλματα όλων των νευρώνων με τους οποίους αυτός συνδέεται απευθείας. Σε αυτήν την περίπτωση όπως είναι κατανοητό αυξάνει σημαντικά η πολυπλοκότητα του αλγόριθμου. Βάσει των εξισώσεων (3.3-14) & (3.3-9) μπορούμε να θεωρήσουμε ότι :

$$\delta_j(n) = -\frac{\partial G(n)}{\partial y_j(n)} \frac{\partial y_j(n)}{\partial u_j(n)} = -\frac{\partial G(n)}{\partial y_j(n)} \phi'_j(u_j(n)) \quad (3.4-1)$$

Επίσης από προηγούμενη αναφορά γνωρίζουμε ότι :

$$G(n) = \frac{1}{2} \sum_{k \in \omega} e_k^2(n) \quad (3.4-2)$$

Διαφορίζοντας ως προς $y_j(n)$ την εξίσωση (3.4-2) παίρνουμε :

$$\frac{\partial G(n)}{\partial y_j(n)} = \sum_k e_k \frac{\partial e_k(n)}{\partial y_j(n)} \quad (3.4-3)$$

Στη συνέχεια σύμφωνα με τον κανόνα της αλυσίδας γράφουμε :

$$\frac{\partial G(n)}{\partial y_j(n)} = \sum_k e_k \frac{\partial e_k(n)}{\partial u_k(n)} \frac{\partial u_k(n)}{\partial y_j(n)} \quad (3.4-4)$$

Όμως γνωρίζουμε ότι :

$$e_k(n) = d_k(n) - y_k(n) = d_k(n) - \phi_k(u_k(n)) \quad (3.4-5)$$

Συνεπώς παίρνουμε :

$$\frac{\partial e_k(n)}{\partial u_k(n)} = -\phi'_k(u_k(n)) \quad (3.4-5)$$

Επίσης το σήμα ενεργοποίησης του νευρώνα k δίνεται από τη σχέση :

$$u_k(n) = \sum_{j=0}^q w_{kj}(n) y_j(n) \quad (3.4-6)$$

όπου q ο συνολικός αριθμός των νευρώνων του προηγούμενου επιπέδου που συνδέονται απευθείας με τον νευρώνα k . Συνεπώς :

$$\frac{\partial u_k(n)}{\partial y_j(n)} = w_{kj}(n) \quad (3.4-7)$$

Από αντικατάσταση των (3.4-6) & (3.4-7) στην (3.4-4) παίρνουμε :

$$\frac{\partial G(n)}{\partial y_j(n)} = -\sum_k e_k(n) \phi'_k(u_k(n)) w_{kj}(n) = -\sum_k \delta_k(n) w_{kj}(n) \quad (3.4-8)$$

Τέλος από τις (3.4-1) & (3.4-8) προκύπτει η τοπική κλίση $\delta_j(n)$ για τον κρυμμένο νευρώνα j :

$$\delta_j(n) = \phi'_j(u_j(n)) \sum_k \delta_k(n) w_{kj}(n) \quad (3.4-9)$$

Σε αυτό το σημείο αξίζει να αναφέρουμε όσον αφορά την παραπάνω εξίσωση (3.4-9), ότι ο παράγοντας $\phi'_j(u_j(n))$ εξαρτάται μόνο από τη συνάρτηση ενεργοποίησης του νευρώνα j , ενώ το άθροισμα, δηλαδή ο δεύτερος παράγοντας, εξαρτάται από την τιμή των τοπικών κλίσεων δ του επόμενου επιπέδου και από όλα τα βάρη που συνδέουν τον νευρώνα j με το επίπεδο αυτό.

Η διόρθωση $\Delta w_{ji}(n)$ του βάρους που συνδέει τον νευρώνα i με τον νευρώνα j του επόμενου επιπέδου, μετά από την εφαρμογή στην είσοδο του n -οστού προτύπου εκπαίδευσης (διάνυσμα $\bar{x}(n)$), ορίζεται από τον κανόνα δέλτα ως εξής :

$$\Delta w_{ji}(n) = \eta \delta_j(n) y_i(n) \quad (3.4-10)$$

όπου η τοπική κλίση $\delta_j(n)$ εξαρτάται από το αν ο νευρώνας j είναι νευρώνας εξόδου ή είναι νευρώνας που ανήκει σε κρυμμένο επίπεδο .

❖ Αν ο νευρώνας j ανήκει στο επίπεδο εξόδου, τότε η τοπική κλίση $\delta_j(n)$ ισούται με το γινόμενο της παραγώγου $\phi'_j(u_j(n))$ και του σφάλματος του νευρώνα $e_j(n)$.

❖ Αν ο νευρώνας j ανήκει σε κρυμμένο επίπεδο, τότε η τοπική κλίση $\delta_j(n)$ ισούται με το γινόμενο της παραγώγου $\phi'_j(u_j(n))$ και του σταθμισμένου αθροίσματος των δ του επόμενου επιπέδου ($j+1$).

➤ 3.5 Τα δύο στάδια του υπολογισμού του αλγορίθμου της ανάστροφης διάδοσης σφάλματος (B-P).

Κατά την εφαρμογή του αλγορίθμου Back-Propagation κι αφού αρχικοποιήσουμε τα βάρη και τις πολώσεις σε μικρές τυχαίες τιμές στο διάστημα $[-1,1]$, διακρίνουμε δύο περάσματα όσον αφορά την εκτέλεση των υπολογισμών. Το πρώτο καλείται *πέρασμα ορθής φοράς* (forward pass) και το δεύτερο *πέρασμα αντίστροφης φοράς* (reverse pass).

Στη διαδικασία του *ευθέως περάσματος* όλα τα βάρη των συνδέσεων παραμένουν αναλλοίωτα, καθ' όλη την διάρκεια και υπολογίζονται οι εξόδοι όλων των νευρώνων. Πιο συγκεκριμένα το σήμα εξόδου του νευρώνα j υπολογίζεται ως εξής :

$$y_j(n) = \phi(u_j(n)) \quad (3.5-1)$$

όπου :

$$u_j(n) = \sum_{i=0}^p w_{ji}(n) y_i(n) \quad (3.5-2)$$

όπου p ο συνολικός αριθμός εισόδων που εφαρμόζονται στον κόμβο j , $w_{ji}(n)$ είναι το βάρος που συνδέει τον νευρώνα i με τον νευρώνα j και $y_i(n)$ είναι η έξοδος του νευρώνα i .

Στην περίπτωση που ο νευρώνας j ανήκει στο πρώτο κρυμμένο επίπεδο, τότε ο δείκτης i αναφέρεται στην i -οστή είσοδο, δηλαδή :

$$y_i(n) = x_i(n) \quad (3.5-3)$$

όπου $x_i(n)$ το i -οστό στοιχείο του πρότυπου διανύσματος εισόδου $\bar{x}(n)$.

Αν αντίθετα ο νευρώνας αποτελεί κόμβο εξόδου, τότε ο δείκτης j με τη σειρά του αναφέρεται στην j -οστή έξοδο, δηλαδή :

$$y_j(n) = o_j(n) \quad (3.5-4)$$

όπου $o_j(n)$ το j -οστό στοιχείο του πρότυπου διανύσματος εξόδου $o(n)$. Η διαφορά της εξόδου $y_j(n)$ από την επιθυμητή έξοδο $d_j(n)$, παράγει το σήμα σφάλματος του νευρώνα εξόδου j , $e_j(n)$.

Ανακεφαλαιώνοντας, το *ευθύ πέρασμα* ξεκινά από το πρώτο κρυμμένο επίπεδο με την παρουσίαση του διανύσματος εισόδου (πρότυπο), στη συνέχεια υπολογίζει όλα τα σήματα εξόδου των νευρώνων βάσει των προαναφερθέντων σχέσεων και προχωράει στα επόμενα επίπεδα του δικτύου, ώσπου να φτάσει τελικά στο τελευταίο επίπεδο (εξόδου) και να υπολογίσει το διάνυσμα εξόδου του δικτύου, καθώς και το σφάλμα του κάθε νευρώνα αυτού του επιπέδου.

Το *αντίστροφο πέρασμα*, αντιθέτως, ξεκινά από το επίπεδο εξόδου περνώντας τα σήματα σφάλματος προς τα πίσω και υπολογίζοντας αναδρομικά την τιμή του δ του κάθε νευρώνα. Με αυτόν τον τρόπο αναπροσαρμόζουμε τα βάρη μεταξύ των συνδέσεων σύμφωνα με τον κανόνα δέλτα -βλ. σχέση (3.4-10).

Εδώ πρέπει να σημειωθεί ότι για τους νευρώνες εξόδου το δ ισούται πάντα με το γινόμενο του σφάλματος και της πρώτης παραγώγου της μη-γραμμικής συνάρτησης ενεργοποίησης. Ακολουθώντας μέσω αυτού του δ υπολογίζουμε τα δ των νευρώνων του προτελευταίου επιπέδου και ούτω καθεξής για τα υπόλοιπα επίπεδα, ώσπου να φτάσουμε τελικά στο πρώτο επίπεδο του δικτύου.

Τέλος αναφέρεται ότι η διαδικασία αυτή των δύο περασμάτων εκτελείται για κάθε πρότυπο του συνόλου εκπαίδευσης με μια κυκλική επανάληψη.

➤ 3.6 Σιγμοειδής μη-γραμμικότητα

Για να υπολογίσουμε το δ κάθε νευρώνα απαιτείται πρώτα απ' όλα η γνώση της πρώτης παραγώγου της συνάρτησης ενεργοποίησης $\phi'(\cdot)$. Προφανώς η ϕ θα πρέπει να παραγωγίσιμη και άρα συνεχής. Ένα παράδειγμα μιας κατηγορίας τέτοιων μη-γραμμικών συναρτήσεων που χρησιμοποιείται συχνά σε πολυεπίπεδα νευρωνικά δίκτυα, είναι οι σιγμοειδείς συναρτήσεις, καθώς και άλλες, τις σημαντικότερες εκ των οποίων βλέπουμε στον παρακάτω πίνακα :

$f(x) = 1/(1 + e^{-ax})$	$f(x) = \tanh(x)$	$f(x) = a \cdot x$
Σιγμοειδής με πεδίο τιμών $[0,1]$	Σιγμοειδής με πεδίο τιμών $[-1,1]$	Γραμμική
$f(x) = \begin{cases} a & , \quad x \geq \gamma \\ b & , \quad x < \gamma \end{cases}$	$f(x) = \exp\left(\frac{-x^2}{v}\right)$	$f(x) = \begin{cases} \gamma & , \quad x \geq \gamma \\ x & , \quad x < \gamma \\ -\gamma & , \quad x \leq -\gamma \end{cases}$
Βηματική	Γκαουσιανή	Αναρρίχησης

Πίνακας 3.6 –1 : Συναρτήσεις Ενεργοποίησης

Πιο συγκεκριμένα η λογιστική συνάρτηση, όπως βλέπουμε παραπάνω, έχει τον τύπο :

$$y = \phi(u) = \frac{1}{1 + \exp(-u)} \quad -\infty < u < +\infty \quad (3.6-2)$$

Η παράγωγος της λογιστικής συνάρτησης είναι :

$$\frac{\partial y}{\partial u} = \frac{\exp(-u)}{[1 + \exp(-u)]^2} \quad (3.6-3)$$

ή αλλιώς :

$$\phi'(u)=y/[1-y] \quad (3.6-4)$$

Για τον νευρώνα j του επιπέδου εξόδου θα ισχύει λόγω της (3.6-3) :

$$\delta_j(n) = e_j(n)\phi'_j(u_j(n))=[d_j(n)-o_j(n)]o_j(n)[1-o_j(n)] \quad (3.6-5)$$

όπου $o_j(n)$ η έξοδος του νευρώνα j και $d_j(n)$ η επιθυμητή έξοδος αυτού.

Επίσης για έναν κρυμμένο νευρώνα j η κλίση εκφράζεται ως εξής :

$$\delta_j(n) = \phi'_j(u_j(n))\sum_k \delta_k(n)w_{kj}(n) = y_j(n)[1-y_j(n)]\sum_k \delta_k(n)w_{kj}(n) \quad (3.6-6)$$

Από τις παραπάνω σχέσεις παρατηρούμε ότι η πρώτη παράγωγος $\phi'_j(u_j(n))$ παίρνει τη μέγιστη τιμή της για $y_j(n) = 0,5$ και την ελάχιστη τιμή (= 0) για $y_j(n) = 0$ ή $y_j(n) = 1$. Επειδή η μεταβολή των βαρών των συνδέσεων είναι ανάλογη της παραγώγου, συμπεραίνουμε ότι τα βάρη μεταβάλλονται περισσότερο σε αυτούς τους νευρώνες όπου το σημείο ενεργοποίησης παίρνει μέτριες τιμές, με άμεσο αποτέλεσμα να εμφανίζεται κάποιο είδος ευστάθειας στην εκπαίδευση του αλγόριθμου.

➤ 3.7 Προσθήκη του όρου ορμής { momentum }

Ο αλγόριθμος Back-Propagation δίνει μια προσέγγιση της τροχιάς των βαρών η οποία υπολογίζεται με τη λεγόμενη μέθοδο της μέγιστης καθόδου. Όσο μικρότερο ρυθμό μάθησης η έχουμε, τόσο μικρότερη θα είναι η μεταβολή των βαρών σε κάθε επανάληψη και άρα τόσο πιο ομαλή θα είναι και η τροχιά της καμπύλης των βαρών. Συνεπώς το κόστος για υψηλό επίπεδο μάθησης είναι ο αργός ρυθμός μάθησης. Αν όμως χρησιμοποιηθεί υψηλός ρυθμός μάθησης για να επιταχυνθεί η διαδικασία, η καμπύλη των βαρών δεν θα είναι ομαλή, καθώς θα έχουμε απότομες μεταβολές, με αποτέλεσμα να εμφανίζεται κίνδυνος αστάθειας του αλγόριθμου ή αλλιώς φαινόμενο zig-zag. Ένας τρόπος εξάλειψης των παραπάνω κινδύνων είναι η τροποποίηση του κανόνα αναπροσαρμογής των βαρών, με την προσθήκη ενός όρου ορμής { momentum term } α , που ονομάζεται και σταθερά ορμής και είναι συνήθως θετικός αριθμός :

$$\Delta w_{ji}(n) = \alpha \Delta w_{ji}(n-1) + \eta \delta_j(n) y_i(n) \quad (3.6-6)$$

Για να παρατηρήσουμε την επίδραση του όρου ορμής, διατυπώνουμε την παραπάνω εξίσωση ως χρονική ακολουθία με δείκτη t , ο οποίος ξεκινάει από την τιμή 0 και φτάνει μέχρι την τρέχουσα τιμή n . Αυτό που θα προκύψει προφανώς θα είναι μια διαφορική εξίσωση πρώτης τάξης της οποίας η λύση είναι :

$$\Delta w_{ji}(n) = \eta \sum_{t=0}^n a^{n-t} \delta_j(t) y_i(t) \quad (3.6-7)$$

η οποία δεν είναι τίποτα άλλο από μια χρονική ακολουθία μήκους $n+1$. Από τις εξισώσεις (3.3-11) & (3.3-14) προκύπτει ότι : $\delta_j(n)y_i(n) = -\frac{\partial G(n)}{\partial w_{ji}(n)}$.

Συνεπώς η (3.6-7) παίρνει τη μορφή :

$$\Delta w_{ji}(n) = -\eta \sum_{t=0}^n a^{n-t} \frac{\partial G(n)}{\partial w_{ji}(n)} \quad (3.6-8)$$

Από την παραπάνω σχέση βγάζουμε τα εξής συμπεράσματα :

1. Η τρέχουσα μεταβολή $\Delta w_{ji}(n)$ ισοδυναμεί με το άθροισμα μιας εκθετικής χρονικής ακολουθίας, η οποία για να συγκλίνει θα πρέπει να ισχύει : $0 \leq |a| \leq 1$, δηλαδή ο όρος ορμής μπορεί να πάρει και αρνητικές τιμές κάτι που στην πράξη δεν συμβαίνει σχεδόν ποτέ.

2. Όταν η μερική παράγωγος $\frac{\partial G(n)}{\partial w_{ji}(n)}$ κρατάει σταθερό πρόσημο σε συνεχόμενες επαναλήψεις, τότε το $\Delta w_{ji}(n)$ αυξάνει σημαντικά και ως εκ τούτου αυξάνει πολύ και το αντίστοιχο βάρος $w_{ji}(n)$. Συνεπώς η εισαγωγή της σταθεράς ορμής σε αυτήν την περίπτωση τείνει κυρίως να επιταχύνει προς μία σταθερή κατεύθυνση την κάθοδο.

3. Όταν η μερική παράγωγος $\frac{\partial G(n)}{\partial w_{ji}(n)}$ μεταβάλλει το πρόσημο της σε κάποιο αριθμό διαδοχικών επαναλήψεων, τότε το $\Delta w_{ji}(n)$ ελαττώνεται σημαντικά και άρα δεν αυξάνει ιδιαίτερα και το αντίστοιχο βάρος $w_{ji}(n)$. Συνεπώς η εισαγωγή του παράγοντα ορμής εδώ έχει σταθεροποιητικό αποτέλεσμα για τον αλγόριθμο.

Συμπερασματικά μπορούμε να πούμε ότι η προσθήκη του όρου ορμής στη διαδικασία ανανέωσης των βαρών μπορεί να είναι εξαιρετικά επωφελής για τη μάθηση. Μάλιστα μπορεί κάλλιστα και ο ρυθμός μάθησης να μεταβάλλεται σε αντίθεση με ότι έχουμε υποθέσει μέχρι τώρα , γεγονός που δίνει περαιτέρω ενδιαφέρον όσον αφορά τον χαρακτήρα του αλγόριθμου.

➤ 3.8 Τρόποι εκπαίδευσης

Η διαδικασία εκπαίδευσης κατά τον αλγόριθμο Back-Propagation (B-P) 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 :

□ 3.8.1 Εκπαίδευση ανά πρότυπο

Στην περίπτωση αυτή 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 $\{ \bar{x}(1), \bar{d}(1) \}, \{ \bar{x}(2), \bar{d}(2) \} \dots \{ \bar{x}(n), \bar{d}(n) \}$, όπου τα $\{ \bar{x}(i), \bar{d}(i) \}$ στην 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. Αν $\Delta w_{ji}(n)$ η μεταβολή του βάρους $w_{ji}(n)$ μετά την παρουσίαση του n -οστού προτύπου, τότε η μέση μεταβολή του ίδιου βάρους για όλο το σύνολο των N προτύπων, μέσω των εξισώσεων (3.3-12) & (3.3-13) αντίστοιχα θα είναι :

$$\Delta \hat{w}_{ji}(n) = \frac{1}{N} \sum_{n=1}^N \Delta w_{ji}(n) = -\frac{\eta}{N} \sum_{n=1}^N \frac{\partial G(n)}{\partial w_{ji}(n)} = -\frac{\eta}{N} \sum_{n=1}^N e_j(n) \frac{\partial e_j(n)}{\partial w_{ji}(n)} \quad (3.8.1-1)$$

□ 3.8.2 Εκπαίδευση ανά εποχή

Σε αυτήν την περίπτωση τα βάρη ενημερώνονται έπειτα από την παρουσίαση στο δίκτυο όλου του συνόλου προτύπων της εποχής. Για μια τέτοια εποχή ορίζουμε τη μέση τιμή των τετραγωνικών σφαλμάτων από τις εξισώσεις (3.3-2) & (3.3-3) ως εξής :

$$G_{av} = \frac{1}{2N} \sum_{n=1}^N \sum_{j \in C} e_j^2(n) \quad (3.8.2-1)$$

Στην παραπάνω εξίσωση το εσωτερικό άθροισμα εφαρμόζεται σε όλους τους νευρώνες εξόδους του δικτύου, ενώ το εξωτερικό σε όλες τις εποχές. Αν ο ρυθμός μάθησης είναι η , τότε η ενημέρωση των βαρών θα καθοριστεί από τον γνωστό κανόνα δέλτα :

$$\Delta w_{ji}(n) = -\eta \frac{\partial G_{av}}{\partial w_{ji}} = -\frac{\eta}{N} \sum_{n=1}^N e_j(n) \frac{\partial e_j(n)}{\partial w_{ji}} \quad (3.8.2-2)$$

Για να υπολογίσουμε τη μερική παράγωγο μέσα στο άθροισμα εργαζόμαστε πανομοιότυπα με πριν. Εδώ λοιπόν η αναπροσαρμογή όλων των βαρών γίνεται μια φορά μετά το τέλος κάθε εποχής.

Αν συγκρίνουμε τώρα τις εξισώσεις (3.8.2-1) & (3.8.2-2), παρατηρούμε ότι για την ίδια τιμή του G_{av} , τα βάρη μεταβάλλονται διαφορετικά ανάλογα με τον τρόπο εκπαίδευσης που χρησιμοποιείται. Ουσιαστικά θα λέγαμε ότι το $\Delta \hat{w}_{ji}(n)$ της εκπαίδευσης ανά πρότυπο, προσεγγίζει το $\Delta w_{ji}(n)$ της εκπαίδευσης ανά εποχή.

Στην πράξη η μέθοδος που προτιμάται είναι η εκπαίδευση ανά πρότυπο, καθώς ενισχύει τον 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.

➤ 3.9 Κριτήρια τερματισμού

Όσον αφορά τον τερματισμό του αλγόριθμου **B-P** 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_o είναι τοπικό ή ολικό ελάχιστο, τότε, για να ισχύει αυτό, θα πρέπει το διάνυσμα κλίσης $\bar{g}(\bar{w})$, δηλαδή η μερική παράγωγος του σφάλματος ως προς το διάνυσμα βαρών $\bar{w}$, να ισούται με το μηδέν, για $\bar{w} = \bar{w}_o$.

Συνεπώς ένα κριτήριο του αλγόριθμου είναι το εξής :

✓ Ο αλγόριθμος **B-P** 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 $\bar{w}_o$ είναι ότι η συνάρτηση $G_{av}(\bar{w})$ είναι στατική, κάτι που μας οδηγεί στο παρακάτω κριτήριο τερματισμού :

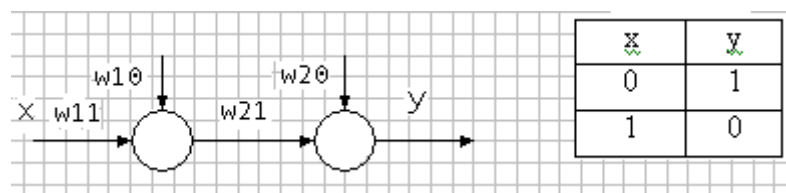
✓ Ο αλγόριθμος **B-P** θεωρείται ότι έχει συγκλίνει, όταν η απόλυτη τιμή του ρυθμού μεταβολής του σφάλματος ανά εποχή είναι ικανοποιητικά μικρή.

Επίσης ένα τελευταίο κριτήριο τερματισμού είναι και το ακόλουθο :

✓ Ο αλγόριθμος **B-P** τερματίζει σε ένα διάνυσμα βαρών $\bar{w}_{final}$, όταν $\|\bar{g}(\bar{w}_{final})\| \leq \varepsilon$, όπου ε ένα αρκετά ικανοποιητικά μικρό κατώφλι κλίσης ή όταν $G_{av}(\bar{w}_{final}) \leq \tau$, όπου τ ένα ικανοποιητικά μικρό κατώφλι για τη μέση τιμή του σφάλματος.

➤ 3.10 Αριθμητικό παράδειγμα – Επικοινωνία σε μια γραμμή κόμβων

Στο παράδειγμα αυτό θα δείξουμε πως ο αλγόριθμος ανάστροφης διάδοσης εφαρμόζεται σε ένα ΤΝΔ, το οποίο θα πρέπει να λειτουργεί ως αναστροφέας, δηλαδή να δίνει έξοδο 0, όταν η είσοδος του είναι 1 και έξοδο 1, όταν η είσοδος είναι 0. Προφανώς το δίκτυο πρέπει να έχει μια είσοδο και μια έξοδο. Έστω ότι χρησιμοποιούμε ένα κρυμμένο στρώμα μόνο, δηλαδή το δίκτυο μας είναι ενός επιπέδου όπως φαίνεται στο παρακάτω σχήμα :



Σχήμα 3.10-1 : Απλό δίκτυο επικοινωνίας με ένα κρυμμένο στρώμα

Επιλέγουμε τη λύση ο κρυμμένος νευρώνας που εκπαιδεύεται, να δρα ως αναστροφέας και ο νευρώνας εξόδου να ακολουθεί τον κρυμμένο νευρώνα.

Βήμα 1

Αρχικοποίηση των βαρών σε μικρές τυχαίες τιμές που να ανήκουν στο $[-1,1]$ και συγκεκριμένα για τις τιμές $w_{11} = -0,2$, $w_{10} = 0,3$, $w_{21} = 0,2$ και $w_{20} = -0,1$.

Βήμα 2

Παρουσίαση των διανυσμάτων εκπαίδευσης στο δίκτυο ($x = 0, d = 1$) και ($x = 1, d = 0$) στο δίκτυο.

Βήμα 3

Για $x = 0$ έχουμε $u_1 = -0,3$ και $y_1 = 1 / (1 + e^{0,3}) = 0,426$ που συνεπάγονται ότι :

$$u_2 = 0,2 * 0,426 + 0,1 \Leftrightarrow u_2 = 0,185 \quad \text{και} \quad y = y_2 = 1 / (1 + e^{-0,185}) = 0,546.$$

Το αντίστοιχο σφάλμα είναι : $e_0 = 1 - 0,546 = 0,454$.

Ο αντίστοιχος υπολογισμός για είσοδο $x = 1$ δίνει σφάλμα $e_1 = -0,544$ και συνεπώς το άθροισμα των τετραγώνων των σφαλμάτων για όλο το σύνολο εκπαίδευσης θα είναι : $e_0^2 + e_1^2 \approx 0,502$

Βήμα 4

Θεωρώντας ότι στο πρόβλημα ο ρυθμός μάθησης είναι ίσος με τη μονάδα και μη χρησιμοποιώντας όρο ορμής, ανανεώνουμε τα βάρη κατά ανάστροφη φορά από το νευρώνα εξόδου στο νευρώνα εισόδου και χρησιμοποιώντας απευθείας τη σχέση (3.6-5) :

$$\Delta w_{20} = 1 * 0,454 * 0,546 * (1 - 0,546) * (-1) \approx -0,112$$

$$\Delta w_{21} = 1 * 0,454 * 0,546 * (1 - 0,546) * (0,426) \approx 0,048$$

Συνεπώς οι ανανεωμένες τιμές των βαρών για το νευρώνα εξόδου θα είναι οι εξής :

$$w_{20, \text{νέο}} = w_{20, \text{παλιό}} + \Delta w_{20} = -0,1 - 0,112 = -0,212$$

$$w_{21, \text{νέο}} = w_{21, \text{παλιό}} + \Delta w_{21} = 0,2 + 0,048 = 0,248$$

Ομοίως για τον κρυμμένο νευρώνα παίρνουμε :

$$\Delta w_{10} \approx 0,027 \text{ και } \Delta w_{11} = 0 \text{ (αφού } x = 0 \text{)} \text{ και } w_{10, \text{νέο}} = 0,3 - 0,027 = 0,273 \text{ και } w_{11, \text{νέο}} = -0,2$$

Οι παραπάνω υπολογισμοί επαναλαμβάνονται για το δεύτερο διάνυσμα εκπαίδευσης ($x = 1, d = 0$), οπότε συμπληρώνεται ο πρώτος κύκλος επαναλήψεων- εποχή. Το συνολικό τετραγωνικό σφάλμα αυτή τη φορά βρίσκεται ίσο με 0,5, δηλαδή είναι λίγο μικρότερο από το αρχικό σφάλμα 0,502, από το οποίο καταλαβαίνουμε ότι μετά από κάθε εποχή υπάρχει βελτίωση (έστω και μικρή).

Για το πληρέστερο του προβλήματος θα πρέπει να πούμε ότι τελικά χρειάζονται να περάσουν 126 εποχές για να φτάσουμε τελικά σε μια τιμή τετραγωνικού σφάλματος ίση με 0,0513 και βάρη :

$$\boxed{w_{11} = -6,05} \quad \boxed{w_{10} = -2,65} \quad \boxed{w_{21} = 3,68} \quad \boxed{w_{20} = 1,66}$$

Οι έξοδοι του δικτύου είναι οι εξής :

✓ $y = 0,856$ για $x = 0$

✓ $y = 0,175$ για $x = 1$

➤ 3.11 Περίληψη της διαδικασίας της ανάστροφης διάδοσης

Τα στάδια της εν σειρά (on-line) λειτουργίας του αλγορίθμου είναι τα εξής :

1) Αρχικοποίηση $\Rightarrow$ Ορίζουμε την αρχιτεκτονική του δικτύου (αριθμός εισόδων-εξόδων και επιπέδων-νευρώνων καθώς και τρόπος διασύνδεσης μεταξύ των νευρώνων) και αρχικοποιούμε τα βάρη των συνδέσεων, σε τυχαίες μικρές τιμές που ανήκουν στο $[-1,1]$.

2) Παρουσίαση των προτύπων εκπαίδευσης $\Rightarrow$ Παρουσιάζουμε στο νευρωνικό δίκτυο μια εποχή προτύπων εκπαίδευσης και για κάθε πρότυπο επαναλαμβάνουμε τα βήματα 3 & 4.

3) Υπολογισμοί ευθείας φοράς $\Rightarrow$ Σε αυτό το στάδιο θεωρούμε το πρότυπο εκπαίδευσης $\{ \bar{x}(n), \bar{d}(n) \}$ όπου $\bar{x}(n)$ το διάνυσμα εισόδου που εφαρμόζεται στο δίκτυο και $\bar{d}(n)$ το επιθυμητό διάνυσμα εξόδου του δικτύου. Υπολογίζουμε τα σήματα που προκύπτουν σε κάθε επίπεδο. Το σήμα ενεργοποίησης του νευρώνα j του l επιπέδου είναι :

$$u_j^{(l)}(n) = \sum_{i=0}^p w_{ji}^{(l)}(n) y_i^{(l-1)}(n)$$

όπου $w_{ji}^{(l)}$ το βάρος που συνδέει τον νευρώνα j του l επιπέδου με τον νευρώνα i του $(l-1)$ επιπέδου, $y_i^{(l-1)}$ το σήμα εξόδου του νευρώνα i του προηγούμενου επιπέδου και p ο συνολικός αριθμός των νευρώνων του $(l-1)$ επιπέδου. Για $i = 0$ προκύπτει $y_o = -1$ και $w_{j0} = \theta_j$ όπου θ_j η τιμή πόλωσης. Το σήμα εξόδου του παραπάνω νευρώνα είναι το εξής :

$$y_j^{(l)}(n) = \varphi(u_j^{(l)}(n)) = \frac{1}{1 + \exp(-u_j^{(l)}(n))}$$

αφού, όπως είναι προφανές, επιλέγουμε τη λογιστική συνάρτηση ως συνάρτηση ενεργοποίησης για τους νευρώνες.

Αν ο νευρώνας j ανήκει στο πρώτο κρυμμένο επίπεδο, δηλαδή $l = 1$, τότε θέτουμε :

$$y_i^{(0)}(n) = x_i(n) \quad \forall i$$

όπου $x_i(n)$ το i -οστό στοιχείο του διανύσματος εισόδου $\bar{x}(n)$.

Αν όμως ο νευρώνας j ανήκει στο επίπεδο εξόδου ($l = L$), τότε θα ισχύει :

$$y_j^{(L)}(n) = o_j(n) \quad \forall j$$

Συνεπώς υπολογίζουμε το σφάλμα εξόδου του νευρώνα j :

$$e_j(n) = d_j(n) - o_j(n)$$

όπου $d_j(n)$ το j -οστό στοιχείο του επιθυμητού διανύσματος εξόδου $\bar{d}(n)$ και $o_j(n)$ το j -οστό στοιχείο του πραγματικού διανύσματος εξόδου $\bar{o}(n)$.

4) Υπολογισμοί αντίστροφης φοράς $\Rightarrow$ Εδώ υπολογίζουμε την τοπική κλίση του κάθε νευρώνα θεωρώντας τις δύο παρακάτω περιπτώσεις :

❖ Αν ο νευρώνας j είναι κόμβος εξόδου ($l=L$), τότε έχουμε :

$$\delta_j^{(L)}(n) = e_j^{(L)}(n) o_j(n) [1 - o_j(n)]$$

❖ Αν ο νευρώνας j ανήκει στο κρυμμένο επίπεδο l , τότε θα ισχύει :

$$\delta_j^{(l)}(n) = y_j^{(l)}(n)[1 - y_j^{(l)}(n)] \sum_k \delta_k^{(l+1)}(n) w_{kj}^{(l+1)}(n)$$

Στη συνέχεια προσαρμόζουμε τα βάρη του επιπέδου l σύμφωνα με τον γνωστό κανόνα δ :

$$w_{ji}^{(l)}(n+1) = w_{ji}^{(l)}(n) + \alpha \{w_{ji}^{(l)}(n) - w_{ji}^{(l-1)}(n-1)\} + \eta \delta_j^{(l)}(n) y_i^{(l-1)}(n)$$

όπου η ο ρυθμός μάθησης και α ο όρος ορμής.

5) Επανάληψη $\Rightarrow$ Στο τελευταίο αυτό στάδιο επαναλαμβάνουμε όλους τους παραπάνω υπολογισμούς για νέες εποχές προτύπων, ώσπου οι ελεύθερες παράμετροι του δικτύου να σταθεροποιηθούν και η μέση τιμή του σφάλματος G_{av} να πάρει την ελάχιστη τιμή της. Τέλος εδώ θα πρέπει να αναφερθεί ότι οι σταθερές μάθησης και ορμής συνήθως βαίνουν μειούμενες κατά το πέρασμα των εποχών.

Στο επόμενο κεφάλαιο θα περιγράψουμε αναλυτικά το πρόγραμμα που υλοποιεί τον αλγόριθμο της ανάστροφης διάδοσης σφάλματος, καθώς και τα σενάρια εκπαίδευσης που δημιουργήσαμε και εκτελέσαμε. Τέλος θα παραθέσουμε και τις ανάλογες παρατηρήσεις – σχόλια στα αποτελέσματα που προέκυψαν.

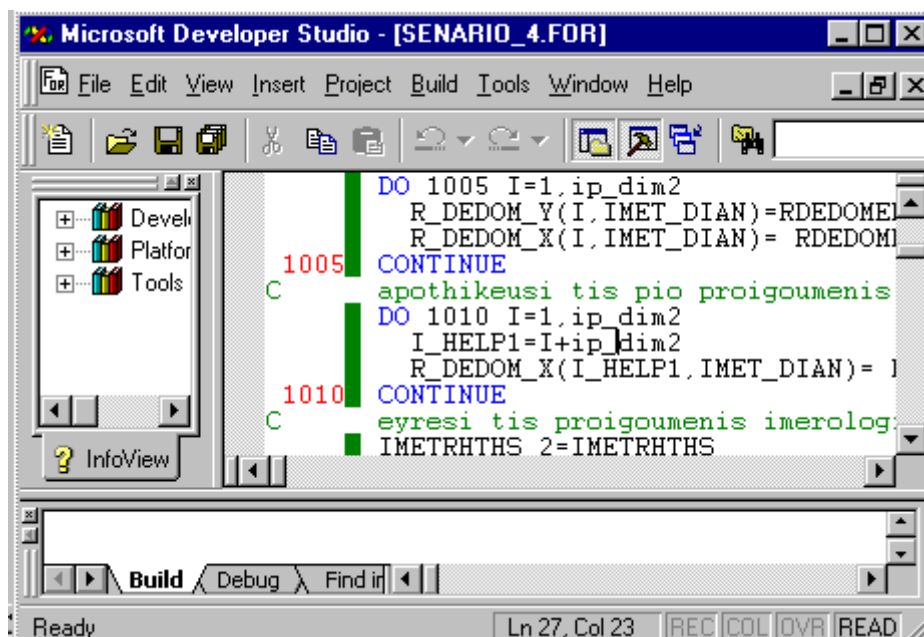
Κεφάλαιο 4

Εφαρμογές & σενάρια του αλγορίθμου της ανάστροφης διάδοσης

➤ 4.1 Γενικά για το πρόγραμμα εκπαίδευσης T.N.Δ.

Στην παρούσα διπλωματική εργασία εργαστήκαμε σε περιβάλλον της Digital Visual Fortran (έκδοση 1997) και το βασικό πρόγραμμα το οποίο χρησιμοποιήσαμε, αφορά τη βραχυπρόθεσμη πρόβλεψη τεχνητού νευρωνικού δικτύου –ΤΝΔ- δύο επιπέδων. Παρακάτω παρουσιάζεται ένα απλό «παράθυρο» του υπολογιστή, το οποίο μας δείχνει το περιβάλλον της Digital Visual Fortran και αμέσως μετά θα κάνουμε μια απλή αναφορά, σχετικά με το περιεχόμενο του προγράμματος.

Στο «παράθυρο», απεικονίζονται τρεις περιοχές εργασίας : μία αριστερά η οποία ονομάζεται *χώρος εργασίας*, μία δεξιά, η οποία αποτελεί το χώρο εγγραφής των εντολών του εκάστοτε προγράμματος και τέλος μία στο κάτω μέρος, η οποία μας εμφανίζει τα τυχόν συντακτικά λάθη του προγράμματος.



Σχήμα 4.1-1 : Παράδειγμα περιβάλλοντος της Visual Fortran

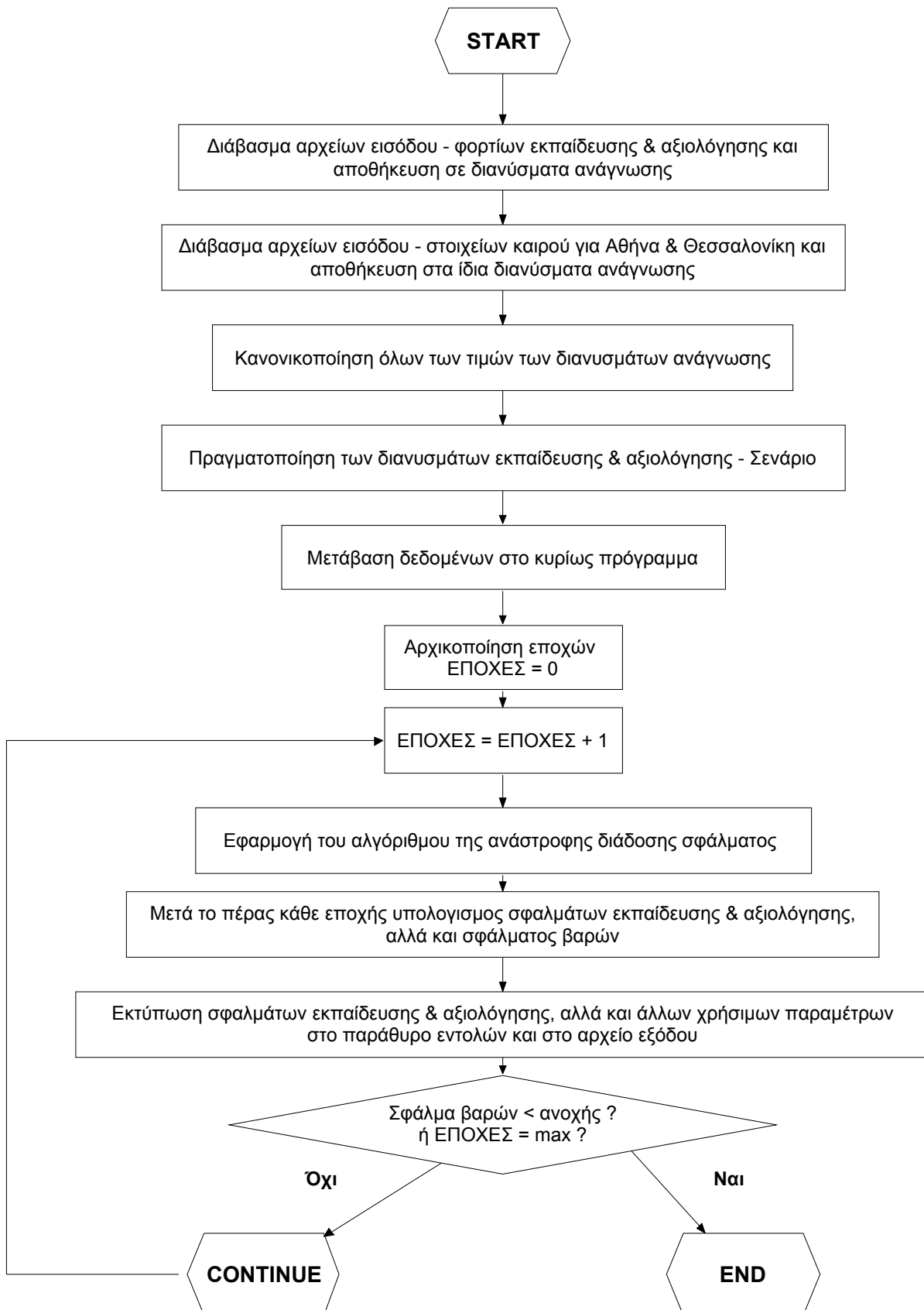
Το πρόγραμμα που χρησιμοποιήσαμε, αρχικά εκπαιδεύει ένα νευρωνικό δίκτυο δύο επιπέδων με τον γνωστό αλγόριθμο της ανάστροφης διάδοσης σφάλματος -Back Propagation, με βάση ένα σύνολο διανυσμάτων εκπαίδευσης που αναφέρονται σε συγκεκριμένα έτη και περιλαμβάνουν κάποιους συνδυασμούς φορτίων, τιμών θερμοκρασιών, αλλά και υγρασιών.

Στη συνέχεια εκτελεί τη λεγόμενη αξιολόγηση, δηλαδή υπολογίζει το συνολικό σφάλμα εξόδου για ένα σύνολο διανυσμάτων αξιολόγησης, που αναφέρονται στο αμέσως επόμενο έτος από αυτό της εκπαίδευσης και έχουν φυσικά ακριβώς την ίδια μορφή με τα διανύσματα εκπαίδευσης.

Ο σκοπός του συγκεκριμένου προγράμματος, είναι η δημιουργία τελικά ενός νευρωνικού δικτύου, με δεδομένες τιμές βαρών, στο οποίο αν εφαρμοστεί στην είσοδο ένα διάνυσμα με τα στοιχεία που αντιστοιχούν στην επόμενη μέρα από αυτή που διανύουμε, τότε θα μπορεί να μας δώσει στην έξοδο και με ένα ικανοποιητικό σφάλμα, την πρόβλεψη για τα φορτία της επόμενης αυτής ημέρας.

Ας αναφερθούμε τώρα πιο αναλυτικά τόσο για το περιεχόμενο του κώδικα του προγράμματος, όσο και για τα αρχεία εισόδου - εξόδου που αυτό επεξεργάζεται.

Αρχικά στην επόμενη σελίδα παραθέτουμε το γενικό διάγραμμα ροής του προγράμματος, για να δούμε σε πρώτη φάση τα βασικά του σημεία.



Γενικό διάγραμμα ροής του προγράμματος εκπαίδευσης T.N.A.

Station ID	Month	Day	Temp 1	Temp 2	Temp 3	Temp 4	Temp 5	Temp 6	Temp 7
716	95	419	11.4	9.8	13.4	18.0	18.6	18.2	15.2
716	95	420	10.6	10.4	14.6	18.6	19.6	18.4	15.8
716	95	421	9.8	9.2	14.6	20.0	19.4	19.8	16.6
716	95	422	11.4	10.0	15.6	21.4	21.8	21.2	18.4
716	95	423	13.8	12.8	17.4	21.6	23.0	23.2	19.6
716	95	424	17.4	16.8	16.8	21.4	25.0	23.0	19.6
716	95	425	15.0	13.6	16.8	19.4	20.6	21.0	19.6
716	95	426	18.2	16.6	17.6	20.8	22.4	22.8	17.2
716	95	427	15.2	14.4	16.4	19.8	20.8	19.4	17.8
716	95	428	15.6	14.4	16.2	20.2	22.4	19.4	17.2

Εικόνα 4.2-2 : Παράδειγμα αρχείου καιρού για Αθήνα

Το τέταρτο και τελευταίο αρχείο καιρού -WEATHER2.txt- είναι ακριβώς ίδιο με το προηγούμενο, με τη μόνη διαφορά ότι αναφέρεται στη Θεσσαλονίκη.

Ας αναφερθούμε τώρα στο περιεχόμενο του προγράμματος ξεκινώντας από την περιγραφή σχετικά με το διάβασμα των αρχείων εισόδου.

Αρχικά ορίζουμε τη μορφοποίηση –FORMAT- των γραμμών που θα χρησιμοποιηθούν. Με τον όρο FORMAT μιας γραμμής, εννοούμε τον ακριβή προσδιορισμό κατά σειρά όλων των μεταβλητών που πρόκειται να διαβάσουμε (για αρχεία εισόδου) ή να εκτυπώσουμε (για αρχεία εξόδου ή πληκτρολόγιο) διαδοχικά σε αυτήν τη γραμμή, όσον αφορά το είδος, αλλά και το μέγεθος τους. Το FORMAT συνεπώς για τα αρχεία εκπαίδευσης και αξιολόγησης των φορτίων είναι το εξής:

15 FORMAT (I4, I2, I2, I2,I2, 24(F8.2))

Η παραπάνω δήλωση μας λέει ότι σε κάθε γραμμή περιμένουμε να διαβάσουμε κατά σειρά έναν ακέραιο αριθμό 4 ψηφίων (έτος), τέσσερις ακέραιες αριθμούς των 2 ψηφίων ο καθένας (μήνας, μέρα, είδος της ημέρας, κατηγορία της ημέρας) και τέλος 24 πραγματικούς αριθμούς συνολικών ψηφίων 8 ο καθένας, εκ των οποίων τα 2 ψηφία είναι δεκαδικά.

Επίσης το FORMAT για τα αρχεία καιρού 1 και 2 είναι το παρακάτω :

16 FORMAT (I3,I3,I2,I2,16(F6.1))

το οποίο με τη σειρά του μας λέει ότι σε κάθε γραμμή αυτών των αρχείων περιμένουμε να διαβάσουμε, πρώτα απ' όλα έναν τριψήφιο ακέραιο που αντιπροσωπεύει τον σταθμό –Αθήνα ή Θεσσαλονίκη, ακολούθως διαβάζουμε άλλον έναν τριψήφιο ακέραιο που είναι το έτος και στη συνέχεια δύο διψήφιους ακέραιους που είναι ο μήνας και η μέρα. Τέλος ακολουθούν οι 8 τρίωρες

τιμές της θερμοκρασίας και οι 8 της υγρασίας η κάθε μια εκ των οποίων έχει 6 ψηφία με τα δύο να είναι δεκαδικά.

Όσον αφορά τις παραμέτρους, αρχικά αναφέρουμε ότι `ip_dim1` (=3000) είναι το μέγιστο των διανυσμάτων ανάγνωσης που μπορούμε να αποθηκεύσουμε και ως εκ τούτου προσδιορίζει το μέγιστο των στηλών οποιωνδήποτε πινάκων ανάγνωσης. Η παράμετρος `ip_dim2` (=24) αφορά φυσικά τον αριθμό των φορτίων της κάθε ημέρας που εμφανίζονται στα αρχεία εκπαίδευσης και αξιολόγησης, ενώ οι παράμετροι `ip_dim_t` (=8) και `ip_dim_y` (=8) αφορούν τις 8 τρίωρες τιμές της θερμοκρασίας και της υγρασίας αντίστοιχα ανά ημέρα, που υπάρχουν στα αρχεία καιρού. Το άθροισμα λοιπόν των τριών αυτών παραμέτρων `ip_dim2_f`, προσδιορίζει το πλήθος των γραμμών του πίνακα διανυσμάτων ανάγνωσης των δεδομένων.

Το πρόγραμμα μας λοιπόν ξεκινάει με το άνοιγμα του αρχείου εισόδου “LOAD_TR”, που είναι το αρχείο των φορτίων εκπαίδευσης. Στη συνέχεια διαβάζει γραμμή-γραμμή το αρχείο με το FORMAT 15 και αποθηκεύει τα στοιχεία κάθε γραμμής πρώτα στις μεταβλητές IYEAR, IMONTH, IDAY, IKDAY, IFLDAY, RLOAD (πίνακας μονοδιάστατος 24 θέσεων) και αμέσως μετά τοποθετεί τους οδηγούς στον δισδιάστατο πίνακα IBSTOIXEIA (`ip_dim4,ip_dim1`), όπου `ip_dim4`=5. Συγκεκριμένα οι οδηγοί κάθε γραμμής αποθηκεύονται διαδοχικά κατά στήλες με την μορφή :

```
IBSTOIXEIA(1,στήλη)=IYEAR,  
IBSTOIXEIA(2,στήλη)=IMONTH,  
IBSTOIXEIA(3,στήλη)=IDAY,  
IBSTOIXEIA(4,στήλη)=IKDAY  
IBSTOIXEIA(5,στήλη)=IFLDAY
```

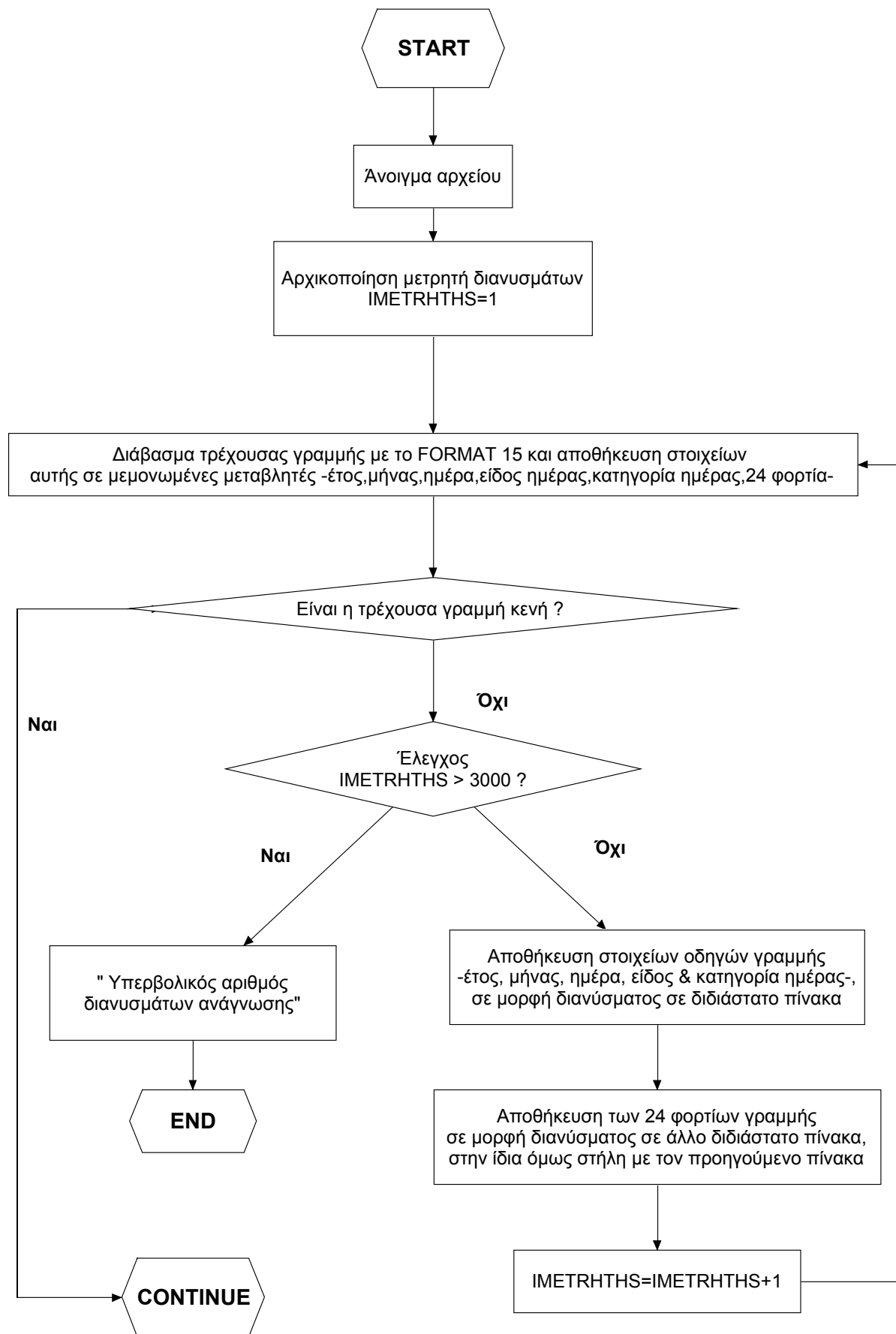
Κατόπιν το πρόγραμμα αποθηκεύει και τα φορτία κατά στήλες στον πίνακα RDEDOMENA (`ip_dim2_f, ip_dim1`), με την στήλη της κάθε ημέρας να είναι η ίδια και στους δύο πίνακες. Μόλις φτάσουμε στην τελευταία γραμμή του αρχείου, τότε σταματάει η αποθήκευση και κλείνουμε το αρχείο.

Στη συνέχεια ανοίγουμε το αρχείο “LOAD_AJ”, που περιέχει τα φορτία αξιολόγησης και επαναλαμβάνουμε πανομοιότυπα την παραπάνω διαδικασία, με τη μόνη διαφορά ότι οι οδηγοί κάθε γραμμής αποθηκεύονται στον πίνακα IBSTOIXEIA_AJ (`ip_dim4,ip_dim_aj`) και ότι τα φορτία αποθηκεύονται στον πίνακα R_DED_AJ (`ip_dim2_f,ip_dim_aj`), όπου η παράμετρος `ip_dim_aj` (=400) έχει να κάνει με το ότι η αξιολόγηση αφορά ένα έτος, άρα οι στήλες του πίνακα αποθήκευσης αποκλείεται ποτέ να ξεπεράσουν τις 400.

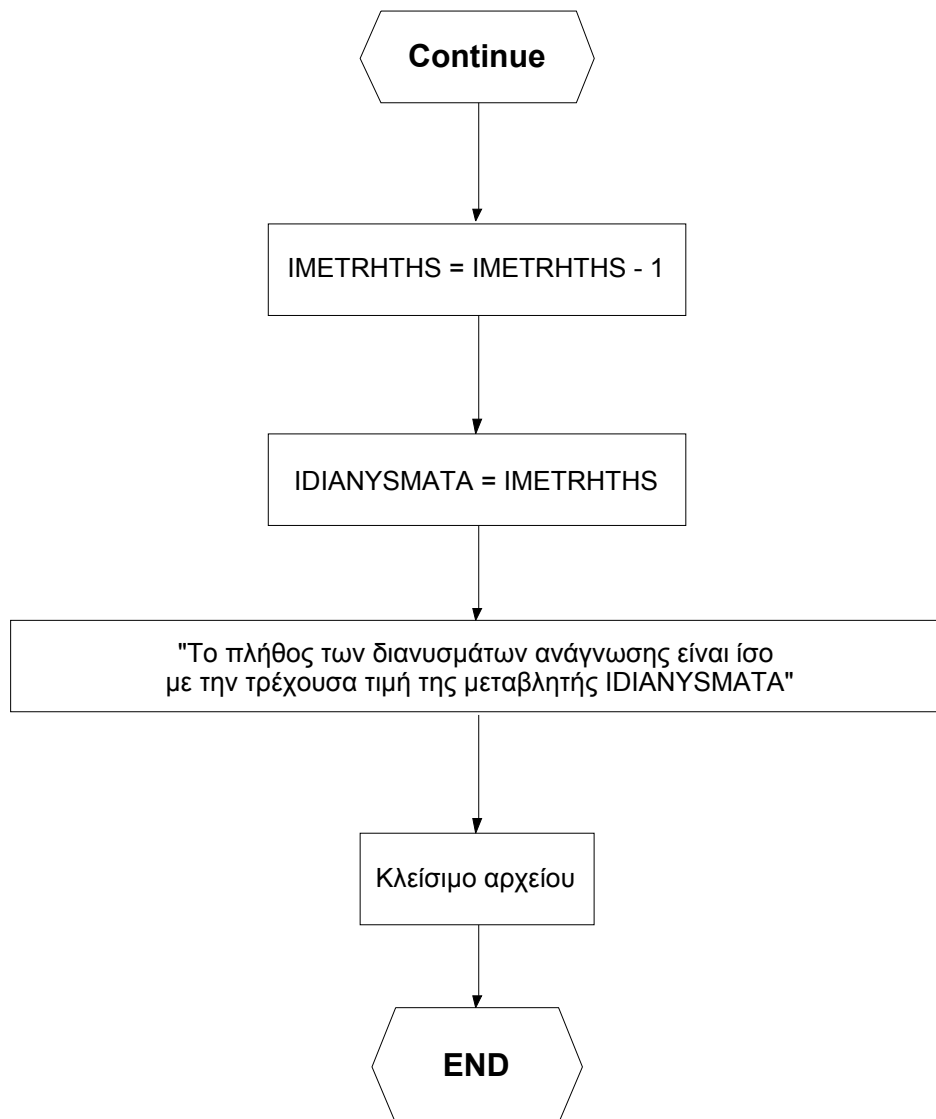
Ύστερα κλείνουμε αυτό το αρχείο και ανοίγουμε το αρχείο “WEATHER1” που αποτελεί το αρχείο καιρού που αντιστοιχεί στην Αθήνα. Διαβάζουμε με το FORMAT 16 κάθε γραμμή και αποθηκεύουμε όλα τα στοιχεία της, προσωρινά, στις μεταβλητές ISTATION, IYEAR, IMONTH, IDAY, R THERMO, RYGRASIA (οι δυο τελευταίοι πίνακες οκτώ θέσεων ο καθένας). Ακολουθώντας διορθώνουμε το έτος, αυξάνοντάς το κατά 1900, ώστε να είναι σε μορφή που να ταυτίζεται με αυτή των αρχείων των φορτίων. Στη συνέχεια για κάθε ημέρα βρίσκουμε με ποια ημέρα από το αρχείο εκπαίδευσης ή αξιολόγησης ταυτίζεται και τότε αποθηκεύουμε στον πίνακα RDEDOMENA (αν ανήκει στην εκπαίδευση) ή στον πίνακα R_DED_AJ (αν ανήκει στην αξιολόγηση), στη στήλη που

αντιστοιχεί σε αυτή τη μέρα, ακριβώς κάτω από τα προϋπάρχοντα φορτία τα 16 στοιχεία θερμοκρασίας και υγρασίας. Αφού προσπελάσουμε όλο το αρχείο, το κλείνουμε και ανοίγουμε το αρχείο “WEATHER2”, που περιγράφει τις καιρικές συνθήκες για τη Θεσσαλονίκη. Επαναλαμβάνουμε με τον ίδιο τρόπο την προαναφερθείσα διαδικασία και αποθηκεύουμε στους ίδιους πίνακες και για την αντίστοιχη ημέρα, τα 16 στοιχεία καιρού για τη Θεσσαλονίκη, ακριβώς μετά από αυτά της Αθήνας.

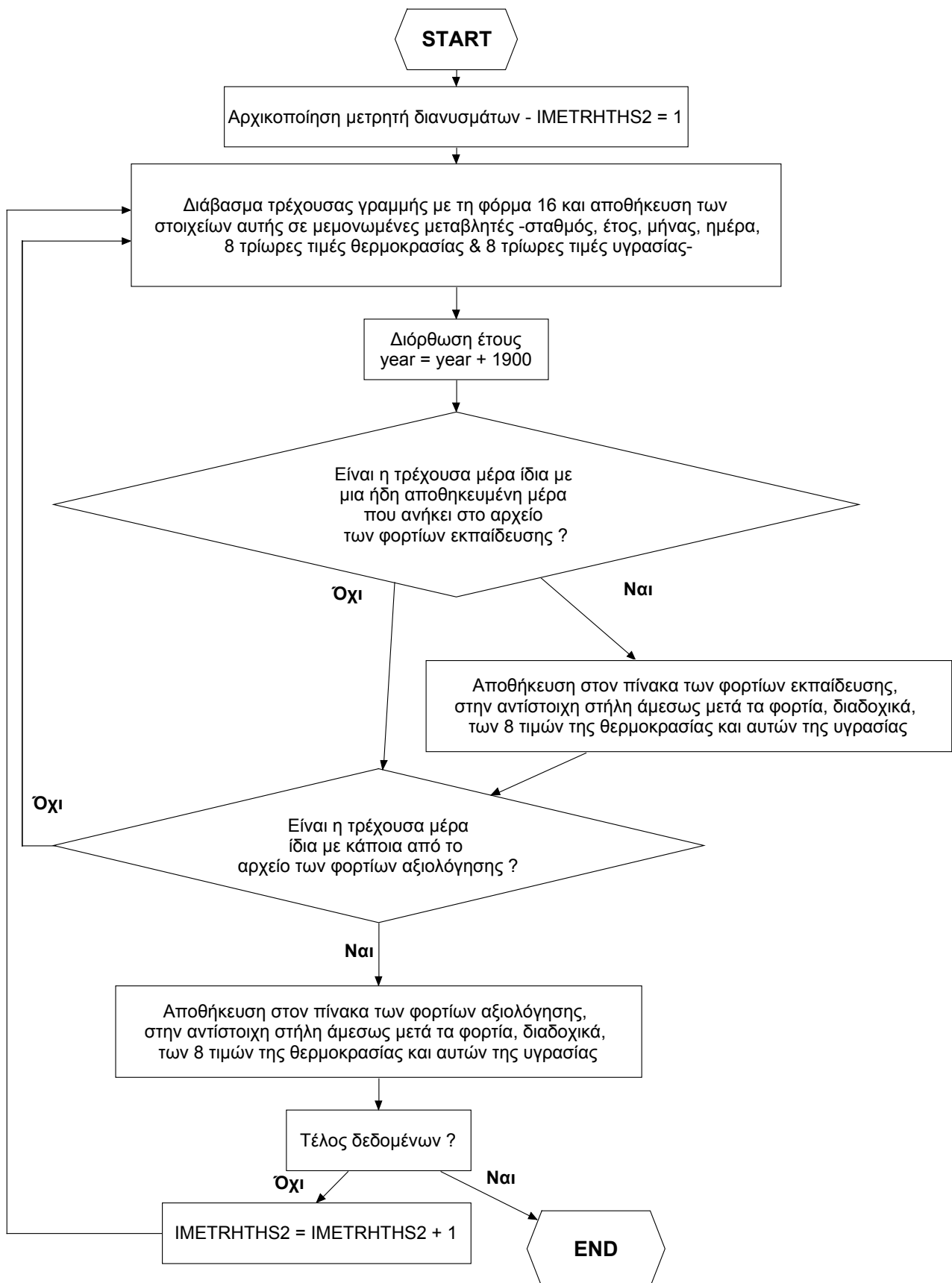
Με αυτόν τον τρόπο, τελικά, έχουμε κατασκευάσει δύο πίνακες για τις μέρες εκπαίδευσης και άλλους δύο για τις μέρες αξιολόγησης, όπου στην ουσία ο ένας είναι προέκταση του άλλου καθώς π.χ. για την ίδια στήλη του IBSTOIXEIA, αλλά και του RDEDOMENA, ο πρώτος περιγράφει τα στοιχεία που χαρακτηρίζουν την μέρα και ο δεύτερος τα φορτία αυτής, αλλά και τις θερμοκρασίες και υγρασίες που την αφορούν. Παρακάτω παραθέτουμε τα διαγράμματα ροής για τη διαδικασία του διαβάσματος από το πρόγραμμα, του αρχείου των φορτίων εκπαίδευσης “LOAD_TR”, αλλά και για τη διαδικασία διαβάσματος του αρχείου καιρού για την Αθήνα “WEATHER1”. Επειδή οι διαδικασίες διαβάσματος για τα αρχεία φορτίων αξιολόγησης “LOAD_AJ” και καιρού για τη Θεσσαλονίκη “WEATHER2” είναι πανομοιότυπες, δεν κρίθηκε σκόπιμο να παρουσιαστούν ανάλογα διαγράμματα.



Διάβασμα αρχείου φορτίων εκπαίδευσης



Διάβασμα αρχείου φορτίων εκπαίδευσης (συνέχεια)



Διάβασμα αρχείου καιρού για Αθήνα

➤ 4.3 Κανονικοποίηση τιμών

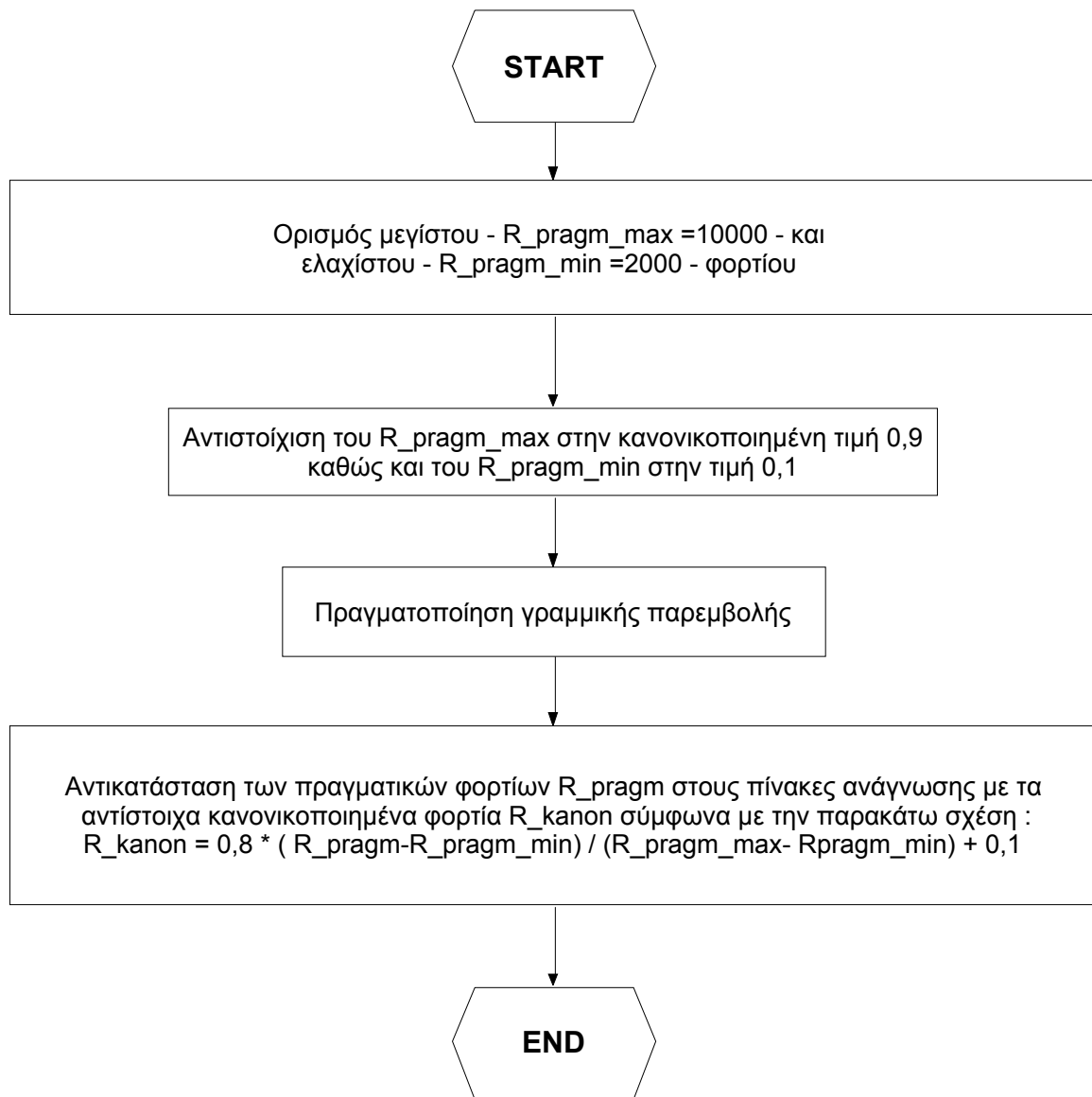
Στη συνέχεια του προγράμματος λοιπόν, δημιουργούμε ένα βοηθητικό αρχείο εξόδου, που το ονομάζουμε “ELEGXOS_EGGRAPHHS” και στο οποίο πρώτα καταγράφουμε σειριακά τις στήλες των δύο πινάκων IBSTOIXEIA και RDEDOMENA και κατόπιν καταγράφουμε τις στήλες των πινάκων IBSTOIXEIA_AJ και του R_DED_AJ. Αυτό γίνεται για να έχουμε τη δυνατότητα να ελέγξουμε στην πράξη, αν όλα τα δεδομένα μας έχουν αποθηκευθεί σωστά, κάτι ιδιαίτερα κρίσιμο φυσικά για την εξέλιξη του προγράμματος.

Ακολούθως προχωράμε στην κανονικοποίηση των τιμών μας, πρώτα για τα φορτία, θεωρώντας ότι η μέγιστη τιμή που μπορούμε να έχουμε είναι 10000 -Rkan_fortia_max- και η ελάχιστη 2000 -Rkan_fortia_min. Η κανονικοποίηση πραγματοποιείται μέσω γραμμικής παρεμβολής για κάθε τιμή, αντιστοιχίζοντας, την ήδη ορισμένη, ελάχιστη τιμή φορτίου στο 0,1 και την, ήδη ορισμένη επίσης, μέγιστη τιμή στο 0,9. Υπολογίζουμε λοιπόν για κάθε στοιχείο των πινάκων RDEDOMENA και R_DED_AJ την αντίστοιχη κανονικοποιημένη τιμή και το αντικαθιστούμε με αυτή. Η ίδια ακριβώς διαδικασία επαναλαμβάνεται για τις τιμές της θερμοκρασίας, όπου επιλέγουμε για μέγιστο Rkan_temp_max =50 και για ελάχιστο Rkan_temp_min = -20, αλλά και για την υγρασία, όπου και επιλέγουμε τις τιμές Rkan_ygr_max =100 και Rkan_temp_min =0.

Σε αυτό το σημείο αξίζει να σημειωθεί, ότι ο σκοπός της κανονικοποίησης είναι η διευκόλυνση του προγράμματος κατά την εκτέλεση των πράξεων αφενός και αφετέρου η καλύτερη εκπαίδευση του νευρωνικού δικτύου, καθώς είναι πιο εύκολο για αυτό να δέχεται τιμές σε ένα δεδομένο και μικρό εύρος, παρά να δέχεται τιμές που να διαφέρουν πάρα πολύ μεταξύ τους, όπως για παράδειγμα η θερμοκρασία 10°C με το φορτίο των 5500 MW.

Αφού ολοκληρωθεί η κανονικοποίηση, τότε ανοίγουμε ένα καινούριο βοηθητικό αρχείο με το όνομα “ELEG_kanoniropoiisis”, όπου εκτυπώνουμε τους πίνακες των δεδομένων, κανονικοποιημένους πλέον, για να διαπιστώσουμε αν όλα έχουν πάει καλά με την κανονικοποίηση.

Στην αμέσως επόμενη σελίδα παρουσιάζεται το διάγραμμα ροής της διαδικασίας κανονικοποίησης των τιμών των φορτίων, μια διαδικασία που, όπως και έχουμε προαναφέρει, είναι εντελώς πανομοιότυπη για τις τιμές της θερμοκρασίας, αλλά και της υγρασίας.



Κανονικοποίηση φορτίων

➤ 4.4 Πραγματοποίηση διανυσμάτων εκπαίδευσης

Στο επόμενο τμήμα του προγράμματος γίνεται λόγος για τα σενάρια, τα οποία δεν είναι τίποτα άλλο, από τον τρόπο πραγματοποίησης των διανυσμάτων εκπαίδευσης. Ουσιαστικά κάθε σενάριο προσδιορίζει με ακρίβεια τα στοιχεία που θα περιλαμβάνει το διάνυσμα εισόδου και εξόδου, τα οποία και απαρτίζουν το κάθε συνολικό διάνυσμα εκπαίδευσης κι ως εκ τούτου προσδιορίζει και το κάθε διάνυσμα αξιολόγησης, αφού είναι πανομοιότυπο με αυτό της εκπαίδευσης, όσον αφορά τον τρόπο δόμησης του. Οι εισοδοί των διανυσμάτων εκπαίδευσης αποθηκεύονται κατά στήλες στον πίνακα R_DEDOM_X (ip_dim_in, ip_dim1) και οι έξοδοι στον πίνακα R_DEDOM_Y (ip_dim_out, ip_dim1), όπου ip_dim_in η διάσταση του διανύσματος

εισόδου και `ip_dim_out` η διάσταση του διανύσματος εξόδου. Ομοίως, όσον αφορά την αξιολόγηση οι αντίστοιχοι πίνακες είναι οι `R_DE_AJ_X` (`ip_dim_in, ip_dim_aj`) και `R_DE_AJ_Y` (`ip_dim_out, ip_dim_aj`).

Ένα απλό παράδειγμα σεναρίου είναι αυτό κατά το οποίο, το κάθε διάνυσμα εκπαίδευσης περιλαμβάνει στην είσοδο του για την τυχαία τρέχουσα μέρα τα 24 ωριαία φορτία της προηγούμενης μέρας, τις 8 τρίωρες θερμοκρασίες της προηγούμενης μέρας, αλλά και της τρέχουσας μέρας για την Αθήνα και για τη Θεσσαλονίκη, ενώ στην έξοδο του περιέχει τα 24 ωριαία φορτία της τρέχουσας μέρας.

Αμέσως μετά την πραγματοποίηση των διανυσμάτων εκπαίδευσης, προχωράμε στη δημιουργία ενός νέου βοηθητικού αρχείου εξόδου που ονομάζουμε “`ELEG_EISODOU001`” και στο οποίο γράφουμε όλες τις εισόδους των διανυσμάτων εκπαίδευσης και αξιολόγησης και κατόπιν τις αντίστοιχες εξόδους, ώστε να μπορεί να γίνει κάποιος σχετικός έλεγχος αν παραστεί ανάγκη.

Παρακάτω θα αναφερθούμε πιο αναλυτικά στα σενάρια εκπαίδευσης του δικτύου, τα οποία κατασκευάσαμε και κατόπιν υλοποιήσαμε στη γλώσσα προγραμματισμού FORTRAN.

➤ 4.5 Μετάβαση δεδομένων στο κυρίως πρόγραμμα

Στην επόμενη φάση του προγράμματος έχουμε τη μετάβαση των δεδομένων στο κυρίως πρόγραμμα πλέον. Εδώ ορίζουμε όλα τα διανύσματα εισόδου και εξόδου για το δίκτυο με έναν τέτοιο τρόπο, ώστε να μπορούμε όποτε θέλουμε, να μην χρησιμοποιούμε όλα τα διανύσματα εκπαίδευσης και αξιολόγησης, αλλά κάποιο μέρος τους. Οι μεταβλητές `IENERGADIANYSMATA` και `IENERGADIAN_AJ` προσδιορίζουν τον ακριβή αριθμό των ενεργών διανυσμάτων εκπαίδευσης και αξιολόγησης, δηλαδή αυτών που θα πάρουν τελικά μέρος στην εκπαίδευση και αξιολόγηση του νευρωνικού δικτύου.

Το επόμενο βήμα είναι η ενεργοποίηση κάποιων συναρτήσεων κατωφλίου, οι οποίες ορίζονται στο τέλος του προγράμματος, στο παράτημα I και ακολούθως η αρχικοποίηση του ρυθμού μάθησης - `trainrate`- και του όρου ορμής -`r_momentum`, παράλληλα με την ανάθεση τιμών των περιόδων μάθησης και ορμής - `trainrate_T`, `r_momentum_T`. Εδώ αξίζει να αναφέρουμε ότι το πρόγραμμα είναι έτσι σχεδιασμένο, ώστε ανάλογα με τις τιμές των παραμέτρων `i_metab_momentum_min`, `i_metab_momentum_max`, `i_T_momentum_min`, `i_T_momentum_max`, `i_metab_trainrate_min`, `i_metab_trainrate_max`, `i_T_trainrate_min`, `i_T_trainrate_max`, που είναι τα άκρα σε κάποιες εντολές μορφής DO, μπορούμε να επαναλάβουμε σχεδόν όλη τη διαδικασία της εκπαίδευσης και αξιολόγησης, για διάφορες τιμές που εμείς επιθυμούμε.

➤ 4.6 Παρουσίαση προτύπων εκπαίδευσης

Στην επόμενη πλέον φάση περνάμε στην αρχικοποίηση των βαρών με τυχαίο τρόπο, χρησιμοποιώντας τη συνάρτηση `RAND(.)` Με `W1, W2` συμβολίζουμε τα βάρη των νευρώνων του πρώτου και δεύτερου επιπέδου αντίστοιχα, ενώ το `pol` σημαίνει βάρη πόλωσης. Η παράμετρος `iqin` είναι ο αριθμός των νευρώνων εισόδου, η `iq1` είναι ο αριθμός των νευρώνων του 1^{ου} επιπέδου, ενώ τέλος η `iqout` ορίζει τον αριθμό των νευρώνων του 2^{ου} επιπέδου -εξόδου. Αφού αρχικοποιήσουμε

τα βάρη, δημιουργούμε ένα νέο αρχείο εξόδου, στο οποίο γράφουμε όλες τις τιμές των βαρών για ένα σχετικό έλεγχο.

Κατόπιν κλείνουμε αυτό το αρχείο και περνάμε στην παρουσίαση των προτύπων εκπαίδευσης που μπορεί να είναι είτε σειριακή, είτε με τυχαίο τρόπο. Σε κάποιες εκτελέσεις του προγράμματος χρησιμοποιήσαμε τον πρώτο τρόπο παρουσίασης και σε κάποιες τον δεύτερο.

Αργότερα θα παρουσιαστούν τα αποτελέσματα με τα ανάλογα σχόλια για τις επιδράσεις στο τελικό αποτέλεσμα και των δύο αυτών τρόπων παρουσίασης προτύπων. Παράλληλα με την παρουσίαση ορίζονται και οι εκθετικές σχέσεις, που μας δείχνουν την εξάρτηση του όρου ορμής και του ρυθμού μάθησης από τις αντίστοιχες περιόδους, αλλά και από τις παρελθούσες εποχές. Μάλιστα από αυτές τις σχέσεις, οι οποίες και παρουσιάζονται παρακάτω, βλέπουμε ότι και οι δύο ρυθμοί φθίνουν με το πέρασμα των εποχών.

$$\text{training_rate} = \text{trainrate_arx} * \exp(-I_EPOCH/\text{trainrate_T}) \quad (4.6-1)$$

$$r_momentum = r_momentum_arx * \exp(-I_EPOCH/r_momentum_T) \quad (4.6-2)$$

όπου I_EPOCH η τρέχουσα εποχή που διανύει το πρόγραμμα.

➤ 4.7 Ο αλγόριθμος της ανάστροφης διάδοσης- Back Propagation

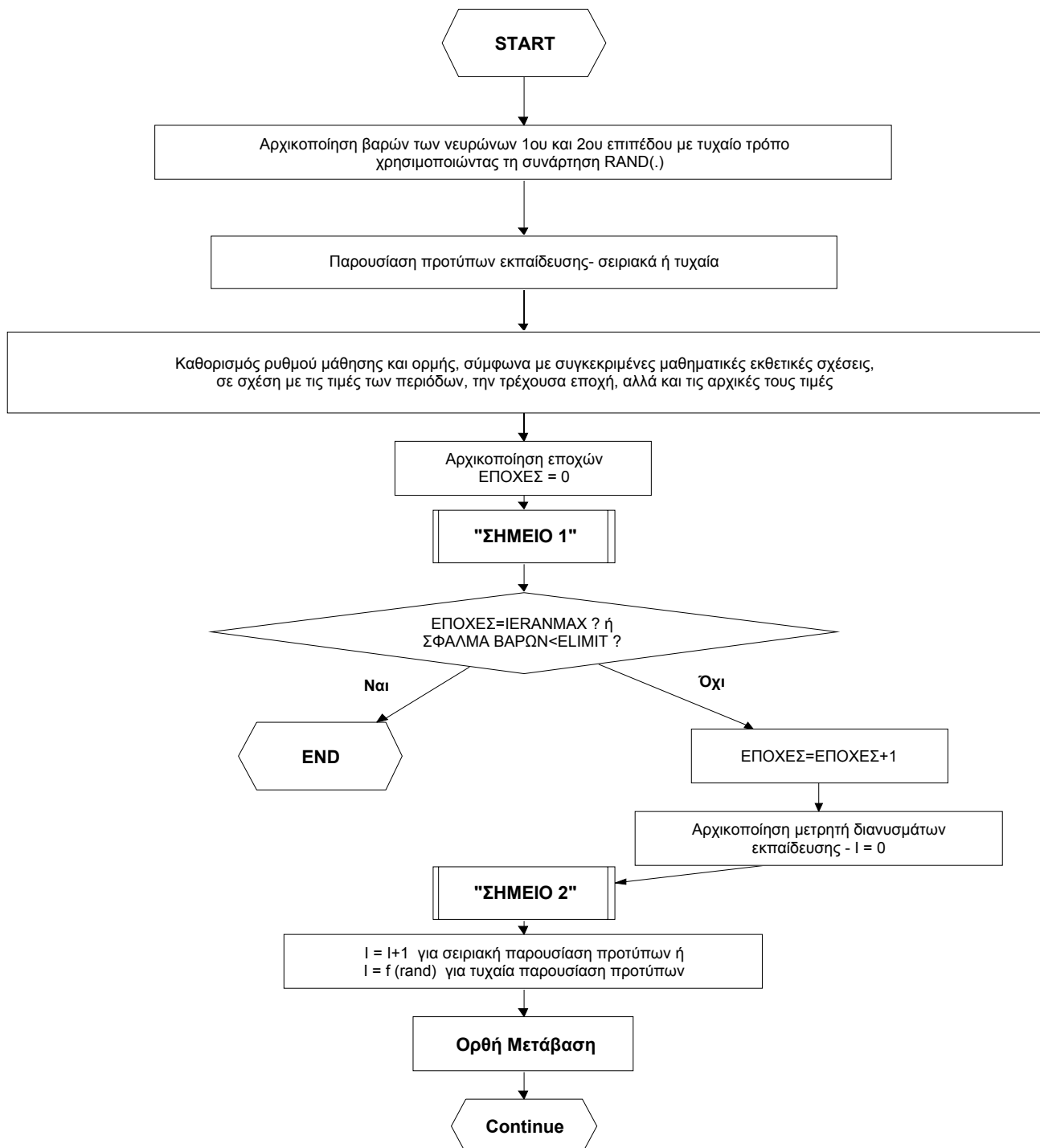
Στο επόμενο πλέον στάδιο περνάμε στο κυρίως μέρος του αλγόριθμου της ανάστροφης διάδοσης -Back Propagation- στο νευρωνικό δίκτυο των δύο επιπέδων. Εδώ ο πίνακας των εισόδων του δικτύου είναι ο XEISODOS (I_iqin , IMET_DIAN), όπου η γραμμή I_iqin εκφράζει το στοιχείο του διάνυσματος εισόδου x^i , ενώ οι στήλες IMET_DIAN μας λένε ποιο συγκεκριμένο διάνυσμα έχουμε επιλέξει. Για κάθε διάνυσμα εισόδου, λοιπόν, υπολογίζεται η απόκριση πρώτα όλων των νευρώνων του 1ου επιπέδου $Y1$ με τη χρήση φυσικά των βαρών $W1$, της αντίστοιχης εισόδου XEISODOS (I_iqin , IMET_DIAN) και της συνάρτησης ενεργοποίησης των νευρώνων SYN_EFARMOGHS($I_SYNFLAG$, $U1(I_q1)$), όπου το $I_SYNFLAG$ εκφράζει το είδος της συνάρτησης, που έχουμε επιλέξει, ενώ $U1(I_q1)$ είναι η αρχική έξοδος του νευρώνα. Στη συνέχεια εφαρμόζεται η ίδια ακριβώς διαδικασία για τους νευρώνες του 2ου επιπέδου, για να υπολογίσουμε αντίστοιχα τις εξόδους $Y2$, χρησιμοποιώντας αυτήν τη φορά τον πίνακα βαρών $W2$ και τις εξόδους $Y1$ των νευρώνων του 1ου επιπέδου.

Ακολούθως υπολογίζουμε όλα τα σφάλματα εξόδου του δικτύου- ERROR_PIN, δηλαδή τη διαφορά των πραγματικών εξόδων $Y2$, από τις επιθυμητές για εμάς, που είναι τα στοιχεία του YEJODOS (I_q2 , IMET_DIAN).

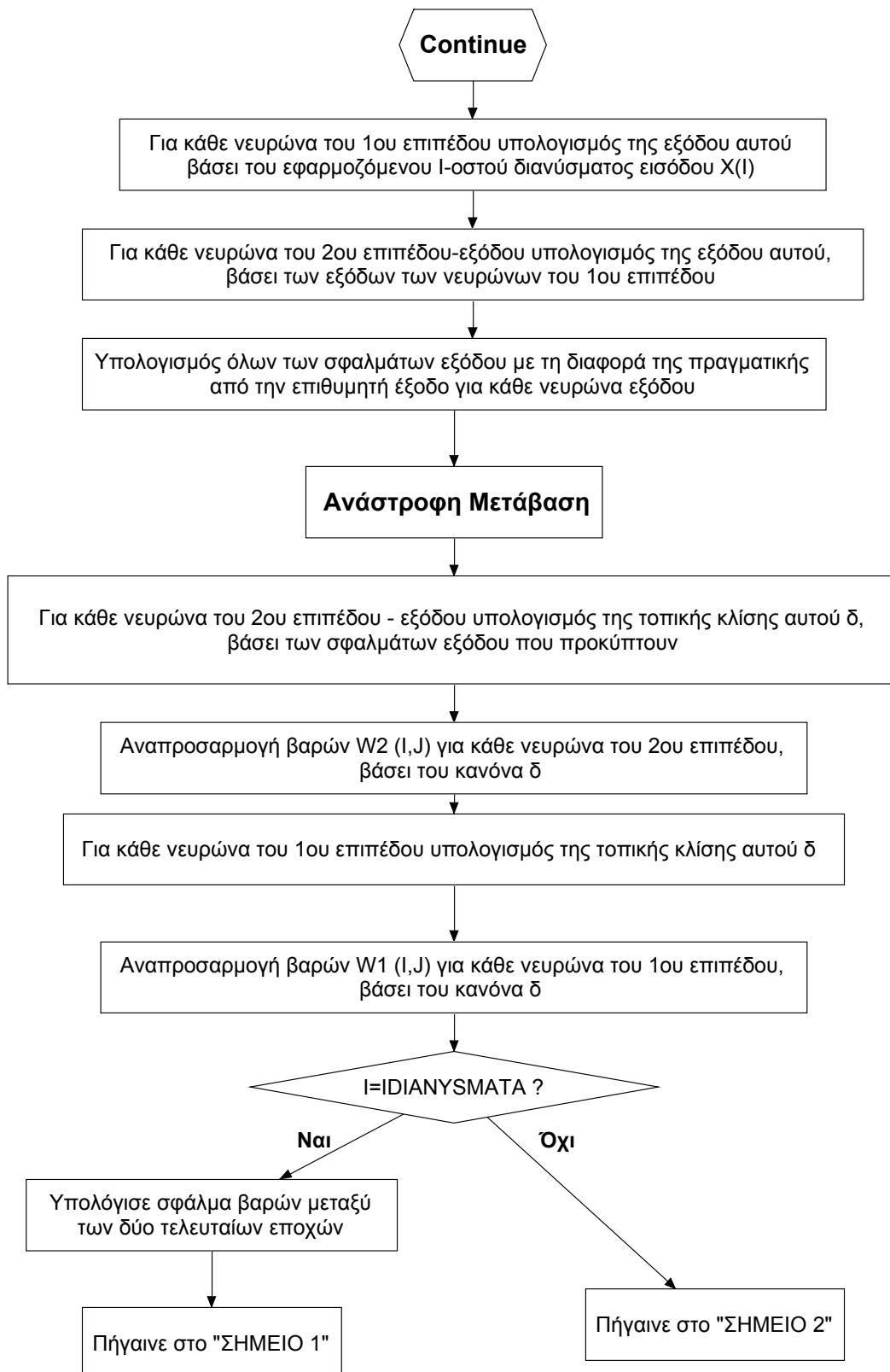
Ύστερα προχωράμε στην αντίστροφη μετάβαση και αρχικά για το 2ο επίπεδο, δηλαδή αυτό της εξόδου, υπολογίζουμε την τοπική κλίση δ για κάθε νευρώνα -Delta2, χρησιμοποιώντας τα σφάλματα εξόδου, αλλά και την παράγωγο της συνάρτησης ενεργοποίησης PARAG_SYN_EF. Ακολούθως περνάμε στην αναπροσαρμογή των βαρών του 2ου επιπέδου, με τη χρήση της γνωστής σχέσης της θεωρίας, όπου $DW2$ είναι ο πίνακας της διαφοράς των βαρών, ο οποίος φυσικά για την πρώτη επανάληψη έχει όλα τα στοιχεία του μηδενικά. Μετά βέβαια επαναλαμβάνουμε όλα τα

παραπάνω και υπολογίζουμε όλα τα δ για το 1^ο επίπεδο -Delta1- και αναπροσαρμόζουμε αντίστοιχα τα βάρη των νευρώνων.

Εδώ και προτού παραθέσουμε το διάγραμμα ροής για τη διαδικασία του αλγόριθμου της ανάστροφης διάδοσης- Back Propagation, θα πρέπει να αναφερθεί ότι στο συγκεκριμένο διάγραμμα τα «κουτάκια», τα οποία περιέχουν τους όρους “ΣΗΜΕΙΟ1” και “ΣΗΜΕΙΟ2”, αντιπροσωπεύουν δύο σημεία του προγράμματος τα οποία και κρίναμε σκόπιμο να ονοματίσουμε. Ο λόγος δεν είναι άλλος, από το ότι πρόκειται για κόμβους στους οποίους το πρόγραμμα επιστρέφει.



Η μέθοδος Back - Propagation



Η μέθοδος Back - Propagation (συνέχεια)

➤ 4.8 Σφάλματα

Στην επόμενη φάση του προγράμματος και αφού έχει ολοκληρωθεί μια εποχή, βρίσκουμε το σφάλμα μεταξύ όλων των βαρών και πολώσεων των νευρώνων, της τωρινής εποχής W1, W1pol, W2, W2pol και όλων των βαρών της προηγούμενης εποχής που έχουν αντίστοιχα αποθηκευθεί στους πίνακες W1palio, W1polpalio, W2palio, W2polpalio. Αν το συνολικό αυτό σφάλμα υπολείπεται του ορίου ανοχής $-ELIMIT=10^{-5}$ που έχουμε θέσει, τότε η διαδικασία εκπαίδευσης σταματάει και οι εποχές δεν προχωρούν άλλο. Επίσης θα πρέπει να πούμε, ότι ο μέγιστος αριθμός των επαναλήψεων ορίζεται από την παράμετρο IERANMAX (=7000).

Επιπρόσθετα μετά το πέρας κάποιας εποχής υπολογίζουμε τα σφάλματα εκπαίδευσης και αξιολόγησης του δικτύου, εφαρμόζοντας ξανά στο δίκτυο τα διανύσματα εκπαίδευσης και επιπλέον αυτά της αξιολόγησης, για τις τρέχουσες τιμές των βαρών. Περιγραφικά τα κυριότερα μεγέθη που υπολογίζονται αλλά και εκτυπώνονται στο αρχείο εξόδου "ELEG_epiloghs001" είναι τα εξής :

- 1) ο αριθμός των εποχών- I_EPOCH,
- 2) η τρέχουσα τιμή του όρου ορμής αλλά και του ρυθμού μάθησης -r_momentum & training_rate,
- 3) η ελάχιστη αλλά και η μέγιστη τιμή του σφάλματος αξιολόγησης -ERROR_SFALMA_TRAIN_MEAN, ERROR_SFALMA_TRAIN_MAX-
- 4) ο αύξων αριθμός του διανύσματος εκπαίδευσης στο οποίο προέκυψε το μέγιστο σφάλμα καθώς και η αντίστοιχη θέση - ώρα αυτού -I_DIANY_ERROR_SFALMA_TRAIN_MAX, I_THESH_ERROR_SFALMA_TRAIN_MAX-
- 5) τα αντίστοιχα μεγέθη για το σφάλμα αξιολόγησης -ERROR_SFALMA_AJ_MEAN, ERROR_SFALMA_AJ_MAX, I_DIANY_ERROR_SFALMA_AJ_MAX, I_THESH_ERROR_SFALMA_AJ_MAX-

Ακόμα σε αυτό το σημείο θα πρέπει να πούμε ότι το πρόγραμμα εκτυπώνει και στο παράθυρο των εντολών του υπολογιστή τις εποχές, καθώς και το μέγιστο και ελάχιστο σφάλμα εκπαίδευσης και αξιολόγησης. Βέβαια υπάρχει η εντολή, ώστε οι εκτυπώσεις να γίνονται ανά 100 εποχές, ώστε τα σημεία που παίρνουμε να μην είναι πάρα πολλά και τα γραφήματα που προκύπτουν να είναι πιο ευανάγνωστα.

Τέλος για την ολοκλήρωση της παρουσίασης του προγράμματος, αμέσως παρακάτω παρουσιάζουμε ένα παράδειγμα «τρεξίματος» μέσα από την εικόνα του παραθύρου των εντολών *Command Prompt* :

```

TND_SENARIO_6
Auto
ENTER INPUT FILE NAME (only 8 characters):
PLHTHOS STOIXEIWN : 1825
ENTER INPUT FILE NAME (only 8 characters):
PLHTHOS STOIXEIWN : 364
ENTER INPUT WEATHER FILE 1 NAME (only 8 characters):
ENTER INPUT WEATHER FILE 2 NAME (only 8 characters):
MEGISTOS ARITHMOS EPANALHFERN : 7000
MEGISTO EPITREPTO SFALMA BARWN : 9.9999997E-05
100 4.659701 33.59733 4.229061 18.87203
200 4.132726 30.98077 3.877539 17.15823
300 3.860289 29.39090 3.707465 16.10243
400 3.687316 28.27610 3.615771 15.36145
500 3.570829 27.43905 3.567769 14.80635
600 3.486024 26.78449 3.540501 14.37362
700 3.422686 26.25819 3.522760 14.02702
800 3.375627 25.82625 3.510902 13.74386
900 3.340230 25.46615 3.503655 13.70656
1000 3.312051 25.16202 3.500091 13.70531
1100 3.289567 24.90263 3.497677 13.70974

```

Εικόνα 4.8-1 : Παράδειγμα παραθύρου εντολών για το πρόγραμμα εκπαίδευσης T.N.Δ.

➤ 4.9 Μεταβολές των ρυθμών μάθησης και ορμής

Αρχικά εκτελέσαμε το πρόγραμμα μας για πλήθος νευρώνων ίσο με 48 και για διάφορες τιμές των παραμέτρων που έχουν να κάνουν με τη μάθηση και την ορμή. Αυτές, όπως και έχουμε προαναφέρει άλλωστε, είναι ο αρχικός ρυθμός μάθησης - trainrate_arx , η περίοδος μάθησης - trainrate_T , ο αρχικός όρος ορμής - r_momentum_arx - και η περίοδος ορμής - r_momentum_T . Το σενάριο που χρησιμοποιήθηκε για αυτές τις δοκιμές ήταν το *σενάριο 1*, ενώ ο τρόπος παρουσίασης των διανυσμάτων εκπαίδευσης ήταν *τυχαίος*.

Συνεπώς κατά το πρακτικό μέρος της εργασίας μας, τρέξαμε το πρόγραμμα πρώτα κρατώντας σταθερές τις παραμέτρους της ορμής και μεταβάλλοντας σημαντικά αυτές της μάθησης, ενώ στη συνέχεια εκτελέσαμε το αντίστροφο με σκοπό πάντα να βελτιστοποιήσουμε, όσο το δυνατόν βέβαια, τις τιμές όλων των προαναφερθέντων.

➤ 4.9.1 Μεταβολές των παραμέτρων μάθησης

Σε αυτήν τη φάση κρατήσαμε σταθερές τις παραμέτρους $\Rightarrow \text{r_momentum_arx} = 0,6$ & $\text{r_momentum_T} = 2000$ και τρέξαμε το πρόγραμμα για όλους τους παρακάτω δυνατούς συνδυασμούς :

- ✓ $\text{trainrate_arx} = 0,3-0,4-0,5-0,6-0,7-0,8$
- ✓ $\text{trainrate_T} = 600-800-1000-1200-1400-1600-1800-2000-2200-2400-2600-2800-3000$

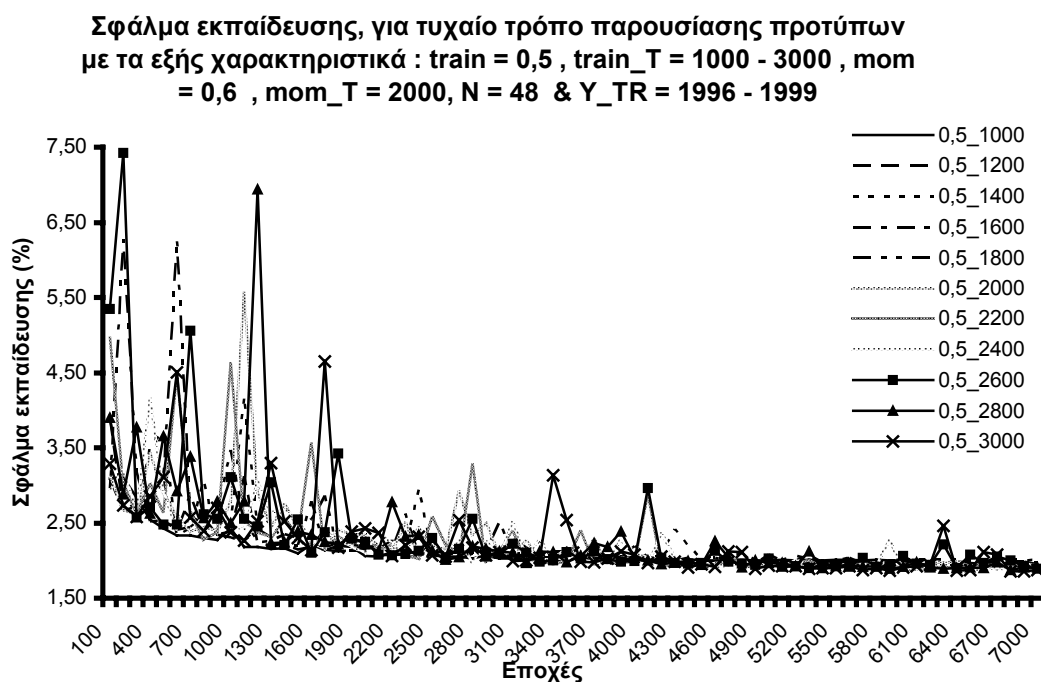
Τα αποτελέσματα που λάβαμε ήταν πάρα πολλά, όπως και τα αντίστοιχα γραφήματα φυσικά και γι' αυτόν το λόγο παρακάτω παρουσιάζονται κάποιες, εκ των γραφικών παραστάσεων

που δείχνουν την μεταβολή του σφάλματος εκπαίδευσης και αξιολόγησης, που είναι άλλωστε αυτά που ενδιαφέρουν, συναρτήσει των εποχών και κρατώντας ως παραμέτρους φυσικά, τις τιμές που αφορούν τους παράγοντες της μάθησης και της ορμής. Τα υπόλοιπα γραφήματα βρίσκονται στο παράρτημα Ι στο τέλος της εργασίας .

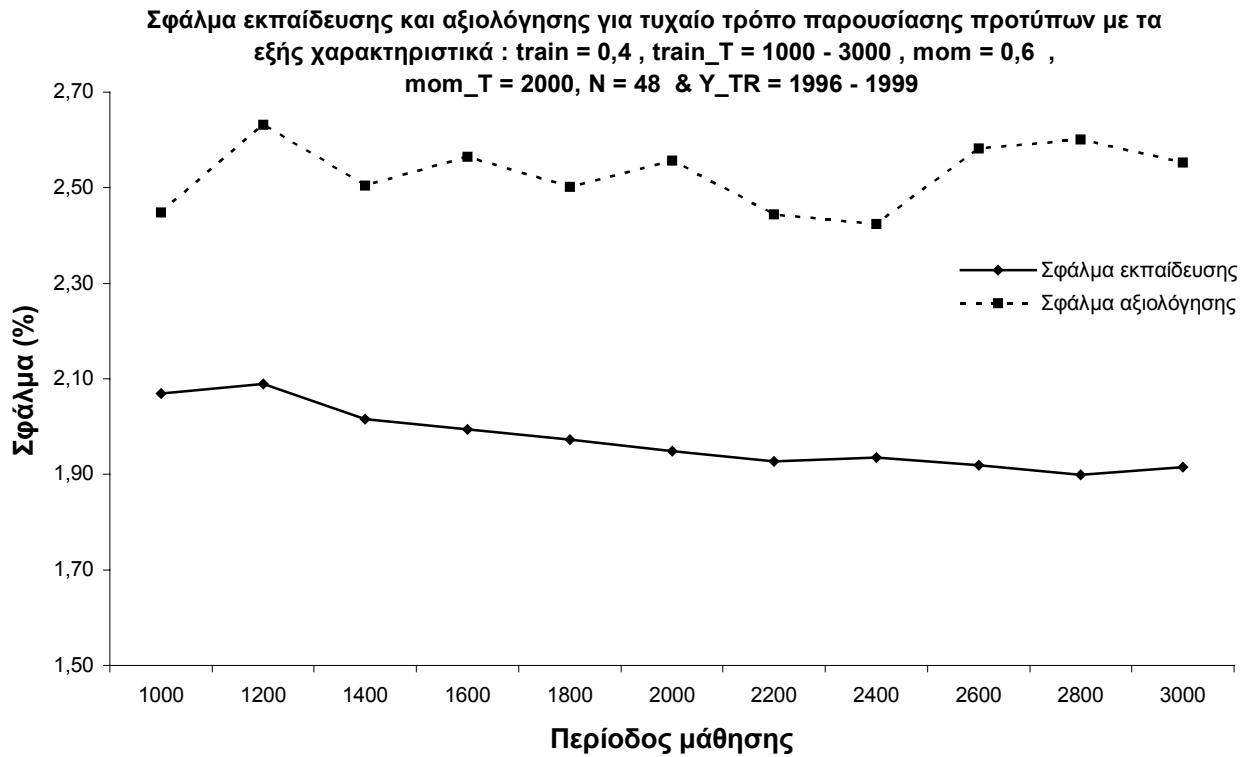
Εδώ θα πρέπει να σημειώσουμε, όσον αφορά τα γραφήματα και τους πίνακες που ακολουθούν από εδώ και πέρα, ότι με *train* και *train_T* συμβολίζουμε τον αρχικό ρυθμό και την περίοδο μάθησης αντίστοιχα, με *mom* και *mom_T* τον αρχικό όρο και την περίοδο ορμής αντίστοιχα, *N* είναι το πλήθος των νευρώνων του κρυμμένου στρώματος και *Y_TR* τα έτη εκπαίδευσης, με το έτος αξιολόγησης να είναι πάντα το 2000.

➤ 4.9.2 Αποτελέσματα για μεταβολή των παραμέτρων μάθησης - Γραφήματα

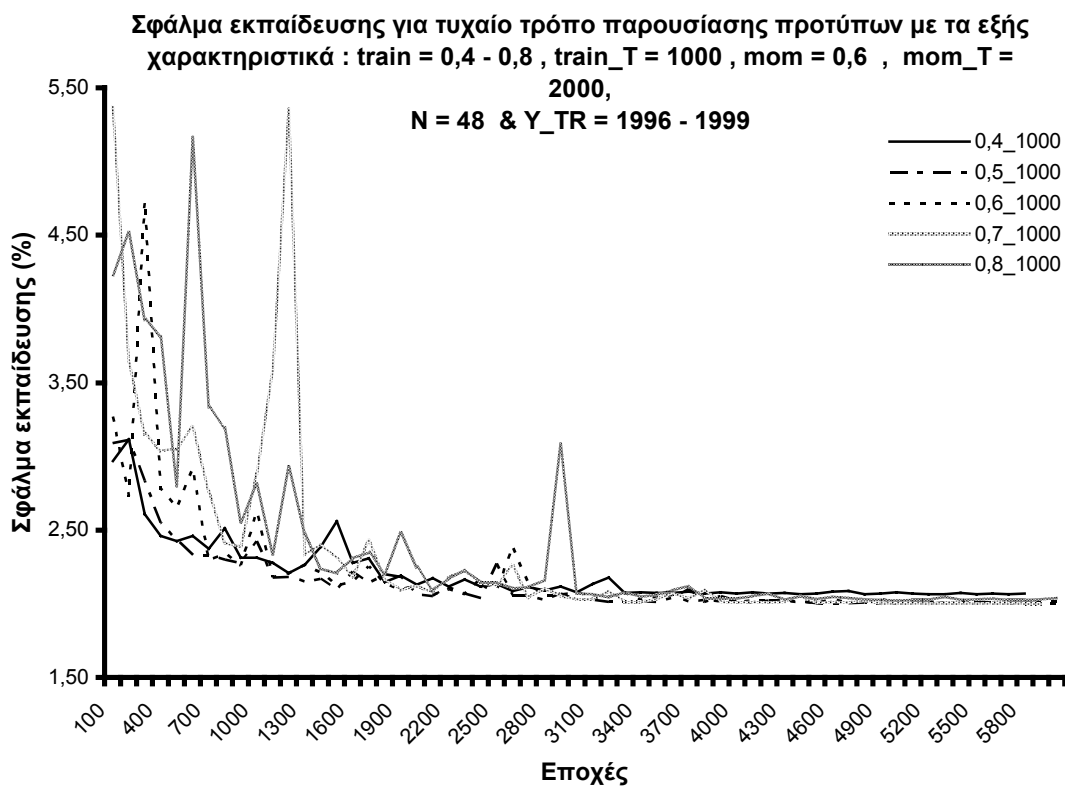
Ακολουθούν ορισμένα ενδεικτικά γραφήματα που προέκυψαν από τη μεταβολή των παραμέτρων μάθησης :



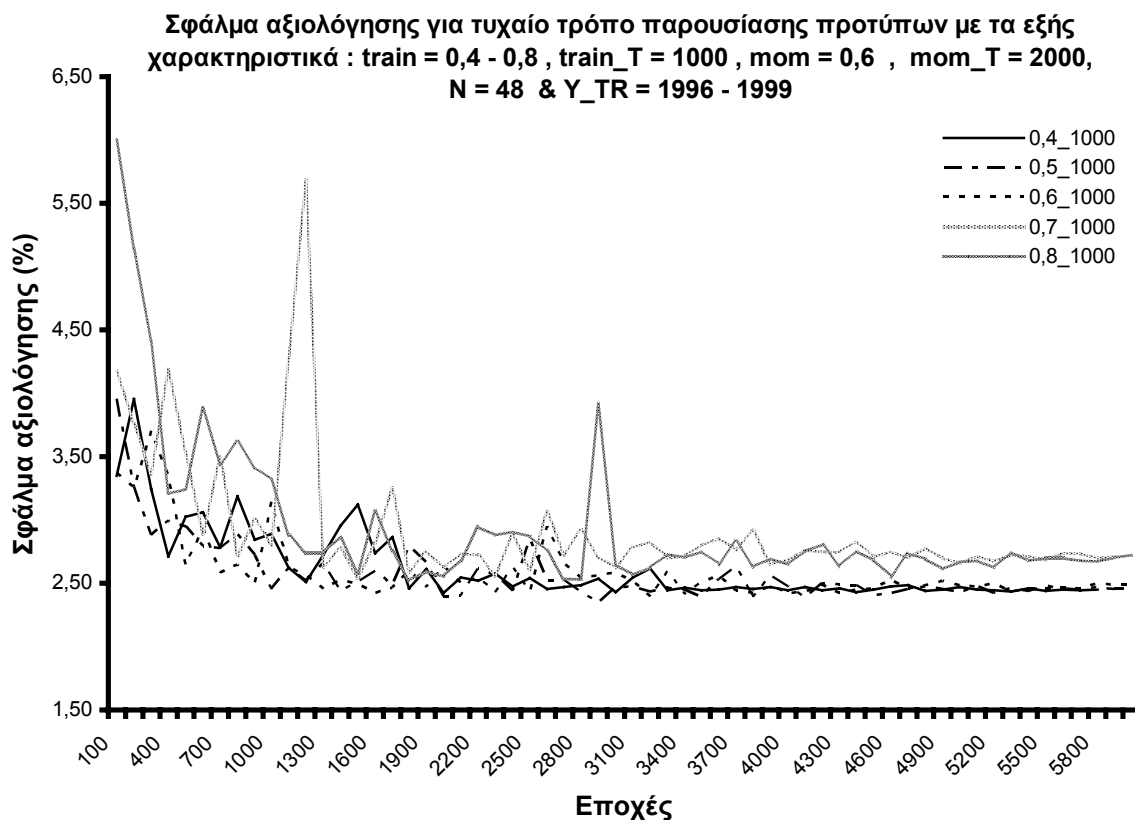
Γράφημα 4.9.2-1



Γράφημα 4.9.2-2



Γράφημα 4.9.2-3



Γράφημα 4.9.2-4

Στη συνέχεια παραθέτουμε κάποιους συγκεντρωτικούς πίνακες δεδομένων :

Μάθησης Περίοδος Μάθησης \ Αρχ.Ρυθμός	0,4	0,5	0,6	0,7	0,8
1000	2,45	2,46	2,50	2,71	2,72
1200	2,63	2,46	3,23	2,51	2,62
1400	2,51	2,40	2,45	2,56	2,61
1600	2,56	2,60	3,22	2,60	2,70
1800	2,50	2,48	2,57	2,73	2,73
2000	2,56	2,49	2,56	3,33	2,65
2200	2,44	2,71	2,59	2,70	2,56
2400	2,42	2,65	2,67	2,71	2,71
2600	2,58	2,71	2,68	2,76	2,63
2800	2,60	2,74	3,03	2,45	2,87
3000	2,55	2,59	2,66	2,85	2,87

Πίνακας 4.9.2-5 : Σφάλμα αξιολόγησης (%) για μεταβλητό αρχικό ρυθμό & περίοδο μάθησης, για τυχαίο τρόπο παρουσίασης προτύπων, με mom = 0,6 , mom_T = 2000, N = 48 & Y_TR = 1997 - 1999

αξιολόγησης είναι 2,4 % και εμφανίζεται για το συνδυασμό τιμών 0,5_1400, ενώ το μέγιστο σφάλμα ξεπερνάει ελάχιστα το 3% και παρουσιάζεται για το συνδυασμό 0,6_2800. Και εδώ έχουμε έντονες διακυμάνσεις και μάλιστα περισσότερες από ότι για το σφάλμα εκπαίδευσης, οι οποίες όμως περιορίζονται όλες αισθητά μετά τις 4000 εποχές και γίνονται ιδιαίτερα αντιληπτές για αρχικούς ρυθμούς μάθησης 0,6 και 0,7.

Γενικά αυτό που θα πρέπει να τονίσουμε για τις δοκιμές που διεξήχθησαν είναι ότι τα αποτελέσματα για τα σφάλματα εκπαίδευσης και αξιολόγησης είναι αρκετά ικανοποιητικά, ενώ κρίνουμε ότι ο βέλτιστος συνδυασμός τιμών είναι 0,5_1400.

➤ 4.9.4 Μεταβολή των παραμέτρων ορμής

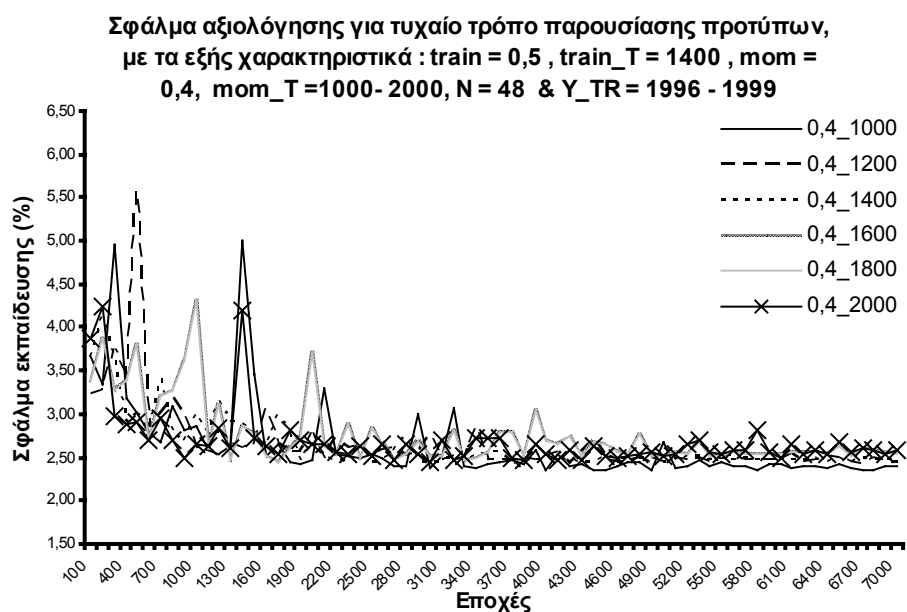
Κατόπιν κρατήσαμε σταθερές τις παραμέτρους $\Rightarrow$ $\text{trainrate_arx} = 0,5$ & $\text{trainrate_T} = 1400$ (ο βέλτιστος συνδυασμός που βρέθηκε πριν) και τρέξαμε το πρόγραμμα για όλους τους παρακάτω δυνατούς συνδυασμούς :

- ✓ $r_momentum_arx = 0,3-0,4-0,5-0,6-0,7-0,8$
- ✓ $r_momentum_T = 500-600-700-800-900-1000-1200-1400-1600-1800-2000$

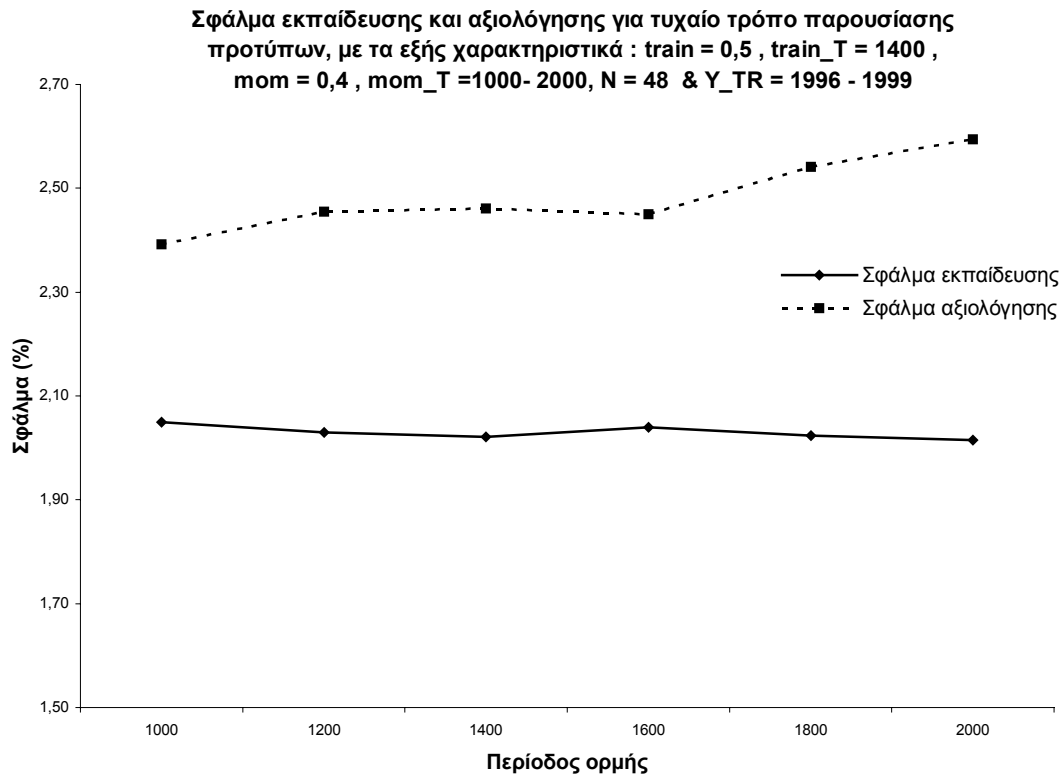
Ακολουθούν μερικές εκ των γραφικών παραστάσεων που προέκυψαν (οι υπόλοιπες βρίσκονται στο παράρτημα I) με τις αντίστοιχες παρατηρήσεις και σχόλια.

➤ 4.9.5 Αποτελέσματα για μεταβολή των παραμέτρων ορμής - Γραφήματα

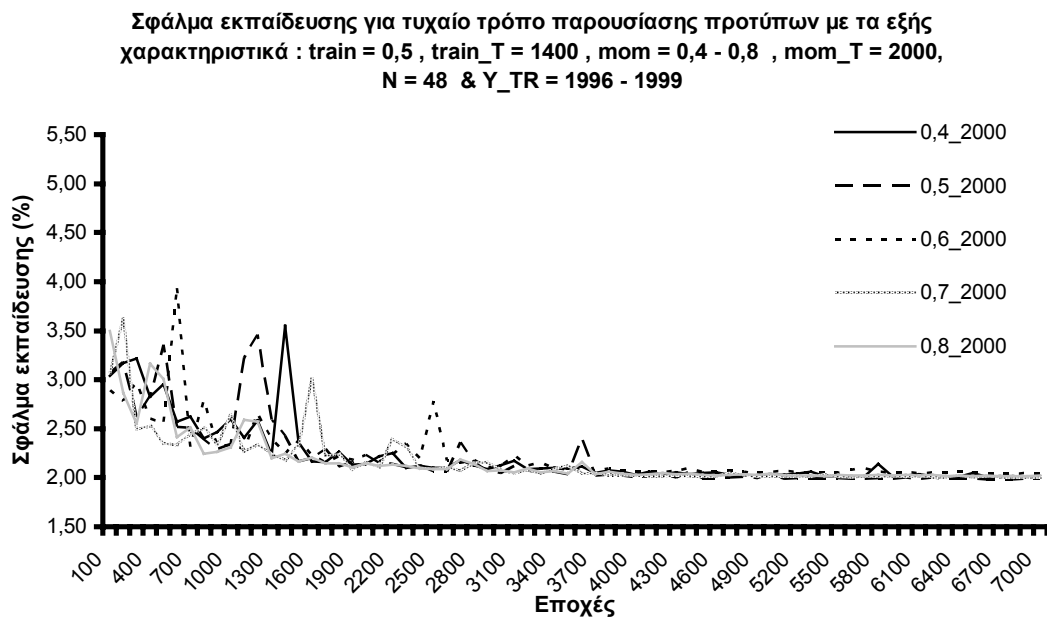
Ακολουθούν ορισμένα ενδεικτικά γραφήματα που προέκυψαν από τη μεταβολή των παραμέτρων ορμής :



Γράφημα 4.9.5-1

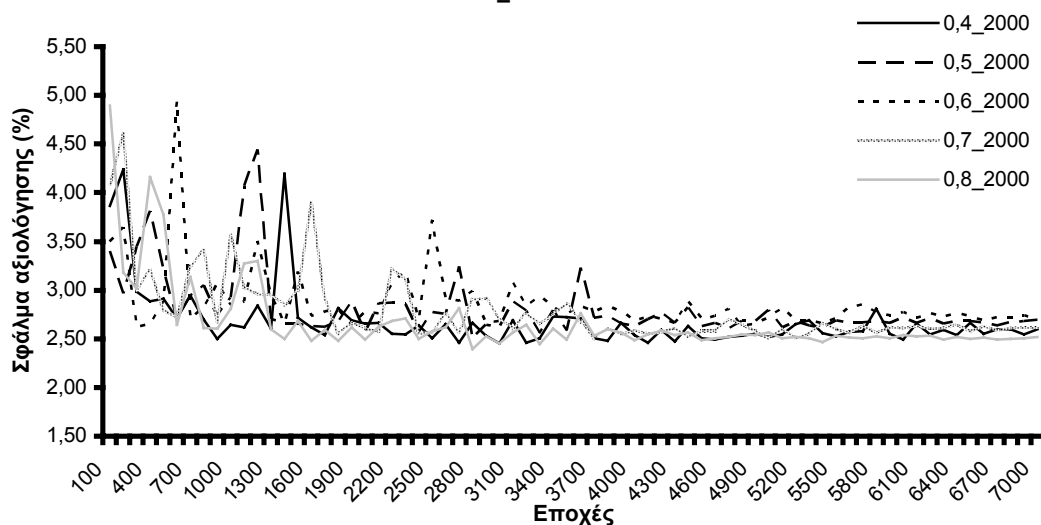


Γράφημα 4.9.5-2



Γράφημα 4.9.5-3

Σφάλμα αξιολόγησης για τυχαίο τρόπο παρουσίασης προτύπων, με τα εξής χαρακτηριστικά : train = 0,5 , train_T = 1400 , mom = 0,4 - 0,8 , mom_T = 2000, N = 48 & Y_TR = 1996 - 1999



Γράφημα 4.9.5-4

Στη συνέχεια παραθέτουμε κάποιους συγκεντρωτικούς πίνακες δεδομένων :

Περίοδος Ορμής \ Αρχ.Ρυθμός Ορμής					
	0,4	0,5	0,6	0,7	0,8
1000	2,39	2,53	2,69	2,62	2,61
1200	2,45	2,49	2,54	2,62	2,61
1400	2,46	2,45	2,63	2,68	2,67
1600	2,45	2,55	2,49	2,69	2,69
1800	2,54	2,59	2,57	2,64	2,60
2000	2,59	2,70	2,69	2,62	2,52

Πίνακας 4.9.5-5 : Σφάλμα αξιολόγησης (%), για τυχαίο τρόπο παρουσίασης προτύπων, για μεταβλητό αρχικό ρυθμό - περίοδο ορμής με N = 48 , train = 0,5 και train_T = 1400

Περίοδος Ορμής \ Αρχ.Ρυθμός Ορμής					
	0,4	0,5	0,6	0,7	0,8
1000	2,05	1,95	2,02	2,03	2,01
1200	2,03	1,98	1,99	2,04	2,02
1400	2,02	1,99	1,99	2,06	2,02
1600	2,04	1,96	2,02	2,03	2,04
1800	2,02	1,97	1,99	2,01	1,99
2000	2,02	1,99	2,04	2,02	2,02

Πίνακας 4.9.5-6 : Σφάλμα εκπαίδευσης (%), για τυχαίο τρόπο παρουσίασης προτύπων, για μεταβλητό αρχ. ρυθμό - περίοδο ορμής με N = 48 , train = 0,5 και train_T = 1400

Όσον αφορά τη συμπεριφορά του σφάλματος αξιολόγησης, έχουμε να σημειώσουμε ότι κρατώντας σταθερό τον αρχικό όρο ορμής και μεταβάλλοντας την περίοδο δεν εμφανίζεται κάποια συγκεκριμένη τάση για το σφάλμα αυτό, καθώς αυξομειώνεται. Για υψηλές αρχικές τιμές του όρου ορμής της τάξης του 0,7-0,8 βλέπουμε ότι το σφάλμα βαίνει αυξανόμενο με σημαντικούς ρυθμούς.

Η πλειονότητα των τιμών του κυμαίνεται από 2,5 % έως 2,7 % με ελάχιστη τιμή να παρουσιάζεται για 0,4_1000 και να αγγίζει το 2,39 %, ενώ η μέγιστη τιμή σφάλματος αγγίζει με τη σειρά της το 2,7 % για το συνδυασμό τιμών 0,8_1600. Και εδώ οι διακυμάνσεις που παρατηρούνται είναι περισσότερες από ότι για το σφάλμα εκπαίδευσης και ενδεικτικά αναφέρουμε ότι η μέγιστη διακύμανση είναι 7,96 % για τις αρχικές επαναλήψεις του συνδυασμού 0,4_900, γεγονός που και αυτό μας φανερώνει κάποιου είδους αστάθεια.

Στη συνέχεια του κεφαλαίου θα αναφερθούμε στα διάφορα σενάρια που υλοποιήσαμε, καθώς και στην κατηγοριοποίηση των ημερών που χρησιμοποιήθηκε, στα σενάρια 4-6.

➤ 4.10 Σενάριο εκπαίδευσης 1

Το πρώτο σενάριο τοποθετεί στην είσοδο του κάθε διανύσματος εκπαίδευσης και αξιολόγησης, για την υπό πρόβλεψη - τρέχουσα ημέρα, τα εξής κατά σειρά δεδομένα :

- ✓ 24 ωριαία φορτία της προηγούμενης μέρας $\Rightarrow$ Θέσεις 1-24
- ✓ 24 ωριαία φορτία της πιο προηγούμενης μέρας $\Rightarrow$ Θέσεις 25-48
- ✓ Μέγιστη και ελάχιστη θερμοκρασία της προηγούμενης μέρας για Αθήνα –σταθμός 1- και Θεσσαλονίκη –σταθμός 2 - $\Rightarrow$ Θέσεις 49-52
- ✓ Μέγιστη και ελάχιστη θερμοκρασία της τρέχουσας μέρας για Αθήνα και Θεσσαλονίκη $\Rightarrow$ Θέσεις 53-56
- ✓ Τετραγωνικές αποκλίσεις των μεγίστων θερμοκρασιών, θεωρώντας αυθαίρετα, ότι $T_{\min}=18$ & $T_{\max}=24$, για Αθήνα και Θεσσαλονίκη, για τρέχουσα και προηγούμενη μέρα $\Rightarrow$ Θέσεις 57-60
- ✓ Διαφορές των μεγίστων θερμοκρασιών της προηγούμενης από την τρέχουσα μέρα για Αθήνα και Θεσσαλονίκη $\Rightarrow$ Θέσεις 61-62
- ✓ Δυαδικός αριθμός 7 ψηφίων του οποίου όλα τα ψηφία είναι 0, πλην της θέσης που αντιστοιχεί στην τρέχουσα για μας ημέρα της εβδομάδας, και εκφράζει την περιοδικότητα της ημέρας ως προς την εβδομάδα. Για παράδειγμα η Τρίτη θα είναι 0100000. $\Rightarrow$ Θέσεις 63-69
- ✓ Αριθμοί που εκφράζουν την περιοδικότητα της ημέρας ως προς το έτος και δίνονται από τις τριγωνομετρικές συναρτήσεις της μορφής $\cos (2\pi/365)$ & $\sin (2\pi/365)$. $\Rightarrow$ Θέσεις 70-71

Το διάνυσμα εισόδου λοιπόν είναι διαστάσεων 71 x 1, ενώ το διάνυσμα εξόδου είναι διαστάσεων 24 x 1, καθώς περιλαμβάνει τα 24 ωριαία φορτία της τρέχουσας μέρας $\Rightarrow$ Θέσεις 1-24. Εδώ αξίζει να αναφερθεί ότι οι μέγιστες και ελάχιστες θερμοκρασίες που χρησιμοποιούνται ως είσοδοι, δεν λαμβάνονται απευθείας από τα δεδομένα των αρχείων καιρού, καθώς αυτά περιέχουν 8 τρίωρες θερμοκρασίες για κάθε ημέρα, αλλά υπολογίζονται μέσω αλγορίθμου στο σενάριο.

Όσον αφορά το σενάριο καθεαυτό, θα λέγαμε ότι βρίσκεται στα πλαίσια της λογικής, καθώς για να προβλέψει τα φορτία της επόμενης ημέρας, θεωρεί ως απαραίτητη προϋπόθεση τη

γνώση όλων των φορτίων της προηγούμενης και της πιο προηγούμενης μέρας, καθώς και κάποια πληροφορία σχετικά με τη μέγιστη και ελάχιστη θερμοκρασία της προηγούμενης μέρας, αλλά και την πρόβλεψη της μέγιστης και ελάχιστης θερμοκρασίας για την επόμενη μέρα. Σε αυτό το σημείο επισημαίνουμε ότι επειδή η πρόβλεψη αφορά το συνολικό φορτίο όλης της Ελλάδας, οποιαδήποτε στοιχεία καιρού –θερμοκρασίες και υγρασίες- που χρησιμοποιούνται προέρχονται από την πρωτεύουσα Αθήνα, αλλά και από την συμπρωτεύουσα Θεσσαλονίκη, όπου υπάρχουν, σε μερικές περιπτώσεις, σημαντικές διαφορές μεταξύ τους. Με αυτόν τον τρόπο προσομοιώνεται στο δίκτυο, ως προς τη συμπεριφορά του καιρού, το τμήμα των καταναλωτών της κεντρικής και βόρειας Ελλάδας που έχει παρόμοιο καιρό με αυτόν της Θεσσαλονίκης, αλλά και το υπόλοιπο τμήμα της κεντρικής και νότιας Ελλάδας, με καιρό που προσεγγίζει αυτόν της Αθήνας.

Επιπλέον στην είσοδο των διανυσμάτων τοποθετούμε και ένα δυαδικό αριθμό που να μας δείχνει ποια μέρα της εβδομάδας είναι η τρέχουσα, ώστε να λαμβάνεται υπ' όψη στο νευρωνικό δίκτυο η περιοδικότητα ως προς την εβδομάδα. Τέλος εισάγονται άλλοι δύο τριγωνομετρικοί αριθμοί $\cos$ & $\sin$, οι οποίοι μας δείχνουν την περιοδικότητα της ημέρας ως προς το έτος αυτήν τη φορά, όταν μιλάμε προφανώς για το σύνολο των διανυσμάτων εκπαίδευσης που έχει μήκος μεγαλύτερο του ενός έτους.

Αρχικά εφαρμόσαμε το σενάριο 1, για *σειριακή* παρουσίαση των προτύπων εκπαίδευσης. Τα έτη εκπαίδευσης ήταν από το 1995-1999 ενώ το έτος αξιολόγησης ήταν το 2000. Εδώ αξίζει να σημειωθεί ότι οι παράμετροι για τις οποίες εκτελέστηκε, στην πρώτη φάση, το σενάριο ήταν οι εξής:

$r_momentum_arx = 0,6$, $r_momentum_T = 2000$, $trainrate_arx = 0,5$ και $trainrate_T = 1400$
οι οποίες και βρέθηκαν, ως ικανοποιητικές με αντίστοιχες δοκιμές.

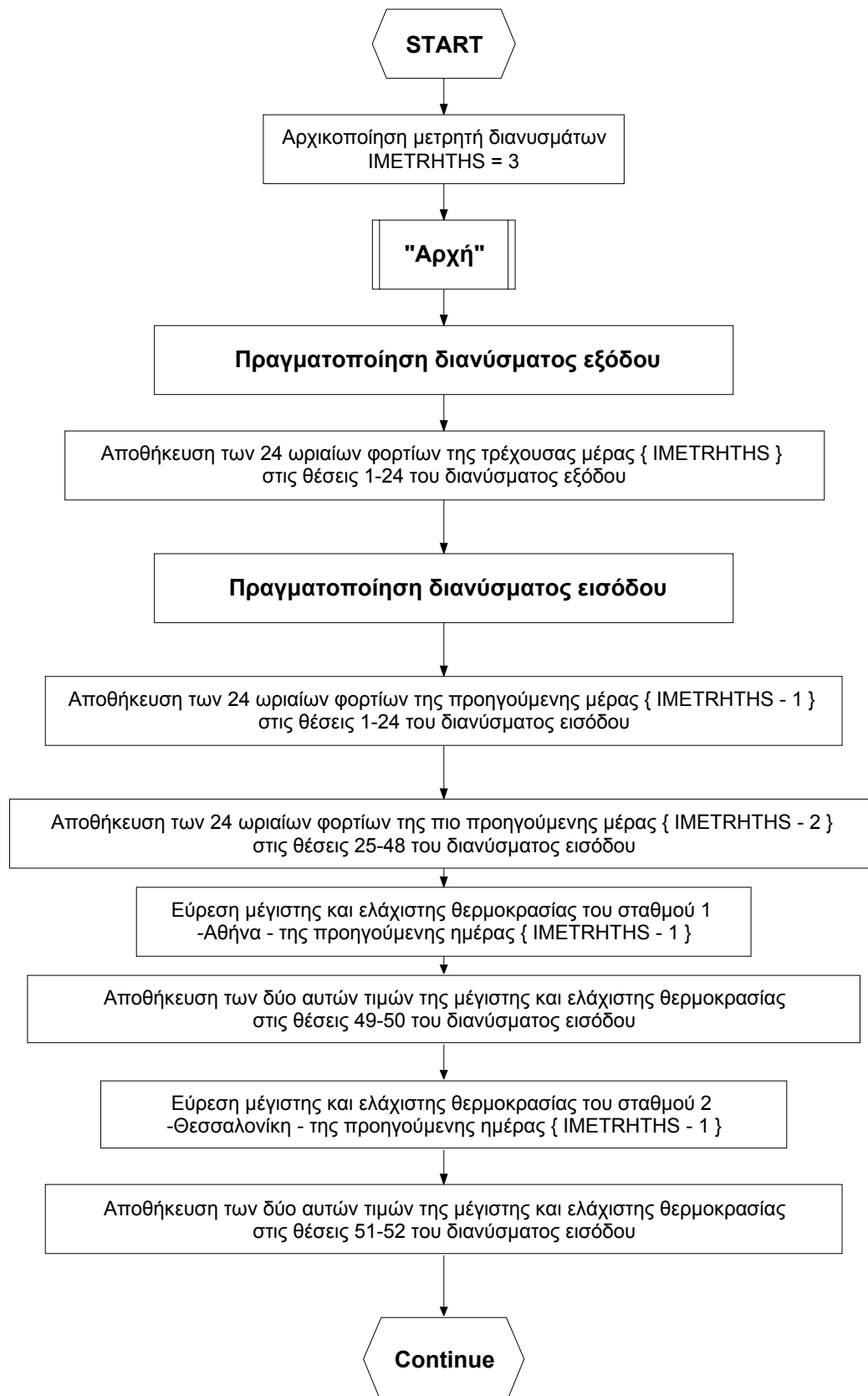
Όσον αφορά τον αριθμό των νευρώνων του κρυμμένου του στρώματος, στην πρώτη φάση “τρέξαμε” το σενάριο, στην πρώτη φάση, για τις τιμές $iq1 = 40-45-48-50-55-60$.

Στη δεύτερη φάση, τρέξαμε το σενάριο για *τυχαία* παρουσίαση προτύπων με τις εξής παραμέτρους :

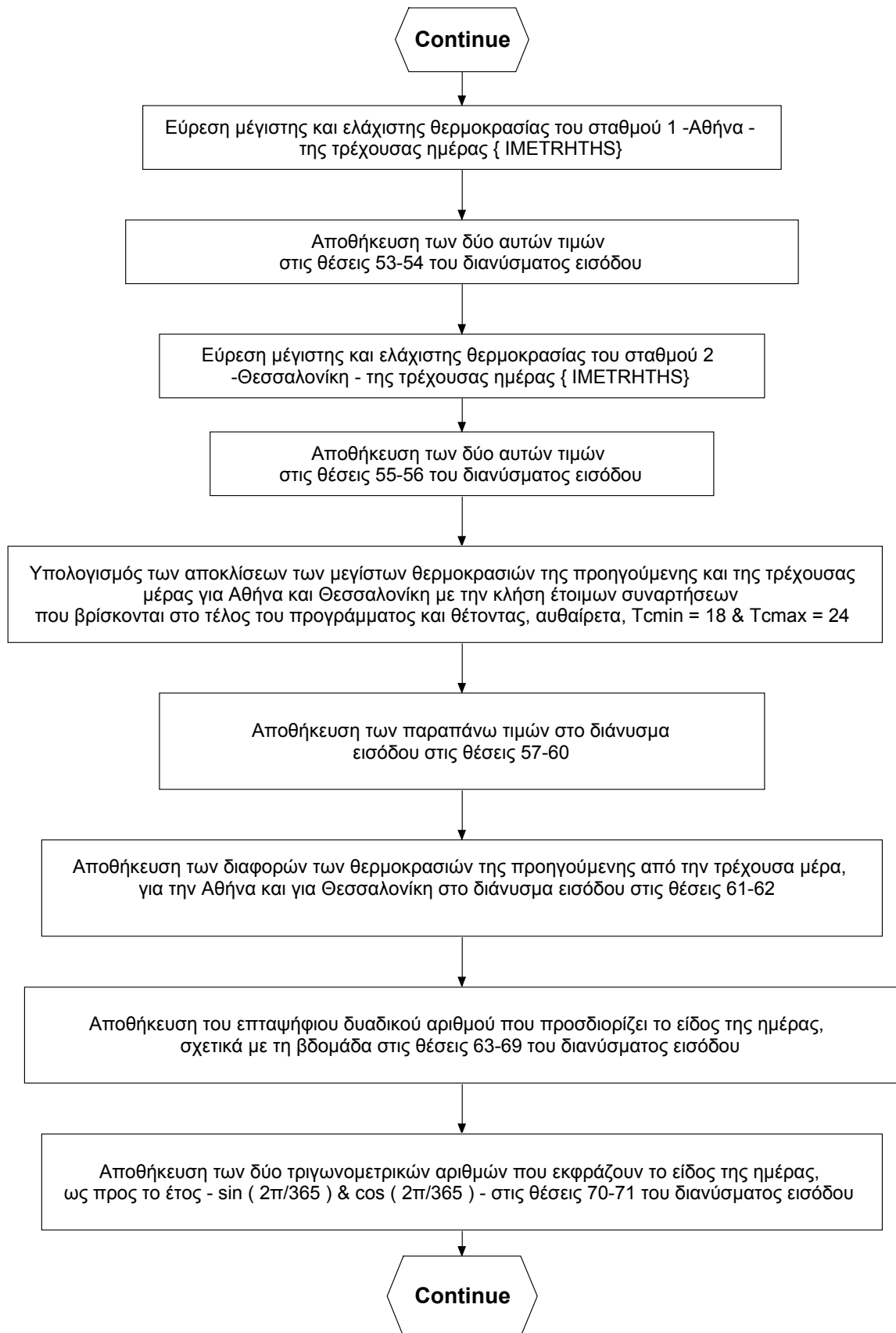
$r_momentum_arx = 0,4$, $r_momentum_T = 1000$, $trainrate_arx = 0,5$ και $trainrate_T = 1400$
και $iq1 = 48$, καθώς από τις προηγούμενες δοκιμές βρέθηκε το καλύτερο πλήθος νευρώνων.

Τα αποτελέσματα καθώς και τα αντίστοιχα σχόλια βρίσκονται στις επόμενες παραγράφους.

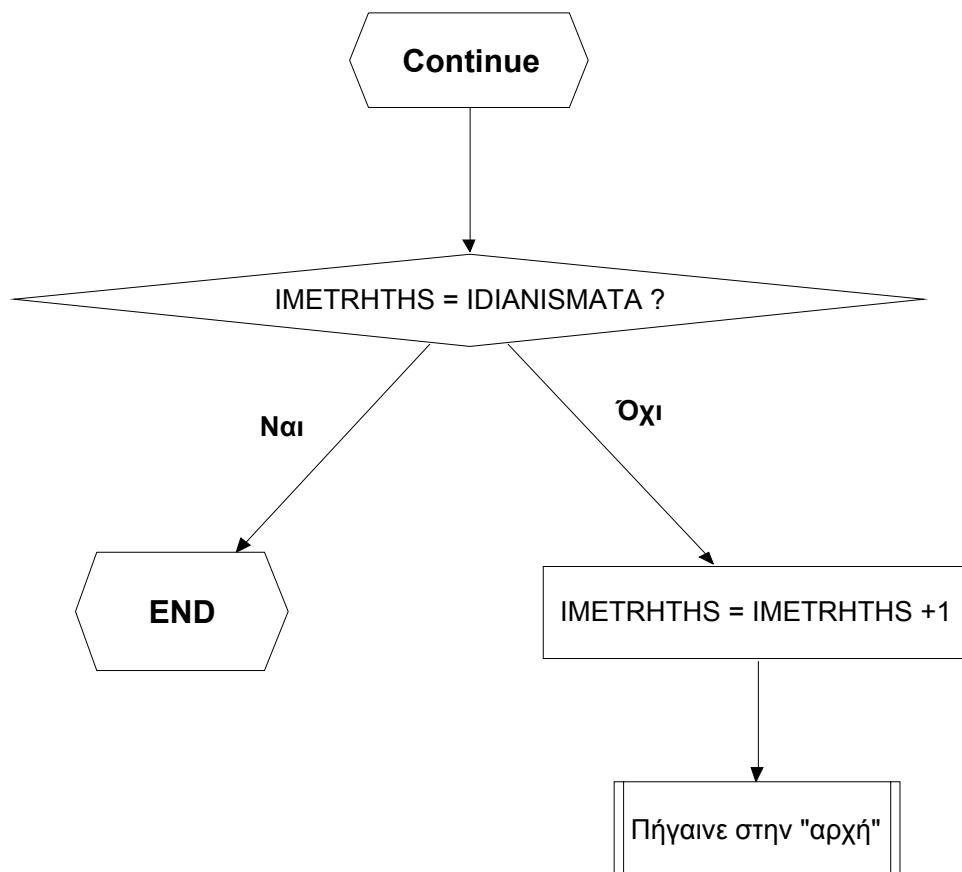
Το διάγραμμα ροής του σεναρίου 1 φαίνεται στην αμέσως επόμενη σελίδα. Όσον αφορά τη διαδικασία δημιουργίας των διανυσμάτων αξιολόγησης αυτή είναι εντελώς ίδια με αυτήν της εκπαίδευσης και ως εκ τούτου δεν κρίθηκε αναγκαία η παρουσίαση αντίστοιχου διαγράμματος.



Πραγματοποίηση διανυσμάτων εκπαίδευσης -Σενάριο 1



Πραγματοποίηση διανυσμάτων εκπαίδευσης- Σενάριο 1 (συνέχεια)



Πραγματοποίηση διανυσμάτων εκπαίδευσης -Σενάριο 1 (συνέχεια)

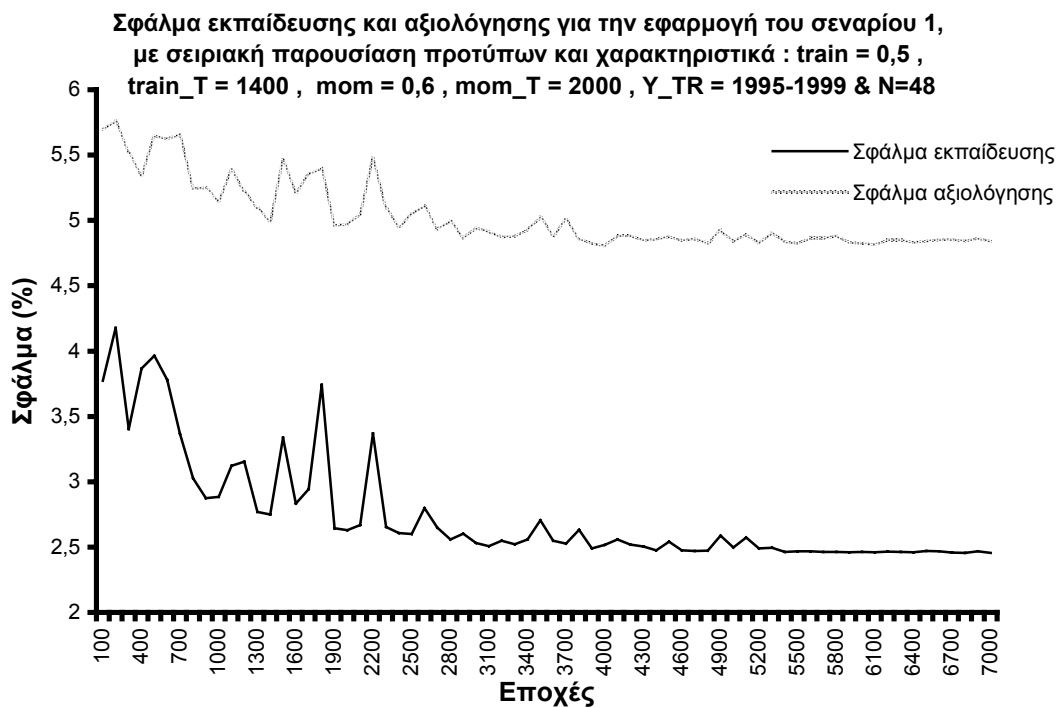
4.10.1 Αποτελέσματα σεναρίου 1 – Γραφήματα

Ακολουθεί συγκεντρωτικός πίνακας για το σενάριο εκπαίδευσης 1 :

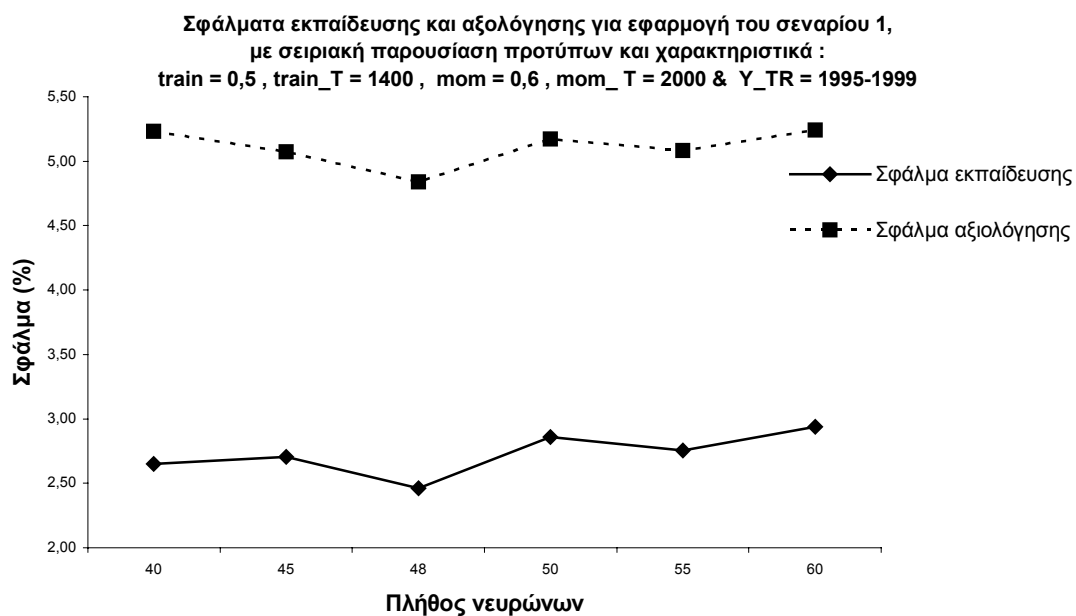
Πλήθος νευρώνων	40	45	48	50	55	60
Σφάλμα εκπαίδευσης (%)	2,65	2,70	2,46	2,86	2,75	2,94
Σφάλμα αξιολόγησης (%)	5,23	5,07	4,84	5,17	5,08	5,24

**Πίνακας 4.10.1-1 : Συγκεντρωτικά αποτελέσματα σεναρίου 1 για σειριακή παρουσίαση προτύπων και για τα χαρακτηριστικά : train = 0,5 train_T = 1400 ,
mom = 0,6 , mom_T = 2000 και Y_TR = 1995-1999**

Στη συνέχεια παρουσιάζουμε και τα αντίστοιχα γραφήματα :

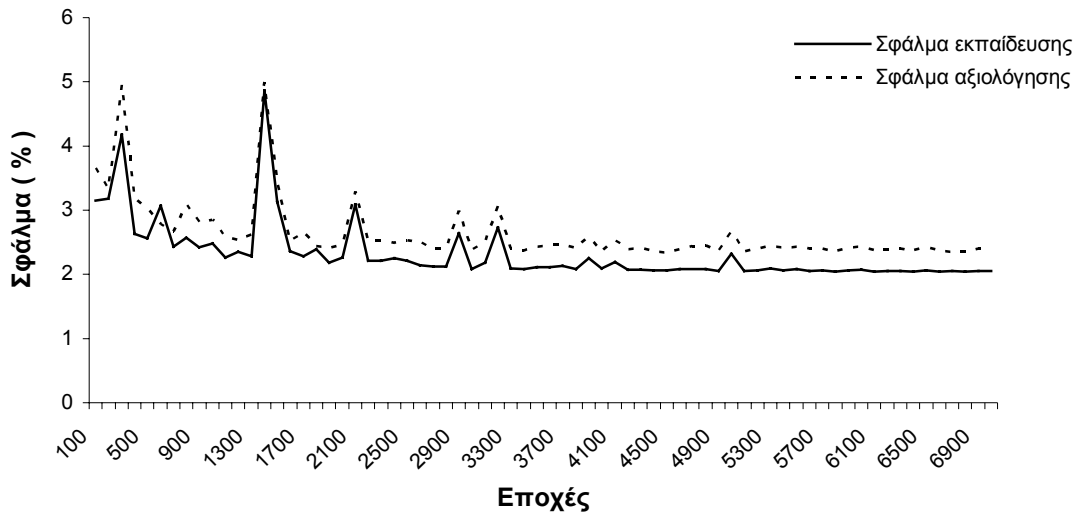


Γράφημα 4.10.1-2



Γράφημα 4.10.1-3

Σφάλμα εκπαίδευσης και αξιολόγησης για εφαρμογή του σεναρίου 1, με τυχαία παρουσίαση προτύπων και χαρακτηριστικά : train = 0,5 , train_T =1400, mom = 0,4 , mom_T =1000 , N = 48 & Y_TR = 1995 - 1999



Γράφημα 4.10.1-4

➤ 4.10.2 Παρατηρήσεις - Σχόλια για το σενάριο 1

Όπως παρατηρούμε από τον παραπάνω πίνακα και το γράφημα (4.10.1-3), βλέπουμε ότι το σενάριο 1 παρουσιάζει ελάχιστο σφάλματος εκπαίδευσης, αλλά και αξιολόγησης ταυτόχρονα για πλήθος νευρώνων κρυμμένου στρώματος ίσο με 48. Στην περιοχή [40,60], ξεκινώντας από το 48 και προχωρώντας προς τα δύο άκρα της, τα δύο σφάλματα εκπαίδευσης και αξιολόγησης αυξομειώνονται, χωρίς όμως να είναι ποτέ μικρότερα από τις τιμές τους για 48 νευρώνες και οι οποίες είναι 2,46 % και 4,84 % αντίστοιχα.

Το γεγονός ότι η διαφορά μεταξύ των δύο σφαλμάτων εκπαίδευσης και αξιολόγησης δεν είναι ιδιαίτερα μεγάλη (πάντα μικρότερη του 2,5 %), καταδεικνύει σε μας ότι η όλη διαδικασία της εκπαίδευσης κύλησε φυσιολογικά και ότι δεν δημιουργήθηκε κάποιο πρόβλημα με την κατασκευή των διανυσμάτων εκπαίδευσης και αξιολόγησης, με τον συγκεκριμένο τρόπο που επιθυμούσαμε. Συνεπώς το σενάριο εκτελέστηκε πετυχημένα όσον αφορά αυτό το σκέλος.

Για τη δεύτερη φάση, όπου ο τρόπος παρουσίασης των προτύπων ήταν τυχαίος, παρατηρούμε πολύ σημαντική βελτίωση καθώς τα σφάλματα αξιολόγησης και εκπαίδευσης καταλήγουν στις τιμές 2,39 % και 2,05 % αντίστοιχα. Η βελτίωση είναι προφανής, ειδικά για το σφάλμα αξιολόγησης, όπου σε σχέση με το σειριακό τρόπο παρουσίασης, έχουμε μείωση κατά 2,4 % περίπου. Το μόνο μειονέκτημα που μπορούμε να παρατηρήσουμε, είναι οι αυξομειώσεις στα σφάλματα εκπαίδευσης και αξιολόγησης, κατά το πέρασμα των εποχών.

Τέλος θα πρέπει σε αυτό το σημείο να αναφέρουμε ότι ο χρόνος για μια εκτέλεση του σεναρίου είναι περίπου 4 ώρες σε ηλεκτρονικό υπολογιστή τα χαρακτηριστικά του οποίου περιγράφονται αναλυτικά παρακάτω και στον οποίο εκτελέσαμε όλα τα σενάρια.

➤ 4.11 Σενάριο εκπαίδευσης 2

Το δεύτερο σενάριο που πραγματοποιείται είναι πανομοιότυπο με το πρώτο, με τη διαφορά όμως ότι στην είσοδο του διανύσματος εκπαίδευσης, ο επταψήφιος δυαδικός αριθμός που εκφράζει τη μέρα σε σχέση με τη βδομάδα, έχει αντικατασταθεί από δύο αριθμούς που έχουν το ίδιο νόημα, αλλά προκύπτουν από τις τριγωνομετρικές σχέσεις της μορφής $\cos (2\pi/7)$ & $\sin (2\pi/7)$.

Οπότε το διάνυσμα εισόδου τώρα θα είναι διαστάσεων 66×1 , ενώ το διάνυσμα εξόδου παραμένει φυσικά διαστάσεων 24×1 . Συνοπτικά περιγράφεται παρακάτω :

- ✓ 24 ωριαία φορτία της προηγούμενης μέρας $\Rightarrow$ Θέσεις 1-24
- ✓ 24 ωριαία φορτία της πιο προηγούμενης μέρας $\Rightarrow$ Θέσεις 25-48
- ✓ Μέγιστη και ελάχιστη θερμοκρασία της προηγούμενης μέρας για Αθήνα –σταθμός 1- και Θεσσαλονίκη –σταθμός 2 - $\Rightarrow$ Θέσεις 49-52
- ✓ Μέγιστη και ελάχιστη θερμοκρασία της τρέχουσας μέρας για Αθήνα και Θεσσαλονίκη $\Rightarrow$ Θέσεις 53-56
- ✓ Τετραγωνικές αποκλίσεις των μεγίστων θερμοκρασιών θεωρώντας, αυθαίρετα, ότι $T_{\min}=18$ & $T_{\max}=24$ για Αθήνα και Θεσσαλονίκη για τρέχουσα και προηγούμενη μέρα $\Rightarrow$ Θέσεις 57-60
- ✓ Διαφορές των μεγίστων θερμοκρασιών της προηγούμενης από την τρέχουσα μέρα για Αθήνα και Θεσσαλονίκη $\Rightarrow$ Θέσεις 61-62
- ✓ Αριθμοί που εκφράζουν και οι δυο την περιοδικότητα της ημέρας, ως προς την εβδομάδα και δίνονται από τις τριγωνομετρικές συναρτήσεις της μορφής $\cos (2\pi/7)$ & $\sin (2\pi/7)$ $\Rightarrow$ Θέσεις 63-64
- ✓ Αριθμοί που εκφράζουν την περιοδικότητα της ημέρας ως προς το έτος και δίνονται από τις τριγωνομετρικές συναρτήσεις της μορφής $\cos (2\pi/365)$ & $\sin (2\pi/365)$. $\Rightarrow$ Θέσεις 65-66

Η αντικατάσταση του δυαδικού αριθμού με τους δύο τριγωνομετρικούς αριθμούς, μπορεί σε κάποιον να μη φαντάζει ιδιαίτερα σημαντική, αλλά για τη γλώσσα προγραμματισμού FORTRAN είναι εντελώς διαφορετική η συμπεριφορά της, καθώς άλλο πράγμα μια ακολουθία ψηφίων 0 & 1 και άλλο δυο μικροί δεκαδικοί αριθμοί, των οποίων το πρόσημο εναλλάσσεται ανάλογα με τη μέρα που βρισκόμαστε. Λόγω αυτής της μικρής, ως προς τη δομή, διαφορά του σεναρίου 2 σε σχέση με το 1, δεν θεωρήθηκε ότι υπάρχει λόγος για παρουσίαση ενός διαγράμματος ροής της συγκεκριμένης διαδικασίας.

Κατά την εφαρμογή του σεναρίου 2 η παρουσίαση των προτύπων εκπαίδευσης ήταν *σειριακή*. Τα έτη εκπαίδευσης ήταν από το 1995-1999 ενώ το έτος αξιολόγησης ήταν το 2000. Οι παράμετροι για τις οποίες εκτελέστηκε το σενάριο ήταν οι εξής:

$r_momentum_arx = 0,6$, $r_momentum_T = 2000$, $trainrate_arx = 0,5$ και $trainrate_T = 1400$.

Όσον αφορά τον αριθμό των νευρώνων του κρυμμένου του στρώματος, αρχικά εκτελέσαμε το σενάριο για $iq1 = 48$ και στη συνέχεια μεταβάλλαμε τους νευρώνες με την εξής σειρά :

iq1 = 40-45-50-55-60.

Τα αποτελέσματα καθώς και τα αντίστοιχα σχόλια βρίσκονται στις επόμενες παραγράφους.

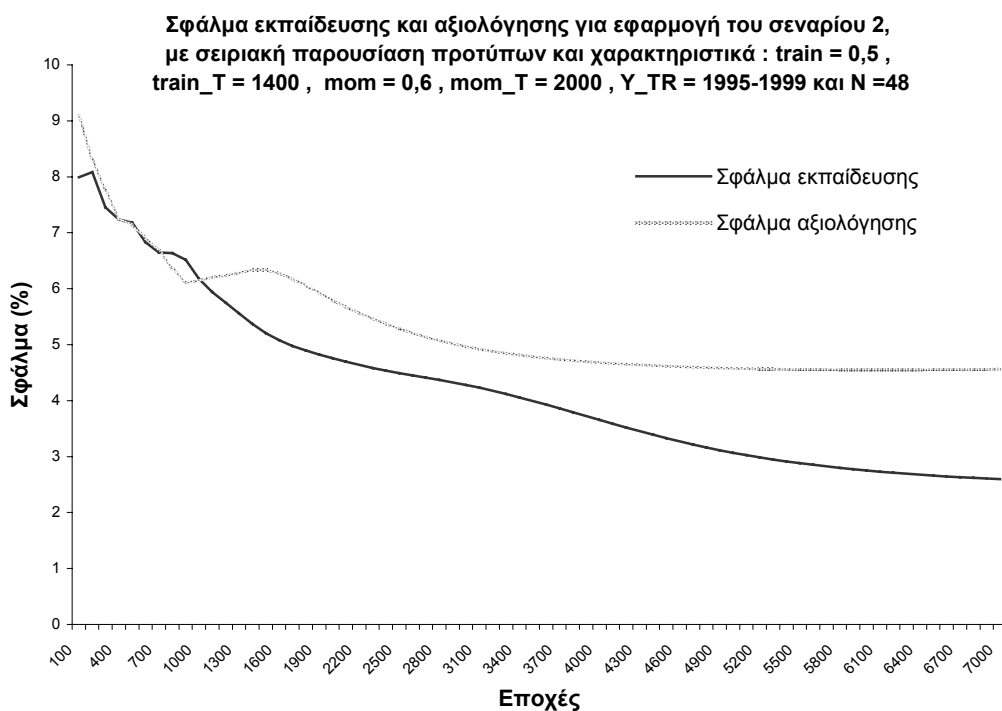
➤ 4.11.1 Αποτελέσματα σεναρίου 2 - Γραφήματα

Ακολουθεί πίνακας συγκεντρωτικών αποτελεσμάτων :

Πλήθος νευρώνων	40	45	48	50	55	60
Σφάλμα εκπαίδευσης (%)	2.62	2.79	2.60	2.70	2.83	3.00
Σφάλμα αξιολόγησης (%)	4.63	4.53	4.56	4.58	4.78	4.81

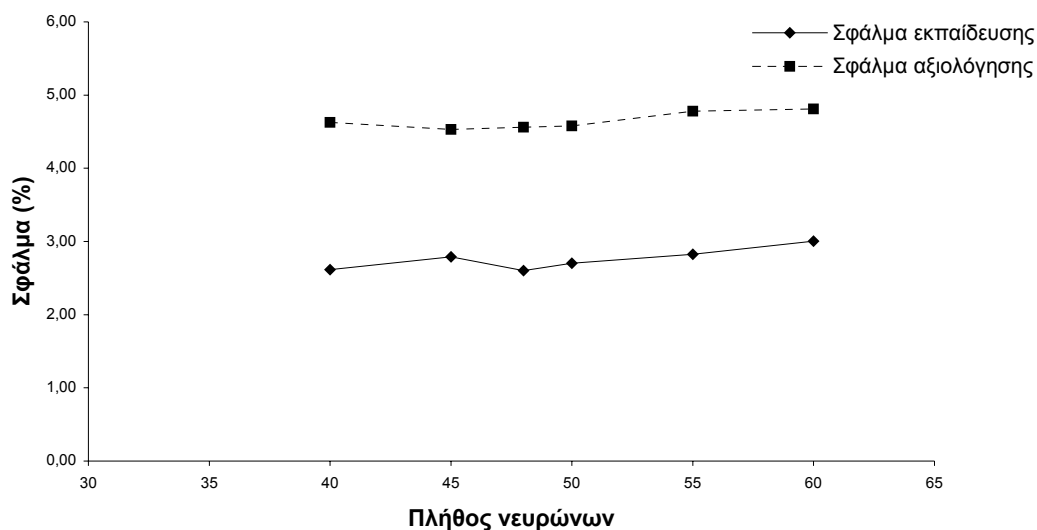
**Πίνακας 4.11.1-1 : Συγκεντρωτικά αποτελέσματα για εφαρμογή του σεναρίου 2, με
σειριακό τρόπο παρουσίασης προτύπων, για $\text{train} = 0,5$, $\text{train_T} = 1400$,
 $\text{mom} = 0,6$, $\text{mom_T} = 2000$ και $Y_TR = 1995-1999$**

Ακολουθούν τα αντίστοιχα γραφήματα :



Γράφημα 4.11.1-2

Σφάλμα εκπαίδευσης και αξιολόγησης για εφαρμογή του σεναρίου 2
με σειριακή παρουσίαση προτύπων και χαρακτηριστικά : train = 0,5 ,
train_T = 1400 , mom = 0,6 , mom_T = 2000 & Y_TR = 1995-1999



Γράφημα 4.11.1-3

➤ 4.11.2 Παρατηρήσεις - Σχόλια για το σενάριο 2

Όσον αφορά το σενάριο 2 βλέπουμε ότι για το διάστημα [45,50] το σφάλμα αξιολόγησης μεταβάλλεται πολύ λίγο και μάλιστα, για πλήθος νευρώνων ίσο με 48 παρατηρούνται οι βέλτιστες, συνολικά, τιμές για τα σφάλματα εκπαίδευσης και αξιολόγησης που είναι 2,6 % και 4,56 % αντίστοιχα. Και εδώ η διαφορά μεταξύ των δύο σφαλμάτων δεν είναι ιδιαίτερα μεγάλη και συνεπώς μπορούμε να θεωρήσουμε ότι και αυτό το σενάριο κύλησε ομαλά.

Επίσης σε σύγκριση με το σενάριο 1 (για σειριακό τρόπο παρουσίασης προτύπων φυσικά) υπάρχει βελτίωση και αυτό που μπορούμε να σημειώσουμε, είναι ότι η αντικατάσταση του επταψήφιου δυαδικού αριθμού, που εκφράζει την κυκλικότητα ως προς τη βδομάδα, με δύο αντίστοιχους τριγωνομετρικούς αριθμούς μειώνει το σφάλμα αξιολόγησης περίπου κατά 0,3 %. Βέβαια εδώ υπάρχουν σαφέστατα περιθώρια βελτίωσης, τα οποία και θα υποδειχθούν στο επόμενο κεφάλαιο.

Τέλος θα πρέπει σε αυτό το σημείο να αναφέρουμε ότι ο χρόνος για μια εκτέλεση του σεναρίου είναι περίπου 4 ώρες.

➤ 4.12 Σενάριο εκπαίδευσης 3

Το τρίτο σενάριο διαφέρει σημαντικά από τα προηγούμενα, ως προς τη δομή του και θέτει ως εισόδους τα εξής στοιχεία :

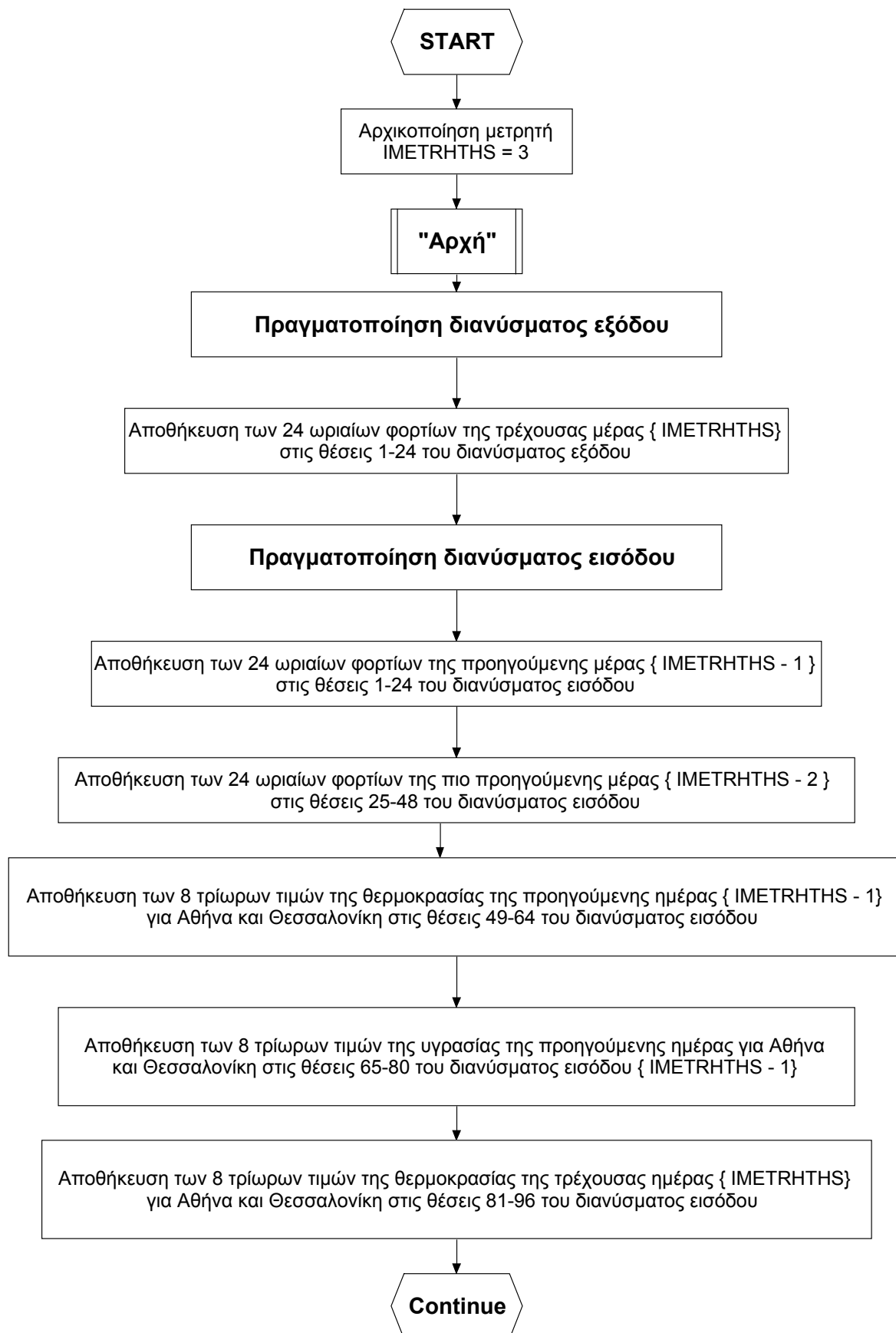
- ✓ 24 ωριαία φορτία της προηγούμενης μέρας $\Rightarrow$ Θέσεις 1-24
- ✓ 24 ωριαία φορτία της πιο προηγούμενης μέρας $\Rightarrow$ Θέσεις 25-48
- ✓ 8 τρίωρες θερμοκρασίες της προηγούμενης μέρας για Αθήνα –σταθμός 1 - και Θεσσαλονίκη – σταθμός 2 - $\Rightarrow$ Θέσεις 49-64
- ✓ 8 τρίωρες υγρασίες της προηγούμενης μέρας για Αθήνα και Θεσσαλονίκη $\Rightarrow$ Θέσεις 65-80
- ✓ 8 τρίωρες θερμοκρασίες της τρέχουσας μέρας για Αθήνα και Θεσσαλονίκη $\Rightarrow$ Θέσεις 81-96
- ✓ Αριθμοί που εκφράζουν την περιοδικότητα της ημέρας ως προς την εβδομάδα και δίνονται από τις τριγωνομετρικές συναρτήσεις της μορφής $\cos (2\pi/7)$ & $\sin (2\pi/7)$. $\Rightarrow$ Θέσεις 97-98
- ✓ Αριθμοί που εκφράζουν την περιοδικότητα της ημέρας ως προς το έτος και δίνονται από τις τριγωνομετρικές συναρτήσεις της μορφής $\cos (2\pi/365)$ & $\sin (2\pi/365)$. $\Rightarrow$ Θέσεις 99-100

Όπως διαπιστώνουμε, σε αυτό το σενάριο η είσοδος έχει διαστάσεις 100 X 1, που είναι οι μεγαλύτερες διαστάσεις για τα σενάρια μας, ενώ η έξοδος έχει διαστάσεις 24 X 1, καθώς περιέχει τα 24 ωριαία φορτία της τρέχουσας μέρας. Είναι ένα σενάριο αρκετά απλό ως προς τη δημιουργία του, που όμως ο μεγάλος αριθμός των στοιχείων εισόδου του νευρωνικού δικτύου δημιουργεί κάποια ερωτηματικά ως προς τα τελικά αποτελέσματα.

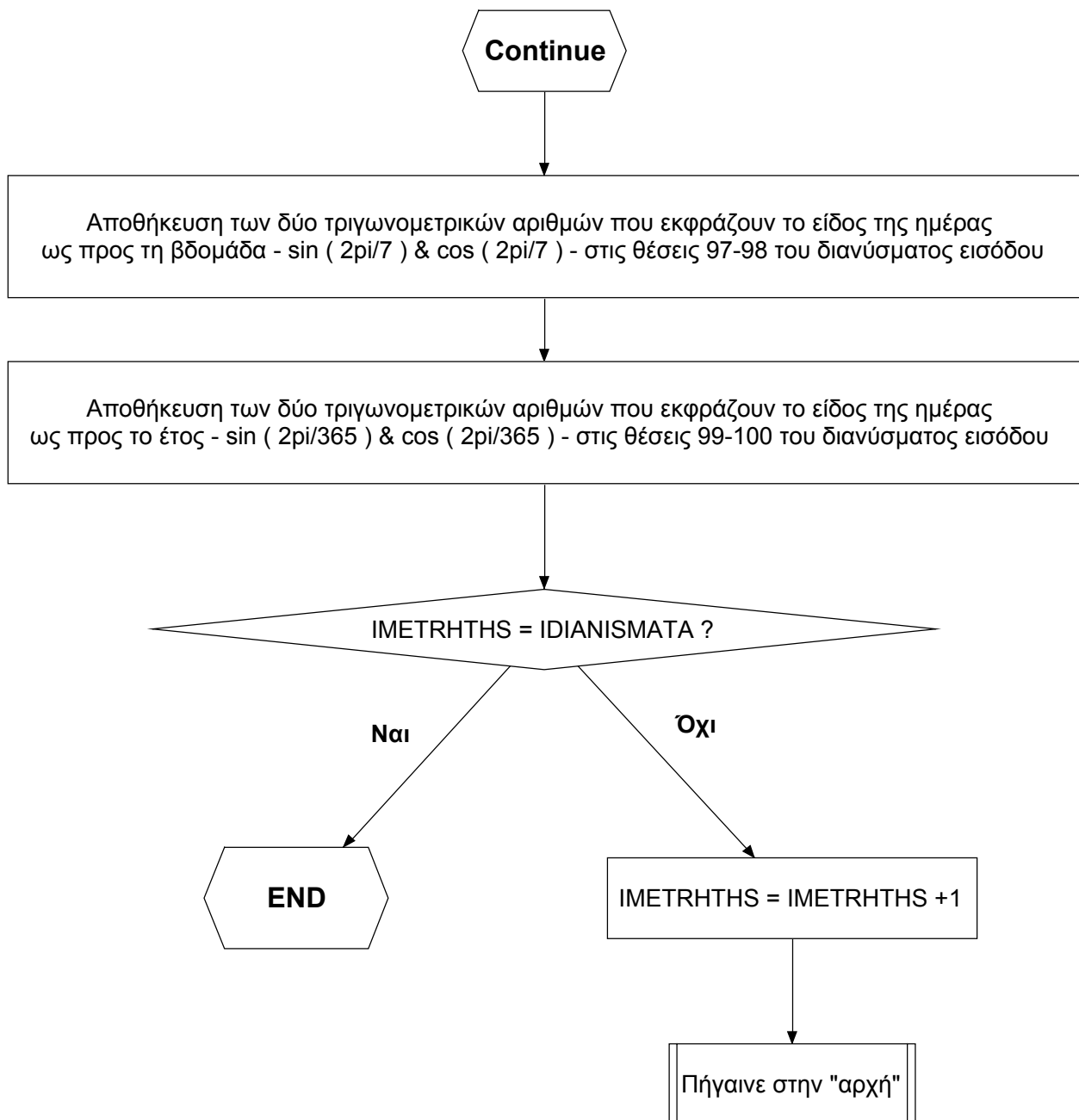
Κατά την εφαρμογή του σεναρίου 3 η παρουσίαση των προτύπων εκπαίδευσης ήταν *σειριακή*. Τα έτη εκπαίδευσης ήταν το χρονικό διάστημα από το 1995-1999, ενώ το έτος αξιολόγησης ήταν το 2000. Οι παράμετροι για τις οποίες εκτελέστηκε το σενάριο ήταν οι εξής:
 $r_momentum_arx = 0,6$, $r_momentum_T = 2000$, $trainrate_arx = 0,5$ και $trainrate_T = 1400$.

Όσον αφορά τον αριθμό των νευρώνων του κρυμμένου του στρώματος, αρχικά εκτελέσαμε το σενάριο για $iq1 = 48$ και στη συνέχεια μεταβάλλαμε τους νευρώνες με την εξής σειρά :
 $iq1 = 50-55-60-65-70$.

Τα αποτελέσματα καθώς και τα αντίστοιχα σχόλια βρίσκονται στις επόμενες παραγράφους.
 Το διάγραμμα ροής αυτού του σεναρίου παρατίθεται στην επόμενη σελίδα.



Πραγματοποίηση διανυσμάτων εκπαίδευσης -Σενάριο 3



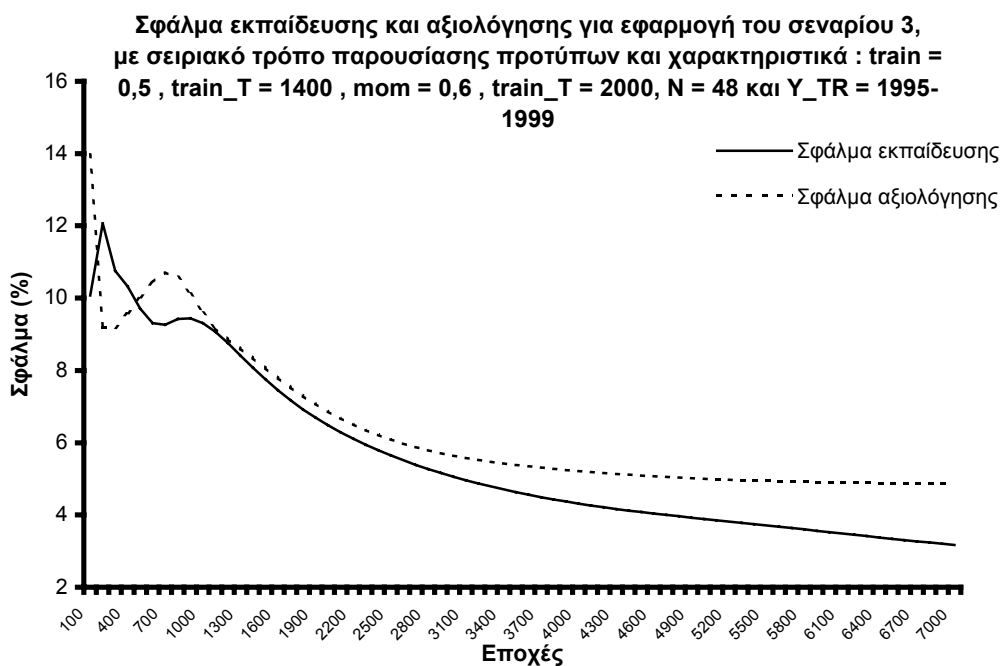
Πραγματοποίηση διανυσμάτων εκπαίδευσης- Σενάριο 3 (συνέχεια)

➤ 4.12.1 Αποτελέσματα σεναρίου 3 – Γραφήματα

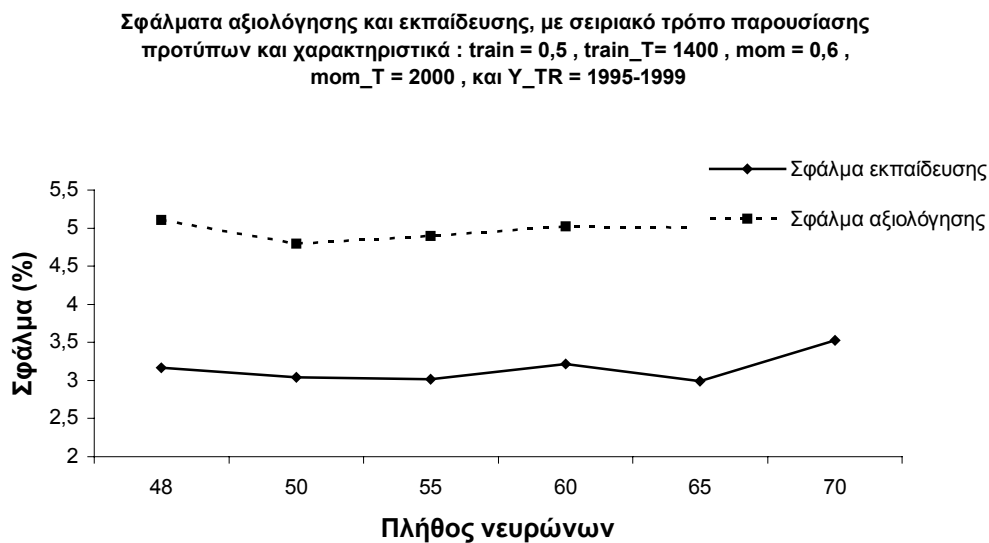
Παρακάτω ακολουθεί πίνακας συγκεντρωτικών αποτελεσμάτων :

Πλήθος νευρώνων	48	50	55	60	65	70
Σφάλμα εκπαίδευσης (%)	3,17	3,04	3,02	3,22	2,99	3,53
Σφάλμα αξιολόγησης (%)	5,11	4,80	4,89	5,02	5,00	5,11

Πίνακας 4.12.1-1 : Συγκεντρωτικά στοιχεία για εφαρμογή του σεναρίου 3 με σειριακό τρόπο παρουσίασης προτύπων, για $\text{train} = 0,5$, $\text{train_T} = 1400$, $\text{mom} = 0,6$, $\text{mom_T} = 2000$ και $Y_TR = 1995-1999$



Γράφημα 4.12.1-2



Γράφημα 4.12.1-3

➤ 4.12.2 Παρατηρήσεις - Σχόλια για το σενάριο 3

Όσον αφορά το σενάριο 3, βλέπουμε ότι η ελάχιστη τιμή σφαλμάτων εμφανίζεται για πλήθος νευρώνων ίσο με 50 και είναι 4,80 % και 3,04 %, για τα σφάλματα εκπαίδευσης και αξιολόγησης αντίστοιχα. Μάλιστα από εκεί και πέρα όσο αυξάνουμε το πλήθος των νευρώνων τα σφάλματα έχουν μια αυξητική τάση.

Επειδή οι διαφορές των δύο σφαλμάτων εκπαίδευσης και αξιολόγησης είναι πάντα μικρές, οδηγούμαστε στο συμπέρασμα ότι το συγκεκριμένο σενάριο εκτελέστηκε επιτυχώς, ως προς την κατασκευή των διανυσμάτων εκπαίδευσης και αξιολόγησης. Βέβαια όσον αφορά το αποτέλεσμα είναι απαραίτητες μερικές τροποποιήσεις, ώστε να έχουμε βελτίωση.

Ο χρόνος εκτέλεσης του σεναρίου ήταν περίπου 4,5 ώρες και είναι μεγαλύτερος από τα δύο προηγούμενα σενάρια, γιατί τα διανύσματα εκπαίδευσης περιλαμβάνουν αυτήν τη φορά 100 στοιχεία το καθένα.

Αμέσως μετά θα αναφερθούμε γενικά στην έννοια της κατηγοριοποίησης των ημερών, αλλά και συγκεκριμένα στην κατηγοριοποίηση ως προς το φορτίο.

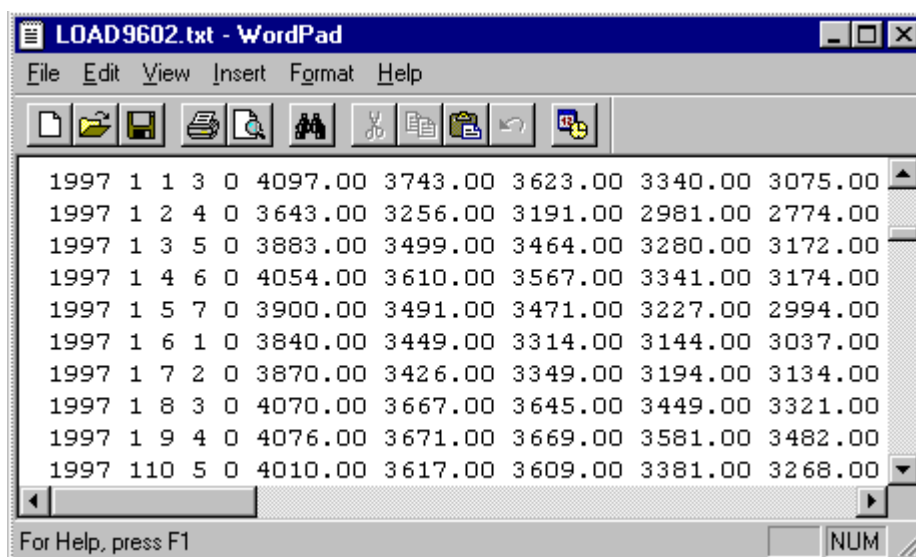
➤ 4.13 Κατηγοριοποίηση των ημερών

Ένα βασικό πρόβλημα, κατά την πρόβλεψη φορτίου και της συνολικά ζητούμενης ηλεκτρικής ενέργειας σ' ένα σύστημα, είναι η διαφορετική συμπεριφορά που παρουσιάζουν οι καταναλωτές κάθε μέρα, ανάλογα με τις δραστηριότητές τους. Αυτή η συμπεριφορά επηρεάζεται, κατά κύριο λόγο, από το είδος της μέρας, τη θερμοκρασία και τον τουρισμό, στην περίπτωση καλοκαιρινών μηνών για τις τουριστικές περιοχές. Γι' αυτόν το λόγο, έχουν προταθεί διάφορες μέθοδοι κατηγοριοποίησης των χρονολογικών ημερησίων καμπυλών του φορτίου, ώστε να βρεθεί ο διαφορετικός αριθμός τύπων «ημερών», τόσο ως προς το φορτίο, όσο ως προς τη θερμοκρασία.

Για τα επόμενα σενάρια χρησιμοποιήθηκε η κατηγοριοποίηση των ημερών, ως προς το φορτίο. Με την έννοια κατηγοριοποίηση των ημερών ως προς το φορτίο, βάσει των προαναφερθέντων, εννοούμε τη διαίρεση ενός συνόλου ημερών σε συγκεκριμένο αριθμό ομάδων, όπου η κάθε μία από αυτές, εκφράζει μία δεδομένη τάση ως προς το φορτίο. Επειδή όμως κάθε μέρα αντιστοιχεί σε μας, σε ένα διάνυσμα εκπαίδευσης, ουσιαστικά μέσω της κατηγοριοποίησης αυτής, χωρίζουμε και το σύνολο των διανυσμάτων εκπαίδευσης στις ίδιες ομάδες. Η κατηγοριοποίηση εκφράζεται πρακτικά στα αρχεία των φορτίων, με το πέμπτο κατά σειρά στοιχείο της κάθε γραμμής – IFLDAY. Για τη δική μας εφαρμογή χρησιμοποιήθηκε η μέθοδος των K-μέσων, η οποία χρησιμοποιεί νευρωνικά δίκτυα και ο αριθμός των ομάδων κατανομής είναι, ενδεικτικά, 15.

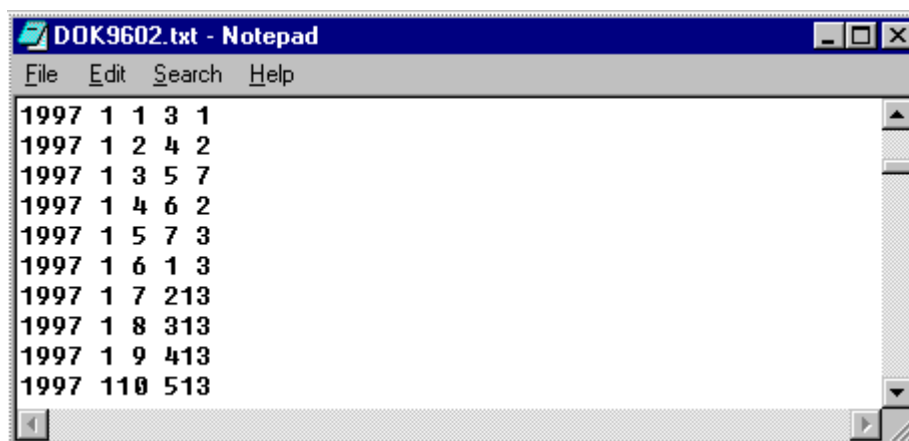
Βέβαια τα δικά μας αρχεία, στην αρχική τους μορφή είχαν το πέμπτο στοιχείο της κάθε γραμμής 0, που σημαίνει ότι δεν χρησιμοποιούσαμε μέχρι τώρα πουθενά, την κατηγοριοποίηση των ημερών.

Η μορφή αυτών των αρχείων των μη κατηγοριοποιημένων ημερών φαίνεται αμέσως παρακάτω :



Εικόνα 4.13-1 : Παράδειγμα αρχείου μη κατηγοριοποιημένων ημερών

Από την άλλη η έξοδος του προγράμματος κατηγοριοποίησης της K-μέσων, είναι ένα αρχείο όπου χρησιμοποιεί την κατηγοριοποίηση, καθώς το πέμπτο στοιχείο –IFLDAY- δεν είναι μηδενικό, αλλά αντίθετα παίρνει τιμές από 1-15 ανάλογα με την ομάδα που ανήκει η συγκεκριμένη μέρα. Όμως δεν περιέχει τα 24 ωριαία φορτία της κάθε ημέρας. Ένα παράδειγμα της μορφής αυτού του αρχείου παρουσιάζεται παρακάτω:



Εικόνα 4.13-2 : Παράδειγμα αρχείου εξόδου της μεθόδου K-μέσων για 15 ομάδες

Συνεπώς χρειάστηκε να γράψουμε ένα επιπλέον πρόγραμμα, το οποίο δέχεται ως εισόδους το αρχείο της κατηγοριοποίησης (K-μέσων) και το αρχείο των μη κατηγοριοποιημένων ημερών. Η έξοδος αυτού του προγράμματος είναι ουσιαστικά το αρχείο των φορτίων, με το μηδενικό πέμπτο στοιχείο –IFLDAY- να έχει αντικατασταθεί, σύμφωνα με την αντίστοιχη τιμή του στο αρχείο κατηγοριοποίησης. Με αυτόν τον τρόπο θα λέγαμε, ότι ενεργοποιούμε την κατηγοριοποίηση των ημερών στο αρχείο των φορτίων. Το παραπάνω πρόγραμμα ονομάστηκε «Ενεργοποίηση κατηγοριοποίησης». Ένα παράδειγμα αρχείου εξόδου του προγράμματος είναι και το παρακάτω :

1997	1	1	3	1	4097.00	3743.00	3623.00	3340.00	3075.00
1997	1	2	4	2	3643.00	3256.00	3191.00	2981.00	2774.00
1997	1	3	5	7	3883.00	3499.00	3464.00	3280.00	3172.00
1997	1	4	6	2	4054.00	3610.00	3567.00	3341.00	3174.00
1997	1	5	7	3	3900.00	3491.00	3471.00	3227.00	2994.00
1997	1	6	1	3	3840.00	3449.00	3314.00	3144.00	3037.00
1997	1	7	213		3870.00	3426.00	3349.00	3194.00	3134.00
1997	1	8	313		4070.00	3667.00	3645.00	3449.00	3321.00
1997	1	9	413		4076.00	3671.00	3669.00	3581.00	3482.00
1997	110	513			4010.00	3617.00	3609.00	3381.00	3268.00

Εικόνα 4.13-3 : Παράδειγμα αρχείου εξόδου του προγράμματος ενεργοποίησης της κατηγοριοποίησης

➤ 4.13.1 Περιγραφή του προγράμματος ενεργοποίησης της κατηγοριοποίησης

Για μια περιγραφή του προγράμματος, αρχικά θα λέγαμε ότι οι κυριότερες παράμετροι που αυτό χρησιμοποιεί είναι οι εξής :

ip_dim1 (=5) ⇒ αφορά τα στοιχεία των οδηγών κάθε γραμμής και για τα δύο αρχεία εισόδου

ip_dim2 (=24) ⇒ αφορά τα 24 ωριαία φορτία της κάθε ημέρας

ip_dim_max (=3000) ⇒ αφορά το μέγιστο αριθμό των γραμμών-ημερών, που μπορεί το πρόγραμμα μας να διαβάσει και να επεξεργαστεί.

Τα FORMAT των γραμμών που χρησιμοποιούμε είναι δύο :

5 FORMAT (I4, I2, I2, I2,I2, 24(F8.2))

που αφορά το αρχείο εισόδου των μη κατηγοριοποιημένων ημερών

10 FORMAT (I4, I2, I2, I2,I2)

που αφορά το αρχείο της κατηγοριοποίησης της K-μέσων.

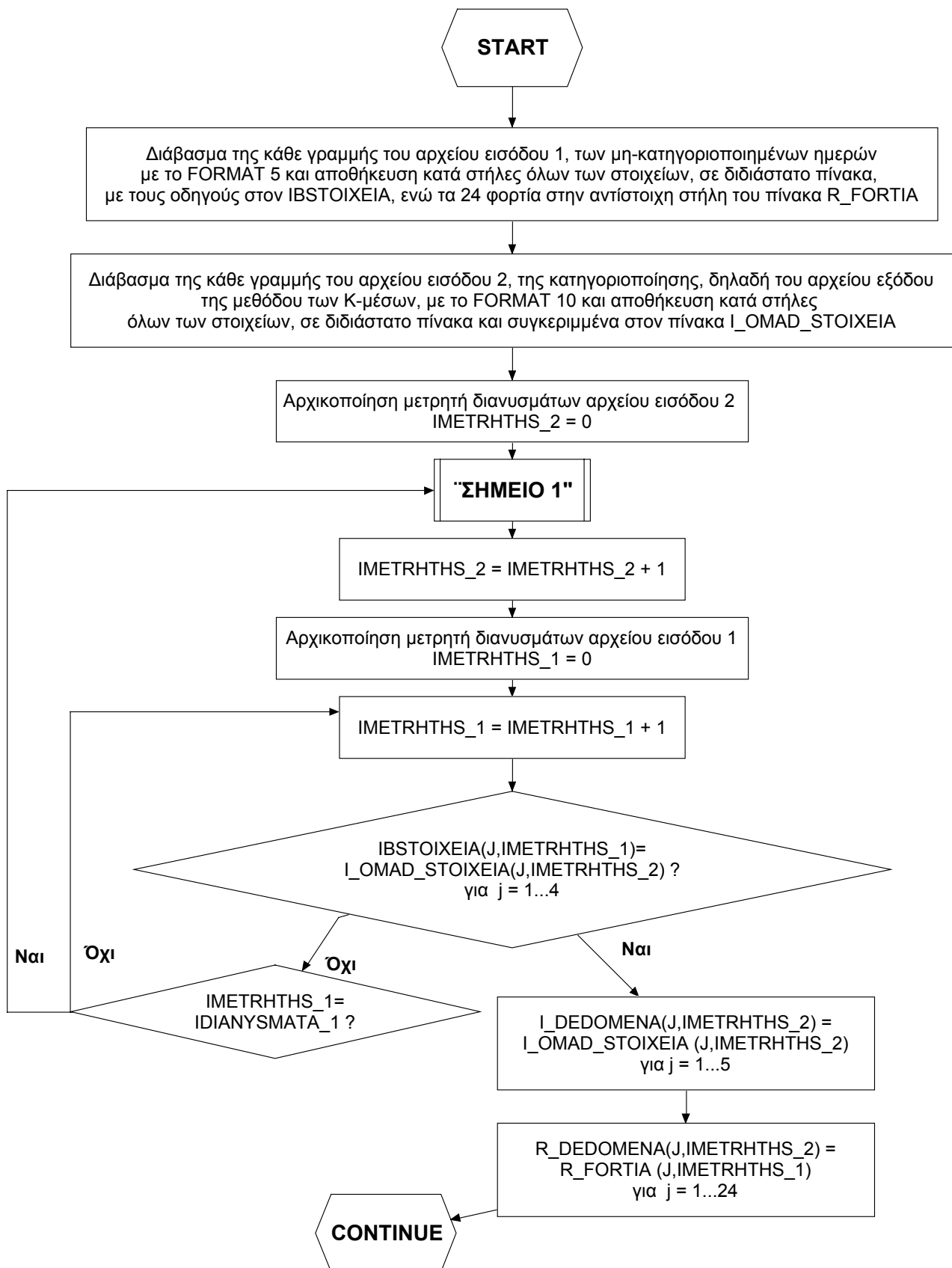
Αρχικά το πρόγραμμα ζητά από τον χρήστη την εισαγωγή του ονόματος του αρχείου εισόδου 1, δηλαδή του αρχείου των μη κατηγοριοποιημένων ημερών, το ανοίγει και διαβάζει με το FORMAT 5 την κάθε γραμμή, αποθηκεύοντας όλα τα στοιχεία της πρώτα στις βοηθητικές μεταβλητές IYEAR, IMONTH, IDAY, IKDAY, IFLDAY, RLOAD (πίνακας μονοδιάστατος 24 στοιχείων) και κατόπιν, τα 5 στοιχεία οδηγούς κατά στήλες σε μορφή διανύσματος στον πίνακα IBSTOIXEIA (ip_dim1, ip_dim_max) και τα 24 φορτία στην αντίστοιχη στήλη του πίνακα R_FORTIA (ip_dim2,ip_dim_max). Ακολούθως, αφού κλείσει το αρχείο εισόδου 1, ζητά το όνομα του αρχείου εισόδου 2 (αρχείο κατηγοριοποίησης της K-μέσων), το ανοίγει και μετά διαβάζει κάθε γραμμή με το FORMAT 10, ενώ αποθηκεύει τα 5 πρώτα στοιχεία, αρχικά στις βοηθητικές μεταβλητές IYEAR, IMONTH, IDAY, IKDAY, IFLDAY και κατόπιν σε μορφή διανύσματος στον πίνακα I_OMAD_STOIXEIA (ip_dim1,ip_dim_max). Μετά κλείνει και το αρχείο εισόδου 2 και ανοίγει το μοναδικό αρχείο εξόδου “LOAD_KATHG” .

Στην επόμενη φάση ξεκινάει η κυρίως διαδικασία όπου για κάθε στήλη του πίνακα I_OMAD_STOIXEIA, προσπαθούμε να βρούμε κάποια στήλη του πίνακα IBSTOIXEIA, που να έχει όλα τα αντίστοιχα στοιχεία της ίδια. Ουσιαστικά ψάχνουμε για κάθε μέρα από το αρχείο της κατηγοριοποίησης (input file2), σε ποιο ακριβώς σημείο βρίσκεται στο αρχείο των μη κατηγοριοποιημένων ημερών (input file1). Αυτό επιτυγχάνεται με τη χρήση δύο μετρητών που αρχικοποιούνται και οι δύο με την τιμή 1, ο πρώτος –IMETRHTHS- είναι ο μετρητής των στηλών του IBSTOIXEIA, ενώ ο δεύτερος –IMETRHTHS_2- είναι ο μετρητής των στηλών του πίνακα I_OMAD_STOIXEIA. Για κάθε δεδομένη τιμή του IMETRHTHS, συγκρίνονται διαδοχικά οι αντίστοιχες στήλες των δύο πινάκων καθώς ο IMETRHTHS_2 παίρνει όλες τις τιμές από 1-IDIANYSMATA (το πλήθος των διανυσμάτων για το αρχείο εισόδου 1), ώσπου να βρεθεί μια ίδια.

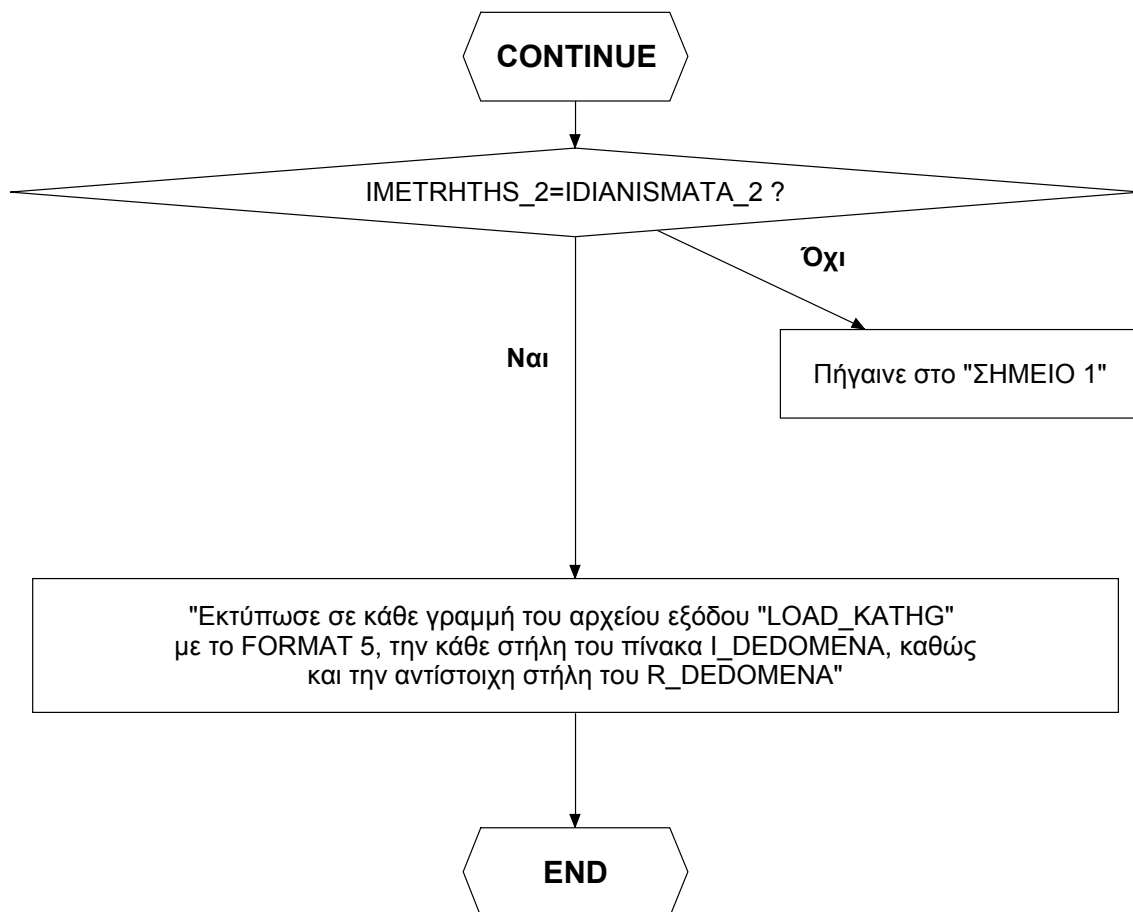
Ακολούθως αποθηκεύουμε τη στήλη του I_OMAD_STOIXEIA που αντιστοιχεί στην τρέχουσα τιμή του IMETRHTHS_2, στον πίνακα I_DEDOMENA(ip_dim1,ip_dim_max), καθώς και την στήλη του IBSTOIXEIA που αντιστοιχεί στην τρέχουσα τιμή του IMETRHTHS στον πίνακα R_DEDOMENA (ip_dim2,ip_dim_max). Με αυτόν τον τρόπο ουσιαστικά κάθε στήλη του I_DEDOMENA και η αντίστοιχη του R_DEDOMENA περιγράφει πλήρως μια ημέρα με όλους τους οδηγούς της, συμπεριλαμβανομένης και της κατηγοριοποίησης, και τα 24 ωριαία φορτία της.

Τέλος στο αρχείο εξόδου εκτυπώνουμε, σε κάθε γραμμή τις αντίστοιχες στήλες των δύο τελευταίων πινάκων, λαμβάνοντας τελικά ένα αρχείο όπου περιγράφει πλήρως μια ημέρα, έχοντας υπόψη του και την κατηγοριοποίηση, που είναι άλλωστε το επιθυμητό αποτέλεσμα. Βέβαια θα πρέπει να τονιστεί ότι το πρόγραμμα ενεργοποιεί φυσικά το IFLDAY, για τις μέρες που βρίσκονται ταυτόχρονα και στα δύο αρχεία εισόδου. Αν δηλαδή μια μέρα περιέχεται μονάχα στο αρχείο εισόδου 1 και όχι στο 2, τότε δεν θα καταγραφεί στο αρχείο εξόδου, κάτι απόλυτα λογικό, αφού η συγκεκριμένη μέρα δεν έχει κατηγοριοποιηθεί.

Για την ολοκλήρωση της παρουσίασης του προγράμματος παρακάτω παραθέτουμε το διάγραμμα ροής του.



Ενεργοποίηση της κατηγοριοποίησης των ημερών στα αρχεία των φορτίων



Ενεργοποίηση της κατηγοριοποίησης των ημερών στα αρχεία των φορτίων (συνέχει

➤ 4.14 Σενάριο εκπαίδευσης 4

Έχοντας λοιπόν πλέον κατηγοριοποιήσει τα φορτία, προχωράμε πλέον στην εφαρμογή του τέταρτου σεναρίου, το οποίο και έχει την εξής δομή, όσον αφορά την είσοδο των διανυσμάτων :

- ✓ 24 ωριαία φορτία της προηγούμενης μέρας $\Rightarrow$ Θέσεις 1-24
- ✓ 24 ωριαία φορτία της προηγούμενης ημερολογιακά μέρας, που ανήκει όμως ταυτόχρονα στην ίδια ομάδα με την τρέχουσα $\Rightarrow$ Θέσεις 25-48
- ✓ 8 τρίωρες θερμοκρασίες της προηγούμενης ημερολογιακά μέρας, η οποία ανήκει στην ίδια ομάδα με την τρέχουσα, για Αθήνα –σταθμός 1 - και Θεσσαλονίκη –σταθμός 2 - $\Rightarrow$ Θέσεις 49-64
- ✓ 8 τρίωρες θερμοκρασίες της προηγούμενης ημερολογιακά μέρας, για Αθήνα και Θεσσαλονίκη $\Rightarrow$ Θέσεις 65-80
- ✓ 8 τρίωρες θερμοκρασίες της τρέχουσας μέρας, για Αθήνα και Θεσσαλονίκη $\Rightarrow$ Θέσεις 81-96
- ✓ Αριθμοί που εκφράζουν και οι δυο την περιοδικότητα της ημέρας ως προς την εβδομάδα και δίνονται από τις τριγωνομετρικές συναρτήσεις της μορφής $\cos (2\pi/7)$ & $\sin (2\pi/7)$. $\Rightarrow$ Θέσεις 97-98

- ✓ Αριθμοί που εκφράζουν την περιοδικότητα της ημέρας ως προς το έτος και δίνονται από τις τριγωνομετρικές συναρτήσεις της μορφής $\cos (2\pi/365)$ & $\sin (2\pi/365)$. $\Rightarrow$ Θέσεις 99-100

Σε αυτήν την περίπτωση το διάνυσμα εισόδου έχει διαστάσεις 100×1 , ενώ η έξοδος περιέχει πάλι τα 24 ωριαία φορτία της τρέχουσας μέρας, άρα είναι 24×1 . Αναφορικά με αυτό το σενάριο, υποθέσαμε ότι, για να προβλέψουμε το φορτίο μιας μέρας, θα πρέπει να λάβουμε υπόψη το γεγονός, ότι κάθε μέρα δεν έχει την ίδια συμπεριφορά με την προηγούμενη της και γι' αυτό ζητάμε επιπλέον πληροφορίες, όχι μόνο για την προηγούμενη μέρα, αλλά και για την τελευταία μέρα που έχει περάσει και ανήκει στην ίδια ομάδα με την ίδια, όσο και αν έπεται χρονικά .

Όσον αφορά τον τρόπο κατασκευής του σεναρίου θα πρέπει να αναφέρουμε ότι, για να μην υπάρχει κανένα πρόβλημα κατά την υλοποίησή του, χρησιμοποιήσαμε έναν μονοδιάστατο πίνακα, $I_FORES (I_OMADES)$, όπου $I_OMADES=15$, ο οποίος έχει αρχικά όλα τα στοιχεία του μηδενικά. Κάθε στοιχείο του στη συνέχεια, παίρνει την τιμή της μονάδας – και την κρατάει μέχρι τέλους -, εφόσον περάσει κάποιο διάνυσμα που να ανήκει στην ομάδα, η οποία αντιστοιχεί στη θέση αυτή. Για παράδειγμα αν περάσει μια ημέρα της ομάδας 5, τότε $I_FORES (5) = 1$. Όμως, πριν από την εντολή αυτή υπάρχει ένας έλεγχος, σύμφωνα με τον οποίο το πρόγραμμα δέχεται κάποια ημέρα-διάνυσμα, μόνο αν το στοιχείο του πίνακα που αντιστοιχεί στην ομάδα της, είναι μη μηδενικό. Με αυτόν τον τρόπο κατορθώνουμε να απορρίψουμε το πρώτο στοιχείο της κάθε ομάδας που εμφανίζεται και να δεχόμαστε στο σενάριο, φυσικά, όλα τα επόμενα στοιχεία. Συνεπώς αποφεύγουμε το πρόβλημα που θα υπήρχε σε αντίθετη περίπτωση, καθώς ο αλγόριθμος προστάζει ότι χρειαζόμαστε για την τρέχουσα ημέρα και την προηγούμενη της ημερολογιακά, η οποία να ανήκει όμως στην ίδια ομάδα.

Επίσης σε αυτό το σημείο θα πρέπει να τονίσουμε, ότι για να βρεθεί αυτή η μέρα, χρησιμοποιούμε έναν μετρητή $I_METRHTHS_2$, ο οποίος ξεκινάει από την τρέχουσα τιμή του $I_METRHTHS$ και μειώνεται σταδιακά κατά ένα, ώσπου να επιτευχθεί η ισότητα : $IBSTOIXEIA(5,IMETRHTHS_2) = IBSTOIXEIA(5,IMETRHTHS)$ και να βρούμε δηλαδή την προηγούμενη μέρα –τρέχουσα τιμή του $IMETRHTHS_2$ -, που να ανήκει στην ίδια ομάδα.

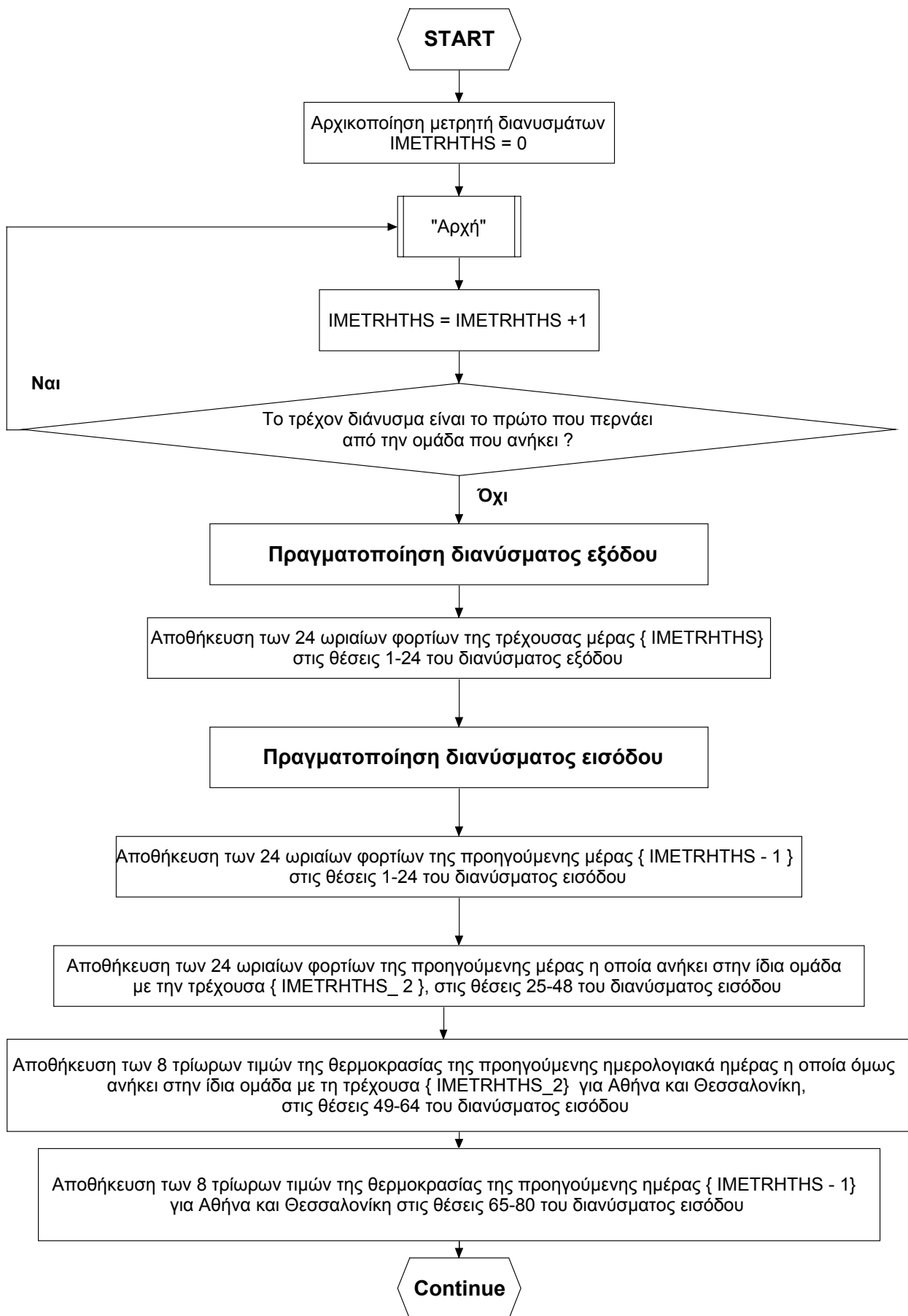
Κατά την εφαρμογή του σεναρίου 4 η παρουσίαση των προτύπων ήταν *σειριακή*, ενώ οι παράμετροι για τις οποίες εκτελέστηκε το σενάριο ήταν οι εξής:

$r_momentum_arx = 0,6$, $r_momentum_T = 2000$, $trainrate_arx = 0,5$ και $trainrate_T = 1400$.

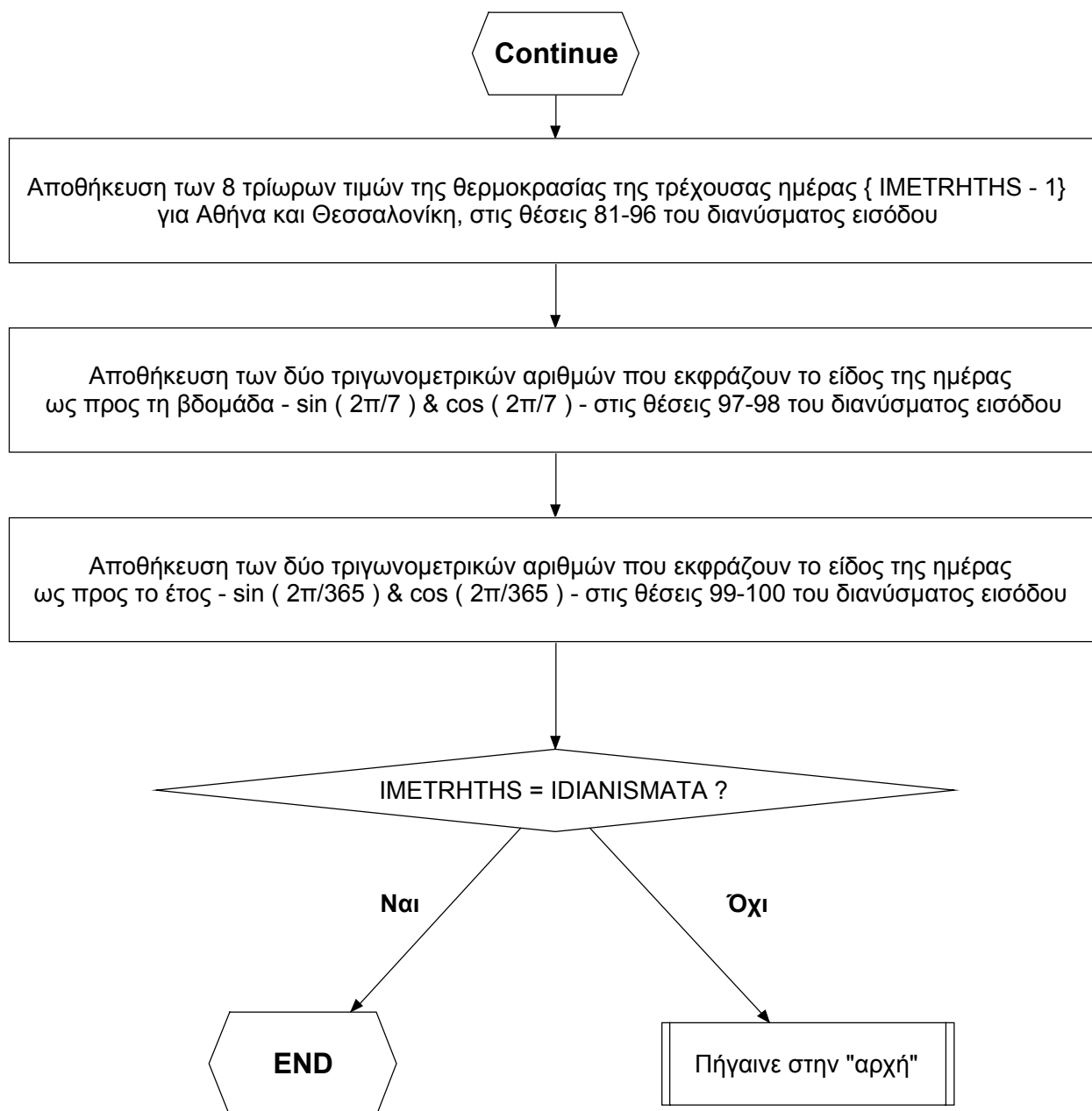
Αρχικά λοιπόν εκτελέσαμε το σενάριο για $iq1 = 48$ με τα έτη εκπαίδευσης να είναι από το 1995-1999 τη μία φορά, ενώ την επόμενη να είναι από το 1996-1999 και με έτος αξιολόγησης πάντα το 2000. Στη συνέχεια μεταβάλλαμε τους νευρώνες με την εξής σειρά :

$iq1 = 50-55-60-65-70-75-80-85-90$ για έτη εκπαίδευσης από 1995-1999.

Τα αποτελέσματα και μερικές εκ των γραφικών παραστάσεων που προέκυψαν (οι υπόλοιπες βρίσκονται στο παράρτημα I), καθώς και τα αντίστοιχα σχόλια βρίσκονται στις επόμενες παραγράφους, ενώ το διάγραμμα ροής αυτού του σεναρίου παρατίθεται στην επόμενη σελίδα.



Πραγματοποίηση διανυσμάτων εκπαίδευσης -Σενάριο 4



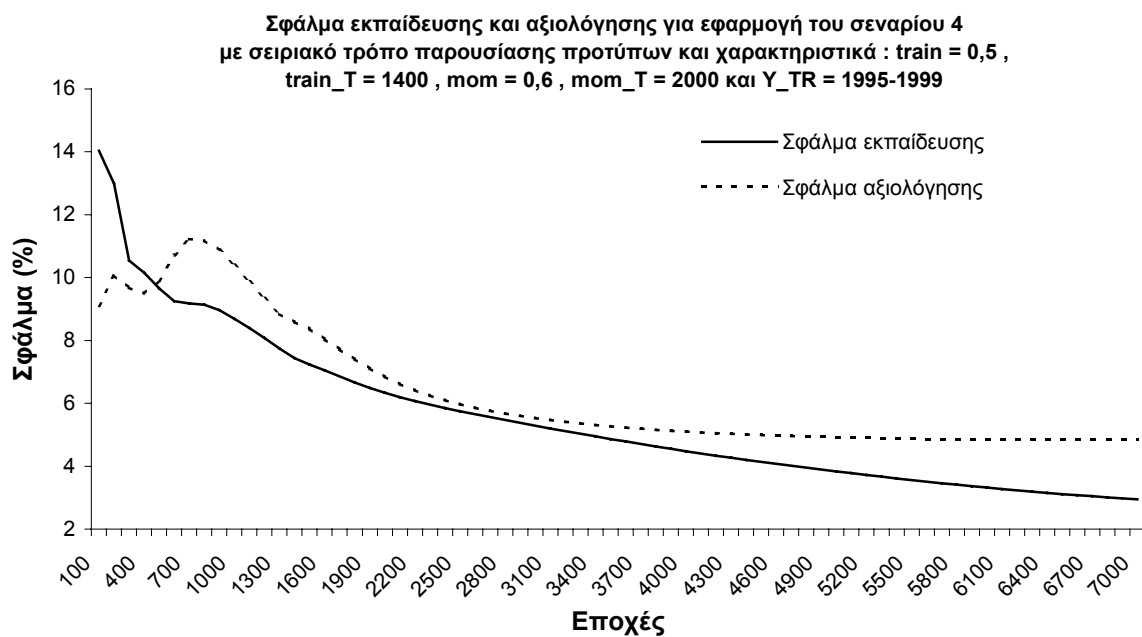
Πραγματοποίηση διανυσμάτων εκπαίδευσης- Σενάριο 4 (συνέχεια)

➤ 4.14.1 Αποτελέσματα σεναρίου 4 -Γραφήματα

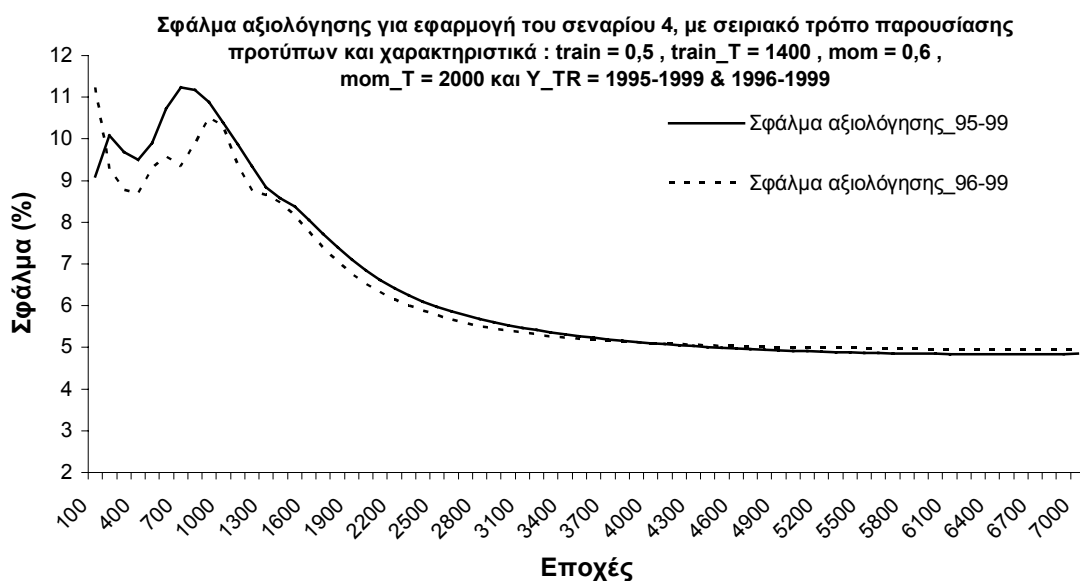
Ακολουθεί πίνακας συγκεντρωτικών δεδομένων για την εφαρμογή του σεναρίου 4 :

Πλήθος νευρώνων	48	50	55	60	65	70	75	80
Σφάλμα εκπαίδευσης (%)	2,52	2,95	2,45	2,70	2,70	2,79	2,97	2,88
Σφάλμα αξιολόγησης (%)	4,38	4,85	4,32	4,37	4,27	4,19	6,91	6,88

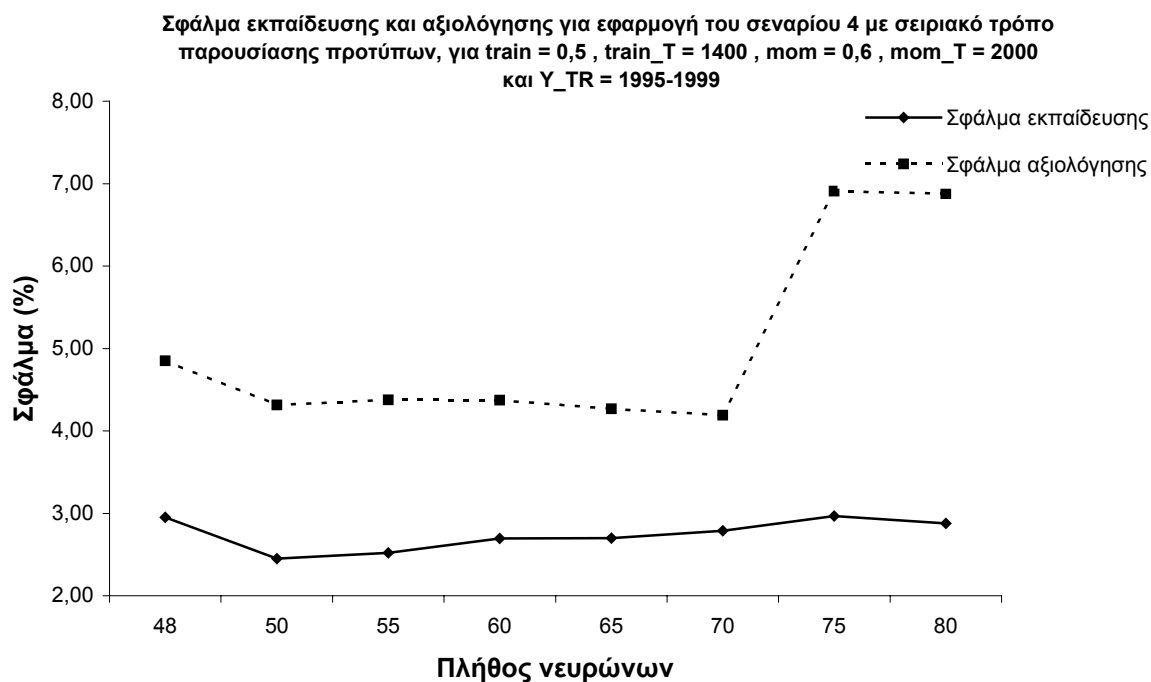
Πίνακας 4.14.1-1 : Συγκεντρωτικά αποτελέσματα για την εφαρμογή του σεναρίου 4 με σειριακό τρόπο παρουσίασης προτύπων, για $\text{train} = 0,5$, $\text{train_T} = 1400$, $\text{mom} = 0,6$, $\text{mom_T} = 2000$ και $\text{Y_TR} = 1995-1999$



Γράφημα 4.14.1-2



Γράφημα 4.14.1-3



Γράφημα 4.14.1-4

➤ 4.14.2 Παρατηρήσεις - Σχόλια για το σενάριο 4

Σχολιάζοντας το σενάριο 4 καταρχήν μπορούμε να πούμε ότι με την αύξηση του αριθμού των νευρώνων του κρυμμένου στρώματος, μειώνουμε σε σημαντικό βαθμό το σφάλμα αξιολόγησης, μέχρι το πλήθος των 70 νευρώνων. Μάλιστα από σφάλμα αξιολόγησης 4,85 % που εμφανίζεται για πλήθος νευρώνων ίσο με 48, καταλήγουμε σε σφάλμα 4,19 % για πλήθος νευρώνων ίσο με 70. Η βελτίωση λοιπόν είναι της τάξης του 0,7 %. Αντίθετα δεν μπορούμε να πούμε όμως το ίδιο για το σφάλμα εκπαίδευσης, το οποίο με την αύξηση του πλήθους των νευρώνων δεν παρουσιάζει κάποια σταθερή τάση, αλλά αυξομειώνεται. Όμως, για πλήθος μεγαλύτερο του 70, τα σφάλματα αρχίζουν να αυξάνουν ραγδαία, για παράδειγμα αναφέρουμε ότι στους 80 νευρώνες το σφάλμα αξιολόγησης, ανεβαίνει στο 6,88 % (!).

Λόγω του ότι η διαφορά μεταξύ των σφαλμάτων αξιολόγησης και εκπαίδευσης δεν είναι, σε κάθε περίπτωση, μεγάλη, θεωρούμε ότι το σενάριο εκτελέστηκε πετυχημένα ως προς το σκέλος της σωστής κατασκευής των διανυσμάτων. Φυσικά κρίνονται απαραίτητες περαιτέρω βελτιώσεις, για ένα πιο ικανοποιητικό αποτέλεσμα σφάλματος.

Όσον αφορά τη μεταβολή των ετών εκπαίδευσης του δικτύου, δείχνει να έχει σχέση με το αποτέλεσμα, καθώς για το σφάλμα εκπαίδευσης η διαφορά είναι 0,1 %, ενώ για το σφάλμα αξιολόγησης είναι 0,11 %.

Τέλος ο χρόνος που απαιτήθηκε για μια εκτέλεση του σεναρίου, είναι περίπου 4,5 ώρες.

➤ 4.15 Σενάριο εκπαίδευσης 5

Το επόμενο σενάριο είναι το σενάριο 5 το οποίο, όσον αφορά τη δομή των διανυσμάτων είναι παρόμοιο με προηγούμενα σενάρια έχει, όμως την σημαντική διαφοροποίηση όσον αφορά την ποσότητα, αλλά και τη σχέση μεταξύ των διανυσμάτων, που λαμβάνουν χώρα στην διαδικασία της εκπαίδευσης. Συγκεκριμένα το σενάριο μας δέχεται στο σύνολο των διανυσμάτων εκπαίδευσης *μονάχα τις ημέρες που ανήκουν σε μια συγκεκριμένη ομάδα*, η οποία ορίζεται από την παράμετρο I_OMADA στην αρχή του προγράμματος μας. Αυτή η σκέψη που οδήγησε στη δημιουργία αυτού του σεναρίου έχει να κάνει με τον επιμερισμό της όλης διαδικασίας της εκπαίδευσης, σε μικρότερες διαδικασίες, εκ των οποίων η καθεμιά θα εκπαιδεύει το νευρωνικό δίκτυο, για μια συγκεκριμένη ομάδα. Προφανώς και η αντίστοιχη αξιολόγηση θα γίνεται από διανύσματα που ανήκουν στην ίδια ομάδα.

Ουσιαστικά με αυτόν τον τρόπο στο τέλος της μεθόδου, αν θέλουμε να προβλέψουμε το φορτίο για κάθε ομάδα, δεν δημιουργούμε ένα, αλλά τόσα δίκτυα, όσα και ο αριθμός των ομάδων. Αυτομάτως τίθεται το ερώτημα λοιπόν, αν η κάθε διαδικασία που εκπαιδεύει ημέρες που έχουν παρόμοια συμπεριφορά ως προς το φορτίο, μας δώσει ή όχι καλύτερα σφάλματα εκπαίδευσης, αλλά και αξιολόγησης από ότι η συνολική διαδικασία εκπαίδευσης.

Παρακάτω παρουσιάζουμε τη δομή του σεναρίου :

- ✓ 24 ωριαία φορτία της προηγούμενης μέρας $\Rightarrow$ Θέσεις 1-24
- ✓ 24 ωριαία φορτία της πιο προηγούμενης μέρας $\Rightarrow$ Θέσεις 25-48
- ✓ Μέγιστη και ελάχιστη θερμοκρασία της προηγούμενης μέρας για Αθήνα –σταθμός 1 - και Θεσσαλονίκη –σταθμός 2 $\Rightarrow$ Θέσεις 49-52
- ✓ Μέγιστη και ελάχιστη θερμοκρασία της τρέχουσας μέρας για Αθήνα και Θεσσαλονίκη $\Rightarrow$ Θέσεις 53-56
- ✓ Τετραγωνικές αποκλίσεις των μεγίστων θερμοκρασιών θεωρώντας, αυθαίρετα, ότι $T_{\min}=18$ & $T_{\max}=24$ για Αθήνα και Θεσσαλονίκη, για τρέχουσα και προηγούμενη μέρα $\Rightarrow$ Θέσεις 57-60
- ✓ Διαφορές των μεγίστων θερμοκρασιών της προηγούμενης από την τρέχουσα μέρα για Αθήνα και Θεσσαλονίκη $\Rightarrow$ Θέσεις 61-62
- ✓ Αριθμοί που εκφράζουν και οι δυο την περιοδικότητα της ημέρας ως προς την εβδομάδα και δίνονται από τις τριγωνομετρικές συναρτήσεις της μορφής $\cos (2\pi/7)$ & $\sin (2\pi/7)$ $\Rightarrow$ Θέσεις 63-64
- ✓ Αριθμοί που εκφράζουν την περιοδικότητα της ημέρας ως προς το έτος και δίνονται από τις τριγωνομετρικές συναρτήσεις της μορφής $\cos (2\pi/365)$ & $\sin (2\pi/365)$. $\Rightarrow$ Θέσεις 65-66

Υπενθυμίζουμε εδώ ότι τα παρακάτω βήματα εκτελούνται *μονάχα για τις ημέρες που ανήκουν στην ομάδα της επιλογής μας*.

Κατά την εφαρμογή του σεναρίου 5, στην πρώτη φάση, η παρουσίαση των προτύπων εκπαίδευσης ήταν *σειριακή* και τα έτη εκπαίδευσης ήταν από το 1995-1999 με το έτος αξιολόγησης να είναι το 2000. Αυτό το σενάριο δοκιμάστηκε για τις ομάδες 7 , 9 , 10 και 13. Η 7^η ομάδα αποτελείται από τα Σαββατοκύριακα του χειμώνα, η 9^η από τις εργάσιμες του χειμώνα, 10^η από τις

εργάσιμες των ημερών εκείνων που έχουμε έντονο καλοκαίρι, ενώ η 13^η από τις εργάσιμες και τα Σαββατοκύριακα του χειμώνα, όπου όμως το κρύο είναι έντονο.

Οι παράμετροι για τις οποίες εκτελέστηκε το σενάριο ήταν οι εξής:

$r_momentum_arx = 0,6$, $r_momentum_T = 2000$, $trainrate_arx = 0,5$ και $trainrate_T = 1400$.

Όσον αφορά τον αριθμό των νευρώνων του κρυμμένου του στρώματος εκτελέσαμε το σενάριο για $iq1 = 48$, όπου έχει βρεθεί ως το καλύτερο πλήθος νευρώνων για το σενάριο 2, που είναι ίδιων εισόδων (66).

Στη δεύτερη φάση των εκτελέσεων, δοκιμάστηκε ο *τυχαίος* τρόπος παρουσίασης των προτύπων εκπαίδευσης, ώστε να είναι εφικτή η σύγκριση με το σειριακό, για τις ίδιες ομάδες και για τις εξής παραμέτρους :

$r_momentum_arx = 0,4$, $r_momentum_T = 1000$, $trainrate_arx = 0,5$ και $trainrate_T = 1400$

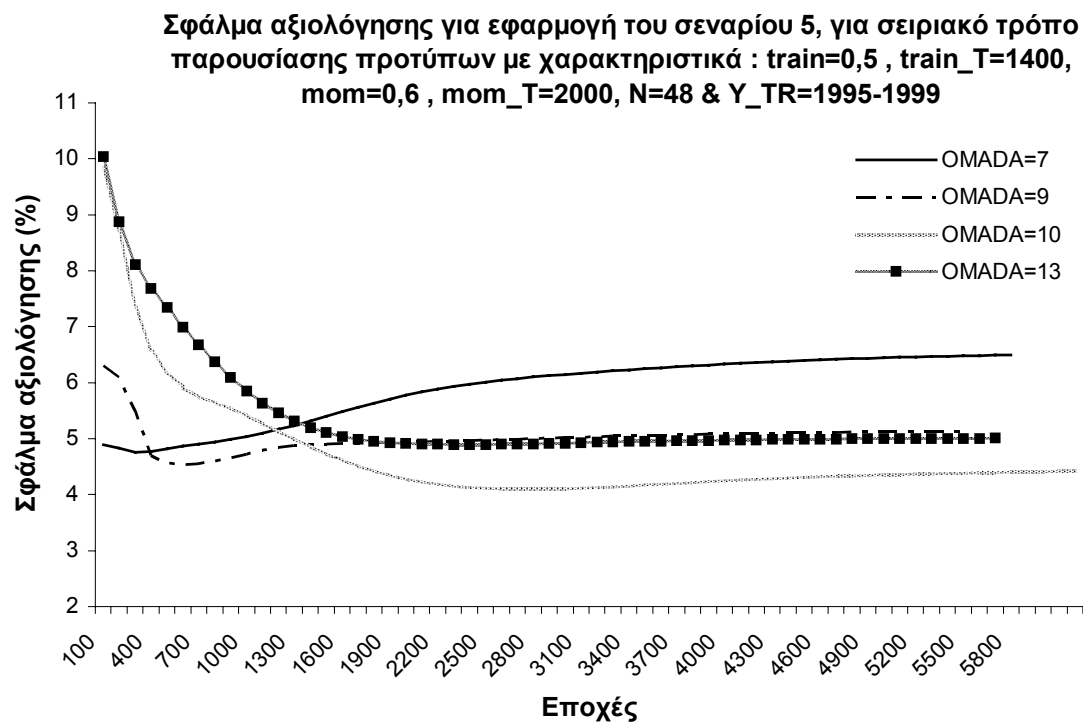
Τα αποτελέσματα και μερικές εκ των γραφικών παραστάσεων που προέκυψαν (οι υπόλοιπες βρίσκονται στο παράρτημα I), καθώς και τα αντίστοιχα σχόλια βρίσκονται στις επόμενες παραγράφους.

➤ 4.15.1 Αποτελέσματα σεναρίου 5 -Γραφήματα

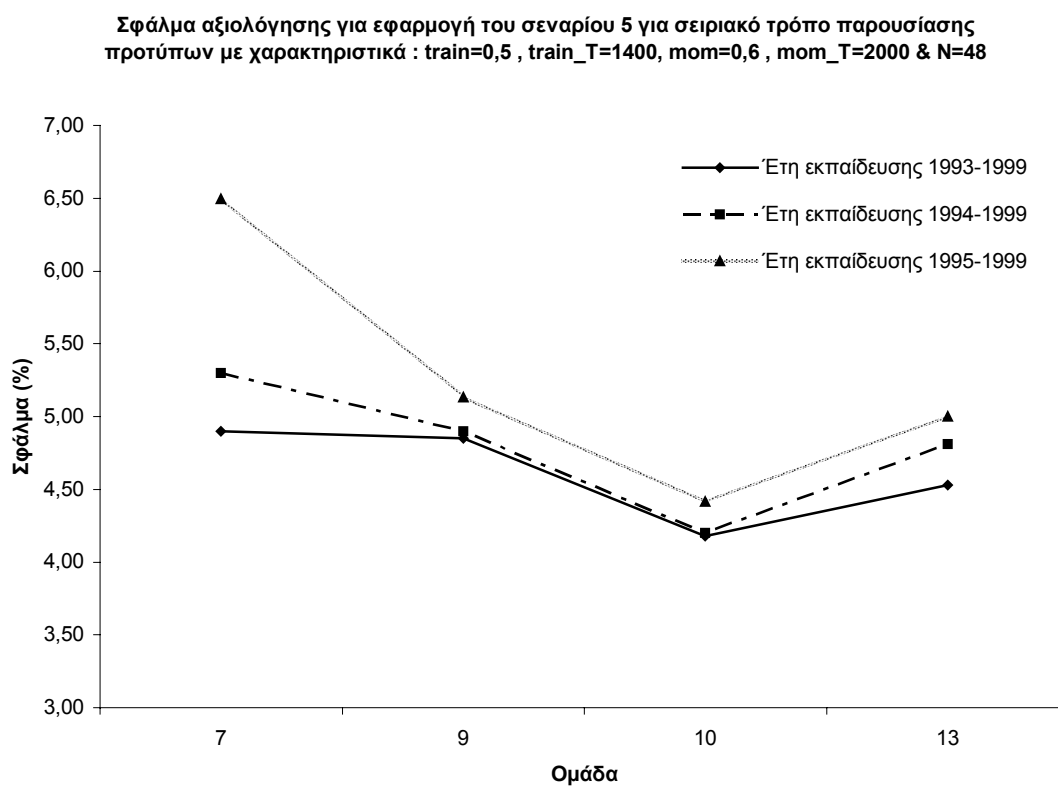
Ακολουθεί πίνακας συγκεντρωτικών αποτελεσμάτων για την εφαρμογή του σεναρίου 5 :

Ομάδα Ετη εκπαίδευσης	7	9	10	13
1993-1999	4,90	4,85	4,18	4,53
1994-1999	5,30	4,90	4,20	4,81
1995-1999	6,49	5,13	4,42	5,00

Πίνακας 4.15.1-1 : Συγκεντρωτικά αποτελέσματα σφάλματος αξιολόγησης για το σενάριο 5, με σειριακό τρόπο παρουσίασης προτύπων, για $train = 0,5$, $train_T = 1400$, $mom = 0,6$, $mom_T = 2000$ και $N=48$



Γράφημα 4.15.1-2



Γράφημα 4.15.1-3



Γράφημα 4.15.1-4

➤ 4.15.2 Παρατηρήσεις - Σχόλια για το σενάριο 5

Για το σενάριο 5, όπου θυμίζουμε ότι τα διανύσματα που συμμετέχουν στην εκπαίδευση ανήκουν σε συγκεκριμένη ομάδα της επιλογής μας, μπορούμε να επισημάνουμε ότι τα αποτελέσματα εξαρτώνται άμεσα από την ομάδα που επιλέγουμε, κάθε φορά. Για παράδειγμα, όσον αφορά την πρώτη φάση των εκτελέσεων, για την ομάδα 7 το σφάλμα αξιολόγησης αγγίζει την, ιδιαίτερα μεγάλη, τιμή του 6,5 %, ενώ για την ομάδα 10 το ίδιο σφάλμα αγγίζει το 4,4 %, μια σαφώς χαμηλότερη τιμή, για έτη εκπαίδευσης από 1995-1999.

Πάντως με την αύξηση των ετών εκπαίδευσης βελτιώνονται τα πράγματα, κάτι απόλυτα λογικό, αφού ο αριθμός των μελών της κάθε ομάδας δεν είναι ιδιαίτερα μεγάλος ανά έτος. Πάραυτα βλέπουμε ότι υπάρχουν περιθώρια σημαντικής βελτίωσης όσον αφορά το μέγεθος των σφαλμάτων. Η επίδραση της επιλογής της ομάδας στο τελικό αποτέλεσμα δικαιολογείται, καθώς κάθε μία περιέχει έναν συγκεκριμένο αριθμό ημερών, τα κοινά χαρακτηριστικά των οποίων μπορούν επιδράσουν, είτε θετικά, είτε αρνητικά στην όλη μέθοδο.

Μιλώντας για τη δεύτερη φάση των εκτελέσεων, όπου έχουμε *τυχαίο* και όχι *σειριακό* τρόπο παρουσίασης προτύπων, παρατηρούμε μια σημαντική βελτίωση στα σφάλματα. Για παράδειγμα η ομάδα 7 έχει σφάλμα αξιολόγησης κοντά στο 4,6 % (βελτίωση της τάξης του 1,9 %, σε σχέση με την 1^η φάση), ενώ στην 10 το ίδιο σφάλμα αγγίζει το 4,1 (βελτίωση της τάξης του 0,3 %). Σε κάθε περίπτωση λοιπόν, παρατηρούμε ότι ο *τυχαίος* τρόπος παρουσίασης των προτύπων επιδρά θετικά, κατά την εφαρμογή του σεναρίου 5.

Τέλος ο χρόνος για μια εκτέλεση του σεναρίου είναι περίπου 20 λεπτά / ομάδα, κάτι που συμβαδίζει απόλυτα με τη λογική, καθώς η κάθε ομάδα από τις 15 περιλαμβάνει ένα μικρό ποσοστό των ημερών ενός ολόκληρου έτους, οπότε ο συνολικός αριθμός των διανυσμάτων εκπαίδευσης είναι μικρός, σε σχέση με τα προηγούμενα σενάρια.

➤ 4.16 Σενάριο εκπαίδευσης 6

Στο έκτο και τελευταίο σενάριο στην ουσία «σπάμε» το νευρωνικό δίκτυο σε επιμέρους, σημαντικά μικρότερων διαστάσεων, καθώς τώρα δεν επιθυμούμε πρόβλεψη για τα 24 ωριαία φορτία της επόμενης μέρας, αλλά για το φορτίο μιας συγκεκριμένης επιθυμητής ώρας –IDHOUR. Εδώ θα πρέπει να σημειωθεί ότι το σύνολο των διανυσμάτων εκπαίδευσης, περιλαμβάνει όλες τις υπάρχουσες ομάδες. Η είσοδος λοιπόν του κάθε διανύσματος εκπαίδευσης, σε αυτήν την περίπτωση, θα απαρτίζεται κατά σειρά από τα εξής στοιχεία :

- ✓ Το φορτίο της ώρας που επιθυμούμε για την προηγούμενη μέρα $\Rightarrow$ Θέση 1
- ✓ Το φορτίο της ώρας που επιθυμούμε για την πιο προηγούμενη μέρα $\Rightarrow$ Θέση 2
- ✓ Το φορτίο της ώρας που επιθυμούμε για την προηγούμενη ημερολογιακά μέρα η οποία όμως ανήκει στην ίδια ομάδα με την τρέχουσα $\Rightarrow$ Θέση 3
- ✓ Το τρίωρο της θερμοκρασίας που αντιστοιχεί στην επιθυμητή ώρα, για την προηγούμενη ημερολογιακά μέρα η οποία όμως ανήκει στην ίδια ομάδα με την τρέχουσα, για Αθήνα και Θεσσαλονίκη $\Rightarrow$ Θέσεις 4-5
- ✓ Το τρίωρο της θερμοκρασίας που αντιστοιχεί στην επιθυμητή ώρα, για την προηγούμενη ημερολογιακά μέρα για Αθήνα και Θεσσαλονίκη $\Rightarrow$ Θέσεις 6-7
- ✓ Το τρίωρο της θερμοκρασίας που αντιστοιχεί στην επιθυμητή ώρα για την τρέχουσα μέρα, για Αθήνα και Θεσσαλονίκη $\Rightarrow$ Θέσεις 8-9
- ✓ Αριθμοί που εκφράζουν και οι δυο την περιοδικότητα της ημέρας ως προς την εβδομάδα και δίνονται από τις τριγωνομετρικές συναρτήσεις της μορφής $\cos (2\pi/7)$ & $\sin (2\pi/7)$
 $\Rightarrow$ Θέσεις 10-11
- ✓ Αριθμοί που εκφράζουν την περιοδικότητα της ημέρας ως προς το έτος και δίνονται από τις τριγωνομετρικές συναρτήσεις της μορφής $\cos (2\pi/365)$ & $\sin (2\pi/365)$. $\Rightarrow$ Θέσεις 12-13

Όπως διαπιστώνουμε στο παραπάνω σενάριο το διάνυσμα εισόδου έχει διαστάσεις 13×1 και είναι από 5 έως και 8 περίπου φορές μικρότερο, σε σχέση με τα προηγούμενα σενάρια, ενώ το διάνυσμα εξόδου διαφέρει αυτήν τη φορά από τις προηγούμενες, καθώς έχει διαστάσεις 1×1 , δηλαδή περιέχει μόνο ένα στοιχείο, που είναι το φορτίο της επιθυμητής ώρας για την τρέχουσα μέρα.

Με αυτόν τον τρόπο θα μπορούσαμε να τρέξουμε τον αλγόριθμο για την κάθε ώρα της ημέρας ξεχωριστά και να διαπιστώσουμε, αν το «σπάσιμο» της διαδικασίας σε αυτήν την περίπτωση βοηθάει στην εξάλειψη μέρους της πολυπλοκότητας του προβλήματος μας και οδηγεί τελικά στην μείωση του μέσου σφάλματος, τόσο για την εκπαίδευση, όσο και για την αξιολόγηση.

Κατά την εφαρμογή του σεναρίου 6 δοκιμάστηκαν ως επιθυμητές ώρες η 4^η, η 12^η και η 22^η. Για την πρώτη φάση και για την καθεμία ώρα μάλιστα εκτελέστηκε το σενάριο από 4 φορές για τα έτη εκπαίδευσης από το 1994-1999, 1995-1999, 1996-1999 και 1997-1999 ενώ το έτος αξιολόγησης είναι πάντα το 2000. Ο αριθμός των νευρώνων για τις παραπάνω εκτελέσεις ήταν $iq1 = 48$.

Ακολουθώς για τη δεύτερη φάση το σενάριο εκτελέστηκε για την 22η ώρα (όπου και είχαμε το βέλτιστο αποτέλεσμα από την πρώτη φάση) και για τα πλήθη των νευρώνων από $iq1 = 10-12-14-16-18-20$. Στη συνέχεια της δεύτερης φάσης εκτελέσαμε το σενάριο για όλες τις ώρες από 1-24 για τον βέλτιστο αριθμό νευρώνων που βρέθηκε ίσος με 14. Τα έτη εκπαίδευσης για όλη αυτήν τη δεύτερη φάση των δοκιμών ήταν από 1995-1999, καθώς μας έδωσαν τα καλύτερα αποτελέσματα από την πρώτη φάση. Οι παράμετροι για την πρώτη και δεύτερη φάση, περιγράφονται παρακάτω, ενώ η παρουσίαση προτύπων ήταν *σειριακή*.

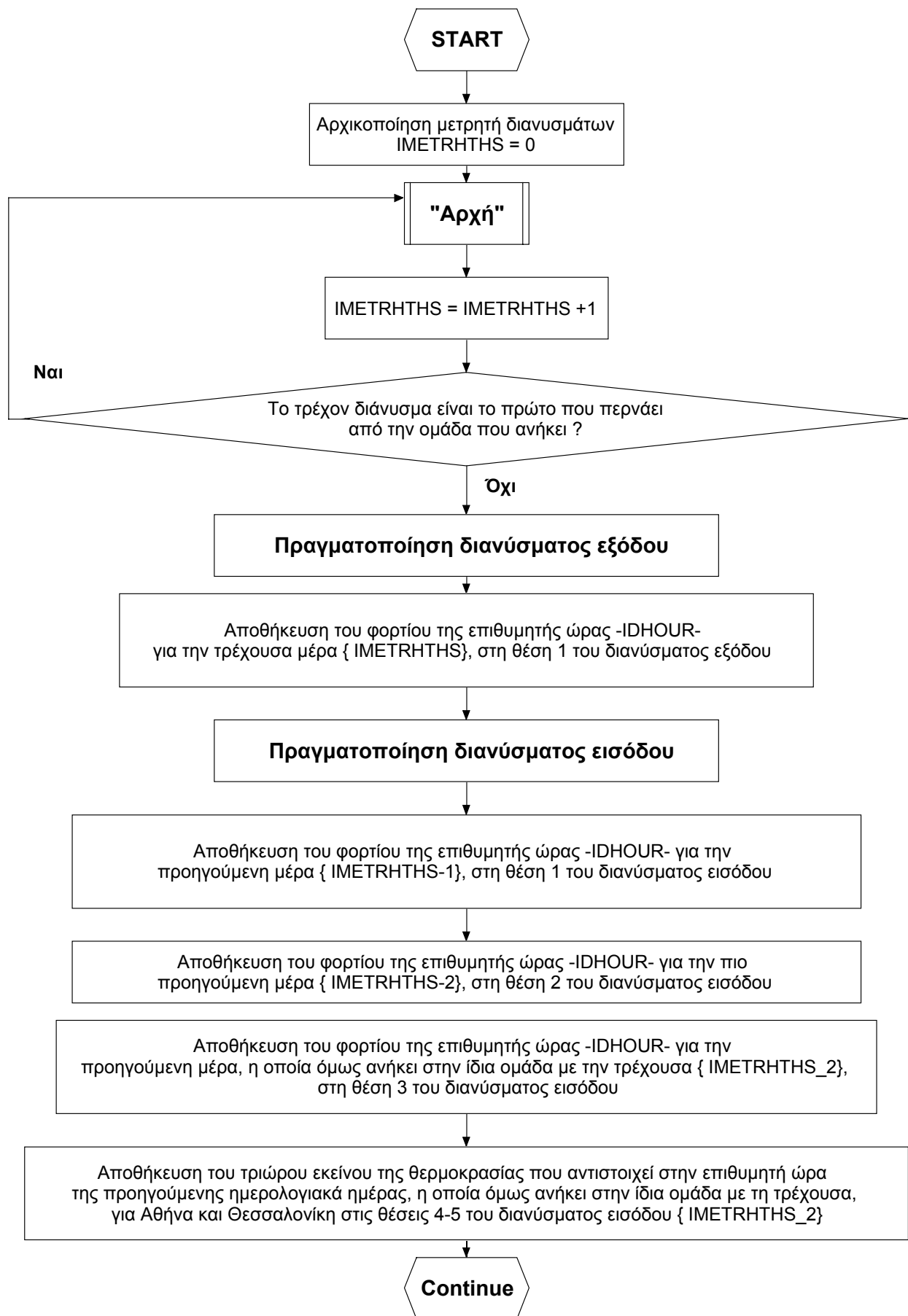
$r_momentum_arx = 0,6$, $r_momentum_T = 2000$, $trainrate_arx = 0,5$ και $trainrate_T = 1400$

Τέλος στην τρίτη φάση, εκτελέσαμε το σενάριο για όλες τις ώρες της ημέρας, από 1-24, με τυχαίο τρόπο παρουσίασης προτύπων, ώστε να είναι εφικτή η σύγκριση με τον σειριακό τρόπο. Το πλήθος νευρώνων κρυμμένου επιπέδου ήταν 14, ενώ οι παράμετροι μάθησης και ορμής :

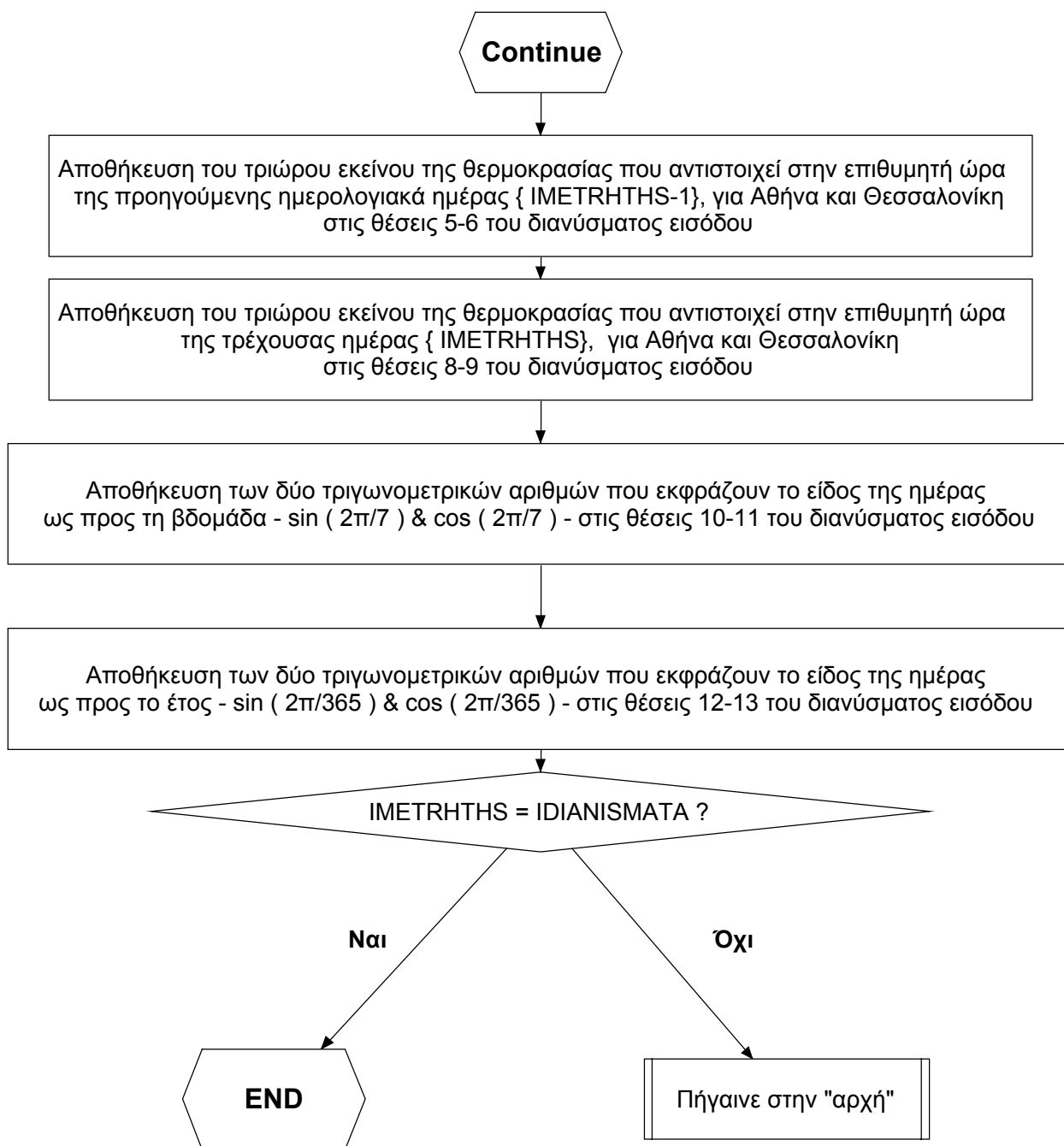
$r_momentum_arx = 0,4$, $r_momentum_T = 1000$, $trainrate_arx = 0,5$ και $trainrate_T = 1400$

Τα αποτελέσματα και μερικές εκ των γραφικών παραστάσεων που προέκυψαν (οι υπόλοιπες βρίσκονται στο παράρτημα I), καθώς και τα αντίστοιχα σχόλια βρίσκονται στις επόμενες παραγράφους.

Στην επόμενη σελίδα παρατίθεται το διάγραμμα ροής για την εφαρμογή του σεναρίου 6.



Πραγματοποίηση διανυσμάτων εκπαίδευσης -Σενάριο 6



Πραγματοποίηση διανυσμάτων εκπαίδευσης- Σενάριο 6 (συνέχεια)

➤ 4.16.1 Αποτελέσματα σεναρίου 6 -Γραφήματα

Ακολουθούν πίνακες συγκεντρωτικών δεδομένων για την εφαρμογή του σεναρίου 6 :

ΩΡΑ ΕΤΗ ΕΚΠΑΙΔΕΥΣΗΣ	4 ^η	12 ^η	22 ^η
1994-1999	3,46 %	3,31 %	2,78 %
1995-1999	2,9 %	2,78 %	2,57 %
1996-1999	3,3 %	4 %	3,16 %
1997-1999	3,9 %	3,45 %	3,18 %

Πίνακας 4.16.1-1 → Αποτελέσματα σεναρίου 6 για σφάλμα αξιολόγησης, με σειριακό τρόπο παρουσίασης προτύπων και για $train = 0,5$, $train_T = 1400$, $mom = 0,6$, $mom_T = 2000$ και $N=48$

ΩΡΑ ΕΤΗ ΕΚΠΑΙΔΕΥΣΗΣ	4 ^η	12 ^η	22 ^η
1994-1999	2,97 %	2,9 %	2,53 %
1995-1999	2,94 %	2,9 %	2,52 %
1996-1999	3,12 %	3,19 %	2,46 %
1997-1999	3,08 %	2,79 %	2,55 %

Πίνακας 4.16.1-2 → Αποτελέσματα σεναρίου 6 για σφάλμα εκπαίδευσης με σειριακό τρόπο παρουσίασης προτύπων, για $train = 0,5$, $train_T = 1400$, $mom = 0,6$, $mom_T = 2000$ και $N=48$

Πλήθος νευρώνων	10	12	14	16	18	20	48
Σφάλμα εκπαίδευσης (%)	2,40	2,43	2,41	2,45	2,46	2,47	2,52
Σφάλμα αξιολόγησης (%)	2,54	2,55	2,52	2,56	2,56	2,55	2,57

Πίνακας 4.16.1-3 → Αποτελέσματα σεναρίου για την 22η ώρα, με σειριακό τρόπο παρουσίασης προτύπων, για $train = 0,5$, $train_T = 1400$, $mom = 0,6$, $mom_T = 2000$ και $Y_TR=1995-1999$

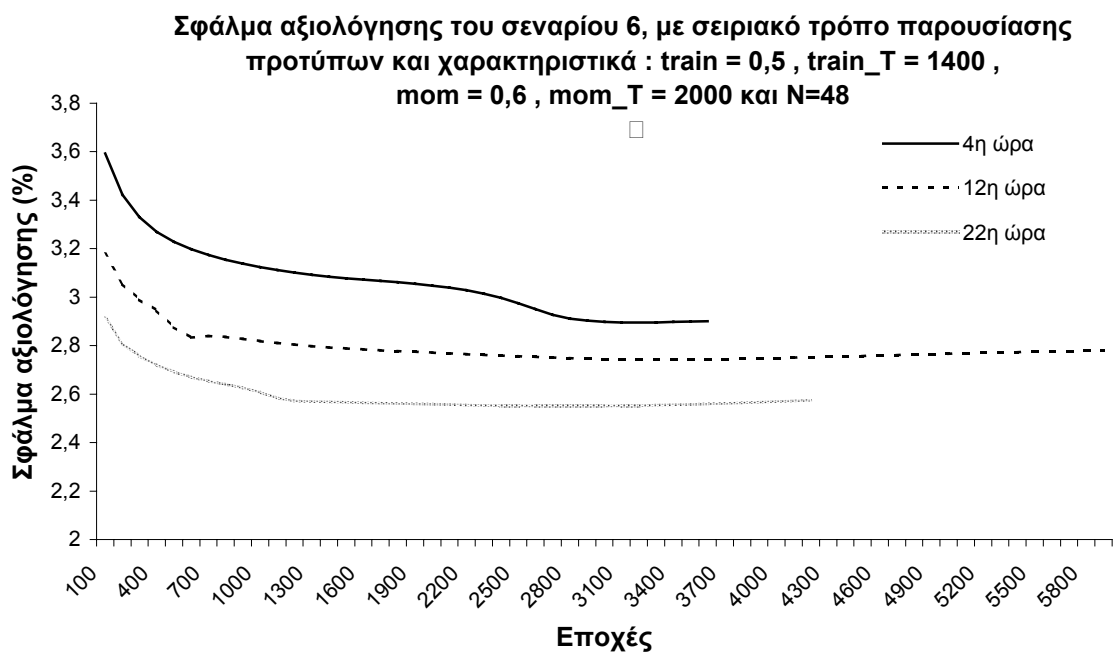
	1	2	3	4	5	6	7	8	9	10	11	12
Εκπ.	2,53	2,87	2,98	2,85	2,83	2,74	2,92	3,55	3,57	3,37	2,95	2,90
Αξ.	2,90	3,09	2,91	2,86	3,05	2,98	3,29	3,46	3,74	3,51	2,96	2,79

Πίνακας 4.16.1-4 → Σφάλματα αξιολόγησης και εκπαίδευσης (%) σεναρίου 6, για εκτέλεση για τις ώρες 1-12, με σειριακό τρόπο παρουσίασης προτύπων και $train = 0,5$, $train_T = 1400$, $mom = 0,6$, $mom_T = 2000$, $N=14$ και $Y_TR=1995-1999$

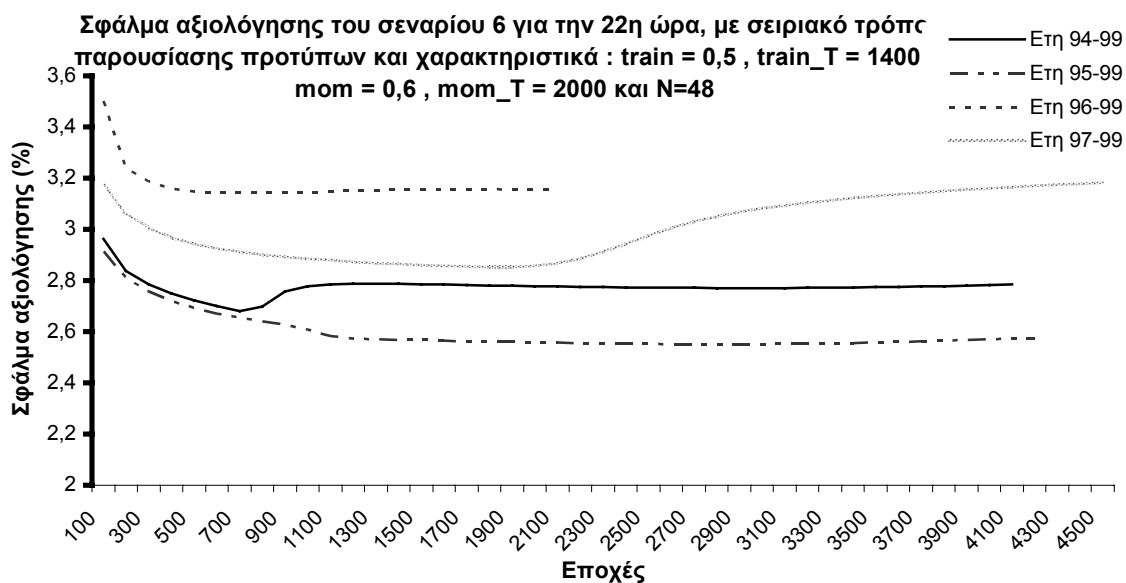
	13	14	15	16	17	18	19	20	21	22	23	24
Εκπ.	3,04	3,30	3,80	3,67	3,47	3,21	3,10	2,86	2,63	2,44	2,36	2,22
Αξιολ.	2,83	3,25	3,83	3,49	3,24	3,29	3,22	3,11	3,01	2,57	2,54	2,49

Πίνακας 4.16.1-5 → Σφάλματα αξιολόγησης και εκπαίδευσης (%) σεναρίου 6 για εκτέλεση για τις ώρες 13-24, με σειριακό τρόπο παρουσίασης προτύπων, για $train = 0,5$, $train_T = 1400$, $mom = 0,6$, $mom_T = 2000$, $N=14$ και $Y_TR=1995-1999$

Ακολουθούν αντίστοιχα γραφήματα :

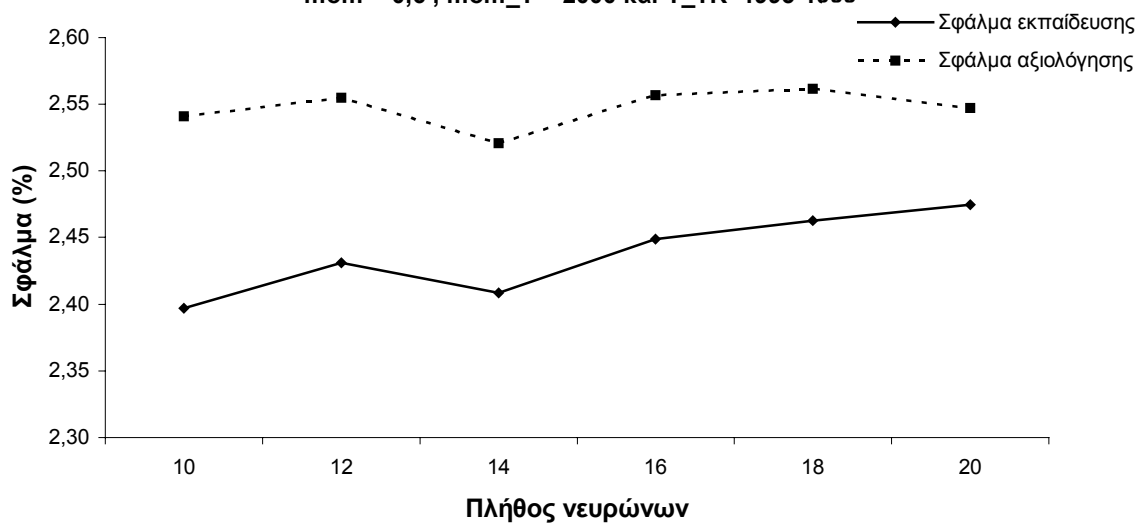


Γράφημα 4.16.1-6



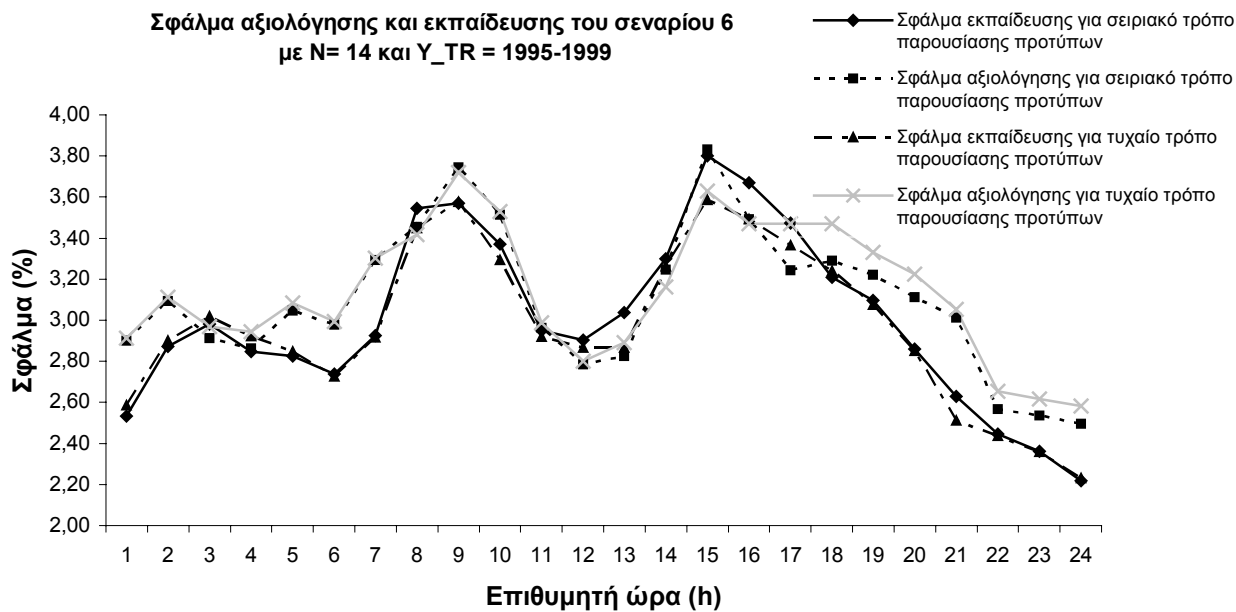
Γράφημα 4.16.1-7

Σφάλματα εκπαίδευσης και αξιολόγησης του σεναρίου 6 για την 22η ώρα, με
σειριακό τρόπο παρουσίασης προτύπων και $\text{train} = 0,5$, $\text{train}_T = 1400$,
 $\text{mom} = 0,6$, $\text{mom}_T = 2000$ και $Y_{TR}=1995-1999$



Γράφημα 4.16.1-8

Σφάλμα αξιολόγησης και εκπαίδευσης του σεναρίου 6
με $N = 14$ και $Y_{TR} = 1995-1999$



Γράφημα 4.16.1-9

➤ 4.16.2 Παρατηρήσεις - Σχόλια για το σενάριο 6

Όσον αφορά το σενάριο 6, όπου υπενθυμίζουμε ότι προβλέπουμε, όχι τα 24 φορτία της επόμενης μέρας, αλλά *μονάχα το φορτίο μια συγκεκριμένης ώρας*, την οποία και επιλέγουμε εμείς, στην πρώτη φάση, μπορούμε να δούμε καλύτερα την επίδραση των ετών εκπαίδευσης, αλλά και της επιλογής της ώρας στη διαδικασία της μεθόδου.

Καταρχήν παρατηρούμε, ότι τα έτη εκπαίδευσης παίζουν ιδιαίτερα σημαντικά ρόλο στο τελικό αποτέλεσμα. Για παράδειγμα αναφέρουμε, ότι για την 12η ώρα το σφάλμα αξιολόγησης αγγίζει το 2,78 % για έτη εκπαίδευσης από 1995-1999, ενώ φτάνει στο 4 % για έτη εκπαίδευσης από 1996-1999. Βλέπουμε λοιπόν ότι η μείωση του συνολικού χρόνου εκπαίδευσης κατά ένα έτος, προκαλεί σε αυτήν την περίπτωση αύξηση του σφάλματος κατά 1,2 % περίπου. Αυτή η επίδραση των ετών εκπαίδευσης, ιδιαίτερα στο σφάλμα αξιολόγησης, φαίνεται πολύ καθαρά στις γραφικές παραστάσεις 4.16.1-6 και 4.16.1-7, αλλά και στους πίνακες 4.16.1-1 και 4.16.1-2.

Παρόμοια επίδραση στα τελικά σφάλματα, και ιδιαίτερα πάλι στο σφάλμα αξιολόγησης, έχει και η επιλογή της επιθυμητής ώρας, για την οποία γίνεται η πρόβλεψη, κάτι που φαίνεται ξεκάθαρα και στο γράφημα 4.16.1-6, αλλά και στους πίνακες 4.16.1-1 και 4.16.1-2. Σαν παράδειγμα μόνο, αναφέρουμε ότι για τα ίδια έτη εκπαίδευσης από 1994-1999, το σφάλμα αξιολόγησης για την 4η ώρα είναι 3,46 %, ενώ για την 22η ώρα είναι 2,78 %, δηλαδή υπάρχει διαφορά της τάξης περίπου του 0,7 %.

Επιπρόσθετα θα πρέπει να αναφέρουμε ότι πάντα στο σενάριο 6 η σύγκλιση των βαρών, που εξετάζεται μετά το πέρας κάθε εποχής, επέρχεται πολύ πριν τις 7000 επαναλήψεις. Μάλιστα η πιο γρήγορη σύγκλιση γενικά έγινε στις 1900 εποχές.

Άλλη μια σημαντική παρατήρηση που έχουμε να κάνουμε, είναι ότι για τις περισσότερες εκτελέσεις, το σφάλμα αξιολόγησης μειώνεται, με σημαντικό ρυθμό, μέχρι κάποια εποχή, αλλά από εκεί και πέρα αυξάνει, με πολύ μικρό ρυθμό βέβαια, ώσπου να πάρει την τελική του τιμή, την εποχή σύγκλισης των βαρών. Το παραπάνω φαινόμενο οφείλεται στην αποστήθιση της συγκεκριμένης παρουσίασης προτύπων, από πλευράς δικτύου.

Για τη δεύτερη φάση αυτό που έχουμε να παρατηρήσουμε, για το συγκεκριμένο σενάριο, είναι ότι η επίδραση των νευρώνων στα τελικά αποτελέσματα δεν είναι ιδιαίτερα σημαντική, καθώς το πλήθος των νευρώνων κινούμενο μέσα στο διάστημα [10,20] αλλά και για την αρχική τιμή 48 προκαλεί μια απόλυτη μεταβολή στα σφάλματα της τάξης του 0,01 %.

Για τη συνέχεια της δεύτερης φάσης, όπου εκτελέσαμε το σενάριο 6 για όλες τις ώρες, από 1-24 και για το βέλτιστο πλήθος νευρώνων ίσο με 14, παρατηρούμε ότι τα μικρότερα σφάλματα αξιολόγησης εμφανίζονται τις ώρες 24, 23, 22 και είναι 2,49 %, 2,54 % και 2,57 % αντίστοιχα, ενώ ταυτόχρονα, εκεί εμφανίζονται και τα μικρότερα σφάλματα εκπαίδευσης. Αντίθετα οι ώρες με τα μεγαλύτερα σφάλματα είναι η 15^η και η 9^η ώρα όπου έχουμε σφάλματα αξιολόγησης 3,83 % και 3,74 % αντίστοιχα, ενώ έχουμε και από τα μεγαλύτερα σφάλματα εκπαίδευσης. Επιπρόσθετα θα πρέπει να πούμε ότι το μέσο σφάλμα εκπαίδευσης και αξιολόγησης για τις 24 ώρες υπολογίστηκε στο 3% και 3,1 % αντίστοιχα.

Τέλος για την τρίτη φάση των εκτελέσεων, όπου «τρέξαμε» το σενάριο 6 ξανά για όλες τις ώρες και για πλήθος νευρώνων ίσο με 14, αλλά με τυχαίο τρόπο παρουσίασης προτύπων αυτήν τη φορά, παρατηρούμε ότι, για την πλειοψηφία των ωρών δεν υπάρχει ουσιαστική διαφορά σε σχέση

με τον σειριακό τρόπο παρουσίασης προτύπων, ενώ για τις ώρες αιχμής, που υπάρχει κάποια διαφορά (η μέγιστη της τάξης, περίπου, του 0,2 % για σφάλμα αξιολόγησης), είναι καλύτερος ο *τυχαίος* τρόπος παρουσίασης προτύπων.

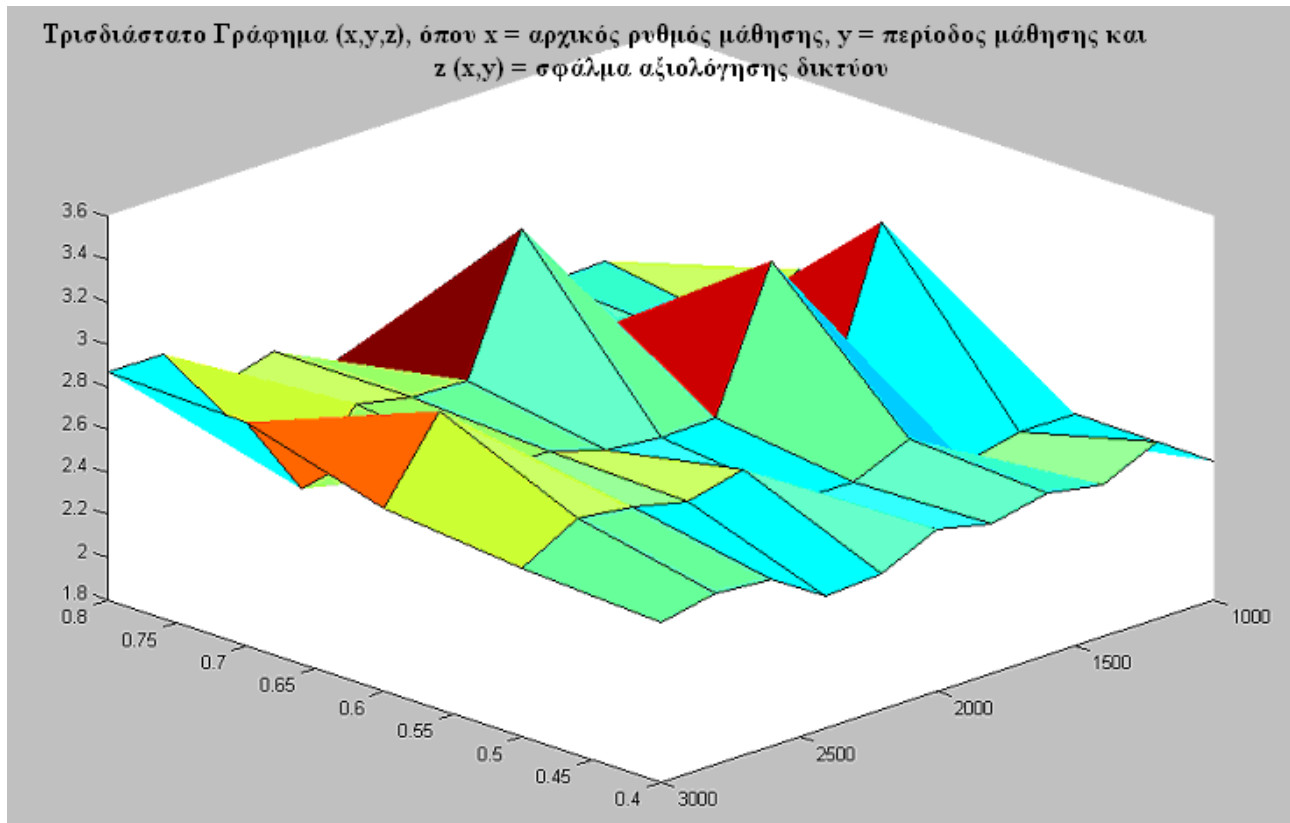
Εν κατακλείδι, συμπεραίνουμε ότι κατά την εφαρμογή του σεναρίου 6, η μέθοδος πάντα συγκλίνει πολύ νωρίτερα από το μέγιστο των 7000 επαναλήψεων και δίνει σχετικά ικανοποιητικά αποτελέσματα, ανεξάρτητα, συνήθως, από το πλήθος των νευρώνων του κρυμμένου επιπέδου. Βέβαια, αναμφίβολα, με κάποιες περαιτέρω τροποποιήσεις, θα μπορούσε να μας δώσει ακόμα πιο ικανοποιητικά αποτελέσματα, για το σύνολο των ωρών της ημέρας.

Επιπρόσθετα, θα πρέπει να πούμε ότι ο χρόνος που απαιτήθηκε για μια εκτέλεση του σεναρίου ήταν περίπου 11 λεπτά και άρα για να προβλέψουμε και τις 24 ώρες της επόμενης ημέρας, απαιτείται χρόνος περίπου ίσος με $24 \cdot 11 = 264$ λεπτά ή αλλιώς 4,4 ώρες. Οπότε ο χρόνος εκτέλεσης του σεναρίου 6 είναι παραπλήσιος με αυτόν, των προηγούμενων σεναρίων.

Ακόμα, έχει ιδιαίτερη σημασία να αναφέρουμε, ότι ο ηλεκτρονικός υπολογιστής στον οποίο και έλαβαν χώρα όλες οι εκτελέσεις, διαθέτει επεξεργαστή, περίπου στα 900 MHz και έχει μνήμη RAM 256 MB.

Τέλος, στην επόμενη παράγραφο, παραθέτουμε ένα ενδεικτικό τρισδιάστατο γράφημα (x,y,z), από προηγούμενες δοκιμές, όπου x : αρχικός ρυθμός μάθησης, y : περίοδος μάθησης και z : σφάλμα αξιολόγησης T.N.Δ.

➤ 4.17 Τρισδιάστατο Γράφημα



Γράφημα 4.17-1

Όπως παρατηρούμε από το παραπάνω γράφημα, οι 3 κορυφές που εμφανίζονται στην επιφάνεια σφάλματος, αντιστοιχούν στα 3 μέγιστα του σφάλματος αξιολόγησης 3,33 %, 3,23 %, 3,22 %. Τα συγκεκριμένα μέγιστα προκύπτουν για τα ζεύγη τιμών (αρχικός ρυθμός μάθησης, περίοδος μάθησης) :

Για 3,33 % $\Rightarrow$ (0,7 , 2000)

Για 3,23 % $\Rightarrow$ (0,6 , 1200)

Για 3,22 % $\Rightarrow$ (0,6 , 1600)

Επίσης παρατηρούμε και την ελάχιστη τιμή του σφάλματος αξιολόγησης 2,4 %, η οποία προκύπτει για (0,5 , 1400).

Κεφάλαιο 5

Τελικά Συμπεράσματα και Προτάσεις για Περαιτέρω Έρευνα

➤ 5.1 Περιγραφή της διπλωματικής εργασίας

Κατά τη διάρκεια της παρούσας διπλωματικής εργασίας, χρησιμοποιήσαμε και κατασκευάσαμε διάφορα σενάρια για το πρόγραμμα εκπαίδευσης T.N.Δ., με τη χρήση του αλγορίθμου της ανάστροφης διάδοσης σφάλματος –Back Propagation. Η όλη διαδικασία της εκπαίδευσης έχει ως σκοπό τη βραχυπρόθεσμη πρόβλεψη φορτίου, δηλαδή την πρόβλεψη των φορτίων των επόμενων 24 ωρών.

Το πρόγραμμα αυτό δέχεται διάφορες παραμέτρους, των οποίων η τιμή παίζει καθοριστικό ρόλο, όσον αφορά τα τελικά αποτελέσματα, δηλαδή τα σφάλματα εκπαίδευσης και αξιολόγησης του νευρωνικού δικτύου και τη σύγκλιση ή όχι των βαρών πριν το μέγιστο των επαναλήψεων-εποχών, που για όλες τις εκτελέσεις μας ήταν 7000. Πέραν όμως, από τις διάφορες παραμέτρους, μεγάλη σημασία έχει και ο τρόπος παρουσίασης των προτύπων εκπαίδευσης, ανά εποχή, ο οποίος μπορεί να είναι, είτε σειριακός (καθορισμένος), είτε τυχαίος (ακαθόριστος).

Κάποιες, ιδιαίτερα σημαντικές, παράμετροι του προγράμματος εκπαίδευσης T.N.Δ., είναι ο αρχικός ρυθμός μάθησης, η περίοδος μάθησης, ο αρχικός όρος ορμής και η περίοδος ορμής, τις οποίες αρχικά μεταβάλλαμε εκ περιτροπής, για τυχαίο τρόπο παρουσίασης των προτύπων εκπαίδευσης, ώστε να βρούμε τις βέλτιστες τιμές τους. Δηλαδή πρώτα θεωρήσαμε σταθερές τις

παραμέτρους της ορμής και μεταβάλλαμε, τότε τον αρχικό ρυθμό μάθησης και τότε την περίοδο μάθησης, ενώ στη συνέχεια εκτελέσαμε, ακριβώς την αντίστροφη διαδικασία. Σε όλες τις παραπάνω δοκιμές, το πλήθος νευρώνων του κρυμμένου επιπέδου ήταν 48. Τα αποτελέσματα που λάβαμε, θα μπορούσαν να χαρακτηριστούν, εν γένει, ενθαρρυντικά, όσον αφορά τις τιμές των σφαλμάτων της μεθόδου.

Ακολουθώντας κατασκευάσαμε ορισμένα σενάρια εκπαίδευσης του δικτύου, τα οποία θεωρούμε ότι μπορούν να συμμετέχουν στη βραχυπρόθεσμη πρόβλεψη φορτίου. Ουσιαστικά υποδείξαμε, στο πρόγραμμα τι ακριβώς στοιχεία θα περιλαμβάνει το κάθε διάνυσμα εκπαίδευσης (είσοδος – έξοδος) και αξιολόγησης φυσικά. Αυτά τα σενάρια, τα εκτελέσαμε για διάφορα πλήθη νευρώνων κρυμμένου επιπέδου, για διάφορα έτη εκπαίδευσης του δικτύου, καθώς και για το σειριακό (σε όλα τα σενάρια) και τυχαίο (σε κάποια σενάρια) τρόπο παρουσίασης των προτύπων.

Παρακάτω θα αναφερθούμε ξεχωριστά, σε όλα τα σενάρια εκπαίδευσης που δοκιμάσαμε και θα περιγράψουμε, τόσο τις γενικές παρατηρήσεις μας, όσο και τις αντίστοιχες προτάσεις μας που αφορούν το καθένα.

➤ 5.2 Τελικά συμπεράσματα και προτάσεις

Για το σενάριο εκπαίδευσης 1, για το οποίο τα διανύσματα εισόδου έχουν 71 στοιχεία, διαπιστώσαμε ότι το βέλτιστο πλήθος νευρώνων είναι 48, καθώς σε αυτό το σημείο τα σφάλματα εκπαίδευσης και αξιολόγησης παρουσιάζουν ελάχιστο. Για το συγκεκριμένο αριθμό νευρώνων, τα σφάλματα του σεναρίου δεν ήταν ικανοποιητικά για τον σειριακό τρόπο παρουσίασης προτύπων, αντιθέτως ήταν αρκετά ικανοποιητικά για την εφαρμογή του, με τυχαίο τρόπο παρουσίασης. Συνεπώς σε αυτήν την περίπτωση, ο τυχαίος τρόπος παρουσίασης προτύπων επέδρασε καταλυτικά ως προς το αποτέλεσμα, κάτι που εξηγείται από τη θεωρία, καθώς με τον τυχαίο τρόπο εξασφαλίζεται η στοχαστικότητα της μεθόδου, ενώ ταυτόχρονα δεν υφίσταται ο κίνδυνος αποστήθισης, από πλευράς δικτύου, μιας συγκεκριμένης σειράς προτύπων. Βέβαια, το μειονέκτημα που παρουσιάζεται είναι ότι δε μπορεί να ελεγχθεί το σφάλμα εκπαίδευσης, εφόσον μάλιστα αυξομειώνεται κατά τη διάρκεια εκπαίδευσης του δικτύου.

Όσον αφορά το σενάριο 2, για το οποίο τα διανύσματα εισόδου των προτύπων περιλαμβάνουν 66 στοιχεία το καθένα, διαπιστώσαμε πάλι, ότι το βέλτιστο πλήθος των νευρώνων είναι 48. Για σειριακό τρόπο παρουσίασης προτύπων, τα αποτελέσματα που παίρνουμε είναι σχετικά, καλύτερα από αυτά του σεναρίου 1, ενώ, βέβαια, υπάρχει περιθώριο για βελτίωση. Αυτή η βελτίωση θα μπορούσε να προκύψει, αντικαθιστώντας τις τριγωνομετρικές συναρτήσεις $\cos$ & $\sin$ -τα τελευταία 4 στοιχεία κάθε διανύσματος εισόδου-, ενδεχομένως με το αντίστοιχο τετράγωνο αυτών των συναρτήσεων ή με άλλες περιοδικές συναρτήσεις, οι οποίες θα εκφράζουν, και αυτές, την περιοδικότητα της ημέρας ως προς την εβδομάδα και το έτος.

Αναφερόμενοι στο σενάριο 3, βλέπουμε ότι το καλύτερο πλήθος νευρώνων είναι 50 και από εκεί και πέρα, όσο αυξάνουμε τους νευρώνες, τόσο τα σφάλματα τείνουν να αυξάνονται. Η παραπάνω διαπίστωση έρχεται σε αντίθεση με το γεγονός, ότι εδώ έχουμε 100 στοιχεία ανά διάνυσμα εισόδου και θα περίμενε κάποιος, το δίκτυο να χρειάζεται περισσότερους των 50 νευρώνων (τα σενάρια 1 και 2 απαιτούσαν 48 νευρώνες, ενώ τα διανύσματα εισόδου τους περιείχαν 71 και 66 στοιχεία αντίστοιχα). Φαίνεται λοιπόν, ότι για το συγκεκριμένο σενάριο, δεν είναι απαραίτητη, ουσιαστικά, η αύξηση του αριθμού των νευρώνων. Για βελτίωση του τελικού σφάλματος, θα μπορούσε κάποιος, είτε να μειώσει τον αριθμό εισόδων του δικτύου, είτε να

αυξήσει το μέγιστο αριθμό επαναλήψεων, καθώς ενδέχεται να καθυστερεί η σύγκλιση της μεθόδου, λόγω των πολλών εισόδων.

Σε αυτό το σημείο και πριν αναφερθούμε παρακάτω, θα πρέπει να τονίσουμε ότι για την πραγματοποίηση των εναπομεινάντων σεναρίων, χρησιμοποιήθηκε η κατηγοριοποίηση των ημερών ως προς το φορτίο. Για τη δική μας εφαρμογή, η κατηγοριοποίηση έγινε μέσω του αλγορίθμου των K-μέσων, ενώ οι κατηγορίες – ομάδες ήταν, ενδεικτικά, 15.

Στο τέταρτο σενάριο, καθορίζεται, το δίκτυο να έχει 100 εισόδους. Το βέλτιστο πλήθος νευρώνων για αυτό το σενάριο, βρέθηκε ίσο με 70 και μάλιστα οποιεσδήποτε αυξομειώσεις των νευρώνων, από το σημείο αυτό, προκαλούν σημαντικές αυξήσεις του συνολικού σφάλματος. Ένα πρόβλημα που πιθανώς εμφανίζεται στο σενάριο 4, σχετίζεται με το γεγονός ότι η ίδια ομάδα, από χρόνο σε χρόνο, μεταβάλλει ελαφρώς τη σύνθεση της, δηλαδή δεν περιέχει ακριβώς τις ίδιες ημέρες. Μάλιστα, με την πάροδο πολλών ετών, η σύνθεση της ίδιας ομάδας αλλάζει σε σημαντικό βαθμό. Συνεπώς, θα πρέπει ο συνολικός χρόνος εκπαίδευσης του δικτύου, να μην περιλαμβάνει μεγάλο αριθμό ετών – όπως και κάναμε στις εκτελέσεις μας. Όσον αφορά ενδεχόμενη βελτίωση των σφαλμάτων στο μέλλον, θα μπορούσε κάποιος να εκτελέσει το παρόν σενάριο, για δεδομένα έτη εκπαίδευσης, ξεχωριστά ανά εποχές. Με αυτόν τον τρόπο κάθε φορά που εντοπίζουμε για μια ημέρα, την προηγούμενη ημερολογιακά που να ανήκει στην ίδια ομάδα, βρίσκουμε πάντα μια ημέρα που να ανήκει στην ίδια εποχή και άρα έχει σίγουρα παρόμοια συμπεριφορά.

Το σενάριο 5 από την άλλη, το οποίο προβλέπει δίκτυο 66 εισόδων, συμπεριλαμβάνει στο σύνολο των διανυσμάτων εκπαίδευσης, τις ημέρες εκείνες που ανήκουν *μόνο σε μια συγκεκριμένη ομάδα*, την οποία και έχουμε επιλέξει εξ' αρχής. Στην ουσία λοιπόν, επιλέγουμε να εκπαιδεύσουμε το δίκτυο *ανά ομάδα*. Η σκέψη που οδήγησε στην υλοποίηση αυτού του σεναρίου, έχει να κάνει με την παρόμοια συμπεριφορά των ημερών που ανήκουν στην ίδια ομάδα και συνεπώς, ίσως με αυτόν τον τρόπο καθίσταται πιο αποτελεσματική η εκπαίδευση του T.N.Δ. Το πρόβλημα που προκύπτει εδώ, αφενός σχετίζεται με όσα προαναφέραμε για το σενάριο 4, αφετέρου έχει να κάνει με τον μικρό αριθμό των στοιχείων που περιλαμβάνει κάθε ομάδα, ανά έτος, με αποτέλεσμα ο συνολικός αριθμός των διανυσμάτων, για δεδομένα έτη εκπαίδευσης, να μην είναι τόσο ικανός για την εκπαίδευση του δικτύου. Συνεπώς εδώ δημιουργείται μια «σύγκρουση» συμφερόντων θα λέγαμε, καθώς από τη μία προτείνεται η αύξηση των ετών εκπαίδευσης, ώστε να αυξηθεί και ο αριθμός των διαθέσιμων διανυσμάτων και από την άλλη προκαλούνται ενδεχόμενα προβλήματα, αφού με το πέρασμα πολλών ετών οι ομάδες μεταβάλλονται σημαντικά. Για βελτίωση των σφαλμάτων του σεναρίου λοιπόν, θα πρέπει να βρεθεί η «χρυσή τομή» για τα έτη της εκπαίδευσης ή να πραγματοποιηθεί το σενάριο ανά εποχές του έτους, όπως προαναφέραμε, άλλωστε.

Επιπλέον, όσον αφορά τις δοκιμές που διεξήχθησαν, διαπιστώσαμε ότι το συνολικό σφάλμα της μεθόδου εξαρτάται άμεσα από την ομάδα που επιλέγουμε και μάλιστα, όσο πιο πολυπληθής είναι μια ομάδα, τόσο μικρότερο σφάλμα δίνει. Για το συγκεκριμένο σενάριο, ο τυχαίος τρόπος παρουσίασης των προτύπων, στις πιο πολλές περιπτώσεις, είναι αρκετά καλύτερος σε σχέση με τον σειριακό τρόπο.

Τελευταίο σενάριο που πραγματοποιήσαμε, είναι το σενάριο 6, το οποίο προβλέπει για το T.N.Δ. 13 εισόδους, δηλαδή 5-8 φορές λιγότερες σε σύγκριση με τα προηγούμενα σενάρια. Η

βασική ιδέα σε αυτό το σενάριο, είναι η εκπαίδευση του δικτύου για την *πρόβλεψη μιας συγκεκριμένης, επιθυμητής, ώρας* της επόμενης ημέρας και *όχι και των 24 ωρών μαζί*. Αν θέλουμε λοιπόν, να προβλέψουμε και τις 24 ώρες της επόμενης μέρας, τότε θα πρέπει να εκτελέσουμε το σενάριο 24 φορές, δηλαδή για τις ώρες 1-24. Τα αποτελέσματα αυτού του σεναρίου είναι από τα καλύτερα που λάβαμε, σε σύγκριση με όλα τα προηγούμενα σενάρια. Επιπρόσθετα αυτό που διαπιστώσαμε, είναι η μεγάλη επίδραση των ετών εκπαίδευσης, αλλά και της επιλογής της ώρας, στα τελικά σφάλματα. Η «αντιπαράθεση» εδώ, των δύο τρόπων παρουσίασης προτύπων δε βγάζει κάποιο ουσιαστικό αποτέλεσμα, καθώς οι μεταβολές των σφαλμάτων για την πλειοψηφία των ωρών είναι ελάχιστες.

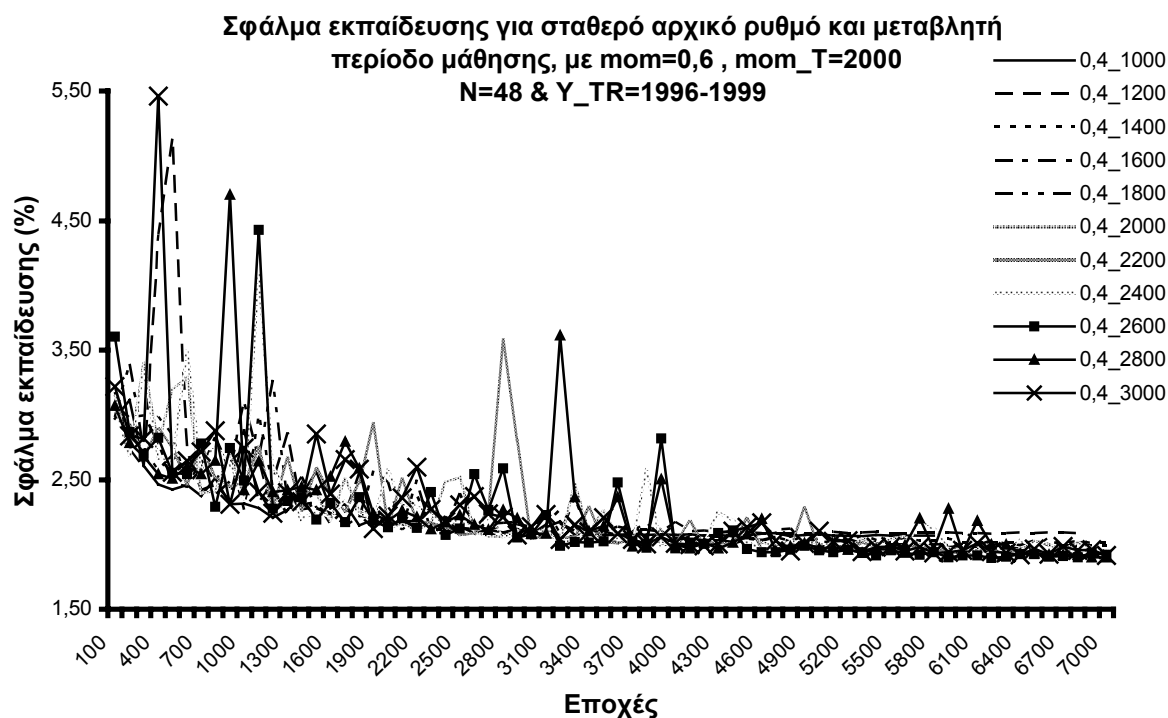
Όσον αφορά, τις προτάσεις μας για ακόμα μικρότερα σφάλματα, για αυτή την περίπτωση, θα μπορούσε κάποιος να αυξήσει, σχετικά τις λίγες εισόδους του δικτύου, καθώς τότε ενδέχεται να είναι πιο αποτελεσματική η εκπαίδευση αυτού.

Συμπερασματικά, δεν μπορέσαμε να θεσπίσουμε κάποιο γενικό κανόνα, σχετικά με την εφαρμογή του αλγορίθμου της ανάστροφης διάδοσης σφάλματος στην βραχυπρόθεσμη πρόβλεψη φορτίου. Αυτό που μπορούμε να διατυπώσουμε, είναι συμπεράσματα για επιμέρους περιπτώσεις, όπως και κάναμε πρωτύτερα, καθώς γενικά υπάρχει μια άμεση και πολύπλοκη αλληλεξάρτηση, μεταξύ όλων των «μεγεθών», όπως είδος σεναρίου εκπαίδευσης, πλήθος νευρώνων κρυμμένου επιπέδου, τρόπος παρουσίασης των διανυσμάτων εκπαίδευσης, καθώς και ρυθμός μάθησης και όρος ορμής.

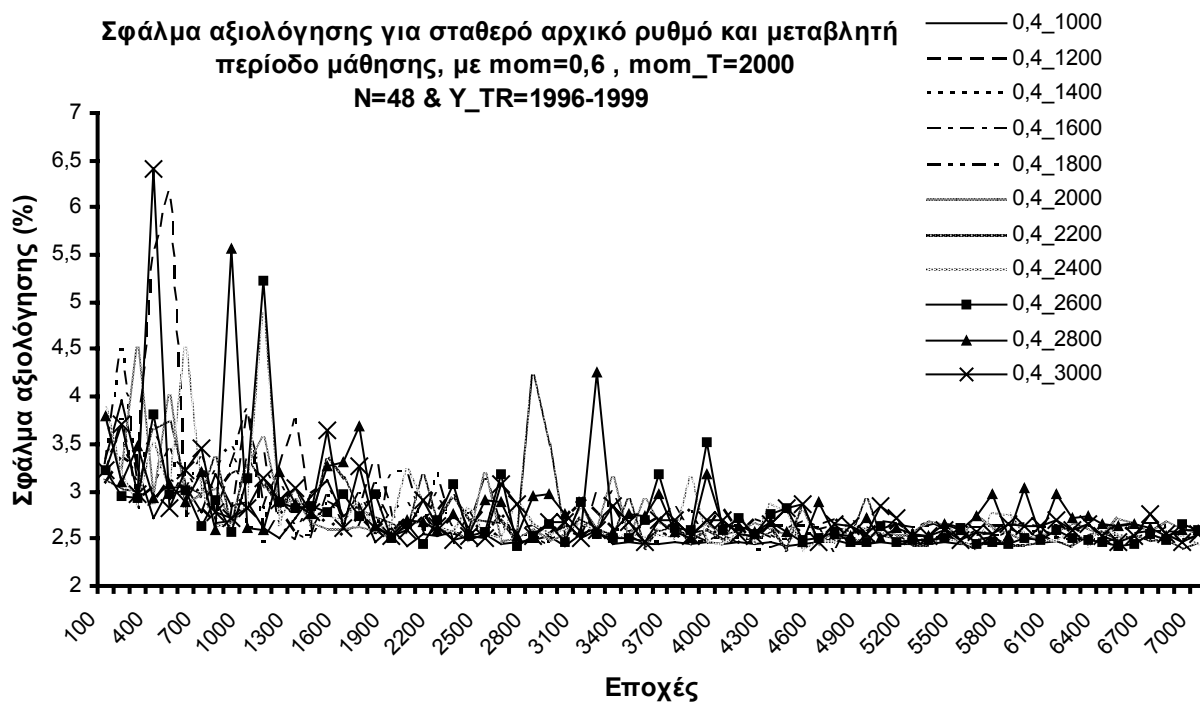
Παράρτημα Ι

Γραφήματα

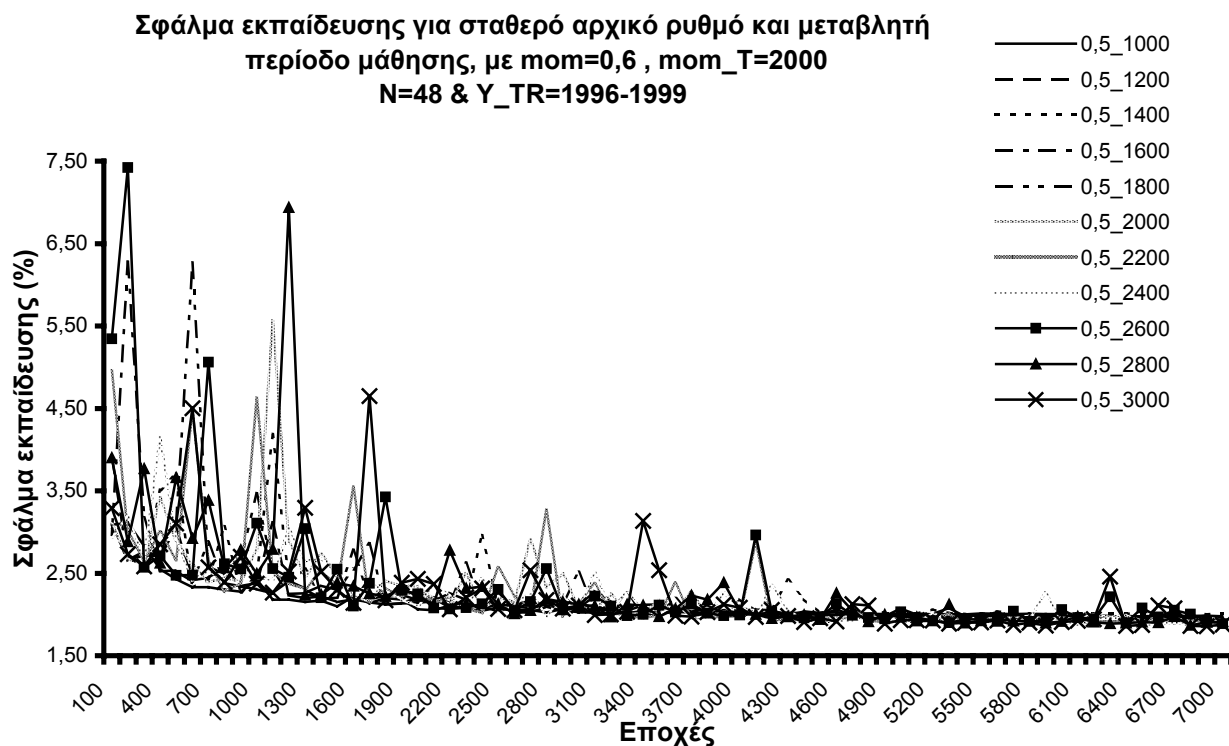
➤ ΠΙ.1 Γραφήματα εκτελέσεων του προγράμματος T.N.Δ. για μεταβολές των παραμέτρων μάθησης και ορμής, με τυχαίο τρόπο παρουσίασης προτύπων



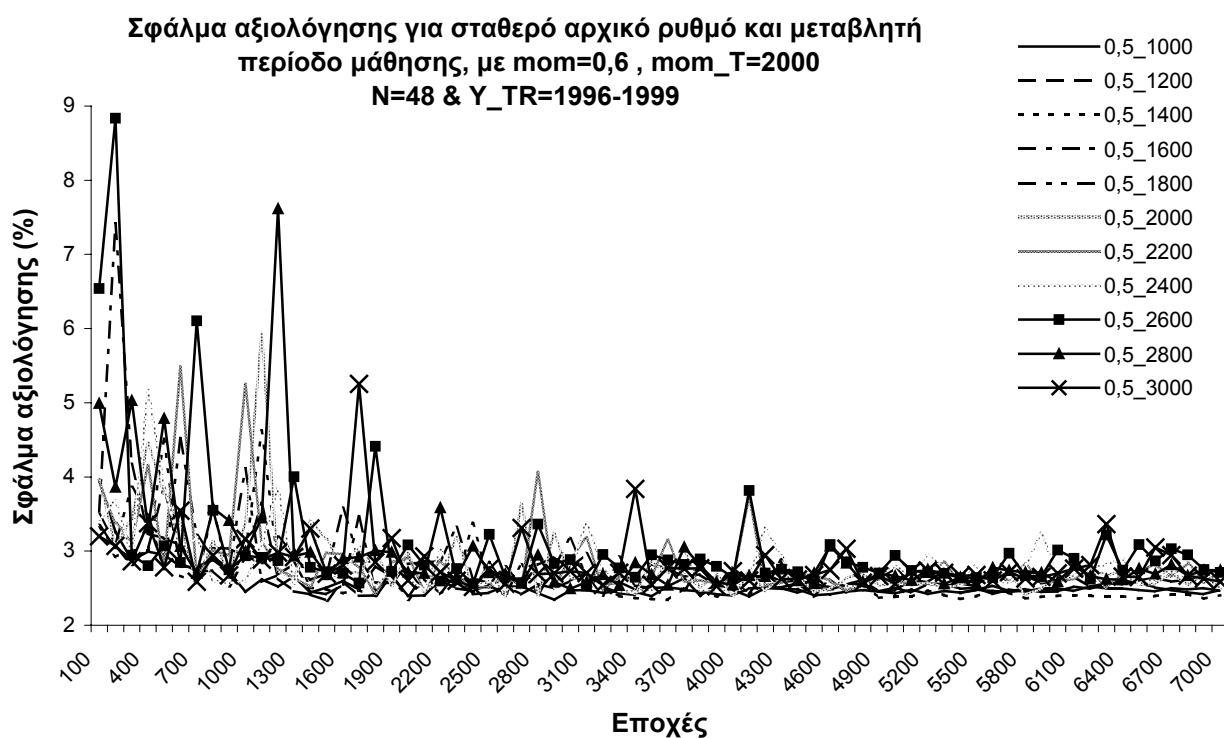
Παράρτημα Ι – Γράφημα 1



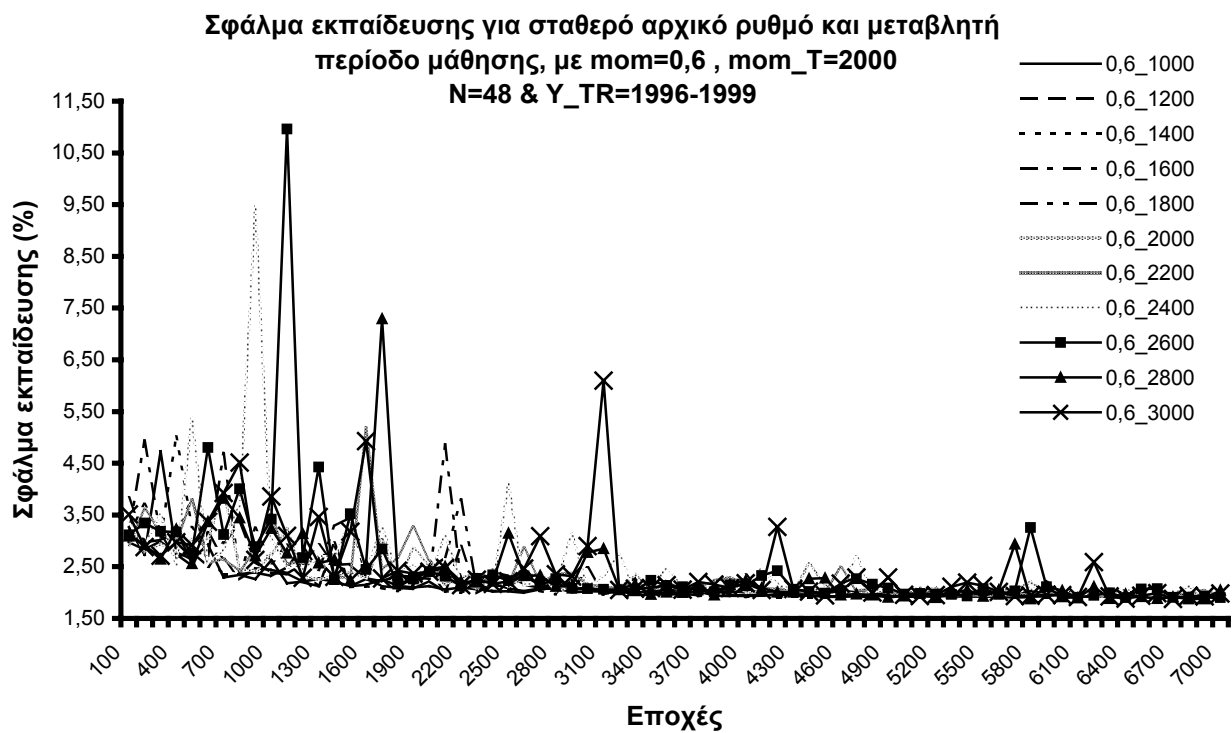
Παράρτημα Ι – Γράφημα 2



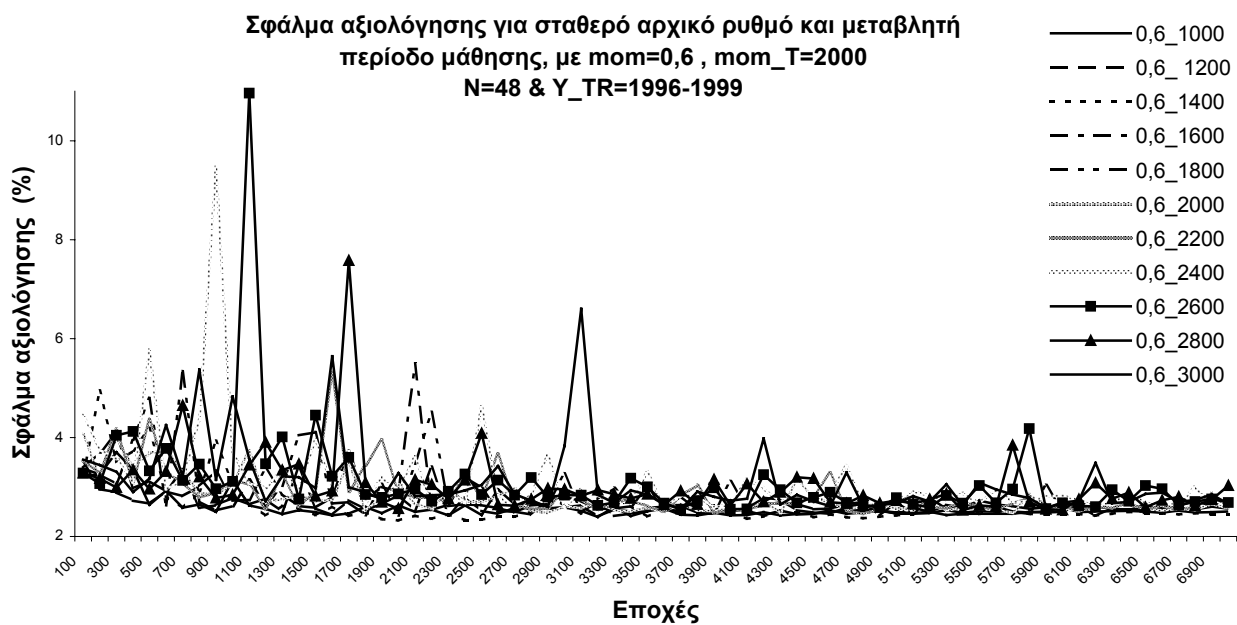
Παράρτημα Ι – Γράφημα 3



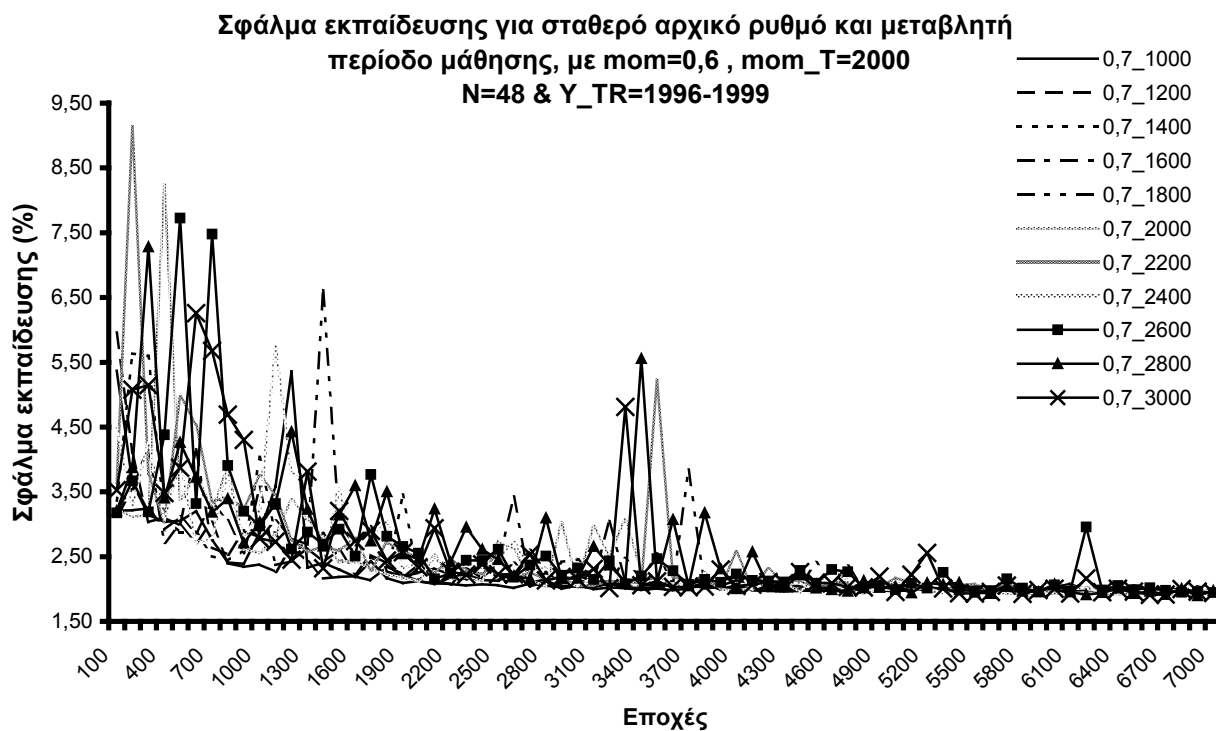
Παράρτημα Ι – Γράφημα 4



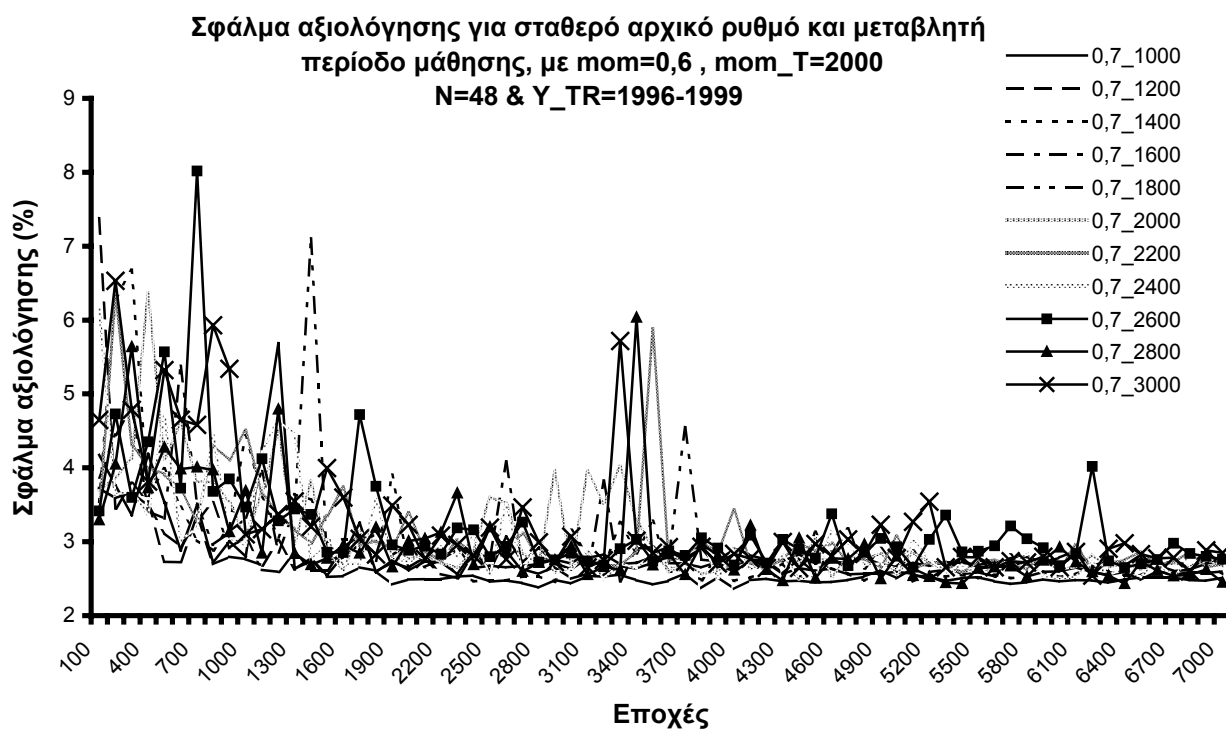
Παράρτημα Ι – Γράφημα 5



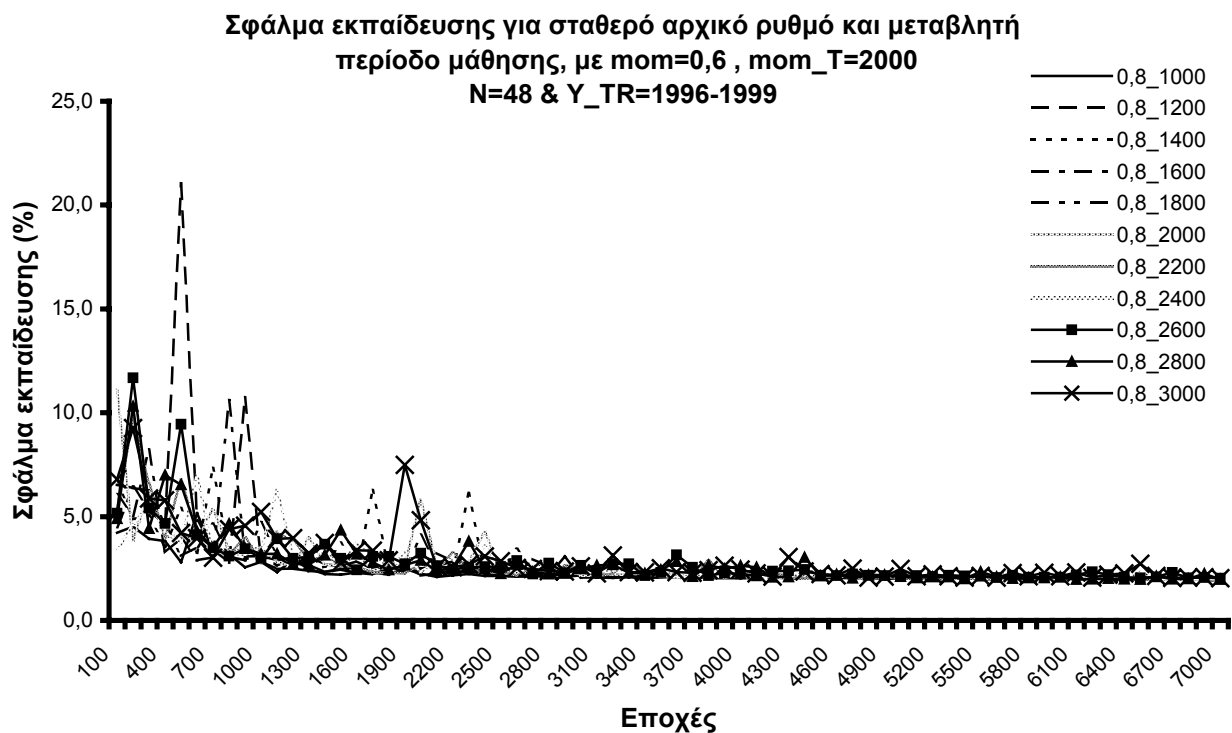
Παράρτημα Ι – Γράφημα 6



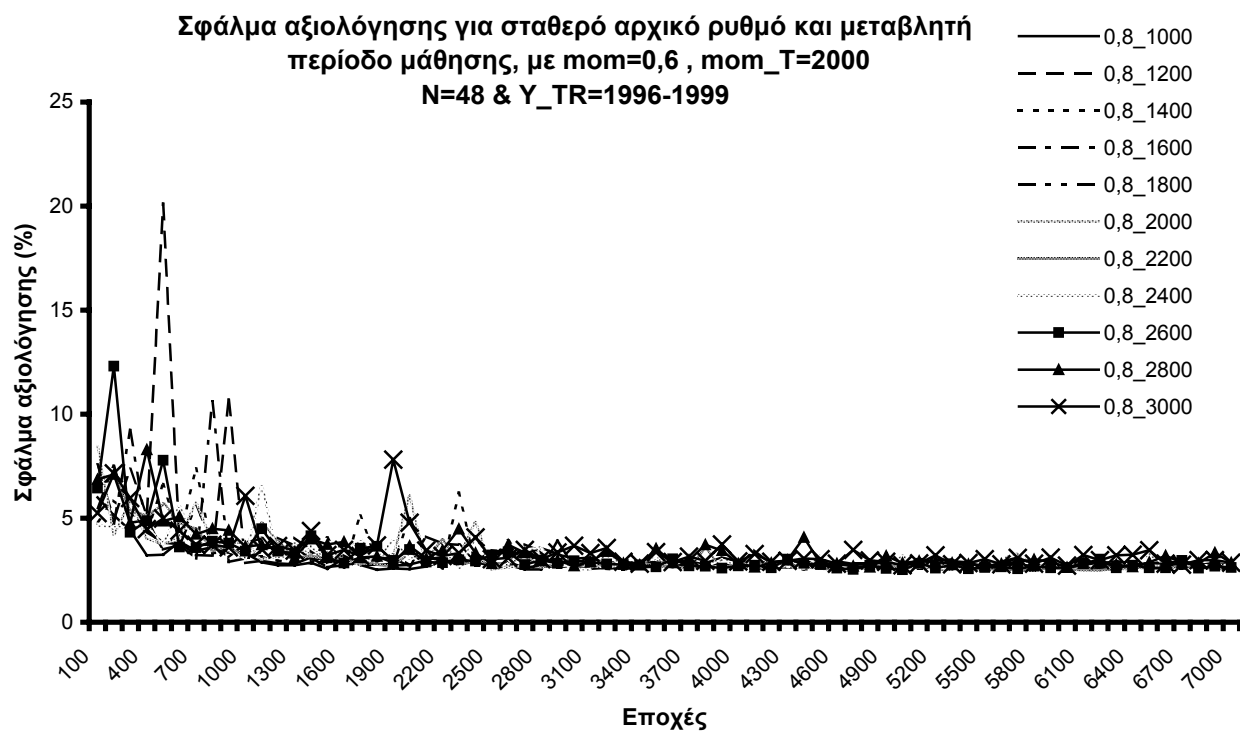
Παράρτημα Ι – Γράφημα 7



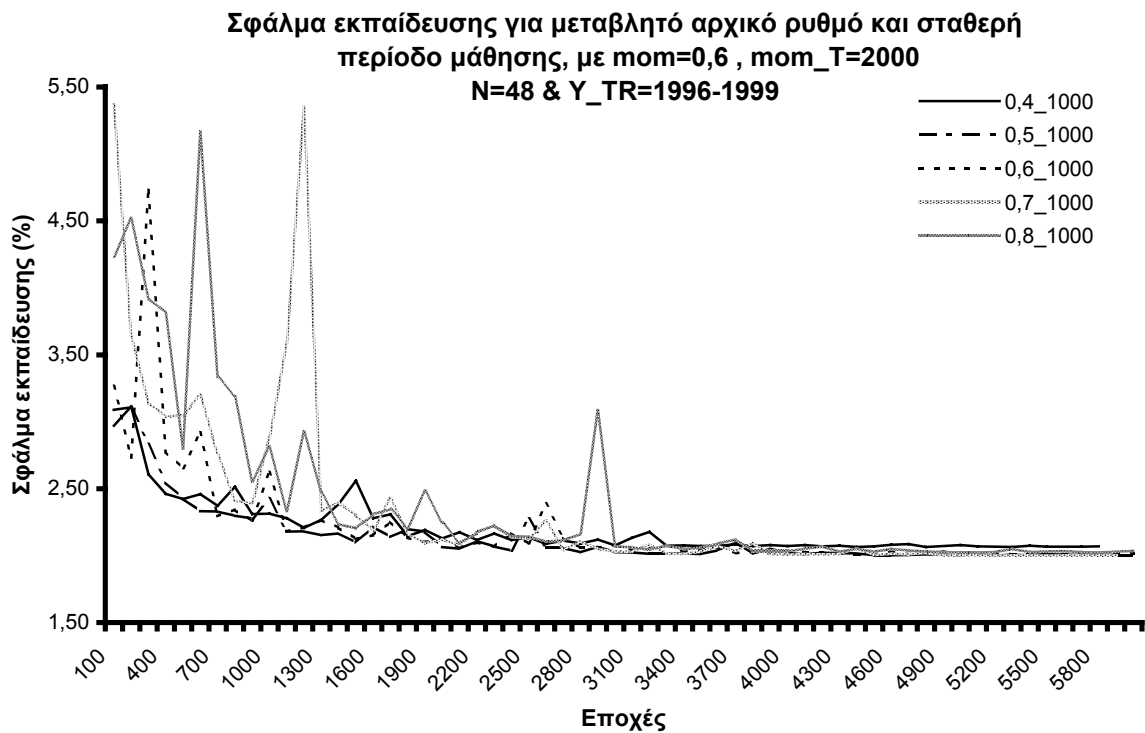
Παράρτημα Ι – Γράφημα 8



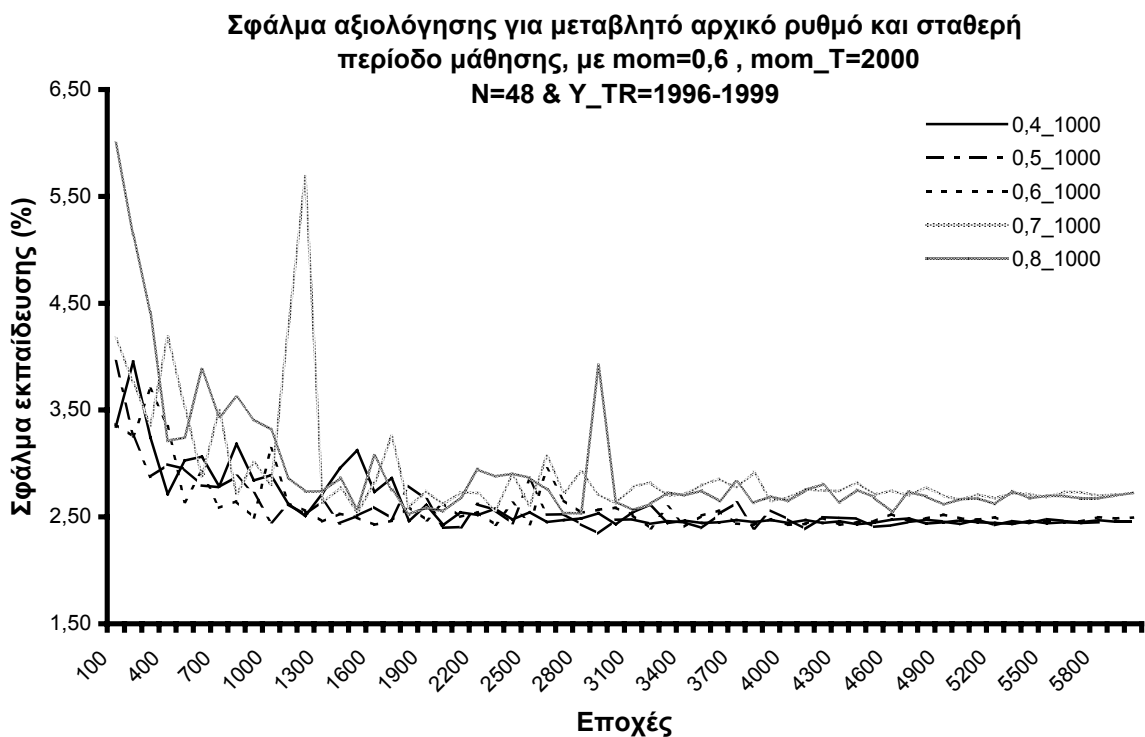
Παράρτημα Ι – Γράφημα 9



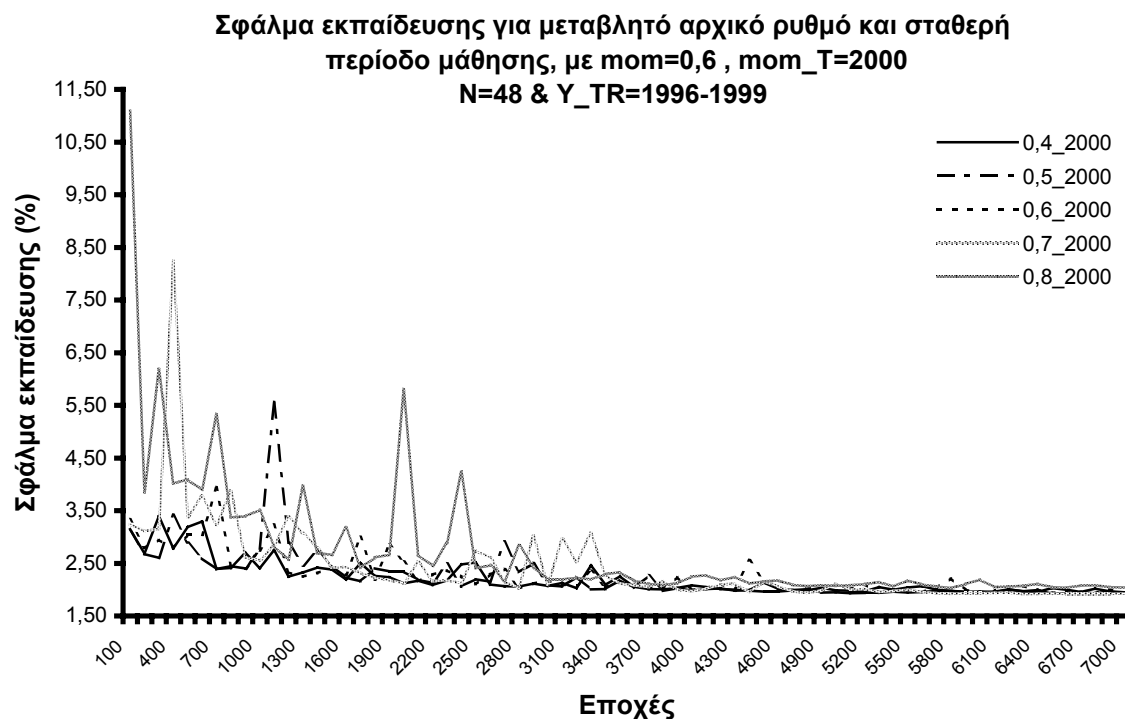
Παράρτημα Ι – Γράφημα 10



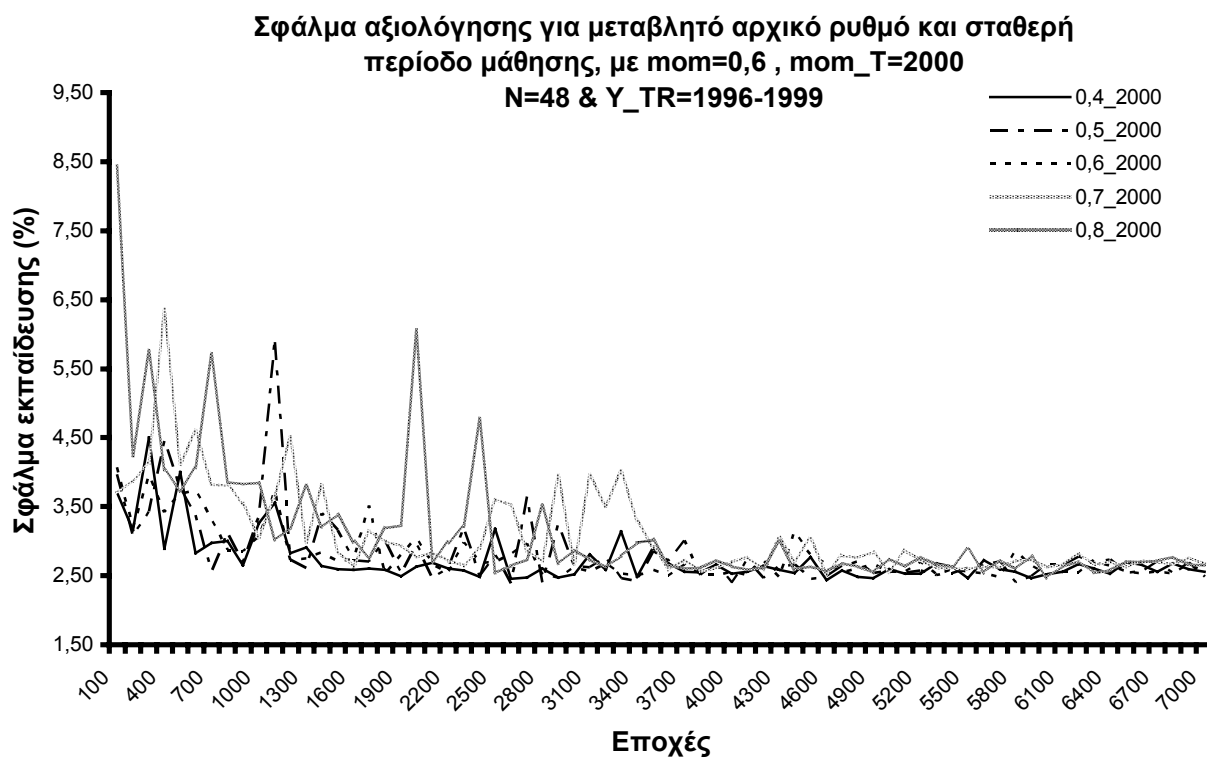
Παράρτημα Ι – Γράφημα 11



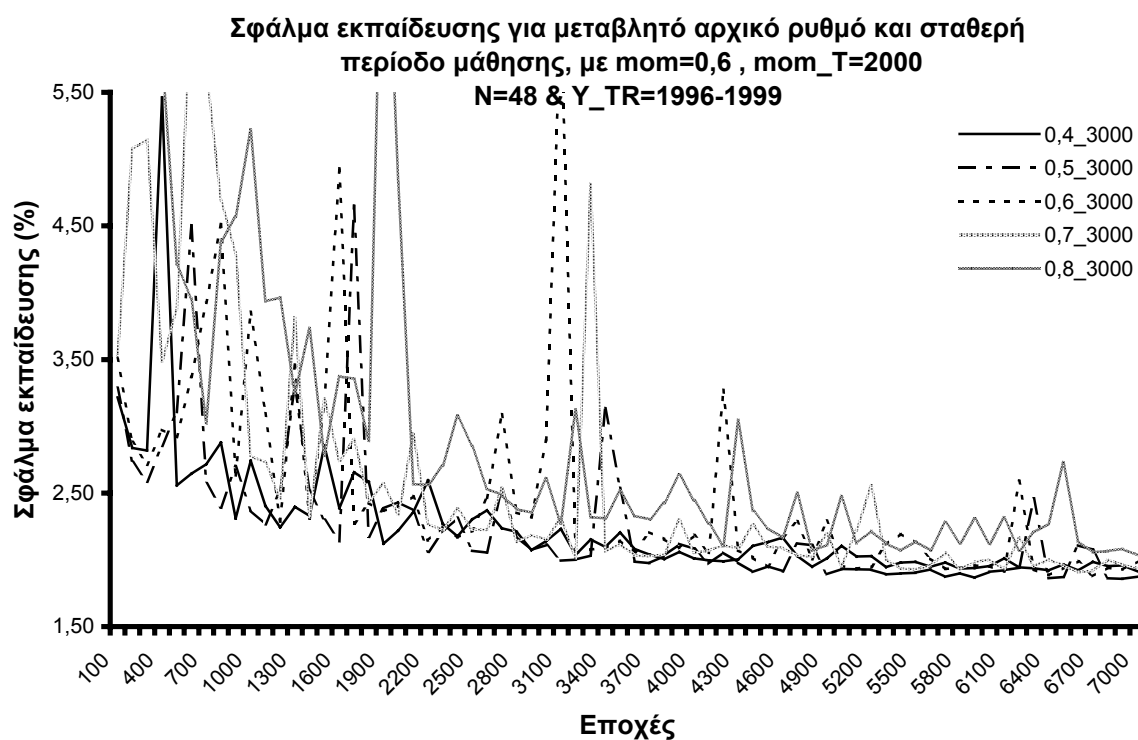
Παράρτημα Ι – Γράφημα 12



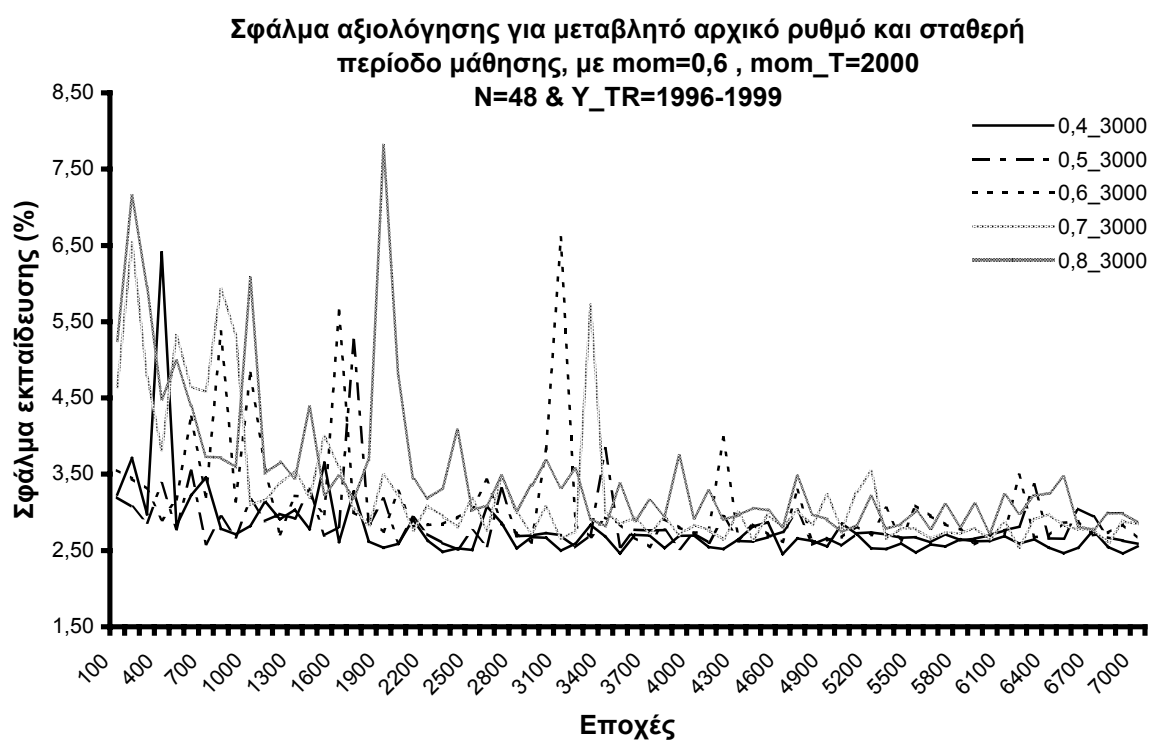
Παράρτημα Ι – Γράφημα 13



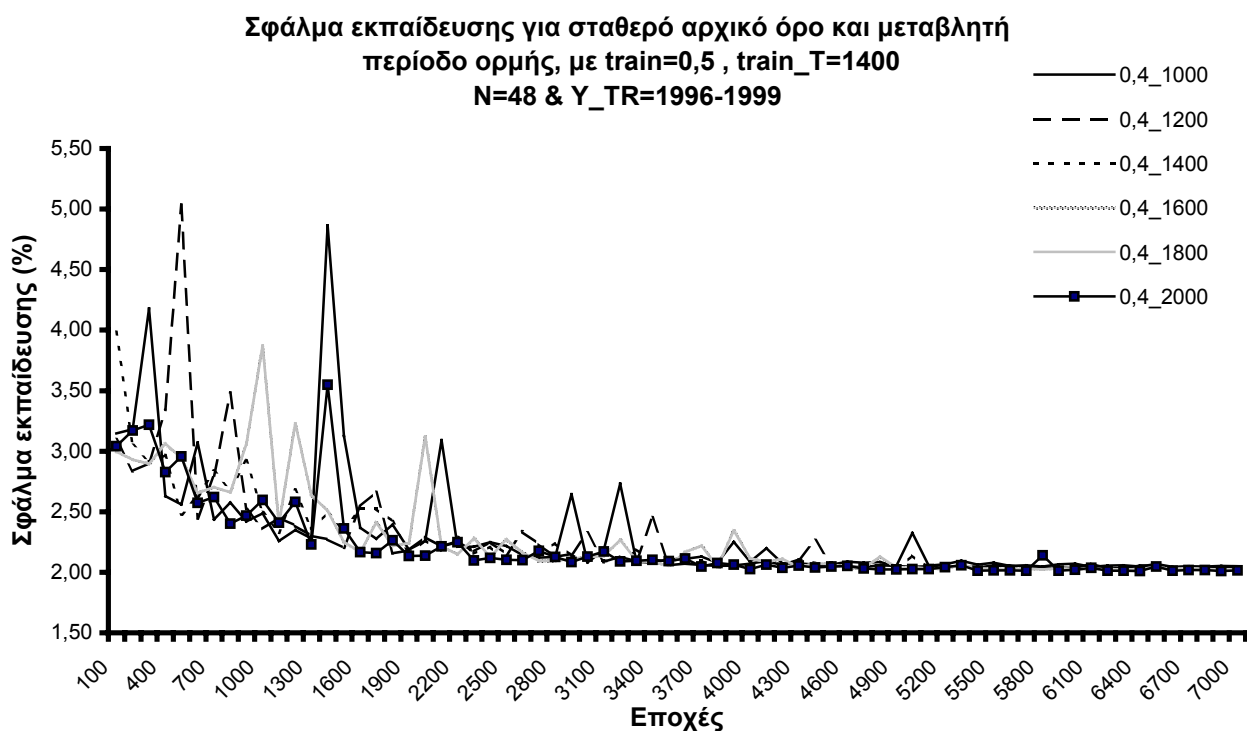
Παράρτημα Ι – Γράφημα 14



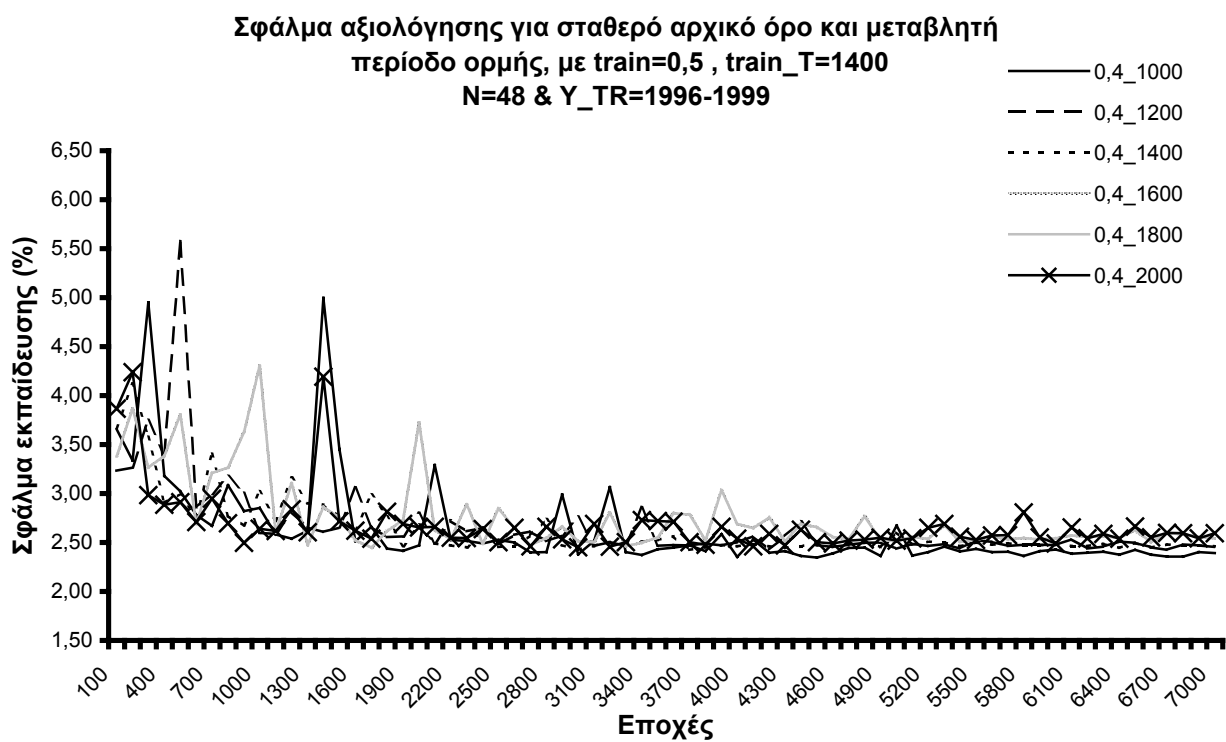
Παράρτημα Ι – Γράφημα 15



Παράρτημα Ι – Γράφημα 16

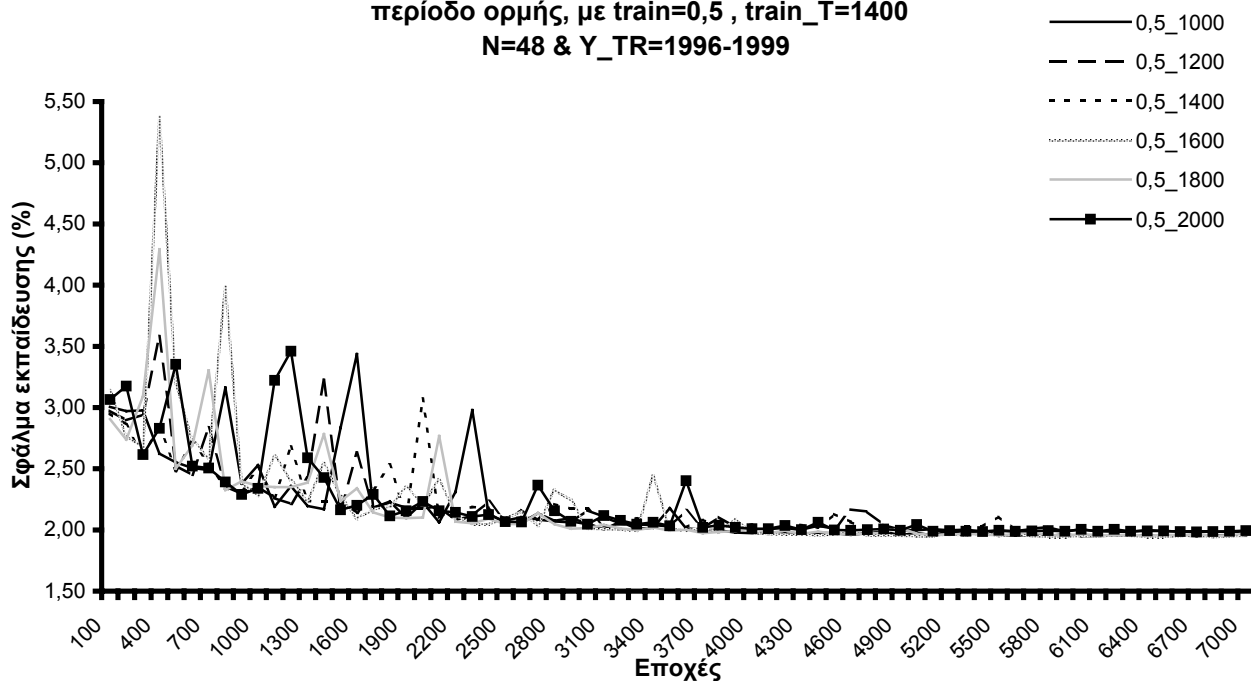


Παράρτημα Ι – Γράφημα 17



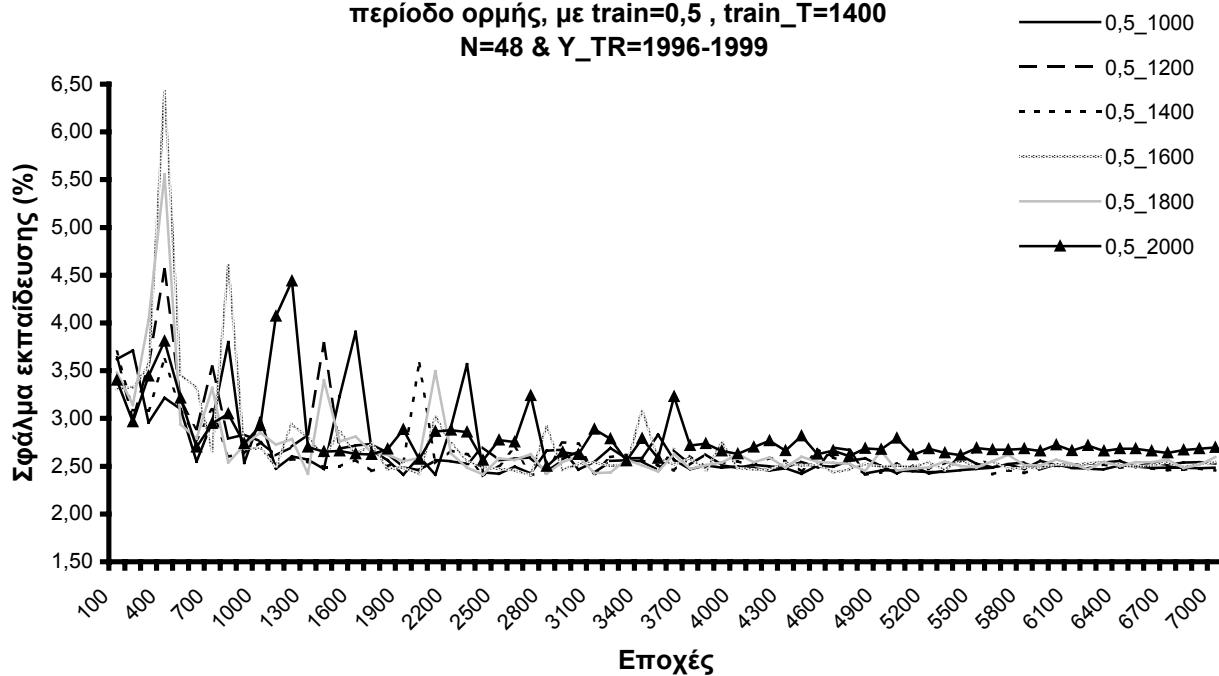
Παράρτημα Ι – Γράφημα 18

Σφάλμα εκπαίδευσης για σταθερό αρχικό όρο και μεταβλητή
περίοδο ορμής, με $\text{train}=0,5$, $\text{train_T}=1400$
 $N=48$ & $Y_TR=1996-1999$



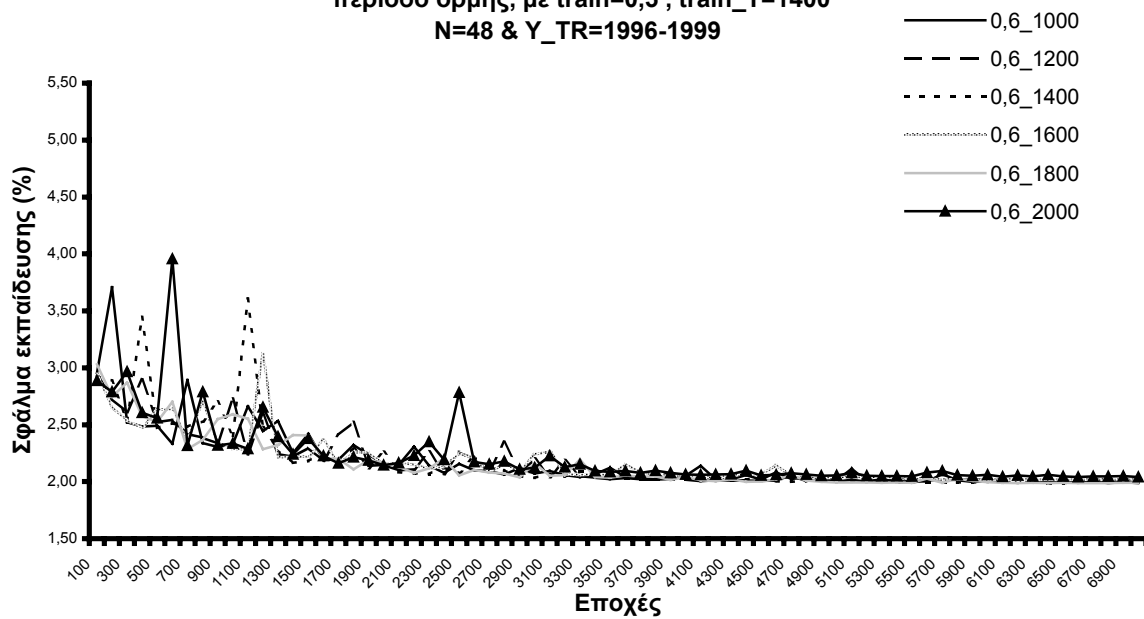
Παράρτημα Ι – Γράφημα 19

Σφάλμα αξιολόγησης για σταθερό αρχικό όρο και μεταβλητή
περίοδο ορμής, με $\text{train}=0,5$, $\text{train_T}=1400$
 $N=48$ & $Y_TR=1996-1999$



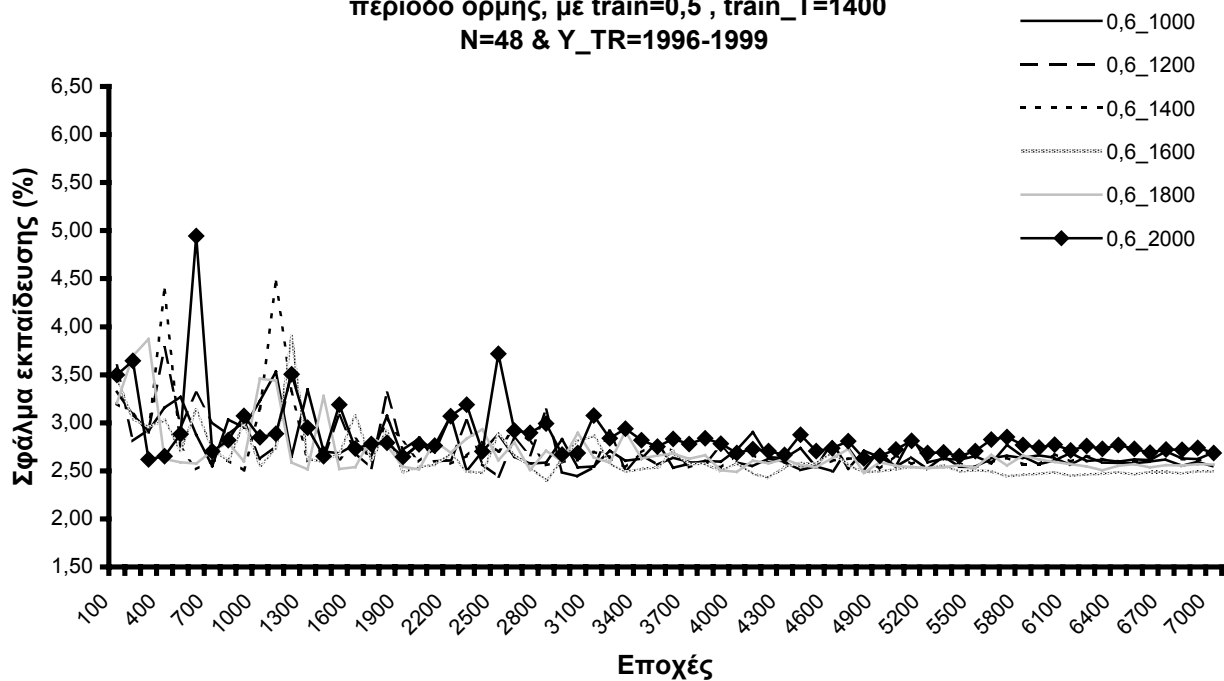
Παράρτημα Ι – Γράφημα 20

Σφάλμα εκπαίδευσης για σταθερό αρχικό όρο και μεταβλητή
περίοδο ορμής, με $\text{train}=0,5$, $\text{train_T}=1400$
 $N=48$ & $Y_TR=1996-1999$

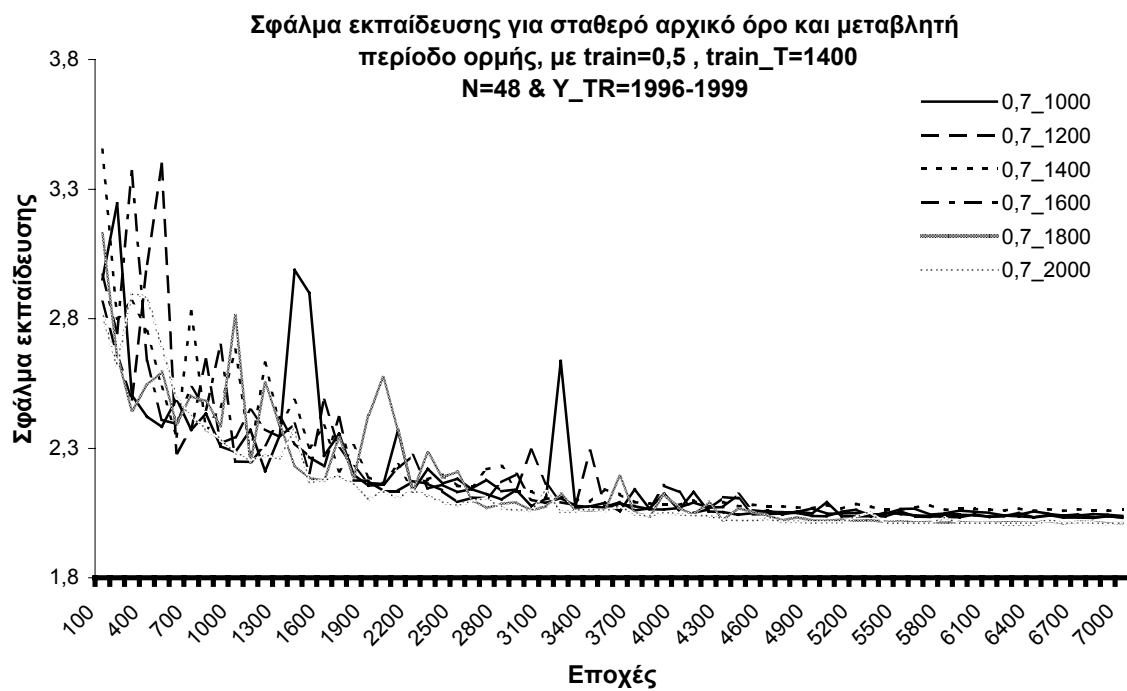


Παράρτημα Ι – Γράφημα 21

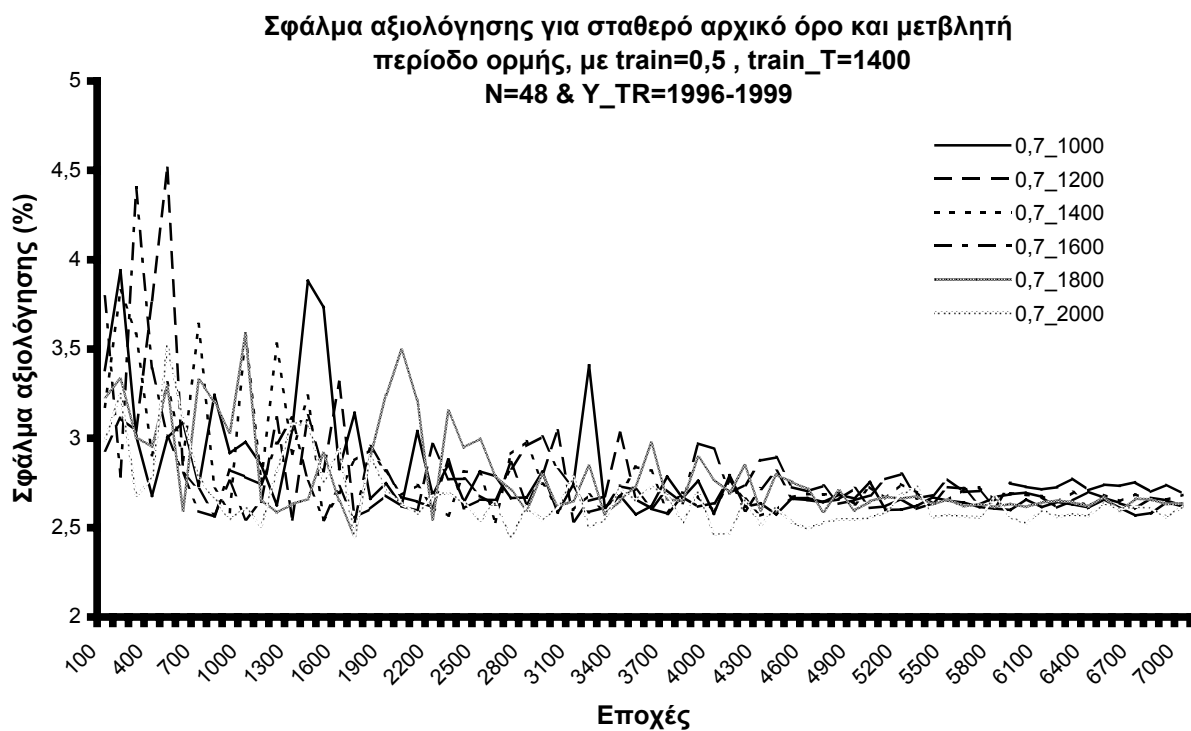
Σφάλμα αξιολόγησης για σταθερό αρχικό όρο και μεταβλητή
περίοδο ορμής, με $\text{train}=0,5$, $\text{train_T}=1400$
 $N=48$ & $Y_TR=1996-1999$



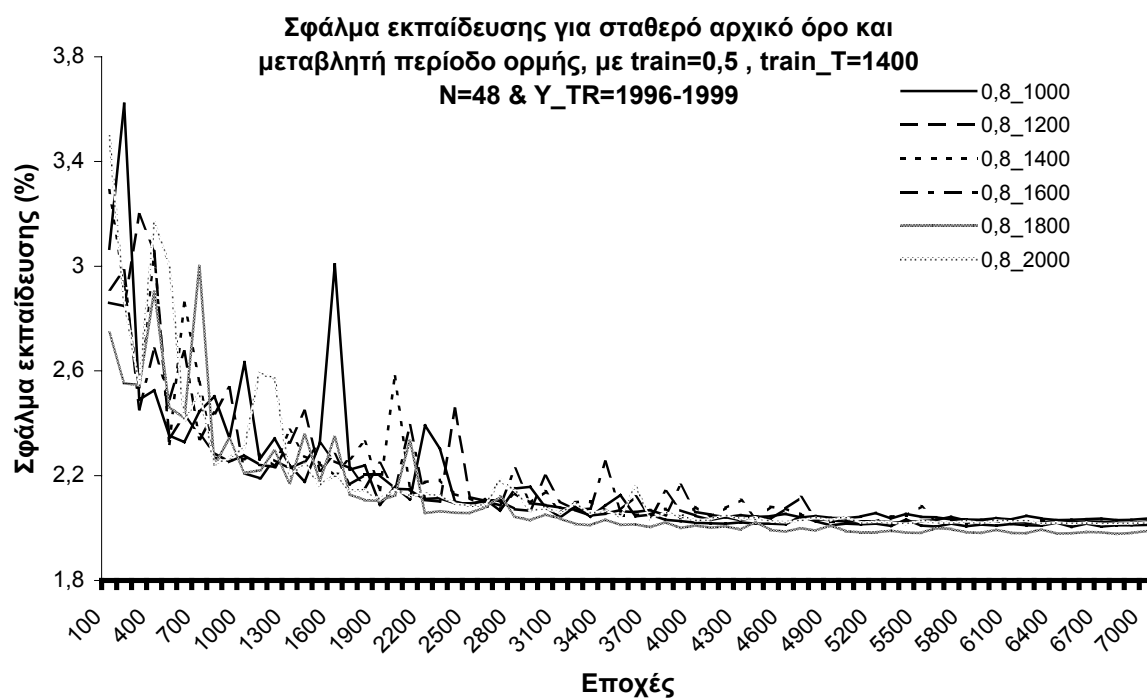
Παράρτημα Ι – Γράφημα 22



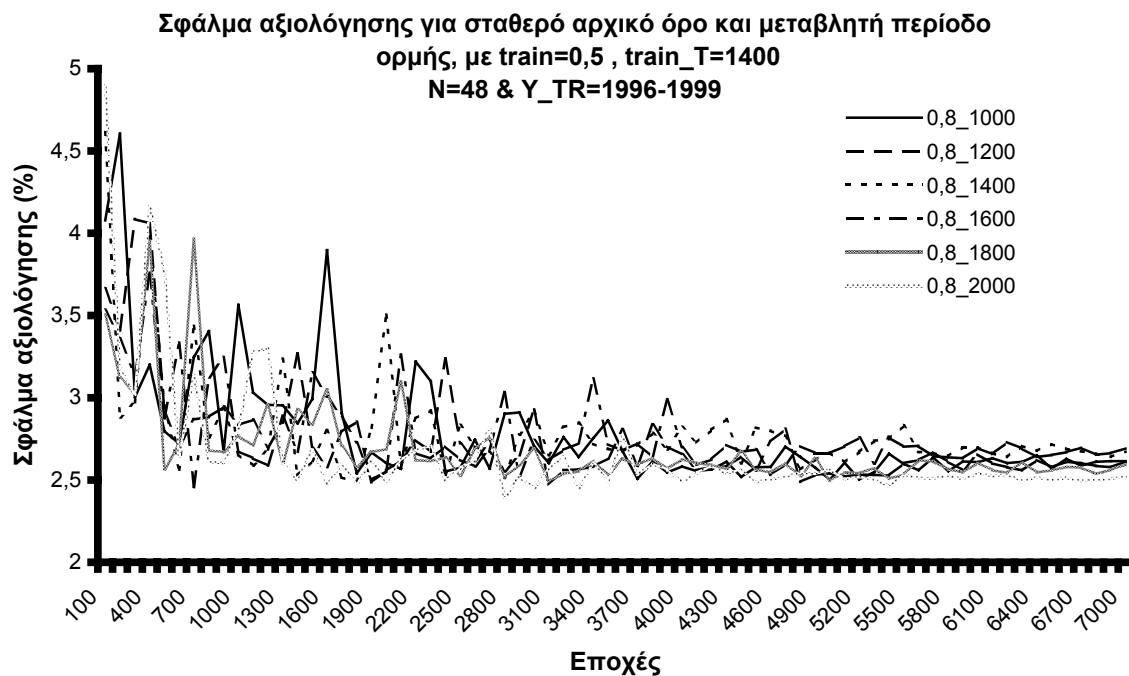
Παράρτημα Ι – Γράφημα 23



Παράρτημα Ι – Γράφημα 24

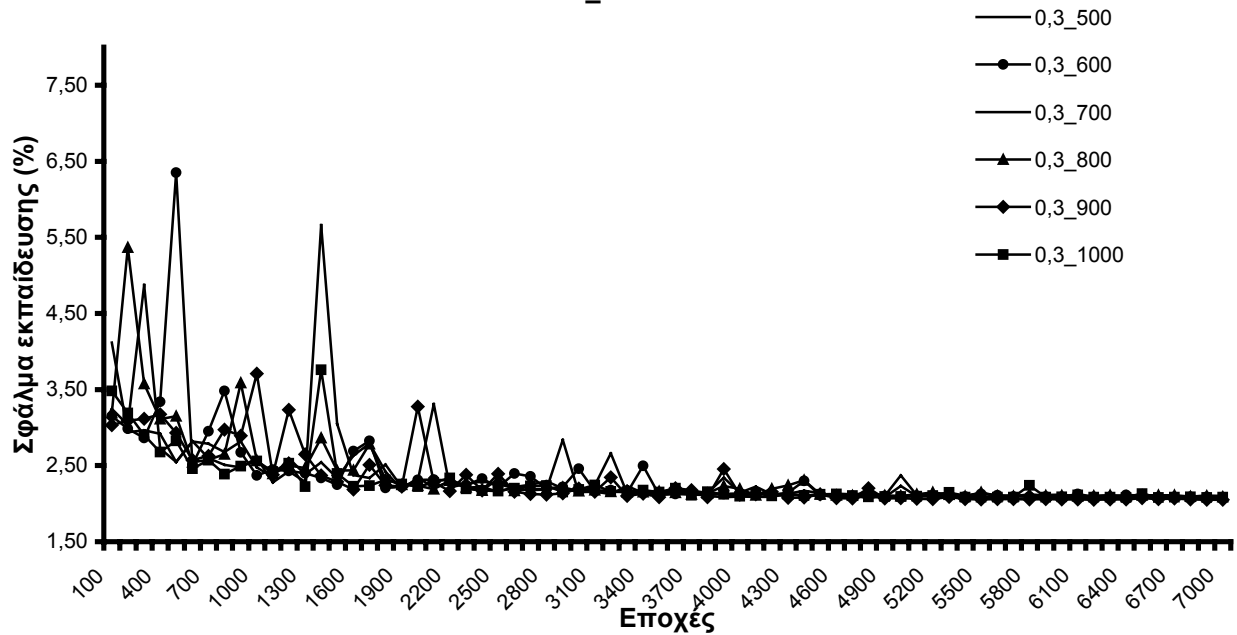


Παράρτημα Ι – Γράφημα 25



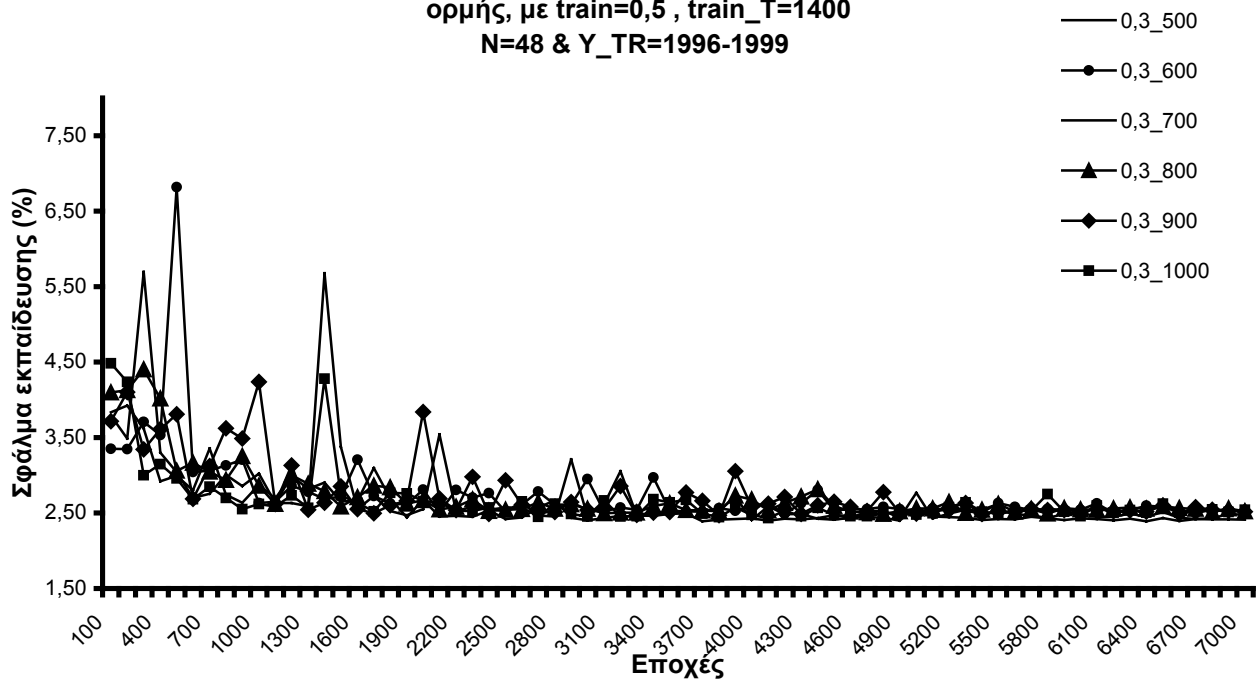
Παράρτημα Ι – Γράφημα 26

Σφάλμα εκπαίδευσης για μεταβλητό αρχικό όρο και σταθερή
περίοδο ορμής, με $\text{train}=0,5$, $\text{train_T}=1400$
 $N=48$ & $Y_TR=1996-1999$

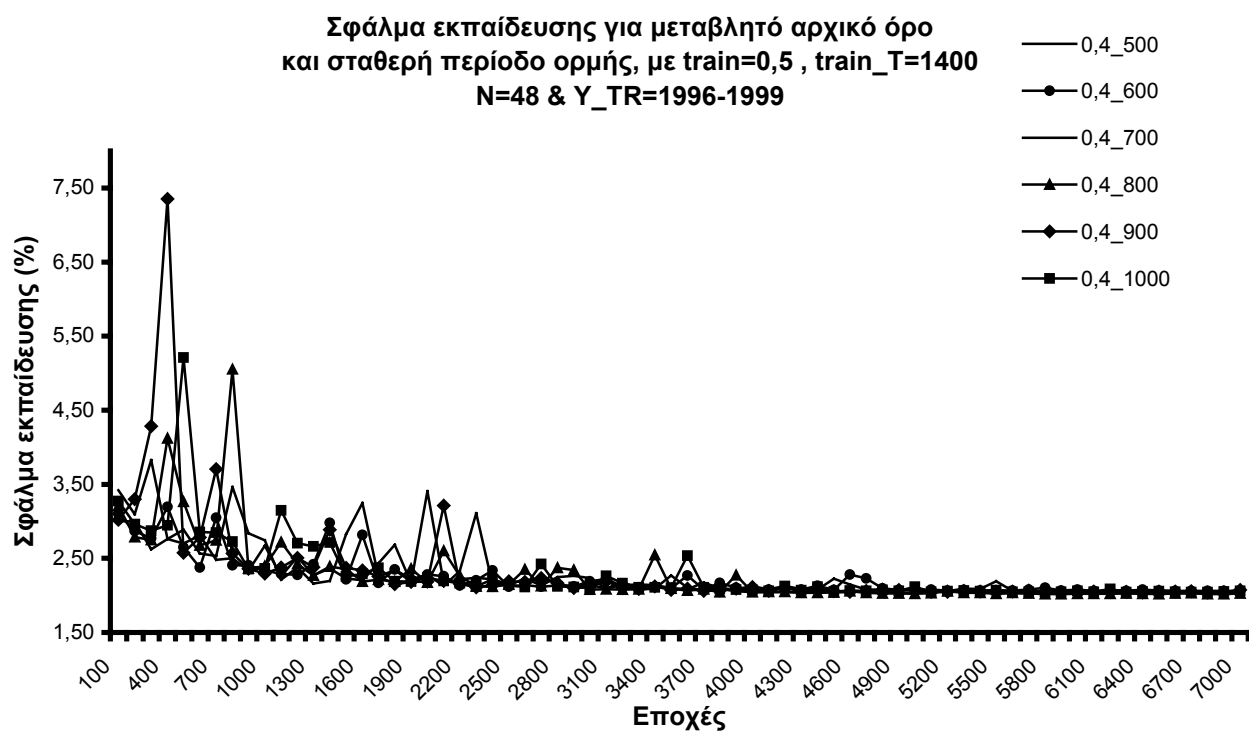


Παράρτημα Ι – Γράφημα 27

Σφάλμα αξιολόγησης για μεταβλητό αρχικό όρο και σταθερή περίοδο
ορμής, με $\text{train}=0,5$, $\text{train_T}=1400$
 $N=48$ & $Y_TR=1996-1999$



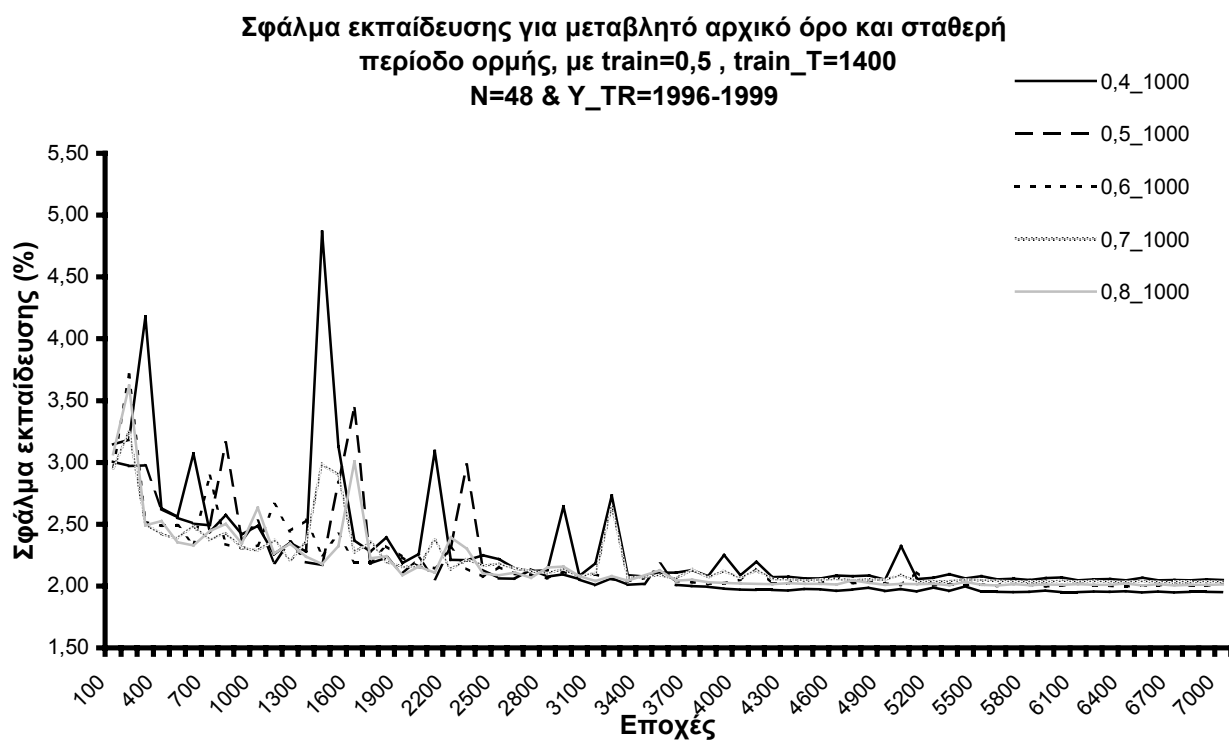
Παράρτημα Ι – Γράφημα 28



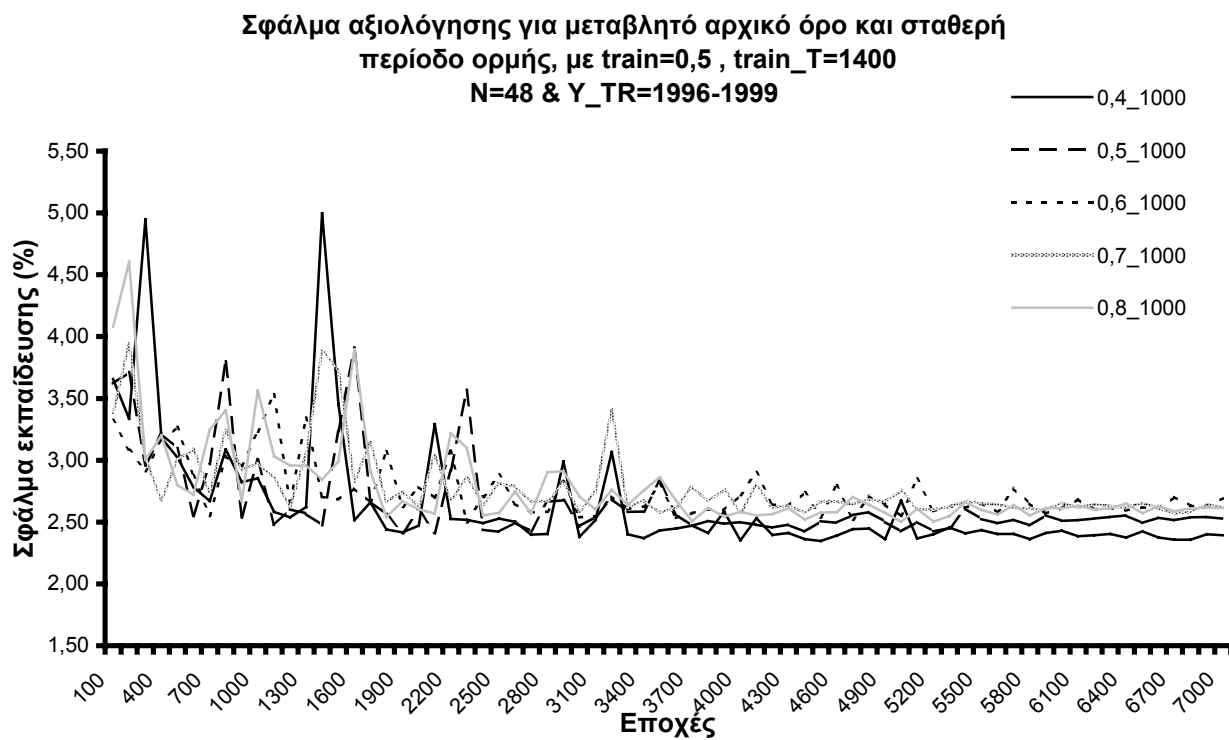
Παράρτημα Ι – Γράφημα 29



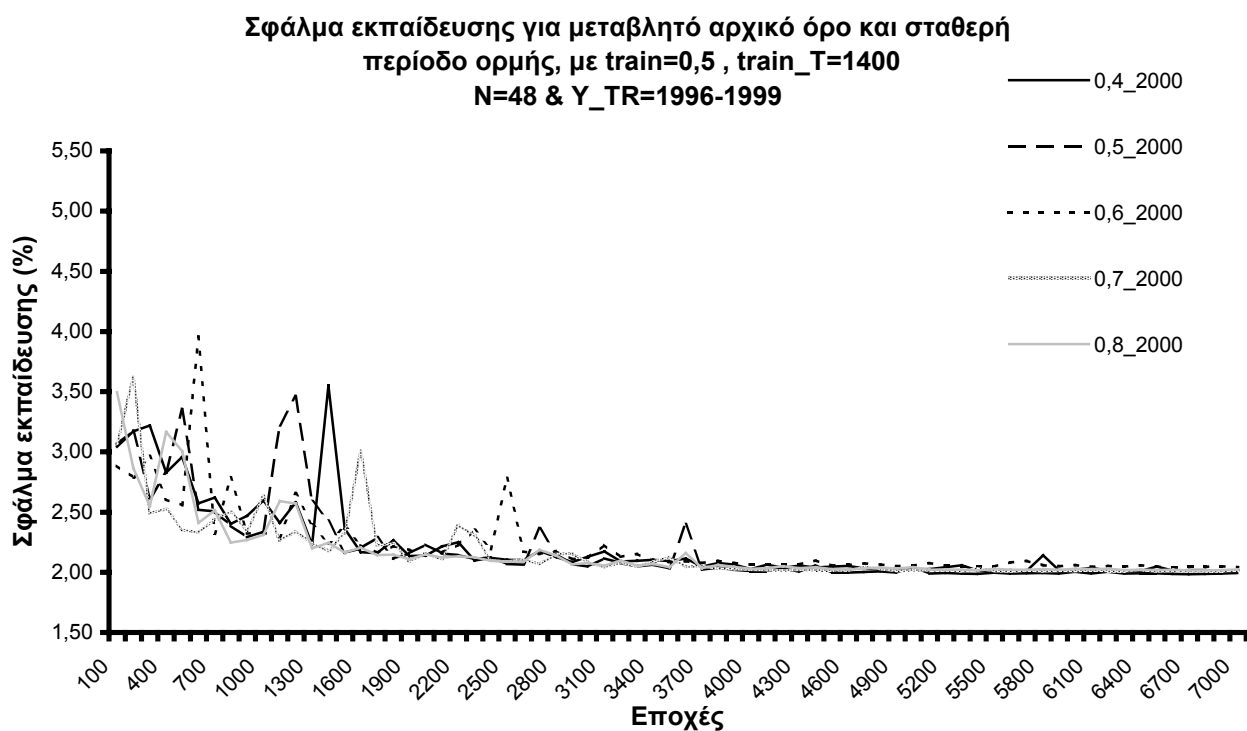
Παράρτημα Ι – Γράφημα 30



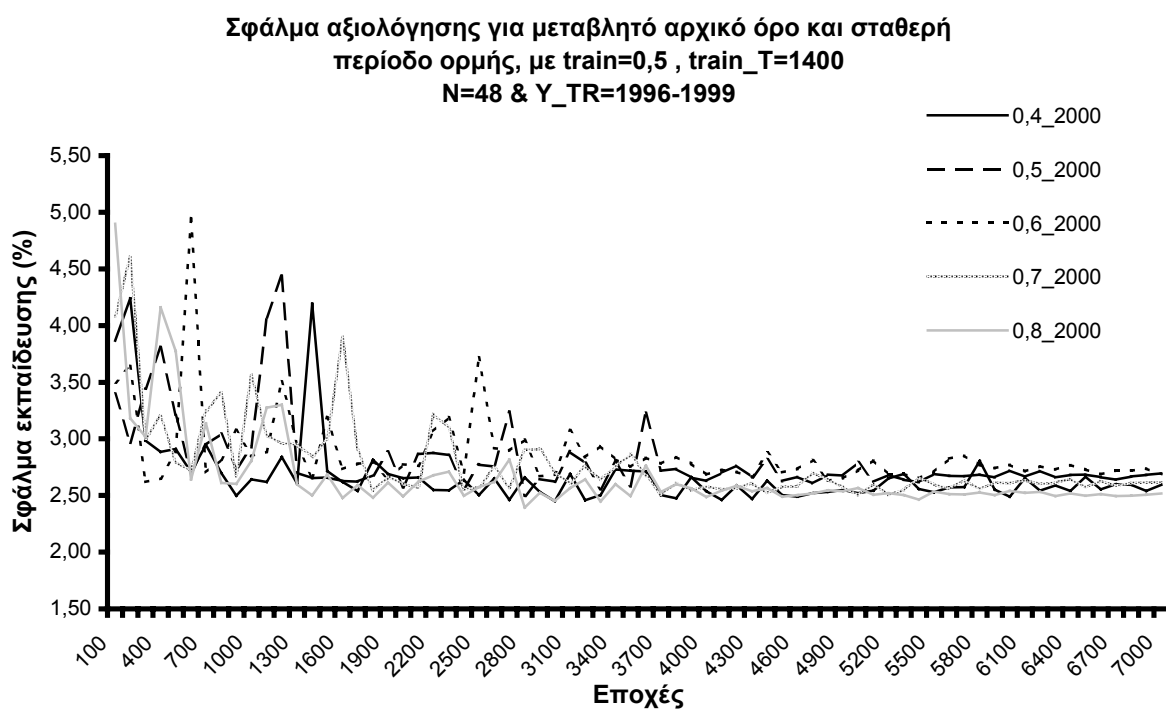
Παράρτημα Ι – Γράφημα 31



Παράρτημα Ι – Γράφημα 32

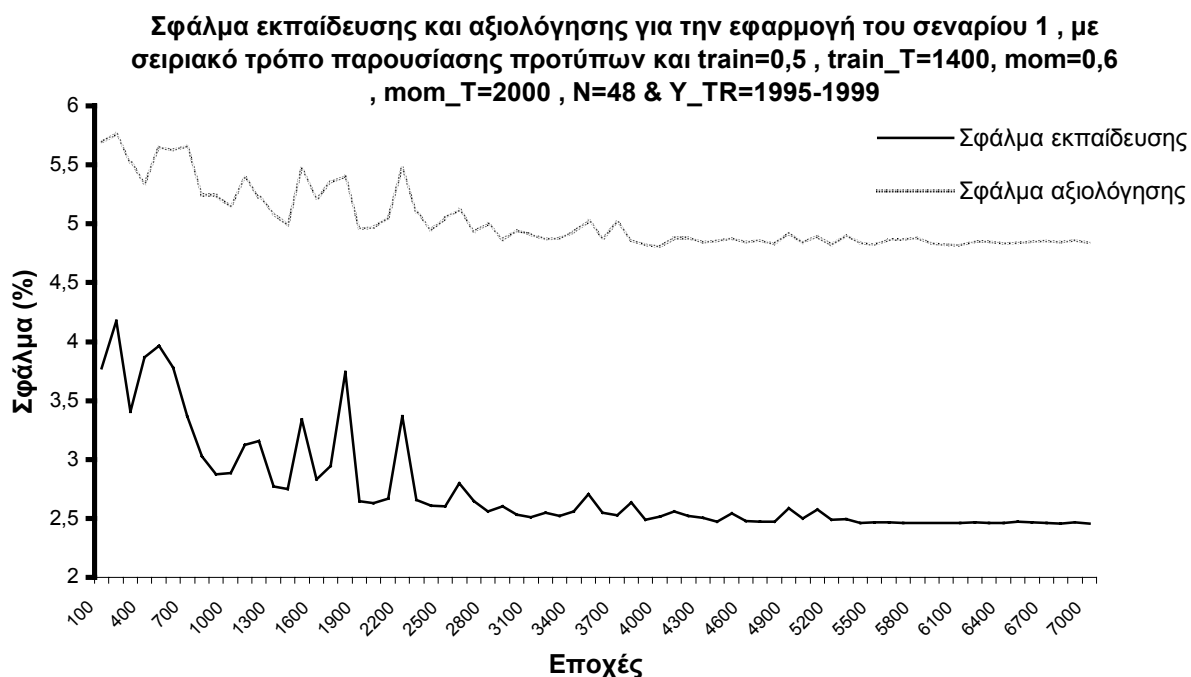


Παράρτημα Ι – Γράφημα 33

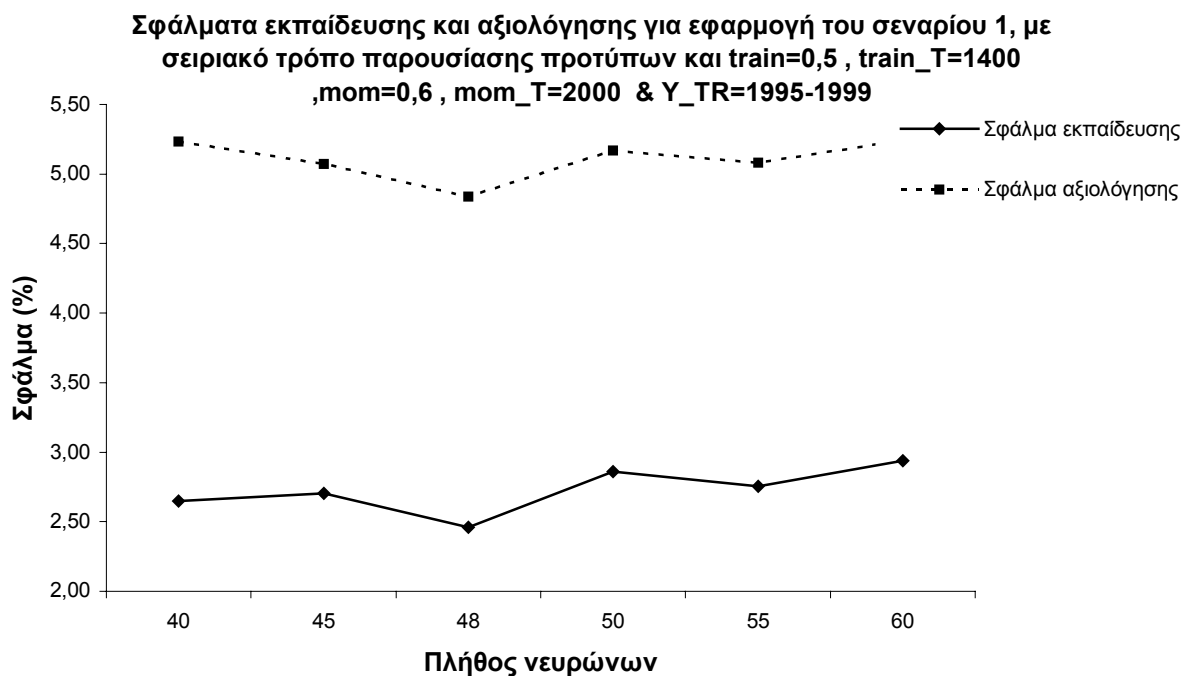


Παράρτημα Ι – Γράφημα 34

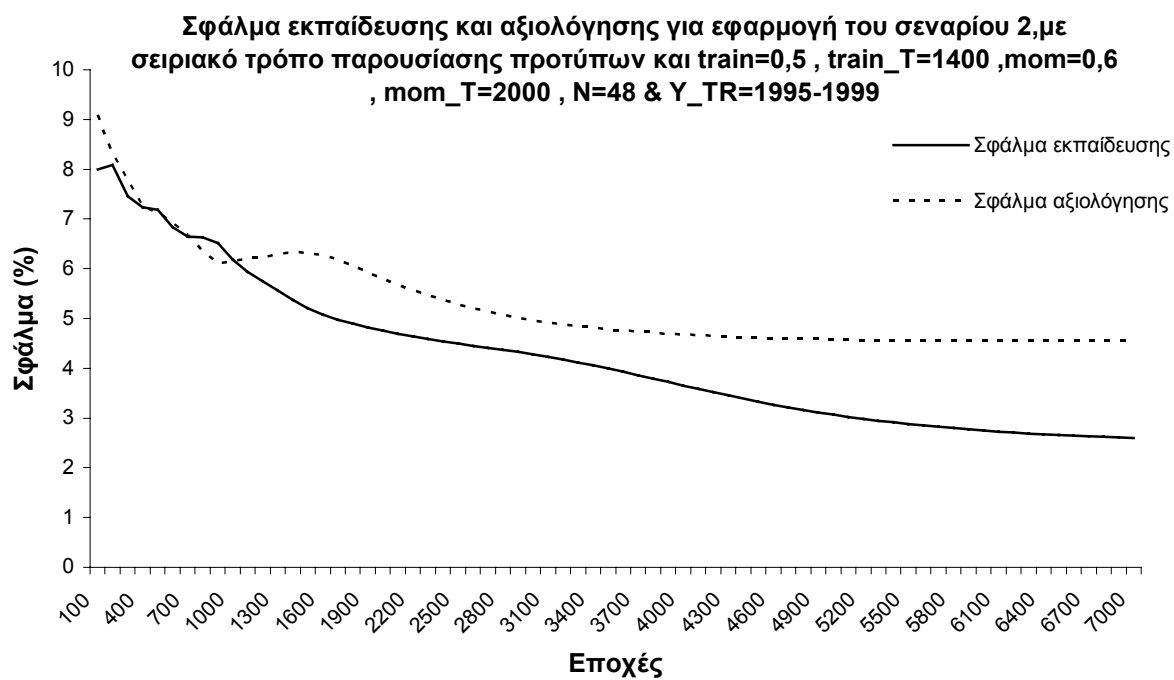
➤ ΠΙ.2 Γραφήματα εκτελέσεων του προγράμματος T.N.Δ. για τα σενάρια εκπαίδευσης με τυχαίο ή σειριακό τρόπο παρουσίασης προτύπων



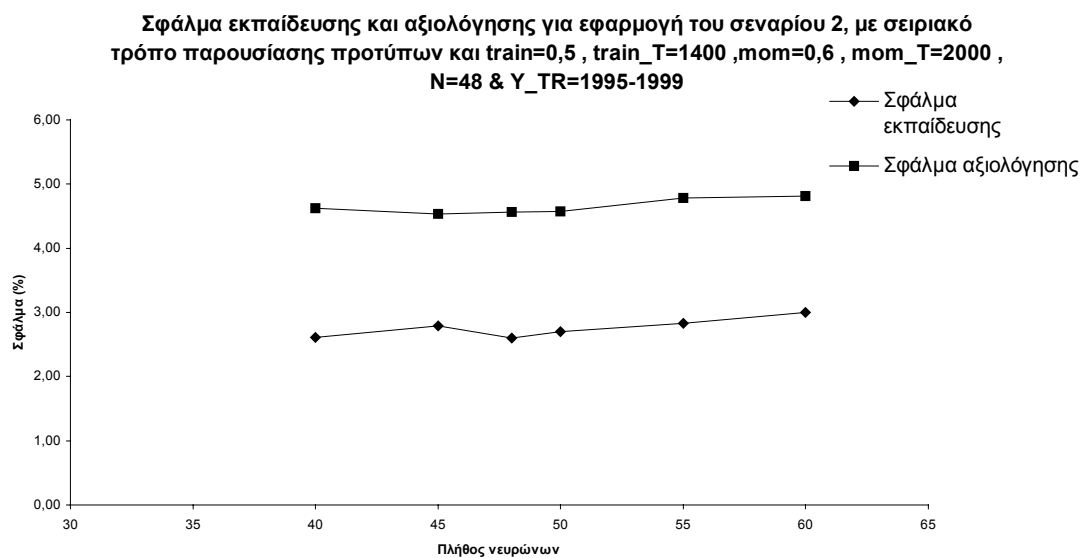
Παράρτημα Ι – Γράφημα 35



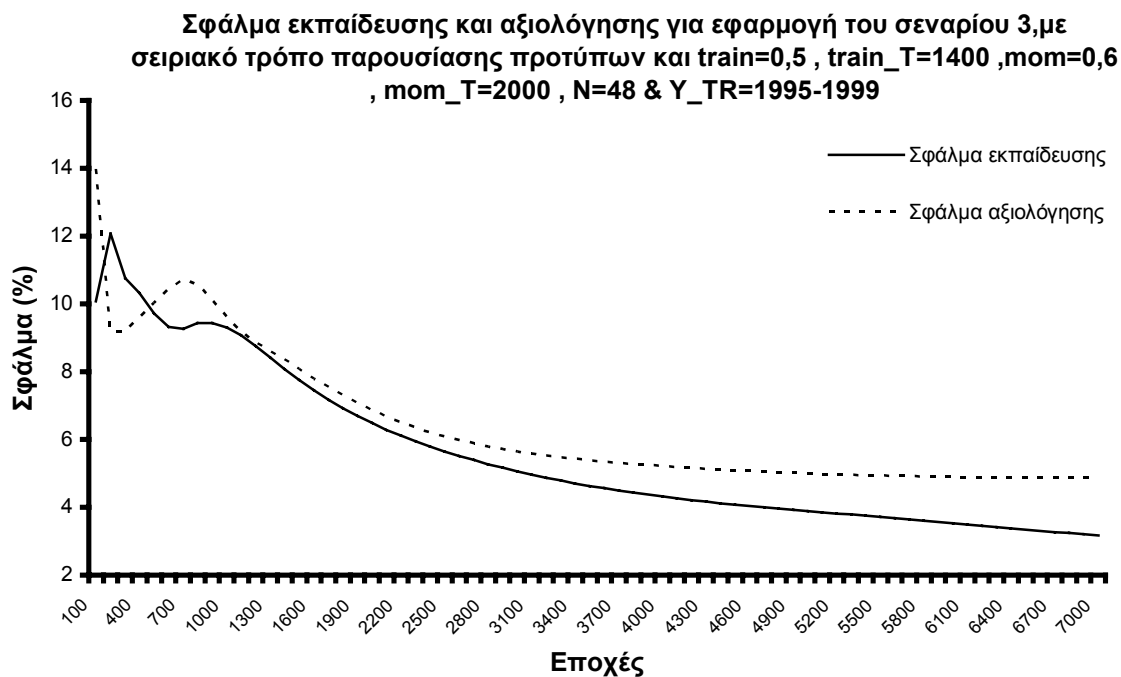
Παράρτημα Ι – Γράφημα 36



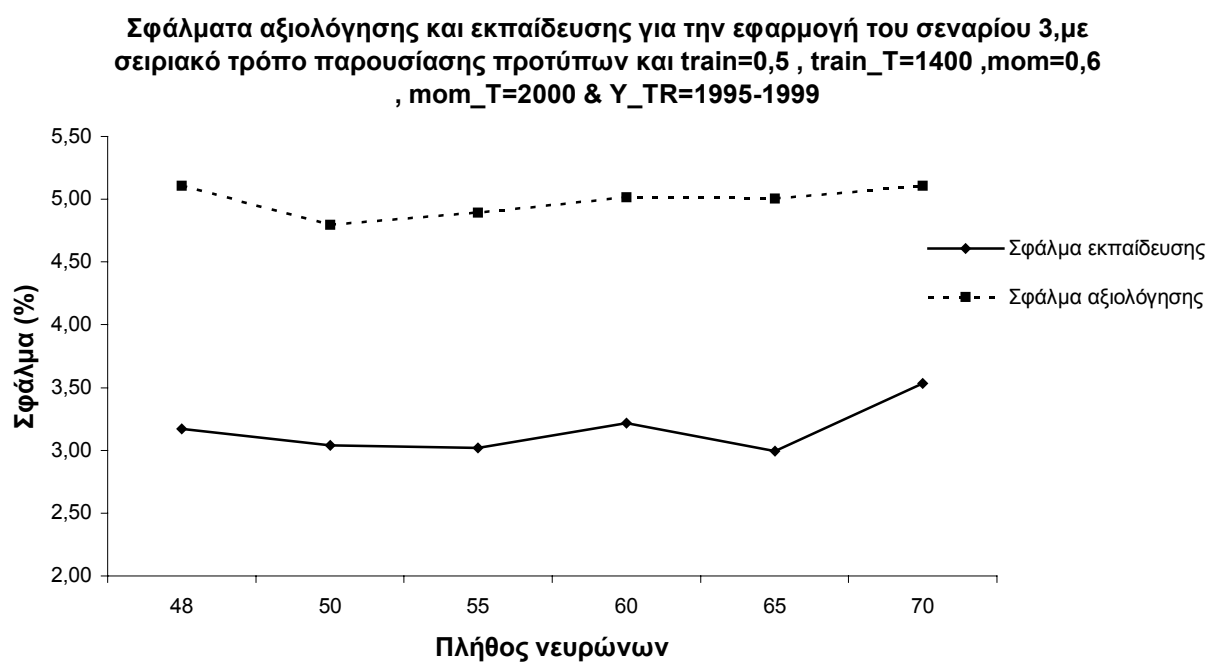
Παράρτημα Ι – Γράφημα 37



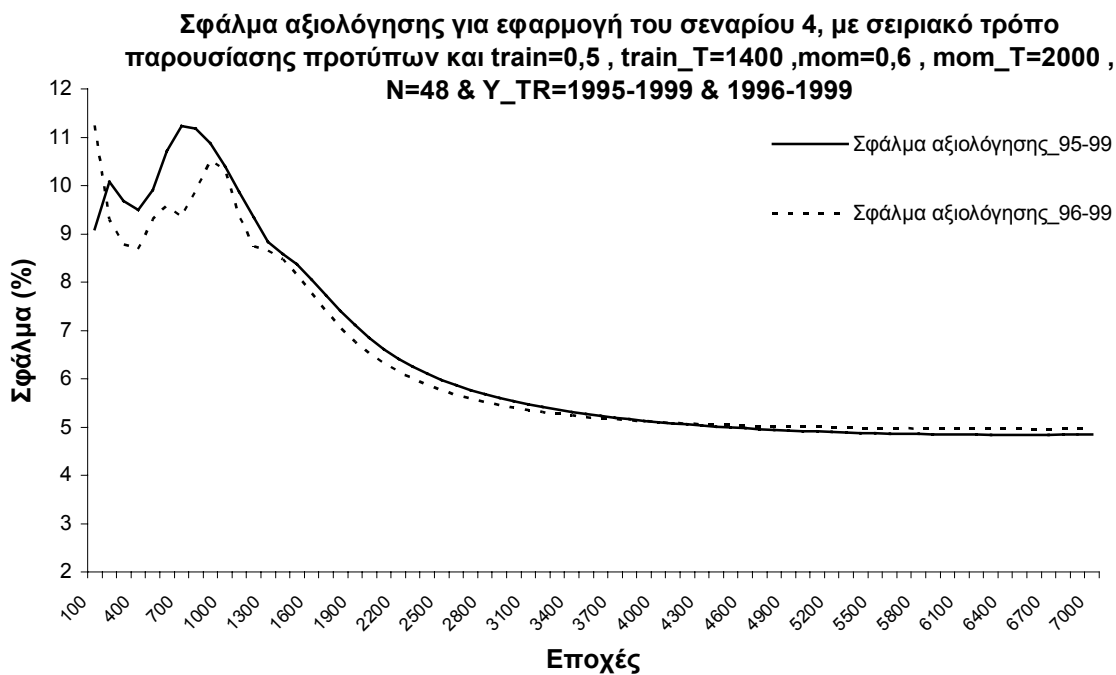
Παράρτημα Ι – Γράφημα 38



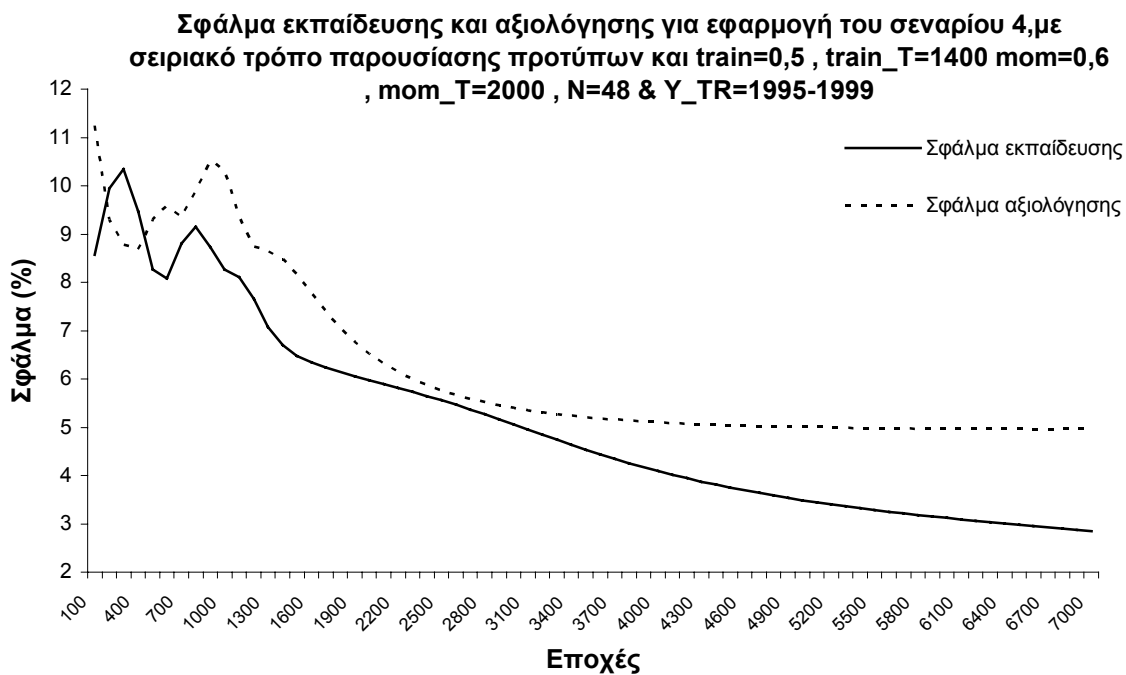
Παράρτημα Ι – Γράφημα 39



Παράρτημα Ι – Γράφημα 40

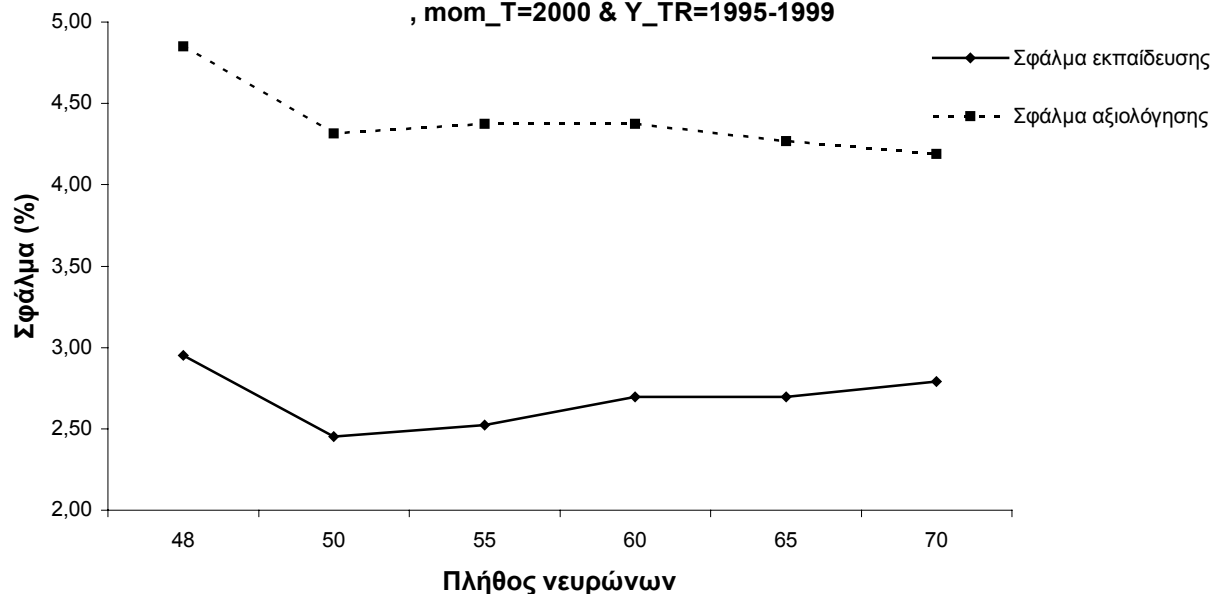


Παράρτημα Ι – Γράφημα 41



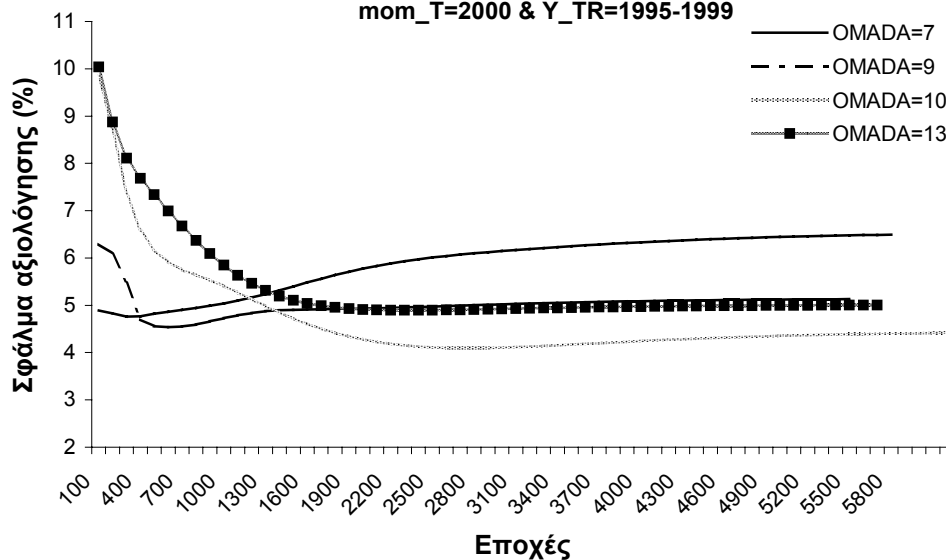
Παράρτημα Ι – Γράφημα 42

**Σφάλμα εκπαίδευσης και αξιολόγησης για εφαρμογή του σεναρίου 4, με
σειριακό τρόπο παρουσίασης προτύπων και $train=0,5$, $train_T=1400$, $mom=0,6$
 , $mom_T=2000$ & $Y_TR=1995-1999$**



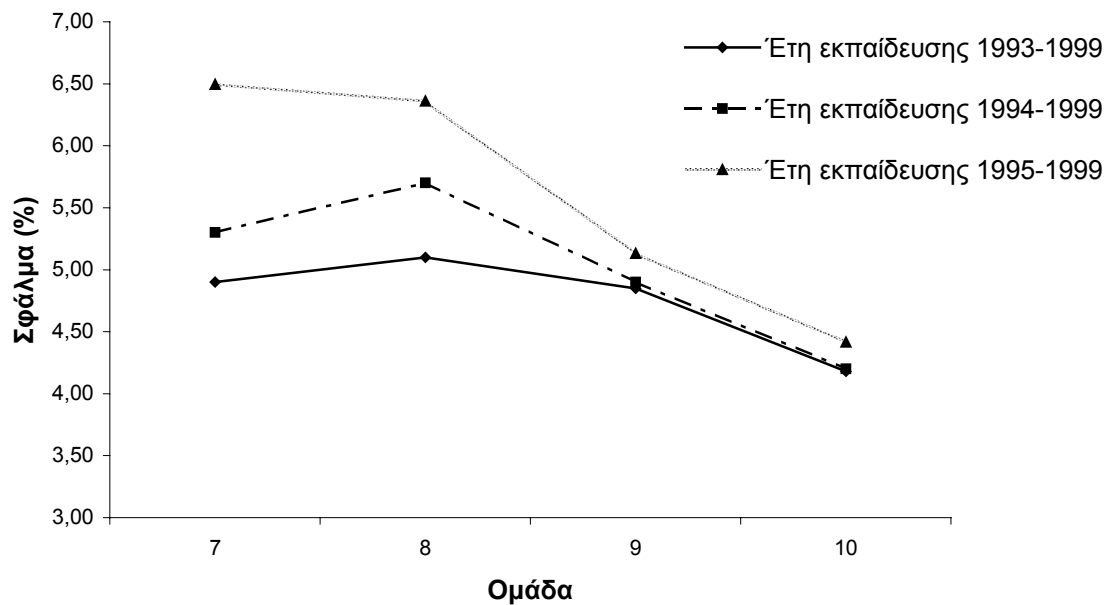
Παράρτημα Ι – Γράφημα 43

**Σφάλμα αξιολόγησης για εφαρμογή του σεναρίου 5, με σειριακό τρόπο
παρουσίασης προτύπων και με $train=0,5$, $train_T=1400$, $mom=0,6$,
 $mom_T=2000$ & $Y_TR=1995-1999$**



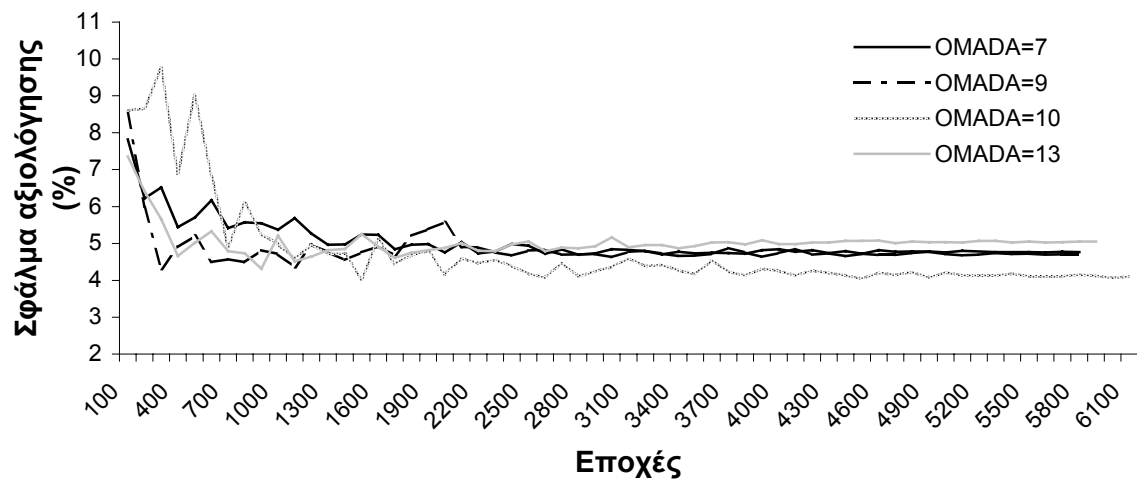
Παράρτημα Ι – Γράφημα 44

Σφάλμα αξιολόγησης για εφαρμογή του σεναρίου 5, με σειριακό τρόπο παρουσίασης προτύπων και με $\text{train}=0,5$, $\text{train_T}=1400$, $\text{mom}=0,6$, $\text{mom_T}=2000$ & $N=48$

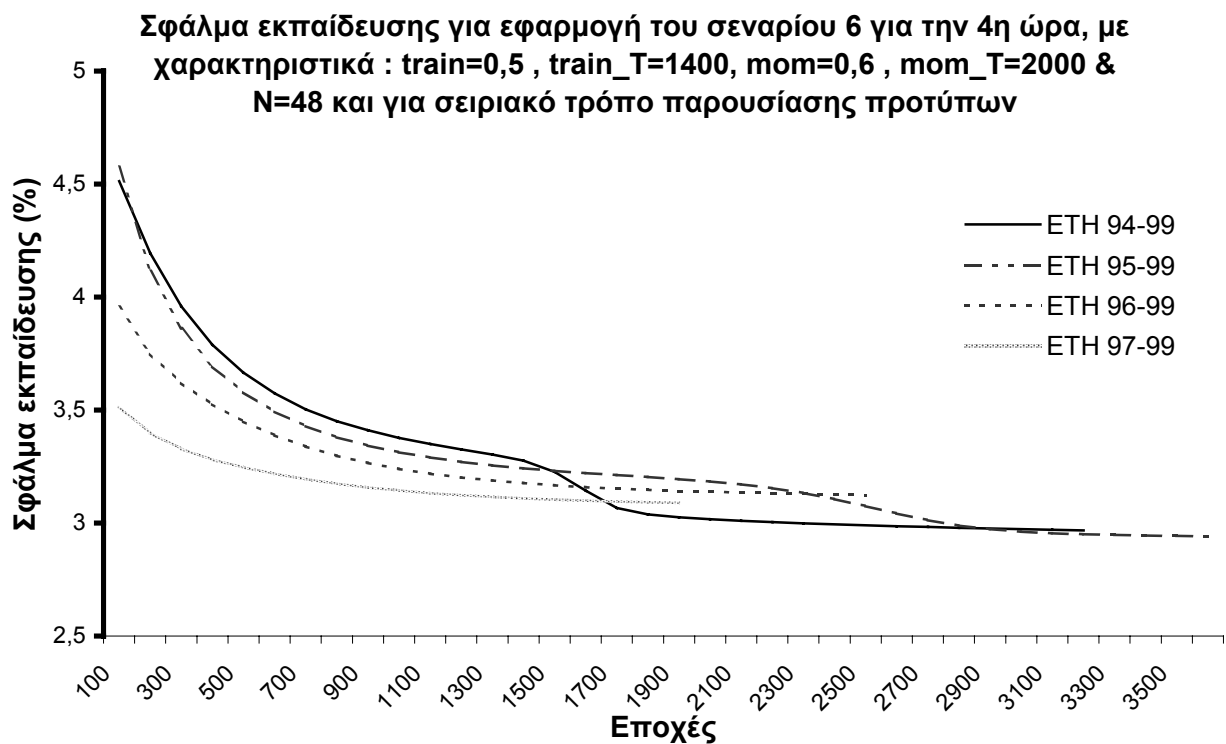


Παράρτημα Ι – Γράφημα 45

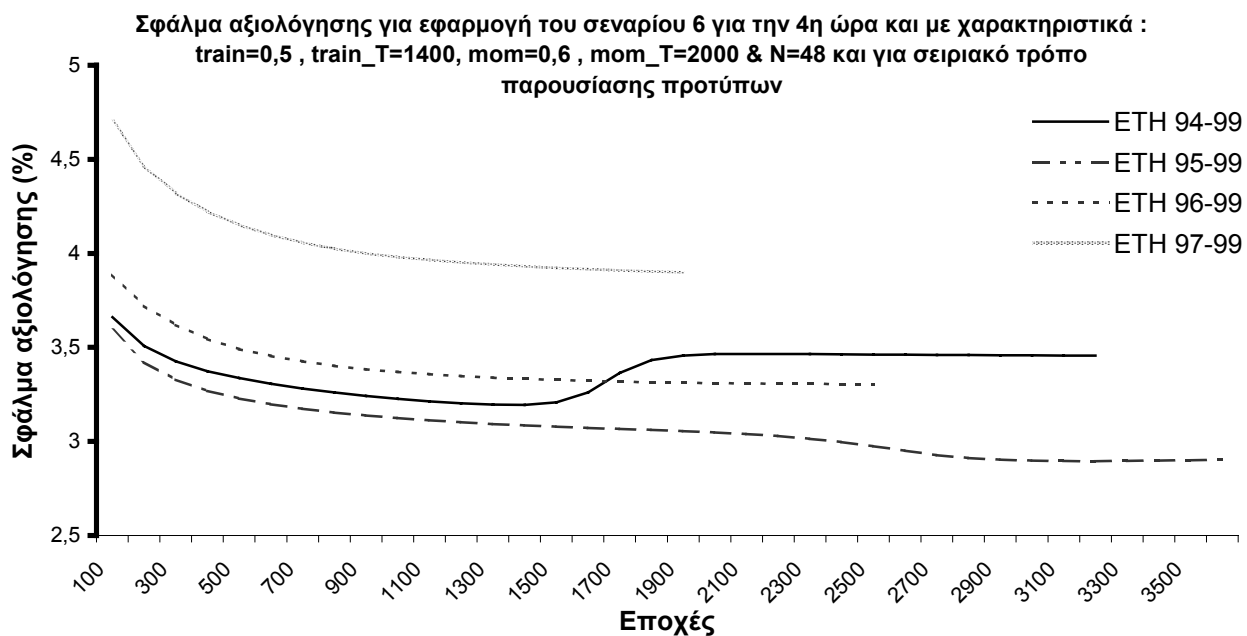
Σφάλμα αξιολόγησης για εφαρμογή του σεναρίου 5 με χαρακτηριστικά : $\text{train}=0,5$, $\text{train_T}=1400$, $\text{mom}=0,4$, $\text{mom_T}=1000$, $Y_TR=1995-1999$ & $N=48$ και για τυχαίο τρόπο παρουσίασης προτύπων



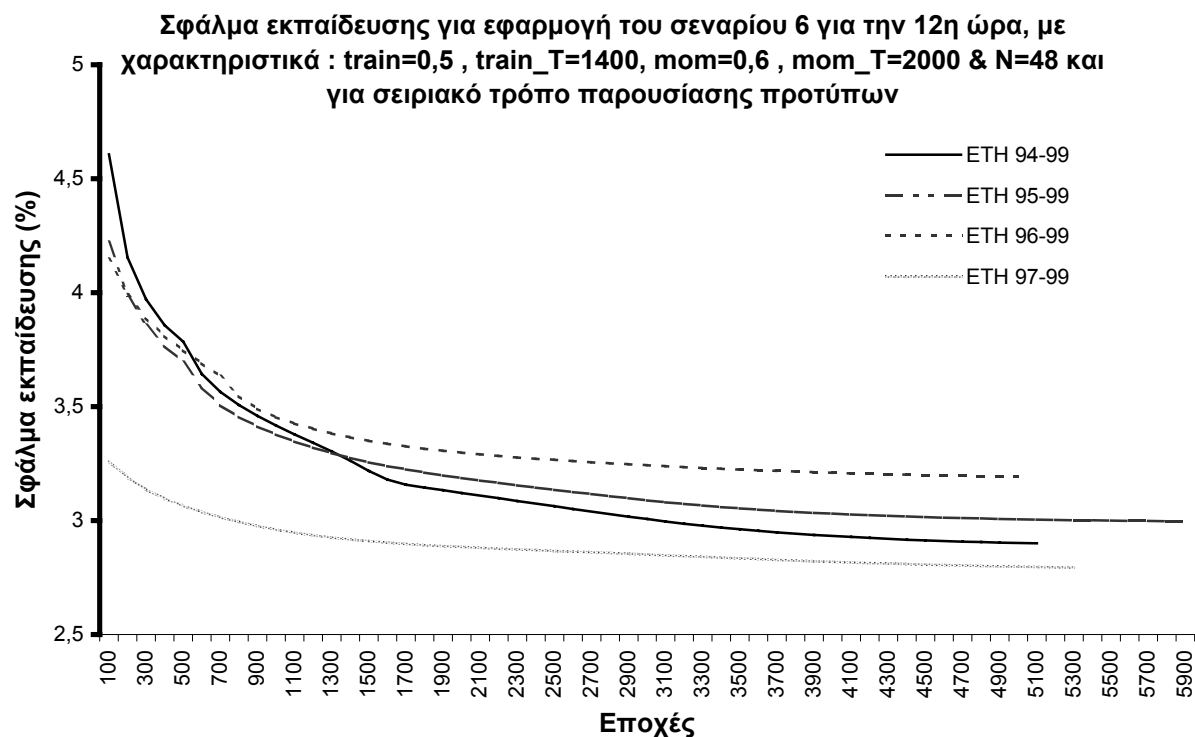
Παράρτημα Ι – Γράφημα 46



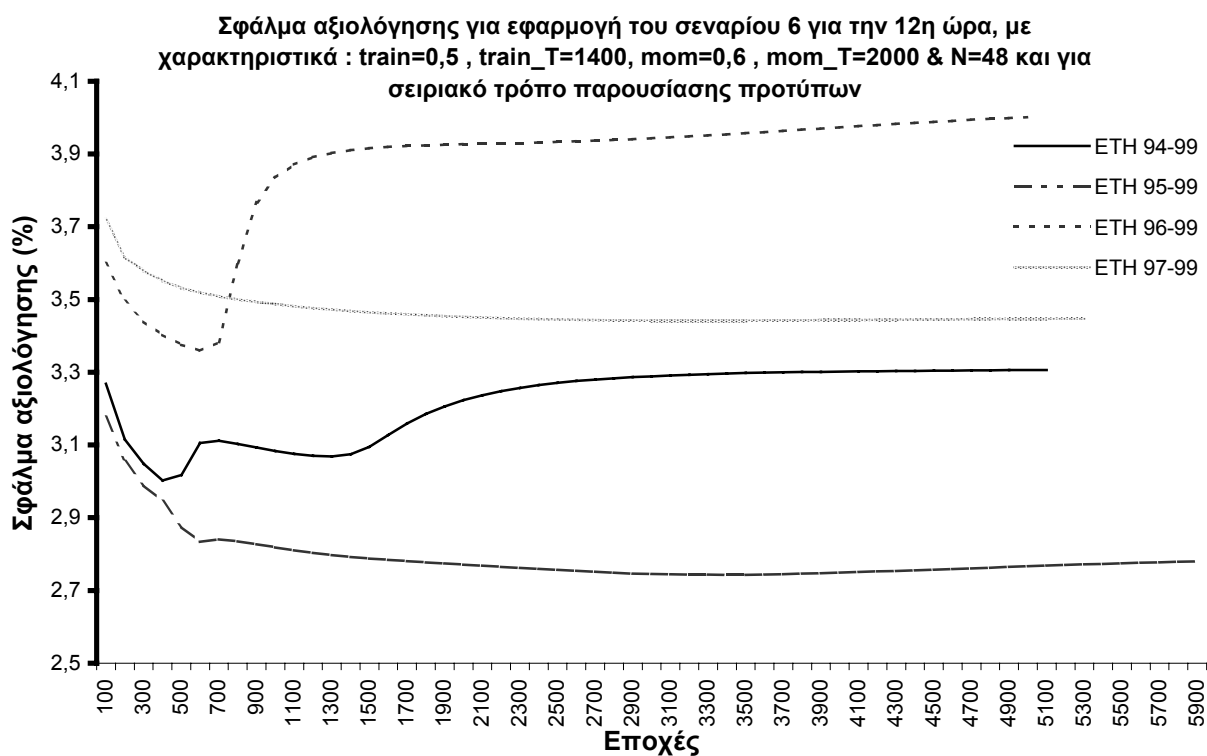
Παράρτημα Ι – Γράφημα 47



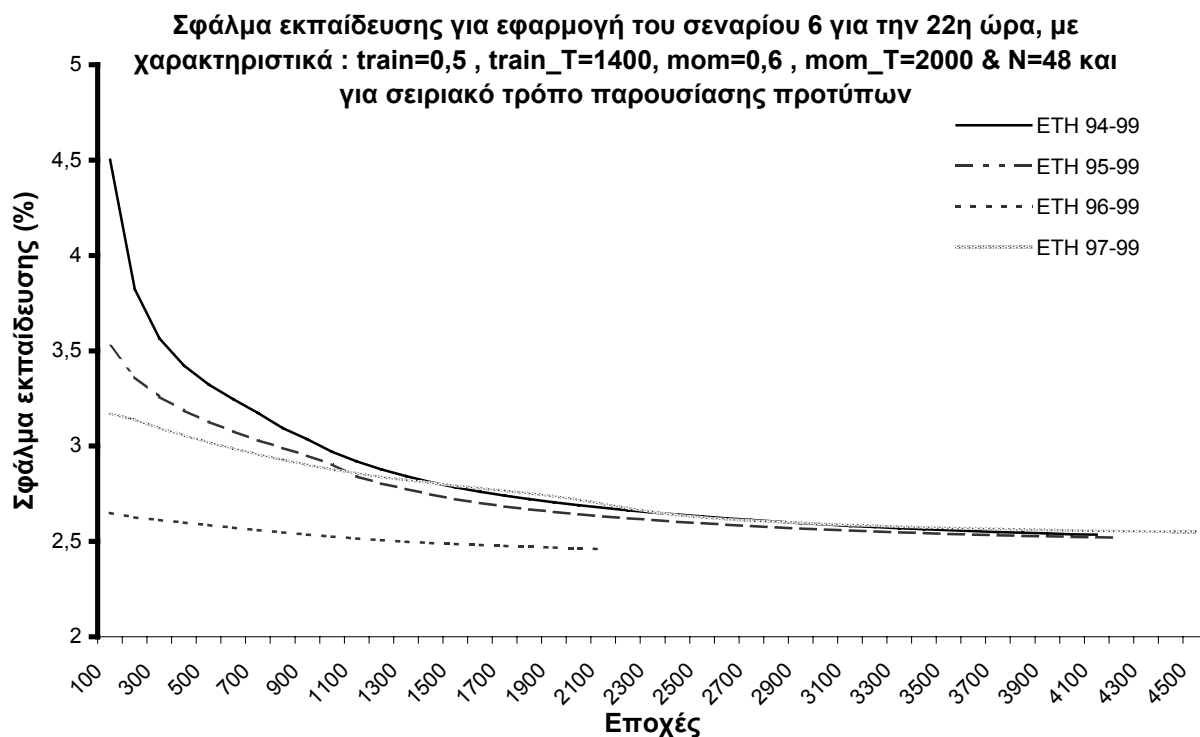
Παράρτημα Ι – Γράφημα 48



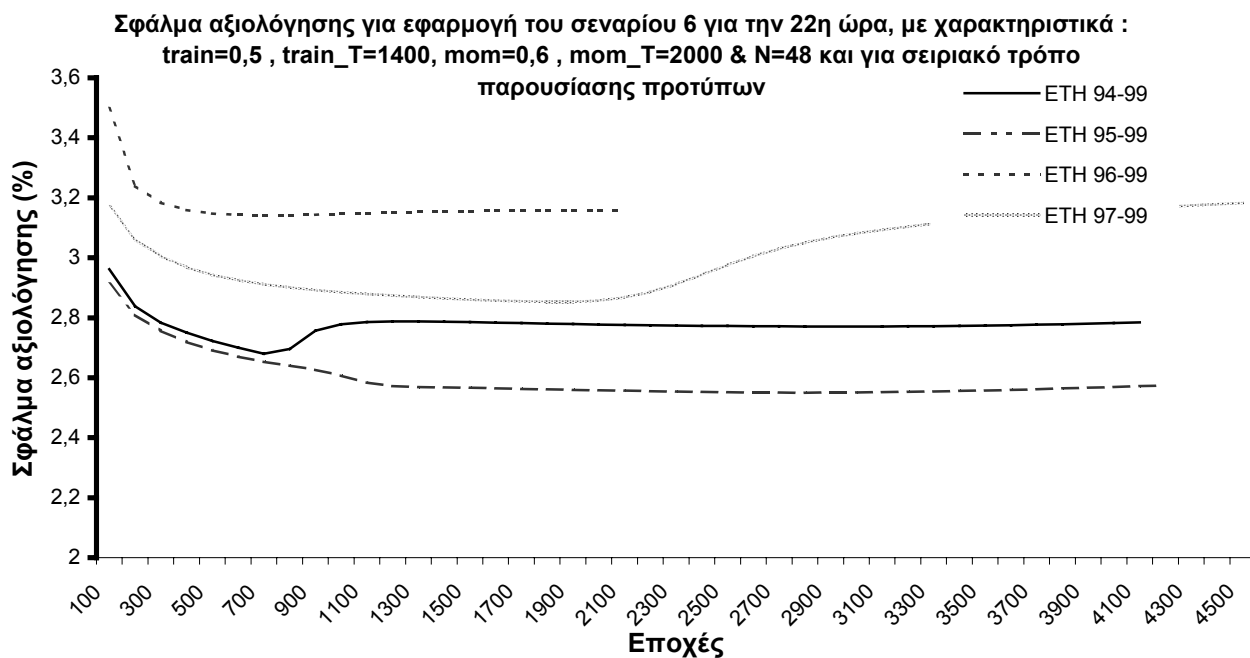
Παράρτημα Ι – Γράφημα 49



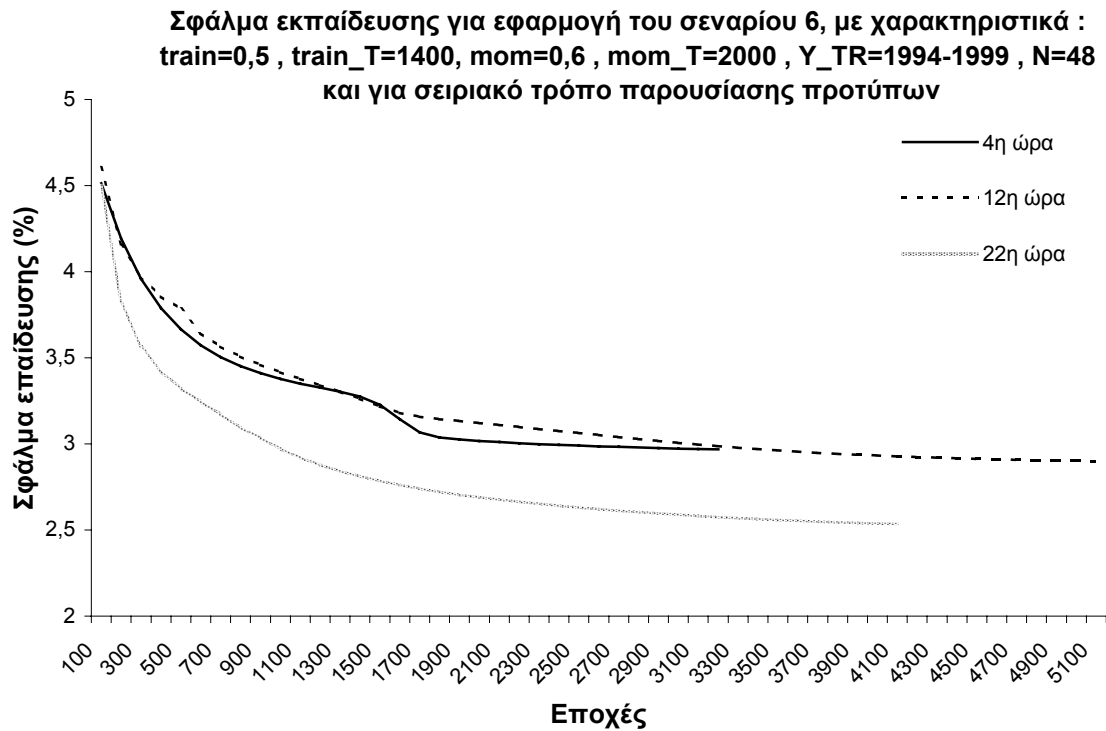
Παράρτημα Ι – Γράφημα 50



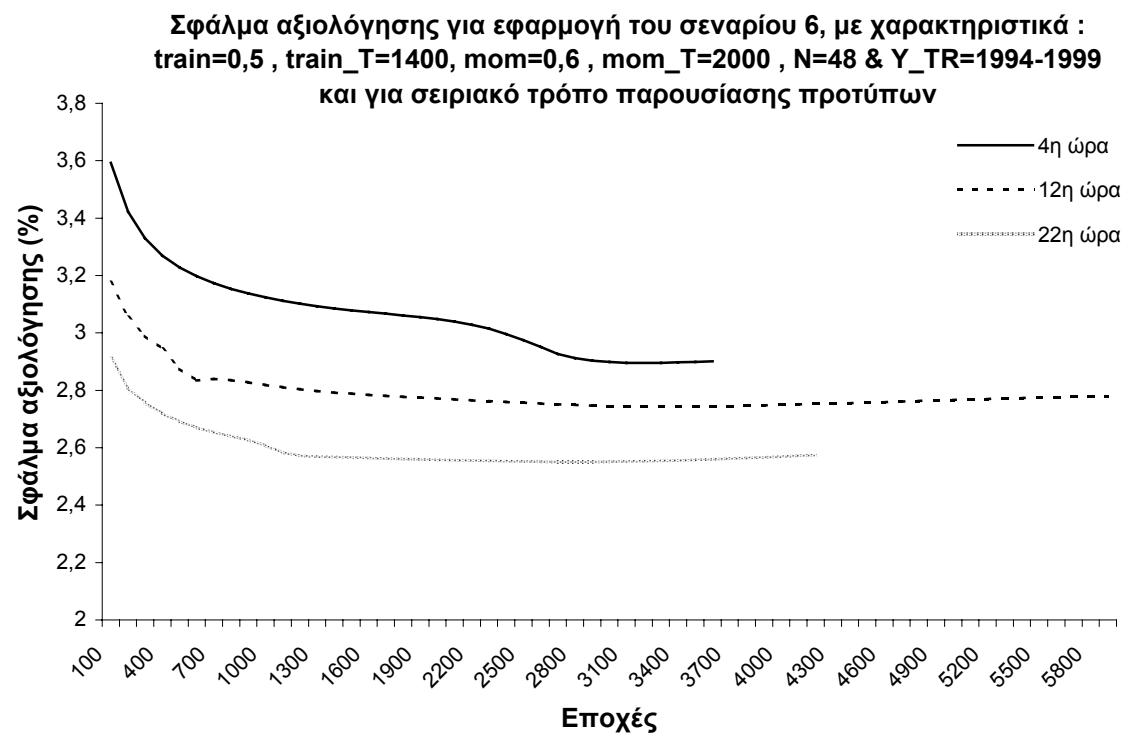
Παράρτημα Ι – Γράφημα 51



Παράρτημα Ι – Γράφημα 52

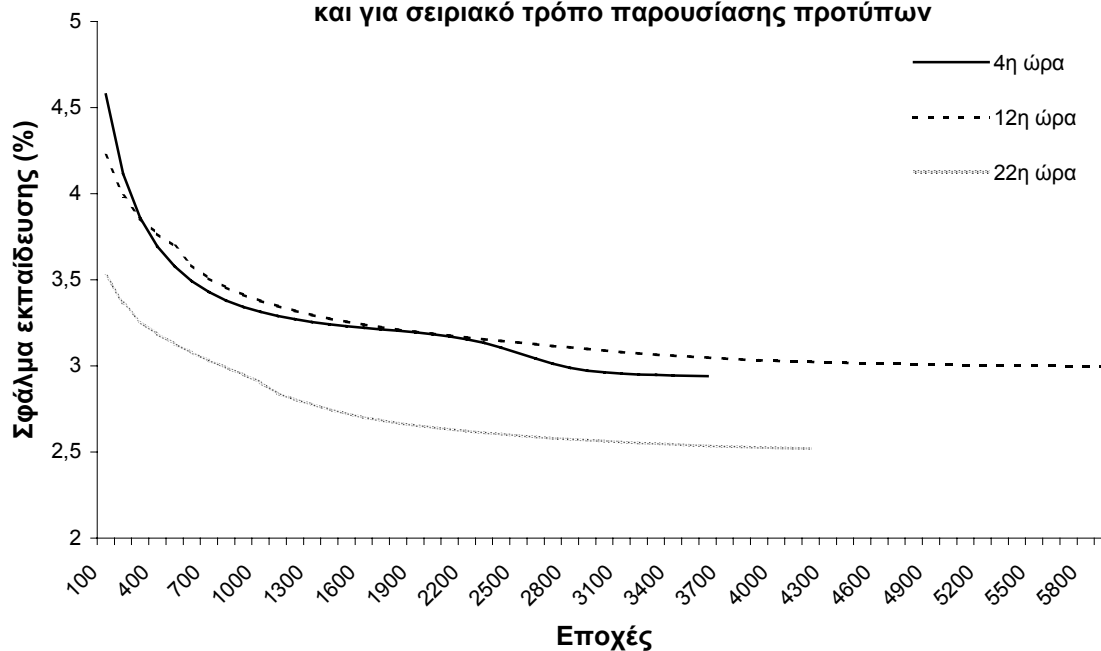


Παράρτημα Ι – Γράφημα 53



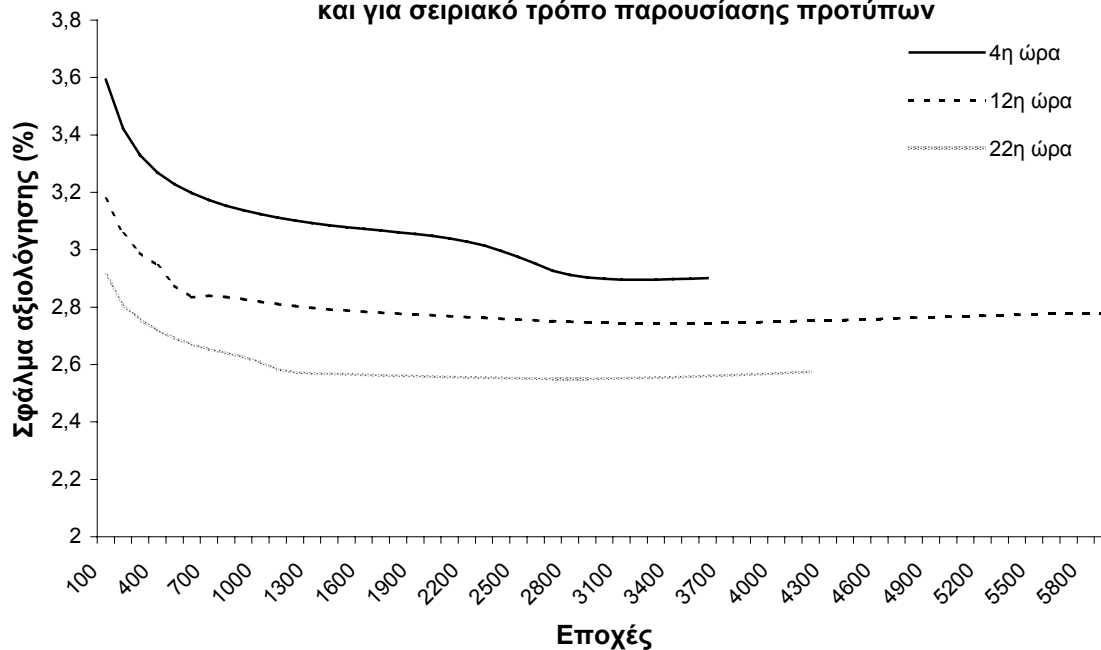
Παράρτημα Ι – Γράφημα 54

Σφάλμα εκπαίδευσης για εφαρμογή του σεναρίου 6, με χαρακτηριστικά :
 $train=0,5$, $train_T=1400$, $mom=0,6$, $mom_T=2000$, $N=48$ & $Y_TR=1995-1999$
 και για σειριακό τρόπο παρουσίασης προτύπων



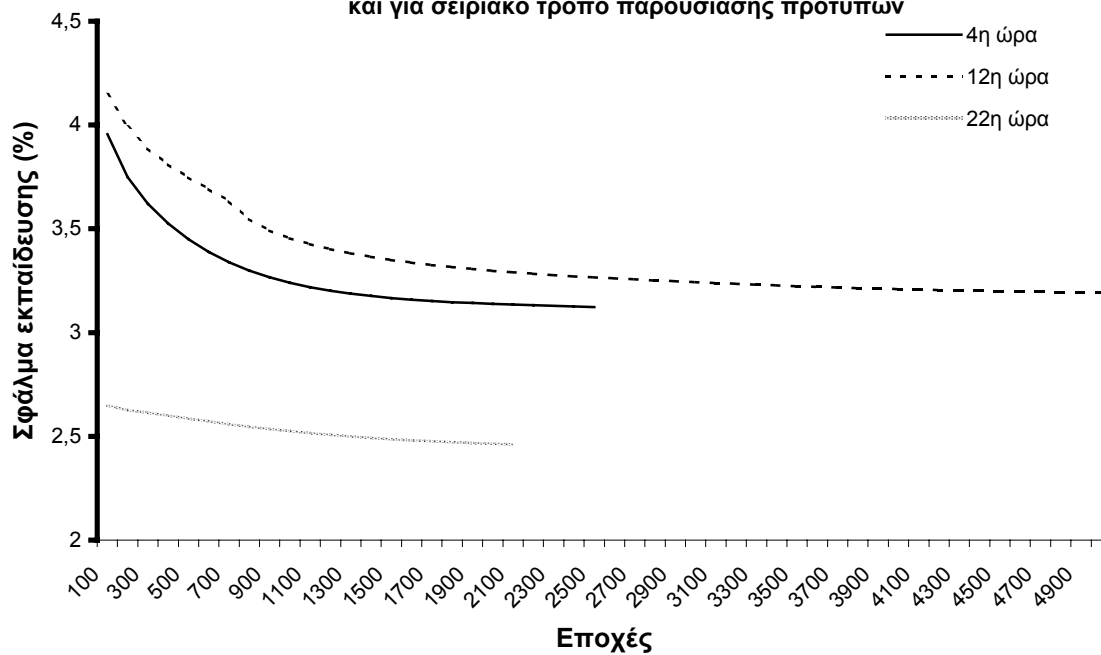
Παράρτημα Ι – Γράφημα 55

Σφάλμα αξιολόγησης για εφαρμογή του σεναρίου 6, με χαρακτηριστικά :
 $train=0,5$, $train_T=1400$, $mom=0,6$, $mom_T=2000$, $N=48$ & $Y_TR=1995-1999$
 και για σειριακό τρόπο παρουσίασης προτύπων



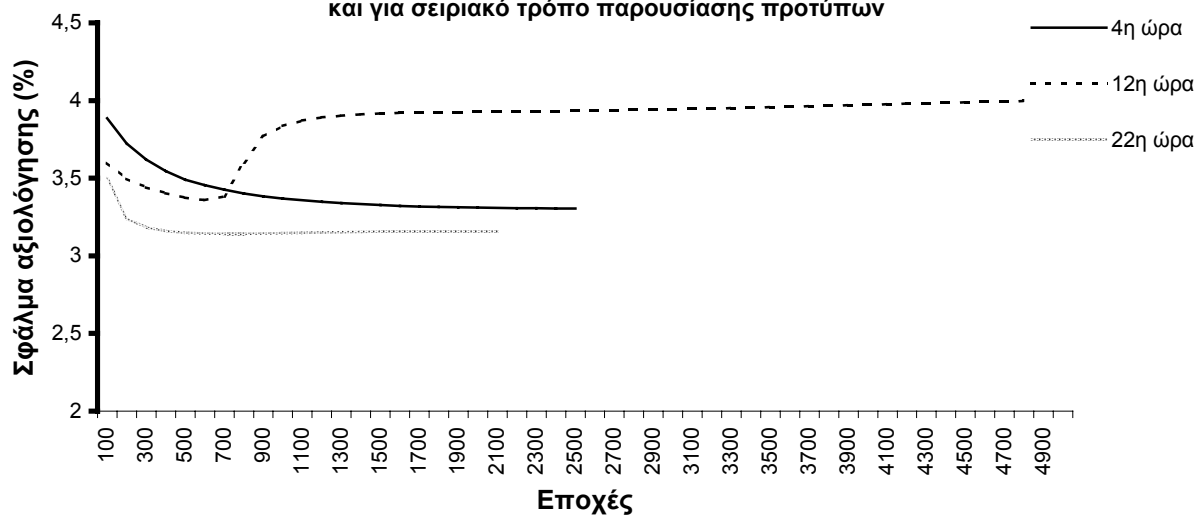
Παράρτημα Ι – Γράφημα 56

Σφάλμα εκπαίδευσης για εφαρμογή του σεναρίου 6, με χαρακτηριστικά :
 $\text{train}=0,5$, $\text{train_T}=1400$, $\text{mom}=0,6$, $\text{mom_T}=2000$, $N=48$ & $Y_TR=1996-1999$
 και για σειριακό τρόπο παρουσίασης προτύπων

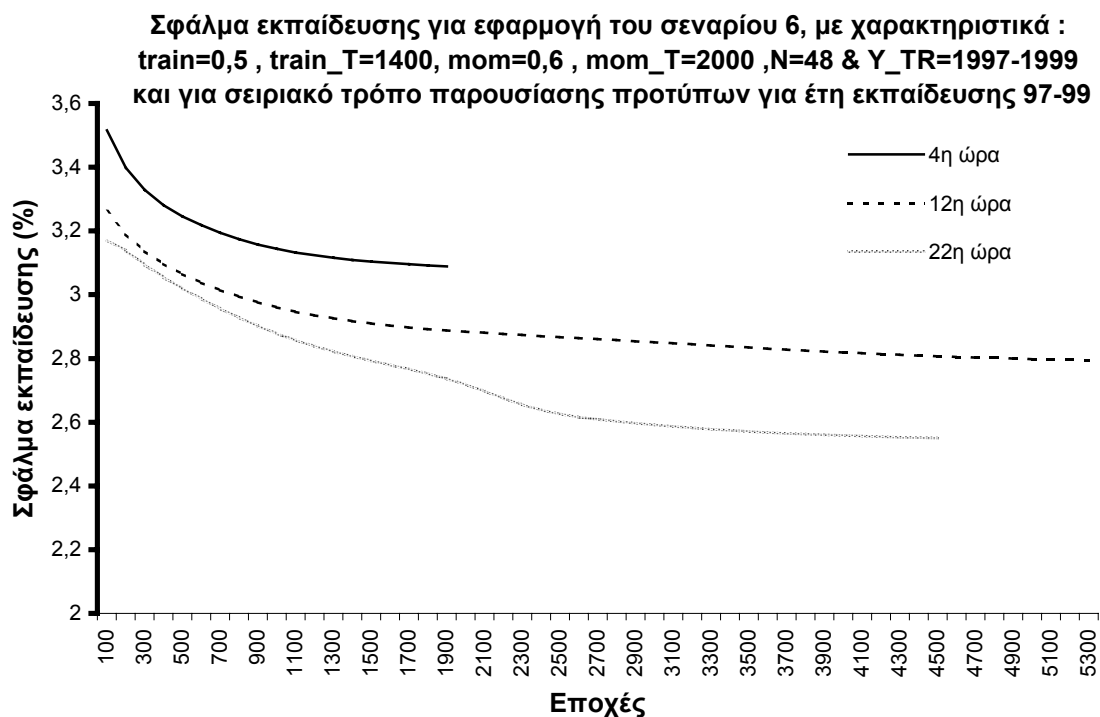


Παράρτημα Ι – Γράφημα 57

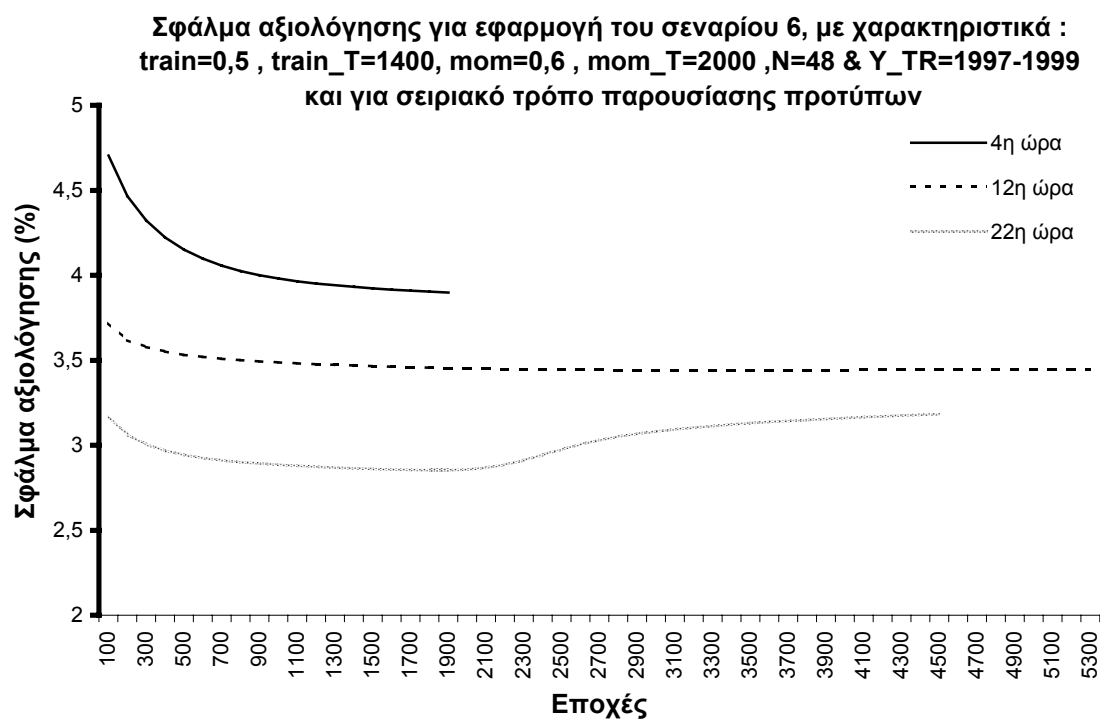
Σφάλμα αξιολόγησης για εφαρμογή του σεναρίου 6 με χαρακτηριστικά : $\text{train}=0,5$,
 $\text{train_T}=1400$, $\text{mom}=0,6$, $\text{mom_T}=2000$, $N=48$ & $Y_TR=1996-1999$
 και για σειριακό τρόπο παρουσίασης προτύπων



Παράρτημα Ι – Γράφημα 58

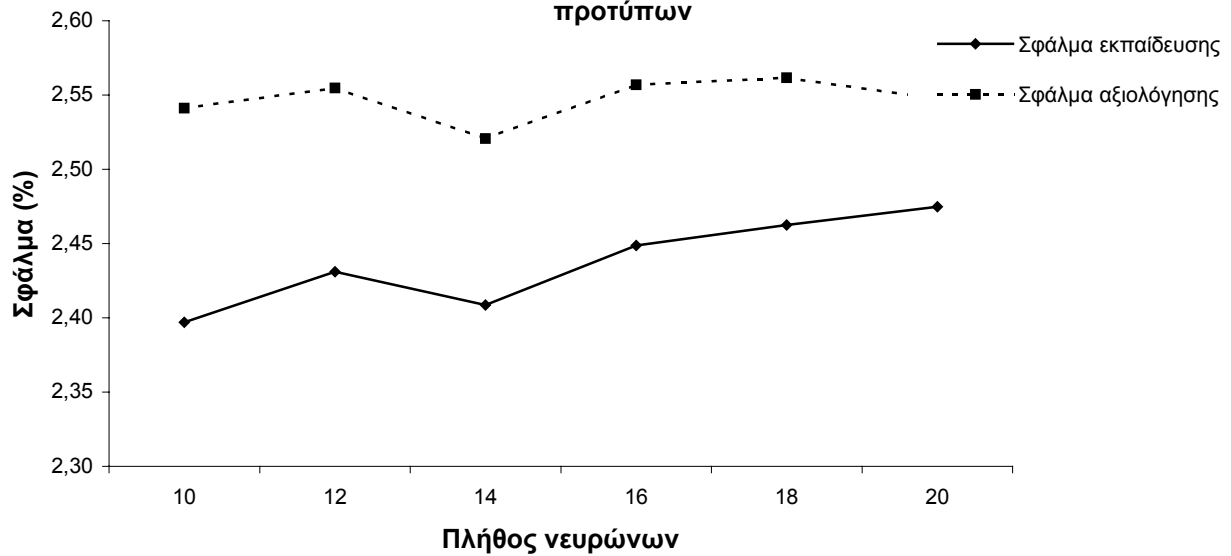


Παράρτημα Ι – Γράφημα 59



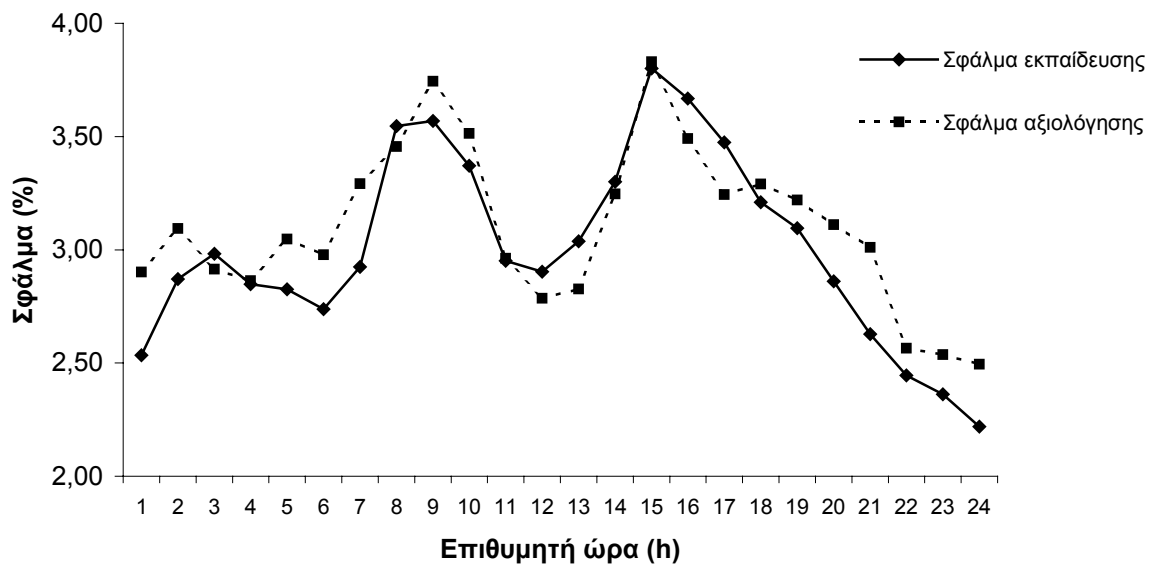
Παράρτημα Ι – Γράφημα 60

Σφάλματα εκπαίδευσης και αξιολόγησης για τεφαρμογή του σεναρίου 6 για την 22η ώραμε χαρακτηριστικά : $\text{train}=0,5$, $\text{train_T}=1400$, $\text{mom}=0,6$, $\text{mom_T}=2000$, $N=48$ & $Y_TR=1995-1999$ και για σειριακό τρόπο παρουσίασης προτύπων



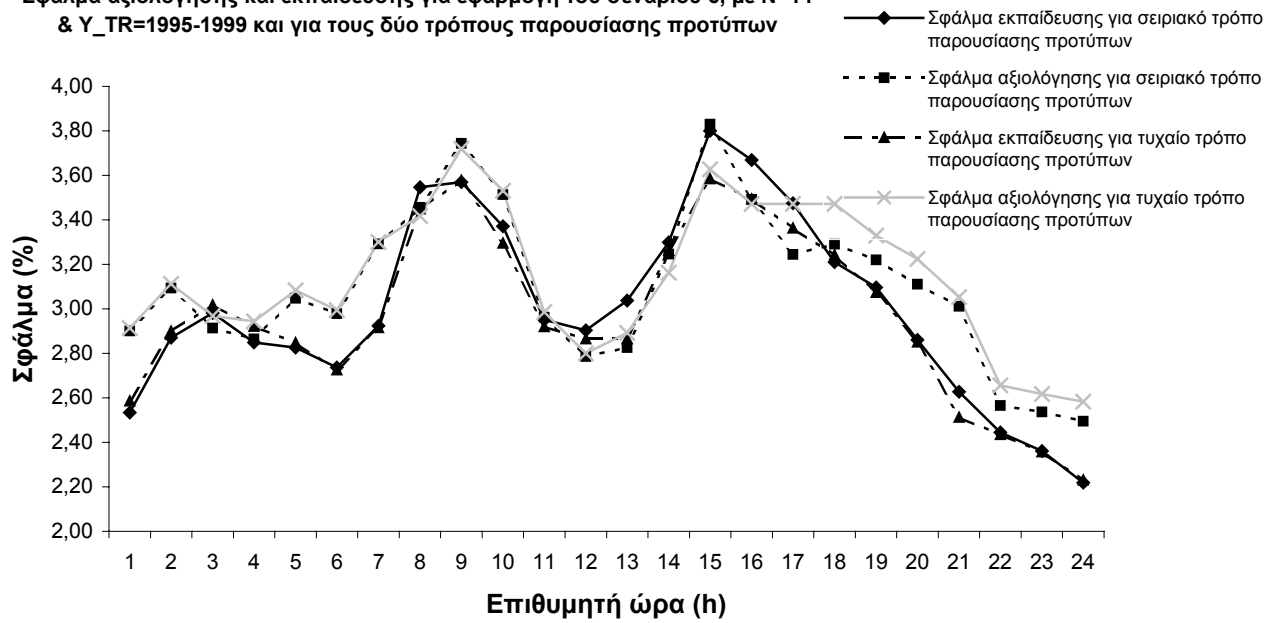
Παράρτημα Ι – Γράφημα 61

Σφάλμα αξιολόγησης και εκπαίδευσης για εφαρμογή του σεναρίου 6 με χαρακτηριστικά : $\text{train}=0,5$, $\text{train_T}=1400$, $\text{mom}=0,6$, $\text{mom_T}=2000$, $N=14$ & $Y_TR=1995-1999$ και για σειριακό τρόπο παρουσίασης προτύπων



Παράρτημα Ι – Γράφημα 62

Σφάλμα αξιολόγησης και εκπαίδευσης για εφαρμογή του σεναρίου 6, με N=14
& Y_TR=1995-1999 και για τους δύο τρόπους παρουσίασης προτύπων



Παράρτημα Ι – Γράφημα 63

Παράρτημα II

Βιβλιογραφία

ΒΙΒΛΙΟΓΡΑΦΙΑ

Σε αυτό το παράρτημα παρουσιάζουμε τα ελληνικά και ξένα συγγράμματα, τα άρθρα, τις διδακτορικές και μεταπτυχιακές εργασίες από τα οποία αντλήθηκαν πληροφορίες κατά τη συγγραφή της παρούσας διπλωματικής εργασίας. Μέσα στο κείμενο, όταν γίνεται αναφορά σε ορισμένα από αυτά, χρησιμοποιούμε την αντίστοιχη αρίθμηση.

Π 2.1 Ελληνική Βιβλιογραφία

- [B1] Κ.Βουρνάς, Γ.Κονταξής: «Εισαγωγή στα Συστήματα Ηλεκτρικής Ενέργειας», Πρώτη Έκδοση, Εκδόσεις Ε.Μ.Π., 1997
- [B2] Κ.Βουρνάς, Β.Κ.Παπαδιάς: «Παραγωγή Ηλεκτρικής Ενέργειας και Έλεγχος Συχνότητας», Δεύτερη Έκδοση, Εκδόσεις Συμμετρία, 1991
- [B3] Γ.Κονταξής, Β.Κ.Παπαδιάς: «Ηλεκτρική Οικονομία», Πρώτη Έκδοση, Εκδόσεις Ε.Μ.Π., 1996
- [B4] Γ.Κονταξής - Ν. Χατζυαργυρίου : «Κέντρα Ελέγχου Ενέργειας», Εκδόσεις Ε.Μ.Π., Αθήνα 2001
- [B5] Σ. Τζαφέστας : «Υπολογιστική Νοημοσύνη : Μεθοδολογίες», Τόμος Α, Εκδόσεις Ε.Μ.Π., Αθήνα 2002
- [B6] Σ. Τζαφέστας : « Ευφυής Αυτόματος Έλεγχος», Εκδόσεις Ε.Μ.Π., Αθήνα 1995
- [B7] Β.Ασημακόπουλος: «Μέθοδοι Προβλέψεων», Πρώτη Έκδοση, Εκδόσεις Ε.Μ.Π., 1994
- [B8] Α.Σταφυλοπάτης, Σ.Κόλλιας: «Εισαγωγή στη Θεωρία των Τεχνητών Νευρωνικών Δικτύων », Πρώτη Έκδοση, Εκδόσεις Ε.Μ.Π., Σεπτέμβριος 2000
- [B9] Αρ. Λύκας: «Υπολογιστική Νοημοσύνη», Πρώτη Έκδοση, Ιωάννινα, Σεπτέμβριος 1999
- [B10] Ε. Ζάχος : «Προγραμματιστικές Τεχνικές - Σημειώσεις», Εκδόσεις Ε.Μ.Π., Αθήνα 1999
- [B11] Α. Μπακόπουλος - Ι. Χρυσοβέργης : «Εισαγωγή στην αριθμητική ανάλυση - με βιβλιοθήκη προγραμμάτων της Fortran», Εκδόσεις Ε.Μ.Π., Αθήνα 1995
- [B12] Ν.Γ. Μαράτος : «Τεχνικές Βελτιστοποιήσεως», Ε.Μ.Π., Αθήνα 1990
- [B13] Adrian Biran, Moshe Breiner: «MATLAB για μηχανικούς», Δεύτερη Έκδοση, Εκδόσεις Τζιόλα, 1999

Π 2.2 Ξένη Βιβλιογραφία

- [Ξ1] Robert Callan: “The essence of Neural Networks”, first edition, Prentice Hall, 1999
- [Ξ2] Simon Hykin: “Neural Networks a comprehensive Foundation”, second edition, Prentice Hall, 1999
- [Ξ3] Howard Demuth, Mark Beale: “Neural Network Toolbox for Use with MATLAB”, First edition, The Math Works Inc., 1995
- [Ξ4] Eva Part-Enander, Anders Sjoberg: “The MATLAB handbook”, First edition, Henry Ling Ltd, 1999

Π 2.3 Διδακτορικές Διατριβές

[Δ1] Δ. Παπαχρήστου: «Συμβολή στην Προσομοίωση Συστημάτων Παραγωγής Ηλεκτρικής Ενέργειας», Διδακτορική Διατριβή, Αθήνα, Ε.Μ.Π., 2000

[Δ2] Σπ. Κιαρτζής: «Εφαρμογές τεχνητής νοημοσύνης στη βραχυπρόθεσμη πρόβλεψη φορτίου», Διδακτορική Διατριβή, Θεσσαλονίκη, Α.Π.Θ., 1998

Π 2.4 Μεταπτυχιακές Εργασίες

[M1] Χ.Ν. Ηλίας : «Ανάπτυξη μοντέλου βραχυπρόθεσμης πρόβλεψης φορτίου με χρήση ασαφούς λογικής και κατηγοριοποίηση τυπικών ημερών», Αθήνα, Οκτώβριος 2002

[E1] Γ.Τσεκούρας: «Καταγραφή & Αναγνώριση Φορτίου, Μελέτη Μέτρων Εξοικονόμησης Ενέργειας & Τιμολογιακών Μέτρων σε μία Χημική Βιομηχανία», Μεταπτυχιακή Εργασία στο μάθημα «Ενεργειακός Προγραμματισμός Ελαχίστου Κόστους», Ε.Μ.Π., Ιούλιος 2000

Π 2.5 Άρθρα

[A1] Roy Billinton, Mahmud Fotuhi-Firuzabad: “A basic framework for generating system operating health analysis”, IEEE Transactions on Power Systems, Vol. 9, No.3, August 1994, pp. 1610 – 1617

[A2] Saleh Aboreshaid, Roy Billinton, Mahmud Fotuhi-Firuzabad: “Probalistic transient stability studies using th method of bisection”, IEEE Transactions on Power Systems, Vol. 11, No.4, November 1996, pp. 1990 – 1995

[A3] G.A.Hamoud, E.G.Neudof, A.L.Hubert: “Reliance on non-utility generation in planning customer delivery systems”, IEEE Transactions on Power Systems, Vol. 9, No.4, November 1994, pp. 1795 – 1802

[A4] M.H. Nehrir , P.S. Dolan , V.Gerez , W.J. Jameson : “Development and Validation of a Physically-based Computer Model for Predicting Winter Electric Heating Loads”, IEEE Transactions on Power Systems, Vol. 10 , No.1, February 1995, pp. 266 – 272

[A5] B.Rautenbach, I.E.Lane: “The multi-objective controller: a novel approach to domestic hot water load control”, IEEE Transactions on Power Systems, Vol. 11, No.4, November 1996, pp. 1832 – 1837

[A6] Mark H. Gravener, Chika Nwanpka: “Available transfer capability and first order sensitivity”, IEEE Transactions on Power Systems, Vol.14, No.2, May 1999, pp. 512 – 518

[A7] Benjamin F.Hobbs, Suradet Jitprapaikulsarn, Dominic J. Maratukulam: “Analysis of the value for unit commitment of imporved load forecasts”, IEEE Transactions on Power Systems, Vol. 14, No.4, November 1999, pp. 1342 – 1348

[A8] H.B.Gooi, D.P.Mendes, K.R.W.Bell, D.S.Kirschen: “Optimal scheduling of spinning reserve”, IEEE Transactions on Power Systems, Vol. 14, No.4, November 1999, pp. 1485-1492

[A9] Mesut E. Baran, Arthur W.Kelley: “State estimation for real-time monitoring of distribution systems”, IEEE Transactions on Power Systems, Vol. 9, No.3, August 1994, pp. 1601 – 1609

[A10] Alex D.Papalexopoulos, Timothy C.Hesterberg: “A regression-based approach to short-term system load forecasting”, IEEE Transactions on Power Systems, Vol. 5, No.4, November 1990, pp. 1535 – 1547

- [A11] J. Y. Fan , J. D. McDonald : “A Real – Time Implementation of Short-Term Load Forecasting for Distribution Power Systems”, IEEE Transactions on Power Systems, Vol. 9, No.2, May 1994, pp. 988 - 994
- [A12] Alireza Khotanzad, Reza Afkhami-Rohani, Dominic Maratukulam: “ANNSTLF – Artificial neural network short-term load forecaster – Generation three”, IEEE Transactions on Power Systems, Vol. 13, No.4, November 1998, pp. 1413 – 1422
- [A13] O. Hyde , P. F. Hodenett : “An Adaptable Automatd Procedure for Short-Term Electricity Load Forecasting”, IEEE Transactions on Power Systems, Vol. 12, No.1, February 1997, pp. 84 – 94
- [A14] Takeshi Haidda: “Regression Based Peak Load Forecasting Using A transformation Technique”, IEEE Transactions on Power Systems, Vol. 9, No.4, November 1994, pp. 1788-1794
- [A15] D.G.Infield, D.C.Hill: “Optimal smoothing for trend removal in short term electricity demand forecasting”, IEEE Transactions on Power Systems, Vol. 13, No.3, August 1998, pp. 1115 – 1120
- [A16] Parech Rupanagunta , Matin L. Baughman , Jerold W. Jones: “Scheduling of Cool Storage Using Non-Linear Programming Techniques”, IEEE Transactions on Power Systems, Vol. 10, No.3, August 1995, pp. 1279 – 1285
- [A17] C. W. Gellings, R. W. Taylor : “Electric Load Curve Synthesis – A Computer Simulation of an Electric Utility Load Shape”, IEEE Power Engineering Society Summer Meeting, 1980 or Talukdar & Gellings : “Load Management”, IEEE Press, pp. 111-116
- [A18] H.R.Kassaei, A.Keyhani, T.Woung, M.Rahman: “A hybrid fuzzy neural network bus load modeling and prediction”, IEEE Transactions on Power Systems, Vol. 14, No.2, May 1999, pp. 718 – 724
- [A19] G.Kariniotakis, E.Nogaret: “Secure wind power penetration in isolated systems: Load & wind power forecasting using artificial intelligence methods”, CARE Workshop, Heraklion Crete, 16-17 July 1999, pages 22
- [A20] Ibrahim Moghran, Saifur Rahman: “Analysis and evaluation of five short-term load forecasting techiques”, IEEE Transactions on Power Systems, Vol. 4, No.4, November 1989, pp. 1484 – 1491
- [A21] J. H. Broehl : “An End-Use Approach to Demand Forecasting”, IEEE Power Engineering Society Summer Meeting, 1980 or Talukdar & Gellings : “Load Management”, IEEE Press, pp. 117-121
- [A22] Alex D.Papalexopoulos, Shangyou Hao, Tie-Mao Peng: “An implementation of a neural network based load forecasting model for the EMS”, IEEE Transactions on Power Systems, Vol. 9, No.4, November 1994, pp. 1956 – 1962
- [A23] A. Cavallini , G. Mazzanti , G. C. Montanari : “Inference of a continuous Autoregressive Model for the Forecasting of Non Stationary Stochastic Processes Deriving from Energy Demand in Electrical Networks”, Melecon Conference 1996, Volume 2, pp.726-730
- [A24] J. Nazarko , W.Zalewski : “The Fussy Regression Approach to Peak Load Estimation in Power Distribution Systems”, IEEE Transactions on Power Systems, Vol. 14, No.3, August 1999, pp. 809 – 814
- [A25] A.G.Bakirtzis, J.B.Theocharis, S.J.Kiartzis, K.J.Satsios: “Short term load forecasting using fuzzy neural networks”, IEEE Transactions on Power Systems, Vol. 10, No.3, August 1995, pp. 1518 – 1524
- [A26] Kwang-Ho Kim, Jong-Keun Park, Kab-Ju Hwang, Sung-Hak Kim: “Implemetation of hybrid short-term load forecasting system using artificial neural networks and fuzzy expert systems”, IEEE Transactions on Power Systems, Vol. 10, No.3, August 1995, pp.1534-1539

- [A27] W. Charytoniuk, M.S.Chen, P.Van Olinda: "Nonparametric Regression based short-term load forecasting", IEEE Transactions on Power Systems, Vol. 13, No.3, August 1998, pp. 725 – 730
- [A28] Alizera Khotanzad, Rey-Chue Hwang, Alireza Abaye, Dominic Maratukulam: "An adaptive modular artificial neural network hourly load forecaster and its implementation at electric utilities", IEEE Transactions on Power Systems, Vol. 10, No.3, August 1995, pp. 1716 – 1722
- [A29] Alizera Khotanzad, Reza Afkhami-Rohani, Tsun-Liang Lu, Alireza Abaye, Malcolm Davis, Dominic Maratukulam: "ANNSTLF- Neural network-based electric load forecasting system", IEEE Transactions on Neural Networks, Vol. 8, No.4, July 1996, pp. 835 – 846
- [A30] O. Mohammed, D. Park, R. Merchant, T. Dinh, C. Tong, A. Azeem: "Practical experiences with an adaptive neural network short-term load forecasting system", IEEE Transactions on Power Systems, Vol. 10, No.1, February 1995, pp. 254 – 265
- [A31] A.G.Bakirtzis, V.Petridis, S.J.Kiartzis, M.C.Alexiadis, A.H.Maissis: "A neural network short term load forecasting model for the Greek power system", IEEE Transactions on Power Systems, Vol. 11, No.2, May 1996, pp. 858 – 863
- [A32] S.J.Kiartzis, C.E.Zournas, J.M.Theocharis, A.G.Bakirtzis, V.Petridis: "Short term load forecasting in an autonomous power system using artificial neural networks", IEEE Transactions on Power Systems, Vol. 12, No.4, November 1997, pp. 1591 – 1596
- [A33] Jerome Connor, Les E.Atlas, Douglas R.Martin: "Recurrent networks and NARMA modeling", Advances in Neural Information Processing Systems NIPS4, 1992, pp. 301 – 308
- [A34] J.Vermaak, E.C.Botha: "Recurrent neural networks for short-term load forecasting", IEEE Transactions on Power Systems, Vol. 13, No.1, February 1998, pp. 126 – 132
- [A35] Shin-Tzo chen, David C.Yu, A.R.Moghaddamjo: "Weather sensitive short-term load forecasting using nonfully connected artificial neural network", IEEE Transactions on Power Systems, Vol. 7, No.3, August 1992, pp. 1098 – 1105
- [A36] A.S.AlFuhaid, M.A.El-Sayed, M.S.Mahmoud: "Cascaded artificial neural networks for short term load forecasting", IEEE Transactions on Power Systems, Vol. 12, No.4, November 1997, pp. 1524 – 1529
- [A37] I.Drezga, S.Rahman: "Input variable selection for ANN-based short-term load forecasting", IEEE Transactions on Power Systems, Vol. 13, No.4, November 1998, pp. 1238 – 1244
- [A38] M.Hisham, Choueiki, Clark A.Mount-Campbell, Stanley C.Ahalt: "Building a 'Quasi optimal' neural network to solve the short-term load forecasting problem", IEEE Transactions on Power Systems, Vol. 12, No.4, November 1997, pp. 1432 – 1439
- [A39] Andrew P. Douglas, Arthur M. Breipohl, Fred N. Lee, Rambabu Adapa: "The Impacts of Temperature Forecast Uncertainty on Bayesian Load Forecasting", IEEE Transactions on Power Systems, Vol. 12, No.4, November 1997, pp. 1507 - 1513
- [A40] E.A.Bossanyi: "Short-term wind prediction using Kalman filters", Wind Engineering, Vol.9, No.1, 1985
- [A41] Djurovic M., Stankovic L.: "Prediction of wind characteristic in short term periods", Energy and the environment into the 1990s, Proceedings of the 1st World Renewable Energy Congress, 1990, pp. 1714-1718
- [A42] L.Lin: "Predicting wind behaviour with neural networks", Proceedings of the 1996 European Union Wind Energy Conference, EUWEC'96, Geteborg, Sweden
- [A43] G.N.Kariniotakis, G.S.Stavarakakis, E.F.Nogaret: "Wind power forecasting using advanced neural networks models", IEEE Transactions on Energy Conversion, Vol.11, No.4, December 1996, pp. 762 – 767

- [A44] G.C.Contaxis, J. Kabouris: "Short-term scheduling in a wind/diesel autonomous system", IEEE Transactions on Power Systems, 91WM 188-3 PWRs
- [A45] H.M. Geerts: "Short range prediction of wind speeds: a system-theoretic approach", Proceedings of the European Wind Energy Conference, EWEC'84, Hamburg, FRG, 22-26 October 1984, pp.594-599
- [A46] G.Kariniotakis, M.Matos, V.Miranda: "Assessment of the benefits from advanced load & wind power forecasting in autonomous power systems", CARE Workshop, Heraklion Crete, 16-17 July 1999, pages 4
- [A47] I.E. Lane, N.Beute: "A Model of the Domestic Hot Water Load", IEEE Transactions on Power Systems, Vol. 11, No.4, November 1996, pp. 1850 – 1855
- [A48] Paresh Rupanagunta, Martin L. Baughman, Jerold W. Jones: "Scheduling of cool storage using non-linear programming techniques", IEEE Transactions on Power Systems, Vol. 10, No.3, August 1995, pp. 1279 – 1285
- [A49] J. L. Small: "Joint estimation of household and heat-pump electricity demand", IEEE Transactions on Power Systems, Vol. 9, No.1, February 1994, pp. 413 – 419
- [A50] H.L.Willis, C.L.Brooks: "An interactive end-use electric load model for microcomputer implementation", IEEE Transactions on Power Apparatus and Systems, Vol. PAS-102, No.11, November 1983, pp. 3693 – 3700
- [A51] Damitha K. Ranaweera, George G. Karady, Richard G. Farmer: "Economic Impact Analysis of Load Forecasting", IEEE Transactions on Power Systems, Vol. 12, No.3, August 1997, pp. 1388 – 1392
- [A52] I.Drezga, S.Rahman: "Short-term load forecasting with local ANN predictors", IEEE Transactions on Power Systems, Vol. 14, No.3, August 1999, pp. 844 - 850
- [A53] Damitha K.Ranaweera, George G.Karady, R.G.Farmer: "Effect of probabilistic inputs on neural network-based electric load forecasting", IEEE Transactions on Neural Networks, Vol. 7, No.6, November 1996, pp. 1528 – 1532
- [A54] Marie Blanchard, Louis Delorme, Christiane Simard, Yvon Nadeau: "Experience with optimization software for distribution system planning", IEEE Transactions on Power Systems, Vol. 11, No.4, November 1996, pp. 1891 - 1898
- [A55] G.Kariniotakis: "Short-term wind power forecasting for the management of Power Systems including WECS", Centre d'Energetique, Position paper on short-term wind forecasting, pages 13.

.....

ΣΑΒΒΑΣ Ι. ΑΝΔΡΟΥΤΣΟΠΟΥΛΟΣ

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

© 2003 – All rights reserved