

**ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ**  
**ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΕΡΙΒΑΛΛΟΝΤΟΣ**  
**ΠΜΣ «ΠΕΡΙΒΑΛΛΟΝΤΙΚΗ ΚΑΙ ΥΓΕΙΟΝΟΜΙΚΗ**  
**ΜΗΧΑΝΙΚΗ»**



**ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΑΤΡΙΒΗ:**

**Βέλτιστος Σχεδιασμός Διαχείρισης Υπογείων Υδάτων του  
Βόρειου Τμήματος της Νήσου Ρόδου με Χρήση Νευρωνικών  
Δικτύων**

**ΣΤΕΡΓΙΑΔΗ ΜΑΡΙΑ**

**ΕΞΕΤΑΣΤΙΚΗ ΕΠΙΤΡΟΠΗ:**

**ΚΑΡΑΤΖΑΣ ΓΕΩΡΓΙΟΣ (επιβλέπων)**

**ΝΙΚΟΛΟΣ ΙΩΑΝΝΗΣ**

**ΝΙΚΟΛΑΪΔΗΣ ΝΙΚΟΛΑΟΣ**

**Χανιά**

**Μάιος 2006**

### ***Ευχαριστίες***

Ολοκληρώνοντας την παρούσα μεταπτυχιακή διατριβή θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή Γεώργιο Καρατζά για την πολύτιμη συνεργασία και το ενδιαφέρον που επέδειξε κατά τη διάρκεια εκπόνησης της εν λόγω διατριβής. Οφείλω επίσης να ευχαριστήσω τον κ. Ιωάννη Νικολό, λέκτορα του τμήματος Μηχανικών Παραγωγής και Διοίκησης του Πολυτεχνείου Κρήτης για την παραχώρηση της μεθοδολογίας που ο ίδιος έχει αναπτύξει, για την προθυμία του καθώς και για την προσφορά βοήθειας όποτε του ζητήθηκε.

## ΠΕΡΙΛΗΨΗ

Στην παρούσα εργασία συνδυάζονται η χρήση Νευρωνικού Δικτύου και Διαφορικού Εξελικτικού Αλγόριθμου με σκοπό την εύρεση των μέγιστων παροχών άντλησης από γεωτρήσεις που βρίσκονται στην περιοχή του βόρειου τμήματος της νήσου Ρόδου, με τον περιορισμό το υδραυλικό ύψος να μην πέφτει κάτω από ορισμένη τιμή σε συγκεκριμένα πηγάδια παρατήρησης. Το πρόβλημα είναι ένα κλασικό πρόβλημα βελτιστοποίησης υπό περιορισμούς, οι οποίοι εισάγονται με την μορφή όρων ποινής στην συνάρτηση κόστους, που πρέπει να ελαχιστοποιηθεί. Τα υδραυλικά ύψη στα πηγάδια παρατήρησης υπολογίζονται από Νευρωνικό Δίκτυο, το οποίο εκπαιδεύεται βάσει των αποτελεσμάτων του τρισδιάστατου μοντέλου ροής υπογείων υδάτων και μεταφοράς ρύπων PTC. Η εργασία ξεκινά παρουσιάζοντας κάποια γενικά στοιχεία της περιοχής μελέτης, ενώ στη συνέχεια παρουσιάζονται κάποιες βασικές έννοιες για την κατανόηση των Νευρωνικών Δικτύων και περιγράφονται ο Perceptron, ο Adaline και ο κανόνας πίσω διάδοσης του σφάλματος που χρησιμοποιείται για την προσαρμογή των συναπτικών βαρών του δικτύου (δηλαδή την εκπαίδευσή του ούτως ώστε εισάγοντας σε αυτό κάποια δεδομένα εισόδου να παράγει τα επιθυμητά - και ήδη γνωστά - δεδομένα εξόδου). Στη συνέχεια, αναφέρονται τα προβλήματα που προκύπτουν κατά το σχεδιασμό ενός τέτοιου δικτύου και με ποιους τρόπους μπορεί αυτό να βελτιωθεί και επεξηγείται γιατί προτιμάται η χρήση Νευρωνικού Δικτύου αντί ενός φυσικού μοντέλου προσομοίωσης για το πρόβλημα βελτιστοποίησης. Ακολουθούν η περιγραφή του χρησιμοποιούμενου Νευρωνικού Δικτύου και μια σύντομη αναφορά στους Εξελικτικούς Αλγόριθμους και στο υπό βελτιστοποίηση πρόβλημα. Τέλος, παρατίθενται τα αποτελέσματα της βελτιστοποίησης βάσει δύο σεναρίων που χρησιμοποιήθηκαν και που διαφέρουν ως προς τα ελάχιστα επιτρεπόμενα υδραυλικά ύψη σε κάποια πηγάδια παρατήρησης. Στα συμπεράσματα διαπιστώνεται ότι η χρήση του Νευρωνικού Δικτύου για την μοντελοποίηση της συμπεριφοράς του υδροφόρου ορίζοντα επιτρέπει την επίλυση απεριόριστων προβλημάτων βελτιστοποίησης με διαφορετικούς περιορισμούς (διαφορετικά σενάρια) σε πολύ μικρό χρόνο, πράγμα που θα ήταν απαγορευτικό με χρήση του μοντέλου PTC λόγω του συνολικού χρόνου που απαιτεί για την προσομοίωση. Έτσι, το Νευρωνικό Δίκτυο που χρησιμοποιήθηκε αποδεικνύεται καλύτερο και ταχύτερο υποκατάστατο του μοντέλου PTC.

## ΠΕΡΙΕΧΟΜΕΝΑ

<b>ΚΕΦΑΛΑΙΟ 1 – ΕΙΣΑΓΩΓΗ.....</b>	<b>7</b>
1.1 Αντικείμενο της Εργασίας .....	7
1.2 Σύντομη Περιγραφή Περιοχής Μελέτης .....	8
1.3 Διαδικασία Μοντελοποίησης με χρήση του Μοντέλου PTC .....	11
<b>ΚΕΦΑΛΑΙΟ 2 – ΕΙΣΑΓΩΓΗ ΣΤΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ.....</b>	<b>16</b>
2.1 Εισαγωγή .....	16
2.2 Από τα Βιολογικά Νευρωνικά Δίκτυα στα Τεχνητά .....	17
2.3 Πλεονεκτήματα Νευρωνικών Δικτύων.....	18
<b>ΚΕΦΑΛΑΙΟ 3 – ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ ΓΙΑ ΤΗΝ ΚΑΤΑΝΟΗΣΗ ΤΩΝ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ .....</b>	<b>20</b>
3.1 Βασικές Συνιστώσες Νευρωνικών Δικτύων .....	20
3.1.1 Μονάδες Επεξεργασίας.....	21
3.1.2 Υπολογισμός Εισόδου και Εξόδου Μονάδας .....	22
3.2 Κατηγορίες Δικτύων Βάσει της Διαμόρφωσής τους .....	23
3.3 Εκπαίδευση Νευρωνικών Δικτύων .....	23
<b>ΚΕΦΑΛΑΙΟ 4 – ΠΕΡΙΓΡΑΦΗ PERCEPTRON ΚΑΙ ADALINE .....</b>	<b>25</b>
4.1 Περιγραφή του Perceptron.....	25
4.1.1 Εισαγωγή .....	25
4.1.2 Κανόνας Εκμάθησης του Perceptron και Θεώρημα Σύγκλισης .....	27
4.2 Περιγραφή του Adaline .....	28
4.2.1 Εισαγωγή .....	28
4.2.2 Κανόνας Εκμάθησης του Adaline .....	29
4.3 Το Πρόβλημα της Exclusive-Or (XOR) .....	30
<b>ΚΕΦΑΛΑΙΟ 5 – ΚΑΝΟΝΑΣ ΠΙΣΩ ΔΙΑΔΟΣΗΣ .....</b>	<b>34</b>
5.1 Εισαγωγή .....	34
5.2 Γενικευμένος Κανόνας Δέλτα - Μαθηματικές Εξισώσεις και Επεξήγηση .....	37
5.3 Ρυθμός Εκμάθησης .....	40
5.4 Τρόποι Εκπαίδευσης του Δικτύου .....	41
5.5 Κριτήρια Σύγκλισης.....	42
5.6 Προβλήματα κατά την Εφαρμογή του Αλγόριθμου Πίσω Διάδοσης.....	43

<b>ΚΕΦΑΛΑΙΟ 6 – ΠΡΟΒΛΗΜΑΤΑ ΣΧΕΔΙΑΣΜΟΥ ΚΑΙ ΒΕΛΤΙΩΣΗ ΠΟΛΥΕΠΙΠΕΔΩΝ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ .....</b>	<b>45</b>
6.1 Καθορισμός Δομής Δικτύου .....	45
6.2 Σφάλματα Προσομοίωσης .....	45
6.2.1 Επίδραση του Αριθμού των Εκπαιδευτικών Προτύπων .....	46
6.2.2 Επίδραση του Αριθμού των Κρυφών Μονάδων .....	48
6.3 Τρόποι Βελτίωσης της Προσομοίωσης .....	50
<b>ΚΕΦΑΛΑΙΟ 7 – ΧΡΗΣΗ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ ΣΤΗΝ ΠΡΟΣΕΓΓΙΣΗ ΦΥΣΙΚΩΝ ΜΟΝΤΕΛΩΝ ΚΑΙ ΣΤΗ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗ ΠΕΡΙΒΑΛΛΟΝΤΙΚΩΝ ΠΡΟΒΛΗΜΑΤΩΝ .....</b>	<b>52</b>
7.1 Εισαγωγή .....	52
7.2 Προσέγγιση Συμβατικών Μοντέλων Προσομοίωσης με Χρήση Νευρωνικών Δικτύων .....	52
7.3 Χρήση Νευρωνικών Δικτύων στη Βελτιστοποίηση Προβλημάτων Μηχανικής Υπογείων Υδάτων .....	53
7.4 Συμπεράσματα .....	54
<b>ΚΕΦΑΛΑΙΟ 8 – ΧΡΗΣΙΜΟΠΟΙΟΥΜΕΝΟ ΝΕΥΡΩΝΙΚΟ ΔΙΚΤΥΟ.....</b>	<b>55</b>
8.1 Περιγραφή Νευρωνικού Δικτύου .....	55
8.2 Εκπαίδευση Νευρωνικού Δικτύου .....	56
<b>ΚΕΦΑΛΑΙΟ 9 – ΕΙΣΑΓΩΓΗ ΣΤΟΥΣ ΕΞΕΛΙΚΤΙΚΟΥΣ ΑΛΓΟΡΙΘΜΟΥΣ....</b>	<b>59</b>
9.1 Ιστορική Αναδρομή και Αναλογία με Βιολογία .....	59
9.2 Λειτουργία Εξελικτικών Αλγορίθμων .....	61
9.3 Μέθοδοι Βελτιστοποίησης και Εξελικτικοί Αλγόριθμοι .....	62
9.4 Ο Διαφορικός Εξελικτικός Αλγόριθμος .....	64
9.4.1 Εισαγωγή .....	64
9.4.2 Λειτουργία Διαφορικού Εξελικτικού Αλγόριθμου .....	64
<b>ΚΕΦΑΛΑΙΟ 10 – ΤΟ ΠΡΟΒΛΗΜΑ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ .....</b>	<b>67</b>
10.1 Σκοπός του Προβλήματος Βελτιστοποίησης .....	67
10.2 Κατασκευή Συνάρτησης Προσαρμογής .....	69
10.3 Εφαρμογή του Διαφορικού Εξελικτικού Αλγόριθμου στο Υπό Μελέτη Πρόβλημα .....	71
<b>ΚΕΦΑΛΑΙΟ 11 – ΑΠΟΤΕΛΕΣΜΑΤΑ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ .....</b>	<b>78</b>
11.1 Πρώτο Σενάριο .....	78
11.2 Δεύτερο Σενάριο .....	81

<b>ΚΕΦΑΛΑΙΟ 12 – ΑΝΑΛΥΣΗ ΑΠΟΤΕΛΕΣΜΑΤΩΝ ΚΑΙ ΣΥΜΠΕΡΑΣΜΑΤΑ</b>	
.....	<b>86</b>
<b>ΒΙΒΛΙΟΓΡΑΦΙΑ</b>	<b>94</b>

**ΠΑΡΑΡΤΗΜΑ Ι: ΧΑΡΤΕΣ ΠΕΡΙΟΧΗΣ ΜΕΛΕΤΗΣ**

**ΠΑΡΑΡΤΗΜΑ ΙΙ: ΚΩΔΙΚΑΣ ΣΕ C++ ΓΙΑ ΤΗ ΔΗΜΙΟΥΡΓΙΑ ΤΩΝ  
ΔΥΝΑΤΩΝ ΣΥΝΔΥΑΣΜΩΝ ΠΑΡΟΧΩΝ ΑΝΤΛΗΣΗΣ**

**ΠΑΡΑΡΤΗΜΑ ΙΙΙ: ΚΩΔΙΚΑΣ ΣΕ VISUAL BASIC ΓΙΑ ΤΗΝ ΚΑΤΑΣΚΕΥΗ  
ΤΗΣ ΣΥΝΑΡΤΗΣΗΣ ΠΡΟΣΑΡΜΟΓΗΣ**

## ΚΕΦΑΛΑΙΟ 1 – ΕΙΣΑΓΩΓΗ

### 1.1 Αντικείμενο της Εργασίας

Το πρόβλημα που αντιμετωπίζεται στην παρούσα εργασία είναι ο προσδιορισμός του βέλτιστου σχεδιασμού διαχείρισης και συγκεκριμένα η εύρεση των μέγιστων παροχών άντλησης από γεωτρήσεις που βρίσκονται στην περιοχή του βόρειου τμήματος της νήσου Ρόδου, με τον περιορισμό το υδραυλικό ύψος να μην πέφτει κάτω από ορισμένη τιμή σε συγκεκριμένα πηγάδια παρατήρησης. Το πρόβλημα είναι ένα κλασικό πρόβλημα βελτιστοποίησης υπό περιορισμούς. Το πρόβλημα βελτιστοποίησης επιλύεται με τη χρήση Διαφορικού Εξελικτικού αλγορίθμου, ενώ οι περιορισμοί του προβλήματος εισάγονται με την μορφή όρων ποινής στην συνάρτηση κόστους, που πρέπει να ελαχιστοποιηθεί. Για κάθε υποψήφια λύση (τιμές παροχών) απαιτείται η γνώση των υδραυλικών υψών στα πηγάδια παρατήρησης. Αυτό επιτυγχάνεται με τη χρήση κατάλληλα εκπαιδευμένου Τεχνητού Νευρωνικού Δικτύου, το οποίο, δρώντας ως προσεγγιστικό μοντέλο, συνεργάζεται με τον Διαφορικό Εξελικτικό αλγόριθμο για την εύρεση της βέλτιστης λύσης.

Η εκπαίδευση του νευρωνικού δικτύου γίνεται με βάση αποτελέσματα του μοντέλου προσομοίωσης PTC για μεγάλο αριθμό προσομοιώσεων. Ο αναγνώστης παραπέμπεται στη διπλωματική εργασία της συγγραφέως (Στεργιάδη, 2004) για αναλυτικά στοιχεία και πληροφορίες σχετικά με

- την αειφόρο διαχείριση των υδατικών πόρων και τα βήματα που έχουν γίνει προς την κατεύθυνση αυτή
- το κοινοτικό και ελληνικό θεσμικό πλαίσιο, που διέπει την προστασία και διαχείριση των υδάτων
- την περιοχή μελέτης και τις γεωλογικές, υδρολογικές και υδρογεωλογικές συνθήκες που τη χαρακτηρίζουν
- τη μοντελοποίηση της κίνησης των υπογείων υδάτων με χρήση αριθμητικών μεθόδων
- το λογισμικό Argus One GIS και το μοντέλο PTC που χρησιμοποιήθηκαν για την προσομοίωση
- τη διαδικασία μοντελοποίησης με τη χρήση του μοντέλου PTC.

Στο παράρτημα Ι παρατίθενται γεωλογικός και τοπογραφικός χάρτης, καθώς και χάρτης απογραφής των σημείων εμφάνισης ύδατος.

## 1.2 Σύντομη Περιγραφή Περιοχής Μελέτης

### *Γενικά στοιχεία*

Η Ρόδος είναι πρωτεύουσα του Νομού Δωδεκανήσου (Εικόνα 1.1). Έχει έκταση 1.400 km<sup>2</sup> περίπου και πληθυσμό 117.000 κατοίκων (απογραφή του 2001), στοιχεία που την καθιστούν το μεγαλύτερο σε έκταση και πληθυσμό νησί του Νομού. Η οικονομία της βασίζεται κατά κύριο λόγο στον τουρισμό, η προσέλκυση του οποίου οφείλεται στον συνδυασμό των ποικίλων πολιτισμικών στοιχείων που έχει να επιδείξει το νησί και στο ιδιαίτερης οικολογικής και επιστημονικής αξίας φυσικό περιβάλλον του.

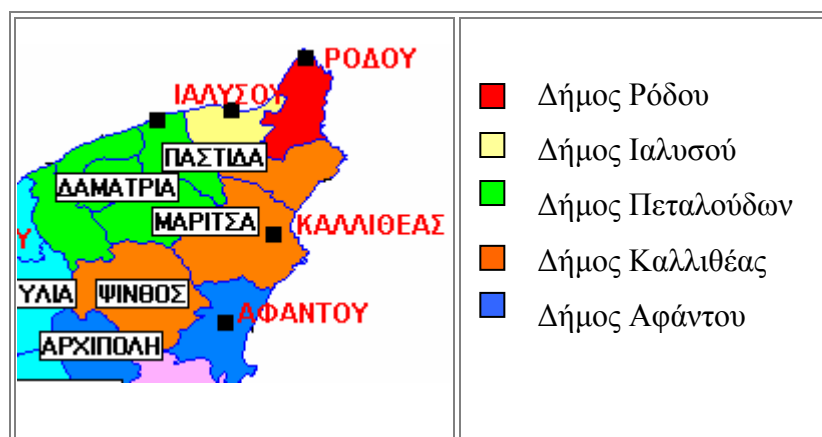


**Εικόνα 1.1:** Νομός Δωδεκανήσου (Στεργιάδη, 2004)

Το νησί δεν αντιμετωπίζει πρόβλημα λειψυδρίας. Η ύδρευση εξασφαλίζεται κυρίως με χρήση γεωτρήσεων, ενώ η ανεπάρκεια δικτύου αποχέτευσης δημιουργεί σημαντικά προβλήματα (υπερχείλιση βόθρων, παράνομες συνδέσεις στο δίκτυο ομβρίων κ.ά.). Στο νησί λειτουργεί χώρος υγειονομικής ταφής απορριμμάτων ενώ το μεγαλύτερο μέρος του οδικού δικτύου αποτελείται κυρίως από επαρχιακές παρά εθνικές οδούς (Στεργιάδη, 2004).



Το βόρειο τμήμα του νησιού που αποτελεί αντικείμενο μελέτης της παρούσας εργασίας περιλαμβάνει τους Δήμους Ρόδου, Ιαλυσού, Πεταλούδων, Καλλιθέας και τμήμα του Δήμου Αφάντου (Εικόνα 1.2).



Εικόνα 1.2: Δήμοι περιοχής μελέτης (Στεργιάδη, 2004)

#### Στοιχεία γεωμορφολογίας – γεωλογίας

Η Ρόδος χαρακτηρίζεται από ανώμαλο τοπογραφικό ανάγλυφο, ενώ οι κύριες τοπογραφικές μονάδες που απαντώνται είναι οι ασβεστολιθικοί όγκοι, οι λοφώδεις περιοχές των τριτογενών κλαστικών σχηματισμών, τα ανωπλειοκαινικά υψίπεδα, οι θαλάσσιες και ποτάμιες αναβαθμίδες και τα παράκτια επίπεδα. Όπως φαίνεται από το χάρτη ισοϋψών των απόλυτων υψομέτρων της περιοχής μελέτης (Εικόνα 1.3), που δημιουργήθηκε με χρήση του μοντέλου PTC, το νησί χαρακτηρίζεται από τοπογραφικές εξάρσεις, που οφείλονται σε δομές τεκτονικών κεράτων και τεκτονικών τάφρων.

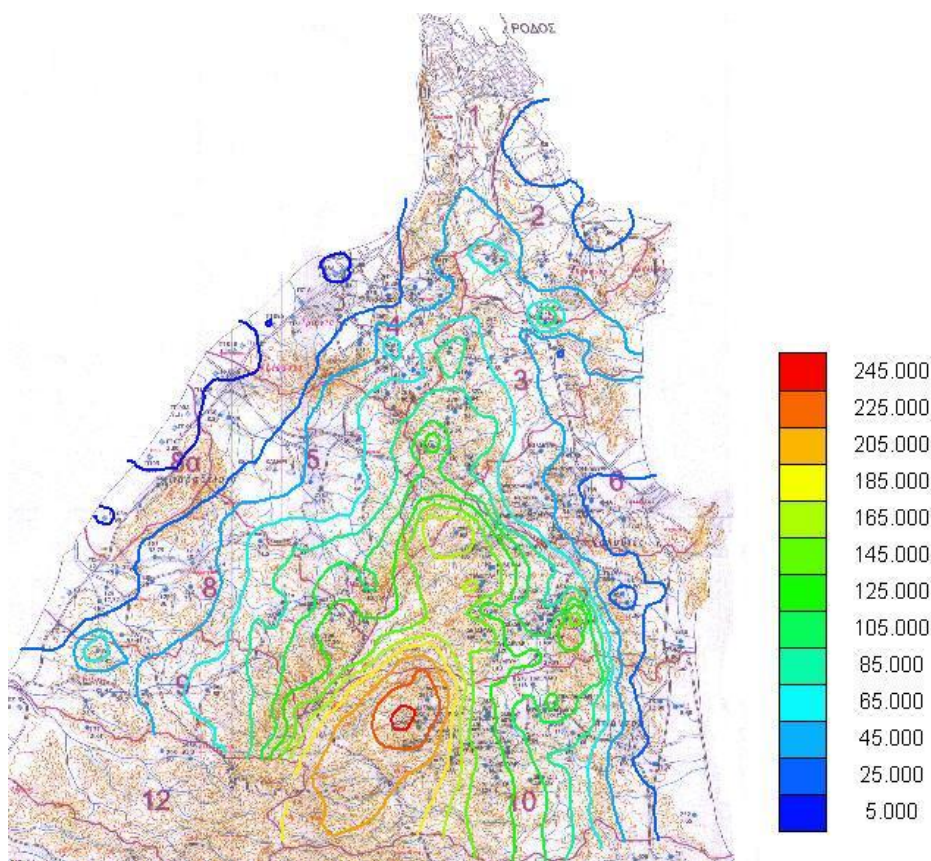
Στην περιοχή παρατηρείται ένας κύριος υδροκρίτης μεταβλητής διεύθυνσης, που τη χωρίζει σε βορειοδυτικό και νοτιοανατολικό τμήμα, ενώ το υδρογραφικό δίκτυο θεωρείται δενδριτικού τύπου (λόγω της ακανόνιστης διακλάδωσης των παραποτάμων προς διάφορες κατευθύνσεις).

Από γεωλογική άποψη, στο νησί εντοπίζονται τρεις μεγάλες τεκτονικές ενότητες, που χαρακτηρίζονται από σαφώς καθορισμένη στρωματογραφική διαδοχή. Οι ενότητες αυτές είναι:

- η κατώτερη ενότητα – αυτόχθονη, που αποτελείται από την ομάδα Αταβύρου και τον φλύσχη Κατταβιάς

- η ενδιάμεση ενότητα – αλλόχθονη, που αποτελείται από την ομάδα Αρχαγγέλου, τον φλύσχη Αρχίπολης, τον ασβεστόλιθο Λίνδου και την ομάδα Προφήτη Ηλία
- η ανώτερη ενότητα – νεοαυτόχθονη, που αποτελείται από την ομάδα Βατί, τα «λεβαντίνια» ιζήματα, τον σχηματισμό Σγουρού και τον Πώρο.

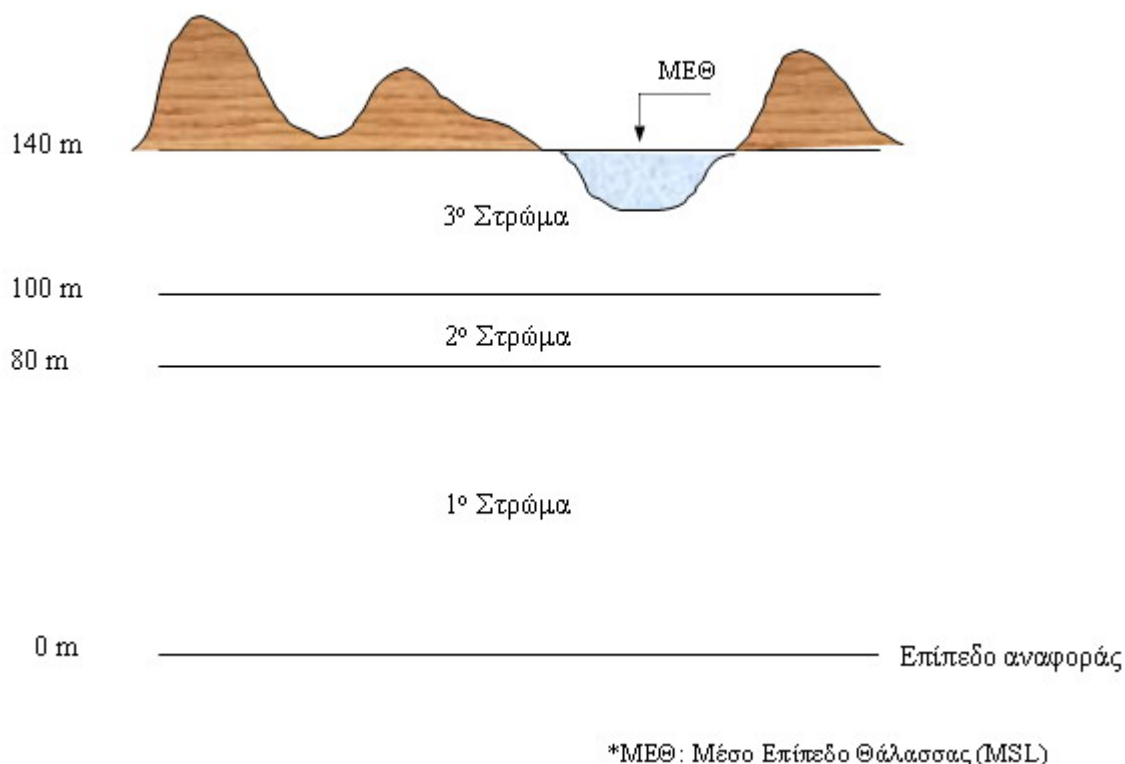
Ακόμη, στην περιοχή εντοπίζονται τρία βασικά συστήματα γεωλογικών δομών: ένα σύστημα πτυχώσεων, πτυχωσιγενών ρηγμάτων, καλυμμάτων και επιπεύσεων, ένα σύστημα ελαφρών πτυχώσεων και ένα σύστημα κανονικών ρηγμάτων, που συνιστούν το εντονότερο δομικό και μορφολογικό χαρακτηριστικό του νησιού και αποτελούν αιτία μερικής καταστροφής των προηγούμενων δομών (Στεργιάδη, 2004).



**Εικόνα 1.3:** Χάρτης ισοϋψών των απόλυτων υψομέτρων (m) της υπό μελέτη περιοχής (Στεργιάδη, 2004)

### 1.3 Διαδικασία Μοντελοποίησης με χρήση του Μοντέλου PTC

Το μοντέλο PTC χρησιμοποιεί για τη μοντελοποίηση τη μέθοδο των πεπερασμένων στοιχείων, με τα υδραυλικά ύψη να υπολογίζονται στους κόμβους του πλέγματος. Η περιοχή μελέτης διαχωρίζεται σε τρία στρώματα, ούτως ώστε να υπάρχει διαχωρισμός ανάμεσα στις βαθιές και ρηχές γεωτρήσεις (Εικόνα 1.4).



**Εικόνα 1.4:** Στρωματοποίηση περιοχής μελέτης (Στεργιάδη, 2004)

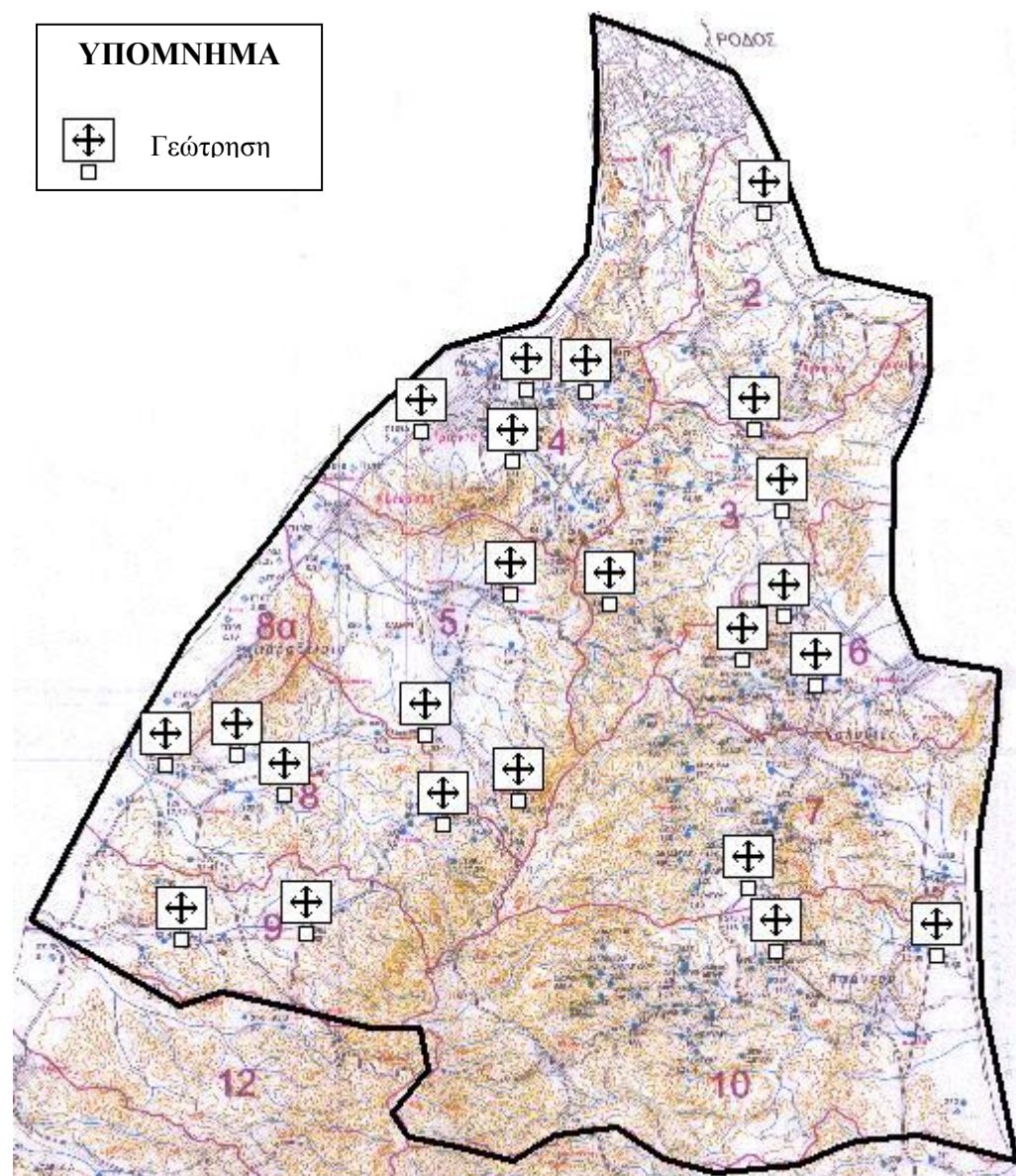
Όπως φαίνεται στην εικόνα 1.5, η περιοχή μελέτης αποτελείται από τις λεκάνες 1 έως 10. Οι γεωτρήσεις της περιοχής, για τις οποίες υπάρχουν βιβλιογραφικά δεδομένα της στάθμης του ύδατος, εισήχθησαν ανάλογα με το βάθος της σωλήνωσης σε ένα από τα τρία στρώματα και θεωρούνται οριακές συνθήκες 2<sup>ου</sup> είδους (σταθερή παροχή άντλησης) (Εικόνα 1.5). Από αυτές, επιλέγονται μόνο πέντε γεωτρήσεις για την εκπαίδευση του νευρωνικού δικτύου και τη διαδικασία βελτιστοποίησης. Ως αρχικά υδραυλικά ύψη στις γεωτρήσεις τίθενται τα ύψη του νερού τον Μάιο του 1997, όπως αυτά δίνονται από το Υπουργείο Γεωργίας, Γενική Διεύθυνση Εγγειοβελτιωτικών Έργων και Γεωργικών Διαρθρώσεων, Δ/νση Γεωλογίας – Υδρολογίας, Τμήμα Υδρογεωλογίας, Γεωτρήσεων και Μαθηματικών Ομοιωμάτων (1999).

Η ακτογραμμή της περιοχής θεωρείται οριακή συνθήκη 1<sup>ου</sup> είδους με σταθερό υδραυλικό ύψος ίσο με 140 μέτρα (Εικόνα 1.6), ενώ στο υπόλοιπο τμήμα της περιμέτρου εφαρμόζονται οριακές συνθήκες 3<sup>ου</sup> είδους (ή διαρροής) (Εικόνα 1.7).

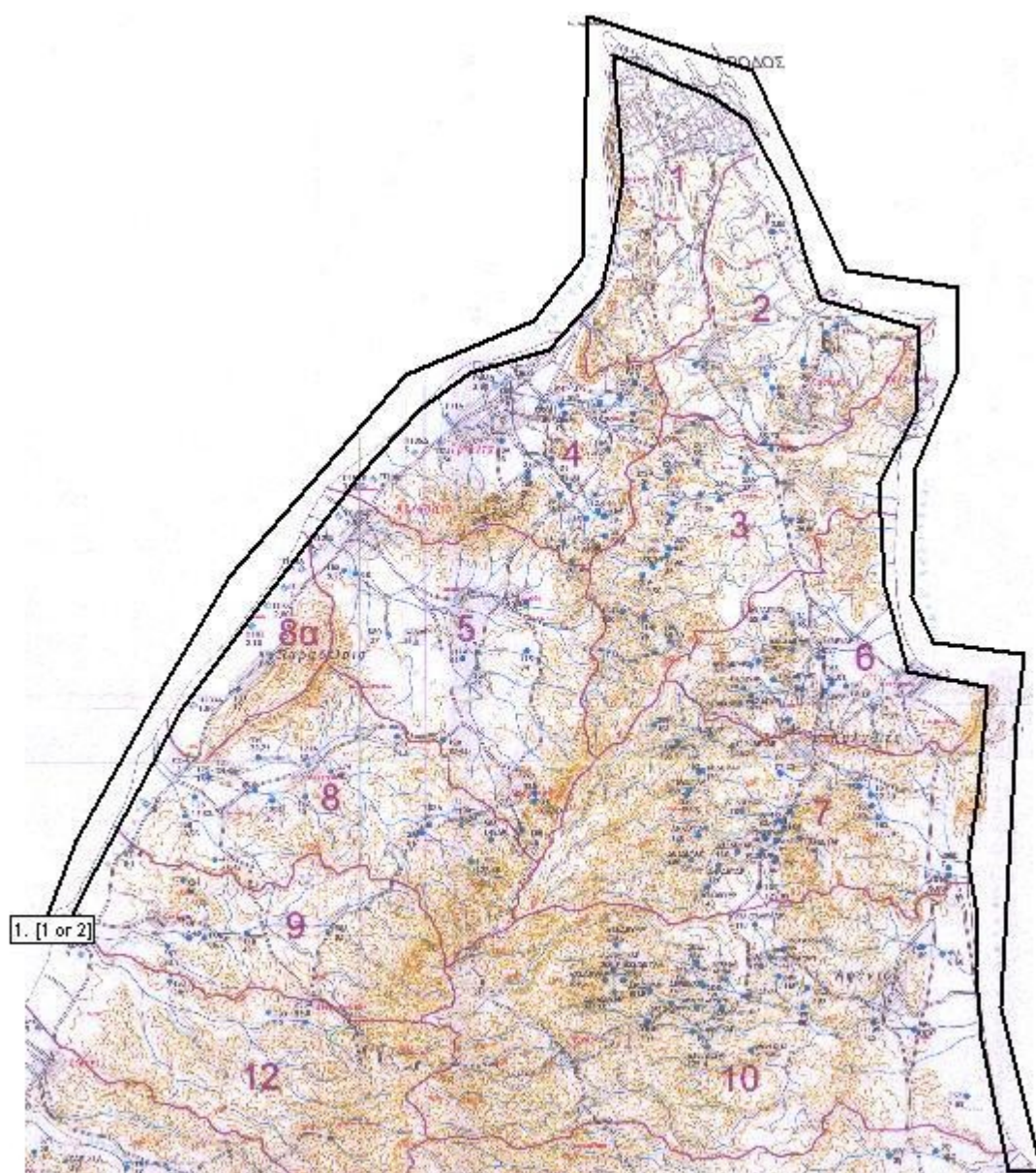
Η βαθμονόμηση του μοντέλου γίνεται συγκρίνοντας τις τιμές του υδραυλικού ύψους τον Μάιο του 1998 (τιμές πεδίου), που δίνονται από το Υπουργείο Γεωργίας, Γενική Διεύθυνση Εγγειοβελτιωτικών Έργων και Γεωργικών Διαρθρώσεων, Δ/ση Γεωλογίας – Υδρολογίας, Τμήμα Υδρογεωλογίας, Γεωτρήσεων και Μαθηματικών Ομοιωμάτων (1999), με τα αποτελέσματα του μοντέλου για τη χρονική περίοδο Μάιος 1997 – Μάιος 1998. Η καλύτερη δυνατή προσέγγιση των τιμών του μοντέλου με τις τιμές του πεδίου εξασφαλίζεται μεταβάλλοντας τόσο τους λόγους  $K/L$  που σχετίζονται με τις οριακές συνθήκες 3<sup>ου</sup> είδους ( $K$  η υδραυλική αγωγιμότητα,  $L$  η απόσταση του σημείου που λαμβάνεται το υδραυλικό ύψος αναφοράς από το όριο), όσο και τις παροχές άντλησης, όπου αυτό ήταν αναγκαίο.

Οι διαδικασίες προετοιμασίας και βαθμονόμησης του μοντέλου παρουσιάζονται αναλυτικότερα στη διπλωματική εργασία της συγγραφέως (Στεργιάδη, 2004).



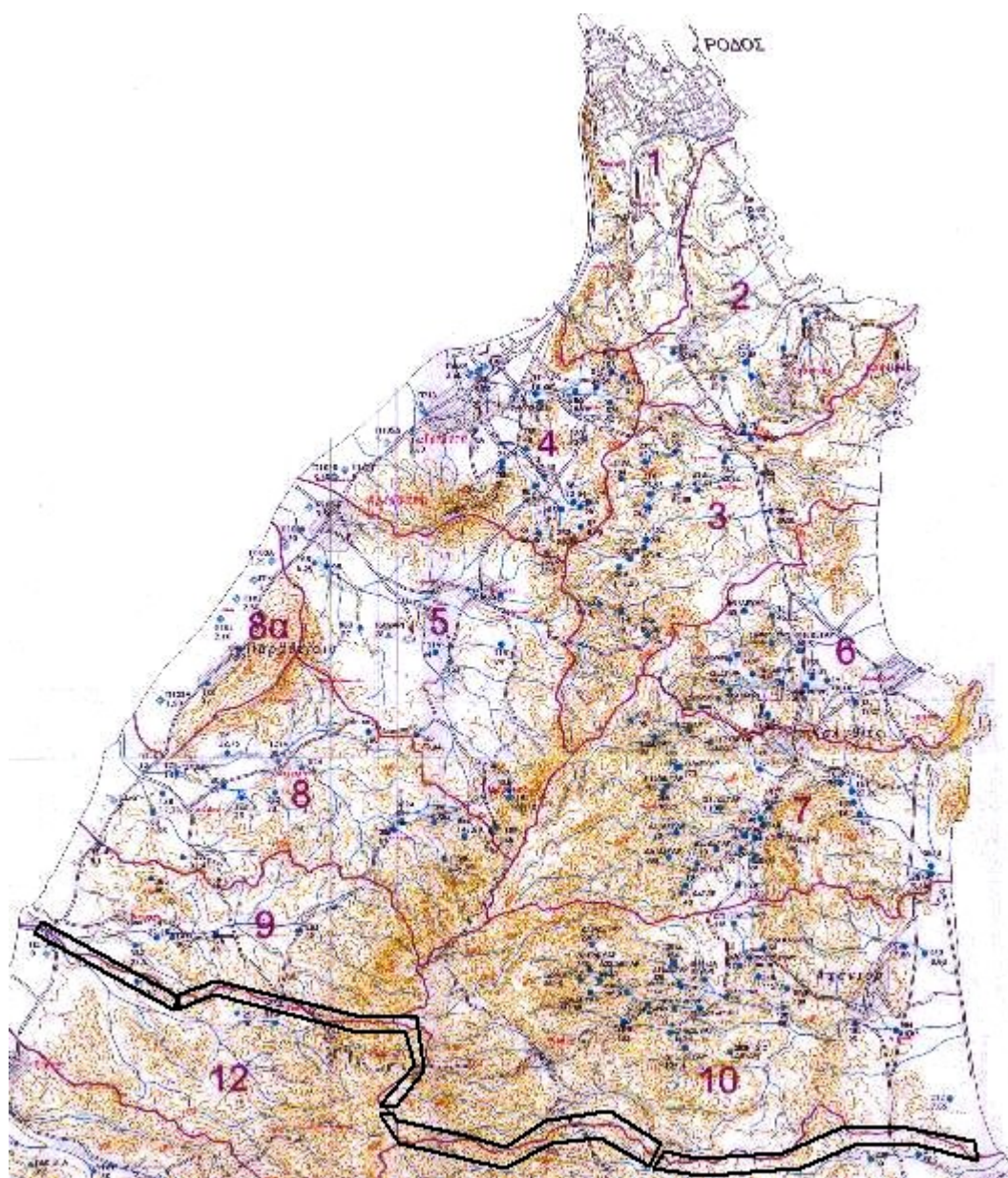


**Εικόνα 1.5:** Περίγραμμα και γεωτρήσεις περιοχής μελέτης, που χρησιμοποιήθηκαν στη διπλωματική εργασία της συγγραφέας για την προσομοίωση της υπόγειας ροής με το μοντέλο PTC (Στεργιάδη, 2004)



**Εικόνα 1.6:** Όριο επιβολής οριακής συνθήκης 1<sup>ο</sup> είδους (Στεργιάδη, 2004)





**Εικόνα 1.7:** Όρια επιβολής οριακών συνθηκών 3<sup>ου</sup> είδους (Στεργιάδη, 2004)

## ΚΕΦΑΛΑΙΟ 2 – ΕΙΣΑΓΩΓΗ ΣΤΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ

### 2.1 Εισαγωγή

Τα τελευταία χρόνια παρατηρείται έντονο το ενδιαφέρον επιστημόνων και μηχανικών για αλγόριθμους που προσομοιώνουν διαδικασίες ανάλογες με τις φυσικές. Η εμφάνιση των μαζικά παράλληλων κατανεμημένων επεξεργαστών (Parallel Distributed Processing, PDP) κατέστησε δυνατή την υλοποίηση τέτοιου είδους αλγόριθμων, που πλέον περιλαμβάνουν τα νευρωνικά δίκτυα, τους γενετικούς αλγόριθμους, τον εξελικτικό προγραμματισμό, τις εξελικτικές στρατηγικές, τις μεθόδους τοπικής αναζήτησης και τα συστήματα ταξινόμησης. Οι αλγόριθμοι αυτοί αναφέρονται στη διεθνή βιβλιογραφία ως αλγόριθμοι υπολογιστικής νοημοσύνης (computational intelligence algorithms).

Παρά τη ραγδαία πρόοδο που έχει σημειώσει η επιστήμη των υπολογιστών τα τελευταία χρόνια, δεν έχει ακόμα καταφέρει να επιτύχει την ουσιαστική επικοινωνία του ανθρώπου με τον υπολογιστή και την απαλλαγή του τελευταίου από τα κλασικά μειονεκτήματα της μηχανής, όπως είναι η ανικανότητα απόκτησης εμπειρίας, γνώσης από τα λάθη καθώς και διανοητικών ικανοτήτων ανάλογων του ανθρώπινου εγκεφάλου. Τα παραπάνω προβλήματα μελετώνται από τον ερευνητικό τομέα γνωστό ως Τεχνητή Νοημοσύνη.

Οι περιοχές έρευνας της Τεχνητής Νοημοσύνης που προσελκύουν το μεγαλύτερο ενδιαφέρον σήμερα είναι τα νευρωνικά δίκτυα (neural networks), οι γενετικοί αλγόριθμοι (genetic algorithms), τα έμπειρα συστήματα (expert systems), τα ασαφή συστήματα (fuzzy systems), οι ευφυείς πράκτορες (intelligent agents) και η μηχανική μάθηση (machine learning) (Λυκοθανάσης, Γεωργόπουλος, 2003).

Το πρώτο κύμα ενδιαφέροντος για τα νευρωνικά δίκτυα εμφανίστηκε μετά από την εισαγωγή της έννοιας του απλού νευρώνα το 1943 από τους McCulloch και Pitts (McCulloch, Pitts, 1943), οι οποίοι παρουσίασαν τους τεχνητούς νευρώνες ως μοντέλα βιολογικών νευρώνων και ως εννοιολογικές συνιστώσες υπολογιστικών βρόγχων.

Το 1969, μετά τη δημοσίευση του βιβλίου «Perceptrons» των Minsky και Papert (Minsky, Papert, 1969), που καταδείκνυε την ανεπάρκεια των perceptrons, πολλοί ερευνητές εγκατέλειψαν τον τομέα των νευρωνικών δικτύων, για να επανέλθουν στις



αρχές της δεκαετίας του '80, όταν πια είχαν προκύψει σημαντικά θεωρητικά αποτελέσματα και κατασκευαστεί ηλεκτρονικοί υπολογιστές με αυξημένες δυνατότητες επεξεργασίας.

Η μέχρι σήμερα εξέλιξη των τεχνητών νευρωνικών δικτύων έχει ως αποτέλεσμα αυτά να μπορούν να χαρακτηριστούν ως υπολογιστικά μοντέλα με ιδιαίτερες ικανότητες, όπως η ικανότητα να προσαρμόζονται ή να μαθαίνουν, να γενικεύουν και να συλλέγουν ή να οργανώνουν δεδομένα (Kröse, Smagt, 1996).

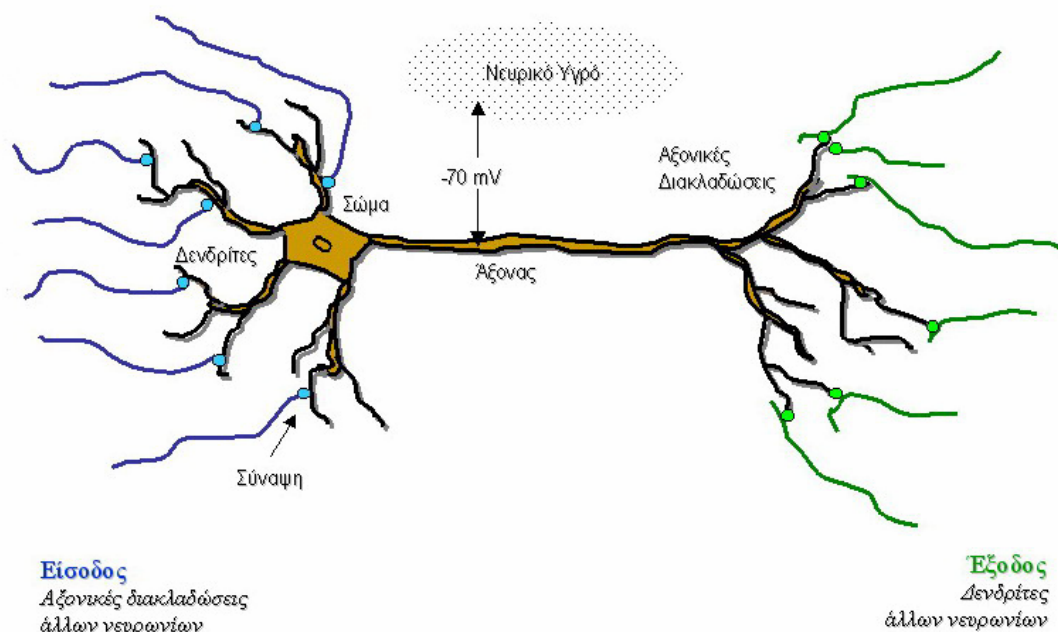
## **2.2 Από τα Βιολογικά Νευρωνικά Δίκτυα στα Τεχνητά**

Η συνειδητοποίηση ότι η λειτουργία του εγκεφάλου είναι εντελώς διαφορετική από τους συμβατικούς ψηφιακούς υπολογιστές άνοιξε το δρόμο για τη μελέτη των τεχνητών νευρωνικών δικτύων, που πλέον καλούνται απλώς νευρωνικά δίκτυα.

Ο εγκέφαλος προσομοιάζει με έναν περίπλοκο, μη γραμμικό, παράλληλο υπολογιστή που δύναται να οργανώνει τους νευρώνες και να εκτελεί υπολογισμούς ταχύτερα από οποιονδήποτε ψηφιακό υπολογιστή. Μπορεί να κατασκευάζει δικούς του κανόνες, δηλαδή να αποκτά εμπειρία, που μεγαλώνει με την πάροδο του χρόνου, ενώ η μέγιστη ανάπτυξή του παρατηρείται κατά τα δύο πρώτα χρόνια της ζωής. Την περίοδο αυτή δημιουργούνται γύρω στις ένα εκατομμύριο συνάψεις το δευτερόλεπτο, οι οποίες αποτελούν τις βασικές λειτουργικές και δομικές μονάδες που συντελούν στην ενδοεπικοινωνία των νευρώνων (Εικόνα 2.1).

Με βάση τα παραπάνω, το νευρωνικό δίκτυο μπορεί να οριστεί ως ένας συμπαγής, παράλληλος, κατανεμημένος επεξεργαστής, που μπορεί να αποθηκεύει γνώση και να τη διαθέτει για χρήση (Λυκοθανάσης, Γεωργόπουλος, 2003).

Τα τεχνητά νευρωνικά δίκτυα παρουσιάζουν δύο κοινά σημεία με τον εγκέφαλο: η γνώση αποκτάται από το δίκτυο μέσα από διαδικασία μάθησης και οι δυνάμεις σύνδεσης των νευρώνων, που καλούνται συναπτικά βάρη, χρησιμοποιούνται για την αποθήκευση γνώσης (Λυκοθανάσης, Γεωργόπουλος, 2003).



Εικόνα 2.1: Το νευρικό κύτταρο (Σπύρου, 1997)

### 2.3 Πλεονεκτήματα Νευρωνικών Δικτύων

Η ικανότητα των νευρωνικών δικτύων να επιλύουν περίπλοκα προβλήματα, η οποία οφείλεται στην παράλληλα κατανομημένη δομή τους και στη δυνατότητα εκπαίδευσής τους, τα καθιστά ένα ιδιαίτερα χρήσιμο εργαλείο. Τα σημαντικότερα πλεονεκτήματά τους είναι τα εξής (Λυκοθανάσης, Γεωργόπουλος, 2003):

- *Μη γραμμικότητα*, καθώς οι νευρώνες που δομούν το νευρωνικό δίκτυο είναι μη γραμμικοί.
- *Προσαρμοστικότητα*. Όταν το νευρωνικό δίκτυο διεγείρεται από το περιβάλλον, υφίσταται αλλαγές ως αποτέλεσμα αυτής της διέγερσης, δηλαδή μεταβάλλει τις ελεύθερες παραμέτρους του. Ωστόσο, μερικές φορές οι προσαρμογές αυτές μπορούν να μειώσουν την απόδοση του συστήματος.
- *Αποδεικτική απόκριση*. Τα νευρωνικά δίκτυα σχεδιάζονται για να παρέχουν πληροφορίες τόσο για το συγκεκριμένο πρότυπο που επιλέγεται όσο και για την εμπιστοσύνη στην απόφαση που λαμβάνεται, πράγμα που αυξάνει την απόδοση του συστήματος.
- *Παροχή συναφούς πληροφορίας*, αφού η γνώση αναπαρίσταται από τη δομημένη και ενεργή κατάσταση του νευρωνικού δικτύου.

- *Ανοχή σε σφάλματα.* Τα νευρωνικά δίκτυα είναι ανεκτικά σε σφάλματα, υπό την έννοια ότι κάτω από αντίξοες συνθήκες λειτουργίας η απόδοσή τους μειώνεται.
- *Δυνατότητα υλοποίησης με χρήση τεχνολογίας VLSI.* Τα VLSI αποτελούν ψηφιακά κυκλώματα, τα λεγόμενα Συστήματα Πολύ Υψηλής Ολοκλήρωσης.
- *Ομοιομορφία Ανάλυσης και Σχεδιασμού,* καθώς όλα τα πεδία στα οποία εφαρμόζονται τα νευρωνικά δίκτυα χρησιμοποιούν τους ίδιους συμβολισμούς.
- *Αναλογία με τη Νευροβιολογία,* η οποία προκύπτει λόγω του σχεδιασμού των νευρωνικών δικτύων σε αναλογία με τον ανθρώπινο εγκέφαλο. Για τους νευροβιολόγους, τα νευρωνικά δίκτυα αποτελούν αντικείμενο έρευνας για την εξήγηση νευροβιολογικών φαινομένων, ενώ για τους μηχανικούς, η νευροβιολογία αποτελεί φυτώριο νέων ιδεών για την επίλυση περίπλοκων προβλημάτων.

## ΚΕΦΑΛΑΙΟ 3 – ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ ΓΙΑ ΤΗΝ ΚΑΤΑΝΟΗΣΗ ΤΩΝ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ

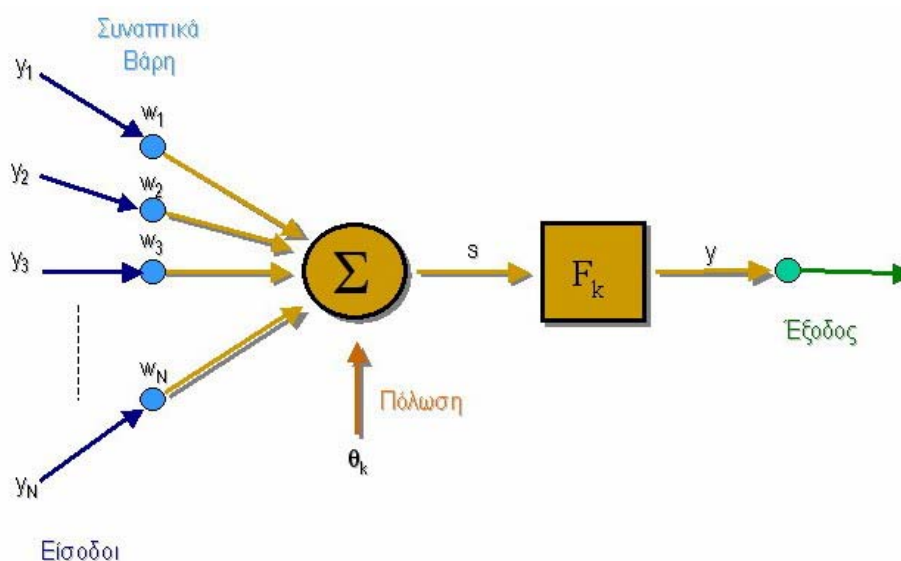
### 3.1 Βασικές Συνιστώσες Νευρωνικών Δικτύων

Τα τεχνητά νευρωνικά δίκτυα αποτελούνται από ένα σύνολο απλών μονάδων επεξεργασίας, που επικοινωνούν μεταξύ τους μέσω των σημάτων που στέλνονται διαμέσου ενός πλήθους σταθμισμένων συνδέσεων (Kröse, Smagt, 1996).

Ένα νευρωνικό δίκτυο αποτελείται από τα εξής στοιχεία (McClelland, Rumelhart, 1986):

- Ένα σύνολο μονάδων επεξεργασίας (οι λεγόμενοι «νευρώνες» ή «κελιά»).
- Μια κατάσταση ενεργοποίησης  $y_k$  για κάθε μονάδα, που αποτελεί και την έξοδο της μονάδας.
- Συνδέσεις μεταξύ των μονάδων, κάθε μία από τις οποίες χαρακτηρίζεται από ένα *συναπτικό βάρος* ή *σύναψη*  $w_{jk}$ , που προσδιορίζει την επίδραση που έχει το σήμα της μονάδας  $j$  στη μονάδα  $k$ .
- Έναν κανόνα διάδοσης (propagation rule), που καθορίζει την ενεργό είσοδο  $s_k$  στη μονάδα  $k$ .
- Μια *συνάρτηση ενεργοποίησης*  $F_k$ , που καθορίζει την ενημέρωση (update) της μονάδας  $k$ , δηλαδή το νέο επίπεδο ενεργοποίησής της, βάσει της ενεργού εισόδου  $s_k(t)$  και της τρέχουσας ενεργοποίησης  $y_k(t)$ .
- Μια *εξωτερική είσοδο*  $\theta_k$  για κάθε μονάδα, που καλείται κατώφλι ή πόλωση (bias) και μπορεί να θεωρηθεί σαν το συναπτικό βάρος μιας μονάδας με τιμή ενεργοποίησης ίση με 1.
- Έναν κανόνα εκμάθησης, μέσω του οποίου συλλέγονται οι πληροφορίες.
- Ένα περιβάλλον μέσα στο οποίο πρέπει να λειτουργήσει το σύστημα, παρέχοντας σήματα εισόδου και, αν χρειαστεί, σήματα λάθους.

Στην εικόνα 3.1 που ακολουθεί παρουσιάζονται οι βασικές συνιστώσες ενός νευρωνικού δικτύου.



**Εικόνα 3.1:** Βασικές συνιστώσες ενός τεχνητού νευρωνικού δικτύου (Σπύρου, 1997)

### 3.1.1 Μονάδες Επεξεργασίας

Οι μονάδες επεξεργασίας ενός νευρωνικού δικτύου έχουν διπλό ρόλο: να δέχονται κάποια είσοδο (input) από γειτονικές ή εξωτερικές πηγές, την οποία χρησιμοποιούν για να υπολογίσουν το σήμα εξόδου (output), το οποίο διαδίδεται στις άλλες μονάδες, καθώς και να ρυθμίζουν τα συναπτικά βάρη.

Οι μονάδες που περιέχονται σε ένα νευρωνικό δίκτυο διακρίνονται σε τρεις κατηγορίες:

- Μονάδες εισόδου, που δέχονται δεδομένα από πηγές που βρίσκονται εκτός του νευρωνικού δικτύου.
- Μονάδες εξόδου, που στέλνουν δεδομένα έξω από το νευρωνικό δίκτυο.
- Κρυφές μονάδες, που στέλνουν σήματα εισόδου και εξόδου εντός του δικτύου.

Καθώς το νευρωνικό δίκτυο λειτουργεί, οι μονάδες μπορούν να ενημερώνονται σύγχρονα (ταυτόχρονη ενημέρωση της ενεργοποίησης των μονάδων) ή ασύγχρονα (υπάρχει κάποια πιθανότητα οι μονάδες να ενημερώσουν την ενεργοποίησή τους τη χρονική στιγμή  $t$ , αλλά συνήθως μόνο μία μονάδα μπορεί να το κάνει αυτό κάποια δεδομένη χρονική στιγμή) (Kröse, Smagt, 1996).

### 3.1.2 Υπολογισμός Εισόδου και Εξόδου Μονάδας

Συνήθως θεωρείται ότι κάθε μονάδα συνεισφέρει αθροιστικά στην είσοδο της μονάδας με την οποία συνδέεται. Έτσι, η ολική είσοδος στη μονάδα  $k$  ισούται με το σταθμισμένο άθροισμα των εξόδων των μονάδων με τις οποίες συνδέεται, συν έναν όρο πόλωσης  $\theta_k$  (bias). Δηλαδή

$$s_k(t) = \sum_j w_{jk}(t) y_j(t) + \theta_k(t) \quad (3.1)$$

Για θετικές τιμές  $w_{jk}$ , η συνεισφορά θεωρείται *διεγερτική* ενώ για αρνητικές τιμές  $w_{jk}$ , η συνεισφορά θεωρείται *ανασταλτική*.

Οι μονάδες που ακολουθούν τον κανόνα διάδοσης της εξίσωσης (3.1) καλούνται *σίγμα μονάδες*.

Η επίδραση της ολικής εισόδου στην ενεργοποίηση της μονάδας καθορίζεται από τη συνάρτηση ενεργοποίησης  $F_k$  η οποία, λαμβάνοντας υπόψη την ενεργό είσοδο  $s_k(t)$  και την τρέχουσα ενεργοποίηση  $y_k(t)$ , παράγει μια νέα τιμή ενεργοποίησης για τη μονάδα  $k$ .

$$y_k(t+1) = F_k(y_k(t), s_k(t)) \quad (3.2)$$

Σε αρκετές περιπτώσεις, η συνάρτηση ενεργοποίησης είναι μη φθίνουσα συνάρτηση της ολικής εισόδου στη μονάδα.

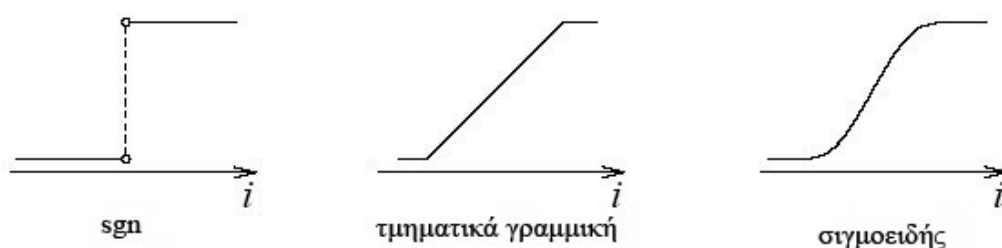
$$y_k(t+1) = F_k(s_k(t)) = F_k\left(\sum_j w_{jk}(t) y_j(t) + \theta_k(t)\right) \quad (3.3)$$

Γενικά, ως συνάρτηση ενεργοποίησης λαμβάνεται κάποια συνάρτηση κατωφλιού (threshold function), όπως η συνάρτηση προσήμου (signum συνάρτηση ή  $\text{sgn}$ ), η γραμμική ή τμηματικά γραμμική συνάρτηση ή μια συνάρτηση κατωφλιού ομαλής μετάβασης (Εικόνα 3.2). Για την τελευταία συνήθως χρησιμοποιείται σιγμοειδής συνάρτηση της μορφής:

$$y_k = F(s_k) = \frac{1}{1 + e^{-s_k}} \quad (3.4)$$

Από τις συναρτήσεις αυτές, η γραμμική (τμηματική ή μη) και η σιγμοειδής δίνουν τιμές εξόδου που κυμαίνονται από 0 έως 1, ενώ η συνάρτηση  $\text{sgn}$  δίνει τιμές εξόδου ίσες με  $-1$  ή  $+1$ .

Σε κάποιες περιπτώσεις, η έξοδος μιας μονάδας μπορεί να είναι στοχαστική συνάρτηση της ολικής εισόδου της μονάδας (Kröse, Smagt, 1996).



Εικόνα 3.2: Είδη συναρτήσεων ενεργοποίησης (Kröse, Smagt, 1996)

### 3.2 Κατηγορίες Δικτύων Βάσει της Διαμόρφωσής τους

Τα νευρωνικά δίκτυα, ανάλογα με τη μορφή των συνδέσεων που διαθέτουν, χωρίζονται σε δύο κατηγορίες (Kröse, Smagt, 1996):

- Τα *πρόσω - τροφοδότησης δίκτυα* (*feed - forward networks*), στα οποία η ροή των δεδομένων από την είσοδο στην έξοδο γίνεται αυστηρά και μόνο προς τα εμπρός, δηλαδή τα σήματα κατευθύνονται από την είσοδο στην έξοδο. Η επεξεργασία των δεδομένων μπορεί να γίνεται σε περισσότερα από ένα επίπεδα (layers), ωστόσο δεν υπάρχουν συνδέσεις που να εκτείνονται από την έξοδο κάποιας μονάδας ως την είσοδο κάποιας άλλης μονάδας του ίδιου επιπέδου ή προηγούμενου. Στην κατηγορία αυτή ανήκουν τα τύπου Perceptron και Adaline.
- Τα *δίκτυα με ανάδραση* (*recurrent networks*), τα οποία περιέχουν ανάδρομες συνδέσεις, δηλαδή οι έξοδοι κάποιων νευρώνων γίνονται είσοδοι σε νευρώνες του ίδιου ή προηγούμενου επιπέδου.

### 3.3 Εκπαίδευση Νευρωνικών Δικτύων

Τα νευρωνικά δίκτυα σχεδιάζονται με τέτοιο τρόπο ώστε όταν τους δίνεται κάποια είσοδος, αυτά να παράγουν την επιθυμητή έξοδο. Αυτό επιτυγχάνεται μεταβάλλοντας τη δύναμη των συνδέσεων μεταξύ των μονάδων, δηλαδή τα συναπτικά βάρη  $w_{jk}$ , πράγμα που μπορεί να γίνει με διάφορους τρόπους. Ένας τρόπος είναι να οριστούν σαφώς τα συναπτικά βάρη, γνωρίζοντας εκ των προτέρων τις τιμές τους. Ένας άλλος τρόπος είναι να εκπαιδευτεί το νευρωνικό δίκτυο, παρέχοντάς του πρότυπα και αφήνοντάς το να αλλάξει τα συναπτικά βάρη βάσει ενός κανόνα εκμάθησης.

Υπάρχουν δύο κατηγορίες εκπαίδευσης:

- *Η επιβλεπόμενη ή συσχετιζόμενη μάθηση (supervised or associative learning)*, κατά την οποία το δίκτυο εκπαιδεύεται τροφοδοτώντας το με πρότυπα ζεύγη εισόδων - εξόδων, τα οποία μπορεί να παρέχονται είτε από κάποιον εξωτερικό εκπαιδευτή (το χρήστη) είτε από το ίδιο το σύστημα που περιλαμβάνει το δίκτυο (αυτο-επιβλεπόμενο). Η επιβλεπόμενη μάθηση εφαρμόζεται όταν το δίκτυο χρησιμοποιείται για την προσέγγιση συνάρτησης, που αντιστοιχίζει την είσοδο με την έξοδο.
- *Η μη - επιβλεπόμενη μάθηση (unsupervised learning) ή αυτο-οργάνωση (self-organization)*, κατά την οποία οι μονάδες εξόδου εκπαιδεύονται να αντιδρούν σε ομάδες προτύπων που παρουσιάζονται στην είσοδο. Σε αντίθεση με την επιβλεπόμενη μάθηση, δε δίνεται στο δίκτυο κάποια επιθυμητή έξοδος, οπότε το σύστημα πρέπει να παράγει δική του αναπαράσταση των ερεθισμάτων εισόδου. Η μη επιβλεπόμενη μάθηση εφαρμόζεται σε δίκτυα με ανάδραση.

Το σύνολο των παραδειγμάτων που χρησιμοποιούνται για την εκπαίδευση ενός νευρωνικού δικτύου ονομάζεται *σύνολο εκπαίδευσης*. Για την εκπαίδευση χρησιμοποιούνται κανόνες που βασίζονται στην ελαχιστοποίηση του λάθους στην έξοδο του δικτύου. Ακολουθεί η *γενίκευση*, δηλαδή το δίκτυο μπορεί να αντιμετωπίζει παραδείγματα για τα οποία δεν έχει εκπαιδευτεί.

Τα τελευταία χρόνια έχουν σχεδιαστεί διάφοροι κανόνες εκμάθησης. Ιδιαίτερα γνωστοί είναι αυτοί που θεωρούνται παραλλαγή του *κανόνα εκμάθησης του Hebb* (Hebbian learning rule), που εισηγήθηκε ο ίδιος στο βιβλίο του “Organization of Behaviour” (Hebb, 1949). Για δύο μονάδες  $j$  και  $k$  που συνδέονται μεταξύ τους, ο κανόνας εκμάθησης του Hebb γράφεται ως

$$\Delta w_{jk} = \gamma y_j y_k \quad (3.5)$$

όπου το  $\gamma$  είναι μια θετική σταθερά αναλογίας, που ονομάζεται *ρυθμός εκμάθησης*.

Ωστόσο, ο περισσότερο χρησιμοποιούμενος κανόνας εκμάθησης είναι ο *κανόνας των Widrow - Hoff* ή *κανόνας δέλτα*, σύμφωνα με τον οποίο η προσαρμογή των συναπτικών βαρών γίνεται βάσει της σχέσης

$$\Delta w_{jk} = \gamma y_j (d_k - y_k) \quad (3.6)$$

όπου  $d_k$  είναι η επιθυμητή ενεργοποίηση (ή έξοδος), που ορίζεται από τον εκπαιδευτή (Kröse, Smagt, 1996).



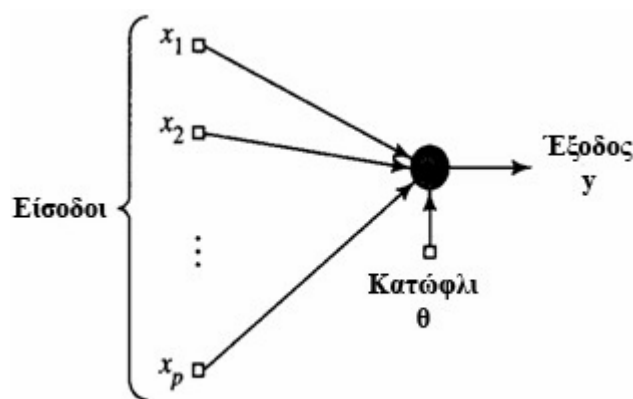
## ΚΕΦΑΛΑΙΟ 4 – ΠΕΡΙΓΡΑΦΗ PERCEPTRON ΚΑΙ ADALINE

Στο κεφάλαιο αυτό γίνεται περιγραφή δύο μοντέλων νευρωνικών δικτύων μοναδικού επιπέδου πρόσω - τροφοδότησης: του *Perceptron*, που προτάθηκε από τον Rosenblatt στο τέλος της δεκαετίας του '50 (Rosenblatt, 1959) και του *Adaline*, που προτάθηκε από τους Widrow και Hoff στις αρχές του '60 (Widrow, Hoff, 1960).

### 4.1 Περιγραφή του Perceptron

#### 4.1.1 Εισαγωγή

Ο Perceptron (Αισθητήρας) (Εικόνα 4.1) αποτελεί την απλούστερη μορφή νευρωνικού δικτύου και βρίσκει χρήση στην ταξινόμηση γραμμικά διαχωρίσιμων προτύπων, δηλαδή προτύπων που βρίσκονται στις αντίθετες πλευρές ενός υπερεπιπέδου, που ορίζει τις περιοχές απόφασης (Λυκοθανάσης, Γεωργόπουλος, 2003).



**Εικόνα 4.1:** Perceptron ενός επιπέδου (Λυκοθανάσης, Γεωργόπουλος, 2003)

Η έξοδος από αυτό το δίκτυο είναι συνάρτηση της εισόδου και δίνεται από τη σχέση

$$y = F\left(\sum_{i=1}^p w_i x_i + \theta\right) \quad (4.1)$$

Μελετώντας την περίπτωση που έχουμε δύο εισόδους, η έξοδος από το δίκτυο θα δίνεται από τη σχέση

$$y = F(w_1x_1 + w_2x_2 + \theta) \quad (4.2)$$

Η συνάρτηση ενεργοποίησης  $F$  μπορεί να είναι γραμμική ή μη γραμμική. Θεωρώντας ότι η  $F$  είναι η συνάρτηση προσήμου  $\text{sgn}$ , έχουμε

$$F(s) = \begin{cases} +1 & \text{αν } s > 0 \\ -1 & \text{αλλιώς} \end{cases} \quad (4.3)$$

Σύμφωνα με τη σχέση (4.3), η έξοδος από το δίκτυο θα ισούται με  $+1$  ή  $-1$ , ανάλογα με την είσοδο. Σκοπός του Perceptron είναι να ταξινομήσει ένα σύνολο προτύπων εισόδου σε μία από δύο κλάσεις. Ο κανόνας απόφασης για την ταξινόμηση είναι ο εξής: αν η είσοδος είναι θετική, το πρότυπο ταξινομείται στην κλάση  $+1$  ενώ αν η είσοδος είναι αρνητική, το πρότυπο ταξινομείται στην κλάση  $-1$ .

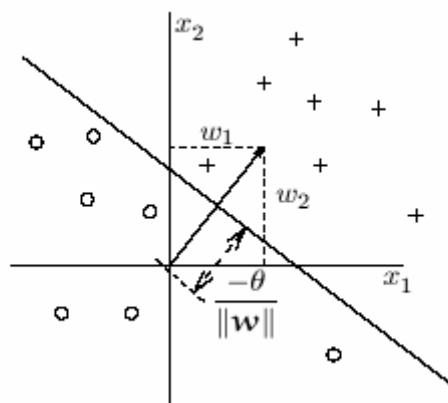
Οι περιοχές απόφασης διαχωρίζονται από μια ευθεία γραμμή, η εξίσωση της οποίας δίνεται από τη σχέση

$$w_1x_1 + w_2x_2 + \theta = 0 \quad (4.4)$$

η οποία μπορεί να ξαναγραφτεί με τη μορφή

$$x_2 = -\frac{w_1}{w_2}x_1 - \frac{\theta}{w_2} \quad (4.5)$$

Γεωμετρική αναπαράσταση της παραπάνω σχέσης δίνεται στην εικόνα 4.2. Από αυτήν προκύπτουν οι εξής αξιολογες παρατηρήσεις: ο λόγος των συναπτικών βαρών  $w_1/w_2$  ισούται με την κλίση της ευθείας, ενώ το κατώφλι  $\theta$  καθορίζει την μετατόπιση (offset), δηλαδή την απόσταση της γραμμής από την αρχή των αξόνων. Ακόμη, το διάνυσμα των συναπτικών βαρών  $\bar{w}$  είναι κάθετο στην διαχωρίζουσα συνάρτηση (Kröse, Smagt, 1996).



**Εικόνα 4.2:** Γραμμικός διαχωρισμός για ένα δισδιάστατο πρόβλημα ταξινόμησης με δύο κλάσεις (Kröse, Smagt, 1996)

#### 4.1.2 Κανόνες Εκμάθησης του Perceptron και Θεώρημα Σύγκλισης

##### Κανόνες εκμάθησης

Το δίκτυο εκπαιδεύεται στον υπολογισμό των τιμών των συναπτικών βαρών και των κατωφλίων βάσει της παρακάτω επαναληπτικής διαδικασίας:

$$w_i(t+1) = w_i(t) + \Delta w_i(t) \quad (4.6)$$

$$\theta(t+1) = \theta(t) + \Delta \theta(t) \quad (4.7)$$

Δηλαδή, ο υπολογισμός κάθε νέας τιμής γίνεται προσθέτοντας στην παλιά τιμή έναν όρο διόρθωσης.

Θεωρώντας ότι έχουμε ένα σύνολο προτύπων που αποτελούνται από το διάνυσμα εισόδου  $\mathbf{x}$  και την επιθυμητή έξοδο  $d(\mathbf{x})$  (όπου το  $d(\mathbf{x})$  όπως είδαμε παίρνει τιμές  $+1$  ή  $-1$ ), ο κανόνας εκμάθησης του Perceptron έχει ως εξής:

1. Ξεκίνα με τυχαίες τιμές συναπτικών βαρών.
2. Επέλεξε ένα διάνυσμα εισόδου  $\mathbf{x}$  από το σύνολο εκπαίδευσης.
3. Αν  $y \neq d(\mathbf{x})$  (το δίκτυο δίνει λάθος απάντηση), τροποποίησε όλα τα συναπτικά βάρη  $w_i$  σύμφωνα με τη σχέση  $\Delta w_i = d(\mathbf{x}) \cdot x_i$ .
4. Επέστρεψε στο βήμα 2.

Από τον παραπάνω κανόνα παρατηρείται ότι όταν το δίκτυο απαντήσει σωστά, τα συναπτικά βάρη δεν υφίστανται περαιτέρω τροποποίηση.

Για την τροποποίηση των τιμών κατωφλιού, θεωρείται ότι το  $\theta$  ισούται με το συναπτικό βάρος  $w_o$  μεταξύ του νευρώνα εξόδου και μιας «ψεύτικης», υποστηρικτικής μονάδας που έχει τιμή ενεργοποίησης  $x_o = 1$ . Έτσι, το κατώφλι μεταβάλλεται σύμφωνα με την παρακάτω σχέση (Kröse, Smagt, 1996):

$$\Delta \theta = \begin{cases} 0 & \text{αν ο Perceptron απαντήσει σωστά} \\ d(\mathbf{x}) & \text{αλλιώς} \end{cases} \quad (4.8)$$

##### Θεώρημα σύγκλισης

Σύμφωνα με το θεώρημα σύγκλισης του Perceptron, αν υπάρχει μια ομάδα συναπτικών βαρών  $\mathbf{w}^*$  που να μπορεί να εκτελέσει τη μετατροπή  $y = d(\mathbf{x})$ , ο κανόνας εκμάθησης του Perceptron θα συγκλίνει σε κάποια λύση (που μπορεί να είναι το  $\mathbf{w}^*$ ,

μπορεί και όχι) μετά από έναν πεπερασμένο αριθμό βημάτων, για οποιαδήποτε αρχική επιλογή συναπτικών βαρών (Kröse, Smagt, 1996).

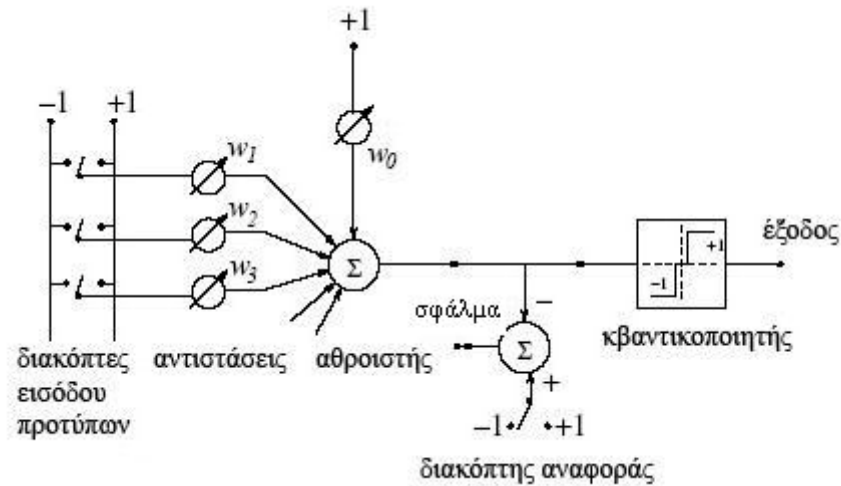
## 4.2 Περιγραφή του Adaline

### 4.2.1 Εισαγωγή

Το Adaline (Adaptive Linear Element, Γραμμικό Προσαρμοζόμενο Στοιχείο), που αναπτύχθηκε από τους Widrow και Hoff, αποτελεί μία προσαρμοζόμενη μηχανή ταξινόμησης προτύπων η οποία χρησιμοποιεί τον κανόνα των *Μέσων Ελαχίστων Τετραγώνων* (Least Mean Square, LMS), ή κανόνα δέλτα, για τη λειτουργία της (Λυκοθανάσης, Γεωργόπουλος, 2003).

Η κύρια διαφορά του κανόνα δέλτα από τον κανόνα εκμάθησης του Perceptron έγκειται στον τρόπο με τον οποίο η έξοδος του συστήματος χρησιμοποιείται στον κανόνα εκμάθησης. Ο κανόνας εκμάθησης του Perceptron χρησιμοποιεί την έξοδο της συνάρτησης κατωφλιού για τη διόρθωση των συναπτικών βαρών (δηλαδή την επιθυμητή απόκριση), ενώ ο κανόνας δέλτα χρησιμοποιεί τη *διαφορά* της πραγματικής εξόδου από την επιθυμητή.

Μια σχηματική αναπαράσταση του Adaline παρουσιάζεται στην εικόνα 4.3. Το δίκτυο αυτό αποτελείται από ένα σύνολο ελεγχόμενων αντιστάσεων, που συνδέονται με έναν αθροιστή (circuit), ο οποίος αθροίζει τα ρεύματα που δημιουργούνται από τη διαφορά δυναμικού των σημάτων εισόδου. Συνήθως μετά τον αθροιστή τοποθετείται ένας κβαντικοποιητής ο οποίος δίνει στην έξοδο τιμές +1 ή -1, ανάλογα με την πολικότητα του αθροίσματος (Kröse, Smagt, 1996).



Εικόνα 4.3: Γραμμικό προσαρμοζόμενο στοιχείο (Adaline) (Kröse, Smagt, 1996)

#### 4.2.2 Κανόνες Εκμάθησης του Adaline

Η έξοδος από τη μονάδα της εικόνας 4.3, δεδομένου ότι στη μονάδα εξόδου χρησιμοποιείται γραμμική συνάρτηση ενεργοποίησης, δίνεται από τη σχέση

$$y = \sum_{j=1}^n w_j x_j + \theta \quad (4.9)$$

όπου το κατώφλι  $\theta$  ταυτίζεται με το  $w_0$ . Το δίκτυο πρέπει να σχεδιαστεί με τέτοιο τρόπο ούτως ώστε όταν του δίνονται στην είσοδο τα στοιχεία του  $p$ -οστού πρότυπου διανύσματος εισόδου  $x_j^p, j = 1, 2, \dots, n$ , αυτό να δίνει την επιθυμητή έξοδο  $y = d^p$ .

Αυτό, όπως ειπώθηκε παραπάνω, επιτυγχάνεται προσαρμόζοντας τις τιμές των συναπτικών βαρών  $w_j, j = 0, 1, \dots, n$  μέσω του κανόνα δέλτα. Σε περίπτωση που δε μπορούμε να έχουμε ακριβώς την επιθυμητή τιμή εξόδου, το μέσο λάθος ελαχιστοποιείται χρησιμοποιώντας τη μέθοδο των μέσων ελαχίστων τετραγώνων.

Για κάθε πρότυπο εισόδου, η πραγματική έξοδος του δικτύου  $y^p$  διαφέρει από την επιθυμητή έξοδο  $d^p$  κατά  $(d^p - y^p)$ . Ο κανόνας δέλτα χρησιμοποιεί μια *συνάρτηση σφάλματος* (ή *συνάρτηση κόστους*), που βασίζεται σε αυτές τις διαφορές, για να ρυθμίσει τα συναπτικά βάρη. Το ολικό σφάλμα δίνεται από τη σχέση

$$E = \sum_p E^p = \frac{1}{2} \sum_p (d^p - y^p)^2 \quad (4.10)$$

όπου το  $E^p$  αντιπροσωπεύει το σφάλμα στο πρότυπο  $p$ .

Ο αλγόριθμος LMS υπολογίζει τις τιμές των συναπτικών βαρών που ελαχιστοποιούν τη συνάρτηση σφάλματος χρησιμοποιώντας τη μέθοδο της μέγιστης κατάβασης, σύμφωνα με την οποία η μεταβολή των συναπτικών βαρών είναι ανάλογη με την αρνητική παράγωγο της συνάρτησης σφάλματος σε ένα δεδομένο πρότυπο, ως προς το αντίστοιχο συναπτικό βάρος:

$$\Delta_p w_j = -\gamma \frac{\partial E^p}{\partial w_j} \quad (4.11)$$

όπου  $\gamma$  μια σταθερά αναλογίας, που ονομάζεται *ρυθμός εκμάθησης* (learning rate). Η παράγωγος της συνάρτησης σφάλματος σε ένα δεδομένο πρότυπο ως προς το αντίστοιχο συναπτικό βάρος ισούται με

$$\frac{\partial E^p}{\partial w_j} = \frac{\partial E^p}{\partial y^p} \cdot \frac{\partial y^p}{\partial w_j} \quad (4.12)$$

Από την εξίσωση (4.9) έχουμε

$$\frac{\partial y^p}{\partial w_j} = x_j \quad (4.13)$$

ενώ από την εξίσωση (4.10) έχουμε

$$\frac{\partial E^p}{\partial y^p} = -(d^p - y^p) \quad (4.14)$$

Αντικαθιστώντας τις εξισώσεις (4.13) και (4.14) στην (4.12) παίρνουμε

$$\frac{\partial E^p}{\partial w_j} = \frac{\partial E^p}{\partial y^p} \cdot \frac{\partial y^p}{\partial w_j} \Rightarrow \frac{\partial E^p}{\partial w_j} = -(d^p - y^p) \cdot x_j \quad (4.15)$$

Τέλος, αντικαθιστώντας την εξίσωση (4.15) στην (4.11), έχουμε

$$\begin{aligned} \Delta_p w_j &= -\gamma \frac{\partial E^p}{\partial w_j} = -\gamma \cdot [-(d^p - y^p) \cdot x_j] \Rightarrow \\ &\Rightarrow \Delta_p w_j = \gamma \cdot \delta^p \cdot x_j \end{aligned} \quad (4.16)$$

όπου  $\delta^p = d^p - y^p$  είναι η διαφορά της επιθυμητής εξόδου από την πραγματική για το πρότυπο  $p$  και ονομάζεται *τοπική κλίση* (Kröse, Smagt, 1996).

### 4.3 Το Πρόβλημα της Exclusive-Or (XOR)

Όπως είδαμε σε προηγούμενη παράγραφο, ο απλός Perceptron περιλαμβάνει μονάδες εισόδου και εξόδου, αλλά όχι κρυφές μονάδες. Απόρροια αυτού είναι να μη μπορεί να

ταξινομήσει πρότυπα εισόδου που δεν είναι γραμμικά διαχωρίσιμα, όπως χαρακτηριστικά συμβαίνει με τη συνάρτηση exclusive-or (XOR).

Θεωρώντας την απλή περίπτωση όπου έχουμε δύο εισόδους και μία έξοδο, η ολική είσοδος ισούται με

$$s = w_1x_1 + w_2x_2 + \theta \quad (4.17)$$

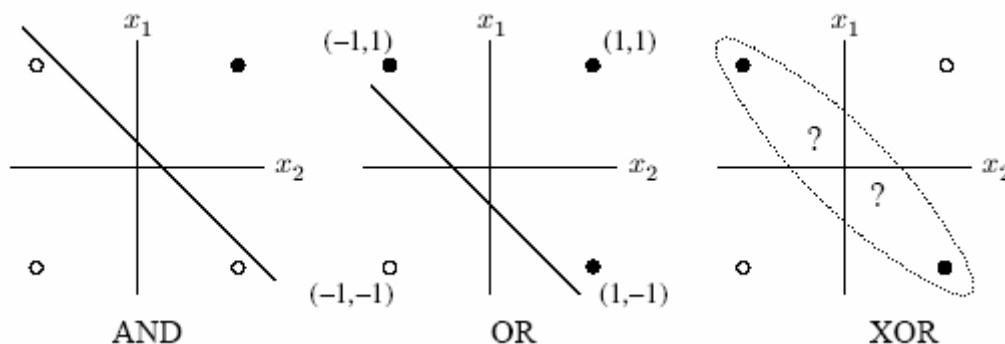
Σύμφωνα με την εξίσωση (4.3), που παρατέθηκε στην παράγραφο 4.1.1, τα πρότυπα εισόδου θα ταξινομούνται σε δύο κλάσεις. Τα πρότυπα (1, 1) και (-1, -1) ανήκουν στην κλάση -1, ενώ τα πρότυπα (-1, 1) και (1, -1) ανήκουν στην κλάση 1, όπως προκύπτει από τις παρακάτω σχέσεις (Λυκοθανάσης, Γεωργόπουλος, 2003):

$$\begin{aligned} -1 \text{ XOR } -1 &= -1 \\ 1 \text{ XOR } 1 &= -1 \\ -1 \text{ XOR } 1 &= 1 \\ 1 \text{ XOR } -1 &= 1 \end{aligned} \quad (4.18)$$

Στον πίνακα που ακολουθεί παρουσιάζονται οι επιθυμητές έξοδοι ( $d$ ) για τα διάφορα πρότυπα εισόδου (πίνακας αληθείας exclusive-or), ενώ στην εικόνα 4.4 παρουσιάζεται η γεωμετρική αναπαράσταση του διαστήματος εισόδου.

**Πίνακας 4.1:** Πίνακας αληθείας exclusive-or (Kröse, Smagt, 1996)

$x_1$	$x_2$	$d$
-1	-1	-1
-1	1	1
1	-1	1
1	1	-1



**Εικόνα 4.4:** Γεωμετρική αναπαράσταση πεδίων ορισμού των συναρτήσεων AND, OR και XOR (Kröse, Smagt, 1996)

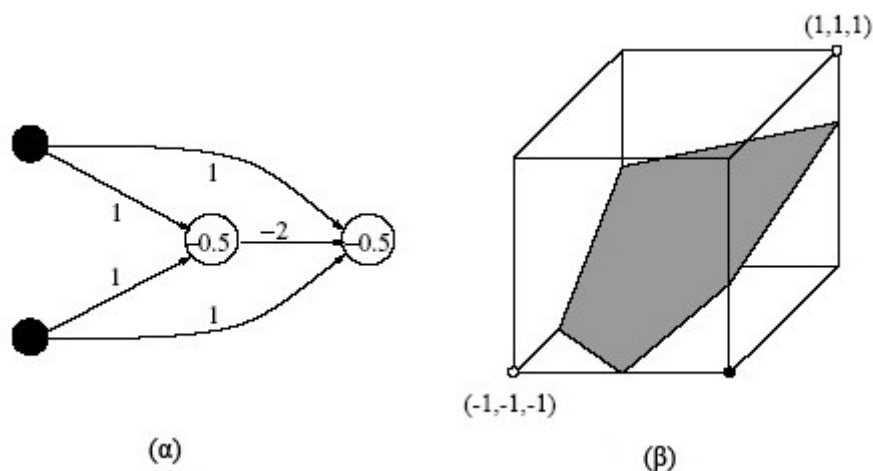
Η ευθεία γραμμή που χωρίζει τις περιοχές απόφασης δίνεται από την εξίσωση (4.4):

$$w_1x_1 + w_2x_2 + \theta = 0$$

Παρατηρώντας την εικόνα 4.4, είναι προφανές ότι δε μπορεί να κατασκευαστεί ευθεία γραμμή για ένα όριο απόφασης έτσι ώστε τα σημεία  $(-1, -1)$  και  $(1,1)$  να βρίσκονται στη μια περιοχή απόφασης και τα σημεία  $(-1,1)$  και  $(1, -1)$  να βρίσκονται στην άλλη. Δηλαδή, ένας απλός Perceptron δε μπορεί να επιλύσει το πρόβλημα της XOR (Λυκοθανάσης, Γεωργόπουλος, 2003).

Το πρόβλημα της exclusive-or μπορεί να επιλυθεί εισάγοντας ένα επίπεδο με κρυφές μονάδες που συνδέονται με τις εισόδους, αυξάνοντας δηλαδή τις διαστάσεις του προβλήματος και μετατρέποντας το δίκτυο από απλό Perceptron σε Perceptron πολλαπλών επιπέδων. Τα τέσσερα σημεία εισόδου βρίσκονται στις κορυφές ενός τετραέδρου και τα πεδίο ορισμού της συνάρτησης είναι ο τρισδιάστατος χώρος που ορίζεται από τις δύο μονάδες εισόδου και την κρυφή μονάδα (Εικόνα 4.5).

Στην εικόνα 4.5α δίνεται η μορφή του δικτύου, μετά την εισαγωγή μιας κρυφής μονάδας. Οι αριθμοί δίπλα στα βέλη αφορούν στα συναπτικά βάρη κάθε σύνδεσης, ενώ οι αριθμοί μέσα στους κύκλους αφορούν στα κατώφλια. Από την εικόνα 4.5β παρατηρείται ότι τα τέσσερα σημεία μπορούν τώρα να διαχωριστούν σε δύο ομάδες μέσω υπερεπιπέδου.



Εικόνα 4.5: Επίλυση του προβλήματος XOR (Kröse, Smagt, 1996)

Γενικά, οι Minsky και Papert απέδειξαν ότι όταν έχουμε πρόσω - τροφοδότσης δίκτυα τύπου Perceptron με δυαδικές μονάδες εισόδου, η προσθήκη επιπέδων με κρυφούς νευρώνες επιτρέπει την επίλυση οποιουδήποτε προβλήματος. Μάλιστα, η



προσθήκη ενός και μόνου κρυφού επιπέδου αρκεί για την προσέγγιση οποιασδήποτε συνάρτησης, αρκεί αυτό να περιλαμβάνει αρκετούς κόμβους (Kröse, Smagt, 1996).

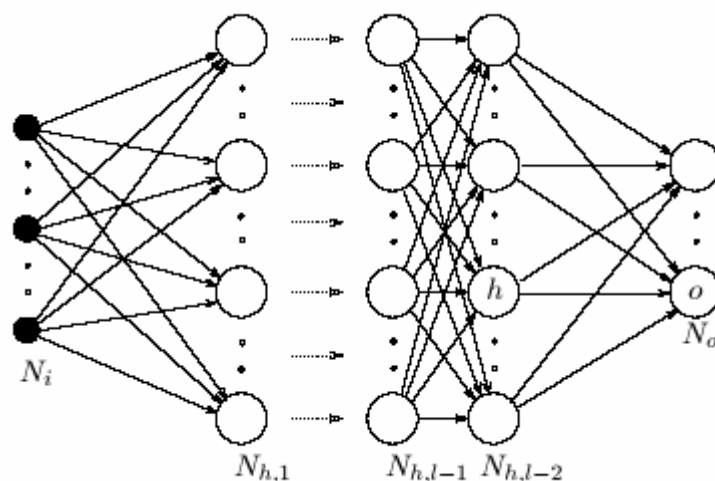
## ΚΕΦΑΛΑΙΟ 5 – ΚΑΝΟΝΑΣ ΠΙΣΩ ΔΙΑΔΟΣΗΣ

### 5.1 Εισαγωγή

Όπως είδαμε στο προηγούμενο κεφάλαιο, οι Minsky και Papert (Minsky, Papert, 1969) απέδειξαν ότι όταν έχουμε πρόσω - τροφοδότησης δίκτυα με περισσότερα από ένα επίπεδα, αυτά μπορούν να λύσουν σχεδόν οποιοδήποτε πρόβλημα. Ωστόσο, παρέλειψαν να κατασκευάσουν κάποιον κανόνα εκμάθησης για την προσαρμογή των συναπτικών βαρών των συνδέσεων μεταξύ των μονάδων εισόδου και των κρυφών μονάδων.

Με το πρόβλημα αυτό ασχολήθηκαν διάφοροι ερευνητές, που κατέληξαν στο γενικό συμπέρασμα ότι τα σφάλματα για τις κρυφές μονάδες μπορούν να καθοριστούν διαδίδοντας προς τα πίσω τα σφάλματα των μονάδων εξόδου. Η μέθοδος ονομάζεται *κανόνας εκμάθησης πίσω - διάδοσης* και θεωρείται γενίκευση του κανόνα δέλτα για μη γραμμικές συναρτήσεις ενεργοποίησης και δίκτυα πολλών επιπέδων.

Γενικά, ένα δίκτυο πρόσω - τροφοδότησης αποτελείται από διάφορα επίπεδα, κάθε ένα από τα οποία αποτελείται από μονάδες που δέχονται σήματα εισόδου από τις μονάδες του ανάντη ευρισκόμενου επιπέδου και στέλνουν σήματα εξόδου στις μονάδες του κατάντη ευρισκόμενου επιπέδου (Kröse, Smagt, 1996). Ένα τέτοιο δίκτυο αποτελείται από ένα σύνολο αισθητήρων (πηγαίοι κόμβοι) που αποτελούν το επίπεδο εισόδου και στο οποίο δε λαμβάνει χώρα επεξεργασία του σήματος, ένα ή περισσότερα κρυφά επίπεδα υπολογιστικών κόμβων και ένα επίπεδο υπολογιστικών κόμβων εξόδου. Το σήμα εισόδου διαδίδεται μέσα στο δίκτυο κατευθυνόμενο προς τα εμπρός, από επίπεδο σε επίπεδο (Λυκοθανάσης, Γεωργόπουλος, 2003). Δηλαδή τα σήματα από τις μονάδες εισόδου κατευθύνονται στο πρώτο επίπεδο κρυφών μονάδων, από εκεί στο δεύτερο κ.ο.κ. μέχρι το τελευταίο επίπεδο κρυφών μονάδων, από το οποίο τα σήματα καταλήγουν στις μονάδες εξόδου (Εικόνα 5.1).

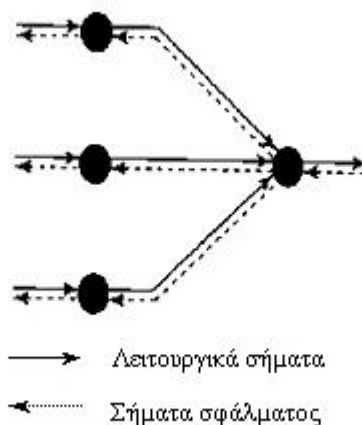


**Εικόνα 5.1:** Πρόσω - τροφοδότησης δίκτυο πολλών επιπέδων (Kröse, Smagt, 1996)

Αυτά τα νευρωνικά δίκτυα είναι γνωστά ως Perceptrons πολλών επιπέδων (Multi - Layer Perceptrons, MLPs) και αποτελούν γενίκευση του απλού Perceptron. Η εκπαίδευσή τους, όπως ειπώθηκε παραπάνω, γίνεται με τον αλγόριθμο πίσω - διάδοσης του σφάλματος, ο οποίος βασίζεται στον κανόνα μάθησης διόρθωσης του σφάλματος (error correction learning rule). Η διαδικασία χαρακτηρίζεται από την ύπαρξη δύο περασμάτων διαμέσου των επιπέδων του δικτύου (Λυκοθανάσης, Γεωργόπουλος, 2003):

- Το προς τα εμπρός πέρασμα, κατά τη διάρκεια του οποίου το σήμα εισόδου διαδίδεται προς τα εμπρός διαμέσου του δικτύου (από επίπεδο σε επίπεδο) και καταλήγει στις μονάδες εξόδου. Το σήμα αυτό καλείται *λειτουργικό σήμα* (function signal) και κατά τη διάδοσή του τα συναπτικά βάρη παραμένουν σταθερά.
- Το προς τα πίσω πέρασμα, κατά τη διάρκεια του οποίου διαδίδονται προς τα πίσω *σήματα σφάλματος* (error signals) από τις μονάδες εξόδου και τα συναπτικά βάρη προσαρμόζονται σύμφωνα με τον κανόνα μάθησης διόρθωσης του σφάλματος.

Τα δύο αυτά σήματα ροής απεικονίζονται στην εικόνα 5.2.



**Εικόνα 5.2:** Λειτουργικά σήματα και σήματα σφάλματος σε ένα νευρωνικό δίκτυο (Λυκοθανάσης, Γεωργόπουλος, 2003)

Τα Perceptrons πολλών επιπέδων χαρακτηρίζονται από τρία σημεία, τα οποία - σε συνδυασμό με τη δυνατότητα εκπαίδευσής τους - τους παρέχουν την ικανότητα επίλυσης περίπλοκων προβλημάτων:

- Η έξοδος των νευρώνων παρουσιάζει μη γραμμικότητα, καθώς η συνάρτηση ενεργοποίησης είναι μη γραμμική. Συνήθως χρησιμοποιείται η σιγμοειδής συνάρτηση της μορφής

$$y^p = F(s^p) = \frac{1}{1 + e^{-s^p}} \quad (5.1)$$

η οποία είναι παραγωγίσιμη σε όλα τα σημεία του πεδίου ορισμού. Η μη γραμμικότητα των συναρτήσεων που χρησιμοποιούνται σε τέτοιου είδους δίκτυα είναι φυσιολογική απόρροια του γεγονότος ότι οι βιολογικοί νευρώνες είναι μη γραμμικοί.

- Το δίκτυο αποτελείται από ένα ή περισσότερα επίπεδα από κρυφούς νευρώνες, οι οποίοι εξάγουν σταδιακά πληροφορίες από τα διανύσματα εισόδου, δίνοντας έτσι στο σύστημα τη δυνατότητα να μαθαίνει και να προσομοιώνει περίπλοκες διεργασίες.
- Οι συνδέσεις του δικτύου έχουν ως αποτέλεσμα αυτό να παρουσιάζει υψηλή διασύνδεση. Μεταβάλλοντας τον αριθμό των συνδέσεων ή τα συναπτικά βάρη, μεταβάλλεται και ο τρόπος διασύνδεσης του δικτύου.

Ωστόσο, στα παραπάνω στοιχεία οφείλεται και η δυσκολία θεωρητικής προσέγγισης των δικτύων πολλών επιπέδων. Η θεωρητική ανάλυση τέτοιων δικτύων είναι εξαιρετικά δύσκολη, λόγω της μη γραμμικότητας και του υψηλού βαθμού

διασύνδεσης που αυτά παρουσιάζουν, ενώ η παρουσία των επιπέδων κρυφών μονάδων έχει ως αποτέλεσμα να είναι δύσκολη η κατανόηση της διαδικασίας μάθησης (Λυκοθανάσης, Γεωργόπουλος, 2003).

## 5.2 Γενικευμένος Κανόνας Δέλτα - Μαθηματικές Εξισώσεις και Επεξήγηση

Η ανάγκη γενίκευσης του κανόνα δέλτα προκύπτει λόγω της χρήσης μη γραμμικών συναρτήσεων σε νευρωνικά δίκτυα με περισσότερα του ενός επιπέδων.

Όταν το πρότυπο διάνυσμα εισόδου  $p$  εφαρμόζεται στο δίκτυο, η είσοδος στο στοιχείο  $k$  θα είναι

$$s_k^p = \sum_j w_{jk} y_j^p + \theta_k \quad (5.2)$$

και η έξοδος (ενεργοποίηση) αυτού θα ισούται με

$$y_k^p = F(s_k^p) \quad (5.3)$$

Θέτουμε

$$\Delta_p w_{jk} = -\gamma \frac{\partial E^p}{\partial w_{jk}} \quad (5.4)$$

όπου ο όρος  $E^p$  αφορά στο τετραγωνικό λάθος στις μονάδες εξόδου για το πρότυπο  $p$  και ισούται με

$$E^p = \frac{1}{2} \sum_{o=1}^{N_o} (d_o^p - y_o^p)^2 \quad (5.5)$$

Στην παραπάνω σχέση, το  $N_o$  συμβολίζει τον αριθμό των μονάδων εξόδου και το  $d_o^p$  την επιθυμητή έξοδο της μονάδας  $o$ .

Το ολικό τετραγωνικό σφάλμα ισούται με

$$E = \sum_p E^p \quad (5.6)$$

Ακόμη έχουμε

$$\frac{\partial E^p}{\partial w_{jk}} = \frac{\partial E^p}{\partial s_k^p} \cdot \frac{\partial s_k^p}{\partial w_{jk}} \quad (5.7)$$

Από την εξίσωση (5.2) έχουμε

$$\frac{\partial s_k^p}{\partial w_{jk}} = y_j^p \quad (5.8)$$

Θέτοντας

$$\delta_k^p = -\frac{\partial E^p}{\partial s_k^p} \quad (5.9)$$

και αντικαθιστώντας τις εξισώσεις (5.7), (5.8) και (5.9) στην εξίσωση (5.4), παίρνουμε

$$\Delta_p w_{jk} = \gamma \cdot \delta_k^p \cdot y_j^p \quad (5.10)$$

Η εξίσωση (5.10) αποτελεί έναν κανόνα ενημέρωσης, που οδηγεί σε σταδιακή μείωση του σφάλματος. Ο υπολογισμός των τιμών του  $\delta$  γίνεται μέσω μιας επαναληπτικής διαδικασίας, κατά την οποία διαδίδονται προς τα πίσω σήματα σφάλματος.


Η εξίσωση (5.9) μπορεί να γραφτεί στη μορφή

$$\delta_k^p = -\frac{\partial E^p}{\partial s_k^p} = -\frac{\partial E^p}{\partial y_k^p} \frac{\partial y_k^p}{\partial s_k^p} \quad (5.11)$$

Ο υπολογισμός του δεύτερου όρου γίνεται παραγωγίζοντας την εξίσωση (5.3), οπότε έχουμε

$$\frac{\partial y_k^p}{\partial s_k^p} = F'(s_k^p) \quad (5.12)$$

Για τον υπολογισμό του πρώτου όρου διακρίνονται δύο περιπτώσεις.

 Η μονάδα k αποτελεί έξοδο ( $k = o$ )

Σε αυτή την περίπτωση και σε αναλογία με την εξίσωση (4.14) του κεφαλαίου 4, έχουμε:

$$\frac{\partial E^p}{\partial y_o^p} = -(d_o^p - y_o^p) \quad (5.13)$$

Αντικαθιστώντας τις εξισώσεις (5.12) και (5.13) στην εξίσωση (5.11) έχουμε:

$$\delta_o^p = (d_o^p - y_o^p) F'_o(s_o^p) \quad (5.14)$$

Θεωρώντας ότι η συνάρτηση ενεργοποίησης  $F$  είναι σιγμοειδής, θα ισχύει

$$y^p = F(s^p) = \frac{1}{1 + e^{-s^p}} \quad (5.15)$$

και η παράγωγός της θα ισούται με:

$$F'(s^p) = \frac{\partial}{\partial s^p} \frac{1}{1 + e^{-s^p}} = \frac{1}{(1 + e^{-s^p})^2} \cdot (-e^{-s^p}) = \frac{1}{(1 + e^{-s^p})} \left( \frac{e^{-s^p}}{1 + e^{-s^p}} \right) = y^p (1 - y^p) \quad (5.16)$$

Έτσι, η εξίσωση (5.14) γίνεται

$$\delta_o^p = (d_o^p - y_o^p) y_o^p (1 - y_o^p) \quad (5.17)$$

➡ Η μονάδα  $k$  αποτελεί κρυφή μονάδα ( $k = h$ )

Σε αυτή την περίπτωση δε γνωρίζουμε τη συνεισφορά της μονάδας στο σφάλμα εξόδου του δικτύου, ωστόσο το σφάλμα μπορεί να θεωρηθεί συνάρτηση των εισόδων από το κρυφό επίπεδο στο επίπεδο εξόδου, δηλαδή  $E^p = E^p(s_1^p, s_2^p, \dots, s_j^p, \dots)$ .

Συνεπώς, σύμφωνα με τον κανόνα της αλυσίδας έχουμε:

$$\frac{\partial E^p}{\partial y_h^p} = \sum_{o=1}^{N_o} \frac{\partial E^p}{\partial s_o^p} \frac{\partial s_o^p}{\partial y_h^p} = \sum_{o=1}^{N_o} \frac{\partial E^p}{\partial s_o^p} \frac{\partial}{\partial y_h^p} \sum_{j=1}^{N_h} w_{ko} y_j^p = \sum_{o=1}^{N_o} \frac{\partial E^p}{\partial s_o^p} w_{ho} = - \sum_{o=1}^{N_o} \delta_o^p w_{ho} \quad (5.18)$$

Αντικαθιστώντας την τελευταία εξίσωση στην εξίσωση (5.11), παίρνουμε

$$\delta_h^p = F'(s_h^p) \sum_{o=1}^{N_o} \delta_o^p w_{ho} \quad (5.19)$$

Θεωρώντας και πάλι ότι η συνάρτηση ενεργοποίησης είναι σιγμοειδής, η τελευταία εξίσωση παίρνει τη μορφή:

$$\delta_h^p = y_h^p (1 - y_h^p) \sum_{o=1}^{N_o} \delta_o^p w_{ho} \quad (5.20)$$

Η διαδικασία που παρουσιάστηκε παραπάνω αποτελεί τη γενίκευση του κανόνα δέλτα. Πέρα όμως από τις μαθηματικές εξισώσεις που την περιγράφουν, η επεξήγηση του τι ακριβώς σημαίνουν αυτές οι εξισώσεις και του τι συμβαίνει στο σύστημα είναι απλή. Όταν χρησιμοποιείται ένα εκπαιδευτικό πρότυπο ως είσοδος του δικτύου, οι τιμές ενεργοποίησης διαδίδονται από τη μία μονάδα στην άλλη μέχρι τις μονάδες εξόδου και η πραγματική έξοδος του δικτύου συγκρίνεται με την επιθυμητή. Καθώς στις περισσότερες περιπτώσεις οι δύο αυτές έξοδοι από κάποια μονάδα  $o$  δεν ταυτίζονται, προκύπτει κάποιο σφάλμα  $e_o$  που ισούται με τη διαφορά τους. Αυτό το σφάλμα πρέπει να ελαχιστοποιηθεί ή, αν είναι εφικτό, να μηδενιστεί. Αυτό επιτυγχάνεται μεταβάλλοντας τις δυνάμεις σύνδεσης του δικτύου (δηλαδή τα συναπτικά βάρη), έτσι ώστε στην επόμενη επανάληψη που εφαρμόζεται το

συγκεκριμένο πρότυπο, το σφάλμα  $e_o$  να μειωθεί. Από τον κανόνα δέλτα γνωρίζουμε ότι αυτή η μεταβολή των συναπτικών βαρών δίνεται από τη σχέση

$$\Delta w_{ho} = (d_o - y_o)y_h \quad (5.21)$$

Ωστόσο, εφαρμόζοντας απλά και μόνο την παραπάνω σχέση, τροποποιούνται οι δυνάμεις σύνδεσης μεταξύ των κρυφών μονάδων και των μονάδων εξόδου, αλλά όχι και οι δυνάμεις σύνδεσης μεταξύ των μονάδων εισόδου και των κρυφών μονάδων. Για να προσαρμοστούν τα συναπτικά βάρη μεταξύ των μονάδων εισόδου και των κρυφών μονάδων πρέπει να εφαρμοστεί και πάλι ο κανόνας δέλτα. Ωστόσο σε αυτή την περίπτωση δε γνωρίζουμε τις τιμές του  $\delta$  (δηλαδή τη διαφορά της επιθυμητής από την πραγματική τιμή) για τις κρυφές μονάδες. Το πρόβλημα αυτό λύνεται εφαρμόζοντας τον κανόνα της αλυσίδας, ο οποίος διαδίδει το σφάλμα της μονάδας εξόδου  $o$  σε όλες τις κρυφές μονάδες με τις οποίες συνδέεται, σταθμίζοντάς το ανάλογα με τη σύνδεση. Δηλαδή, μια κρυφή μονάδα  $h$  δέχεται ένα δέλτα από κάθε μονάδα εξόδου  $o$ , το οποίο είναι σταθμισμένο ανάλογα με το βάρος της σύνδεσης και πολλαπλασιασμένο με τη συνάρτηση ενεργοποίησης της κρυφής μονάδας, όπως δηλώνεται από την εξίσωση (5.19) (Kröse, Smagt, 1996).

### 5.3 Ρυθμός Εκμάθησης

Σύμφωνα με τη μέθοδο της μέγιστης κατάβασης που αναφέρθηκε παραπάνω, η μεταβολή των συναπτικών βαρών δίνεται από τη σχέση (5.4):

$$\Delta_p w_{jk} = -\gamma \frac{\partial E^p}{\partial w_{jk}}$$

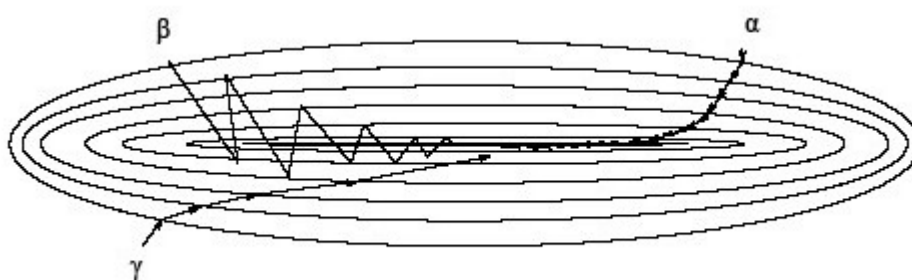
Προκειμένου η μέθοδος αυτή να οδηγήσει σε σταθερή σύγκλιση, πρέπει οι διορθώσεις των συναπτικών βαρών σε κάθε βήμα να είναι απειροελάχιστες, πράγμα που συνεπάγεται πολύ μικρή τιμή του ρυθμού εκμάθησης  $\gamma$ . Κάτι τέτοιο όμως δεν είναι πρακτικά εφικτό, λόγω του αυξημένου χρόνου που απαιτείται για τη σύγκλιση της μεθόδου. Από την άλλη, αν χρησιμοποιηθεί πολύ μεγάλη τιμή του ρυθμού εκμάθησης ώστε να επιταχυνθεί η διαδικασία, οι μεγάλες αλλαγές που θα υπάρξουν στα βάρη μπορεί να οδηγήσουν σε ταλαντώσεις και αστάθεια του δικτύου. Ένας απλός τρόπος για να επιταχυνθεί η διαδικασία μάθησης και ταυτόχρονα να αποφευχθεί ο κίνδυνος της αστάθειας, είναι η τροποποίηση του κανόνα δέλτα έτσι ώστε οι διορθώσεις των συναπτικών βαρών σε κάθε βήμα να εξαρτώνται από τις



διορθώσεις των συναπτικών βαρών του προηγούμενου βήματος. Αυτό επιτυγχάνεται προσθέτοντας στο δεξί μέλος της εξίσωσης (5.10) έναν όρο ορμής (momentum term), οπότε αυτή θα έχει τη μορφή

$$\Delta w_{jk}(t+1) = \gamma \delta_k^p y_j^p + a \Delta w_{jk}(t) \quad (5.22)$$

όπου  $a$  η σταθερά ορμής, η οποία καθορίζει πόσο μεγάλη είναι η αλλαγή του βάρους του προηγούμενου υπολογισμού (Εικόνα 5.3) (Kröse, Smagt, 1996).



**Εικόνα 5.3:** Επίδραση του όρου ορμής στη σύγκλιση της μεθόδου α) με μικρό ρυθμό εκμάθησης β) με μεγάλο ρυθμό εκμάθησης γ) με μεγάλο ρυθμό εκμάθησης και προσθήκη όρου ορμής (Kröse, Smagt, 1996)

Παρόλο που στην παραπάνω διαδικασία ο ρυθμός εκμάθησης  $\gamma$  θεωρήθηκε σταθερή παράμετρος, στην πραγματικότητα μπορεί να διαφέρει από σύνδεση σε σύνδεση στα διάφορα μέρη του δικτύου. Αυτό δίνει τη δυνατότητα ορισμένα συναπτικά βάρη να τροποποιούνται και άλλα να παραμένουν σταθερά, καθώς τα σήματα σφάλματος διαδίδονται προς τα πίσω στο δίκτυο. Το τελευταίο μπορεί να επιτευχθεί θέτοντας τον ρυθμό εκμάθησης  $\gamma_{jk}$  που αντιστοιχεί στο συναπτικό βάρος  $w_{jk}$  ίσο με μηδέν (οπότε και  $\Delta w_{jk} = 0$ ) (Λυκοθανάσης, Γεωργόπουλος, 2003).

#### 5.4 Τρόποι Εκπαίδευσης του Δικτύου

Όπως είδαμε, η εκπαίδευση ενός νευρωνικού δικτύου πραγματοποιείται εφαρμόζοντας ένα σύνολο πρότυπων διανυσμάτων εισόδου, η προβολή (παρουσίαση) των οποίων στο δίκτυο καλείται *κύκλος* ή *εποχή* (*epoch*). Έτσι, η διαδικασία μάθησης προχωρά από κύκλο σε κύκλο, έως ότου τα συναπτικά βάρη και τα κατώφλια του δικτύου φτάσουν σε μία σταθερή τιμή και το μέσο τετραγωνικό σφάλμα ελαχιστοποιηθεί.

Η εκπαίδευση του δικτύου μπορεί να γίνει με δύο τρόπους:

- *Τρόπος προτύπων (pattern mode)*, σύμφωνα με τον οποίο τα πρότυπα διανύσματα παρουσιάζονται ένα - ένα στο δίκτυο. Δηλαδή, αρχικά προβάλλεται το πρώτο διάνυσμα, οπότε και εκπέμπονται τα λειτουργικά σήματα (προς τα εμπρός) και τα σήματα σφάλματος (προς τα πίσω) και τροποποιούνται τα συναπτικά βάρη και τα κατώφλια. Στη συνέχεια προβάλλεται το δεύτερο διάνυσμα, με αποτέλεσμα να επαναλαμβάνεται η παραπάνω διαδικασία και να τροποποιούνται περαιτέρω τα συναπτικά βάρη και τα κατώφλια του δικτύου. Η διαδικασία συνεχίζεται μέχρι να εφαρμοστεί και το τελευταίο διάνυσμα εισόδου.
- *Σωρηδόν τρόπος (batch mode)*, σύμφωνα με τον οποίο όλα τα πρότυπα διανύσματα εισόδου προβάλλονται στο δίκτυο και στη συνέχεια γίνεται η τροποποίηση των συναπτικών βαρών και των κατωφλίων.

Ο πρώτος τρόπος πλεονεκτεί σε αρκετά σημεία σε σχέση με τον δεύτερο, καθώς απαιτεί λιγότερη μνήμη και θεωρείται εμπειρικά ότι οδηγεί ταχύτερα σε σύγκλιση. Ακόμη, καθώς τα διανύσματα εφαρμόζονται με τυχαία σειρά στην είσοδο του δικτύου από κύκλο σε κύκλο, η τροποποίηση των βαρών σε κάθε κύκλο μάθησης γίνεται με στοχαστικό τρόπο, αποφεύγοντας έτσι τον κίνδυνο να γίνουν λιγότεροι κύκλοι από ό,τι πρέπει ή να παγιδευτεί ο αλγόριθμος σε τοπικά ελάχιστα (Λυκοθανάσης, Γεωργόπουλος, 2003). Ωστόσο, ιδιαίτερη προσοχή πρέπει να δίνεται στη σειρά με την οποία προβάλλονται τα διανύσματα, υπό την έννοια ότι η συνεχής χρήση της ίδιας αλληλουχίας προτύπων στο δίκτυο θα έχει ως αποτέλεσμα αυτό να «συγκεντρωθεί» μόνο στα λιγοστά πρώτα πρότυπα. Το πρόβλημα αυτό αντιμετωπίζεται με τη χρήση μεθόδων εκπαίδευσης που αντιμετωπίζουν τα διανύσματα εισόδου και τα συνδυάζουν με όλους τους τρόπους (Kröse, Smagt, 1996).

## 5.5 Κριτήρια Σύγκλισης

Προκειμένου να συγκλίνει ο αλγόριθμος πίσω διάδοσης, υπάρχουν διάφορα κριτήρια τερματισμού, τα οποία στηρίζονται στις μοναδικές ιδιότητες του ελαχίστου (τοπικού ή ολικού) της υπερ-επιφάνειας σφάλματος. Μερικά από αυτά τα κριτήρια είναι τα εξής:

- Ο αλγόριθμος πίσω διάδοσης συγκλίνει όταν η ευκλείδεια νόρμα του διανύσματος κλίσης  $g(w)$  της υπερ-επιφάνειας σφάλματος φτάσει σε μια πολύ

μικρή τιμή κατώφλιού κλίσης, δηλαδή όταν  $\|g(w_{\text{τελικό}})\| \leq \varepsilon$ , όπου  $\varepsilon$  είναι ένα πολύ μικρό κατώφλι κλίσης. Το κριτήριο αυτό υστερεί σε δύο σημεία: στο ότι απαιτείται ο υπολογισμός του διανύσματος κλίσης  $g(w)$  και στο ότι ο χρόνος εκπαίδευσης μπορεί να είναι αρκετός.

- Ο αλγόριθμος πίσω διάδοσης συγκλίνει όταν ο ρυθμός μεταβολής του μέσου τετραγωνικού σφάλματος ανά κύκλο είναι πολύ μικρός (κατ' απόλυτη τιμή). Το κριτήριο αυτό προκύπτει από το γεγονός ότι, αν  $w^*$  είναι κάποιο τοπικό ή ολικό ελάχιστο, τότε η συνάρτηση σφάλματος  $E(w)$  στο σημείο  $w = w^*$  θα είναι σταθερή. Ο ρυθμός της μεταβολής του μέσου τετραγωνικού σφάλματος θεωρείται αρκετά μικρός όταν κυμαίνεται από 0,1% έως 1% ανά κύκλο εκπαίδευσης.
- Ο αλγόριθμος πίσω διάδοσης συγκλίνει όταν η μέγιστη τιμή του μέσου τετραγωνικού σφάλματος  $E(w)$  είναι ίση ή μικρότερη από ένα πολύ μικρό κατώφλι ενέργειας σφάλματος, δηλαδή όταν  $E(w) \leq \tau$ , όπου  $\tau$  ένα πολύ μικρό κατώφλι.

Γενικά, η διαδικασία εκπαίδευσης του δικτύου μπορεί να σταματήσει όταν η γενική απόδοσή του κριθεί ότι είναι ικανοποιητική (Λυκοθανάσης, Γεωργόπουλος, 2003).

## 5.6 Προβλήματα κατά την Εφαρμογή του Αλγόριθμου Πίσω Διάδοσης

Όπως είδαμε στην εισαγωγή του κεφαλαίου αυτού, ο αλγόριθμος πίσω διάδοσης επιτρέπει στα πρόσω - τροφοδότησης δίκτυα με περισσότερα από ένα επίπεδα να λύσουν οποιοδήποτε πρόβλημα ταξινόμησης προτύπων. Ωστόσο, ο αλγόριθμος αυτός ενέχει και κάποια μειονεκτήματα, από τα οποία σημαντικότερο είναι ο μεγάλος χρόνος εκπαίδευσης που απαιτείται (Kröse, Smagt, 1996). Κάποια άλλα προβλήματα που προκύπτουν κατά την εκπαίδευση ενός νευρωνικού δικτύου είναι τα εξής:

### Το πρόβλημα της αρχικοποίησης

Το πρόβλημα της αρχικοποίησης αφορά στις αρχικές τιμές που λαμβάνουν οι ελεύθερες παράμετροι του δικτύου. Η επιλογή των αρχικών αυτών τιμών πρέπει να γίνεται με ιδιαίτερη προσοχή, κυρίως αν δεν έχουμε προηγούμενες πληροφορίες και είναι προτιμότερο να είναι ομοιόμορφα καταναμημένες σε μια μικρή περιοχή. Η επιλογή ακατάλληλων τιμών για τα συναπτικά βάρη μπορεί να οδηγήσει σε

φαινομενική σταθεροποίηση του ολικού τετραγωνικού σφάλματος κατά τα διάρκεια μιας χρονικής περιόδου του χρόνου εκπαίδευσης, χωρίς όμως να έχει επιτευχθεί και τοπικό ελάχιστο. Το φαινόμενο αυτό ονομάζεται *πρόωρος κορεσμός*, καθώς το τετραγωνικό σφάλμα στο τέλος της χρονικής αυτής περιόδου συνεχίζει να μειώνεται (Λυκοθανάσης, Γεωργόπουλος, 2003).

#### *Η παράλυση του δικτύου*

Η παράλυση του δικτύου αφορά στη στασιμότητά του και την ανικανότητά του να ρυθμίσει περαιτέρω τα συναπτικά βάρη. Μια τέτοια κατάσταση προκύπτει ως εξής: κατά τη διαδικασία εκπαίδευσης, τα συναπτικά βάρη μπορεί να λάβουν πολύ μεγάλες τιμές (απόλυτα), με αποτέλεσμα η ολική είσοδος σε έναν νευρώνα (κρυφό ή εξόδου) να λάβει επίσης πολύ μεγάλες τιμές, θετικές ή αρνητικές. Χρησιμοποιώντας μια σιγμοειδή συνάρτηση ενεργοποίησης, η έξοδος από τον νευρώνα θα τείνει στο 1 ή στο 0, πράγμα που από τις εξισώσεις (5.17) και (5.20) συνεπάγεται ότι η τοπική κλίση  $\delta$ , άρα και η μεταβολή των συναπτικών βαρών  $\Delta w$ , θα είναι μηδενικές (Kröse, Smagt, 1996). Όταν η έξοδος ενός νευρώνα πλησιάζει στο 1 ή στο 0, ο νευρώνας καλείται *κορεσμένος*, ενώ αν η έξοδος τείνει στο 1 ενώ η επιθυμητή έξοδος είναι 0, ή το αντίστροφο, ο νευρώνας καλείται *κατά λάθος κορεσμένος*. Ο κατά λάθος κορεσμός μπορεί να αποφευχθεί επιλέγοντας αρχικές τιμές συναπτικών βαρών και κατωφλιών ομοιόμορφα κατανομημένες σε μια μικρή περιοχή τιμών και έχοντας λίγους κρυφούς νευρώνες και ικανοποιητική λειτουργία του δικτύου σε γραμμικές περιοχές (Λυκοθανάσης, Γεωργόπουλος, 2003).

#### *Η παγίδευση σε τοπικά ελάχιστα*

Σε περίπλοκα δίκτυα η επιφάνεια σφάλματος αποτελείται από όρη και κοιλάδες, με αποτέλεσμα το δίκτυο να μπορεί να παγιδευτεί σε κάποιο τοπικό ελάχιστο ενώ δίπλα του υπάρχει μεγαλύτερο ελάχιστο. Το πρόβλημα αυτό μπορεί να αντιμετωπιστεί είτε με στοχαστικές μεθόδους, οι οποίες όμως είναι αργές, είτε με αύξηση του αριθμού των κρυφών νευρώνων. Το τελευταίο μειώνει την πιθανότητα παγίδευσης του δικτύου, ωστόσο υπάρχει κάποιος μέγιστος αριθμός κρυφών μονάδων που μπορούν να χρησιμοποιηθούν, πέρα από τις οποίες το σύστημα πάλι παγιδεύεται σε τοπικά ελάχιστα (Kröse, Smagt, 1996).

## **ΚΕΦΑΛΑΙΟ 6 – ΠΡΟΒΛΗΜΑΤΑ ΣΧΕΔΙΑΣΜΟΥ ΚΑΙ ΒΕΛΤΙΩΣΗ ΠΟΛΥΕΠΙΠΕΔΩΝ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ**

### **6.1 Καθορισμός Δομής Δικτύου**

Από θεωρητική άποψη, η ικανότητα των πολυ - επίπεδων πρόσω - τροφοδότησης δικτύων να προσεγγίζουν οποιαδήποτε συνάρτηση τα καθιστά ένα ιδιαίτερα χρήσιμο εργαλείο. Ωστόσο, στην πράξη, κατά την υλοποίηση ενός νευρωνικού δικτύου προκύπτουν διάφορα προβλήματα, με κυριότερο αυτό του καθορισμού της βέλτιστης διαμόρφωσης και δομής του, δηλαδή του αριθμού των κρυφών επιπέδων και των κόμβων σε κάθε κρυφό επίπεδο.

Το πρόβλημα αυτό μπορεί να αντιμετωπιστεί με τη χρήση της μεθόδου δοκιμής και σφάλματος (trial and error method), η οποία όμως είναι εξαιρετικά χρονοβόρα και επίπονη. Αρχικά και με τη χρήση κάποιων εμπειρικών κανόνων κατασκευάζεται δίκτυο συγκεκριμένης αρχιτεκτονικής, το οποίο εκπαιδεύεται. Αν μετά την εκπαίδευση το σφάλμα στην έξοδο του δικτύου είναι μεγάλο, μεταβάλλεται η δομή του δικτύου (αλλάζοντας τον αριθμό κρυφών επιπέδων ή/και των κόμβων) και το δίκτυο εκπαιδεύεται ξανά. Έτσι, με διαδοχικές δοκιμές μειώνεται όλο και περισσότερο το σφάλμα της προσέγγισης και καταλήγουμε σε κάποια βέλτιστη αρχιτεκτονική δικτύου. Τα τελευταία χρόνια βέβαια έχουν εμφανιστεί κάποιες καλύτερες μέθοδοι για την αντιμετώπιση του προβλήματος, όπως η μέθοδος του κλαδέματος (pruning), της ανάπτυξης (growing) και της εξέλιξης νευρωνικών δικτύων (Λυκοθανάσης, Γεωργόπουλος, 2003).

### **6.2 Σφάλματα Προσομοίωσης**

Κατά την εκπαίδευση ενός νευρωνικού δικτύου πολλών επιπέδων, το σφάλμα κατά τη διαδικασία προσέγγισης εξαρτάται από διάφορους παράγοντες, όπως

- ο αλγόριθμος εκπαίδευσης και ο αριθμός των επαναλήψεων, που καθορίζουν πόσο καλά ελαχιστοποιείται το σφάλμα του συνόλου των εκπαιδευτικών προτύπων.
- ο αριθμός των εκπαιδευτικών προτύπων, που καθορίζει κατά πόσο αυτά είναι αντιπροσωπευτικά της πραγματικής συνάρτησης.

- ο αριθμός των κρυφών μονάδων, που καθορίζει την ικανότητα του δικτύου να εκφράζει τη συνάρτηση. Όσο πιο ομαλή είναι αυτή, τόσο μικρότερος είναι ο αριθμός των απαιτούμενων κρυφών μονάδων, ενώ όσο πιο απότομες διακυμάνσεις παρουσιάζει, τόσο περισσότεροι κρυφοί νευρώνες απαιτούνται.

Τα νευρωνικά δίκτυα εκπαιδεύονται έτσι ώστε να ελαχιστοποιούν το σφάλμα του συνόλου των εκπαιδευτικών προτύπων που χρησιμοποιούνται. Το μέσο σφάλμα ανά εκπαιδευτικό πρότυπο, το οποίο υπολογίζεται κατά τη διαδικασία εκπαίδευσης, λέγεται *μέσο σφάλμα εκπαίδευσης (learning error rate)* και δίνεται από τη σχέση

$$E_{learning} = \frac{1}{P_{learning}} \sum_{p=1}^{P_{learning}} E^p \quad (6.1)$$

όπου  $E^p$  είναι η διαφορά της επιθυμητής από την πραγματική έξοδο και ισούται με

$$E^p = \frac{1}{2} \sum_{o=1}^{N_o} (d_o^p - y_o^p)^2 \quad (6.2)$$

Το πραγματικό σφάλμα του δικτύου προκύπτει από την ολοκλήρωση των διαφορών επιθυμητής και πραγματικής εξόδου σε όλο το πεδίο εισόδου και μπορεί να υπολογιστεί θεωρώντας ότι έχουμε έναν πολύ μεγάλο αριθμό προτύπων, το λεγόμενο σύνολο δοκιμής (test set). Έτσι, μπορεί να οριστεί το μέσο σφάλμα του συνόλου δοκιμής, που ισούται με

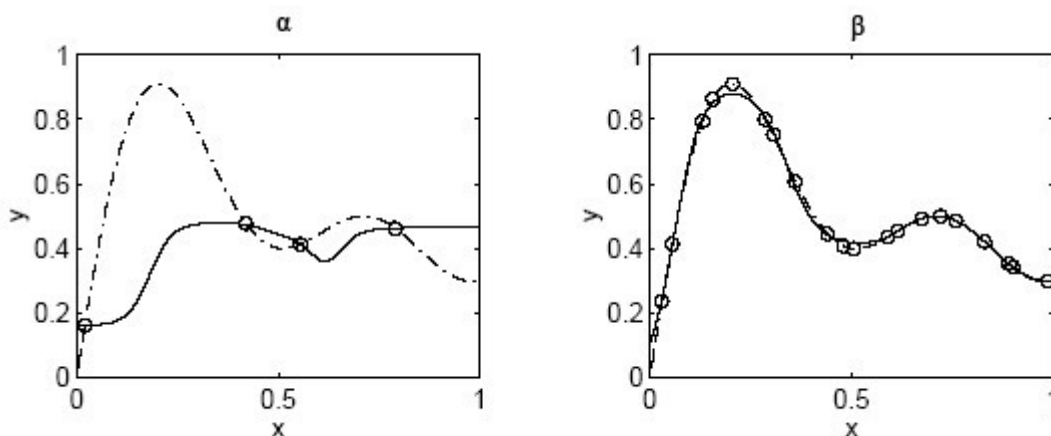
$$E_{test} = \frac{1}{P_{test}} \sum_{p=1}^{P_{test}} E^p \quad (6.3)$$

και ονομάζεται *μέσο σφάλμα δοκιμής (test error rate)* (Köröse, Smagt, 1996). Παρακάτω παρουσιάζεται η επίδραση που έχουν ο αριθμός των εκπαιδευτικών προτύπων και των κρυφών μονάδων στα προαναφερόμενα σφάλματα.

### 6.2.1 Επίδραση του Αριθμού των Εκπαιδευτικών Προτύπων

Για να μελετηθεί η επίδραση του αριθμού των προτύπων διανυσμάτων εισόδου, που χρησιμοποιούνται για την εκπαίδευση του δικτύου, στο μέσο σφάλμα εκπαίδευσης και στο μέσο σφάλμα δοκιμής, παρατίθεται το εξής παράδειγμα: ένα νευρωνικό δίκτυο, αποτελούμενο από μία είσοδο, πέντε κρυφές μονάδες με σιγμοειδή συνάρτηση ενεργοποίησης και μία γραμμική έξοδο, καλείται να προσεγγίσει μια συνάρτηση της μορφής  $y = f(x)$ . Στην εικόνα που ακολουθεί παρουσιάζονται η πραγματική συνάρτηση και η προσέγγιση του δικτύου στις περιπτώσεις που

εφαρμόζονται 4 εκπαιδευτικά πρότυπα (Εικόνα 6.1α) και 20 εκπαιδευτικά πρότυπα (Εικόνα 6.1β). Η πραγματική συνάρτηση παριστάνεται με τη διακεκομμένη γραμμή, η προσέγγιση του δικτύου με τη συνεχή και τα εκπαιδευτικά πρότυπα με τους κύκλους.

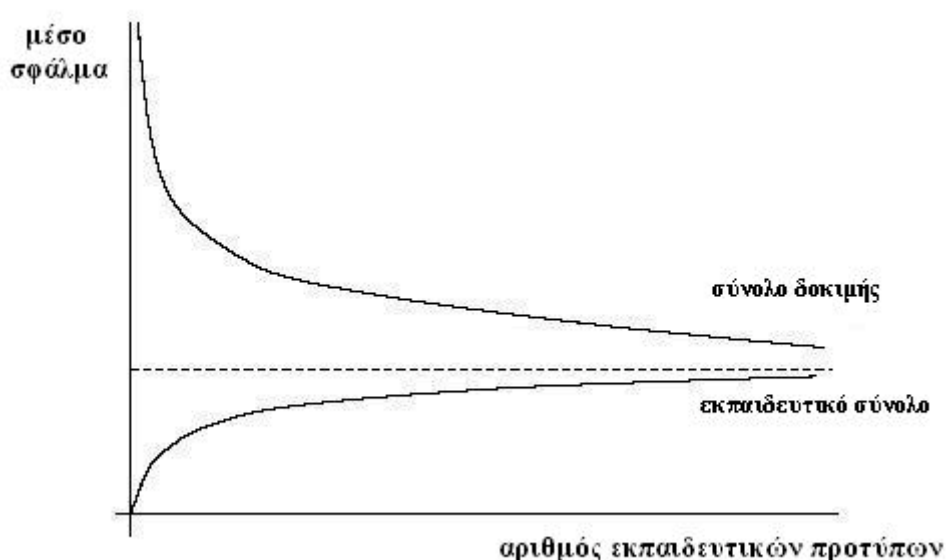


**Εικόνα 6.1:** Επίδραση του αριθμού των εκπαιδευτικών προτύπων στην προσέγγιση του δικτύου α) για 4 εκπαιδευτικά πρότυπα β) για 20 εκπαιδευτικά πρότυπα (Kröse, Smagt, 1996)

Όπως φαίνεται στην εικόνα 6.1α, το μέσο σφάλμα εκπαίδευσης  $E_{\text{learning}}$  είναι μικρό καθώς η συνεχής γραμμή, δηλαδή η έξοδος του δικτύου, διέρχεται ακριβώς από τα εκπαιδευτικά πρότυπα. Αντιθέτως, το μέσο σφάλμα δοκιμής  $E_{\text{test}}$  είναι μεγάλο. Στην εικόνα 6.1β παρουσιάζεται η προσέγγιση του δικτύου όταν έχουμε 20 εκπαιδευτικά πρότυπα. Στην περίπτωση αυτή, το μέσο σφάλμα εκπαίδευσης  $E_{\text{learning}}$  είναι μεγαλύτερο από όταν έχουμε 4 εκπαιδευτικά πρότυπα, ενώ το μέσο σφάλμα δοκιμής  $E_{\text{test}}$  είναι μικρότερο.

Στην εικόνα 6.2 παρουσιάζεται η επίδραση του αριθμού των εκπαιδευτικών προτύπων στα μέσα σφάλματα, όπως αυτά προέκυψαν μετά από πειράματα, κατά τα οποία χρησιμοποιήθηκαν διάφορα σύνολα εκπαιδευτικών προτύπων και τα οποία επαναλήφθηκαν δέκα φορές για κάθε σύνολο προτύπων. Παρατηρείται ότι το μέσο σφάλμα εκπαίδευσης αυξάνεται με την αύξηση του μεγέθους του συνόλου των εκπαιδευτικών προτύπων, ενώ το μέσο σφάλμα δοκιμής μειώνεται με την αύξηση του μεγέθους του συνόλου των εκπαιδευτικών προτύπων, πράγμα που σημαίνει ότι ένα μικρό μέσο σφάλμα εκπαίδευσης, όταν έχουμε μικρό αριθμό προτύπων μάθησης, δε συνεπάγεται και καλή απόδοση του δικτύου. Αυξάνοντας τον αριθμό των

εκπαιδευτικών προτύπων, τα δύο μέσα σφάλματα συγκλίνουν σε μία τιμή, η οποία εξαρτάται από την αντιπροσωπευτικότητα του δικτύου, δηλαδή από την ικανότητά του να προσεγγίζει καλά τη συνάρτηση. Αν το μέσο σφάλμα εκπαίδευσης και το μέσο σφάλμα δοκιμής δε συγκλίνουν, τότε κατά τη διαδικασία μάθησης δεν έχει βρεθεί σφαιρικό ελάχιστο (Kröse, Smagt, 1996).

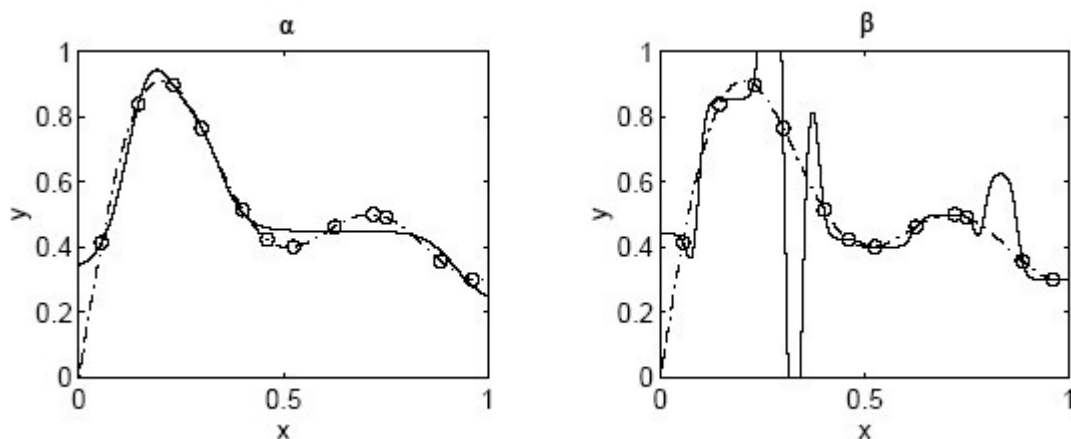


**Εικόνα 6.2:** Επίδραση του αριθμού των εκπαιδευτικών προτύπων στα μέσα σφάλματα εκπαίδευσης και δοκιμής (Kröse, Smagt, 1996)

### 6.2.2 Επίδραση του Αριθμού των Κρυφών Μονάδων

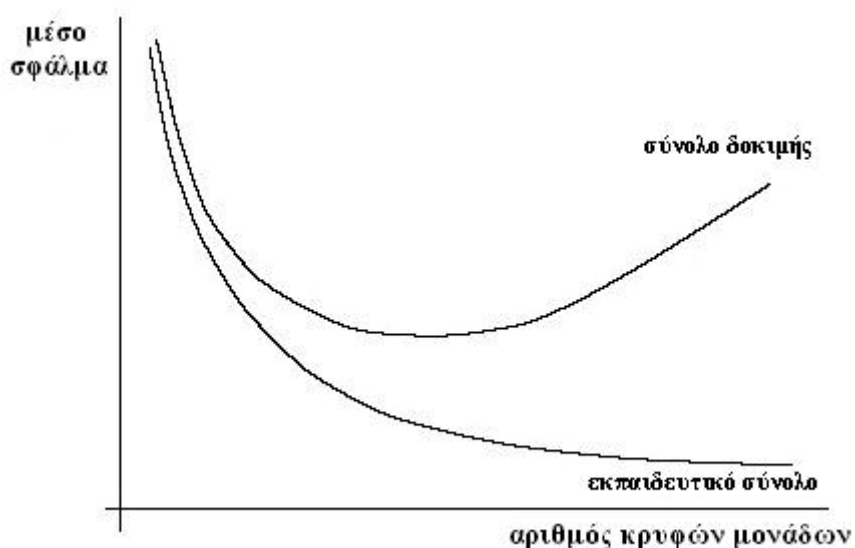
Και σε αυτή την περίπτωση χρησιμοποιείται το παράδειγμα της προηγούμενης παραγράφου, μόνο που τώρα χρησιμοποιείται σταθερός αριθμός εκπαιδευτικών προτύπων (12 εκπαιδευτικά πρότυπα) και μεταβάλλεται ο αριθμός των κρυφών μονάδων. Στην εικόνα 6.3 που ακολουθεί παρουσιάζονται η πραγματική συνάρτηση και η προσέγγιση του δικτύου στις περιπτώσεις που έχουμε 5 κρυφές μονάδες (Εικόνα 6.3α) και 20 κρυφές μονάδες (Εικόνα 6.3β). Όπως παρατηρείται στην εικόνα 6.3β, η έξοδος του δικτύου ταιριάζει ακριβώς στα εκπαιδευτικά πρότυπα, ωστόσο λόγω του μεγάλου αριθμού κρυφών μονάδων, η συνάρτηση που παράγει το δίκτυο παρουσιάζει πολύ περισσότερες «διακυμάνσεις» από την πραγματική. Το φαινόμενο αυτό ονομάζεται υπερ - εκπαίδευση (overtraining).





**Εικόνα 6.3:** Επίδραση του αριθμού των κρυφών μονάδων στην προσέγγιση του δικτύου α) για 5 κρυφές μονάδες β) για 20 κρυφές μονάδες (Kröse, Smagt, 1996)

Γενικά, η προσθήκη κρυφών μονάδων μειώνει το μέσο σφάλμα εκπαίδευσης αλλά δε μειώνει απαραίτητα και το μέσο σφάλμα δοκιμής, το οποίο αρχικά μειώνεται αλλά στη συνέχεια αυξάνεται. Το φαινόμενο αυτό ονομάζεται *peaking effect* (Εικόνα 6.4) (Kröse, Smagt, 1996).

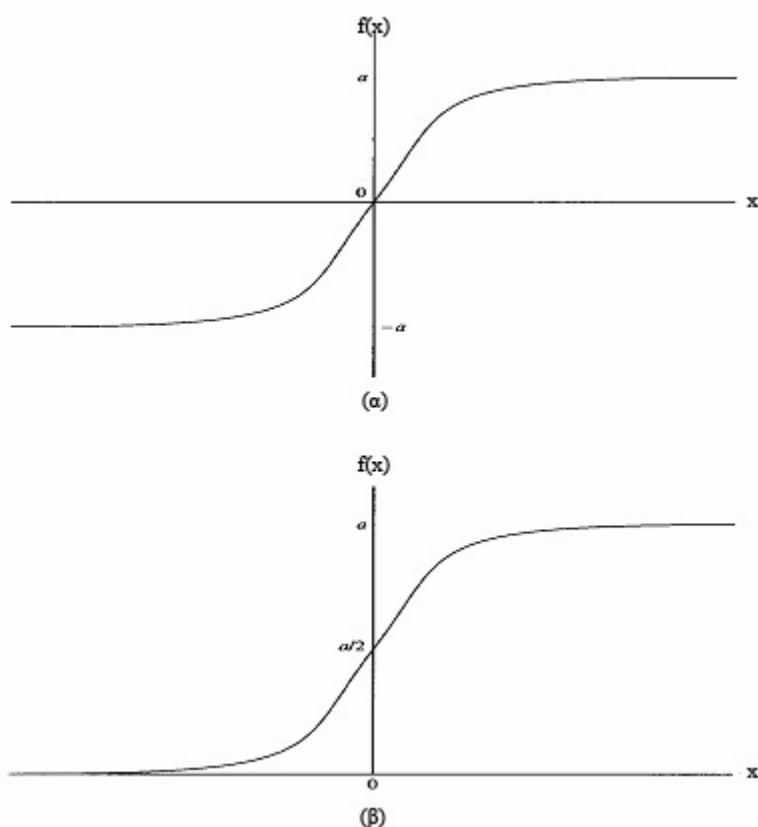


**Εικόνα 6.4:** Επίδραση του αριθμού των κρυφών μονάδων στα μέσα σφάλματα εκπαίδευσης και δοκιμής (Kröse, Smagt, 1996)

### 6.3 Τρόποι Βελτίωσης της Προσομοίωσης

Κατά το σχεδιασμό ενός νευρωνικού δικτύου πολλών επιπέδων που χρησιμοποιεί τον αλγόριθμο πίσω διάδοσης, υπεισέρχονται πολλοί και διάφοροι παράγοντες, που καθιστούν τη διαδικασία αυτή αρκετά δύσκολη και περίπλοκη. Ωστόσο, υπάρχουν κάποιοι τρόποι που εξασφαλίζουν την καλύτερη λειτουργία του αλγόριθμου πίσω διάδοσης, οι οποίοι παρουσιάζονται ακριβώς παρακάτω (Λυκοθανάσης, Γεωργόπουλος, 2003).

- Η διαδικασία εκπαίδευσης ενός νευρωνικού δικτύου πολλών επιπέδων μπορεί να επιταχυνθεί (δηλαδή να απαιτούνται λιγότερες επαναλήψεις) όταν η στιγμοειδής συνάρτηση ενεργοποίησης που χρησιμοποιείται στις μονάδες είναι ασυμμετρική παρά μη συμμετρική (Εικόνα 6.5). Για μια ασυμμετρική συνάρτηση ισχύει  $f(-x) = -f(x)$ .



**Εικόνα 6.5:** (α) Ασυμμετρική συνάρτηση (β) μη συμμετρική συνάρτηση (Λυκοθανάσης, Γεωργόπουλος, 2003)

- Οι επιθυμητές εξοδοι του δικτύου πρέπει να βρίσκονται μέσα στο διάστημα της σιγμοειδούς συνάρτησης ενεργοποίησης. Συγκεκριμένα, για κάποια μονάδα εξόδου  $k$ , η επιθυμητή έξοδος  $d_k$  θα πρέπει να αυξηθεί ή να μειωθεί κατά μία τιμή  $\varepsilon$ , αλλιώς ο αλγόριθμος πίσω διάδοσης θα οδηγεί τις ελεύθερες παραμέτρους του δικτύου στο άπειρο, επιβραδύνοντας έτσι τη διαδικασία μάθησης. Έτσι, στην περίπτωση της ασυμμετρικής συνάρτησης της εικόνας 6.5α και για μια τιμή  $a$ , η επιθυμητή έξοδος τίθεται ίση με  $d_k = a - \varepsilon$ , όπου  $\varepsilon$  μια κατάλληλη θετική σταθερά.
- Προκειμένου να μειωθεί η πιθανότητα κορεσμού των νευρώνων και παραγωγής μικρών κλίσεων σφάλματος, οι αρχικές τιμές των ελεύθερων παραμέτρων (συναπτικών βαρών και κατωφλιών) πρέπει να είναι τυχαίες και ομοιόμορφα κατανεμημένες σε ένα μικρό διάστημα. Το διάστημα αυτό ωστόσο δεν πρέπει να είναι πάρα πολύ μικρό, καθώς τότε οι πολύ μικρές κλίσεις σφάλματος καθιστούν τη διαδικασία μάθησης εξαιρετικά αργή.
- Όλοι οι νευρώνες ενός νευρωνικού δικτύου πρέπει να μαθαίνουν με την ίδια ταχύτητα. Ωστόσο, σε ένα νευρωνικό δίκτυο τα τελευταία επίπεδα έχουν την τάση να παράγουν μεγαλύτερες τοπικές κλίσεις σε σχέση με τα πρώτα επίπεδα. Το πρόβλημα αυτό αντιμετωπίζεται θέτοντας μικρότερη τιμή ρυθμού εκμάθησης ( $\gamma$ ) στα τελευταία επίπεδα σε σχέση με τα προηγούμενα. Επίσης, στους νευρώνες με πολλές εισόδους θα πρέπει να τίθεται μικρότερη τιμή ρυθμού εκμάθησης σε σχέση με νευρώνες που έχουν λίγες εισόδους.
- Για εφαρμογές πραγματικού χρόνου, η εκπαίδευση του δικτύου με τον τρόπο προτύπων (δηλαδή πρότυπο προς πρότυπο) είναι προτιμότερη σε σχέση με τον σωρηδόν τρόπο, καθώς είναι ταχύτερη.
- Τα πρότυπα εκπαίδευσης πρέπει να εφαρμόζονται στο δίκτυο με τυχαία σειρά.
- Κατά τη διαδικασία μάθησης από ένα σύνολο προτύπων εκπαίδευσης, το δίκτυο εξάγει τις πληροφορίες που περιέχονται στα πρότυπα και κατασκευάζει μια συνάρτηση η οποία αντιστοιχίζει την είσοδο με την έξοδο.

## **ΚΕΦΑΛΑΙΟ 7 – ΧΡΗΣΗ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ ΣΤΗΝ ΠΡΟΣΕΓΓΙΣΗ ΦΥΣΙΚΩΝ ΜΟΝΤΕΛΩΝ ΚΑΙ ΣΤΗ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗ ΠΕΡΙΒΑΛΛΟΝΤΙΚΩΝ ΠΡΟΒΛΗΜΑΤΩΝ**

### **7.1 Εισαγωγή**

Στον τομέα της μηχανικής των υπογείων υδάτων (groundwater engineering) χρησιμοποιούνται κατά κόρον φυσικά μοντέλα, που προσομοιώνουν την υπόγεια ροή και τη μεταφορά ρύπων στο υπέδαφος, υπολογίζοντας υδραυλικά ύψη, ταχύτητα του υπογείου ύδατος, συγκέντρωση και χρόνο μεταφοράς των ρύπων κλπ. Τα μοντέλα αυτά χρησιμοποιούν αριθμητικές μεθόδους, με συνηθέστερη τη Μέθοδο των Πεπερασμένων Στοιχείων (Finite Element Method, FEM), πράγμα που σημαίνει ότι καλούνται να επιλύσουν πολυδιάστατα συστήματα μερικών διαφορικών εξισώσεων με πολλές μεταβλητές προκειμένου να δώσουν ακριβή και σωστά αποτελέσματα (Arndt et al, 2005). Σε πολλές μάλιστα πραγματικές εφαρμογές, τα μοντέλα αυτά συνδυάζονται με κάποιον αλγόριθμο βελτιστοποίησης (Global Optimization Algorithm, GOA) (Maskey et al, 2000) με σκοπό π.χ. να βρεθεί ο βέλτιστος αριθμός και η βέλτιστη θέση γεωτρήσεων άντλησης/εμπλουτισμού με σκοπό την εξυγίανση ρυπασμένου υδροφορέα (όπως συμβαίνει στη μέθοδο pump and treat – άντλησης και επεξεργασίας) ή ο μέγιστος επιτρεπόμενος ρυθμός άντλησης νερού από παράκτιες γεωτρήσεις, ούτως ώστε να μη δημιουργηθεί ή ενταθεί το φαινόμενο της υφαλμύρωσης.

### **7.2 Προσέγγιση Συμβατικών Μοντέλων Προσομοίωσης με Χρήση Νευρωνικών Δικτύων**

Η χρήση συμβατικών μοντέλων απλά και μόνο για την προσομοίωση του φυσικού συστήματος απαιτεί αρκετό χρόνο προκειμένου αυτά να «τρέξουν» και να δώσουν αποτελέσματα, λόγω του μεγάλου υπολογιστικού φόρτου. Λύση σε αυτό το πρόβλημα δίνουν τα νευρωνικά δίκτυα, τα οποία μπορούν να προβλέψουν την απόκριση του συστήματος χωρίς να χρειάζεται να γνωρίζουν ή να κατανοούν τη φύση των φαινομένων τα οποία μοντελοποιούν. Έτσι, μπορούν να προσεγγίσουν ακόμα και μη γραμμικές σχέσεις μεταξύ των δεδομένων εισόδου και εξόδου, χωρίς

να λαμβάνουν υπόψη φυσικές διαδικασίες και μαθηματικές εξισώσεις που χαρακτηρίζουν το σύστημα ροής. Ως εκ τούτου, ένα νευρωνικό δίκτυο είναι πολύ πιο γρήγορο από το φυσικό μοντέλο το οποίο προσεγγίζει (Maskey et al, 2000).

### **7.3 Χρήση Νευρωνικών Δικτύων στη Βελτιστοποίηση Προβλημάτων Μηχανικής Υπογείων Υδάτων**

Η χρήση τεχνικών βελτιστοποίησης για την κατάστροψη βέλτιστης στρατηγικής διαχείρισης ενέχει το μειονέκτημα ότι χρειάζεται μοντέλα άμεσης ανταπόκρισης για την εξαγωγή αποτελεσμάτων σε σχετικά μικρά χρονικά διαστήματα. Στο σημείο αυτό τα συμβατικά μοντέλα και πάλι υστερούν, καθώς για να ενοποιηθούν με το πρόγραμμα βελτιστοποίησης χρειάζονται εκτελέσιμα προγράμματα που να μετατρέπουν τα αρχεία εξαγωγής του μοντέλου βελτιστοποίησης σε αρχεία εισαγωγής στο συμβατικό μοντέλο και να διαβάζουν τη φόρμα εξαγωγής των αποτελεσμάτων του συμβατικού μοντέλου και να την τροφοδοτούν στο μοντέλο βελτιστοποίησης. Ως εκ τούτου, απαιτείται αρκετός χρόνος για την εύρεση βέλτιστης λύσης.

Στην περίπτωση όμως που τα φυσικά μοντέλα αντικαθίστανται από νευρωνικά δίκτυα, η διαδικασία βελτιστοποίησης είναι πολύ πιο γρήγορη αφού δε χρειάζονται εκτελέσιμα προγράμματα. Το νευρωνικό δίκτυο είναι φτιαγμένο με τέτοιο τρόπο ώστε να δέχεται απευθείας ως είσοδο τα αρχεία εξόδου του προγράμματος βελτιστοποίησης και να τα διαμορφώνει έτσι ώστε αυτά να είναι αποδεκτά σαν δεδομένα εισόδου σε αυτό (Maskey et al, 2000). Γενικά, τα νευρωνικά δίκτυα χρησιμοποιούνται ευρέως για την προσέγγιση συναρτήσεων «μαύρου κουτιού» (black box functions), δηλαδή συναρτήσεων για τις οποίες υπάρχει ελάχιστη (ή καθόλου) γνώση των μαθηματικών τους ιδιοτήτων. Έτσι, προσεγγίζουν αρκετά καλά την αντικειμενική συνάρτηση και τους περιορισμούς του προβλήματος βελτιστοποίησης, τα οποία συνήθως είναι άγνωστα (Arndt et al, 2005).

#### **7.4 Συμπεράσματα**

Από τα παραπάνω γίνεται κατανοητή η υπεροχή των νευρωνικών δικτύων έναντι στα φυσικά μοντέλα προσομοίωσης, καθώς η απλότητά τους, η ικανότητα γρήγορης προσομοίωσης και η δυνατότητά τους να συνεργάζονται εύκολα με μοντέλα βελτιστοποίησης αποτελούν τα πιο σημαντικά τους πλεονεκτήματα. Παρ' όλα αυτά, είναι ξεκάθαρο ότι τα νευρωνικά δίκτυα πρέπει να εκπαιδεύονται με όσο το δυνατόν περισσότερα και ακριβή αποτελέσματα, έτσι ώστε να παρουσιάζουν υψηλή απόδοση. Ιδιαίτερης σημασίας τυγχάνει το γεγονός ότι η διαδικασία βελτιστοποίησης μπορεί να γίνει πιο γρήγορη και ακριβής με χρήση αρχικά ενός νευρωνικού δικτύου για την εύρεση του διαστήματος αναζήτησης όπου υπάρχει μεγαλύτερη πιθανότητα να βρεθεί η βέλτιστη λύση κι έπειτα ενός φυσικού μοντέλου προσομοίωσης για περαιτέρω έρευνα σε αυτή την περιοχή. Ακόμη, νευρωνικό δίκτυο μπορεί να χρησιμοποιηθεί σε πειράματα εκτίμησης της απόδοσης διάφορων μοντέλων βελτιστοποίησης, στα οποία η διαδικασία βελτιστοποίησης πρέπει να επαναληφθεί αρκετές φορές (Maskey et al, 2000).

## ΚΕΦΑΛΑΙΟ 8 – ΧΡΗΣΙΜΟΠΟΙΟΥΜΕΝΟ ΝΕΥΡΩΝΙΚΟ ΔΙΚΤΥΟ

### 8.1 Περιγραφή Νευρωνικού Δικτύου

Το νευρωνικό δίκτυο που χρησιμοποιείται στην παρούσα εργασία είναι ένας πολυ-επίπεδος Perceptron (Multi-layer Perceptron, MLP) που εκπαιδεύεται μέσω επιβλεπόμενης μάθησης με χρήση του αλγόριθμου πίσω – διάδοσης του σφάλματος (error back-propagation algorithm) (Nikolos, 2004).

Όσον αφορά τη δομή του, το δίκτυο αποτελείται από:

- το επίπεδο εισόδου, το οποίο αποτελείται από 5 κόμβους, όσες δηλαδή και οι ανεξάρτητες μεταβλητές απόφασης του προβλήματος βελτιστοποίησης (παροχές άντλησης των πέντε επιλεχθέντων γεωτρήσεων)
- το επίπεδο εξόδου, στο οποίο υπάρχουν 22 κόμβοι, όσα δηλαδή και τα πηγάδια παρατήρησης, στα οποία λαμβάνονται οι μη ρητοί περιορισμοί του προβλήματος βελτιστοποίησης (αναλυτικά η επιλογή των πηγαδιών παρατήρησης και των περιορισμών παρουσιάζεται στην παράγραφο 10.1)
- ένα κρυφό επίπεδο που περιλαμβάνει 40 κόμβους. Γενικά, τα κρυφά επίπεδα πρέπει να παρουσιάζουν μεγαλύτερη πολυπλοκότητα από τα επίπεδα εισόδου και εξόδου και αρκετά μεγάλη μη γραμμικότητα.

Το δίκτυο είναι πλήρως διασυνδεδεμένο, δηλαδή οι κόμβοι σε κάθε επίπεδο συνδέονται με όλους τους κόμβους του προηγούμενου επιπέδου.

Όλες οι μονάδες περιλαμβάνουν μια σιγμοειδή συνάρτηση ενεργοποίησης και η ολική είσοδος σε κάθε μονάδα ισούται με το σταθμισμένο άθροισμα των εξόδων των μονάδων με τις οποίες συνδέεται.

Η εκπαίδευση του δικτύου γίνεται με την εφαρμογή ενός συνόλου διανυσμάτων εισόδου και των αντίστοιχων διανυσμάτων εξόδου. Στην περίπτωση μας, τα διανύσματα εισόδου είναι οι παροχές άντλησης και τα διανύσματα εξόδου τα υδραυλικά ύψη στα πηγάδια παρατήρησης. Τα συναπτικά βάρη προσαρμόζονται βάσει του γενικευμένου κανόνα δέλτα (Nikolos, 2004).

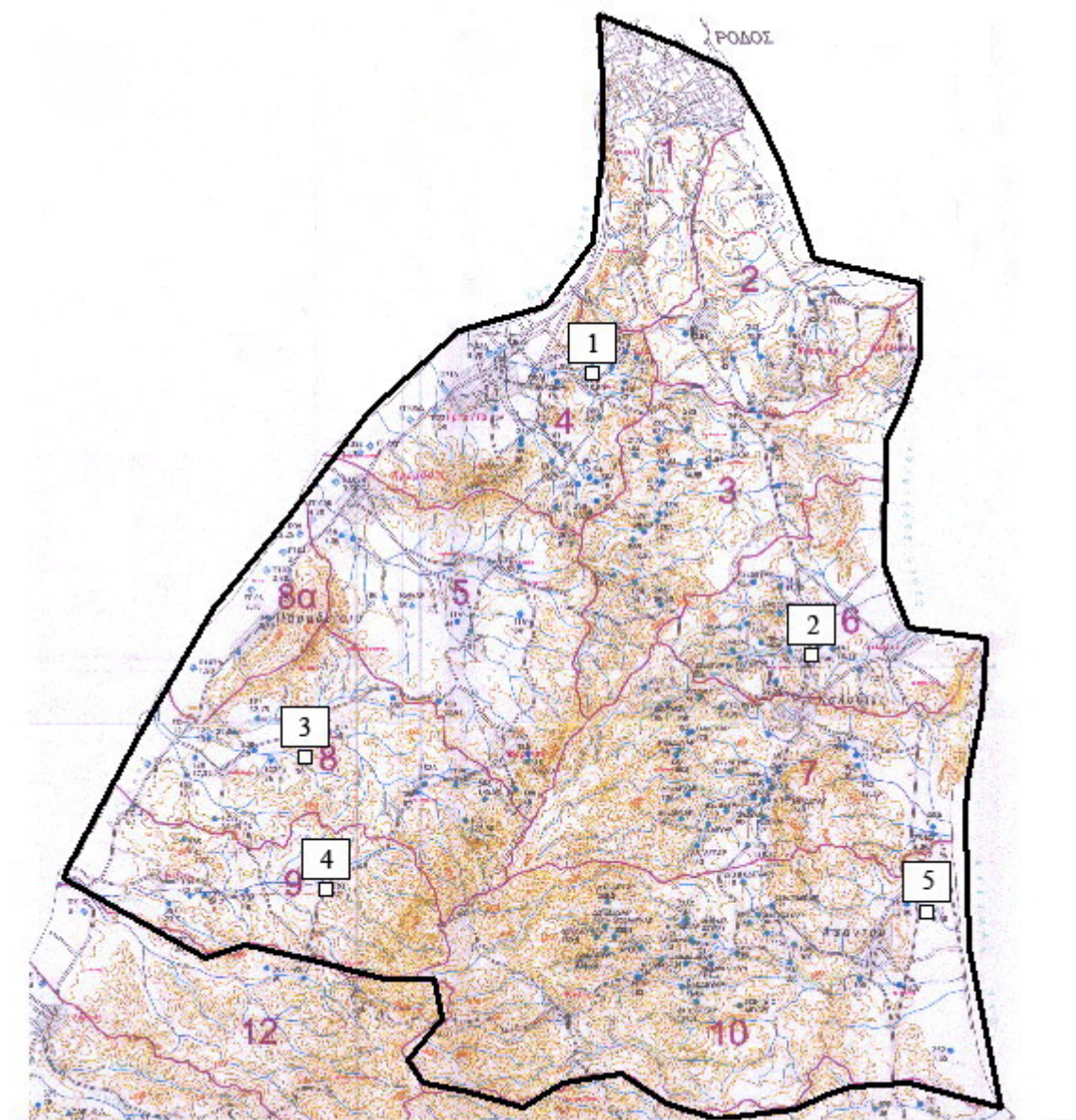
## 8.2 Εκπαίδευση Νευρωνικού Δικτύου

Όπως αναφέρθηκε στην παράγραφο 1.2, από τις 23 γεωτρήσεις της περιοχής μελέτης για τις οποίες υπάρχουν δεδομένα της στάθμης του ύδατος, επιλέγονται 5 για την εκπαίδευση του νευρωνικού δικτύου με σκοπό την εξοικονόμηση χρόνου και κόπου. Οι πέντε γεωτρήσεις που επιλέγονται έχουν τις μεγαλύτερες τιμές παροχών άντλησης και είναι διεσπαρμένες στην υπό μελέτη περιοχή, ούτως ώστε να είναι οι πιο αντιπροσωπευτικές του συνόλου των γεωτρήσεων. Στην εικόνα 8.1 παρουσιάζονται οι επιλεγθείσες γεωτρήσεις, ενώ στον πίνακα 8.1 παρουσιάζεται η αντιστοιχία αρίθμησης των γεωτρήσεων του χάρτη Ι – 2 του παραρτήματος Ι με τις γεωτρήσεις της εικόνας 8.1.

**Πίνακας 8.1:** Παλιά και νέα αρίθμηση γεωτρήσεων περιοχής μελέτης

Αρίθμηση γεώτρησης στον χάρτη του παραρτήματος Ι	Νέα αρίθμηση γεώτρησης
88Α	1
4 (Φ4ΔΕΥΑΡ)	2
122	3
383	4
413	5





**Εικόνα 8.1:** Γεωτρήσεις περιοχής μελέτης, που χρησιμοποιήθηκαν για την εκπαίδευση του νευρωνικού δικτύου και νέα αρίθμηση αυτών

Η προσομοίωση με το μοντέλο PTC γίνεται για 16 περιόδους των 180 ημερών η κάθε μία, δηλαδή για 8 χρόνια (1997 – 2005). Τα πρότυπα διανύσματα για την εκπαίδευση του νευρωνικού δικτύου δημιουργούνται βάσει του αρχείου PTC\_Mesh\_heads\_s16.fin που δημιουργεί το PTC για τους διάφορους συνδυασμούς των παροχών άντλησης των πέντε γεωτρήσεων και αφορά τα υδραυλικά ύψη στο τέλος της 16<sup>ης</sup> περιόδου (τον Μάιο του 2005).

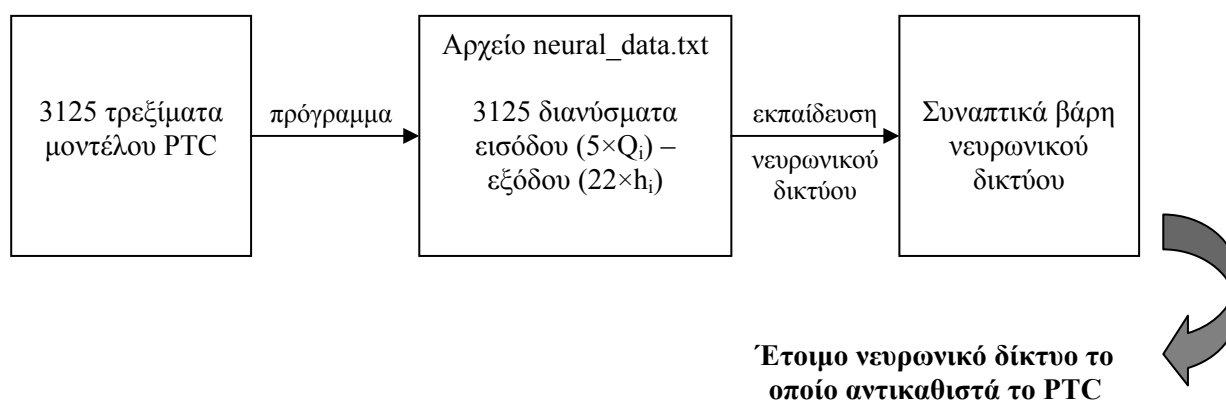
Η μέγιστη παροχή άντλησης των γεωτρήσεων (που σημειώνεται κατά τη θερινή περίοδο) κυμαίνεται από 2200 m<sup>3</sup>/d έως 3880 m<sup>3</sup>/d. Έτσι, ως μέγιστη παροχή άντλησης όλων των γεωτρήσεων τίθεται η τιμή 4000 m<sup>3</sup>/d, θεωρώντας ότι στις

γεωτρήσεις με αρκετά μικρότερη παροχή άντλησης συμπεριλαμβάνεται η επίδραση γειτονικών γεωτρήσεων.

Η δημιουργία διανυσμάτων εκπαίδευσης γίνεται θεωρώντας ότι η παροχή άντλησης κάθε γεώτρησης μπορεί να πάρει πέντε πιθανές τιμές: 0, 1000, 2000, 3000 και 4000  $\text{m}^3/\text{d}$ , πράγμα που σημαίνει ότι υπάρχουν  $5^5 = 3125$  δυνατοί συνδυασμοί παροχών άντλησης για το σύνολο των γεωτρήσεων της περιοχής. Με χρήση της γλώσσας προγραμματισμού C++ καταστρώνεται ο αλγόριθμος του παραρτήματος II, ο οποίος παράγει τους δυνατούς συνδυασμούς παροχών άντλησης.

Η προβολή των διανυσμάτων εκπαίδευσης στο δίκτυο γίνεται σε 30 κύκλους (epochs). Από τα 3125 διανύσματα εισόδου - εξόδου που δημιουργούνται από το PTC, αρχικά χρησιμοποιείται το 70% αυτών για την εκπαίδευση του νευρωνικού δικτύου ενώ το υπόλοιπο 30% χρησιμοποιείται για τη δοκιμή της αποδοτικότητας του δικτύου, το λεγόμενο testing. Δηλαδή, εισάγονται στο δίκτυο κάποια διανύσματα εισόδου για τα οποία δεν έχει εκπαιδευτεί και τα αποτελέσματα στην έξοδο συγκρίνονται με τα αποτελέσματα του μοντέλου PTC για τις δεδομένες παροχές άντλησης. Στη συνέχεια, αφού το μέσο σφάλμα δοκιμής που προκύπτει είναι αρκετά μικρό, χρησιμοποιείται 99% των δεδομένων για την εκπαίδευση του δικτύου και 1% για testing. Στην περίπτωση αυτή, το μέσο σφάλμα εκπαίδευσης που προκύπτει ισούται με  $1,55 \cdot 10^{-2}$  ή 1,55%, πράγμα που συνεπάγεται ότι το δίκτυο έχει εκπαιδευτεί αρκετά καλά.

Σχηματικά, η διαδικασία εκπαίδευσης του νευρωνικού δικτύου παρουσιάζεται στην εικόνα που ακολουθεί.



**Εικόνα 8.2:** Σχηματική αναπαράσταση εκπαίδευσης νευρωνικού δικτύου με χρήση των αποτελεσμάτων του μοντέλου PTC

## ΚΕΦΑΛΑΙΟ 9 – ΕΙΣΑΓΩΓΗ ΣΤΟΥΣ ΕΞΕΛΙΚΤΙΚΟΥΣ ΑΛΓΟΡΙΘΜΟΥΣ

### 9.1 Ιστορική Αναδρομή και Αναλογία με Βιολογία

Οι Εξελικτικοί Αλγόριθμοι (Evolutionary Algorithms, EA) αποτελούν μια κατηγορία μεθόδων βελτιστοποίησης, των οποίων η λειτουργία στηρίζεται στη μίμηση των διαδικασιών της φυσικής εξέλιξης. Παρόλο που εφευρέθηκαν στο τέλος της δεκαετίας του 1950, η επιστημονική κοινότητα δεν ασχολήθηκε με τον τομέα αυτό για περίπου τριάντα χρόνια, τόσο λόγω της απουσίας ηλεκτρονικών υπολογιστών, που θα μπορούσαν να τους εκμεταλλευτούν, όσο και λόγω των ελλείψεων των πρώτων προσεγγίσεων σε αυτό το θέμα. Ωστόσο, μετά τη δεκαετία του 1960, επιστήμονες όπως οι Holland (Holland, 1962), Rechenberg (Rechenberg, 1965), Schwefel (Schwefel, 1968) και Fogel (Fogel, 1962) παρουσίασαν κάποιες διαφορετικές προσεγγίσεις του θέματος, που όμως είχαν ως κοινό άξονα τη χρησιμοποίηση ενός πληθυσμού πιθανών λύσεων οι οποίες, με χρήση μεθόδων προερχόμενων από τη γενετική, εξελίσσονταν και κατέληγαν σε σχεδόν βέλτιστες λύσεις.

Παρόλο που σήμερα εμφανίζονται διάφορες τροποποιήσεις των Εξελικτικών Αλγορίθμων, υπάρχουν τρεις βασικές κατηγορίες οι οποίες αναπτύσσονται εντελώς ξεχωριστά αλλά συνδέονται άρρηκτα μεταξύ τους:

- οι Γενετικοί Αλγόριθμοι (Genetic Algorithms, GA), παρακλάδι των οποίων αποτελεί ο Γενετικός Προγραμματισμός (Genetic Programming, GP)
- ο Εξελικτικός Προγραμματισμός (Evolutionary Programming, EP)
- οι Εξελικτικές Στρατηγικές (Evolution Strategies, ES)

Συγκεκριμένα, οι Γενετικοί Αλγόριθμοι, που εφευρέθηκαν από τον Holland (Holland, 1962), μοιάζουν περισσότερο από τις υπόλοιπες τεχνικές στη φυσική γενετική διαδικασία και χρησιμοποιούνται κυρίως στη βελτιστοποίηση προβλημάτων (Νικολός, 2005).

Καθώς οι Εξελικτικοί Αλγόριθμοι στηρίζονται στη μίμηση των διαδικασιών της εξέλιξης και της κληρονομικότητας, πολλές διαδικασίες που μοιάζουν με τις βιολογικές έχουν δανειστεί την αντίστοιχη ορολογία της βιολογίας.

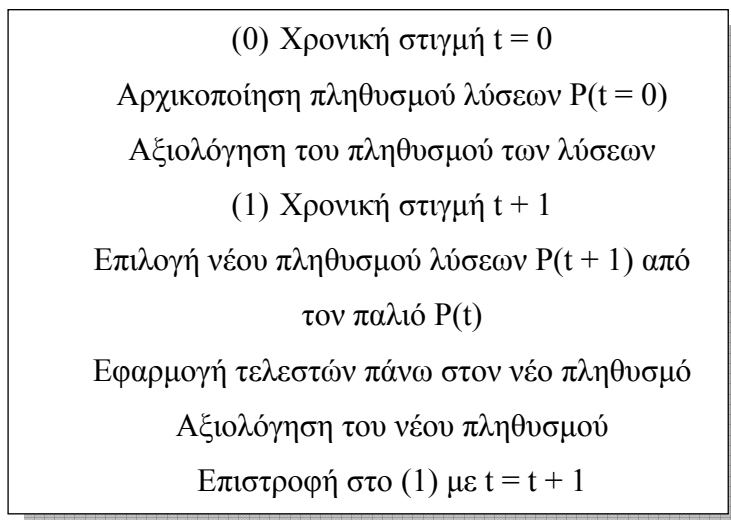
Στη φύση γίνεται κωδικοποίηση των εξωτερικών χαρακτηριστικών του οργανισμού, μέσω της οποίας λαμβάνει χώρα η διεργασία της κληρονομικότητας και επιτυγχάνεται η ποικιλομορφία. Η φυσική επιλογή έχει άμεση επίδραση στον φαινότυπο, δηλαδή στα εξωτερικά χαρακτηριστικά του ατόμου, τα οποία και αλληλεπιδρούν με το περιβάλλον και έμμεση επίδραση στον γονότυπο, δηλαδή στην κωδικοποίηση αυτών των εξωτερικών χαρακτηριστικών μέσω της επιβίωσης των ατόμων που παρουσιάζουν μεγαλύτερη προσαρμοστικότητα. Η εφαρμογή των παραπάνω σε ένα πρόβλημα βελτιστοποίησης απαιτεί τη διασύνδεση ανάμεσα στον γονότυπο και στον φαινότυπο, μια διαδικασία αξιολόγησης, μια διαδικασία επιλογής και κάποιες διαδικασίες που να έχουν ως αποτέλεσμα την ποικιλομορφία του πληθυσμού.

Προϋπόθεση για τη βελτιστοποίηση ενός προβλήματος είναι η μοντελοποίηση του συστήματος ούτως ώστε αυτό να περιγράφεται από έναν πεπερασμένο αριθμό παραμέτρων, τις *ανεξάρτητες μεταβλητές σχεδίασης* (independent design variables). Βάσει του μοντέλου που έχει επιλεγεί και της απόδοσης κάποιων συγκεκριμένων τιμών των ανεξάρτητων μεταβλητών, μπορεί κάθε φορά να δημιουργείται ένα διαφορετικό σύστημα, δηλαδή μια διαφορετική λύση του προβλήματος βελτιστοποίησης.

Η κωδικοποιημένη μορφή του συνόλου των ανεξάρτητων μεταβλητών σχεδίασης αποτελεί το *χρωμόσωμα* και όταν σε αυτό οι ανεξάρτητες μεταβλητές πάρουν συγκεκριμένες τιμές δημιουργείται μια *υποψήφια λύση*. Η κωδικοποίηση ενός συγκεκριμένου συστήματος στη μορφή του χρωμοσώματος αποτελεί τον *γονότυπο*, ενώ όταν αυτός εφαρμοστεί στο μοντέλο του συστήματος, το σύστημα που προκύπτει αποτελεί τον *φαινότυπο*. Ο φαινότυπος αξιολογείται με κάποια συνήθως ασαφή και υποκειμενικά κριτήρια, τα οποία πρέπει να μετατραπούν σε μετρήσιμα μεγέθη ούτως ώστε η «ποιότητα» του συστήματος να δύναται να απεικονιστεί με έναν αριθμό. Ο αριθμός αυτός, που αντιστοιχεί στη βαθμολόγηση της ποιότητας, ονομάζεται *συνάρτηση προσαρμογής* (fitness function) και αποτελεί την αντικειμενική συνάρτηση του προβλήματος η οποία και βελτιστοποιείται. Η βελτιστοποίηση της συνάρτησης προσαρμογής αντιστοιχεί στην προσαρμογή του ατόμου στο περιβάλλον κατά τη διαδικασία της φυσικής επιλογής (Νικολός, 2005).

## 9.2 Λειτουργία Εξελικτικών Αλγορίθμων

Σε έναν εξελικτικό αλγόριθμο ακολουθούνται τα εξής βήματα:



Αναλυτικότερα, η διαδικασία πραγματοποιείται ως εξής: η αρχικοποίηση του πληθυσμού των λύσεων γίνεται παράγοντας τυχαία έναν αριθμό χρωμοσωμάτων (υποψήφιων λύσεων) μέσα από το χώρο των λύσεων, με χρήση γεννήτριας τυχαίων αριθμών. Η αξιολόγηση των λύσεων γίνεται υπολογίζοντας τη συνάρτηση προσαρμογής κάθε χρωμοσώματος, η οποία στη διαδικασία της φυσικής επιλογής παίζει το ρόλο του περιβάλλοντος. Βάσει της κατανομής της πιθανότητας βασισμένης στις τιμές αυτής της συνάρτησης, επιλέγεται ο νέος πληθυσμός λύσεων.

Πιο αναλυτικά, υπολογίζεται η συνολική τιμή της συνάρτησης προσαρμογής του πληθυσμού καθώς και η πιθανότητα να επιλεχθεί το συγκεκριμένο χρωμόσωμα, η οποία ισούται με το κλάσμα της τιμής της συνάρτησης προσαρμογής του χρωμοσώματος προς το άθροισμα των τιμών των συναρτήσεων προσαρμογής όλου του πληθυσμού. Έπειτα, υπολογίζεται η αθροιστική πιθανότητα για κάθε χρωμόσωμα, πράγμα που σημαίνει ότι τα χρωμοσώματα κατανέμονται το ένα μετά το άλλο στο διάστημα μεταξύ 0 και 1. Για να γίνει πιο κατανοητή αυτή η διαδικασία, μπορεί να προσομοιαστεί με την τοποθέτηση όλων των ατόμων μιας γενιάς (δηλαδή των πιθανών λύσεων) πάνω σε έναν τροχό ρουλέτας με μοναδιαίο μήκος περιφέρειας τροχού. Κάθε άτομο καταλαμβάνει τόσο χώρο όσο η σχετική του συνάρτηση προσαρμογής (η πιθανότητα επιλογής του). Η διαδικασία επιλογής γίνεται

παράγοντας τυχαίους αριθμούς ομοιόμορφα κατανεμημένους μεταξύ 0 και 1 τόσες φορές όσες και ο πληθυσμός κάθε γενιάς.

Ανάλογα με τον αριθμό που παράγεται τυχαία κάθε φορά, επιλέγεται και το αντίστοιχο χρωμόσωμα από τον τροχό βάσει του εξής κανόνα: Έστω ότι με  $v_i$  συμβολίζονται τα χρωμοσώματα και με  $r$  ο τυχαίος αριθμός που παράγεται. Αν  $r \leq q_1$ , όπου  $q_1$  η αθροιστική πιθανότητα του χρωμοσώματος  $v_1$ , τότε επιλέγεται το πρώτο χρωμόσωμα  $v_1$ . Αλλιώς, επιλέγεται το χρωμόσωμα  $v_i$  έτσι ώστε  $q_{i-1} \leq r \leq q_i$ .

Από τα παραπάνω γίνεται αντιληπτό ότι ορισμένα χρωμοσώματα θα επιλεγθούν περισσότερες φορές από κάποια άλλα, ενώ κάποια χρωμοσώματα μπορεί να μην επιλεγθούν καθόλου. Αυτό που είναι βέβαιο είναι ότι τα καλύτερα χρωμοσώματα παράγουν περισσότερα αντίγραφα, τα μέτρια παραμένουν όπως έχουν και τα χειρότερα απομακρύνονται.

Στη συνέχεια γίνεται εφαρμογή των τελεστών πάνω στον ενδιαμέσο πληθυσμό που έχει δημιουργηθεί. Οι τελεστές αυτοί είναι η διασταύρωση ή επιχiasμός (crossover) και η μετάλλαξη (mutation). Συνοπτικά αναφέρεται ότι επιχiasμός καλείται η δημιουργία νέων οργανισμών, στα χρωμοσώματα των οποίων τα μισά γονίδια προέρχονται από τη μητέρα και τα άλλα μισά από τον πατέρα, ενώ μετάλλαξη ονομάζεται η μεταβολή της δομής των χρωμοσωμάτων, συνήθως λόγω λανθασμένης αντιγραφής μορίων ή λόγω εξωγενών παραγόντων, με αποτέλεσμα την αλλαγή κάποιου χαρακτηριστικού.

Για να καθοριστούν οι προαναφερθέντες τελεστές πρέπει να οριστούν οι αντίστοιχες παράμετροί τους, οι οποίες αντιστοιχούν στην πιθανότητα να συμβεί ένα από τα δύο φαινόμενα. Η πιθανότητα να υποστεί κάποιο χρωμόσωμα διασταύρωση συμβολίζεται με  $p_c$  ενώ η πιθανότητα να υποστεί μετάλλαξη με  $p_m$ . Συνήθως ισχύει  $p_c \geq p_m$ . Η διασταύρωση ωθεί τις λύσεις προς το βέλτιστο ενώ η μετάλλαξη παρέχει τη δυνατότητα εξερεύνησης του χώρου των λύσεων (Νικολός, 2005).

### 9.3 Μέθοδοι Βελτιστοποίησης και Εξελικτικοί Αλγόριθμοι

Κατά τη βελτιστοποίηση προβλημάτων δύο είναι τα ζητούμενα: η διαδικασία βελτίωσης και η κατεύθυνση στην οποία πρέπει να αναζητηθεί η βέλτιστη λύση.

Οι υπάρχουσες μέθοδοι αναζήτησης είναι αυτές που στηρίζονται στη *μαθηματική ανάλυση* (calculus – based), οι *απαριθμητικές* (enumerative) και αυτές που στηρίζονται σε *τυχαία αναζήτηση* (random search). Ωστόσο οι συμβατικές μέθοδοι

βελτιστοποίησης, στις οποίες υπάγονται οι δύο πρώτες κατηγορίες, παρόλο που αποδίδουν αρκετά καλά σε πολλά προβλήματα, δε μπορούν να ανταπεξέλθουν σε περισσότερο περίπλοκες εφαρμογές. Σε τέτοιες περιπτώσεις, οι αλγόριθμοι τυχαίας αναζήτησης ανταποκρίνονται εκπληκτικά καλά, καθώς με αυτούς αναζητείται και αποθηκεύεται η βέλτιστη λύση που συναντάται σε τυχαίες διαδρομές μέσα στο χώρο των λύσεων. Στην τελευταία κατηγορία ανήκουν οι εξελικτικοί αλγόριθμοι, οι οποίοι διαφέρουν από τις συμβατικές μεθόδους αναζήτησης στα εξής σημεία:

- Η αναζήτηση βέλτιστης λύσης γίνεται μέσα από έναν πληθυσμό σημείων και όχι από ένα μόνο σημείο.
- Οι πληροφορίες που χρειάζονται για τη βελτιστοποίηση της αντικειμενικής συνάρτησης λαμβάνονται απευθείας από αυτήν και όχι από παραγώγους ή άλλα βοηθήματα.
- Οι κανόνες μετάβασης που χρησιμοποιούνται είναι μη αιτιοκρατικοί και βασίζονται σε πιθανότητες.
- Στην περίπτωση των γενετικών αλγορίθμων, αυτοί δουλεύουν με κωδικοποίηση κάποιων παραμέτρων και όχι με τις ίδιες τις παραμέτρους.

Οι σημαντικότεροι λόγοι που οι εξελικτικοί αλγόριθμοι αποδίδουν τόσο καλά είναι οι ακόλουθοι:

- Σε σχέση με τους υπόλοιπους αλγόριθμους, παρουσιάζουν πρωτοφανή ισορροπία ανάμεσα στην ικανότητα εξερεύνησης του πεδίου λύσεων και στην πίεση για εύρεση βέλτιστης λύσης.
- Είναι παράλληλοι αλγόριθμοι, δηλαδή μπορούν να δουλεύουν με παράλληλο τρόπο εξοικονομώντας χρόνο και κόπο.
- Είναι ιδιαίτερα προσαρμοστικοί στα διάφορα προβλήματα που καλούνται να επιλύσουν.
- Οι τελεστές διαφοροποίησης των χρωμοσωμάτων μπορούν εύκολα να τροποποιηθούν και να προσαρμοστούν στα διάφορα προβλήματα.

Οι εξελικτικοί αλγόριθμοι σήμερα τυγχάνουν ευρείας εφαρμογής σε διάφορα πεδία όπως τα μαθηματικά, η ιατρική, η μηχανική και η πολιτική επιστήμη (Νικολός, 2005).

## 9.4 Ο Διαφορικός Εξελικτικός Αλγόριθμος

### 9.4.1 Εισαγωγή

Οι αλγόριθμοι αναζήτησης που βασίζονται στον πληθυσμό (population – based search algorithms), όπως οι εξελικτικοί αλγόριθμοι, παρά τα πολλά πλεονεκτήματα που παρουσιάζουν, υστερούν στο ότι χρειάζονται αρκετό χρόνο για την διεκπεραίωση των πολυάριθμων αξιολογήσεων των υποψήφιων λύσεων. Προκειμένου να μειωθεί ο υπολογιστικός φόρτος μπορούν να χρησιμοποιηθούν διάφορες μέθοδοι όπως η παράλληλη επεξεργασία, η χρήση ειδικών τελεστών και η χρήση μη ακριβούς αξιολόγησης του πληθυσμού.

Ιδιαίτερα ικανοποιητικά ως μοντέλα προσέγγισης αποδεικνύονται τα τεχνητά νευρωνικά δίκτυα (artificial neural networks), τα οποία μπορούν να συνδυαστούν αποτελεσματικά με έναν Διαφορικό Εξελικτικό αλγόριθμο (differential evolution algorithm), επιταχύνοντας έτσι τη διαδικασία σύγκλισης. Ο Διαφορικός Εξελικτικός Αλγόριθμος αποτελεί μια πολλά υποσχόμενη καινοτομία στους εξελικτικούς αλγόριθμους, καθώς μπορεί να χειριστεί συνεχείς, διακριτές και ακέραιες μεταβλητές απόφασης και πολλούς περιορισμούς με ιδιαίτερη ευρωστία και αποτελεσματικότητα (Nikolos, 2004).

### 9.4.2 Λειτουργία Διαφορικού Εξελικτικού Αλγόριθμου

Έστω ότι το πρόβλημα βελτιστοποίησης είναι της μορφής

$$\min_x = f(X) \quad (9.1)$$

όπου  $X$  είναι το διάνυσμα – στήλη των μεταβλητών απόφασης με  $n$  παραμέτρους (γονίδια), που αποτελεί το χρωμόσωμα και  $f$  η συνάρτηση κόστους που πρέπει να ελαχιστοποιηθεί. Σε κάθε γενιά  $G$ , ο πληθυσμός των  $N_p$  υποψήφιων λύσεων υποβάλλεται σε συγκεκριμένους τελεστές, οπότε η υποψήφια λύση μπορεί να γραφτεί ως

$$X_{i,G}, \quad i = 1, \dots, N_p, \quad G = 1, \dots, mgens$$

όπου  $mgens$  είναι ο μέγιστος αριθμός γενεών. Ο πληθυσμός των υποψήφιων λύσεων  $N_p$  παραμένει σταθερός κατά τη διαδικασία βελτιστοποίησης.



Ο Διαφορικός Εξελικτικός αλγόριθμος ξεκινά παράγοντας τυχαία και με ομοιόμορφη πιθανότητα τον αρχικό πληθυσμό χρωμοσωμάτων (διανυσμάτων μεταβλητών απόφασης), με τα γονίδια (μεταβλητές απόφασης) να παίρνουν τιμές μέσα στο διάστημα των ρητών περιορισμών. Θεωρούμε ότι όλες οι τυχαίες μεταβλητές απόφασης παρουσιάζουν ομοιόμορφη πιθανότητα κατανομής.

Ο Διαφορικός Εξελικτικός αλγόριθμος στηρίζεται σε μια μέθοδο μετάλλαξης, σύμφωνα με την οποία παράγονται δοκιμαστικά διανύσματα μεταβλητών απόφασης, προσθέτοντας ένα διάνυσμα σταθμισμένων διαφορών μεταξύ δύο ατόμων του πληθυσμού σε ένα τρίτο άτομο, που λέγεται δότης (donor). Η διαδικασία της μετάλλαξης εφαρμόζεται σε όλα τα άτομα του πληθυσμού. Για κάθε άτομο, η διαδικασία μετάλλαξης ξεκινά επιλέγοντας τυχαία τρία άτομα του τρέχοντος πληθυσμού που αποτελούν μια τριάδα. Σε αυτή την τριάδα, επιλέγεται τυχαία ένα άτομο να είναι ο δότης ενώ τα άλλα δύο άτομα προκαλούν διαταραχή στο δότη. Έτσι, το  $i$  – οστό διαταραγμένο άτομο που γεννιέται δίνεται από τη σχέση

$$V_{i,G+1} = X_{r3,G} + F(X_{r1,G} - X_{r2,G}) \quad (9.2)$$

όπου τα

$$r1, r2, r3 \in \{1, \dots, N_p\} \quad r1 \neq r2 \neq r3 \neq i$$

επιλέγονται τυχαία από τις υποψήφιες λύσεις του τρέχοντος πληθυσμού. Η παράμετρος στάθμισης  $F$  αποτελεί την πρώτη παράμετρο ελέγχου, που ορίζεται από τον χρήστη και παίρνει τιμές στο διάστημα  $[0, 1+]$ . Αυτή παραμένει σταθερή καθ' όλη τη διαδικασία και ελέγχει την ενίσχυση της διαταραχής που προστίθεται στον δότη.

Στη συνέχεια, το διαταραγμένο άτομο  $V_{i,G+1}$  και το άτομο  $X_{i,G}$  του αρχικού πληθυσμού υπόκεινται σε επιχiasμό (διασταύρωση), με αποτέλεσμα τη γέννηση ενός ενδιάμεσου πληθυσμού δοκιμαστικών διανυσμάτων  $U_{i,G+1}$ .

Αν για το άτομο  $i$  του πληθυσμού στη γενιά  $G$  συμβολίσουμε το αρχικό, το διαταραγμένο και το δοκιμαστικό διάνυσμα αντίστοιχα με

$$X_{i,G} = (x_{1,i,G}, \dots, x_{n,i,G})^T, \quad V_{i,G+1} = (v_{1,i,G+1}, \dots, v_{n,i,G+1})^T, \quad U_{i,G+1} = (u_{1,i,G+1}, \dots, u_{n,i,G+1})^T$$

τότε κάθε γονίδιο  $j$  του δοκιμαστικού διανύσματος προκύπτει με επιχiasμό ως:

$$u_{j,i,G+1} = \begin{cases} v_{j,i,G+1} & \text{αν } \text{rand}_j \leq C_r \quad \forall \quad j = k \\ x_{j,i,G} & \text{αλλιώς} \end{cases} \quad (9.3)$$

όπου  $j = 1, \dots, n$  αφορά τα γονίδια του χρωμοσώματος και  $k \in \{1, \dots, n\}$  είναι ένας τυχαίος δείκτης ο οποίος επιλέγεται μία φορά για όλα τα  $N_p$  μέλη του πληθυσμού. Με τη χρήση του εξασφαλίζεται ότι τουλάχιστον ένα γονίδιο του διαταραγμένου διανύσματος θα περάσει στο δοκιμαστικό διάνυσμα. Η παράμετρος επιχιασμού  $C_r$ , που παίρνει τιμές στο διάστημα  $[0,1]$ , αποτελεί τη δεύτερη παράμετρο ελέγχου που ορίζεται από τον χρήστη.

Μετά το τέλος του επιχιασμού, τα δοκιμαστικά διανύσματα αξιολογούνται και υπολογίζεται η συνάρτηση κόστους τους. Τα άτομα που θα απαρτίσουν την επόμενη γενιά επιλέγονται μεταξύ του τρέχοντος πληθυσμού και των αντίστοιχων δοκιμαστικών διανυσμάτων, σύμφωνα με τον κανόνα που ακολουθεί (για πρόβλημα ελαχιστοποίησης):

$$X_{i,G+1} = \begin{cases} U_{i,G+1} & \text{αν } f(U_{i,G+1}) \leq f(X_{i,G}) \\ X_{i,G} & \text{αλλιώς} \end{cases} \quad (9.4)$$

Δηλαδή, τα άτομα του τρέχοντος και του ενδιάμεσου πληθυσμού αξιολογούνται βάσει της τιμής της συνάρτησης προσαρμογής τους και επιλέγονται τα καλύτερα. Έτσι, ο πληθυσμός κάθε γενιάς περιέχει τουλάχιστον ίσα ή καλύτερα μέλη σε σχέση με την προηγούμενη γενιά (Nikolos, 2004).

## ΚΕΦΑΛΑΙΟ 10 – ΤΟ ΠΡΟΒΛΗΜΑ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ

### 10.1 Σκοπός του Προβλήματος Βελτιστοποίησης

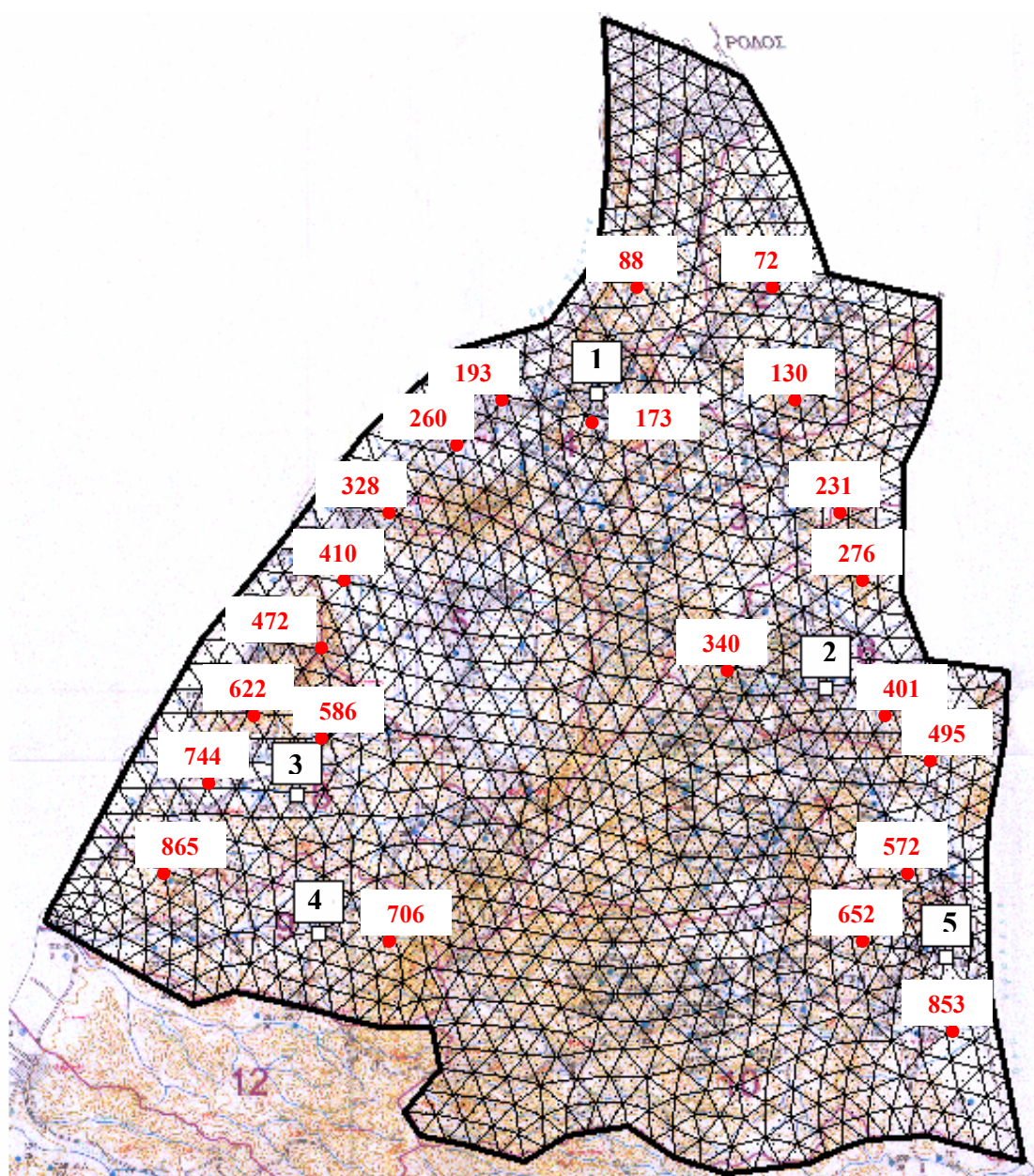
Σκοπός του προβλήματος βελτιστοποίησης είναι η μεγιστοποίηση των παροχών άντλησης των γεωτρήσεων, χωρίς όμως το υδραυλικό ύψος σε συγκεκριμένα σημεία της παραλιακής ζώνης να πέσει κάτω από μια ελάχιστη τιμή. Το πρακτικό νόημα αυτής της διατύπωσης είναι το εξής: να γίνεται άντληση όσο το δυνατόν περισσότερου νερού από τις γεωτρήσεις, ούτως ώστε να καλύπτονται οι ανάγκες του πληθυσμού (σε ύδρευση, άρδευση κλπ.), αλλά και να εξασφαλίζεται ότι δε θα εξαπλωθεί το φαινόμενο της υφαλμύρωσης, που ήδη παρατηρείται στην περιοχή. Σύμφωνα με τη διπλωματική εργασία της συγγραφέως, «η υπερ-άντληση νερού, κυρίως από τις παράκτιες γεωτρήσεις, έχει ως αποτέλεσμα την υφαλμύρωση του υδροφορέα, δηλαδή τη διείσδυση θαλασσινού νερού μέσα σε αυτόν. Το φαινόμενο αυτό εμφανίζεται σε όλη την παράκτια ζώνη, ωστόσο σε μεγαλύτερο βαθμό εντοπίζεται στο βόρειο τμήμα της περιοχής μελέτης, όπου η διεπιφάνεια αλμυρού – γλυκού νερού έχει μετατοπιστεί κάποια χιλιόμετρα προς την ξηρά» (Στεργιάδη, 2004).

Γενικά, τα προβλήματα βελτιστοποίησης αποτελούνται από ρητούς και μη ρητούς περιορισμούς. *Ρητοί περιορισμοί (explicit constraints)* είναι αυτοί που διατυπώνουν τα επιτρεπτά όρια τιμών των ανεξάρτητων μεταβλητών σχεδίασης (στην περίπτωση που εξετάζουμε είναι οι παροχές άντλησης) και ορίζονται εύκολα στον Διαφορικό Εξελικτικό αλγόριθμο. Οι *μη ρητοί ή πεπλεγμένοι περιορισμοί (implicit constraints)*, που στο συγκεκριμένο πρόβλημα αφορούν τα ελάχιστα επιτρεπόμενα υδραυλικά ύψη σε κάποιες παράκτιες γεωτρήσεις, δε μπορούν να συνδεθούν εύκολα με τις ανεξάρτητες μεταβλητές σχεδίασης, καθώς δεν υπάρχει κάποια απλή σχέση που να τα συνδέει. Καθώς αυτοί οι περιορισμοί δε μπορούν να τεθούν ως εξωτερικοί περιορισμοί στο Διαφορικό Εξελικτικό αλγόριθμο, δεν ορίζονται στον αλγόριθμο αλλά ενσωματώνονται στη συνάρτηση προσαρμογής ως *συνάρτηση τιμωρίας (penalty function)*, όπως επεξηγείται παρακάτω (Νικολός, 2005).

Οι ρητοί περιορισμοί του προβλήματος αφορούν τις παροχές άντλησης, οι οποίες πρέπει να είναι θετικές και να μην ξεπερνούν τη μέγιστη τιμή των 4000 m<sup>3</sup>/d.

Οι μη ρητοί περιορισμοί του προβλήματος βελτιστοποίησης λαμβάνονται στους παρακάτω 22 κόμβους παρατήρησης (ή πηγάδια παρατήρησης), οι περισσότεροι από

τους οποίους εντοπίζονται στην παράκτια ζώνη: 72, 88, 130, 173, 193, 231, 260, 276, 328, 340, 401, 410, 472, 495, 572, 586, 622, 652, 706, 744, 853 και 865 (Εικόνα 10.1). Ο υπολογισμός του ελάχιστου επιτρεπόμενου υδραυλικού ύψους σε αυτούς τους κόμβους στην παρούσα εργασία πραγματοποιήθηκε ως εξής: με το μοντέλο PTC υπολογίζονται τα υδραυλικά ύψη έχοντας όλες τις γεωτρήσεις κλειστές ( $Q = 0 \text{ m}^3/\text{d}$ ) και όλες τις γεωτρήσεις να αντλούν με τη μέγιστη παροχή ( $Q = 4000 \text{ m}^3/\text{d}$ ). Ο μέσος όρος των υδραυλικών υψών των δύο αυτών περιπτώσεων τέθηκε ίσος με το ελάχιστο επιτρεπόμενο υδραυλικό ύψος στα πηγάδια παρατήρησης (Πίνακας 10.1).



**Εικόνα 10.1:** Κόμβοι παρατήρησης, στους οποίους λαμβάνονται οι μη ρητοί περιορισμοί του προβλήματος βελτιστοποίησης

**Πίνακας 10.1:** Εύρεση ελάχιστου επιτρεπόμενου υδραυλικού ύψους στους κόμβους παρατήρησης

Κόμβοι	h (m) (όλα Q = 0 m <sup>3</sup> /d)	h (m) (όλα Q = 4000 m <sup>3</sup> /d)	h <sub>μέσο</sub> (m)
72	148,75	148,09	148,42
88	147,32	146,03	146,68
130	155,34	154,45	154,90
173	157,68	149,63	153,66
193	149,61	147,56	148,59
231	151,94	151,64	151,79
260	152,26	151,48	151,87
276	153,45	153,25	153,35
328	154,96	154,59	154,78
340	180,98	180,61	180,79
401	157,83	157,58	157,70
410	158,25	157,68	157,97
472	165,39	163,90	164,65
495	155,44	155,05	155,25
572	163,51	161,99	162,75
586	180,69	176,32	178,50
622	161,71	158,89	160,30
652	180,57	178,44	179,51
706	208,84	204,94	206,89
744	162,56	159,22	160,89
853	155,70	153,93	154,81
865	160,60	158,72	159,66

## 10.2 Κατασκευή Συνάρτησης Προσαρμογής

Η συνάρτηση προσαρμογής (fitness function) υλοποιήθηκε με χρήση κώδικα σε Visual Basic (Παράρτημα ΙΙΙ) και αποτελεί τη συνάρτηση κόστους η οποία πρέπει να ελαχιστοποιηθεί. Η συνάρτηση αυτή αποτελείται ουσιαστικά από δύο επιμέρους συναρτήσεις: μια συνάρτηση τιμωρίας για την ελαχιστοποίηση της διαφοράς των τιμών των υδραυλικών υψών που δίνει το νευρωνικό δίκτυο από τις ελάχιστες επιτρεπόμενες τιμές, όσον αφορά στα υδραυλικά ύψη στους κόμβους παρατήρησης ( $f_1$ ) και μία συνάρτηση για τη μεγιστοποίηση του αθροίσματος των παροχών άντλησης ( $f_2$ ).

Η συνάρτηση  $f_1$  δίνεται από τη σχέση

$$f_1 = \frac{\sum_{i=1}^{22} |(true\ heights)_i - (lower\ heights)_i|}{22}$$

όπου:

$i = 1, \dots, 22$  αφορά στα πηγάδια παρατήρησης

*true heights* είναι οι τιμές που δίνει το νευρωνικό δίκτυο για τα υδραυλικά ύψη στους κόμβους παρατήρησης

*lower heights* είναι τα ελάχιστα επιτρεπόμενα υδραυλικά ύψη στους κόμβους παρατήρησης, όπως δίνονται στον πίνακα 10.1 παραπάνω.

Ο αριθμητής της παραπάνω σχέσης διαιρείται με τον αριθμό 22 για να αδιαστατοποιηθεί (να έρθει κοντά στη μονάδα).

Η συνάρτηση  $f_1$  αποτελεί συνάρτηση τιμωρίας, καθώς μετρά το βαθμό μη ικανοποίησης του συγκεκριμένου περιορισμού και προσθέτει κάποιο ποσό στη συνάρτηση κόστους. Έτσι, όσο πιο κοντά βρίσκονται τα υδραυλικά ύψη στα ελάχιστα επιτρεπόμενα, τόσο λιγότερο «τιμωρείται» η συνάρτηση προσαρμογής (προστίθεται σε αυτήν μικρότερο ποσό). Όταν τα υδραυλικά ύψη συμπίσουν με τα ελάχιστα επιτρεπόμενα υδραυλικά ύψη, η συνάρτηση προσαρμογής παίρνει την ελάχιστη, άρα και βέλτιστη τιμή της (αφού η συνάρτηση τιμωρίας μηδενίζεται). Η απόλυτη τιμή στη συνάρτηση  $f_1$  χρησιμοποιήθηκε για να ωθήσει τις παροχές σε τιμές κοντά στα όρια των περιορισμών. Έτσι, λύσεις που απέχουν ακόμη και θετικά από το όριο των περιορισμών είναι λιγότερο επιθυμητές, αφού αυτές κατά κανόνα αντιστοιχούν σε μικρότερες παροχές άντλησης.

Η συνάρτηση  $f_2$  δίνεται από τη σχέση

$$f_2 = \frac{1}{\left(\sum_{i=1}^5 Q_i\right) + 1000}$$

όπου

$i = 1, \dots, 5$  αφορά στις γεωτρήσεις άντλησης

$Q_i$  είναι οι παροχές άντλησης

Εφόσον το άθροισμα των παροχών πρέπει να μεγιστοποιηθεί ενώ έχουμε πρόβλημα ελαχιστοποίησης, λαμβάνεται το αντίστροφο αυτού. Η τιμή 1000 προστίθεται στον παρονομαστή για να μην πάρει αυτός ποτέ μηδενική τιμή.

Τελικά, η συνολική συνάρτηση προσαρμογής, που ελαχιστοποιείται, ορίζεται ως:

$$f = 5f_1 + f_2 \Rightarrow$$

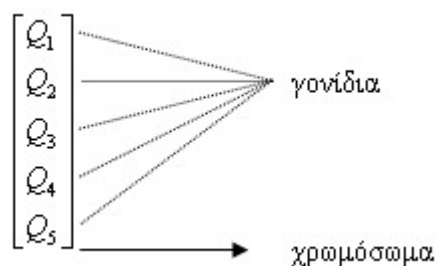
$$\Rightarrow f = 5 \cdot \frac{\sum_{i=1}^{22} |(true\ heights)_i - (lower\ heights)_i|}{22} + \frac{1}{\left(\sum_{i=1}^5 Q_i\right) + 1000}$$

Η συνάρτηση  $f_1$  πολλαπλασιάζεται με τον συντελεστή 5 για να δοθεί περισσότερη βαρύτητα στη βελτιστοποίηση της διαφοράς των υδραυλικών υψών από τα ελάχιστα επιτρεπόμενα υδραυλικά ύψη παρά στη βελτιστοποίηση του αθροίσματος των παροχών. Με τον τρόπο αυτό επιτυγχάνεται καλύτερη ικανοποίηση των περιορισμών του προβλήματος.

### 10.3 Εφαρμογή του Διαφορικού Εξελικτικού Αλγόριθμου στο Υπό Μελέτη Πρόβλημα

Η βελτιστοποίηση του προβλήματος γίνεται με χρήση Διαφορικού Εξελικτικού αλγόριθμου. Τα άτομα κάθε γενιάς υπόκεινται σε μετάλλαξη και επιχiasμό με αποτέλεσμα τη δημιουργία μιας ενδιάμεσης γενιάς. Στη συνέχεια τα άτομα της τρέχουσας και της ενδιάμεσης γενιάς αξιολογούνται βάσει της τιμής της συνάρτησης προσαρμογής τους και τα καλύτερα σχηματίζουν την επόμενη γενιά.

Τα διανύσματα των ανεξάρτητων μεταβλητών απόφασης αποτελούν τα χρωμοσώματα (υποψήφιες λύσεις) ενώ τα επιμέρους στοιχεία κάθε διανύσματος αποτελούν τα γονίδια, κάθε ένα από τα οποία μπορεί να λάβει τιμή από 0 έως 4000 (Εικόνα 10.2). Ο γονότυπος ταυτίζεται με το χρωμόσωμα ενώ ο φαινότυπος είναι το αποτέλεσμα του γονότυπου στο σύστημα, δηλαδή τα υδραυλικά ύψη στους κόμβους του πλέγματος.



**Εικόνα 10.2:** Χρωμόσωμα και γονίδια του προβλήματος βελτιστοποίησης

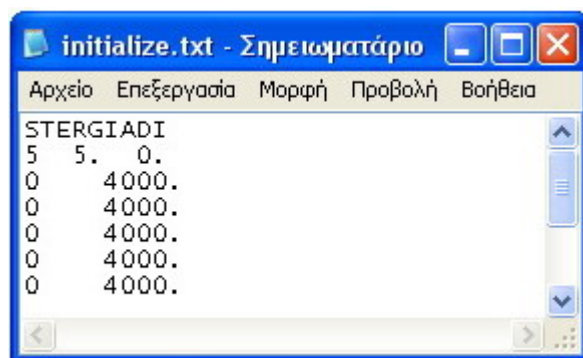
Από το σύνολο των χρωμοσωμάτων, επιλέγονται με γεννήτρια τυχαίων αριθμών 20 χρωμοσώματα που απαρτίζουν τον αρχικό πληθυσμό. Ο πληθυσμός παραμένει σταθερός καθ' όλη τη διαδικασία βελτιστοποίησης και η τιμή του πρέπει να είναι πολλαπλάσια του αριθμού των μεταβλητών απόφασης και μεγαλύτερη ή ίση από τον διπλάσιο αριθμό αυτών. Δηλαδή:

Πληθυσμός κάθε γενιάς  $\geq 2 \cdot (\text{αριθμός μεταβλητών απόφασης}) \Rightarrow$

$\Rightarrow$  Πληθυσμός κάθε γενιάς  $\geq 2 \cdot 5 \Rightarrow$  Πληθυσμός κάθε γενιάς  $\geq 10$

Η εισαγωγή των παραμέτρων του προβλήματος αλλά και η εξαγωγή των αποτελεσμάτων γίνεται με χρήση επιμέρους αρχείων ASCII, που παρουσιάζονται παρακάτω:

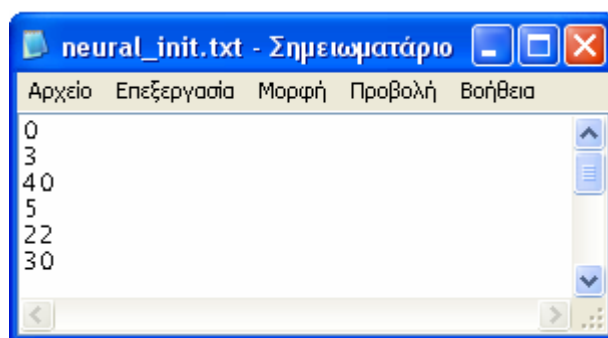
• *initialize.txt*: στο αρχείο αυτό ορίζεται ο αριθμός των ανεξάρτητων μεταβλητών απόφασης (5), η μέγιστη και η ελάχιστη αναμενόμενη τιμή της συνάρτησης προσαρμογής, που προσεγγιστικά τίθενται ίσες με 5 και 0 αντίστοιχα, καθώς και τα ανώτατα και κατώτατα όρια ανάμεσα στα οποία μπορούν να κυμαίνονται οι μεταβλητές απόφασης (από 0 έως 4000 m<sup>3</sup>/d) (Εικόνα 10.3). Δηλαδή, αυτό αποτελεί το αρχείο στο οποίο ορίζονται οι ρητοί περιορισμοί του προβλήματος.



**Εικόνα 10.3:** Το αρχείο initialize.txt



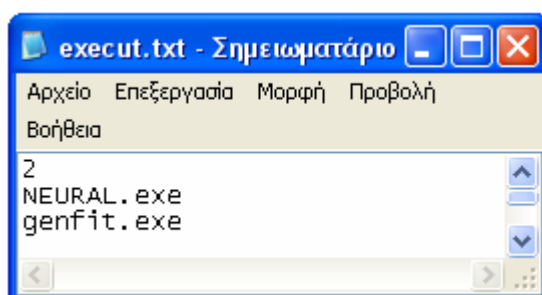
- *chromosome.txt*: το αρχείο αυτό ουσιαστικά αποτελεί το εκάστοτε άτομο του πληθυσμού κάθε γενιάς που αξιολογείται και σε αυτό αναγράφονται ο αριθμός των ανεξάρτητων μεταβλητών απόφασης και οι τιμές που αυτές παίρνουν. Στο τέλος της διαδικασίας βελτιστοποίησης, το αρχείο αυτό αφορά το τελευταίο άτομο του πληθυσμού της τελευταίας γενιάς που αξιολογήθηκε.
- *neural\_init.txt*: το αρχείο αυτό αφορά στη δομή του νευρωνικού δικτύου, το οποίο καλεί ο Διαφορικός Εξελικτικός αλγόριθμος για τον υπολογισμό των υδραυλικών υψών. Σε αυτό αναγράφονται αν το νευρωνικό δίκτυο χρησιμοποιείται για εκπαίδευση (1) ή για αξιολόγηση (0), όπως συμβαίνει στην περίπτωση μας, ο αριθμός των επιπέδων του δικτύου (3, δηλαδή το επίπεδο εισόδου, το επίπεδο εξόδου και ένα κρυφό επίπεδο), ο αριθμός των κόμβων στο κρυφό επίπεδο (40), ο αριθμός των κόμβων στο επίπεδο εισόδου (5), ο αριθμός των κόμβων στο επίπεδο εξόδου (22) και ο αριθμός των epochs (30) (Εικόνα 10.4).



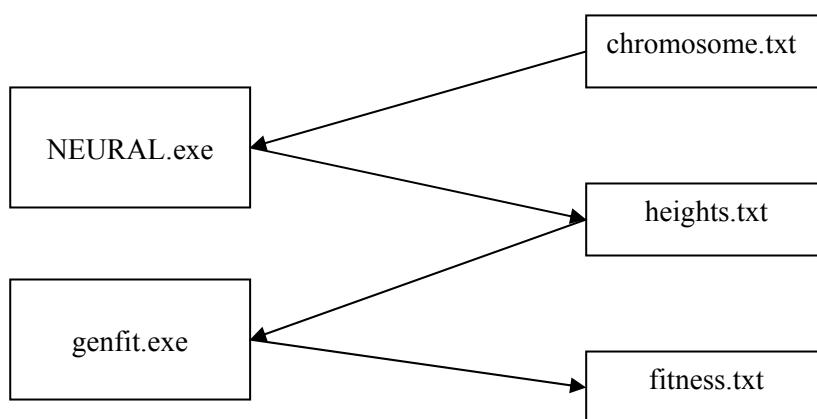
Εικόνα 10.4: Το αρχείο neural\_init.txt

- *neuralerror.txt*: στο αρχείο αυτό δίνονται το μέσο σφάλμα εκπαίδευσης και το μέσο σφάλμα δοκιμής του νευρωνικού δικτύου.
- *lowerheights.txt*: στο αρχείο αυτό αναγράφονται οι κόμβοι παρατήρησης και τα ελάχιστα επιτρεπόμενα υδραυλικά ύψη σε αυτούς.
- *heights.txt*: το αρχείο αυτό δημιουργείται από το νευρωνικό δίκτυο, το οποίο για κάθε γενιά καλεί τα 20 άτομα του πληθυσμού (αρχεία chromosome.txt) και βάσει των αναγραφόμενων παροχών υπολογίζει τα υδραυλικά ύψη στους κόμβους παρατήρησης.
- *execut.txt*: στο αρχείο αυτό καταγράφονται ο αριθμός και τα ονόματα των εκτελέσιμων αρχείων τα οποία χρησιμοποιεί ο Διαφορικός Εξελικτικός αλγόριθμος, με τη σειρά που αυτά εκτελούνται. Τα αρχεία αυτά είναι το

NEURAL.exe, δηλαδή το νευρωνικό δίκτυο, το οποίο κάθε φορά διαβάσει το αρχείο chromosome.txt (που αντιστοιχεί στο άτομο κάθε γενιάς) και δημιουργεί το αντίστοιχο αρχείο heights.txt και το genfit.exe, δηλαδή το πρόγραμμα υπολογισμού της συνάρτησης προσαρμογής, το οποίο καλεί το αρχείο heights.txt και υπολογίζει την τιμή της συνάρτησης προσαρμογής του συγκεκριμένου ατόμου. Η βέλτιστη λύση της συνάρτησης προσαρμογής στο τέλος της διαδικασίας βελτιστοποίησης αποθηκεύεται στο αρχείο fitness.txt (Εικόνες 10.5 και 10.6).

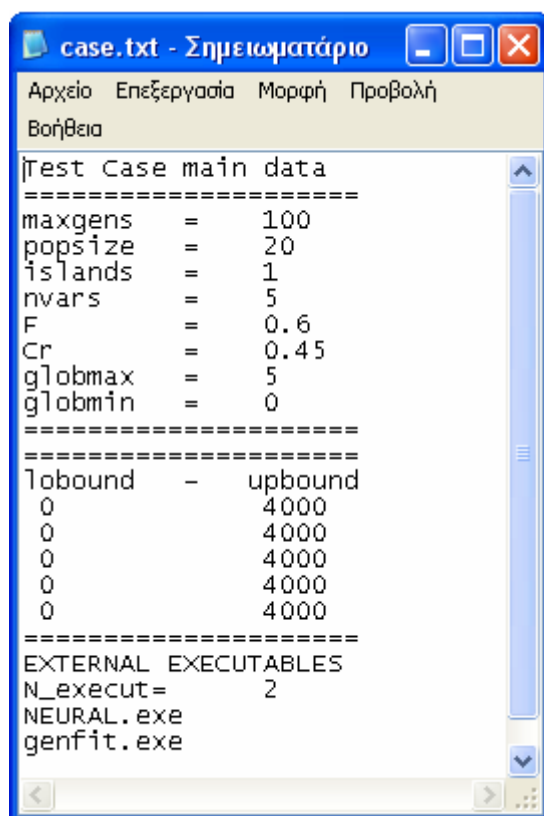


Εικόνα 10.5: Το αρχείο execut.txt



Εικόνα 10.6: Λειτουργία των εκτελέσιμων αρχείων NEURAL.exe και genfit.exe

- *case.txt*: στο αρχείο αυτό αναφέρονται λεπτομέρειες του προβλήματος, όπως ο αριθμός των γενεών (100), ο πληθυσμός ανά γενιά (20), ο αριθμός των μεταβλητών απόφασης (5), η τιμή της παραμέτρου στάθμισης  $F$  (0,6), η τιμή της παραμέτρου επιχιασμού  $C_r$  (0,45), η μέγιστη και ελάχιστη τιμή της συνάρτησης προσαρμογής κ.ά. (Εικόνα 10.7).



Εικόνα 10.7: Το αρχείο case.txt

- *dokgen.txt*: το αρχείο αυτό δίνει πληροφορίες σχετικά με τη διαδικασία σύγκλισης του αλγόριθμου στη βέλτιστη λύση. Για κάθε γενιά παρατίθενται η καλύτερη και η χειρότερη λύση, ο αριθμός των αξιολογήσεων και οι συνολικές αξιολογήσεις που έχουν γίνει μέχρι εκείνη τη γενιά.
- *curentbest.txt*: το αρχείο αυτό δίνει τη βέλτιστη λύση του προβλήματος βελτιστοποίησης.

Το γραφικό περιβάλλον του Διαφορικού Εξελικτικού αλγόριθμου παρουσιάζεται στην εικόνα 10.8. Οι παράμετροι που ορίζονται στον αλγόριθμο είναι:

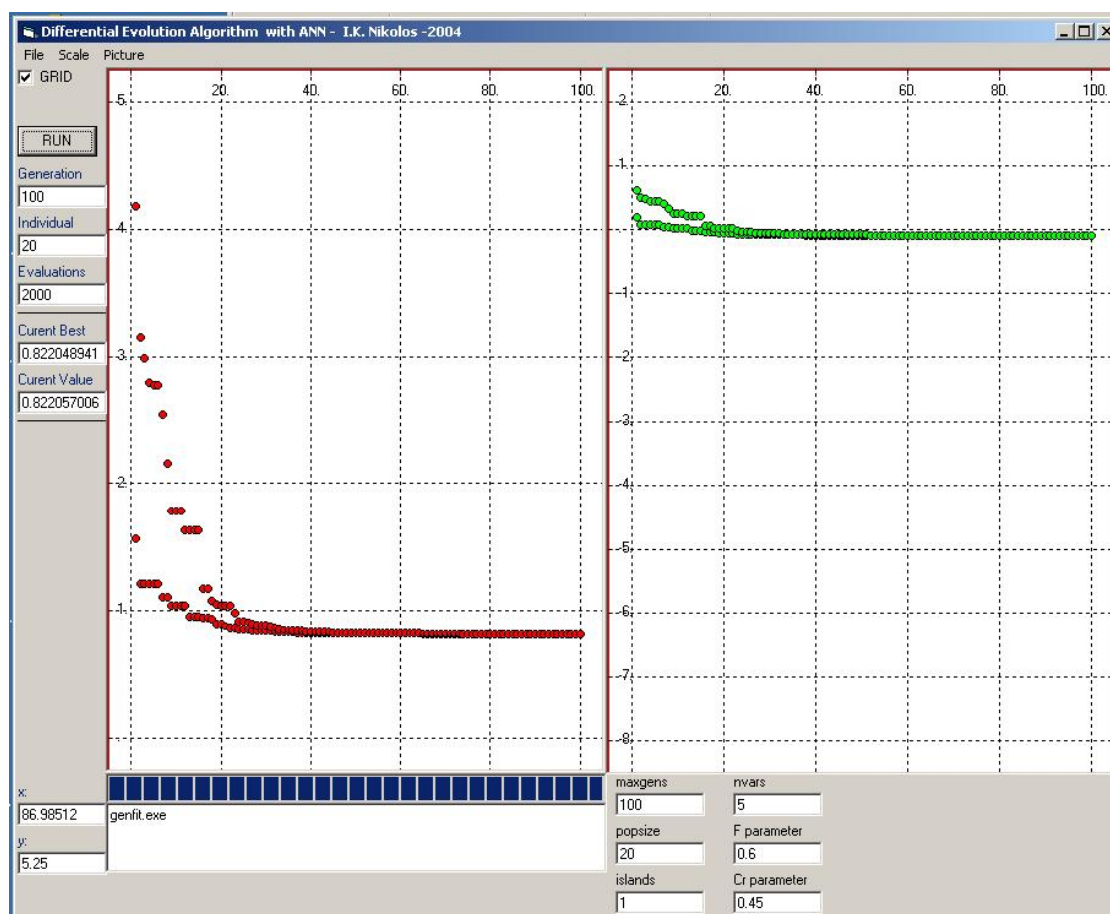
- η γραφική σχεδίαση σε κάναβο (GRID)
- ο αριθμός των γενεών που αξιολογούνται ( $\text{maxgens} = 100$ )
- ο πληθυσμός ανά γενιά ( $\text{popsize} = 20$ )
- ο αριθμός των νησίδων ( $\text{islands} = 1$ ). Σε κάποια προβλήματα βελτιστοποίησης ο χώρος των υποψήφιων λύσεων δεν είναι ενιαίος καθώς κάποιες λύσεις μπορεί να βρίσκονται διεσπαρμένες σε διαφορετικά σημεία, τα οποία καλούνται

νησίδες. Στη συγκεκριμένη περίπτωση, θέτοντας ως αριθμό νησίδων τη μονάδα, υποδηλώνεται ότι ο χώρος των λύσεων είναι ενιαίος.

- ο αριθμός των ανεξάρτητων μεταβλητών απόφασης ( $nvars = 5$ )
- η παράμετρος στάθμισης ( $F = 0,6$ )
- η παράμετρος επιχιασμού ( $C_r = 0,45$ )

Παρακολουθώντας την πρόοδο του αλγόριθμου, στα πεδία που βρίσκονται αριστερά αναγράφονται η γενιά (generation) και το άτομο της γενιάς που αξιολογείται εκείνη τη στιγμή (individual), ο αριθμός των αξιολογήσεων που έχουν γίνει μέχρι εκείνη τη στιγμή (evaluations), η βέλτιστη τιμή της συνάρτησης προσαρμογής για την προηγούμενη γενιά (curent best) και η τιμή της συνάρτησης προσαρμογής για το προηγούμενο άτομο της γενιάς που αξιολογείται (curent value). Δεδομένου ότι έχουν οριστεί 100 γενιές με πληθυσμό 20 ατόμων, ο συνολικός αριθμός των αξιολογήσεων θα ισούται με  $100 \times 20 = 2000$ . Μετά τον τερματισμό της διαδικασίας, στο πεδίο curent value αναγράφεται η τιμή της συνάρτησης προσαρμογής του τελευταίου ατόμου της τελευταίας γενιάς που αξιολογήθηκε, ενώ στο πεδίο curent best αναγράφεται η τιμή της συνάρτησης προσαρμογής για τη βέλτιστη λύση.

Στο αριστερό παράθυρο παρουσιάζονται γραφικά η καλύτερη και η χειρότερη τιμή της συνάρτησης προσαρμογής για κάθε γενιά, ενώ το γράφημα στο δεξί παράθυρο είναι το ίδιο, μόνο που η τιμή της συνάρτησης προσαρμογής είναι σε λογαριθμική κλίμακα.



**Εικόνα 10.8:** Γραφικό περιβάλλον Διαφορικού Εξελικτικού αλγόριθμου και σύγκλιση αυτού στη βέλτιστη λύση

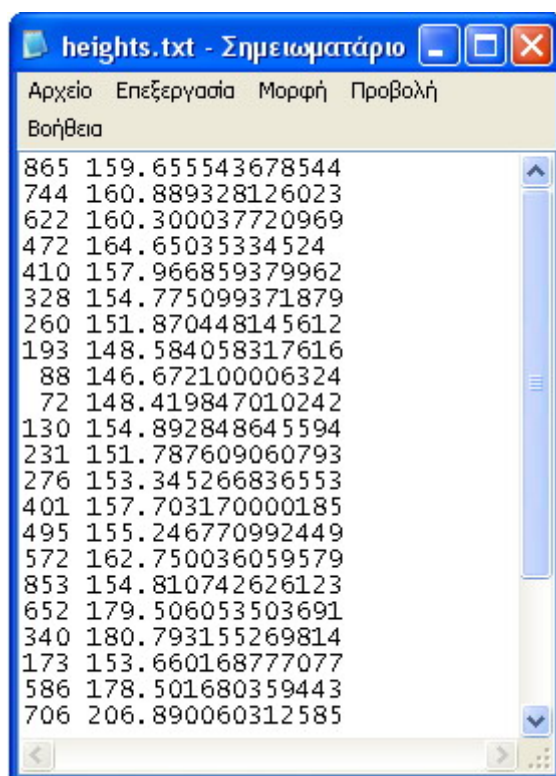
## **ΚΕΦΑΛΑΙΟ 11 – ΑΠΟΤΕΛΕΣΜΑΤΑ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗΣ**

Η βελτιστοποίηση του προβλήματος γίνεται βάσει δύο διαφορετικών σεναρίων, που διαφοροποιούνται μόνο ως προς τα ελάχιστα επιτρεπόμενα υδραυλικά ύψη σε κάποιους κόμβους παρατήρησης.

### **11.1 Πρώτο Σενάριο**

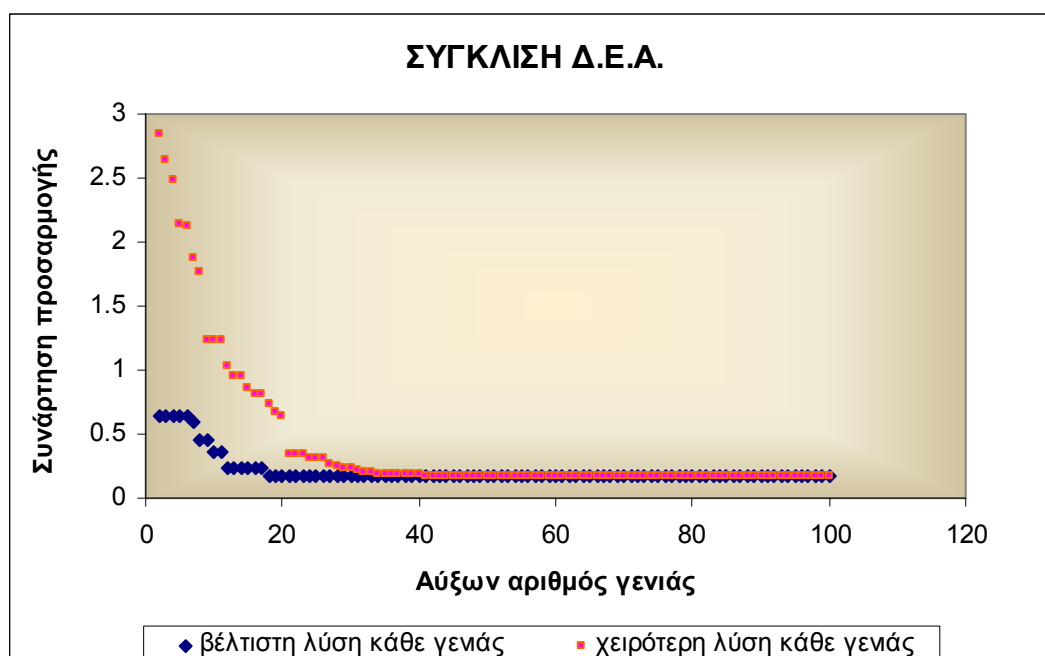
Στην περίπτωση αυτή τα ελάχιστα επιτρεπόμενα υδραυλικά ύψη στους κόμβους παρατήρησης είναι αυτά που δίνονται στον πίνακα 10.1 του κεφαλαίου 10 τα οποία, όπως ειπώθηκε στο εν λόγω κεφάλαιο, προκύπτουν ως οι μέσοι όροι των υδραυλικών υψών θεωρώντας όλες τις γεωτρήσεις κλειστές και όλες τις γεωτρήσεις να αντλούν με τη μέγιστη παροχή.

Το νευρωνικό δίκτυο βάσει των διάφορων τιμών των παροχών άντλησης (chromosome.txt) υπολογίζει τα υδραυλικά ύψη στα πηγάδια παρατήρησης (Εικόνα 11.1), τα οποία χρησιμοποιεί ο Διαφορικός Εξελικτικός αλγόριθμος για να υπολογίσει την τιμή της συνάρτησης προσαρμογής κάθε ατόμου του τρέχοντος και του ενδιάμεσου πληθυσμού. Τα αποτελέσματα της βελτιστοποίησης παρουσιάζονται στο γράφημα 11.1, όπου φαίνονται η καλύτερη και η χειρότερη λύση κάθε γενιάς και η σύγκλιση του αλγόριθμου στη βέλτιστη λύση.



Node	Hydraulic Head (m)
865	159.655543678544
744	160.889328126023
622	160.300037720969
472	164.65035334524
410	157.966859379962
328	154.775099371879
260	151.870448145612
193	148.584058317616
88	146.672100006324
72	148.419847010242
130	154.892848645594
231	151.787609060793
276	153.345266836553
401	157.703170000185
495	155.246770992449
572	162.750036059579
853	154.810742626123
652	179.506053503691
340	180.793155269814
173	153.660168777077
586	178.501680359443
706	206.890060312585

**Εικόνα 11.1:** Υδραυλικά ύψη (m) στους κόμβους παρατήρησης όπως δίνονται από το νευρωνικό δίκτυο βάσει του πρώτου σεναρίου



**Γράφημα 11.1:** Χειρότερη και καλύτερη λύση κάθε γενιάς και σύγκλιση Διαφορικού Εξελικτικού αλγορίθμου στη βέλτιστη λύση βάσει του πρώτου σεναρίου

Η βέλτιστη λύση του προβλήματος, που ελαχιστοποιεί τη συνάρτηση προσαρμογής, δίνεται στο αρχείο curentbest.txt και είναι

$$Q_1 = 2049,9 \text{ m}^3/\text{d}$$

$$Q_2 = 3997,1 \text{ m}^3/\text{d}$$

$$Q_3 = 2042,6 \text{ m}^3/\text{d}$$

$$Q_4 = 2059,3 \text{ m}^3/\text{d}$$

$$Q_5 = 1972,6 \text{ m}^3/\text{d}$$

οπότε το άθροισμα των παροχών άντλησης ισούται με  $Q_{ολ} = \sum_{i=1}^5 Q_i = 12121,5 \text{ m}^3/\text{d}$ .

Βάσει των τιμών αυτών, η ελάχιστη (βέλτιστη) τιμή της συνάρτησης προσαρμογής, όπως δίνεται στο αρχείο fitness.txt, ισούται με 0,1657.

Εισάγοντας τις βέλτιστες παροχές άντλησης στο μοντέλο PTC, τα αποτελέσματα της προσομοίωσης, όσον αφορά τα υδραυλικά ύψη στους 22 κόμβους παρατήρησης είναι αυτά που παρουσιάζονται στον πίνακα 11.1, όπου επίσης δίνονται οι αντίστοιχες τιμές που έχει υπολογίσει το νευρωνικό δίκτυο (ANN) και τα ελάχιστα επιτρεπόμενα όρια. Συγκρίνοντας τα αποτελέσματα του νευρωνικού δικτύου με αυτά του μοντέλου, παρατηρείται ότι το νευρωνικό δίκτυο έχει προσεγγίσει πολύ καλά το μοντέλο PTC, καθώς διαφορές υφίστανται ως επί το πλείστον στο δεύτερο δεκαδικό ψηφίο. Έτσι, βλέπουμε και πρακτικά ότι το νευρωνικό δίκτυο έχει εκπαιδευτεί πολύ καλά και μπορεί πλέον να αντικαταστήσει το PTC στην πρόβλεψη των τιμών των υδραυλικών υψών της περιοχής μελέτης. Αξιοσημείωτη είναι και η διαπίστωση ότι τα υπολογιζόμενα, τόσο από το νευρωνικό δίκτυο όσο και από το μοντέλο, υδραυλικά ύψη «συμμορφώνονται» πλήρως στον περιορισμό που τους απαγορεύει να πέσουν κάτω από κάποια όρια. Διαφορές, βέβαια, υφίστανται και πάλι στο δεύτερο δεκαδικό ψηφίο, αλλά αυτές ασφαλώς είναι αμελητέες για την τάξη μεγέθους στην οποία αναφερόμαστε.

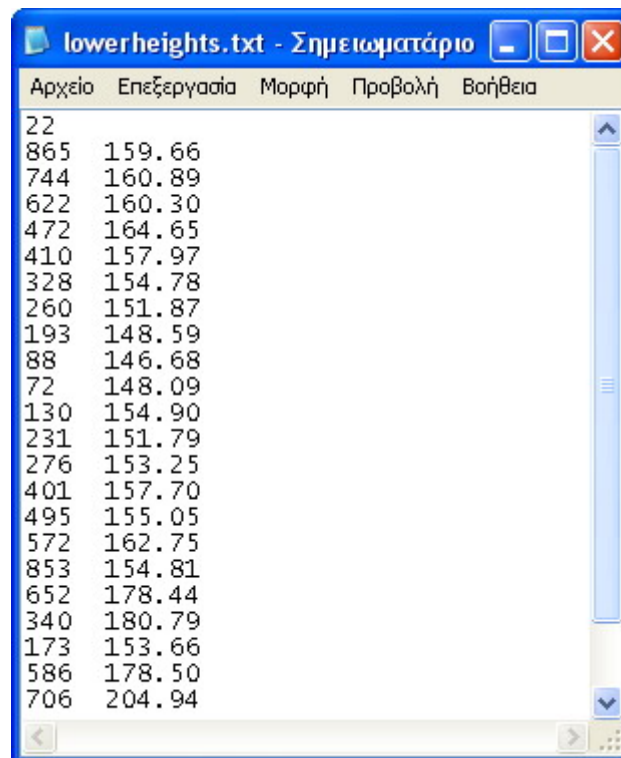


**Πίνακας 11.1:** Υδραυλικά ύψη στους κόμβους παρατήρησης όπως δίνονται από το μοντέλο PTC και το νευρωνικό δίκτυο για το πρώτο σενάριο

Αριθμός κόμβου	Υδραυλικά ύψη (m)		
	Αποτελέσματα PTC	Αποτελέσματα ANN	Κατώτερα επιτρεπόμενα όρια
865	159,63	159,66	159,66
744	160,86	160,89	160,89
622	160,27	160,30	160,30
472	164,62	164,65	164,65
410	157,95	157,97	157,97
328	154,76	154,78	154,78
260	151,85	151,87	151,87
193	148,55	148,58	148,59
88	146,65	146,67	146,68
72	148,41	148,42	148,43
130	154,88	154,89	154,90
231	151,77	151,79	151,79
276	153,33	153,35	153,35
401	157,69	157,70	157,70
495	155,23	155,25	155,25
572	162,75	162,75	162,75
853	154,81	154,81	154,81
652	179,51	179,51	179,51
340	180,78	180,79	180,79
173	153,60	153,66	153,66
586	178,46	178,50	178,50
706	206,84	206,89	206,89

## 11.2 Δεύτερο Σενάριο

Στην περίπτωση αυτή τα ελάχιστα επιτρεπόμενα υδραυλικά ύψη στους κόμβους παρατήρησης 72, 276, 495, 652 και 706 τίθενται ίσα με τα υδραυλικά ύψη που προκύπτουν όταν οι γεωτρήσεις αντλούν με τη μέγιστη παροχή ( $4000 \text{ m}^3/\text{d}$ ). Έτσι, το αρχείο lowerheights.txt είναι αυτό που φαίνεται στην εικόνα 11.2.

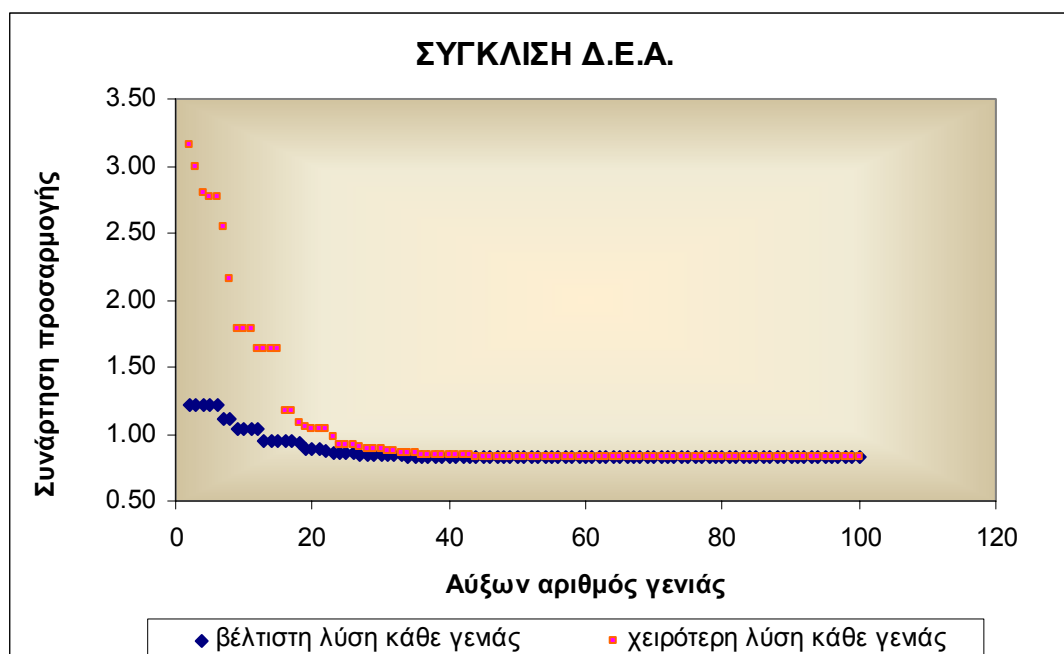


Αρχείο	Επεξεργασία	Μορφή	Προβολή	Βοήθεια
22				
865	159.66			
744	160.89			
622	160.30			
472	164.65			
410	157.97			
328	154.78			
260	151.87			
193	148.59			
88	146.68			
72	148.09			
130	154.90			
231	151.79			
276	153.25			
401	157.70			
495	155.05			
572	162.75			
853	154.81			
652	178.44			
340	180.79			
173	153.66			
586	178.50			
706	204.94			

**Εικόνα 11.2:** Ελάχιστα επιτρεπόμενα υδραυλικά ύψη (m) στους κόμβους παρατήρησης βάσει του δεύτερου σεναρίου

Τα υδραυλικά ύψη στα πηγάδια παρατήρησης δίνονται στην εικόνα 11.3, ενώ τα αποτελέσματα της βελτιστοποίησης παρουσιάζονται στο γράφημα 11.2, όπου φαίνονται η καλύτερη και η χειρότερη λύση κάθε γενιάς και η σύγκλιση του αλγόριθμου στη βέλτιστη λύση.

**Εικόνα 11.3:** Υδραυλικά ύψη (m) στους κόμβους παρατήρησης όπως δίνονται από το νευρωνικό δίκτυο βάσει του δεύτερου σεναρίου



**Γράφημα 11.2:** Χειρότερη και καλύτερη λύση κάθε γενιάς και σύγκλιση Διαφορικού Εξελικτικού αλγορίθμου στη βέλτιστη λύση βάσει του δεύτερου σεναρίου

Η βέλτιστη λύση του προβλήματος, που ελαχιστοποιεί τη συνάρτηση προσαρμογής, δίνεται στο αρχείο curentbest.txt και είναι

$$Q_1 = 2029,5 \text{ m}^3/\text{d}$$

$$Q_2 = 3998,3 \text{ m}^3/\text{d}$$

$$Q_3 = 1601,6 \text{ m}^3/\text{d}$$

$$Q_4 = 3999,5 \text{ m}^3/\text{d}$$

$$Q_5 = 1883,8 \text{ m}^3/\text{d}$$

οπότε το άθροισμα των παροχών άντλησης ισούται με  $Q_{ολ} = \sum_{i=1}^5 Q_i = 13512,7 \text{ m}^3/\text{d}$ .

Βάσει των τιμών αυτών, η ελάχιστη (βέλτιστη) τιμή της συνάρτησης προσαρμογής, όπως δίνεται στο αρχείο fitness.txt, ισούται με 0,8220.

Τρέχοντας το μοντέλο PTC για τις βέλτιστες παροχές άντλησης προκύπτουν τα αποτελέσματα του πίνακα 11.2. Και σε αυτή την περίπτωση παρατηρείται ότι τα αποτελέσματα του νευρωνικού δικτύου προσεγγίζουν αρκετά καλά τα αποτελέσματα του PTC, ενώ και στις δύο περιπτώσεις (του νευρωνικού δικτύου και του μοντέλου) τηρούνται τα κατώτερα επιτρεπόμενα όρια.

**Πίνακας 11.2:** Υδραυλικά ύψη στους κόμβους παρατήρησης όπως δίνονται από το μοντέλο

PTC και το νευρωνικό δίκτυο για το δεύτερο σενάριο

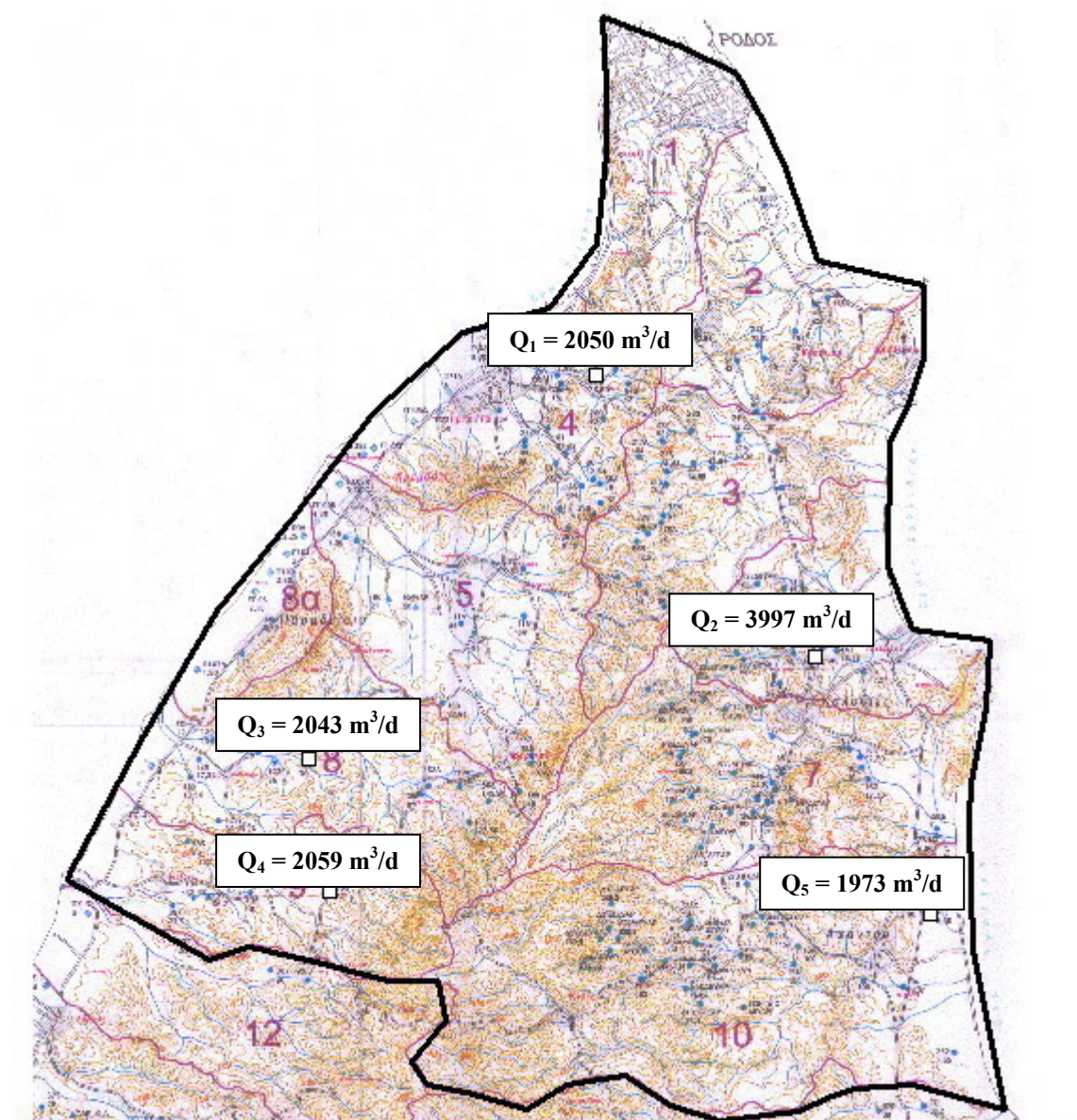
Αριθμός κόμβου	Υδραυλικά ύψη (m)		
	Αποτελέσματα PTC	Αποτελέσματα ANN	Κατώτερα επιτρεπόμενα όρια
865	159,36	159,34	159,66
744	160,79	160,78	160,89
622	160,33	160,36	160,30
472	164,65	164,66	164,65
410	157,97	157,98	157,97
328	154,78	154,79	154,78
260	151,87	151,88	151,87
193	148,56	148,58	148,59
88	146,62	146,62	146,68
72	148,35	148,35	148,09
130	154,80	154,80	154,90
231	151,75	151,75	151,79
276	153,32	153,33	153,25
401	157,68	157,69	157,70
495	155,22	155,22	155,05
572	162,72	162,72	162,75
853	154,81	154,81	154,81
652	179,44	179,42	178,44
340	180,74	180,75	180,79
173	153,61	153,66	153,66
586	178,49	178,50	178,50
706	205,58	205,62	204,94

## ΚΕΦΑΛΑΙΟ 12 – ΑΝΑΛΥΣΗ ΑΠΟΤΕΛΕΣΜΑΤΩΝ ΚΑΙ ΣΥΜΠΕΡΑΣΜΑΤΑ

Οι τιμές των βέλτιστων παροχών άντλησης των πέντε κυριότερων γεωτρήσεων της περιοχής μελέτης βάσει των δύο σεναρίων δίνονται στον πίνακα 12.1, ενώ στις εικόνες 12.1 και 12.2 που ακολουθούν υπενθυμίζεται στον αναγνώστη η θέση των γεωτρήσεων, στις οποίες επίσης αναγράφονται οι βέλτιστες τιμές των παροχών άντλησης για κάθε μία από τις δύο περιπτώσεις. Τα υδραυλικά ύψη της περιοχής μελέτης βάσει των βέλτιστων στρατηγικών άντλησης των δύο σεναρίων, όπως δημιουργούνται από το μοντέλο PTC, παρουσιάζονται στα γραφήματα 12.1 και 12.2.

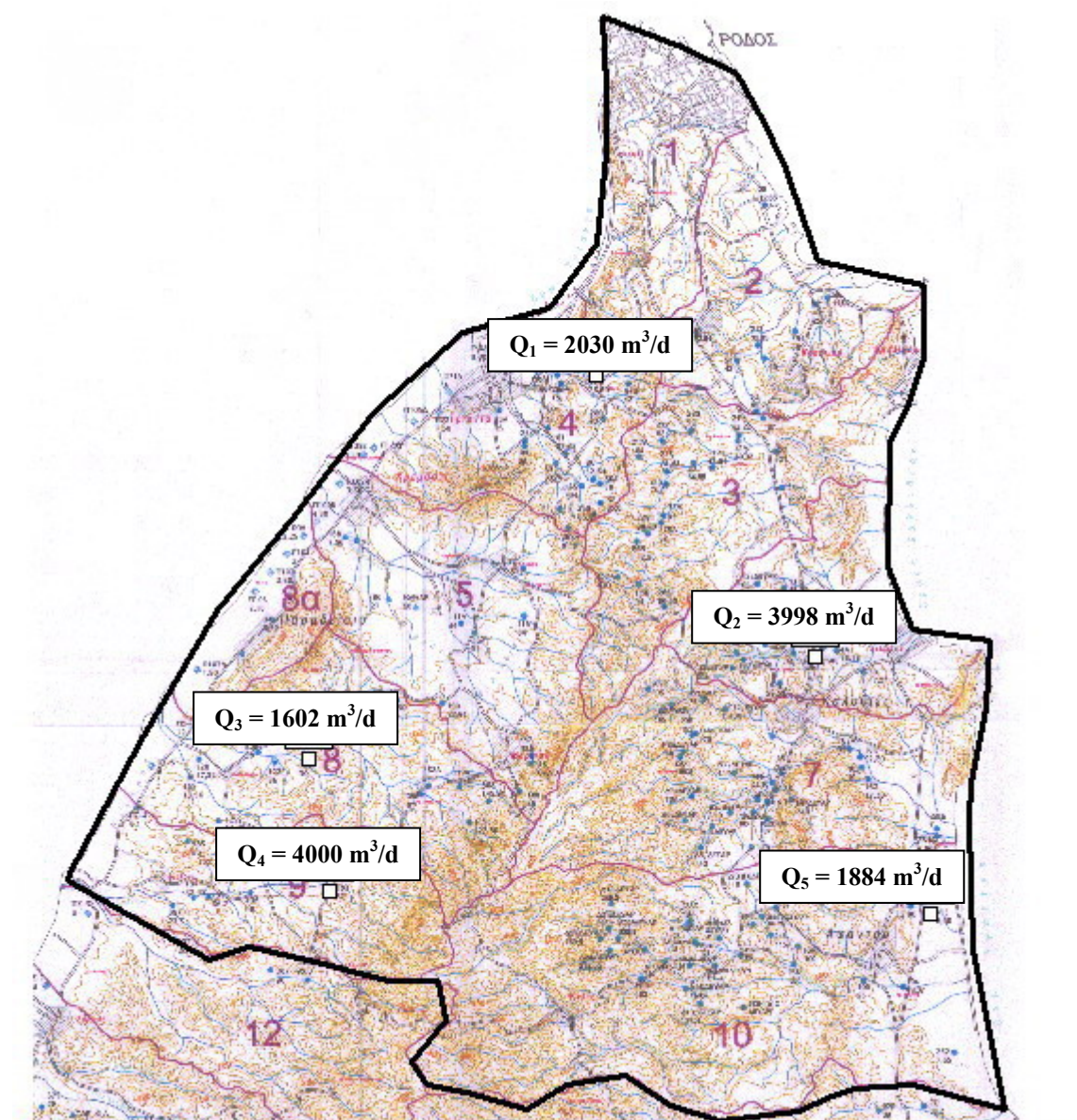
**Πίνακας 12.1:** Βέλτιστες παροχές άντλησης περιοχής μελέτης, βάσει των δύο χρησιμοποιούμενων σεναρίων

Αριθμός γεώτρησης	Βέλτιστες παροχές άντλησης (m <sup>3</sup> /d)	
	1 <sup>ο</sup> σενάριο	2 <sup>ο</sup> σενάριο
1	2050	2030
2	3997	3998
3	2043	1602
4	2059	4000
5	1973	1884
<i>Άθροισμα</i>	<i>12122</i>	<i>13514</i>



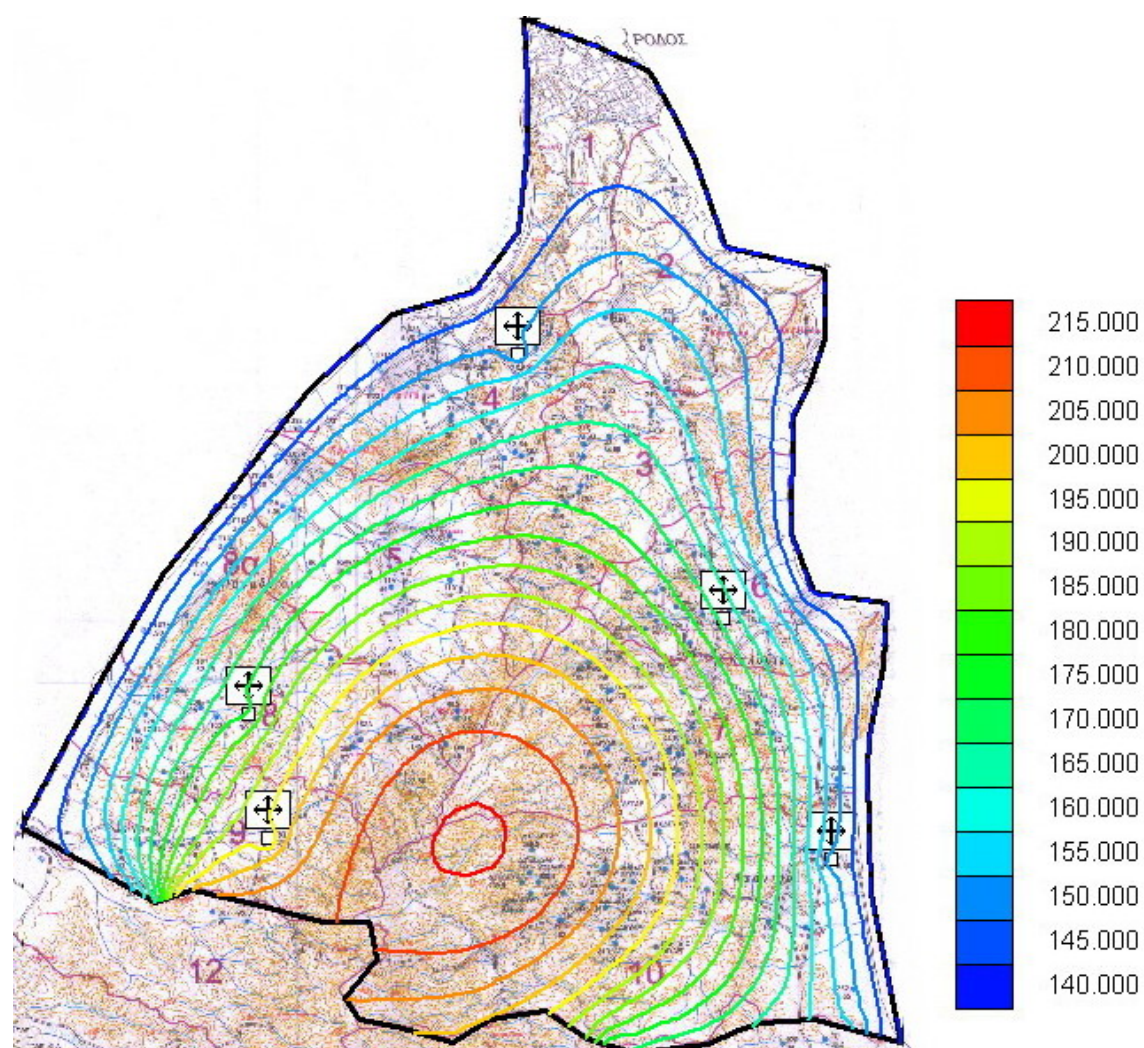
**Εικόνα 12.1:** Βέλτιστες παροχές άντλησης των πέντε γεωτρήσεων βάσει του πρώτου σεναρίου



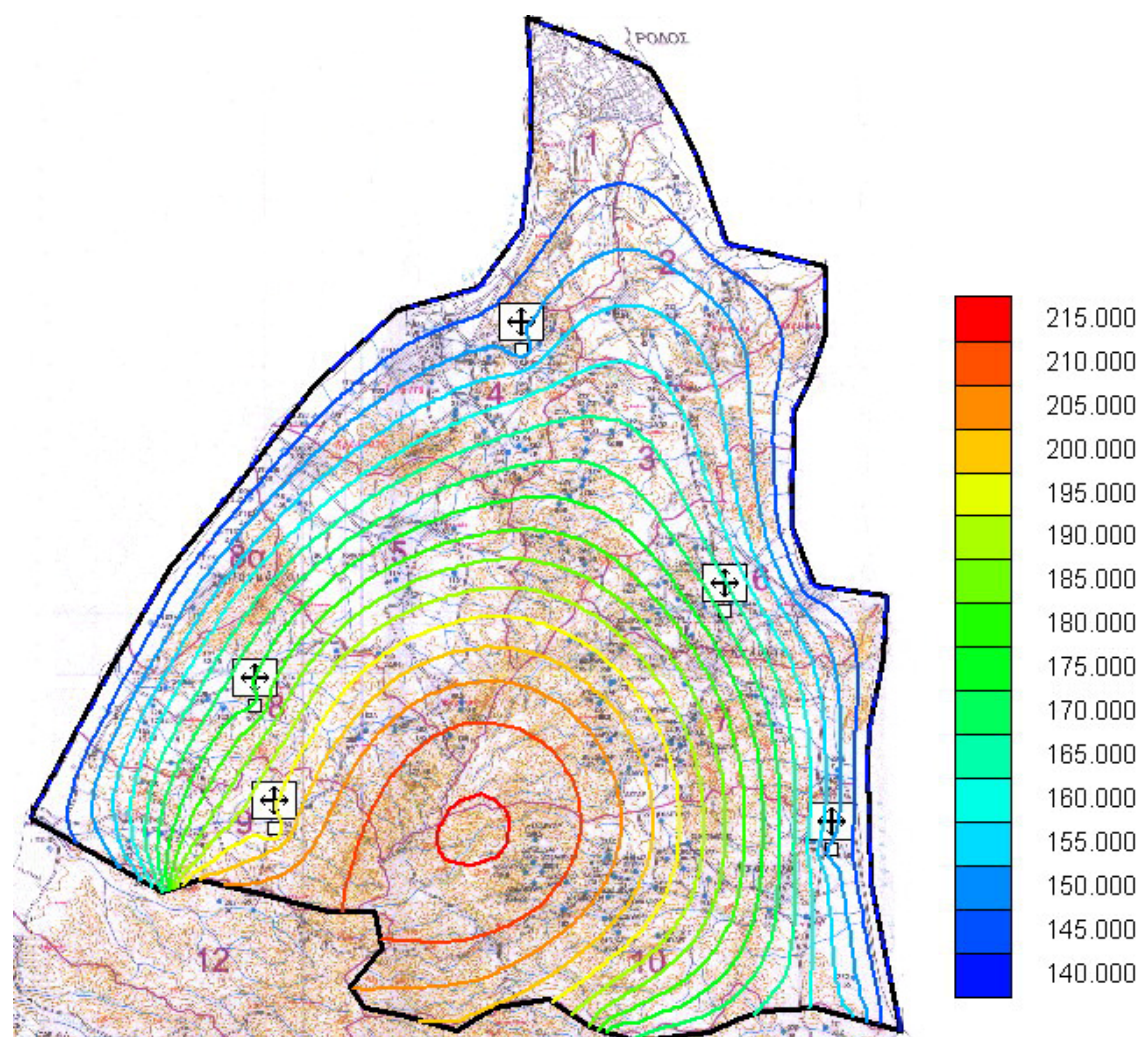


**Εικόνα 12.2:** Βέλτιστες παροχές άντλησης των πέντε γεωτρήσεων βάσει του δεύτερου σεναρίου





**Γράφημα 12.1:** Υδραυλικά ύψη της περιοχής μελέτης όπως δίνονται από το PTC βάσει των βέλτιστων παροχών άντλησης του πρώτου σεναρίου



**Γράφημα 12.2:** Υδραυλικά ύψη της περιοχής μελέτης όπως δίνονται από το PTC βάσει των βέλτιστων παροχών άντλησης του δεύτερου σεναρίου

### Σύγκριση σεναρίων και συμπεράσματα

Συγκρίνοντας τα δύο σενάρια παρατηρείται ότι οι παροχές  $Q_1$  και  $Q_2$  δε μεταβάλλονται ιδιαίτερα, καθώς οι κόμβοι παρατήρησης για τους οποίους διαφοροποιούνται τα ελάχιστα επιτρεπόμενα υδραυλικά ύψη βρίσκονται σχετικά μακριά από τις γεωτρήσεις 1 και 2 (δηλαδή εκτός της ακτίνας επιρροής των εν λόγω γεωτρήσεων).

Αντίθετα, οι παροχές  $Q_3$  και  $Q_4$  παρουσιάζουν μεγάλη διαφοροποίηση από το ένα σενάριο στο άλλο. Από την εικόνα 10.1 (Κεφάλαιο 10), παρατηρείται ότι ο κόμβος 706, στον οποίο κατά το δεύτερο σενάριο το ελάχιστο επιτρεπόμενο υδραυλικό ύψος μειώνεται κατά 2 μέτρα περίπου, βρίσκεται πολύ κοντά στη γεώτρηση 4. Έτσι, μειώνοντας το υδραυλικό ύψος σε αυτόν, «επιτρέπουμε» στη γεώτρηση 4 να αντλεί περισσότερο νερό, πράγμα που συνεπάγεται ότι η διπλανή γεώτρηση 3 θα αναγκάζεται να αντλεί με μικρότερη παροχή.

Στην περίπτωση της γεώτρησης 5, αυτή δε δείχνει να επηρεάζεται σημαντικά από τη μείωση του ελάχιστου επιτρεπόμενου υδραυλικού ύψους στον γειτονικό κόμβο παρατήρησης 652 (η οποία είναι περίπου 1 μέτρο). Μάλιστα, ενώ το αναμενόμενο θα ήταν η μείωση του ελάχιστου επιτρεπόμενου υδραυλικού ύψους στον εν λόγω κόμβο να προκαλέσει αύξηση της παροχής άντλησης  $Q_5$ , αυτή τελικά μειώνεται κατά  $100 \text{ m}^3/\text{d}$  σχεδόν. Το φαινόμενο αυτό οφείλεται στο γεγονός ότι στη γειτονιά της γεώτρησης 5 υπάρχουν δύο επιπλέον κόμβοι παρατήρησης (572, 853) των οποίων τα ελάχιστα επιτρεπόμενα υδραυλικά ύψη είναι τα ίδια και στα δύο σενάρια. Έτσι, ενδεχόμενη αύξηση της παροχής άντλησης  $Q_5$  θα μείωνε τα υδραυλικά ύψη σε αυτούς τους κόμβους. Προκειμένου αυτά να παραμείνουν σταθερά η παροχή άντλησης  $Q_5$  κατά το δεύτερο σενάριο μειώνεται, αντισταθμίζοντας έτσι την επίδραση του κόμβου 652. Δηλαδή ουσιαστικά οι κόμβοι παρατήρησης 572, 652 και 853 και η γεώτρηση 5 αποτελούν ένα αλληλεπιδρών σύστημα που προκειμένου να διατηρήσει τις τιμές των υδραυλικών υψών στους κόμβους 572 και 853 σταθερές, το υδραυλικό ύψος στον κόμβο 652 και η παροχή άντλησης  $Q_5$  λειτουργούν ανταγωνιστικά.

Σημειώνεται ότι η μείωση των ελάχιστων επιτρεπόμενων υδραυλικών υψών σε κάποιους κόμβους παρατήρησης, όπως συμβαίνει στο δεύτερο σενάριο, ωθεί τη βέλτιστη λύση σε μεγάλες παροχές, δηλαδή στα όρια του χώρου αναζήτησης όπου το νευρωνικό δίκτυο παρουσιάζει μικρότερη ακρίβεια.

Από την άλλη, όπως φαίνεται στα γραφήματα 12.1 και 12.2, τα υδραυλικά ύψη της περιοχής μελέτης δε μεταβάλλονται ιδιαίτερα από το ένα σενάριο στο άλλο, πράγμα αναμενόμενο καθώς οι παροχές άντλησης όλων των γεωτρήσεων (εκτός από την 3 και την 4) δεν αλλάζουν σημαντικά. Έτσι, διαφορές της τάξης των 1 – 2 μέτρων αναμένεται να υπάρχουν στις περιοχές γύρω από τις γεωτρήσεις 3 και 4, οι οποίες όμως και πάλι δεν είναι σπουδαίες.

Παρά τις μικρές διαφοροποιήσεις ανάμεσα στα δύο σενάρια, η επιλογή βέλτιστης στρατηγικής διαχείρισης μεταξύ αυτών επαφίεται στον εκάστοτε φορέα διαχείρισης αλλά και στις ανάγκες της τοπικής κοινωνίας. Έτσι, όταν το επιθυμητό είναι η αναχαίτιση του μετώπου της υφαλμύρωσης σε παράκτιες περιοχές, περισσότερο ενδεικτική είναι η χρήση του πρώτου σεναρίου σύμφωνα με το οποίο η ολική παροχή άντλησης ισούται με 12120 m<sup>3</sup>/d περίπου. Αντίθετα, όταν δίνεται προτεραιότητα στην κάλυψη των αναγκών του πληθυσμού σε νερό για ύδρευση, άρδευση κλπ., δύναται να χρησιμοποιηθεί το δεύτερο σενάριο βάσει του οποίου αντλούνται συνολικά 13510 m<sup>3</sup>/d περίπου.

Η ανάλυση και η μεθοδολογία που χρησιμοποιήθηκαν στην παρούσα εργασία αποτελούν ένα ισχυρό εργαλείο για την κατάστρωση πολυάριθμων περαιτέρω στρατηγικών διαχείρισης των υπογείων υδάτων της υπό μελέτη περιοχής στο μέλλον, σταθμίζοντας τις πραγματικές απαιτήσεις της τοπικής κοινωνίας και την αναγκαιότητα για αειφόρο διαχείριση των υπογείων υδάτων και βιωσιμότητα των υδροφορέων. Έτσι, μεταβάλλοντας τα ελάχιστα επιτρεπόμενα υδραυλικά ύψη στα πηγάδια παρατήρησης ή τη θέση και τον αριθμό τους με σκοπό τη μεγιστοποίηση της αντλούμενης παροχής ή την παρεμπόδιση της μετακίνησης της διεπιφάνειας γλυκού – αλμυρού νερού προς την ξηρά, δημιουργούνται διάφορα πιθανά σενάρια άντλησης τα οποία μπορούν να αξιολογηθούν με στόχο την επιλογή και εφαρμογή του προτιμότερου στην εν λόγω περιοχή.

Η χρήση του Νευρωνικού Δικτύου για την μοντελοποίηση της συμπεριφοράς του υδροφόρου ορίζοντα επιτρέπει την επίλυση απεριόριστων προβλημάτων βελτιστοποίησης με διαφορετικούς περιορισμούς (διαφορετικά σενάρια) σε πολύ μικρό χρόνο. Αντίθετα, αν χρειαζόταν η κλήση του PTC για την αξιολόγηση κάθε υποψήφιας λύσης, η αξιολόγηση διαφορετικών σεναρίων θα ήταν απαγορευτική.

Όπως φαίνεται και από τα αποτελέσματα που παρουσιάστηκαν, το Τεχνητό Νευρωνικό Δίκτυο που χρησιμοποιήθηκε αποδείχθηκε πολύ καλό υποκατάστατο του μοντέλου PTC και βεβαίως πολύ ταχύτερο, αφού με τη χρήση του ο Διαφορικός



Εξελικτικός αλγόριθμος χρειάστηκε για να ολοκληρωθεί περίπου 2 λεπτά. Με τη χρήση του ακριβούς μοντέλου θα χρειαζόταν χρόνο της τάξης των 100 ωρών.

Αν και ο χρόνος υπολογισμού όλων των δεδομένων εκπαίδευσης του νευρωνικού δικτύου ήταν μεγάλος (περίπου ένας μήνας), η δυνατότητα προσομοίωσης διαφορετικών σεναρίων στη συνέχεια με τη χρήση του νευρωνικού δικτύου σε ελάχιστο χρόνο είναι πολύ σημαντική και αντισταθμίζει τον απαιτούμενο χρόνο για τη δημιουργία των δεδομένων εκπαίδευσης.

## ΒΙΒΛΙΟΓΡΑΦΙΑ

Λυκοθανάσης, Σ. & Γεωργόπουλος, Ε. (2003). Ηλεκτρονικές Εκπαιδευτικές Σημειώσεις Μαθήματος «*Βασικές Αρχές Νευρωνικών Δικτύων*». Τμήμα Μηχανικών Η/Υ και Πληροφορικής, Πανεπιστήμιο Πάτρας (<http://prlab.ceid.upatras.gr/>).

Νικολός, Ι.Κ. (2005). Σημειώσεις Μεταπτυχιακού Μαθήματος «*Ανάπτυξη Συστημάτων και Σχεδίαση Προϊόντων με Χρήση Ευφύων Μεθόδων Βελτιστοποίησης*». Τμήμα Μηχανικών Παραγωγής και Διοίκησης, Πολυτεχνείο Κρήτης.

Σπύρου, Σ. (1997). *Νευρωνικά Δίκτυα Ευθείας Τροφοδότησης με Ανάστροφη Διάδοση Σφάλματος. Ανάπτυξη Αλγόριθμων Εξομοίωσης. Ενδεικτική Εφαρμογή OCR*. ([http://platon.teipir.gr/new/ecs/pelab\\_1/ppt/neural.ppt](http://platon.teipir.gr/new/ecs/pelab_1/ppt/neural.ppt)).

Στεργιάδη, Μ. (2004). Διπλωματική Εργασία: «*Υδρογεωλογική Μελέτη του Βόρειου Τμήματος της Νήσου Ρόδου – Προσομοίωση Υπόγειας Ροής με χρήση του Τρισδιάστατου Μοντέλου Ροής Υπογείων Υδάτων και Μεταφοράς Ρύπων PTC*». Τμήμα Μηχανικών Περιβάλλοντος, Πολυτεχνείο Κρήτης.

Υπουργείο Γεωργίας, Γενική Διεύθυνση Εγγειοβελτιωτικών Έργων και Γεωργικών Διαρθρώσεων, Δ/ση Γεωλογίας – Υδρολογίας, Τμήμα Υδρογεωλογίας, Γεωτρήσεων και Μαθηματικών Ομοιωμάτων. (1999). *Υδρογεωλογική Μελέτη Νήσου Ρόδου και Σύνταξη Μαθηματικού Μοντέλου*.

Arndt, O., Barth, T., Freisleben, B., Grauer, M., (2005). *Approximating a Finite Element Model by Neural Network Prediction for Facility Optimization in Groundwater Engineering*. European Journal of Operational Research, 166. (p. 769 – 781).

Fogel, L.J. (1962). *Autonomous Automata*. Ind. Res., vol. 4. (p. 14 – 19).

Hebb, D.O. (1949). *The Organization of Behaviour*. New York: Wiley.

Holland, J.H. (1962). *Outline for a Logical Theory of Adaptive Systems*. J. Assoc. Comput. Mach., vol. 3. (p. 297 – 314).

Kröse, B. & Smagt, P. van der. (1996). *An Introduction to Neural Networks*. University of Amsterdam.

Maskey, S., Dibike, Y.B., Jonoski, A., Solomatine, D.P. (2000). *Groundwater Model Approximation with Artificial Neural Network for Selecting Optimum Pumping Strategy for Plume Removal*. In: AI Methods in Civil Engineering Applications (O. Schleider, A. Zijderfeld, eds), Cottbus. (p. 67 – 80).

McClelland, J.L. & Rumelhart, D.E. (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. The MIT Press.

McCulloch, W.S. & Pitts, W. (1943). *A Logical Calculus of the Ideas Immanent in Nervous Activity*. Bulletin of Mathematical Biophysics, 5. (p. 115 – 133).

Minsky, M. & Papert, S. (1969). *Perceptrons: An Introduction to Computational Geometry*. The MIT Press.

Nikolos, I.K. (2004). *Inverse Design of Aerodynamic Shapes Using Differential Evolution Coupled with Artificial Neural Network*. Design Optimization International Conference, Athens.

Rechenberg, I. (1965). *Cybernetic Solution Path of an Experimental Problem*. Royal Aircraft Establishment, Library Translation No. 1122, Farnborough, Hants, U.K.

Rosenblatt, F. (1959). *Principles of Neurodynamics*. New York: Spartan Books.

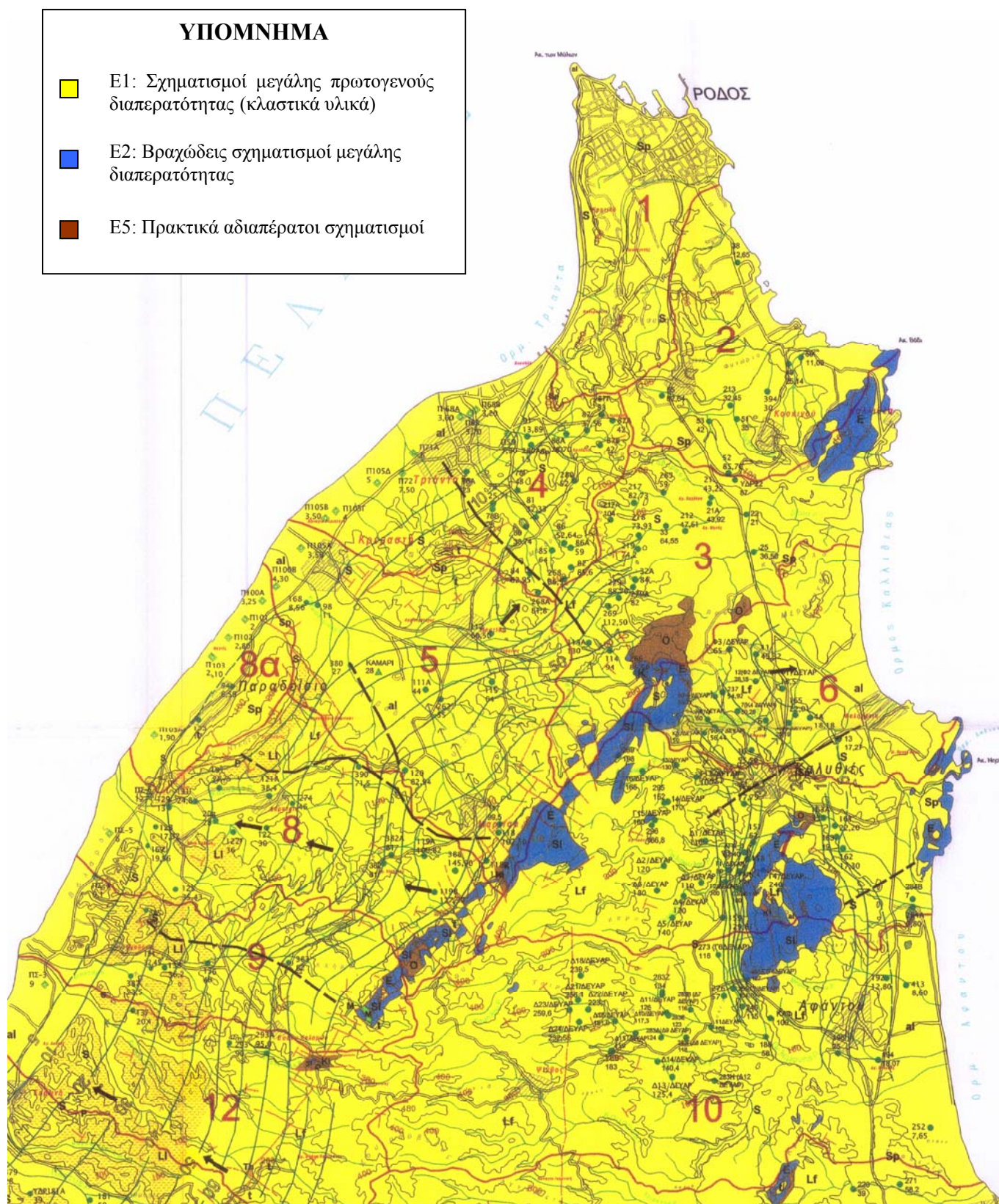
Schwefel, H.-P. (1968). *Projekt MHD – Staustahlrohr: Experimentelle Optimierung einer Zweiphasenduse, Teil I*. Technischer Bericht 11.034/68, 35 AEG Forschungsinstitut, Berlin, Germany.

Widrow, B. & Hoff, M.E. (1960). *Adaptive switching circuits*. In 1960 IRE WESCON Convention Record (p. 96 – 104). DUNNO.



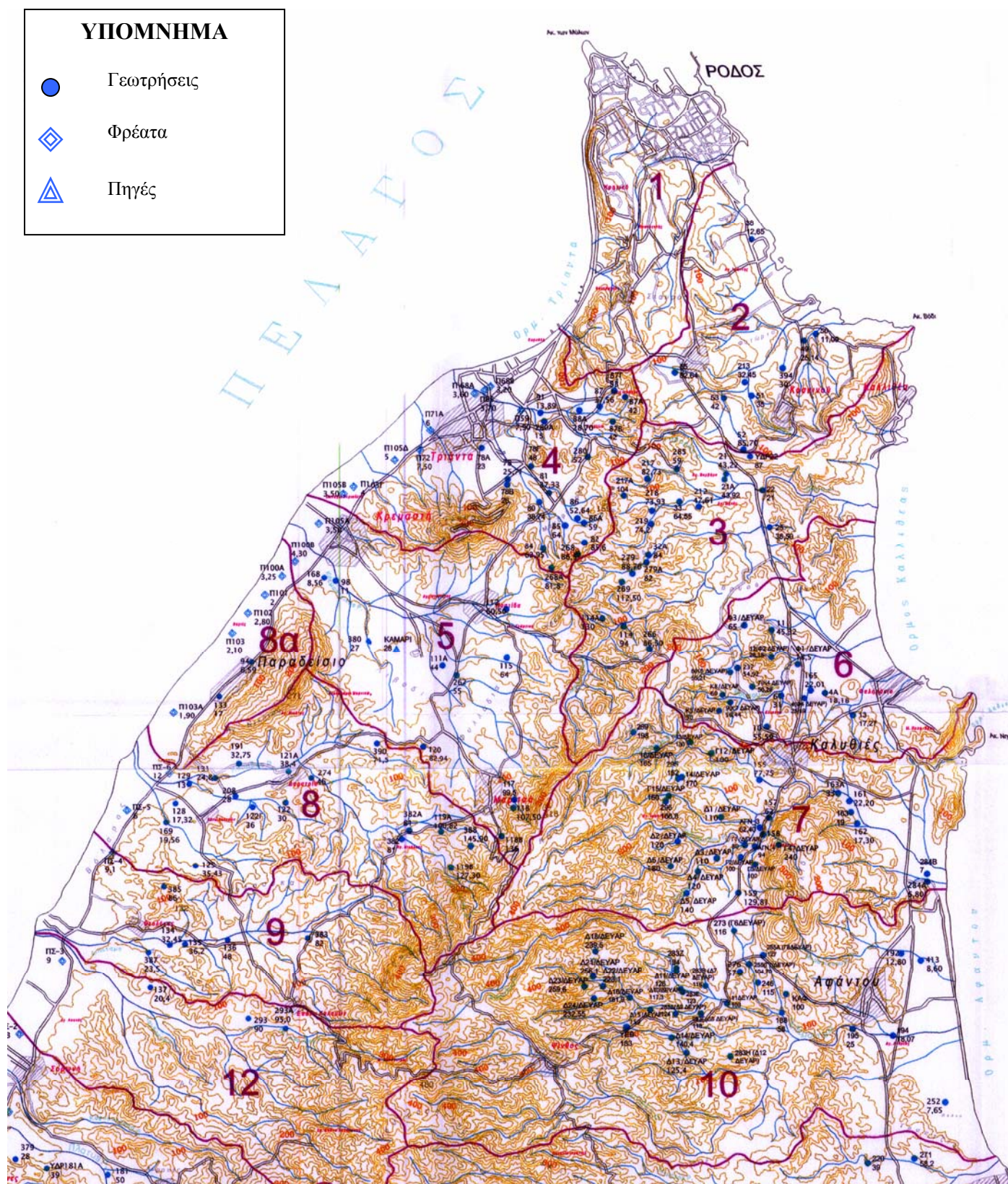
# **ΠΑΡΑΡΤΗΜΑ Ι**

## **Χάρτες περιοχής μελέτης**



**ΧΑΡΤΗΣ Ι – 1:** Γεωλογικός χάρτης και λεκάνες περιοχής μελέτης (Στεργιάδη, 2004)





**ΧΑΡΤΗΣ Ι – 2:** Χάρτης απογραφής σημείων εμφάνισης ύδατος περιοχής μελέτης  
(Στεργιάδη, 2004)





ΧΑΡΤΗΣ Ι – 3: Τοπογραφικός χάρτης περιοχής μελέτης (Στεργιάδη, 2004)

## **ΠΑΡΑΡΤΗΜΑ II**

### **Κώδικας σε C++ για τη δημιουργία των δυνατών συνδυασμών παροχών άντλησης**

```
#include <fstream>
using namespace std;

enum { SZ = 5 };

void compute(double *X, ofstream &out)
{
    int counter = 0;
    for(int i = 0; i < SZ; i++) {
        double t1 = X[i];
        for(int j = 0; j < SZ; j++) {
            double t2 = X[j];
            for (int k = 0; k < SZ; k++) {
                double t3 = X[k];
                for (int l = 0; l < SZ; l++) {
                    double t4 = X[l];
                    for (int m = 0; m < SZ; m++) {
                        double t5 = X[m];
                        out<<++counter<<":
"<<t1<<"\t"<<t2<<"\t"<<t3<<"\t"<<t4<<"\t"<<t5<<endl<<endl;
                    }
                }
            }
        }
    }
}

int main(int argc, char *argv[])
{
    double xvec[SZ] = { 0, 1000, 2000, 3000, 4000 };
    ofstream out("results.txt");

    compute(xvec, out);
    return 0;
}
```

## **ΠΑΡΑΡΤΗΜΑ ΙΙΙ**

### **Κώδικας σε Visual Basic για την κατασκευή της συνάρτησης προσαρμογής**

Attribute VB\_Name = "Module1"

'Dr. Ioannis K. Nikolos, 9-3-2006

'Program for postprocessing the NEURAL DATA

Public flowrate(10) As Double

Public t\_heights(50) As Double

Public l\_heights(50) As Double

Public nvars As Integer

Public npoints As Integer

'main sub of the code

Sub Main()

Call inputfile

Call outputfile

End Sub

Sub inputfile()

Dim i As Integer

Dim idum As Integer

Open ("chromosome.txt") For Input As #2

Input #2, nvars

For i = 1 To nvars

    Input #2, flowrate(i)

Next i

Close #2

Open ("lowerheights.txt") For Input As #3

Input #3, npoints

For i = 1 To npoints



```
    Input #3, idum, l_heights(i)
Next i
Close #3

Open ("heights.txt") For Input As #1
For i = 1 To npoints
    Input #1, t_heights(i)
Next i
Close #1

End Sub

'for writing the data to output file
Sub outputfile()
Dim i As Integer
Dim f1 As Double
Dim f2 As Double
Dim fitness As Double

f1 = 0#
For i = 1 To npoints
f1 = f1 + Abs(t_heights(i) - l_heights(i))
Next i
f1 = f1 / npoints

f2 = 0#
For i = 1 To nvars
f2 = f2 + flowrate(i)
Next i

f2 = f2 + 1000#
f2 = f2 / 2000#
f2 = 1# / f2
```

fitness = 5# \* f1 + f2

Open ("fitness.txt") For Output As #9

Print #9, fitness

Close #9

End Sub