



ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ  
ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ ΟΡΥΚΤΩΝ ΠΟΡΩΝ  
ΕΡΓΑΣΤΗΡΙΟ ΑΝΑΛΥΣΗΣ ΡΕΥΣΤΩΝ ΚΑΙ ΠΥΡΗΝΩΝ

Διπλωματική Εργασία :

Ανάπτυξη νευρωνικού δικτύου για την πρόβλεψη αναλυτικής σύστασης υπογείων ταμιευτήρων χρησιμοποιώντας υπόγειες μετρήσεις DFA.

Φοιτήτρια : Νομική Κωττάκη.

Επιβλέπων καθηγητής : Νικόλαος Βαρότσης.

Εξεταστική Επιτροπή : Νικόλαος Βαρότσης, Μιχάλης Γαλετάκης.

Βασίλειος Γαγάνης

Χανια

Οκτώβριος 2014

## ΕΥΧΑΡΙΣΤΙΕΣ

Αισθάνομαι την ανάγκη να ευχαριστήσω θερμά τον καθηγητή μου και επιβλέποντα της παρούσης εργασίας κ. Νικόλαο Βαρότση για την πολύτιμη συμπαράστασή του σε όλα τα στάδια της εκπόνησης της εργασίας αυτής. Οι συμβουλές και η καθοδήγησή του υπήρξαν πραγματικός πλούτος για την προσωπική μου πορεία.

Επίσης θα ήθελα να ευχαριστήσω τον κ. Βασίλειο Γαγάνη για την πολύτιμη βοήθειά του, την υπομονή του, και την εξαιρετική μεταδοτικότητα των γνώσεών του καθ'όλη τη διάρκεια της εκπόνησης της παρούσης εργασίας.

Τέλος, θέλω να ευχαριστώ τον κ. Δημήτριο Μαρινάκη που συνέβαλε κι αυτός με τις συμβουλές και την προσωπική βοήθειά του, στην διεκπεραίωση της διπλωματικής αυτής.

Το παρόν σύγγραμμα αφιερώνεται στους γονείς μου.



## Περίληψη.

Αντικείμενο της παρούσης διπλωματικής εργασίας είναι η ανάπτυξη ενός μοντέλου, που με την εφαρμογή των τεχνητών νευρωνικών δικτύων θα επιτρέπει την ασφαλή πρόβλεψη της γραμμομοριακής σύστασης πετρελαϊκών ρευστών της μορφής  $CO_2$ ,  $C_1$ ,  $C_2$ ,  $C_3$ ,  $C_4$ ,  $C_5$ ,  $C_{6-8}$ ,  $C_{9-11}$  και  $C_{12+}$ , αξιοποιώντας τις μετρήσεις του οπτικού αναλυτή ρευστών OFA (Optical Fluid Analyzer) που λαμβάνονται στο πεδίο κατά τη διάρκεια της αξιολόγησης μίας ερευνητικής γεώτρησης. Μέχρι σήμερα το OFA, που αποτελεί τμήμα του οργάνου διαγραφιών παραγωγής MDT, παρέχει κατά τη διάρκεια του τεστ μόνο την επί τοις εκατό κατά μάζα περιεκτικότητα του ρευστού σε  $CO_2$ ,  $C_1$ ,  $C_2$ ,  $C_{3-5}$ , και  $C_{6+}$ .

Για την ανάπτυξη του μοντέλου χρησιμοποιήθηκαν επιπλέον, πέραν των κατά μάζα περιεκτικότητων, μετρήσεις πεδίου που είναι πάντα διαθέσιμες όπως ο δείκτης πτητικότητας GOR, ο δείκτης πυκνότητας πετρελαίου API, και τέλος η σχετική πυκνότητα της αερίου φάσης  $S_g$ . Η εκπαίδευση των τεχνητών νευρωνικών δικτύων πραγματοποιήθηκε χρησιμοποιώντας μια βάση δεδομένων 752 δειγμάτων πετρελαϊκών ρευστών. Οι προβλέψεις του μοντέλου για την επί τοις εκατό κατά mole περιεκτικότητα των ρευστών σε  $C_3$ ,  $C_4$ ,  $C_5$  και  $C_{12+}$  κρίνονται ιδιαίτερα ικανοποιητικές, όπως αποδεικνύουν τα αποτελέσματα επί ενός σετ ελέγχου με δεδομένα γεωτρήσεων τα οποία δεν ελήφθησαν υπόψιν από το μοντέλο κατά τη διαδικασία της εκπαίδευσής του, ενώ απλώς ικανοποιητικές κρίνονται οι προβλέψεις του για την επί τοις εκατό κατά mole περιεκτικότητα των ρευστών στα κλάσματα  $C_{6-8}$  και  $C_{9-11}$ . Η επιτυχής ανάπτυξη του μοντέλου μπορεί να αποδειχθεί ιδιαίτερα χρήσιμη στην βιομηχανία πετρελαίου καθώς οι προβλέψεις του διευρύνουν την πληροφορία επί της σύστασης του ρευστού του ταμειυτήρα σε ένα πολύ πρώιμο στάδιο σε σχέση με ότι ισχύει μέχρι σήμερα. Η πληροφορία που παρέχει σήμερα ο οπτικός αναλυτής ρευστών OFA στην βιομηχανία πετρελαίου δεν δύναται να αξιοποιηθεί αμέσως από τα μοντέλα προσομοίωσης PVT των ταμειυτήρων καθώς αυτά απαιτούν να εισαχθεί ο

αναλυτικός χαρακτηρισμός κατά mole κι όχι κατά μάζα. Η διαθεσιμότητα των δεδομένων που παρέχει το μοντέλο ενισχύει την εμπιστοσύνη για την ακρίβεια των μελετών που θα εκπονηθούν από τους μηχανικούς πετρελαίων κατά τη διάρκεια του πρώιμου αυτού σταδίου κατά την αξιολόγηση μίας γεώτρησης.

## ΠΕΡΙΕΧΟΜΕΝΑ

## ΠΕΡΙΕΧΟΜΕΝΑ

ΚΕΦΑΛΑΙΟ 1ο .....	<b>Error! Bookmark not defined.</b>
1.0 Εισαγωγή .....	7
1.1 Γεωτρητικές διαγραφίες (well logging).....	7
1.2 Ιστορική αναδρομή.....	9
1.3 Διαγραφίες παραγωγής. ....	10
1.4 Τεστ παραγωγής και mdt. ....	11
1.5 τι είναι ο οπτικός αναλυτής ρευστων (OFA). ....	14
1.6 Θεωρητικό υπόβαθρο των μετρήσεων του OFA. ....	16
1.7 Διάταξη του OFA και αρχές λειτουργίας. ....	18
1.7.1 Σύστημα ανίχνευσης και διάκρισης των υγρών. ....	19
1.7.2 Σύστημα ανίχνευσης του αερίου. ....	22

1.8 Πλεονεκτήματα της μεθόδου. ....	24
1.9 Μειονεκτήματα της μεθόδου.....	25
1.10 Στόχος και σκοπός της εργασίας. ....	25
ΚΕΦΑΛΑΙΟ 2ο .....	28
2.0 Εισαγωγή. ....	28
2.1 Αντίστροφα προβλήματα. ....	28
2.2 Παλινδρομική ανάλυση. ....	30
2.3 Τεχνητά Νευρωνικά Δίκτυα. ....	32
2.3.1 Ιστορική αναδρομή.....	32
2.3.2 Τι είναι το Τεχνητό Νευρωνικό Δίκτυο. ....	33
2.3.3 Μοντέλο ενός τεχνητού νευρώνα. ....	34
2.3.4 Δίκτυα πρόσθιας τροφοδότησης. ....	37
2.3.5 Χαρακτηριστικά εκπαίδευσης.....	41
2.3.6 Εφαρμογές Τεχνητών Νευρωνικών Δικτύων στην βιομηχανία πετρελαίου. ....	42
ΚΕΦΑΛΑΙΟ 3 .....	47
3.0 Εισαγωγή. ....	47
3.1 Περιγραφή του περιβάλλοντος προγραμματισμού matlab. ....	47
3.2 Τι είναι ο συντελεστής απόκλισης. ....	48
3.2.1 Διάγραμμα Standing και Katz.....	49
3.2.2 Θεωρητικό υπόβαθρο διαγράμματος Standing-Katz.....	53
3.3 Η εξίσωση Beggs & Brill.....	54
3.4 Περιγραφή και αξιολόγηση του νευρωνικού δικτύου που αναπτύχθηκε για την πρόβλεψη του συντελεστή z.....	55
ΚΕΦΑΛΑΙΟ 4ο .....	62
4.0 Εισαγωγή. ....	62
4.1 Βάση δεδομένων και επεξεργασία. ....	63
4.2 Εφαρμογή των Τεχνητών Νευρωνικών Δικτύων. ....	71

4.3 Υπολογισμός γραμμομοριακών κλασμάτων. ....	107
4.4 Συμπεράσματα.....	122
5. Βιβλιογραφία : .....	123

## ΚΕΦΑΛΑΙΟ 1ο

## **1.0 ΕΙΣΑΓΩΓΗ**

Η διαδικασία της εκμετάλλευσης ενός ταμιευτήρα υδρογονανθράκων συνίσταται από τρία βασικά στάδια. Αρχικά από την ανίχνευση και τον εντοπισμό του σχηματισμού που ενδέχεται να φιλοξενεί υδρογονάνθρακες, έπειτα από το στάδιο των δοκιμαστικών γεωτρήσεων κατά το οποίο επιβεβαιώνεται η ύπαρξη ή μη οικονομικά εκμεταλλεύσιμης ποσότητας υδρογονανθράκων, και τέλος από το στάδιο της παραγωγής κατά το οποίο πραγματοποιείται η ανάπτυξη του ταμιευτήρα. Στο παρόν κεφάλαιο περιγράφονται κάποιες διαδικασίες που λαμβάνουν χώρα κατά το δεύτερο στάδιο των δοκιμαστικών γεωτρήσεων, και σχετίζονται με την αξιολόγηση του ταμιευτήρα, και της παραγωγικής του δυνατότητας [1].

### **1.1 ΓΕΩΤΡΗΤΙΚΕΣ ΔΙΑΓΡΑΦΙΕΣ (WELL LOGGING).**

Ως διαγραφή γεώτρησης (well logging) χαρακτηρίζεται η διαδικασία καταγραφής πληροφορίας, με διάφορες μεθόδους, που αφορά στην μέτρηση ιδιοτήτων των γεωλογικών σχηματισμών καθώς και των περιεχομένων σε αυτά ρευστών, που συναντώνται κατά το βάθος μιας γεώτρησης. Οι μετρήσεις αυτές γίνονται είτε προτού σωληνωθεί και κατά συνέπεια τσιμεντωθεί η γεώτρηση (open hole) είτε και μετά (cased hole) καθώς και σε κάποιες περιπτώσεις κατά τη διάρκεια της διάτρησης (logging while drilling ή LWD), και ως κύριο στόχο έχουν τον προσδιορισμό του όγκου και της ποιότητας των υδρογονανθράκων καθώς και την αξιολόγηση των σχηματισμών, με σκοπό την βέλτιστη ανάπτυξη του ταμιευτήρα. Οι μετρήσεις που καταγράφονται αναπαρίστανται συνήθως ως συναρτήσεις του βάθους, η επεξεργασία των οποίων καθώς και ο συνδυασμός τους οδηγεί στην αξιολόγηση των υπό έρευνα σχηματισμών [2].

Τα εργαλεία των τεστ αυτών (logging tools), κάποια εκ των οποίων παρίστανται στην Εικόνα 1.1, εισέρχονται εντός της γεώτρησης για να



μετρήσουν ακουστικές, ηλεκτρικές, ηλεκτρομαγνητικές ή ακόμα και ραδιενεργές ιδιότητες του υπεδάφους. Κάποιες φορές εκτελούν τις μετρήσεις τους όντας ενσωματωμένα στη διατρητική στήλη, ενώ άλλες κατέρχονται ανεξάρτητα εντός της γεώτρησης με ρυθμό κατάλληλο ώστε να καταγράφουν τα δεδομένα τους στα επιθυμητά βάθη.



**Εικόνα 1.1 :** Εργαλεία διαγραφιών. Το *dipmeter*, αριστερά, φέρει αισθητήρες σε 4 κινούμενους βραχίονες, οι οποίοι φαίνονται στην εκτεταμένη τους θέση. Σε έναν από τους 4 αυτούς βραχίονες βρίσκεται επισυναπτόμενη μια συστοιχία ηλεκτροδίων ενσωματωμένη σε καοτσούκ. Στη δεύτερη θέση από αριστερά βρίσκεται ένα ηχητικό καταγραφικό εργαλείο του οποίου χαρακτηριστικό είναι το διάτρητο περίβλημα, μια θέση δεξιά από αυτό φαίνεται μια συσκευή πυκνότητας με τον υδραυλικό της εφεδρικό βραχίονα πλήρως εκτεταμένο. Τέλος, τέρμα δεξιά βλέπετε μια άλλη έκδοση του *dipmeter* με πολλαπλές συστοιχίες ηλεκτροδίων. (Courtesy of Schlumberger) [2]

Τα όργανα αυτά συνδέονται μέσω καλωδίου με υπολογιστή που βρίσκεται στην επιφάνεια, έτσι ώστε κατά την διεξαγωγή του τεστ και με τη χρήση της ψηφιακής τηλεμετρίας, όλα τα δεδομένα της εκάστοτε μέτρησης να μεταδίδονται και να καταγράφονται επί τόπου στον υπολογιστή αυτόν, ο οποίος με τη σειρά του τα επεξεργάζεται και τα εμφανίζει ως συναρτήσεις του βάθους. Οι ταχύτητες της διεξαγωγής των τεστ ποικίλουν ανάλογα με τη φύση της συσκευής [2], [3].

## **1.2 ΙΣΤΟΡΙΚΗ ΑΝΑΔΡΟΜΗ.**

Εμπνευστές της ιδέας των γεωτρητικών διαγραφιών ήταν οι αδερφοί Conrad και Marcel Schlumberger. Πειραματιζόμενοι ήδη με επιφανειακά τεστ ηλεκτρικής αντίστασης από το 1912 αλλά και ηλεκτρομαγνητικά τεστ αργότερα, άρχισαν να παρέχουν τις υπηρεσίες τους το 1919. Η έμπνευσή τους βέβαια να κατεβάσουν το σύστημα ηλεκτροδίων κάθετα μέσα στη γεώτρηση, αλλάζει ριζικά το τοπίο στην άντληση πληροφορίας για το υπέδαφος κι έτσι, το πρώτο γεωτρητικό τεστ στην ιστορία φαίνεται να λαμβάνει χώρα στην Αλσατία (Γαλλία) το 1927 για τις ανάγκες της εταιρείας πετρελαίου Pechlebronn. Το τεστ αυτό ήταν τεστ ηλεκτρικής αντίστασης και στόχο είχε την κατανόηση της υποκείμενης στρωματογραφίας. Δύο χρόνια μετά, στις 19 Αυγούστου του 1929 το τεστ αυτό διεξάγεται για πρώτη φορά στην Αμερική, για λογαριασμό της εταιρίας Shell στην Καλιφόρνια [3], [4].

Από τότε μέχρι σήμερα, με την πρόοδο της τεχνολογίας αλλά και τις συνεχώς αυξανόμενες απαιτήσεις της βιομηχανίας πετρελαίου για βελτιστοποίηση της παραγωγικής διαδικασίας, ολοένα και πιο εξελιγμένα τεστ προκύπτουν, όπως διαγραφίες (logs) ακτίνων γάμμα που μετρούν τη ραδιενέργεια των πετρωμάτων με σκοπό να καθορίσουν την περιεκτικότητά τους σε σχιστόλιθο, ηχητικά τεστ που μετρούν το πορώδες από την ταχύτητα με την οποία ταξιδεύουν τα ηχητικά κύματα μέσα στα πετρώματα, τεστ ηλεκτρικής αντίστασης από τα οποία καθορίζονται οι τύποι των υγρών που εμπεριέχονται στο πέτρωμα, καθώς και τεστ οπτικής φασματοσκοπίας που αναλύουν ποσοτικά και ποιοτικά τα ρευστά του ταμιευτήρα [3]. Για τα τελευταία

προαναφερθέντα γίνεται λόγος παρακάτω, μιας και αποτελέσματα των μετρήσεων τους είναι αυτά που αξιοποιούνται για την ανάπτυξη της εν λόγω διπλωματικής εργασίας.

Στα πλαίσια της Μηχανικής Πετρελαίου, τα δεδομένα των διαγραφιών χρησιμοποιούνται από τους μηχανικούς για υπολογισμούς όπως η ακριβής τοποθεσία της κορυφής των σχηματισμών, το πορώδες, το πάχος τους, η θερμοκρασία, ο κορεσμός τους σε νερό και υδρογονάνθρακες, η εκτιμώμενη διαπερατότητά τους, η παρουσία πετρελαίου ή φυσικού αερίου, η αλατότητα του περιεχομένου νερού, καθώς και η εκτίμηση τυχόν αποθεμάτων τους. Με τις πληροφορίες αυτές λαμβάνεται η απόφαση για το κατά πόσο μια γεώτρηση θα αξιολογηθεί θετικά ή θα εγκαταλειφθεί. Γίνεται σαφές λοιπόν πως η διαδικασία των τεστ γεωτρήσεων δεν είναι απλώς μια καταγραφή του τι συμβαίνει κάτω από την επιφάνεια, αλλά συντελεί σημαντικά στη διαχείριση του ρίσκου της διαδικασίας παραγωγής ή στην βελτιστοποίησή της όταν ήδη αυτή εκτελείται [2].

### **1.3 ΔΙΑΓΡΑΦΙΕΣ ΠΑΡΑΓΩΓΗΣ.**

Οι διαγραφίες παραγωγής (production logging) αναφέρονται στην καταγραφή πληροφορίας που αφορά αποκλειστικά και μόνο στην διαδικασία της άντλησης των υδρογονανθράκων από τον ταμιευτήρα. Πιο συγκεκριμένα, εκτελούνται όταν έχει ήδη εκτιμηθεί από τις διαγραφίες γεώτρησης πως στους υποκείμενους σχηματισμούς εμπεριέχεται οικονομικά εκμεταλλεύσιμη ποσότητα υδρογονανθράκων, κι αφού το φρέαρ έχει υποστεί τσιμέντωση. Τα όργανα των τεστ αυτών (production logging tools) χρησιμοποιούνται για να εξακριβώσουν τόσο την φύση των ρευστών όσο και την συμπεριφορά τους κατά την ροή μέσα στο φρέαρ ή γύρω από αυτό, στα διάφορα στάδια της παραγωγής. Ανάλογα με την χρήση των οργάνων αυτών, οι διαγραφίες παραγωγής ομαδοποιούνται σε έξι βασικές κατηγορίες :

- Διάγνωση προβλημάτων παραγωγής.
- Αξιολόγηση της τσιμέντωσης του φρέατος.
- Έλεγχος της διάβρωσης.

- Καθορισμός των διεπιφανειών των ρευστών του ταμιευτήρα.
- Επιλογή ζωνών για τσιμέντωση.
- Μέτρηση της ροής.

Οι ερμηνείες που προκύπτουν από τις μετρήσεις τέτοιων οργάνων, σχετίζονται για παράδειγμα με την παραγωγική ικανότητα διαφορετικών ζωνών του ταμιευτήρα, τη διαπερατότητα και την μέση πίεσή τους, την ποσοτική σύσταση του παραγόμενου ρευστού, τις παροχές της κάθε φάσης, την κατανόηση πολυφασικών ροών σε περίπλοκα περιβάλλοντα γεωτρήσεων μεγάλης κλίσης, την αξιολόγηση των αποτελεσμάτων της έγχυσης νερού (water injection), καθώς και με την διάγνωση προβληματικών γεωτρήσεων παραγωγής. Για το σημαντικότερο όργανο αυτής της κατηγορίας γίνεται λόγος αμέσως μετά [5],[6].

#### **1.4 ΤΕΣΤ ΠΑΡΑΓΩΓΗΣ ΚΑΙ MDT.**

Τα αποτελέσματα των τεστ παραγωγής ή DST (Drill Stem Test) δίνουν τις πρώτες πληροφορίες για την παραγωγική δυναμικότητα του ταμιευτήρα, καθώς εξετάζουν την σχέση της παροχής του ρευστού κατά την άντλησή του, με την πτώση πίεσης που συνεπάγεται η άντληση αυτή. Η πραγματοποίηση των τεστ αυτών προϋποθέτει την τσιμέντωση του φρέατος. Κατά την εκτέλεση του τεστ παραγωγής απομονώνεται μία ζώνη του φρέατος με την βοήθεια μονάδων διπλής στεγανοποίησης (dual-packer module), και με τη χρήση βαλβίδων επιτρέπεται η ροή ρευστού του ταμιευτήρα μόνο μέσα στη σωλήνωση της γεώτρησης (drillpipe). Κατά τη διάρκεια της ροής αυτής καταγράφεται η πτώση πίεσης συναρτήσει του χρόνου.

Το MDT (Modular Formation Dynamics Tester, © trademark of Schlumberger) είναι το σημαντικότερο όργανο που χρησιμοποιείται προτού πραγματοποιηθεί τεστ παραγωγής DST για να παράσχει τα πρώτα δεδομένα παραγωγής ώστε σε περίπτωση που τα αποτελέσματα που θα δώσει δεν είναι ενθαρρυντικά, να αποφευχθεί η χρονική καθυστέρηση και το κόστος που συνεπάγεται η διεξαγωγή ενός DST. Επιτυγχάνεται έτσι και χρόνος, λόγω της άμεσης πρόσβασης σε πληροφορίες και δεδομένα, κι από την άλλη προκύπτει

οικονομικό όφελος για την επένδυση, αφού μικραίνει ο χρόνος ενοικίασης των εγκαταστάσεων της πλατφόρμας.

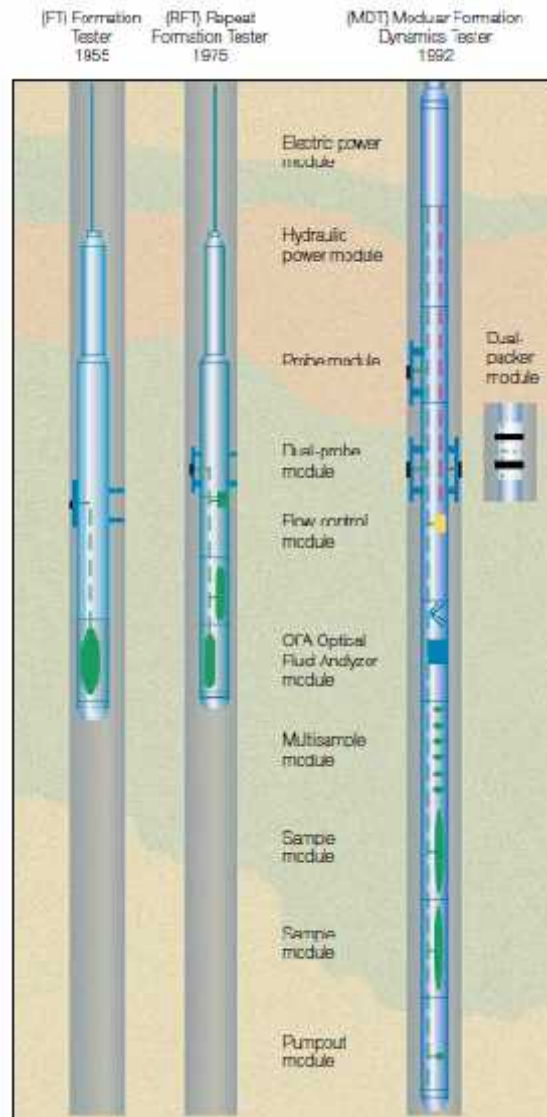
Οι μετρήσεις του MDT αφορούν κυρίως στην παραγωγική δυνατότητα του ταμιευτήρα. Οι κύριες εφαρμογές του είναι :

- Ο υπολογισμός της πίεσης ταμιευτήρα.
- Η ποσοτική ανάλυση του ρευστού καθώς και η δειγματοληψία αυτού με ταυτόχρονη παρακολούθηση της ποιότητας του δείγματος, ως προς ενδεχόμενη μόλυνσή του από διήθημα γεωτρητικού ρευστού.
- Η εκτίμηση του λόγου αερίου/πετρελαίου (GOR).
- Η εκτίμηση της κάθετης διαπερατότητας του πετρώματος.
- Η εξακρίβωση περιοχών που παρεμποδίζουν τη ροή.
- Ο υπολογισμός της πυκνότητας του ρευστού.
- Ο υπολογισμός του pH του ρευστού.
- Η εξακρίβωση της θέσης των διεπιφανειών νερού-πετρελαίου και πετρελαίου-αερίου.

Το MDT ως η πιο εξελιγμένη και αναβαθμισμένη έκδοση των ενσύρματων εργαλείων αξιολόγησης του σχηματισμού (wireline formation testers) (Εικόνα 1.2), παρέχει τη δυνατότητα πολλαπλής δειγματοληψίας με μία μόνο κάθοδο στη γεώτρηση, και χάρις στην χρήση ενός μανομέτρου χαλαζία τελευταίας γενιάς (quartz gauge), δίνει σύντομη και ακριβή μέτρηση της πίεσης σχηματισμού. Η ομαλή εκτέλεση της δειγματοληψίας εξασφαλίζεται με τον έλεγχο της πτώσης της πίεσης μέσα στον δειγματοληπτικό θάλαμο, από ένα εξελιγμένο ηλεκτρο-υδραυλικό σύστημα βαλβίδων. Ως προς τη διάταξή του, τέσσερα είναι τα βασικά μέρη που το συνθέτουν :

- Η ηλεκτρική μονάδα (electrical module), που τροφοδοτεί όλα τα ηλεκτρονικά και ηλεκτρουδραυλικά συστήματα.
- Η μονάδα υδραυλικής ισχύος (hydraulic power module), που τροφοδοτεί τις μονάδες που λειτουργούν με υδραυλική πίεση.
- Το μανόμετρο (single-probe module), που μετρά την πίεση του σχηματισμού.

- Οι θάλαμοι δειγματοληψίας (sample chamber modules), που ποικίλουν σε αριθμό και μεγέθη και συναρμολογούνται πάνω στον βασικό σκελετό.



**Εικόνα 1.2 :** Εξέλιξη των ενσύρματων εργαλείων αξιολόγησης σχηματισμού. Από αριστερά προς δεξιά : Το FT (Formation Tester) είχε ένα μανόμετρο για να καταγράψει την δειγματοληψία μέσα σε ένα μονό κελί. Με το RFT (Repeat Formation Tester) εισάγονται επιπλέον κελιά, για να εξασφαλιστεί μια ένδειξη της διαπερατότητας του σχηματισμού πριν την δειγματοληψία του ρευστού. Τέλος με το MDT (Modular Formation Dynamics Tester) εισάγονται η μονάδα διπλής στεγανοποίησης (dual probe module) που επιτρέπει την απομόνωση ζώνης που παράγει, και η μονάδα εκτροπής της ροής (pumpout module) που σε συνδυασμό με την μονάδα οπτικής ανάλυσης ρευστών (OFA Optical Fluid Analyzer module) επιτρέπει την δειγματοληψία μόνο όταν ρεύσει το επιθυμητό ρευστό [7].

Πέραν όμως των βασικών μονάδων του MDT υπάρχουν και κάποια συστήματα που του προσδίδουν επιπλέον δυνατότητες. Κάποια από αυτά είναι η μονάδα πολλαπλής δειγματοληψίας (multisample module), που του παρέχει τη δυνατότητα λήψης έξι δειγμάτων των 450 ml, από διαφορετικά βάθη της γεώτρησης, η μονάδα διπλής στεγανοποίησης (dual-packer module), που επιτρέπει την απομόνωση μίας ζώνης του φρέατος από τη στήλη της λάσπης, με στόχο τη διεξαγωγή του τεστ ροής σε συγκεκριμένη ζώνη του ταμιευτήρα, η μονάδα ελέγχου ροής (flow control module), που παρέχει τη δυνατότητα ελέγχου της πτώσης πίεσης κατά τη ροή του δείγματος στο θάλαμο δειγματοληψίας καθώς και η μονάδα εκτροπής της κατεύθυνσης ροής (rumpout module), που επιτρέπει την απόρριψη μολυσμένου δείγματος πριν την δειγματοληψία [8].

### **1.5 ΤΙ ΕΙΝΑΙ Ο ΟΠΤΙΚΟΣ ΑΝΑΛΥΤΗΣ ΡΕΥΣΤΩΝ (OFA).**

Το OFA (optical fluid analyzer) ή LFA (live fluid analyzer) ή InSitu Fluid Analyser (© trademark of Schlumberger), δηλαδή ο οπτικός αναλυτής ρευστών, είναι μία διάταξη που, προσαρμοζόμενη στο MDT, επιτρέπει την αξιολόγηση του προς δειγματοληψία ρευστού, ελέγχοντας εάν σε αυτό περιέχεται γεωτρητικό διήθημα. Όπως θα εξηγηθεί παρακάτω, χρησιμοποιώντας την φασματοσκοπία οπτικής απορρόφησης το OFA παρέχει τη δυνατότητα της, σε πραγματικό χρόνο, επί τόπου ποσοτικής ανάλυσης της σύστασης του ρέοντος ρευστού από τον ταμιευτήρα στον δειγματολήπτη. Για την ακρίβεια δίνει την επί τοις εκατό κατά μάζα σύσταση του ρευστού σε  $CO_2$ ,  $C_1$ ,  $C_2$ ,  $C_{3-5}$  και  $C_{6+}$ .

Η απόκτηση δείγματος ρευστού του ταμιευτήρα είναι εξέχουσας σημασίας, αφού τα δείγματα αυτά αποστέλλονται για την εργαστηριακή μελέτη της ισορροπίας φάσεων του ρευστού. Η μελέτη αυτή είναι άκρως απαραίτητη καθώς από αυτήν εξάγονται τα συμπεράσματα για την θερμοδυναμική συμπεριφορά των αποθεμάτων κατά την παραγωγή και παρέχει τις τιμές των φυσικών ιδιοτήτων που απαιτούνται για τη μοντελοποίηση της παραγωγής.

Πριν τη χρήση του OFA η διαδικασία της δειγματοληψίας δεν εξασφάλιζε την απόκτηση καθαρού δείγματος, αφού δεν μπορούσε να εκτιμηθεί σε συνθήκες φρέατος η παρουσία ή μη λάσπης στο δείγμα. Σε περίπτωση λήψης μολυσμένου δείγματος η διαδικασία δειγματοληψίας έπρεπε να επαναληφθεί με ευνόητη σοβαρή οικονομική επιβάρυνση. Με το OFA εξακριβώνεται η παρουσία γεωτρητικού διηθήματος στο ρευστό κατά την ροή του από τον σχηματισμό στο MDT, κι έτσι το ρευστό εκτρέπεται με το pump out module από τον θάλαμο δειγματοληψίας μέχρι να εισρεύσει καθαρό δείγμα και να επιτραπεί η κατεύθυνση του ρευστού προς τον δειγματολήπτη. Την εκτροπή αυτή του ρευστού κατά το δοκούν, επιτρέπει μία διάταξη βαλβίδων που ελέγχεται ηλεκτρονικά από το εργοτάξιο κατά την διεξαγωγή του τεστ.

Επίσης, πριν την εμφάνιση του OFA η μόνη πληροφορία σχετικά με τη σύσταση του δείγματος, προέκυπτε μόνο με χρήση της αέριας χρωματογραφίας, μετά την εξαγωγή του δείγματος από το φρέαρ και την αποστολή του στο εργαστήριο. Με τη χρήση όμως του OFA προσφέρεται η πληροφορία της επί τοις εκατό κατά μάζα σύστασης του ρευστού εντός του φρέατος, και σε πραγματικό χρόνο, σε ένα τόσο πρώιμο ερευνητικό στάδιο (open hole).

Το OFA είναι έργο της επιστημονικής ομάδας OAS (Optics and Sensors) που είναι υπεύθυνη για τον σχεδιασμό και την ανάπτυξη των oil logging tools, οπτικών αισθητήρων και τεχνολογιών αισθητήρων, του ομίλου εταιρειών Schlumberger. Στις αρχές του 1990, με την δημιουργία του OFA η επιστημονική αυτή ομάδα καταφέρνει πρώτη στον κόσμο να επιτύχει φασματοσκοπική ανάλυση δείγματος υδρογονανθράκων μέσα στις γεωτρήσεις, σε συνθήκες υψηλής πίεσης και θερμοκρασίας. Το OFA είναι το πρώτο οπτικό εργαλείο της Schlumberger που κυκλοφόρησε στην αγορά της βιομηχανίας πετρελαίου, και χαίρει εξαιρετικής φήμης ανάμεσα στις πετρελαϊκές εταιρείες που χρησιμοποιούν τις υπηρεσίες του ευρύτατα [6].

Το OFA έχει εφαρμοστεί μέχρι τώρα με εξαιρετική επιτυχία. Ένα χαρακτηριστικό παράδειγμα επιτυχούς εφαρμογής του, σε συνδυασμό με τεχνικές που αποσκοπούν στην ελαχιστοποίηση της μόλυνσης του δείγματος, αποτελεί η απόκτηση δείγματος υψηλής καθαρότητας (low contamination



sample) από ταμειευτήρα στην Αγκόλα [9]. Στην περίπτωση αυτή υπήρχε η ανάγκη απόκτησης δείγματος ρευστού προς αποστολή του για PVT ανάλυση, για τον χαρακτηρισμό των αποθεμάτων σε εξωχώριο περιβάλλον. Κατά τη διάρκεια του τεστ, γινόταν παρακολούθηση του επιπέδου μόλυνσης του δείγματος από γεωτρητικό διήθημα, μέσω των καταγραφών της μονάδας του OFA σε πραγματικό χρόνο ροής του ρευστού από τον σχηματισμό στο MDT, μέχρι να κριθεί πότε πρέπει να ξεκινήσει η δειγματοληψία. Αποτέλεσμα αυτού του τεστ ήταν η ανάκτηση δύο δειγμάτων πετρελαίου με μόλις 3.3% και 4.2% κατά βάρος περιεκτικότητα σε γεωτρητικό διήθημα, ενώ ο αρχικός σχεδιασμός περιελάμβανε ανοχή 10%.

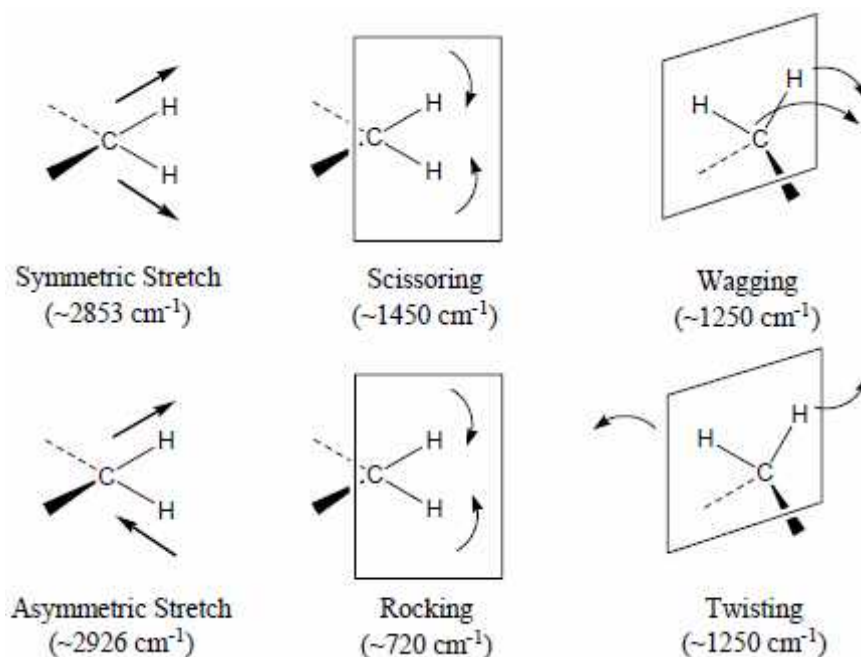
Αρχικά το OFA προέκυψε ως σύνεργο του MDT (Modular Formation Dynamics Tester) αφενός για να εξασφαλίσει πως τα κελιά δειγματοληψίας θα πληρωθούν μόνο όταν υπάρχει το επιθυμητό δείγμα ρευστού, επαληθεύοντας ότι ο σχηματισμός παράγει, κι αφετέρου, για να προσφέρει την πληροφορία περί του κατά πόσον η πίεση δειγματοληψίας βρίσκεται πάνω από το σημείο φυσαλίδας (bubblepoint) ανιχνεύοντας την ύπαρξη ή μη αερίου φάσεως μέσα στη ροή που επιτυγχάνεται με μία ειδική μέτρηση του ανακλώμενου πολωμένου φωτός η οποία περιγράφεται στη συνέχεια [7], [6].

## **1.6 ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ ΤΩΝ ΜΕΤΡΗΣΕΩΝ ΤΟΥ OFA.**

Στα πλαίσια της Φυσικής, η Φασματοσκοπία αποτελεί κλάδο της Κυματικής Οπτικής και σχετίζεται με τη μελέτη της αλληλεπίδρασης της ύλης με την ηλεκτρομαγνητική ακτινοβολία. Η αλληλεπίδραση αυτή έγκειται στην απορρόφηση της ενέργειας του ταλαντευόμενου ηλεκτρομαγνητικού πεδίου του φωτός, από τα άτομα της ύλης. Πιο συγκεκριμένα αυτό που συμβαίνει όταν ένα μόριο τεθεί σε ηλεκτρομαγνητικό πεδίο, εάν φωτιστεί για παράδειγμα με ακτινοβολία κατάλληλης συχνότητας, είναι η μεταφορά ενέργειας από το πεδίο στο μόριο.

Η απορρόφηση της ενέργειας αυτής, έχει ως αποτέλεσμα την μετάβαση των μορίων σε υψηλότερες ενεργειακές στάθμες. Οι απορροφήσεις που λαμβάνουν χώρα στην περιοχή του υπερύθρου, οφείλονται κυρίως στις

δονήσεις και τις κάμψεις των δεσμών των μορίων και δευτερευόντως σε κάποια άλλα είδη παραμόρφωσής τους όπως τα φαινόμενα κλυδωνισμού (rocking), στρέβλωσης (twisting) ή ψαλιδωτής κίνησης (scissoring).



**Εικόνα 1.3 :** Αναπαράσταση δονήσεων των μοριακών δομών και αναφορά περιοχών απορρόφησης στην περιοχή του υπέρυθρου φάσματος της ηλεκτρομαγνητικής ακτινοβολίας [10].

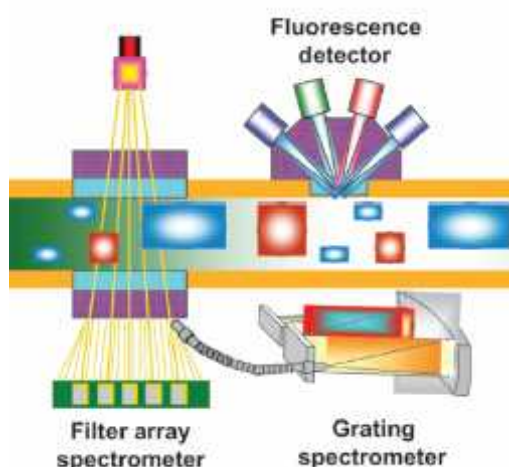
Η τεχνική της υπέρυθρης φασματοσκοπίας ή αλλιώς IR βρίσκει εφαρμογή στην ποσοτική ανάλυση οργανικών ενώσεων ή μίγματος χημικών ουσιών όταν δεν μπορούν να εφαρμοστούν άλλες ακριβέστερες μέθοδοι όπως είναι η υπεριώδης φασματοσκοπία ή η φασματοσκοπία πυρηνικού μαγνητικού συντονισμού [10], [11]. Είναι μία από τις φασματοσκοπικές τεχνικές, που μελετά την απορρόφηση ενέργειας στην περιοχή υπέρυθρου του φάσματος της ηλεκτρομαγνητικής ακτινοβολίας. Η χρήση της τεχνικής αυτής στην Οργανική Χημεία είναι ευρεία, καθώς η λήψη των φασμάτων απορρόφησης αλλά και η σύγκριση αυτών με φάσματα γνωστών οργανικών ενώσεων είναι εύκολη.

Αυτό που μετράται στην φασματοσκοπία υπερύθρου είναι η σχετική ένταση του φωτός, ή αλλιώς η οπτική διαπερατότητα  $T$  (transmittance) του μέσου που εξετάζεται, που ορίζεται ως ο λόγος της έντασης του εκπεμπόμενου φωτός (από την πηγή που παράγει την ηλεκτρομαγνητική ακτινοβολία) προς την ένταση του φωτός από το σώμα που ακτινοβολείται και προσπίπτει στους ανιχνευτές.

## 1.7 ΔΙΑΤΑΞΗ ΤΟΥ OFA ΚΑΙ ΑΡΧΕΣ ΛΕΙΤΟΥΡΓΙΑΣ.

Η μονάδα του OFA, μέσα από την οποία περνά η γραμμή ροής του δείγματος κατά τη διάρκεια του τεστ του MDT, αποτελείται από δύο ανεξάρτητα συστήματα αισθητήρων όπως φαίνεται και στην Εικόνα 1.4.

Το πρώτο χρησιμοποιείται για την ανίχνευση και ανάλυση των υγρών φάσεων του μίγματος σε νερό,  $CO_2$ ,  $C_1$ ,  $C_2$ ,  $C_{3-5}$  και  $C_{6+}$  μέσω της φασματοσκοπίας απορρόφησης, και το δεύτερο για την ανίχνευση αερίου μέσω ενός ειδικού τύπου μέτρησης της οπτικής ανάκλασης. Επειδή οι δύο αυτές διατάξεις διαφέρουν τόσο ως προς τις αρχές λειτουργίας όσο και ως προς τα φυσικά φαινόμενα που διέπουν τις μετρήσεις που εκτελούν, κρίνεται ορθό να εξηγηθούν χωριστά.



**Εικόνα 1.4 :** Αναπαράσταση της μονάδας του OFA με τα δύο συστήματα αισθητήρων. Το πρώτο αποτελείται από δύο φασματομέτρα (*Filter array spectrometer & Grating spectrometer*), και το δεύτερο από τον ανιχνευτή αερίου φάσεως (*Fluorescence detector*) [12]

### 1.7.1 ΣΥΣΤΗΜΑ ΑΝΙΧΝΕΥΣΗΣ ΚΑΙ ΔΙΑΚΡΙΣΗΣ ΤΩΝ ΣΥΣΤΑΤΙΚΩΝ.

Η φυσική ιδιότητα των διαφόρων ενώσεων, που απαρτίζουν το μίγμα υδρογονανθράκων, που εκμεταλλεύεται το OFA για να πετύχει τη διάκρισή τους, είναι η διαφορά της οπτικής μετάδοσης του φωτός στο ορατό και κοντά στο υπέρυθρο μήκος κύματος. Η διαφορά αυτή στην οπτική μετάδοση του φωτός οφείλεται στην απορρόφηση ενέργειας από τα συστατικά στην περιοχή του υπέρυθρου του φάσματος της ηλεκτρομαγνητικής ακτινοβολίας.

Πιο συγκεκριμένα η απορρόφηση ενέργειας από τα μόρια των υδρογονανθράκων οφείλεται κυρίως στον δεσμό  $C-H$ . Τα μόρια του μεθανίου έχουν τέσσερις πανομοιότυπους δεσμούς  $C-H$ , τα μόρια των ελαφριών υδρογονανθράκων (αιθάνιο, προπάνιο, βουτάνιο, πεντάνιο) έχουν ως κύριο χαρακτηριστικό τη μεθυλική ομάδα  $C-H_3$ , ενώ τα βαρύτερα μόρια των υδρογονανθράκων περιέχουν κυρίως την ομάδα  $-CH_2-$  και ενίοτε αρωματικές ενώσεις όπως ρητίνες και ασφαλτένια. Η απορρόφηση της ενέργειας από τα μόρια του νερού οφείλεται στους δύο πανομοιότυπους δεσμούς  $H-O$ , και αυτή του διοξειδίου του άνθρακα στους δύο πανομοιότυπους δεσμούς  $C-O$ . Αυτές οι διαφορές στις δομές των μορίων συνεπάγονται και μια διαφορετική συμπεριφορά του κάθε συστατικού κατά την αλληλεπίδρασή του με την ηλεκτρομαγνητική ακτινοβολία, δηλαδή το κάθε συστατικό απορροφά ενέργεια σε συγκεκριμένα μήκη κύματος της περιοχής του υπέρυθρου του φάσματος της ακτινοβολίας [12].

Για να πραγματοποιηθεί η ταυτοποίηση των συστατικών από τον οπτικό αναλυτή ρευστών, αυτό που μετράται από τους ανιχνευτές σε διάφορα μήκη κύματος είναι η σχετική ένταση του φωτός ή οπτική διαπερατότητα του μέσου  $T$  (transmittance), που ορίζεται ως ο λόγος της έντασης του εκπεμπόμενου, από την πηγή, φωτός προς την ένταση του φωτός που ακτινοβολείται από το σώμα και προσπίπτει στους ανιχνευτές. Ένας όρος που έχει πρακτική σημασία για την ερμηνεία των μετρήσεων αυτών είναι η οπτική πυκνότητα, που ορίζεται ως εξής :

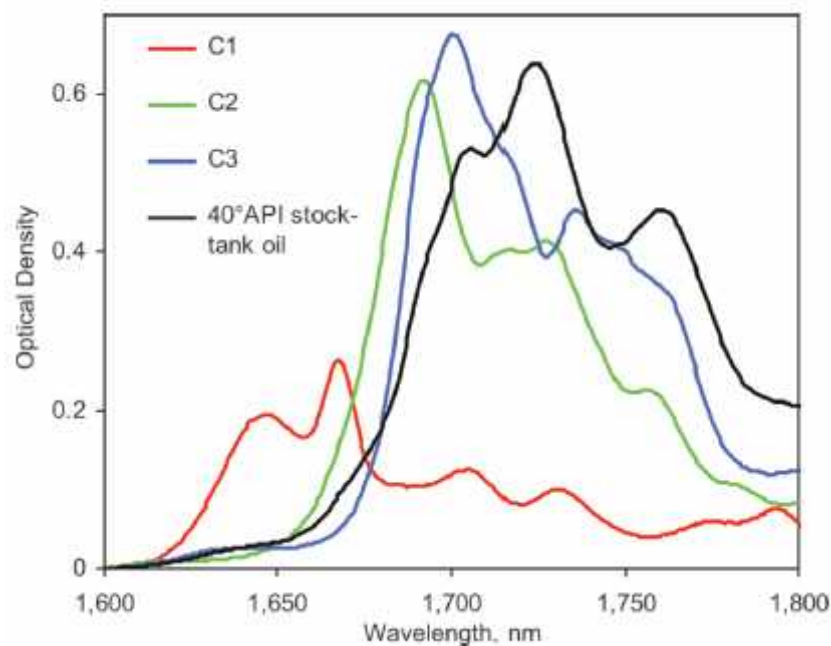
$$d = \log\left(\frac{1}{T}\right)$$

Όπου :

$d$  : η οπτική πυκνότητα.

$T$  : οπτική διαπερατότητα.

Έτσι η οπτική πυκνότητα είναι μέτρο της απορρόφησης ενέργειας του ταλαντευόμενου ηλεκτρομαγνητικού πεδίου του φωτός από τα διάφορα μόρια του προς εξέταση δείγματος. Το διάγραμμα που προκύπτει από την οπτική πυκνότητα ως προς τα μήκη κύματος καλείται φάσμα απορρόφησης (Εικόνα 1.5), και είναι οι χαρακτηριστικές κορυφές των καμπυλών του, που καθιστούν εφικτή τη διάκριση των διαφορετικών υγρών.



**Εικόνα 1.5 :** Στο διάγραμμα αυτό αναπαρίστανται οι χαρακτηριστικές καμπύλες του φάσματος απορρόφησης που αντιστοιχούν στο  $C_1$  (κόκκινο χρώμα), στο  $C_2$  (πράσινο χρώμα), στο  $C_3$  (μπλε χρώμα) και στο  $C_6$  (μαύρο χρώμα). Στον άξονα των  $y$  βρίσκεται η οπτική πυκνότητα, ενώ στον άξονα των  $x$  τα μήκη κύματος εκπομπής της ηλεκτρομαγνητικής ακτινοβολίας [7].

Ο λόγος για τον οποίον ο προσδιορισμός της οπτικής πυκνότητας έχει πρακτική σημασία, είναι γιατί από την τιμή της οπτικής πυκνότητας ενός συστατικού υπολογίζεται η συγκέντρωσή του στο μίγμα που περιέχεται, όπως υποδεικνύεται από τον νόμο Beer-Lambert [12]. Ο νόμος αυτός προβλέπει γραμμική συσχέτιση της συγκέντρωσης ενός συστατικού με την οπτική πυκνότητά του ως εξής [14]:

$$d = E^{1\%} * b * c$$

Όπου :

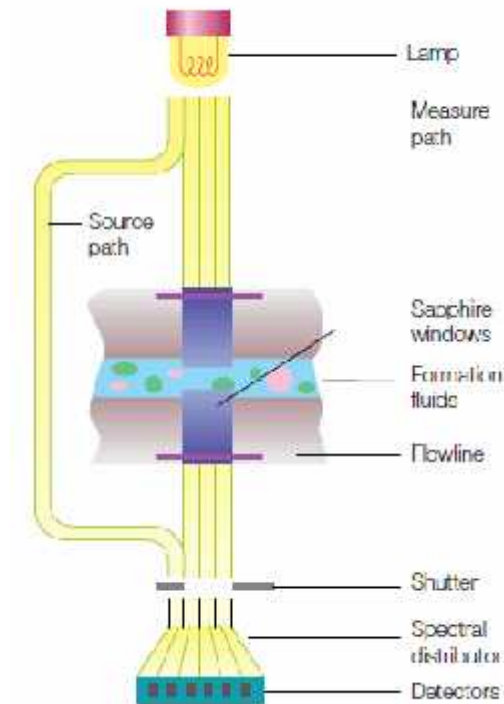
$d$  = η οπτική πυκνότητα.

$E$  = ο συντελεστής μοριακής απορρόφησης που εξαρτάται από το μήκος κύματος της ακτινοβολίας κι έχει μονάδες  $\frac{ml}{gr * cm}$ .

$b$  = το μήκος της διαδρομής που διανύει το ρευστό κατά τη μέτρηση σε  $cm$ .

$c$  = η συγκέντρωση του εν διαλύσει συστατικού σε  $\frac{gr}{100ml}$ .

Η μέτρηση, μέσα στη γραμμή ροής του MDT, πραγματοποιείται ως εξής : Από μία λυχνία αλογόνου παράγεται υπέρυθρη ακτινοβολία η οποία κατευθύνεται σε δύο διαδρομές, στη μία για να γίνει η βαθμονόμηση του οργάνου στο περιβάλλον του φρέατος, και στην άλλη για να γίνει η μέτρηση στη γραμμή ροής. Το φως, διασχίζει κάθετα τη γραμμή ροής κατευθυνόμενο από ένα παράθυρο ζαφειριού προς έναν διανομέα φωτός, (Εικόνα 1.6) όπου ανιχνευτές φωτός, έκαστος ρυθμισμένος σε διαφορετικό μήκος κύματος, μετρούν τη σχετική ένταση του λαμβανόμενου φωτός η οποία μετατρέπεται σε οπτική πυκνότητα για να κατασκευαστούν στη συνέχεια τα φάσματα απορρόφησης. Από τα φάσματα αυτά πραγματοποιείται η διάκριση των συστατικών.



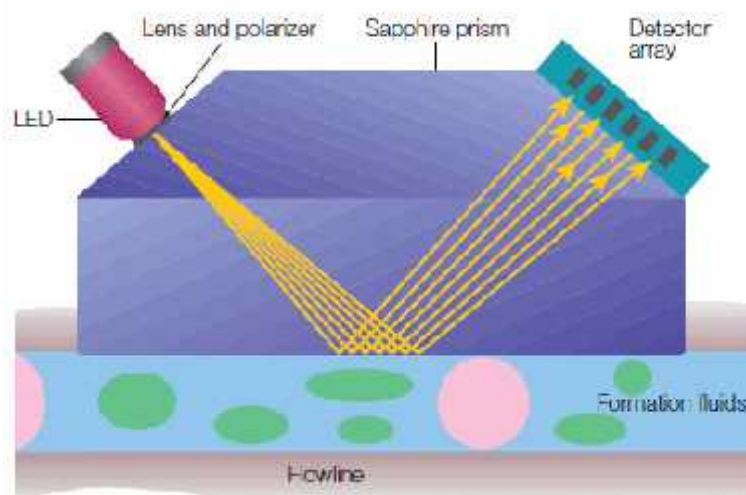
**Εικόνα 1.6 :** Διάταξη φασματοσκοπικής ανάλυσης της ρευστής φάσης. Λυχνία (Lamp), διαδρομή μέτρησης (Measure path), διαδρομή βαθμονόμησης (Source path), παράθυρα ζαφειριού (Sapphire windows), ρευστό του σχηματισμού (Formation fluids), γραμμή ροής (Flowline), διανομέας φωτός (Spectral distributor), ανιχνευτές φωτός (Detectors) [7].

### 1.7.2 ΣΥΣΤΗΜΑ ΑΝΙΧΝΕΥΣΗΣ ΤΟΥ ΑΕΡΙΟΥ.

Η ανίχνευση του αερίου επιτυγχάνεται με τη μέτρηση του ανακλώμενου πολωμένου φωτός από τη διεπιφάνεια ρευστού και παραθύρου ζαφειριού. Η ανάκλαση του φωτός διέπεται από την κρίσιμη γωνία, που είναι η γωνία πρόσπτωσης πέραν της οποίας η επιφάνεια που δέχεται το εκπεμπόμενο φως δρα ως κάτοπτρο και επιτρέπει την ολική ανάκλασή του. Έτσι, σε γωνίες πρόσπτωσης μεγαλύτερες από την κρίσιμη, όλο το φως ανακλάται, ενώ σε μικρότερες αυτής κάποιο ποσοστό του φωτός σκεδάζεται. Στην περίπτωση τώρα, που το φως είναι πολωμένο παράλληλα με το επίπεδο πρόσπτωσης, υπάρχει μία γωνία μικρότερη της κρίσιμης, που επιτρέπει την εκατό τοις εκατό σκέδαση του φωτός. Η γωνία αυτή ονομάζεται γωνία Brewster, και καθώς οι τιμές της μαζί με τις τιμές των κρίσιμων γωνιών διαφέρουν σημαντικά για τα διάφορα συστατικά, η μέτρηση των σχετικών εντάσεων του ανακλώμενου

φωτός σε ένα συγκεκριμένο εύρος γωνιών επιτρέπει την ταυτοποίηση της αέριας φάσης. Ο εντοπισμός του αερίου πραγματοποιείται στιγμιαία μέσω της χρήσης ενός απλού αλγορίθμου, που αξιοποιεί τόσο το σχήμα όσο και την τοποθεσία της καμπύλης ανακλαστικότητας ανάμεσα σε δύο διαφορετικές καμπύλες, αυτής του εκατό τοις εκατό αερίου κι εκείνη του εκατό τοις εκατό νερού.

Η μέτρηση της ανάκλασης γίνεται με τη χρήση μονοχρωματικού υπέρυθρου φωτός που εκλύεται από μία διοδική λυχνία εκπομπής φωτός (Εικόνα 1.7). Το φως αυτό πολώνεται περνώντας από έναν κυλινδρικό πρισματικό φακό, και κατευθύνεται προς τη γραμμή ροής μέσω ενός παραθύρου ζαφειριού. Μία διάταξη έξι ανιχνευτών μετρά στη συνέχεια την ένταση του φωτός που ανακλάται από τη διεπιφάνεια ρευστού-ζαφειριού σε διακριτές γωνίες, συγκεκριμένα από γωνία μικρότερη της γωνίας Brewster του αερίου μέχρι και λίγο κάτω από την κρίσιμη γωνία του νερού. Οι βαθμονομήσεις που πραγματοποιούνται γίνονται με ξηρό αέρα και με νερό στη γραμμή ροής [7].



**Εικόνα 1.7 :** Διάταξη ανιχνευτή αερίου φάσεως. Διοδική λυχνία εκπομπής φωτός (LED), κυλινδρικός πρισματικός φακός (Lens and polarizer), πρίσμα ζαφειριού (Sapphire prism), ανιχνευτές φωτός (Detector array), ρευστό του σχηματισμού (Formation fluids), γραμμή ροής (Flowline) [7].



## 1.8 ΠΛΕΟΝΕΚΤΗΜΑΤΑ ΤΗΣ ΜΕΘΟΔΟΥ.

Ήδη η ποσοτική μέτρηση της σύστασης του ρευστού, σε πραγματικό χρόνο ροής του από τον ταμιευτήρα στον δειγματολήπτη, και σε πραγματικές συνθήκες ταμιευτήρα, είναι μια δυνατότητα που οδηγεί σε μία άντληση πληροφορίας εξαιρετικά σημαντικής όσον αφορά το ρευστό του ταμιευτήρα. Σημαντικό πλεονέκτημα της μεθόδου είναι η άμεση γνώση της ποσοτικής σύστασης, σε σύγκριση με την χρονοβόρα διαδικασία της αποστολής του δείγματος για εργαστηριακή ανάλυση.

Άλλο πλεονέκτημα της μεθόδου είναι πως μετά από χρόνια διεξαγωγής των τεστ του OFA, υπάρχουν πλέον διαθέσιμες μετρήσεις του για έναν πολύ μεγάλο αριθμό πετρελαϊκών ρευστών, τέτοιου ώστε να θεωρείται το πλήθος αυτό αντιπροσωπευτικό του συνόλου των πετρελαίων. Αυτό καθιστά την τράπεζα αυτή των δεδομένων σε ένα θησαυρό πληροφορίας προς ανάλυση και επεξεργασία, που δύναται να χρησιμοποιηθεί στην ανάπτυξη μαθηματικών μοντέλων που στόχο έχουν την καλύτερη κατανόηση θεμάτων της Μηχανικής Πετρελαίων αλλά και την υποβοήθηση των υπολογισμών της.

Τέλος, με την ενσωμάτωση του OFA στο MDT, και τη βοήθεια μαθηματικών μοντέλων, επιτυγχάνεται για πρώτη φορά σε συνθήκες open hole ο υπολογισμός του λόγου αερίου-πετρελαίου GOR, της πυκνότητας του πετρελαίου, της κατά μάζα σύστασης του ρευστού σε διοξείδιο του άνθρακα ( $CO_2$ ), του pH του περιεχόμενου νερού, της εξακρίβωσης παρουσίας αερίου φάσης μέσα στη ροή, της πίεσης και θερμοκρασίας της γραμμής ροής, καθώς και η εξακρίβωση της παρουσίας λάσπης στο ρευστό έτσι ώστε να αποφευχθεί η λήψη μολυσμένου δείγματος κατά τη δειγματοληψία. [7], [13] . Παράδειγμα μιας τέτοιας προσπάθειας είναι η εφαρμογή των τεχνητών νευρωνικών δικτύων σε μετρήσεις του OFA (Application of Artificial Neural Networks to Downhole Fluid Analysis) , κατά την οποία μετρήσεις του OFA αποτέλεσαν τις εισόδους νευρωνικού δικτύου που εκπαιδεύτηκε για να προβλέπει τον λόγο αερίου/πετρελαίου GOR [15].

## 1.9 ΜΕΙΟΝΕΚΤΗΜΑΤΑ ΤΗΣ ΜΕΘΟΔΟΥ.

Κατά την διεξαγωγή της μέτρησης του OFA, ενδέχεται να παρουσιαστούν κάποιες επιπλοκές. Κατά την διάρκεια της δειγματοληψίας του ρευστού του σχηματισμού ίσως τύχει να εμπεριέχονται στο ρευστό του ταμιευτήρα σωματίδια της γεωτρητικής λάσπης, ή μέρος του ρευστού να έχει υποστεί γαλακτωματοποίηση, φαινόμενο κατά το οποίο το ρευστό μετατρέπεται σε ένα κολλώδες μίγμα, επειδή σταγονίδια νερού διασπείρονται μέσα στο πετρέλαιο. Αποτέλεσμα και των δύο προαναφερθεισών διαδικασιών, είναι σκέδαση του φωτός που συνεπάγεται μία μείωση της έντασης του διαδιδόμενου φωτός, κι αν δε τύχει η συνεισφορά της σκέδασης στην οπτική πυκνότητα να είναι μεγάλη, τότε οι κορυφές των καμπυλών του φάσματος απορρόφησης του νερού και του πετρελαίου γίνονται δυσδιάκριτες και είναι δύσκολο να ερμηνευθούν σωστά. Σε μία τέτοια περίπτωση το ρευστό ερμηνεύεται ως γεωτρητική ιλύς [7].

Ένα άλλο μειονέκτημα της μεθόδου ως έχει, είναι πως προβλέπει την επί τοις εκατό κατά μάζα σύσταση κατά ομάδες συστατικών κι όχι ανά καθαρό συστατικό, με εξαίρεση βέβαια τα  $CO_2$ ,  $C_1$  και  $C_2$ , γεγονός που αφήνει περιθώριο μεγάλης αβεβαιότητας για την πραγματική κατά μάζα σύσταση των υπολοίπων συστατικών. Επίσης, η επί τοις εκατό κατά μάζα σύσταση που προβλέπεται δεν είναι άμεσα αξιοποιήσιμη στην Μηχανική Ταμιευτήρων καθώς είναι η γραμμομοριακή σύσταση αυτή που εμπλέκεται στους διάφορους υπολογισμούς.

## 1.10 ΣΤΟΧΟΣ ΚΑΙ ΣΚΟΠΟΣ ΤΗΣ ΕΡΓΑΣΙΑΣ.

Στόχος της εργασίας αυτής, είναι η μέσω της ανάπτυξης και εκπαίδευσης μοντέλου νευρωνικών δικτύων, αξιοποίηση των μετρήσεων του OFA για την ασφαλή πρόβλεψη περαιτέρω δεδομένων σχετικά με τη σύσταση του ρευστού. Πιο συγκεκριμένα αφορά τον προσδιορισμό της επί τοις εκατό κατά μάζα σύστασης του ρευστού σε  $C_3$ ,  $C_4$ ,  $C_5$ ,  $C_{6-8}$ ,  $C_{9-11}$  και  $C_{12+}$ , σε συνδυασμό με τα δεδομένα πεδίου όπως η σχετική πυκνότητα της αέριας

φάσης (Sg), ο δείκτης πυκνότητας πετρελαίου (API) και ο λόγος αερίου/πετρελαίου (GOR) κάθε ρευστού αντιστοίχως. Το μοντέλο αυτό δίνει την δυνατότητα για αναλυτική πρόβλεψη της επί τοις εκατό γραμμομοριακής σύστασης των ρευστών αυτών σε  $C_3$ ,  $C_4$ ,  $C_5$ ,  $C_{6-8}$ ,  $C_{9-11}$  και  $C_{12+}$ , αναβαθμίζοντας την αξία των μετρήσεων του OFA, καθώς οι εταιρείες παραγωγής υδρογονανθράκων μπορούν να έχουν πληροφορία για την γραμμομοριακή ανάλυση του ρευστού του ταμιευτήρα σε ένα τόσο πρώιμο ερευνητικό στάδιο, άρα και πρόσβαση στη σύσταση που χρησιμοποιείται από τους προσομοιωτές ταμιευτήρων.

Το μοντέλο αναπτύχθηκε στα εξής στάδια :

- Εκπαίδευση πρώτου νευρωνικού δικτύου για την πρόβλεψη του ποσοστού του  $C_3$  στο κλάσμα  $C_{3-5}$ , και υπολογισμός της επί τοις εκατό κατά μάζα περιεκτικότητας του ρευστού σε  $C_3$  από τον πολλαπλασιασμό του ποσοστού αυτού με το κλάσμα  $C_{3-5}$ .
- Εκπαίδευση δεύτερου νευρωνικού δικτύου για την πρόβλεψη του ποσοστού του  $C_4$  στο κλάσμα  $C_{4-5}$ , και υπολογισμός της επί τοις εκατό κατά μάζα περιεκτικότητας του ρευστού σε  $C_4$  από τον πολλαπλασιασμό του ποσοστού αυτού με το κλάσμα  $C_{4-5}$ .
- Υπολογισμός της επί τοις εκατό κατά μάζα περιεκτικότητας του ρευστού σε  $C_5$  με την αφαίρεση των  $C_3$  και  $C_4$  από το κλάσμα  $C_{3-5}$ .
- Εκπαίδευση τρίτου νευρωνικού δικτύου για την πρόβλεψη του ποσοστού του  $C_{6-8}$  στο κλάσμα  $C_{6+}$ , και υπολογισμός της επί τοις εκατό κατά μάζα περιεκτικότητας του ρευστού σε  $C_{6-8}$  από τον πολλαπλασιασμό του ποσοστού αυτού με το κλάσμα  $C_{6+}$ .
- Εκπαίδευση τέταρτου νευρωνικού δικτύου για την πρόβλεψη του ποσοστού του  $C_{9-11}$  στο κλάσμα  $C_{9+}$ , και υπολογισμός της επί τοις εκατό κατά μάζα περιεκτικότητας του ρευστού σε  $C_{9-11}$  από τον πολλαπλασιασμό του ποσοστού αυτού με το κλάσμα  $C_{9+}$ .

- Υπολογισμός της επί τοις εκατό κατά μάζα περιεκτικότητας του ρευστού σε  $C_{12^+}$  με την αφαίρεση των  $C_{6-8}$  και  $C_{9-11}$  από το κλάσμα  $C_{6^+}$ .
- Εκπαίδευση πέμπτου νευρωνικού δικτύου για την πρόβλεψη του μέσου μοριακού βάρους του  $C_{6^+}$ .
- Εφαρμογή του ισοζυγίου μάζας στο κλάσμα  $C_{6^+}$  με τη χρήση των μοριακών βαρών κατά Katz, με στόχο τον υπολογισμό του μέσου μοριακού βάρους του κλάσματος  $C_{12^+}$ .
- Μετατροπή της επί τοις εκατό κατά μάζα περιεκτικότητας σε επί τοις εκατό κατά mole περιεκτικότητα των ρευστών σε  $CO_2$ ,  $C_1$ ,  $C_2$ ,  $C_3$ ,  $C_4$ ,  $C_5$ ,  $C_{6-8}$ ,  $C_{9-11}$  και  $C_{12^+}$ .

Η ανάπτυξη του μοντέλου αυτού επιτυγχάνεται με την εφαρμογή της παλινδρομικής ανάλυσης, μέσω της εκπαίδευσης πέντε τεχνητών νευρωνικών δικτύων των οποίων οι λεπτομέρειες ανάπτυξης και εκπαίδευσης αναλύονται εκτενώς στο επόμενο κεφάλαιο. Τελικός σκοπός της εργασίας αυτής, είναι η μεγιστοποίηση του πλεονεκτήματος της μεθόδου του OFA, δηλαδή της άμεσης πρόσβασης σε δεδομένα σύστασης του ρευστού του ταμειυτήρα κατά το πρώιμο στάδιο της έρευνας, με την παροχή αξιόπιστης μοριακής ανάλυσης που μπορεί να χρησιμοποιηθεί κτευθείαν στους υπολογισμούς για την αξιολόγηση του κοιτάσματος.

## ΚΕΦΑΛΑΙΟ 2ο

Παλινδρομική ανάλυση και Τεχνητά Νευρωνικά Δίκτυα.

### 2.0 ΕΙΣΑΓΩΓΗ.

Η μοντελοποίηση, ως βασικό εργαλείο ανάπτυξης μαθηματικών προτύπων, συνιστά αναπόσπαστο και απαραίτητο τμήμα του πεδίου της Μηχανικής για δύο κυρίως λόγους. Ο πρώτος είναι η δυνατότητα πρόβλεψης της τιμής κάποιων μεταβλητών όταν για αυτές δεν διατίθεται μέτρηση κι ο δεύτερος είναι η συγκέντρωση μεγάλου όγκου πληροφορίας σε ένα σύνολο παραμέτρων. Στο παρόν κεφάλαιο αναπτύσσεται η περιγραφή των Τεχνητών Νευρωνικών Δικτύων (ΤΝΔ) ως εργαλείων μοντελοποίησης, στα πλαίσια της επίλυσης αντίστροφων προβλημάτων με τη μέθοδο της παλινδρόμησης.

### 2.1 ΑΝΤΙΣΤΡΟΦΑ ΠΡΟΒΛΗΜΑΤΑ.

Τα αντίστροφα προβλήματα εμφανίζονται συχνά στο πεδίο της Μηχανικής, και η επίλυσή τους αποτελεί για τους μηχανικούς μια πολύ σημαντική ερευνητική περιοχή. Οι κλασσικές αριθμητικές μέθοδοι, σε συνδυασμό με μια ραγδαία εξέλιξη των ηλεκτρονικών υπολογιστικών συστημάτων, έχουν αναβαθμίσει σημαντικά την προσέγγιση πληθώρας πρακτικών προβλημάτων της Μηχανικής.

Δοθείσας της πλήρους περιγραφής ενός φυσικού συστήματος μέσα από ένα μαθηματικό πρότυπο, ο υπολογισμός της τιμής κάποιου από τα φυσικά μεγέθη του συστήματος αυτού, θα προκύψει με την απλή αντικατάσταση των τιμών των υπολοίπων μεγεθών στο μαθηματικό αυτό πρότυπο. Αυτός ο τρόπος πρόβλεψης της τιμής ενός μεγέθους ονομάζεται ευθύ πρόβλημα.

Κατά το αντίστροφο πρόβλημα αυτό που συμβαίνει είναι, πως έχοντας τις πειραματικές μετρήσεις ενός φυσικού μεγέθους, και γνωρίζοντας κάποιες ή όλες τις παραμέτρους που σχετίζονται με το φυσικό σύστημα στο οποίο ενυπάρχει το μέγεθος αυτό, επιδιώκεται η προσέγγιση της εξίσωσης που περιγράφει καλύτερα τη συσχέτιση παραμέτρων και μεγέθους. Κατά αυτή την προσέγγιση, ενώ ένα ευθύ πρόβλημα έχει μοναδική λύση, ένα αντίστροφο πρόβλημα έχει πάντα πάνω από μία λύσεις [16].

Προς εξέταση των παραπάνω, δίνεται ως παράδειγμα ένα απλούστατο πρόβλημα από τη Μηχανική Πετρελαίου, όπου το μέσο μοριακό βάρος ενός κλάσματος πετρελαίου, περιγράφεται από την παρακάτω εξίσωση του ισοζυγίου μάζας :

$$MW_n = \sum_{i=1}^n y_i * MW_i \quad (2.1)$$

όπου :

$MW_n$  = το μέσο μοριακό βάρος του κλάσματος  $n$  που απαρτίζεται από τα συστατικά  $i$ .

$y_i$  = το γραμμομοριακό κλάσμα του συστατικού  $i$  στο κλάσμα  $n$ .

$MW_i$  = το μοριακό βάρος της ένωσης  $i$ .

Κατά το ευθύ πρόβλημα του υπολογισμού του μέσου μοριακού βάρους του κλάσματος  $n$ , αυτό που πρέπει να συμβεί είναι η αντικατάσταση των τιμών των γραμμομοριακών κλασμάτων ( $y_i$ ) στην εξίσωση (2.1) δεδομένου πως τα μοριακά βάρη των ενώσεων ( $MW_i$ ) είναι γνωστά για κάθε συστατικό. Η αντίστροφη προσέγγιση του παραπάνω προβλήματος θα ίσχυε εάν ήταν γνωστό, το μέσο μοριακό βάρος του κλάσματος  $n$  ( $MW_n$ ), και το ζητούμενο ήταν το γραμμομοριακό κλάσμα του κάθε συστατικού ( $y_i$ ). Κατά την

προσέγγιση αυτή οι λύσεις του προβλήματος είναι πάνω από μία, καθώς περισσότεροι από ένας συνδυασμοί των παραμέτρων  $y_i$  μπορούν να δώσουν την ίδια τιμή στο μετρούμενο μέγεθος ( $MW_n$ ).

Αυτό το χαρακτηριστικό της πολλαπλότητας των λύσεων ενός αντίστροφου προβλήματος, καθιστά αναγκαία σε κάθε περίπτωση, τόσο την αξιοποίηση όλων των δυνατών πληροφοριών που υπάρχουν για ένα φυσικό σύστημα, όσο και την προσοχή σε ενδεχόμενη αβεβαιότητα επί των δεδομένων.

## 2.2 ΠΑΛΙΝΔΡΟΜΙΚΗ ΑΝΑΛΥΣΗ.

Η Παλινδρομική ανάλυση (Regression Analysis) είναι μια πολυμεταβλητή στατιστική τεχνική, και ως ειδική μορφή προσέγγισης συναρτήσεων αποτελεί ισχυρό εργαλείο στην επίλυση των αντίστροφων προβλημάτων. Κατά την Παλινδρομική ανάλυση, το ζητούμενο είναι η εύρεση ενός μαθηματικού προτύπου, στο οποίο θα προσαρμόζονται πειραματικά δεδομένα ενός ή και περισσοτέρων μεγεθών, με τον βέλτιστο δυνατό τρόπο. Στην περίπτωση όπου η ανεξάρτητη μεταβλητή του προς εξέταση φυσικού συστήματος είναι μία, η παλινδρομική ανάλυση χαρακτηρίζεται ως απλή, ενώ ως πολλαπλή χαρακτηρίζεται όταν οι ανεξάρτητες μεταβλητές είναι περισσότερες από μία. Όταν η σχέση των ανεξαρτήτων μεταβλητών με τις εξαρτημένες είναι γραμμική, η παλινδρόμηση ονομάζεται γραμμική, ενώ σε οποιαδήποτε άλλη περίπτωση η παλινδρόμηση ονομάζεται μη γραμμική [16][17].

Η μέθοδος που εφαρμόζεται στην παλινδρομική ανάλυση συνήθως είναι αυτή των ελαχίστων τετραγώνων που προτάθηκε από τον Gauss το 1795. Η μέθοδος αυτή εφαρμόζεται με στόχο την δημιουργία της γραφικής παράστασης που περιγράφει τη συσχέτιση των παραμέτρων ενός φυσικού συστήματος, όταν είναι γνωστές οι πειραματικές τιμές κάποιων μεγεθών του συστήματος αυτού κι όχι η μεταξύ τους σχέση. Σκοπός της μεθόδου είναι ο προσδιορισμός των παραμέτρων μιας σχέσης μεταξύ των μεταβλητών, ώστε οι προβλέψεις να ταιριάζουν καλύτερα με τα πειραματικά δεδομένα (στην παλινδρομική ανάλυση ο χρήστης προδιαγράφει τη σχέση π.χ. ευθεία, επίπεδο πολυώνυμο 3<sup>ου</sup> βαθμού κλπ). Ο προσδιορισμός αυτός επιτυγχάνεται

μέσω του υπολογισμού των σταθερών συντελεστών της επιλεγείσας σχέσης, που προκύπτει από την ελαχιστοποίηση του αθροίσματος των τετραγωνικών σφαλμάτων των πειραματικών δεδομένων ως προς της προβλέψεις της ζητούμενης σχέσης για τα δεδομένα αυτά, την ελαχιστοποίηση δηλαδή της ποσότητας :

$$\sum_{i=1}^M (y_i - Y_i)^2 \quad (2.2)$$

Όπου :

$y_i$  = η πειραματική μέτρηση του δείγματος  $i$ .

$Y_i$  = η πρόβλεψη της ζητούμενης εξίσωσης για την τιμή του δείγματος  $i$ .

$M$  = ο αριθμός των δειγμάτων.

Η κατασκευή μοντέλων, δηλαδή μαθηματικών προτύπων, μέσω της παλινδρόμησης βρίσκει εφαρμογή σε πεδία όπου εξετάζεται η ανάλυση πλήθους στατιστικών δεδομένων. Σε ένα γενικό πλαίσιο η φιλοσοφία της διαδικασίας είναι η εξής :

Έστω ένα σετ δεδομένων μέσα στο οποίο πρέπει να εξεταστεί η ύπαρξη ή μη σχέσεων μεταξύ των μεταβλητών. Κάποια από αυτές τις μεταβλητές ορίζεται ως εξαρτημένη μεταβλητή ή απόκριση. Αυτομάτως οι υπόλοιπες καθίστανται ως οι ανεξάρτητες μεταβλητές ή αλλιώς παλινδρομητές, και σκοπός τους είναι η πρόβλεψη ή η επεξήγηση της στατιστικής συμπεριφοράς της εξαρτημένης μεταβλητής. Τέλος μέσω της παλινδρομικής ανάλυσης ορίζεται η σχέση της απόκρισης με τους παλινδρομητές. Εάν η σχέση αυτή προκύψει γραμμική τότε και το μοντέλο παλινδρόμησης θα καλείται γραμμικό, ενώ μη γραμμικό θα χαρακτηριστεί το μοντέλο σε αντίθετη περίπτωση [17] [18].

Όπως αναπτύσσεται στις παρακάτω ενότητες η παλινδρόμηση βρίσκει εφαρμογή και σε τομείς της τεχνητής νοημοσύνης μέσα από εργασίες μάθησης των τεχνητών νευρωνικών δικτύων.



## **2.3 ΤΕΧΝΗΤΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ.**

Τα τεχνητά νευρωνικά δίκτυα ή ΤΝΔ (ANN : Artificial Neural Networks) αποτελούν κλάδο της υπολογιστικής επιστήμης. Προέκυψαν έπειτα από έρευνες, που στόχο είχαν την κατανόηση της επεξεργασίας πολύπλοκων δεδομένων που εκτελεί ο εγκέφαλος. Αυτή η ανάγκη προέκυψε καθώς, ενώ οι ηλεκτρονικοί υπολογιστές μπορούσαν να εκτελούν με ταχύτητα και ακρίβεια μαθηματικές πράξεις, ή μια σειρά εντολών μέσω του προγραμματισμού, αδυνατούσαν να επεξεργάζονται πολύπλοκα δεδομένα που εμπεριείχαν θόρυβο ή μικρή ακρίβεια, ή να αναγνωρίζουν συσχετισμούς ανάμεσα σε μεγάλους όγκους φαινομενικά ασύνδετων δεδομένων. Αυτό το στοιχείο που απουσίαζε από τα συμβατικά υπολογιστικά συστήματα, βρισκόταν μέσα στις δομικές μονάδες του εγκεφάλου, τους νευρώνες, των οποίων οι διασυνδέσεις μεταξύ τους αλλά και με το περιβάλλον της πληροφορίας, σε συνδυασμό με την εμπειρία, αποδίδουν στον εγκέφαλο την ικανότητα της μάθησης και της εκτέλεσης εργασιών, όπως η αναγνώριση προτύπων μέσα σε ένα πλήθος πληροφοριών, η ανοχή σε λάθη επί των δεδομένων καθώς και η παράλληλη επεξεργασία μεγάλου όγκου πληροφοριών. Η διαδικασία με την οποία καταφέρνει ο εγκέφαλος, μέσα από την συσχέτιση και τον συνδυασμό των ερεθισμάτων ή αλλιώς την εμπειρία, να παράγει μοντέλα ή πρότυπα καλείται μάθηση [20]. Για να επιτευχθεί λοιπόν η ανάπτυξη ενός υπολογιστικού συστήματος, το οποίο θα μπορούσε να προσεγγίζει τη γνώση μέσα από την εμπειρία, αυτό που έπρεπε να προσομοιωθεί αλγοριθμικά, ήταν αυτές οι διασυνδέσεις των νευρώνων καθώς κι ο τρόπος που αυτές αλληλεπιδρούν με σκοπό την επίτευξη ενός μαθησιακού στόχου.

### **2.3.1 ΙΣΤΟΡΙΚΗ ΑΝΑΔΡΟΜΗ.**

Η εξέλιξη των Ηλεκτρονικών Υπολογιστών και η Μηχανική Λογισμικού, φαίνεται ουσιαστικά πως καθυστέρησαν την υλοποίηση της ιδέας των τεχνητών νευρωνικών δικτύων. Η πρώτη ιδέα μοντελοποίησης της λειτουργίας των νευρώνων του εγκεφάλου, εμφανίζεται το 1943 από τον νευροφυσιολόγο Warren McCulloch και τον μαθηματικό Walter Pitts, οι οποίοι κατασκευάζουν το πρώτο απλό 'αναλογικό' νευρωνικό δίκτυο,

χρησιμοποιώντας ηλεκτρικά κυκλώματα. Με τη δουλειά τους αυτή καταφέρνουν να συστήσουν τα νευρωνικά δίκτυα ως υπολογιστικές μηχανές. Ακολούθησε πλήθος εργασιών μέχρι να τεθούν τα θεμέλια των τεχνητών νευρωνικών δικτύων όπως τα γνωρίζουμε σήμερα. Στην εργασία του Hebb (1949) γίνεται η παρατήρηση πως οι διασυνδέσεις των νευρώνων ενισχύονταν κάθε φορά που χρησιμοποιούνταν, ενώ ο Rosenblatt (1958) παρουσιάζει το πρώτο μοντέλο επιβλεπόμενης εκπαίδευσης, το λεγόμενο perceptron. Το 1969 οι Minsky και Papert αποδεικνύουν μαθηματικά πως τα τεχνητά νευρωνικά δίκτυα ενός επιπέδου αδυνατούν να επιλύσουν κάποια προβλήματα, ενώ το 1986 οι McClelland και Rumelhart προτείνουν την μέθοδο οπισθοδιάδοσης του σφάλματος ως αλγόριθμο μάθησης [18][19] .

Η ευρεία χρήση των νευρωνικών δικτύων ξεκίνησε στα μέσα του 1980, και βρίσκει σήμερα εφαρμογές σε πλήθος επιστημονικών πεδίων όπως η Φυσική, η Γενετική, η Μηχανική, οι Τηλεπικοινωνίες, τα Οικονομικά, η Αεροδιαστημική, η Ιατρική, η Ρομποτική κ.α.

### **2.3.2 ΤΙ ΕΙΝΑΙ ΤΟ ΤΕΧΝΗΤΟ ΝΕΥΡΩΝΙΚΟ ΔΙΚΤΥΟ.**

Ως τεχνητό νευρωνικό δίκτυο, ορίζεται μία αλγοριθμική αρχιτεκτονική διάταξη παράλληλης επεξεργασίας πληροφορίας, που μέσω της μηχανικής μάθησης επιτυγχάνει την αποθήκευση εμπειρικής γνώσης αλλά και την διάθεση αυτής προς χρήση [18]. Στην πράξη ένα νευρωνικό δίκτυο είναι μια συνάρτηση πολλών διαστάσεων που περιγράφει την συσχέτιση μεταξύ παραμέτρων. Οι διαστάσεις ενός δικτύου καθορίζονται από τον αριθμό των εισόδων του, όπου η κάθε είσοδος είναι με τη σειρά της μία παράμετρος. Ως έξοδοι του δικτύου αναφέρονται οι μαθησιακοί του στόχοι. Τα τεχνητά νευρωνικά δίκτυα ή νευρωνικά δίκτυα αποτελούν μια ιδιαίτερη προσέγγιση ανάπτυξης συστημάτων νοημοσύνης για την εκτέλεση υπολογισμών. Το βασικό δομικό στοιχείο ενός νευρωνικού δικτύου είναι ο νευρώνας (neuron) που αποτελεί τη βασική μονάδα επεξεργασίας της πληροφορίας, κι είναι οι διασυνδέσεις των νευρώνων ενός δικτύου με τις εισόδους του, αυτές που οδηγούν στην υλοποίηση των διαφόρων εργασιών μέσω της εκπαίδευσης. Οι ομοιότητες

που παρουσιάζει ένα νευρωνικό δίκτυο με τον ανθρώπινο εγκέφαλο είναι οι εξής :

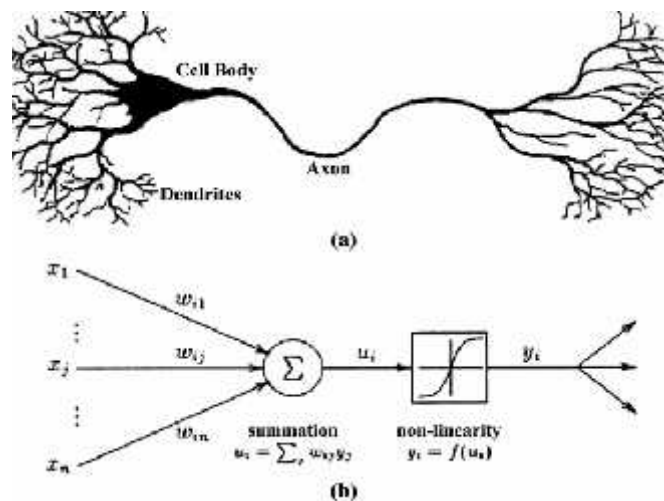
- Τόσο ο ανθρώπινος εγκέφαλος όσο κι ένα νευρωνικό δίκτυο επεξεργάζονται πληροφορία που υπάρχει στο περιβάλλον τους, και μέσω της διαδικασίας της μάθησης αποκτούν τη γνώση.
- Η ισχύς των διασυνδέσεων των νευρώνων ή αλλιώς συναπτικό βάρος, είναι ο χώρος αποθήκευσης της αποκτηθείσας γνώσης.

Ο τρόπος που εκτελεί την διαδικασία της μάθησης ένα νευρωνικό δίκτυο ονομάζεται αλγόριθμος μάθησης. Ο αλγόριθμος αυτός έχει ως σκοπό την τροποποίηση των συναπτικών βαρών, με τρόπο τέτοιο που να επιτυγχάνεται με το μικρότερο δυνατό σφάλμα ο μαθησιακός στόχος, να προσεγγίζεται δηλαδή με τον βέλτιστο δυνατό τρόπο η έξοδος του δικτύου. Ο τρόπος αυτός είναι ουσιαστικά ίδιου τύπου διαδικασία με αυτή που εκτελείται κατά την εφαρμογή της μεθόδου των ελαχίστων τετραγώνων για την εύρεση των σταθερών συντελεστών της ζητούμενης συνάρτησης. Στην περίπτωση ενός νευρωνικού δικτύου τα συναπτικά βάρη είναι οι σταθεροί συντελεστές της πολυδιάστατης ζητούμενης συνάρτησης. Κάποιες από τις εργασίες μάθησης που υλοποιεί ένα νευρωνικό δίκτυο είναι : η συσχέτιση προτύπων, η κατηγοριοποίηση, η προσέγγιση συνάρτησης, η αναγνώριση προτύπων, ο έλεγχος διεργασιών κλπ. Η εργασία μάθησης που εκτελείται από τα τεχνητά νευρωνικά δίκτυα στην παρούσα διπλωματική εργασία, είναι η προσέγγιση συνάρτησης.

### **2.3.3 ΜΟΝΤΕΛΟ ΕΝΟΣ ΤΕΧΝΗΤΟΥ ΝΕΥΡΩΝΑ.**

Ένας τεχνητός νευρώνας είναι το μαθηματικό μοντέλο ενός βιολογικού νευρώνα (Εικόνα 2.1). Ένας βιολογικός νευρώνας αποτελείται από το σώμα του κυττάρου, τους δενδρίτες και τον άξονα (Εικόνα 2.1.(α)). Ο δενδρίτης λαμβάνει ηλεκτρικό σήμα από τους άξονες άλλων νευρώνων, στον τεχνητό νευρώνα αυτά τα σήματα αντιπροσωπεύονται από αριθμητικές τιμές. Στις συνάψεις ανάμεσα στον δενδρίτη και τους άξονες άλλων νευρώνων το σήμα ποικίλει σε ισχύ, αυτό μοντελοποιείται από τον τεχνητό νευρώνα με την

δημιουργία των συναπτικών βαρών. Ο βιολογικός νευρώνας ενεργοποιεί το σήμα εξόδου μόνο όταν η συνολική ισχύς του σήματος εισόδου ξεπεράσει μια συγκεκριμένη τιμή. Αυτό το φαινόμενο μοντελοποιείται από τον τεχνητό νευρώνα από τον υπολογισμό του αθροίσματος των συναπτικών βαρών και την εφαρμογή μιας βηματικής συνάρτησης στο άθροισμα αυτό, που καθορίζει και την έξοδο, δηλαδή τον μαθησιακό στόχο. Όπως και στον βιολογικό νευρώνα αυτή η έξοδος τροφοδοτείται ως πληροφορία σε άλλους τεχνητούς νευρώνες [19].



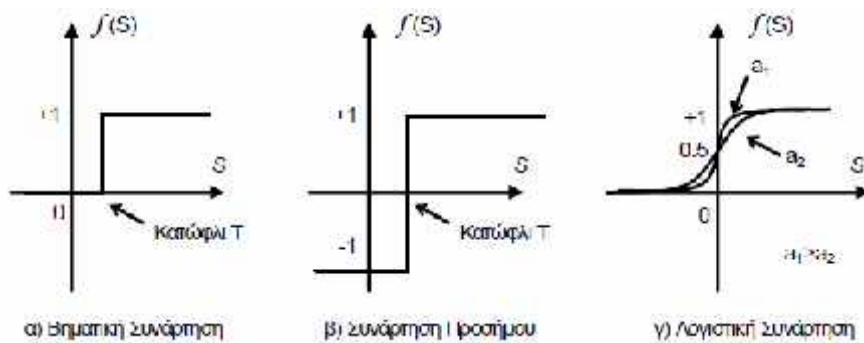
Εικόνα 2.1 : (α) αναπαράσταση βιολογικού νευρώνα, (β) αναπαράσταση τεχνητού νευρώνα που αποτελεί και το μοντέλο ενός νευρώνα [19].

Τα νευρωνικά δίκτυα, όπως αναπτύσσεται στην επόμενη ενότητα, παρουσιάζουν μια ποικιλία αρχιτεκτονικών δομών. Η απλούστερη δομική μονάδα από την οποία προκύπτουν όλες οι διαφορετικές αρχιτεκτονικές δομές, είναι το μοντέλο ενός νευρώνα (Εικόνα 2.1(b)) . Οι συνάψεις του νευρώνα, ο κόμβος άθροισης και μία συνάρτηση ενεργοποίησης είναι τα τρία θεμελιακά στοιχεία που συνθέτουν το μοντέλο ενός νευρώνα, και καθένα από αυτά εκτελεί τις παρακάτω λειτουργίες :

- Η κάθε σύναψη εμφανίζει το δικό της 'βάρος' το οποίο προκύπτει ως εξής : στην σύναψη για παράδειγμα της εισόδου  $x_1$  με τον νευρώνα  $i$ , το σήμα  $x_1$  της εισόδου πολλαπλασιάζεται με το συναπτικό βάρος  $w_{i1}$ .

- Στον κόμβο άθροισης αθροίζονται όλες οι επιμέρους είσοδοι πολλαπλασιαζόμενες με τα αντίστοιχα συναπτικά τους βάρη.
- Η συνάρτηση ενεργοποίησης περιορίζει το πλάτος του σήματος της εξόδου ενός νευρώνα σε κάποια πεπερασμένη τιμή.

Υπάρχουν διάφοροι τύποι συνάρτησης ενεργοποίησης που εφαρμόζονται σε ένα νευρωνικό δίκτυο, και η επιλογή τους σχετίζεται με την εργασία μάθησης που θα εκτελεστεί από το εκάστοτε δίκτυο (π.χ. κατηγοριοποίηση ή εύρεση συνάρτησης) καθώς και με τον αλγόριθμο μάθησης που θα εφαρμοστεί σε κάθε περίπτωση. Οι κύριοι τύποι των συναρτήσεων ενεργοποίησης είναι η βηματική συνάρτηση, η συνάρτηση προσήμου και η σιγμοειδής συνάρτηση, και οι γραφικές τους παραστάσεις αναπαρίστανται στην Εικόνα 2.2.

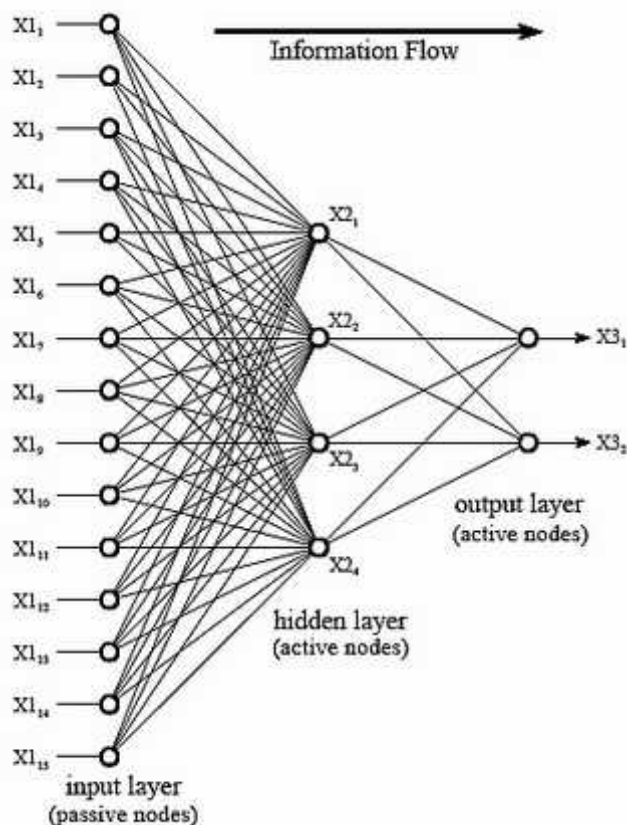


**Εικόνα 2.2 :** α) Γραφική παράσταση της βηματικής συνάρτησης ενεργοποίησης, β) Γραφική παράσταση της συνάρτησης κατωφλίου ή προσήμου, γ) Γραφική παράσταση της λογιστικής συνάρτησης που ανήκει στην κατηγορία των σιγμοειδών συναρτήσεων ενεργοποίησης, όπου  $a_1, a_2$  είναι οι διαφορετικές τιμές της παραμέτρου κλίσης [20].

Στην παρούσα διπλωματική εργασία για όλα τα νευρωνικά δίκτυα που αναπτύχθηκαν επιλέχθηκε η σιγμοειδής συνάρτηση ενεργοποίησης μιας και για την υλοποίηση του αλγορίθμου μάθησης που εφαρμόστηκε (Levenberg-Marquardt), απαιτείται η διαφορισιμότητα της συνάρτησης ενεργοποίησης.

### 2.3.4 ΔΙΚΤΥΑ ΠΡΟΣΘΙΑΣ ΤΡΟΦΟΔΟΤΗΣΗΣ.

Η ποικιλία των αρχιτεκτονικών δομών των νευρωνικών δικτύων είναι μεγάλη και προκύπτει από τον τρόπο με τον οποίο δομούνται και συνδέονται οι νευρώνες των δικτύων, αλλά και από τον αλγόριθμο μάθησης που εκτελούν τα δίκτυα κατά την εκπαίδευσή τους. Στην ενότητα αυτή περιγράφονται τα νευρωνικά δίκτυα πρόσθιας τροφοδότησης (Εικόνα 2.3) που εκτελούν την εκπαίδευσή τους με τον αλγόριθμο μάθησης Levenberg-Marquardt, καθώς είναι αυτά που εφαρμόστηκαν για την εκπόνηση της παρούσης διπλωματικής εργασίας.



**Εικόνα 2.3 :** Δίκτυο πρόσθιας τροφοδότησης πλήρους διασύνδεσης ενός κρυφού επιπέδου. Ο όρος πρόσθια τροφοδότηση επισημαίνει την ροή της πληροφορίας από το ένα επίπεδο στο επόμενο κι όχι αντίστροφα, ενώ το κρυφό επίπεδο αποτελεί το επίπεδο οργάνωσης των νευρώνων. Information flow (ροή πληροφορίας), input layer (επίπεδο εισόδου), hidden layer (κρυφό επίπεδο), output layer (επίπεδο εξόδου) [20].

Η εργασία μάθησης που κλήθηκαν να εκτελέσουν τα νευρωνικά δίκτυα στην παρούσα εργασία, ήταν αυτή της προσέγγισης συνάρτησης. Για την εκπαίδευσή τους τα δίκτυα χρειάζονται ένα σύνολο χαρακτηρισμένων παραδειγμάτων, το οποίο εκπροσωπείται από τις εισόδους και τις εξόδους του δικτύου. Έστω το σύνολο παραδειγμάτων  $\mathfrak{Z}$ :

$$\mathfrak{Z} = \{x_i, d_i\}_{i=1}^N \quad (2.3)$$

Όπου :

$x_i$  = η είσοδος του παραδείγματος  $i$ .

$d_i$  = η έξοδος του παραδείγματος  $i$ .

.

$N$  = ο αριθμός των δειγμάτων της εκπαίδευσης.

Τα  $x$  και  $d$  είναι διανύσματα και μπορούν να έχουν πάνω από μία συνιστώσες. Η κάθε συνιστώσα αντιπροσωπεύει και μία πληροφορία για κάθε δείγμα. Παραδείγματος χάριν έστω  $Z = \{X_i, D_i\}_{i=1}^M$  ο πληθυσμός  $M$  χαρακτηρισμένων παραδειγμάτων ρευστών υδρογονανθράκων, από τον οποίον πρέπει να διεξαχθεί η εξίσωση συσχέτισης των  $X_i$  και  $D_i$ . Το διάνυσμα εισόδου  $X = (X_1, X_2, X_3)$  του πληθυσμού αποτελείται από τρεις συνιστώσες οι οποίες αντιστοιχούν στις εξής πληροφορίες για κάθε παράδειγμα  $X_1$ : δείκτης πτητικότητας πετρελαίου GOR,  $X_2$ : δείκτης πυκνότητας πετρελαίου API,  $X_3$ : η σχετική πυκνότητα της αερίου φάσης του δείγματος  $S_g$ , και το διάνυσμα εξόδου  $D = (D_1)$  αποτελείται από μία συνιστώσα που αντιστοιχεί στην πληροφορία του μέσου μοριακού βάρους του κλάσματος  $C_{7+}$  (κλάσμα επτανίων και πάνω) για κάθε ρευστό. Αφού λοιπόν

για κάθε είσοδο  $X$  υπάρχει και η αντιστοίχιση της επιθυμητής απόκρισης  $D$ , η συνάρτηση που καλείται να προσεγγίσει το δίκτυο είναι η :

$$D = f(X) \quad (2.4)$$

Το ζητούμενο από την διαδικασία εκπαίδευσης του δικτύου είναι η προσέγγιση της άγνωστης συνάρτησης  $f(X)$ . Έστω  $F(X)$  η συνάρτηση που υλοποιεί το δίκτυο προσεγγίζοντας την άγνωστη  $f(X)$ , η εκπαίδευση του δικτύου έχει ως στόχο την βέλτιστη προσαρμογή της  $F(X)$  στην  $f(X)$  υπό την Ευκλείδεια έννοια, όπως υποδεικνύεται από τη σχέση :

$$\|F(X) - f(X)\| < \varepsilon \text{ για κάθε } X \quad (2.5)$$

Όπου  $\varepsilon$  είναι ένας πολύ μικρός θετικός αριθμός.

Δεδομένου πως για κάθε παράδειγμα εκπαίδευσης τα  $X$  και  $D$  είναι σταθερά η αναζήτηση της  $F(X)$  έγκειται στην εύρεση των σταθερών συντελεστών της  $F(X)$ , που είναι στην ουσία τα συναπτικά βάρη για κάθε συνιστώσα της εισόδου  $X$ . Τα συναπτικά βάρη, όπως αναφέρθηκε και στην προηγούμενη ενότητα, αντιπροσωπεύουν το 'βάρος' της σύναψης κάθε στοιχείου της εισόδου με τους νευρώνες του δικτύου. Κατά αυτή την έννοια η διαδικασία εκπαίδευσης που εκτελεί ένα νευρωνικό δίκτυο, είναι στην ουσία η μέσω ενός αλγορίθμου μάθησης, αναζήτηση των βέλτιστων συναπτικών βαρών, με στόχο την ελαχιστοποίηση της παρακάτω συνάρτησης κόστους :

$$E(w) = \frac{1}{2M} \sum_{i=1}^M [D(i) - F(X(i); w)]^2 \quad (2.6)$$

Όπου :

$w$  = το διάνυσμα των βαρών.

$E(w)$  = η συνάρτηση κόστους.



$\{X(i), D(i)\}_{i=1}^M$  = ο πληθυσμός εκπαίδευσης.

$F(X(i); w)$  = η συνάρτηση προσέγγισης που υλοποιεί το δίκτυο.

Ο αλγόριθμος μάθησης που εκτελέστηκε για την εκπαίδευση των νευρωνικών δικτύων της παρούσης εργασίας είναι ο αλγόριθμος Levenberg-Marquardt. Η μέθοδος των Levenberg-Marquardt, είναι μία τεχνική επίλυσης μη γραμμικών προβλημάτων ελαχίστων τετραγώνων, κι αποτελεί συνδυασμό δύο άλλων μεθόδων βελτιστοποίησης, της μεθόδου Newton και της μεθόδου βαθμωτής κατάβασης. Ο τρόπος υπολογισμού της βέλτιστης προσαρμογής των βαρών  $\Delta w$  που προτείνει η μέθοδος αυτή είναι :

$$\Delta w = [H + \lambda I]^{-1} g \quad (2.7)$$

Όπου :

$\Delta w$  = η βέλτιστη προσαρμογή των βαρών.

$H$  = ο πίνακας Hessian της συνάρτησης  $F(x)$ . Πιο συγκεκριμένα

$H(w) = \frac{\partial^2 E(w)}{\partial w^2}$  , όπου  $E(w)$  η συνάρτηση κόστους.

$g$  = το διάνυσμα κλίσεων. Πιο συγκεκριμένα  $g(w) = \frac{\partial E(w)}{\partial w}$  .

$I$  = ο μοναδιαίος πίνακας ίδιων διαστάσεων με τον  $H$ .

$\lambda$  = μία παράμετρος ομαλοποίησης που αναγκάζει το άθροισμα  $(H + \lambda I)$  να είναι ένας θετικά ορισμένος πίνακας.

Μετά το πέρας της εκπαίδευσης, το προϊόν του νευρωνικού δικτύου είναι η  $F(X)$ , η οποία μπορεί πλέον να χρησιμοποιηθεί ως ευθεία εξίσωση συσχέτισης των  $X$  και  $D$ . Να σημειωθεί στο σημείο αυτό, πως η ποιότητα της πληροφορίας που εμπεριέχεται στο σύνολο των χαρακτηρισμένων

παραδειγμάτων κάθε περίπτωσης, παίζει καθοριστικό ρόλο στην αξιολόγηση της εργασίας μάθησης ενός νευρωνικού δικτύου. Εάν για παράδειγμα δεν υφίσταται καμία εκ των προτέρων συσχέτιση των εισόδων  $X$  με τις εξόδους  $D$ , εάν δεν υπάρχει δηλαδή κάποια αρχική  $D=f(X)$ , τότε και το δίκτυο θα ανταποκριθεί ως προς την εκπαίδευσή του, αναδεικνύοντας την απουσία της συσχέτισης αυτής. Αυτό εφιστά την προσοχή ως προς την κριτική που πρέπει να ασκείται επί της τροφοδοτούμενης στο δίκτυο πληροφορίας.

### 2.3.5 ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΕΚΠΑΙΔΕΥΣΗΣ.

Η εκπαιδευτική διαδικασία των δικτύων που εκτελούν τον αλγόριθμο Levenberg-Marquardt ολοκληρώνεται σε κύκλους, όπου κάθε κύκλος αντιστοιχεί σε μία εφαρμογή του αλγορίθμου για όλον τον πληθυσμό εκπαίδευσης, για το εκάστοτε σύνολο χαρακτηρισμένων παραδειγμάτων δηλαδή (π.χ.  $\mathfrak{S} = \{x_i, d_i\}_{i=1}^N$ ). Τα παραδείγματα του συνόλου  $\mathfrak{S}$ , δεν διαδραματίζουν όλα τον ίδιο ρόλο κατά την εκπαίδευση. Πιο συγκεκριμένα τα παραδείγματα αυτά χωρίζονται σε τρεις διαφορετικές κατηγορίες :

- Δεδομένα εκπαίδευσης (training data). Αποτελούν το μεγαλύτερο ποσοστό του πληθυσμού εκπαίδευσης (συνήθως γύρω στο 70%) και είναι τα δεδομένα όπου το δίκτυο καλείται να 'μάθει'.
- Δεδομένα επιβεβαίωσης (validation data). Το ποσοστό τους επί του πληθυσμού εκπαίδευσης είναι συνήθως γύρω στο 15%. Στα δεδομένα αυτά το δίκτυο δεν εκπαιδεύεται, υπό την έννοια ότι δεν τα 'μαθαίνει', αλλά ο ρόλος τους είναι να περιορίζουν το δίκτυο να προβλέπει τις εξόδους του σε έναν αναμενόμενο κατά τα δεδομένα τους χώρο. Με τον τρόπο αυτό, τα δεδομένα επιβεβαίωσης περιορίζουν τον βαθμό ελευθερίας του δικτύου έτσι ώστε σε δεδομένα επί των οποίων δεν έχει εκπαιδευτεί, οι προβλέψεις του να ακολουθούν όσο είναι δυνατό, τη γενική τάση του συνόλου των παραδειγμάτων.
- Δεδομένα ελέγχου (test data). Χρησιμοποιούνται για να αξιολογήσουν την έκβαση κάθε κύκλου εκπαίδευσης. Είναι σημεία επί των οποίων το δίκτυο δεν εκπαιδεύεται, αλλά σε κάθε κύκλο εκπαίδευσης, εφαρμόζει

την βέλτιστη  $F(X)$  για τις εισόδους των σημείων αυτών ( $X_{test}$ ), εκτιμώντας την απόκλιση των προβλέψεων του από τις εξόδους τους ( $D_{test}$ ), ως προς το μέσο τετραγωνικό σφάλμα.

Ο αριθμός των κύκλων εκπαίδευσης που θα εκτελέσει ένα νευρωνικό δίκτυο κάθε φορά, εξαρτάται από το πόσο 'σύντομα' θα εξαχθεί η βέλτιστη προσαρμογή των βαρών  $\Delta w$  κατά την Levenberg-Marquardt, και ως κανόνας τερματισμού των επαναλήψεων επιλέγεται το σημείο όπου μία προσαρμογή του διανύσματος των βαρών  $w$ , βελτιώνει την συνάρτηση κόστους  $E(w)$  κατά ένα πολύ μικρό ποσό, αμελητέας στατιστικής σπουδαιότητας [18].

### **2.3.6 ΕΦΑΡΜΟΓΕΣ ΤΕΧΝΗΤΩΝ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ ΣΤΗΝ ΒΙΟΜΗΧΑΝΙΑ ΠΕΤΡΕΛΑΙΟΥ.**

Τα τεχνητά νευρωνικά δίκτυα, εφαρμόζονται ευρέως στην βιομηχανία πετρελαίου από την δεκαετία του 1970. Το σημαντικό στοιχείο των εφαρμογών αυτών, είναι πως μέσω της μηχανικής μάθησης ενισχύεται η αξία ενός τεράστιου όγκου σύνθετων ιστορικών δεδομένων αερίων και πετρελαίων, τα οποία αναλύονται εκ νέου παράγοντας καινούρια πληροφορία. Με τον τρόπο αυτό επιτυγχάνεται η βελτιστοποίηση της παραγωγής, η ποσοτικοποίηση της αβεβαιότητας επί των δεδομένων κατά την εκμετάλλευση ενός ταμειυτήρα, η βελτίωση των λογισμικών προσομοίωσης Μηχανικής Πετρελαίου, και ο έλεγχος διαδικασιών (process control) [21]. Ακολουθούν κάποια παραδείγματα εφαρμογών των τεχνητών νευρωνικών δικτύων από την βιβλιογραφία, με έμφαση σε εφαρμογές που σχετίζονται με τις ιδιότητες των ρευστών των υδρογονανθράκων :

1. Εφαρμογή των Νευρωνικών δικτύων με στόχο την πρόβλεψη της καθίζησης των ασφατενίων [22].

Ένα από τα σημαντικότερα φαινόμενα στην βιομηχανία πετρελαίου είναι η καθίζηση των βαρέων οργανικών αλύσεων όπως τα ασφατενία. Μία τέτοια καθίζηση επηρεάζει άμεσα την παραγωγική διαδικασία καθώς μπορεί να προκαλέσει αλλαγή στην διαβρεχτότητα και την διαπερατότητα ενός ταμειυτήρα, και να οδηγήσει κατά συνέπεια στην μείωση του ανακτήσιμου

όγκου των υδρογονανθράκων. Ο προσδιορισμός των συνθηκών αλλά και του βαθμού της καθίζησης των ασφατενίων, είναι για τους παραπάνω λόγους μία εξέχουσας σημασίας πληροφορία.

Στην εργασία αυτή σχεδιάστηκε και εφαρμόστηκε ένα Τεχνητό Νευρωνικό Δίκτυο, με στόχο την πρόβλεψη του ποσού καθίζησης των ασφατενίων, σε δεδομένες συνθήκες. Για την εκπαίδευση του δικτύου αυτού χρησιμοποιήθηκε πλήθος πειραματικών μετρήσεων του ποσού καθίζησης των ασφατενίων σε διάφορες θερμοκρασίες (293-343 K). Τα αποτελέσματα της εφαρμογής συγκρίθηκαν επίσης και με τις προβλέψεις κάποιων προτεινόμενων εξισώσεων, οι οποίες αποτύγχαναν να προβλέπουν ικανοποιητικά το ποσό της καθίζησης σε συγκεκριμένο εύρος θερμοκρασιών, σε αντίθεση με το δίκτυο, που προέβλεπε ικανοποιητικά σε όλο το εύρος των θερμοκρασιών.

## 2. Νέα μέθοδος πρόβλεψης της θερμοδυναμικής συμπεριφοράς ρευστών ταμιευτήρων [23].

Οι πιο διαδεδομένοι τρόποι υπολογισμού της θερμοδυναμικής συμπεριφοράς των ρευστών των ταμιευτήρων, είναι υπολογιστικά μοντέλα που βασίζονται είτε σε απλούς συσχετισμούς πειραματικών δεδομένων, γεγονός που συνεπάγεται μικρής ακρίβειας αποτελέσματα, είτε μοντέλα που βασίζονται στην καταστατική εξίσωση και η ακρίβεια των προβλέψεών τους εξαρτάται σε μεγάλο βαθμό από τη φύση των υπό μελέτη ρευστών, και από την επιλογή της εξίσωσης κάθε φορά καθώς και από την ποιότητα των πειραματικών δεδομένων.

Στην εργασία αυτή παρουσιάζεται μια νέα προσέγγιση για την πρόβλεψη της πλήρους θερμοδυναμικής συμπεριφοράς πετρελαίων και συμπυκνωμάτων. Η μέθοδος χρησιμοποιεί βασικές μετρήσεις, που μπορούν να εκτελεστούν γρήγορα είτε στο εργαστήριο είτε και στο εργοτάξιο, οι οποίες εισέρχονται ως είσοδοι σε ένα Τεχνητό νευρωνικό δίκτυο. Το δίκτυο εκπαιδεύτηκε επί μίας βάσης δεδομένων 650 ρευστών. Η εφαρμογή είχε ως αποτέλεσμα ένα μέσο σχετικό σφάλμα των εκτιμήσεων των θερμοδυναμικών ιδιοτήτων 0.5-2.5%, σε ρευστά για τα οποία δεν είχε εκπαιδευτεί. Ένα τέτοιο μέσο σχετικό σφάλμα είναι πολύ μικρό και μάλιστα καλύτερο από αυτό που θα προέκυπτε από την

εφαρμογή μοντέλου βασισμένου στην καταστατική εξίσωση. Το μοντέλο της εργασίας αυτής ονομάστηκε PVT Expert (© trademark of Schlumberger) και αξιολογείται ως ένα ισχυρό ακριβές και γρήγορο εργαλείο πρόβλεψης θερμοδυναμικών ιδιοτήτων πετρελαϊκών ρευστών.

### 3. Εργαλείο ελέγχου ποιότητας, επί των προβλέψεων προσομοιωτών θερμοδυναμικής συμπεριφοράς ρευστών υδρογονανθράκων [24].

Οι προσομοιωτές θερμοδυναμικής συμπεριφοράς ή αλλιώς προσομοιωτές PVT που διατίθενται προς χρήση στην βιομηχανία πετρελαίου, προσεγγίζουν τις προβλέψεις τους μέσω εξισώσεων που βασίζονται στην καταστατική εξίσωση. Οι προσομοιωτές αυτοί παράγουν προβλέψεις κυμαινόμενης ακρίβειας, η οποία εξαρτάται από τον τύπο του μοντέλου που εφαρμόζεται κάθε φορά, από τη φύση του προς εξέταση ρευστού, καθώς κι από τις επικρατούσες συνθήκες. Ένα σημαντικό μειονέκτημα όλων αυτών των προσομοιωτών, είναι πως δεν παρέχουν τη δυνατότητα εκτίμησης της αβεβαιότητας επί των προβλέψεών τους.

Οι προσομοιωτές PVT που έχουν αναπτυχθεί μέσω των νευρωνικών δικτύων παρέχουν προβλέψεις καλής ακρίβειας κυρίως γιατί έχουν εκπαιδευτεί μέσω της εμπειρίας. Η ακρίβεια των προβλέψεων που δίνουν εξαρτάται από το κατά πόσο οι συντεταγμένες του προς εξέταση ρευστού, βρίσκονται σε περιοχή μεγάλης πυκνότητας, άρα και αρκετής πληροφορίας, ως προς το πλήθος των δειγμάτων με τα οποία εκπαιδεύτηκε το δίκτυο. Έτσι, για ρευστά που δεν έχουν αποτελέσει δείγματα εκπαίδευσης και οι συντεταγμένες τους βρίσκονται έξω από τις περιοχές όπου το δίκτυο έχει εκπαιδευτεί, αναμένονται προβλέψεις μικρής ακρίβειας.

Στην εργασία αυτή παρουσιάζεται ένα εργαλείο αξιολόγησης επί των προβλέψεων ενός μοντέλου PVT βασισμένου στα τεχνητά νευρωνικά δίκτυα, του PVT Expert (© trademark of Schlumberger), του οποίου η εφαρμογή αναφέρθηκε στην προηγούμενη ενότητα. Το εργαλείο αυτό εκτιμά την αβεβαιότητα της πρόβλεψης που θα κάνει ο PVT Expert για ένα καινούριο ρευστό, αξιολογώντας τη θέση του ρευστού ως προς την περιοχή όπου το μοντέλο έχει εκπαιδευτεί, κι άρα γενικεύει καλά. Το εργαλείο αξιολόγησης

εφαρμόστηκε επιτυχώς σε ένα μεγάλο αριθμό δειγμάτων επί των οποίων το μοντέλο δεν είχε εκπαιδευτεί, και με τον τρόπο αυτό το μοντέλο (PVT Expert) αναβάθμισε την απόδοσή του παρέχοντας αξιόπιστη κριτική επί των προβλέψεών του.

#### 4. Εφαρμογή των τεχνητών νευρωνικών δικτύων για την πρόβλεψη της θερμοκρασίας σχηματισμού κεριού παραφίνης [25].

Ο σχηματισμός κεριού παραφίνης αποτελεί προβληματικό φαινόμενο όσον αφορά την διαδικασία παραγωγής πετρελαίου, ιδιαίτερα στις περιπτώσεις των παράκτιων εκμεταλλεύσεων μεγάλου βάθους. Το φαινόμενο αυτό συμβαίνει όταν η θερμοκρασία του πετρελαίου πέσει κάτω από μια συγκεκριμένη τιμή (σημείο θόλωσης), καθώς αυτό κατευθύνεται από τον ταμιευτήρα στην επιφάνεια. Έτσι, καθώς οι παραφινικές ενώσεις εναποτίθενται με τη μορφή κεριού στις σωληνώσεις της παραγωγής δυσχεραίνεται η παραγωγική διαδικασία είτε γιατί απαιτείται αύξηση της ενέργειας άντλησης ώστε να διατηρείται η ομαλή ροή του πετρελαίου, είτε γιατί η εναπόθεση κεριού είναι τέτοια που αποφράσσονται τελείως οι γραμμές ροής με αποτέλεσμα την καθυστέρηση της παραγωγής ή ακόμα και την εγκατάλειψη της γεώτρησης παραγωγής. Μια αξιόπιστη πρόβλεψη των συνθηκών σχηματισμού των κεριών είναι αναγκαία ώστε να αποφευχθεί ή έστω να μετριαστεί μια επιβλαβής ως προς την διαδικασία παραγωγής, εναπόθεση κεριού.

Τα περισσότερα μοντέλα πρόβλεψης των συνθηκών σχηματισμού των κεριών διαφέρουν σημαντικά από πειραματικές μετρήσεις, καθώς απαιτούνται για την εφαρμογή τους διάφορες απλουστευτικές υποθέσεις. Σκοπός της εργασίας αυτής είναι η εκπαίδευση ενός Τεχνητού Νευρωνικού Δικτύου, που θα προβλέπει την θερμοκρασία καθίζησης των παραφινικών ενώσεων για ένα πλήθος διαφορετικών μιγμάτων υδρογονανθράκων. Το δίκτυο εκπαιδεύτηκε χρησιμοποιώντας πειραματικές μετρήσεις της θερμοκρασίας σχηματισμού των κεριών παραφίνης, δώδεκα διαφορετικών ρευστών υδρογονανθράκων, σε διάφορους συνδυασμούς θερμοδυναμικών χαρακτηριστικών. Τα αποτελέσματα της εργασίας έδειξαν πως το δίκτυο προέβλεπε την

θερμοκρασία σχηματισμού του κεριού με μεγαλύτερη ακρίβεια από τα προϋπάρχοντα μοντέλα, ως προς το μέσο απόλυτο σφάλμα.

## **ΚΕΦΑΛΑΙΟ 3**

### **3.0 ΕΙΣΑΓΩΓΗ.**

Στο κεφάλαιο αυτό περιγράφεται η ανάπτυξη μοντέλου για την πρόβλεψη του συντελεστή απόκλισης  $z$  με χρήση ενός κατάλληλα εκπαιδευμένου νευρωνικού δικτύου. Κύριος σκοπός του πρώτου αυτού σκέλους της διπλωματικής εργασίας είναι η εκμάθηση του περιβάλλοντος του MATLAB, καθώς και η εξοικείωση με τη λειτουργία και τη συμπεριφορά των αλγορίθμων νευρωνικών δικτύων. Γίνεται αναφορά στη σημασία και τη χρησιμότητα του συντελεστή αυτού στη Μηχανική Πετρελαίου, και γίνεται επίσης μια σύντομη περιγραφή του εργαλείου που χρησιμοποιήθηκε για την εκπόνηση της εργασίας.

### **3.1 ΠΕΡΙΓΡΑΦΗ ΤΟΥ ΠΕΡΙΒΑΛΛΟΝΤΟΣ ΠΡΟΓΡΑΜΜΑΤΙΣΜΟΥ MATLAB.**

Όλη η επεξεργασία των δεδομένων, και ο κώδικας που χρειάστηκε να αναπτυχθεί για την εκπόνηση της εργασίας αυτής έγινε στο περιβάλλον του MATLAB. Το MATLAB (Matrix Laboratory : Εργαστήριο Πινάκων) ως ισχυρό εργαλείο αριθμητικής υπολογιστικής αλλά και προγραμματισμού, είναι ένα



ευρέως διαδεδομένο λογισμικό πακέτο σε πλήθος επιστημονικών πεδίων και πεδίων της Μηχανικής. Είναι δημιούργημα του καθηγητή Cleve Moler και σχεδιάστηκε αρχικά για τις ανάγκες των φοιτητών του τη δεκαετία του 1970. Έκτοτε αποτελεί εμπορικό προϊόν, και είναι διαθέσιμο σε διάφορες εκδόσεις.

Μέσα στο διαδραστικό του περιβάλλον, το MATLAB μπορεί να υποστηρίξει την επίλυση μαθηματικών προβλημάτων υλοποιώντας διαφόρων ειδών συναρτήσεις όπως πραγματικές, μιγαδικές, πεπλεγμένες συναρτήσεις δύο μεταβλητών και άλλες. Στα πλαίσια της στατιστικής ανάλυσης παρέχει τη δυνατότητα απεικόνισης δεδομένων υλοποιώντας ιστογράμματα, εμβαδογράμματα, τομεογράμματα κλπ. Επίσης, το περιβάλλον του προσφέρεται για την ανάπτυξη αλγορίθμων μέσα από την γραφή κώδικα προγραμματισμού, με στόχο την επεξεργασία μεγάλου όγκου δεδομένων.

Πέραν όμως των βασικών του λειτουργιών, το πακέτο του MATLAB προσφέρει ένα πλήθος εργαλειοθηκών (toolboxes) ειδικού ενδιαφέροντος, που απευθύνονται στην ανάπτυξη εξειδικευμένων επιστημονικών εφαρμογών στα πεδία των Μαθηματικών, της Στατιστικής, της Βιολογίας, των Οικονομικών, κλπ. Στο παρόν τεύχος έγινε χρήση της εργαλειοθήκης των Νευρωνικών Δικτύων (Neural Network Toolbox). Η εργαλειοθήκη αυτή παρέχει εξισώσεις και εφαρμογές που σχετίζονται με την μοντελοποίηση μη γραμμικών συστημάτων, προσφέροντας στον χρήστη τις επιλογές που σχετίζονται με τα χαρακτηριστικά του δικτύου αναλόγως με τις εφαρμογές που θα εκτελεστούν, όπως η επιλογή επιβλεπόμενης ή μη μάθησης, ο αριθμός των επιπέδων (layers), ο αριθμός των νευρώνων (neurons), η επιλογή του αλγορίθμου μάθησης κλπ.

### **3.2 ΤΙ ΕΙΝΑΙ Ο ΣΥΝΤΕΛΕΣΤΗΣ ΑΠΟΚΛΙΣΗΣ.**

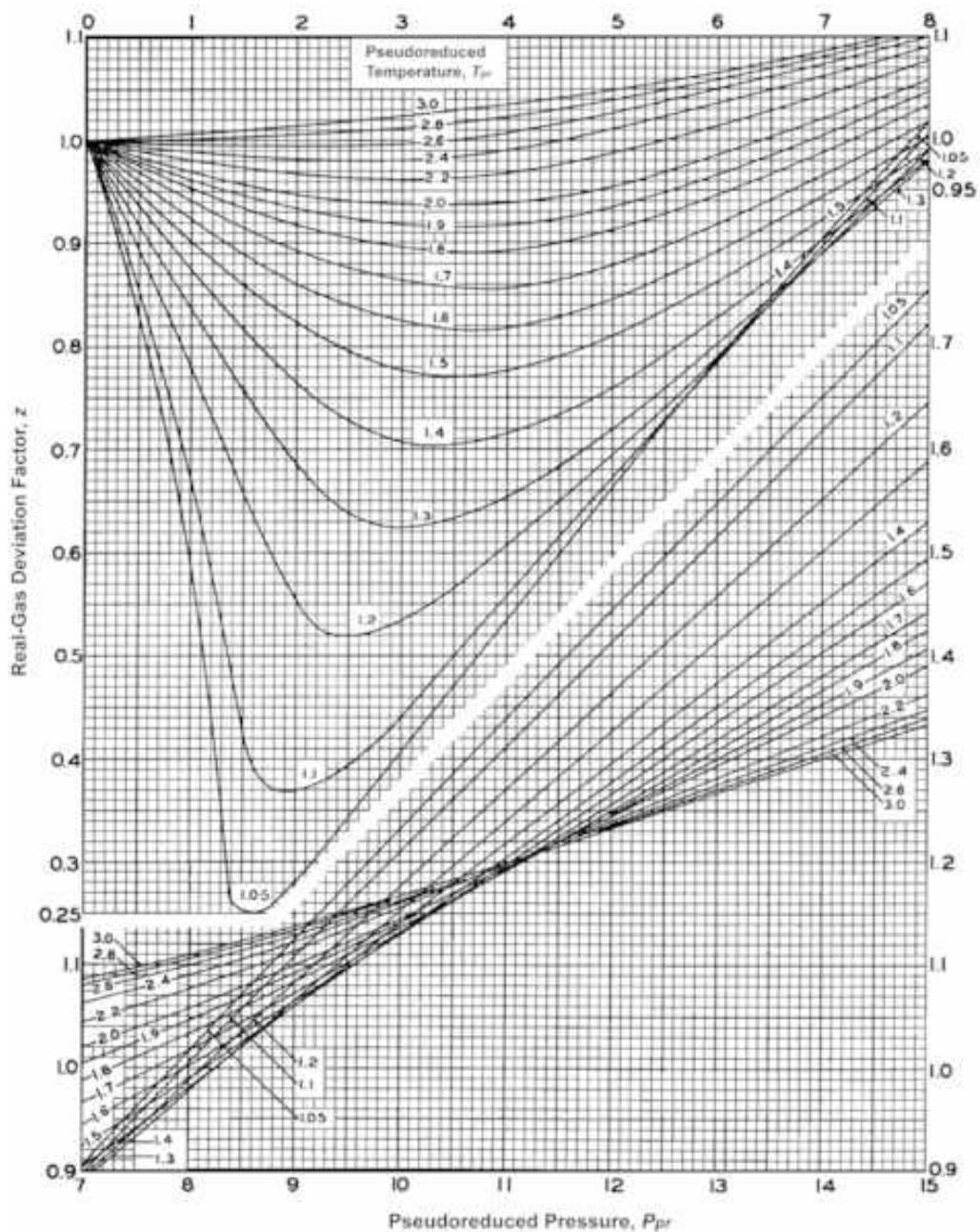
Ο συντελεστής συμπιεστότητας (compressibility factor) ή συντελεστής απόκλισης (deviation factor) ή συντελεστής  $z$ , ορίζεται ως ο λόγος του όγκου ενός πραγματικού αερίου σε συγκεκριμένες συνθήκες πίεσης και

θερμοκρασίας, προς τον όγκο που θα καταλάμβανε αυτό εάν ήταν ιδανικό, στις ίδιες συνθήκες. Είναι εξ'ορισμού μία αδιάστατη παράμετρος η οποία ερμηνεύεται ως το μέτρο της απόκλισης της συμπεριφοράς ενός πραγματικού αερίου από αυτήν του ιδανικού.

Ο συντελεστής αυτός προέκυψε ως ανάγκη στη βιομηχανία του πετρελαίου, με στόχο την επέκταση του πεδίου εφαρμογής της καταστατικής εξίσωσης των ιδανικών αερίων, σε μίγματα αερίων υδρογονανθράκων (πραγματικά αέρια) σε συνθήκες παραγωγής. Για να εκτιμηθεί η σημασία του ακριβούς προσδιορισμού του, αρκεί να αναφερθεί πως χρησιμοποιείται από όλα τα λογισμικά προσομοίωσης της Μηχανικής Πετρελαίου που ασχολούνται με τη συμπεριφορά της φάσης του φυσικού αερίου (Reservoir simulation software) σε διαφόρων ειδών υπολογισμούς όπως η εκτίμηση του κορεσμού σε αέριο ενός σχηματισμού, ή η συμπίεστότητά του αερίου, ή σε υπολογισμούς που σχετίζονται με τον σχεδιασμό διαχωριστήρων, καθώς και αγωγών μεταφοράς πετρελαίου και φυσικού αερίου.

### **3.2.1 ΔΙΑΓΡΑΜΜΑ STANDING ΚΑΙ KATZ.**

Το διάγραμμα που ακολουθεί (Εικόνα 3.1) είναι έργο των Standing και Katz (1942) και αναπτύχθηκε από πειραματικά δεδομένα μεθανίου και αερίων μιγμάτων υδρογονανθράκων. Απεικονίζει τη μεταβολή του  $z$  συναρτήσει της πίεσης και της θερμοκρασίας εκπεφρασμένων σε ψευδοανηγμένες συνθήκες. Χρησιμοποιείται ευρέως στη βιομηχανία πετρελαίου μέχρι και σήμερα κυρίως πλέον σε ψηφιοποιημένη μορφή.



Εικόνα 3.1 : Διάγραμμα Standing & Katz (1942).

Οι συνθήκες πίεσης και θερμοκρασίας του διαγράμματος έχουν αναχθεί ως προς τις κρίσιμες συνθήκες των αερίων ως εξής :

$$P_r = \frac{P}{P_c} \quad \text{και} \quad T_r = \frac{T}{T_c}$$

Όπου :

$P$  = πίεση

$T$  = θερμοκρασία

$P_r$  = ανηγμένη πίεση

$P_c$  = κρίσιμη πίεση

$T_r$  = ανηγμένη θερμοκρασία

$T_c$  = κρίσιμη θερμοκρασία

Οι κρίσιμες συνθήκες των αερίων ενός συστατικού είναι γνωστές μιας και αποτελούν φυσικές σταθερές για το κάθε συστατικό. Άρα και ο υπολογισμός των αντιστοίχων ανηγμένων συνθηκών τους είναι εύκολος.

Η πλειονότητα όμως των πραγματικών αερίων υδρογονανθράκων είναι μίγματα συστατικών, κι έτσι για τον υπολογισμό των τιμών που θα χρησιμοποιηθούν αντί των κρίσιμων συνθηκών τους χρησιμοποιείται ο νόμος του Amagat. Ο νόμος αυτός αναφέρει πως ο συνολικός όγκος ενός αερίου μίγματος είναι το άθροισμα των μερικών όγκων των συστατικών που το αποτελούν, μετρημένων στις ίδιες συνθήκες. Οι μερικοί όγκοι των συστατικών των αερίων μιγμάτων όμως, είναι με τη σειρά τους τα αντίστοιχα μοριακά κλάσματα. Έτσι, όταν γνωρίζουμε το μοριακό κλάσμα κάθε συστατικού ενός μίγματος αερίου υπολογίζουμε τις κρίσιμες ιδιότητες του μίγματος σύμφωνα με τις παρακάτω εξισώσεις που προτάθηκαν από τον W.B. Kay (1936) :

$$P_{pc} = \sum_{i=1}^n y_i \times P_{ci} \quad \text{και} \quad T_{pc} = \sum_{i=1}^n y_i \times T_{ci}$$

Όπου :

$P_{pc}$  = ψευδοκρίσιμη πίεση του αερίου μίγματος

$T_{pc}$  = ψευδοκρίσιμη θερμοκρασία του αερίου μίγματος

$y_i$  = το μοριακό κλάσμα του συστατικού  $i$

$P_{ci}$  = η κρίσιμη πίεση του συστατικού  $i$

$T_{ci}$  = η κρίσιμη θερμοκρασία του συστατικού  $i$

Να σημειωθεί εδώ πως οι εξισώσεις αυτές δεν υπολογίζουν τις πραγματικές κρίσιμες συνθήκες ενός αερίου μίγματος, γι'αυτό αναφέρονται και ως ψευδοκρίσιμες συνθήκες. Είναι όμως παράμετροι που μπορούν να χρησιμοποιηθούν για τον υπολογισμό των ψευδοανηγμένων συνθηκών, και του  $z$  κατ'επέκταση.

Έτσι οι αντίστοιχες ψευδοανηγμένες συνθήκες είναι :

$$P_{pr} = \frac{P}{P_{pc}} \quad \text{και} \quad T_{pr} = \frac{T}{T_{pc}}$$

Σε περίπτωση που η μόνη γνωστή πληροφορία για το αέριο μίγμα, είναι η σχετική του πυκνότητα, τότε για τον υπολογισμό των ψευδοκρίσιμων συνθηκών, έχουν προταθεί από τον Standing οι παρακάτω εμπειρικές σχέσεις :

Για αέριο :

$$P_{pc} = 677 + 15 * \gamma - 37,5 * \gamma^2$$

$$T_{pc} = 168 + 325 * \gamma - 12,5 * \gamma^2$$

Για συμπύκνωμα :

$$P_{pc} = 706 - 51,7 * \gamma - 11,1 * \gamma^2$$

$$T_{pc} = 187 - 330 * \gamma - 715 * \gamma^2$$

Όπου η πίεση μετράται σε psi και η θερμοκρασία σε Rankin.

Η ακρίβεια των παραπάνω εξισώσεων επηρεάζεται σημαντικά από την παρουσία σεβαστών ποσοτήτων αερίων όπως το άζωτο ή το διοξείδιο του άνθρακα οπότε και απαιτούνται διορθώσεις.

### 3.2.2 ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ ΔΙΑΓΡΑΜΜΑΤΟΣ STANDING-KATZ.

Η γενικευμένη χρήση του διαγράμματος αυτού καθίσταται εφικτή, χάρις στον νόμο των αντιστοίχων καταστάσεων ο οποίος επιτρέπει τον συσχετισμό της συμπεριφοράς πραγματικών αερίων με όμοια θερμοδυναμικά χαρακτηριστικά, με εργαλείο τις ψευδοανηγμένες εντατικές τους ιδιότητες. Ο νόμος αυτός προβλέπει πως εάν δύο παρόμοιας φύσης αέρια παρεκκλίνουν το ίδιο από την κρίσιμή τους κατάσταση, όταν δηλαδή δύο από τις εντατικές ανηγμένες ιδιότητες τους (όγκος, πίεση, θερμοκρασία) είναι ίδιες, τότε κατ'ανάγκη και η τρίτη ανηγμένη ιδιότητά τους είναι ίδια. Η πρόταση αυτή επιβεβαιώνεται από την παρακάτω εξίσωση του Van der Waals εκφρασμένη με τις ανηγμένες συνθήκες :

$$(P_r + \frac{3}{V_r^2})(3V_r - 1) = 8T_r \quad (P_r + \frac{3}{V_r^2})(3V_r - 1) = 8T_r$$

Όπου :

$P_r$  = ανηγμένη πίεση

$V_r$  = ανηγμένος όγκος

$T_r$  = ανηγμένη θερμοκρασία

Η ιδέα για τον νόμο των αντιστοίχων καταστάσεων αναπτύχθηκε έπειτα από την εξής παρατήρηση: η γραφική απεικόνιση της συμπεριφοράς των καθαρών αερίων (ως προς την πίεση και την θερμοκρασία) ήταν ποιοτικά ίδια, παρόλο που ποσοτικά οι τιμές της πίεσης και της θερμοκρασίας διέφεραν. Έτσι λοιπόν, οι ιδιότητες των συστατικών θα μπορούσαν να συσχετιστούν, εάν συγκρίνονταν σε έναν κοινό τόπο. Με κοινό σημείο αναφοράς λοιπόν, τις

κρίσιμες εντατικές ιδιότητες του κάθε συστατικού, και δεδομένου ότι η κρίσιμη κατάσταση είναι ποιοτικά ίδια για όλα τα συστατικά, προέκυψε ο κοινός αυτός τύπος των ανηγμένων συνθηκών.

### 3.3 Η ΕΞΙΣΩΣΗ BEGGS & BRILL.

Πολλές εξισώσεις έχουν κατά καιρούς αναπτυχθεί με στόχο τον υπολογισμό του  $z$  μέσω μοντέλων προσομοίωσης των ισοθερμοκρασιακών καμπυλών του διαγράμματος των Standing και Katz. Κάποιες από αυτές τις προσεγγίσεις είναι τα μοντέλα των Dranchuk και Abou-Kassem (DAK), των Nishiumi–Saito, και των Brill–Beggs. Η ανάπτυξη τέτοιων μοντέλων πραγματοποιήθηκε λόγω κάποιων ανακριβών προβλέψεων του  $z$  σύμφωνα με το διάγραμμα των Standing–Katz σε κάποιες οριακές περιοχές. Στην πράξη κάθε ένα από τα προτεινόμενα μοντέλα τείνει να ‘προσαρμόζεται’ καλύτερα σε συγκεκριμένες περιοχές των ισοθερμοκρασιακών καμπυλών, αποτυγχάνοντας όμως κοντά και πέρα από τα όριά του να δώσει καλής ακρίβειας πρόβλεψη του  $z$ . Στην συγκεκριμένη εργασία χρησιμοποιήθηκε η εξίσωση των Beggs και Brill για την παραγωγή των διαφόρων τιμών του  $z$  που αποτέλεσε όπως θα δούμε και στη συνέχεια την έξοδο του τεχνητού νευρωνικού δικτύου.

Εξίσωση Beggs και Brill :

$$z = A + (1 - A)e^{-B} + CP_{pr}^D$$

$$A = 1.39(T_{pr} - 0.92)^{0.5} - 0.36T_{pr} - 0.101$$

$$B = (0.62 - 0.23T_{pr})P_{pr} + \left[ \left( \frac{0.066}{T_{pr} - 0.86} \right) - 0.037 \right] P_{pr}^2 + \left[ \frac{0.32}{10^9(T_{pr} - 1)} \right] P_{pr}^6$$

$$C = (0.132 - 0.32 \log T_{pr})$$

$$D = 10^{0.3106 - 0.49T_{pr} + 0.1824T_{pr}^2}$$

### **3.4 ΠΕΡΙΓΡΑΦΗ ΚΑΙ ΑΞΙΟΛΟΓΗΣΗ ΤΟΥ ΝΕΥΡΩΝΙΚΟΥ ΔΙΚΤΥΟΥ ΠΟΥ ΑΝΑΠΤΥΧΘΗΚΕ ΓΙΑ ΤΗΝ ΠΡΟΒΛΕΨΗ ΤΟΥ ΣΥΝΤΕΛΕΣΤΗ Z.**

Στην ενότητα αυτή σκοπός είναι η ανάπτυξη ενός τεχνητού νευρωνικού δικτύου που μετά την εκπαίδευσή του θα είναι σε θέση να προβλέπει με ακρίβεια τον συντελεστή συμπίεστότητας  $z$  μιγμάτων φυσικών αερίων υδρογονανθράκων.

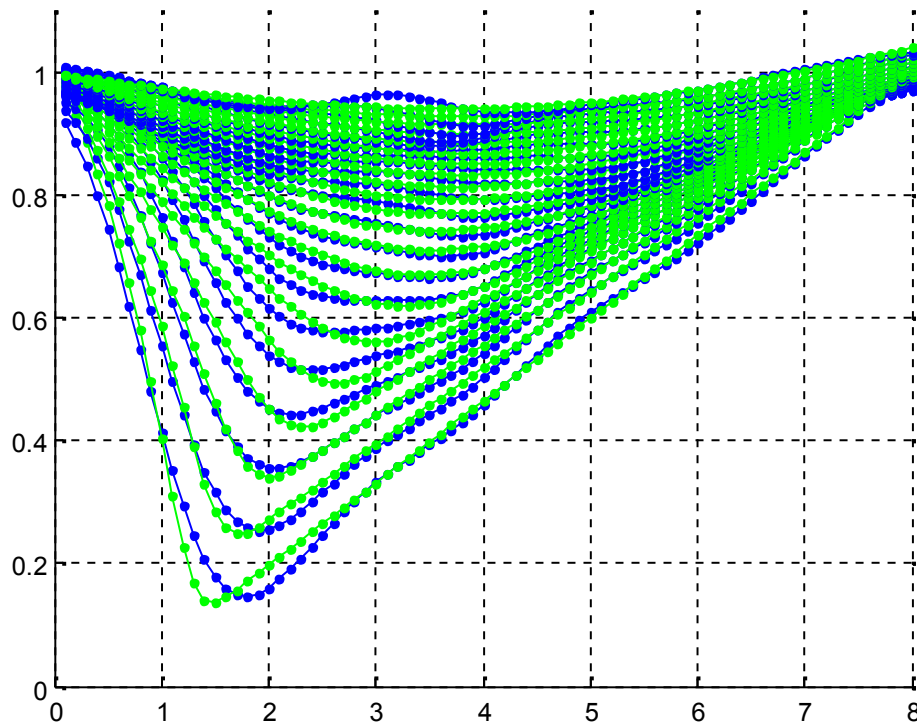
Ο πληθυσμός της εκπαίδευσης παρήχθηκε με οδηγό την φόρμουλα των Beggs και Brill στο περιβάλλον του MATLAB. Τις εισόδους (inputs) του μοντέλου αποτέλεσε ένα πλήθος εκατό χιλιάδων τυχαίων ζευγών ανηγμένης πίεσης και θερμοκρασίας στα εύρη 0,1-8 και 0,95-2,00 αντίστοιχα, και την έξοδο αποτέλεσε ο υπολογισμός του  $z$  για κάθε τυχαίο ζεύγος, όπως αυτός προβλέπονταν από την προαναφερθείσα εξίσωση.

Έγιναν διάφορες δοκιμές ως προς την αρχιτεκτονική του δικτύου μέχρι την τελική του μορφή. Κριτήριο αξιολόγησης της κάθε δοκιμής σε πρώτη φάση, ήταν η πιστότητα των ισοθερμοκρασιακών καμπυλών του διαγράμματος Standing & Katz όπως αυτή προβλεπόταν από το εκπαιδευμένο μοντέλο σε σχέση με αυτήν των Beggs και Brill. Η αναπαράσταση αυτή των ισοθερμοκρασιακών καμπυλών γίνεται με χρήση εισόδων επί των οποίων το δίκτυο δεν είχε εκπαιδευτεί. Ενδεικτικά ακολουθούν τρία από αυτά τα διαγράμματα (Εικόνα 3.2, Εικόνα 3.3, Εικόνα 3.4) με το τελευταίο να αντιπροσωπεύει την αναπαράσταση των ισοθερμοκρασιακών καμπυλών που παρήγαγε το τελικό δίκτυο.

Τα στοιχεία που αλλάζουν σε κάθε δοκιμή εκπαίδευσης είναι το πλήθος των δεδομένων που χρησιμοποιούνται, και ο αριθμός των νευρώνων του νευρωνικού δικτύου. Ενώ σταθερά μένουν τα υπόλοιπα χαρακτηριστικά του δικτύου που είναι η επιλογή ενός κρυφού επιπέδου (hidden layer), οι αναλογίες των δεδομένων εκπαίδευσης (training), επιβεβαίωσης (validation) και τεστ (test) που είναι αντιστοίχως 70%, 15% και 15% του πλήθους των ζευγών ψευδοανηγμένης πίεσης και θερμοκρασίας, και η επιλογή του



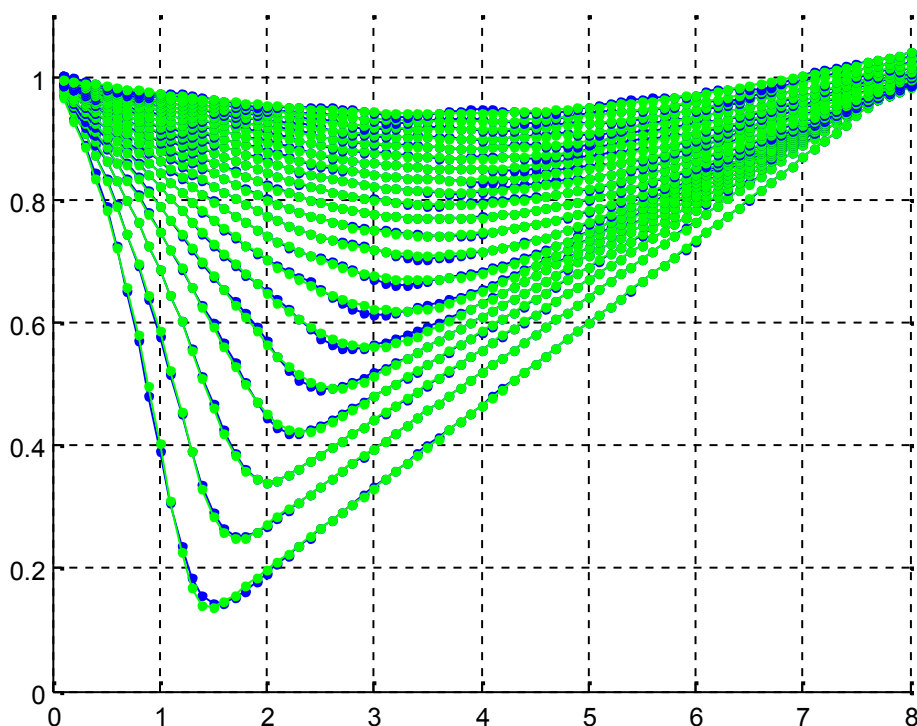
αλγορίθμου μάθησης με οπισθοδιάδοση του σφάλματος (back-propagation training algorithm).



Εικόνα 3.2 : Το πράσινο χρώμα αντιπροσωπεύει τις ισοθερμοκρασιακές καμπύλες της φόρμουλας Begg & Brill, ενώ το μπλε αυτές που παράγει το νευρωνικό δίκτυο με τα εξής χαρακτηριστικά : 20.000 σημεία εκπαίδευσης, αριθμός νευρώνων 10. Ο άξονας των  $x$  αντιστοιχεί στις τιμές της ψευδοανηγμένης πίεσης  $P_{pr}$ , ενώ ο άξονας των  $y$  στις τιμές του  $z$ .

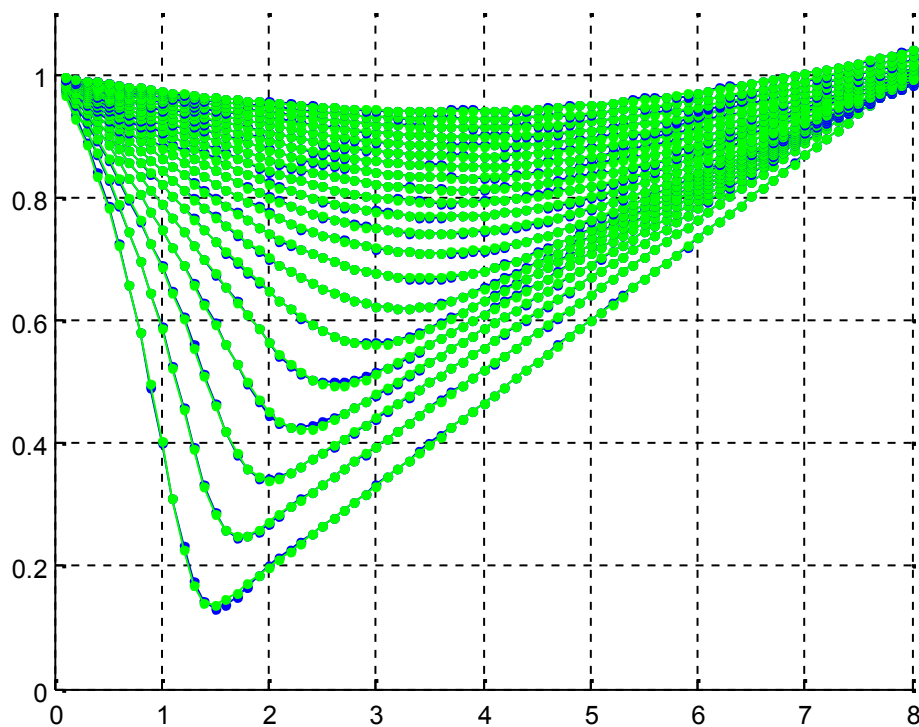
Όπως φαίνεται στο διάγραμμα 3.2 οι ισοθερμοκρασιακές καμπύλες που παράγονται από το δίκτυο (μπλε χρώμα), τείνουν να μην προσαρμόζονται καλά επί των αντιστοίχων ισοθερμοκρασιακών των Begg & Brill (πράσινο χρώμα), και ιδιαίτερος στα άνω και κάτω όρια του διαγράμματος, όπου βέβαια η συγκέντρωση πληροφορίας είναι μικρότερη. Η επιλογή λοιπόν του μεγέθους του διανύσματος των εισόδων (20.000) σε συνδυασμό με τον

αριθμό των νευρώνων (10) που επιλέγη δεν αποδείχθηκε ικανή για να επιτύχει το δίκτυο μια ικανοποιητική εκπαίδευση.



Εικόνα 3.3 : Το πράσινο χρώμα αντιπροσωπεύει τις ισοθερμοκρασιακές καμπύλες της φόρμουλας Begg & Brill, ενώ το μπλε αυτές που παράγει το νευρωνικό δίκτυο με τα εξής χαρακτηριστικά : 50.000 σημεία εκπαίδευσης, αριθμός νευρώνων 20. Ο άξονας των  $x$  αντιστοιχεί στις τιμές της ψευδοανηγμένης πίεσης  $P_{pr}$ , ενώ ο άξονας των  $y$  στις τιμές του  $z$ .

Όπως φαίνεται στο διάγραμμα 3.3 η προσαρμογή των ισοθερμοκρασιακών καμπυλών που παράγει το δίκτυο επί των καμπυλών των Beggs & Brill, τείνει να είναι σαφώς καλύτερη από την προηγούμενη δοκιμή (3.2). Στην δοκιμή αυτή ο αριθμός των εισόδων αυξάνεται κατά 150% (50.000 σημεία έναντι των 20.000 σημείων της πρώτης δοκιμής), ενώ διπλασιάζεται και ο αριθμός των νευρώνων (20 νευρώνες έναντι των 10 νευρώνων της πρώτης δοκιμής). Φαίνεται να υπάρχουν όμως ακόμα περιθώρια βελτίωσης της εκπαίδευσης κι έτσι ακολουθεί και τρίτη δοκιμή (εικόνα 3.4).



Εικόνα 3.4 : (Τελική δοκιμή) Το πράσινο χρώμα αντιπροσωπεύει τις ισοθερμοκρασιακές καμπύλες της φόρμουλας Begg & Brill, ενώ το μπλε αυτές που παράγει το νευρωνικό δίκτυο με τα εξής χαρακτηριστικά : 100.000 σημεία εκπαίδευσης, αριθμός νευρώνων 25. Ο άξονας των  $\chi$  αντιστοιχεί στις τιμές της ψευδοανηγμένης πίεσης  $P_{pr}$ , ενώ ο άξονας των  $\psi$  στις τιμές του  $z$ .

Όπως φαίνεται στο διάγραμμα 3.4 η προσαρμογή των καμπυλών του δικτύου επί των καμπυλών των Beggs & Brill, είναι εξαιρετικά καλή και όπως εξετάζεται κι από τα σφάλματα στη συνέχεια η δοκιμή αυτή κρίνεται αρκετά ικανοποιητική.

Το τελικό δίκτυο αποτελείται από 25 νευρώνες κι ένα κρυφό επίπεδο (hidden layer), και εκτέλεσε την εκπαίδευσή του με την χρήση του αλγορίθμου Levenberg-Marquardt. Τα σφάλματα του μοντέλου εξετάστηκαν σε ένα συγκεκριμένο εύρος του  $z$ , το οποίο ανταποκρίνεται σε αναμενόμενες τιμές μιγμάτων αερίων υδρογονανθράκων. Το εύρος αυτό έχει τιμές από 0,5 έως

1,1 και το μοντέλο στην περιοχή αυτή παρουσιάζει μετά την εκπαίδευσή του τα παρακάτω χαρακτηριστικά.

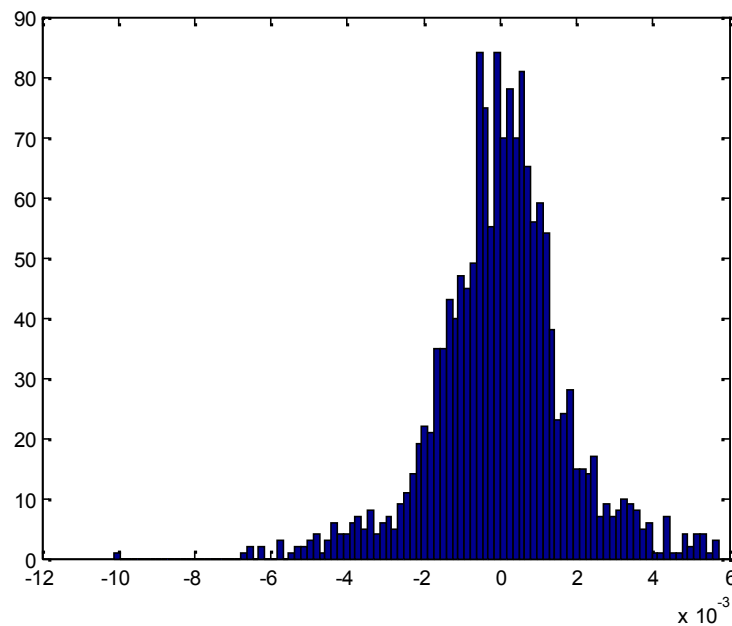
Απόλυτο σφάλμα  $E$  :

$$E_i = z_{f_i} - z_{m_i}$$

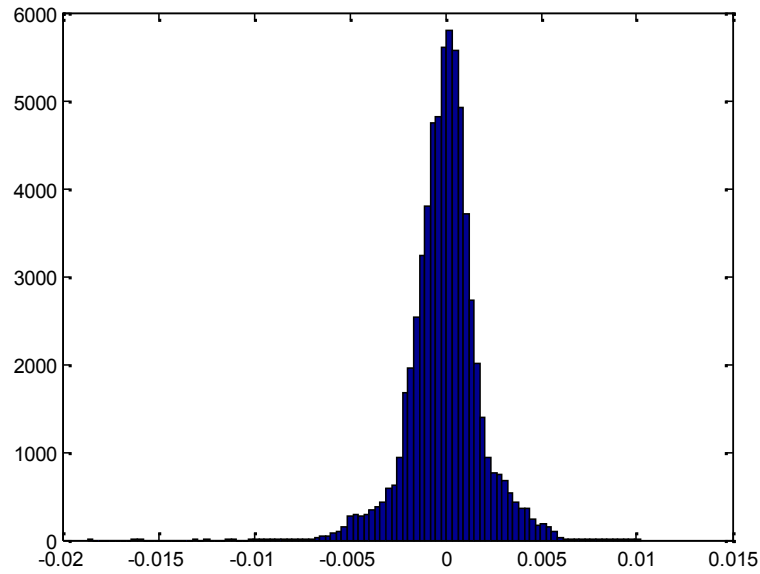
Όπου :

$z_{f_i}$  = η τιμή του συντελεστή  $z$  που προβλέπει η φόρμουλα Beggs & Brill για το ζεύγος ανηγμένης πίεσης και θερμοκρασίας  $P_{pr_i}, T_{pr_i}$ .

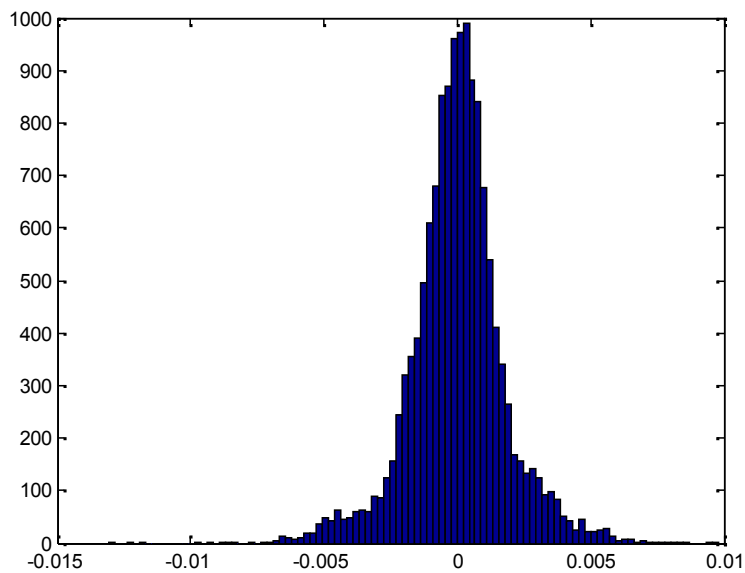
$z_{m_i}$  = η τιμή του συντελεστή  $z$  που προβλέπει το μοντέλο μετά την εκπαίδευσή του για το ίδιο ζεύγος  $P_{pr_i}, T_{pr_i}$ .



**Εικόνα 3.5 :** Ιστόγραμμα απολύτων σφαλμάτων του τελικού δικτύου για σημεία που το δίκτυο δεν έχει εκπαιδευτεί και για εύρος του  $z$  (0.5-1.1). Ο άξονας των  $x$  αντιπροσωπεύει τις τιμές των απολύτων σφαλμάτων  $E_i$ , ενώ ο άξονας των  $y$  το πλήθος των δειγμάτων.



Εικόνα 3.6 : Ιστόγραμμα απολύτων σφαλμάτων του τελικού δικτύου για τα σημεία που αποτέλεσαν σημεία εκπαίδευσης (training) στο εύρος του  $z$  (0.5-1.1). Ο άξονας των  $x$  αντιπροσωπεύει τις τιμές των απολύτων σφαλμάτων  $E_i$ , ενώ ο άξονας των  $y$  το πλήθος των δειγμάτων.



Εικόνα 3.7 : Ιστόγραμμα απολύτων σφαλμάτων του τελικού δικτύου για τα σημεία που αποτέλεσαν σημεία επιβεβαίωσης (validation) στο εύρος του  $z$  (0.5-1.1). Ο άξονας των  $x$  αντιπροσωπεύει τις τιμές των απολύτων σφαλμάτων  $E_i$ , ενώ ο άξονας των  $y$  το πλήθος των δειγμάτων.

Όπως φαίνεται από το ιστόγραμμα των απολύτων σφαλμάτων (Εικόνα 3.4) η δειγματοσυνάρτηση του απολύτου σφάλματος ακολουθεί την κανονική κατανομή. Έτσι σύμφωνα με την πιθανότητα η τιμή του απολύτου σφάλματος να περιέχεται σε ένα συγκεκριμένο διάστημα τιμών, ορίστηκαν τα διαστήματα εμπιστοσύνης του τελικού μοντέλου ως εξής :

Η τιμή του απολύτου σφάλματος για μία πρόβλεψη της τιμής του  $z$  από το τελικό μοντέλο, έχει 68.27% πιθανότητα να βρεθεί στο διάστημα  $(-0.0017, 0.0017)$ , 95.45% πιθανότητα να βρεθεί στο διάστημα  $(-0.0035, 0.0035)$  και 99.73% πιθανότητα να βρεθεί στο διάστημα  $(-0.0052, 0.0052)$ .

## ΚΕΦΑΛΑΙΟ 4ο

### 4.0 ΕΙΣΑΓΩΓΗ.

Στο παρόν κεφάλαιο περιγράφονται όλα τα στάδια της ανάπτυξης ενός μοντέλου που χρησιμοποιώντας τα τεχνητά νευρωνικά δίκτυα θα προβλέπει τη γραμμομοριακή σύσταση πετρελαϊκών ρευστών στα κλάσματα  $CO_2$ ,  $C_1$ ,  $C_2$ ,  $C_3$ ,  $C_4$ ,  $C_5$ ,  $C_{6-8}$ ,  $C_{9-11}$  και  $C_{12+}$ , αξιοποιώντας τις μετρήσεις του οπτικού αναλυτή ρευστών OFA, που είναι η επί τοις εκατό κατά μάζα περιεκτικότητα του ρευστού σε  $CO_2$ ,  $C_1$ ,  $C_2$ ,  $C_{3-5}$ , και  $C_{6+}$ , σε συνδυασμό με κάποιες μετρήσεις πεδίου όπως είναι ο δείκτης πτητικότητας GOR, ο δείκτης πυκνότητας πετρελαίου API, και τέλος η σχετική πυκνότητα της αερίου φάσης  $S_g$  στην επιφάνεια. Γίνεται αναφορά στη βάση δεδομένων που χρησιμοποιήθηκε, καθώς και στην επεξεργασία που αυτή υπέστη ώστε να προκύψουν τα δεδομένα εκπαίδευσης των νευρωνικών δικτύων. Τέλος παρατίθεται η αξιολόγηση του μοντέλου μαζί με σχόλια, συμπεράσματα και προτάσεις για περαιτέρω έρευνα.

#### 4.1 ΒΑΣΗ ΔΕΔΟΜΕΝΩΝ ΚΑΙ ΕΠΕΞΕΡΓΑΣΙΑ.

Η βάση δεδομένων που παραλήφθηκε είναι η γραμμομοριακή επί τοις εκατό ανάλυση 752 μιγμάτων πετρελαίων και φυσικών αερίων, όπως αυτή προέκυψε από την εργαστηριακή τους ανάλυση. Έτσι τα αρχικά δεδομένα πριν την επεξεργασία είναι :

- Η επί τοις εκατό κατά mol σύσταση του κάθε μίγματος σε :

- Διοξείδιο του άνθρακα ( $CO_2$ ).

- Υδρόθειο ( $H_2S$ ).

- Άζωτο ( $N_2$ ).

- Μεθάνιο ( $C_1$ ).

- Αιθάνιο ( $C_2$ ).

- Προπάνιο ( $C_3$ ).

- Βουτάνιο ( $C_4$ ).

- Ισομερές βουτάνιο.

- Πεντάνιο ( $C_5$ ).

- Ισομερή πεντάνια.

- Ψευδοεξάνια ( $C_6$ ).

- Ψευδοεπτάνια ( $C_7$ ).



- Ψευδοοκτάνια ( $C_8$ ).
- Ψευδοεννιάνια ( $C_9$ ).
- Ψευδοδεκάνια ( $C_{10}$ ).
- Ψευδοεντεκάνια ( $C_{11}$ ).
- Κλάσμα ψευδοδωδεκανίων και πάνω ( $C_{12+}$ ).
- Το μέσο μοριακό βάρος MW του κάθε ρευστού σε  $\frac{gr}{mole}$ .
- Η πυκνότητα της υγρής φάσεως  $d_{STO}$  σε  $\frac{gr}{cm^3}$ .
- Η σχετική πυκνότητα της αερίου φάσεως στην επιφάνεια  $S_g$ .
- Ο δείκτης πτητικότητας GOR  $\frac{scf}{STB}$ .

Η επεξεργασία που υπέστησαν τα παραπάνω δεδομένα, με τελικό στόχο να προκύψουν οι είσοδοι των νευρωνικών δικτύων που εφαρμόστηκαν στη συνέχεια, χωρίστηκε στα εξής βήματα :

1. Πρόσθεση των επί τοις εκατό κατά mole περιεκτικότητων των ισομερών βουτανίων ( $i-C_4$ ) με τα βουτάνια ( $n-C_4$ ), και των ισομερών πεντανίων ( $i-C_5$ ) με τα πεντάνια ( $n-C_5$ ).

Το βήμα αυτό κρίνεται πως δεν έχει σημαντική επίδραση επί των αποτελεσμάτων της παρούσης εργασίας. Η μόνη διαφορά των ισομερών βουτανίων με τα βουτάνια και των ισομερών πεντανίων με τα πεντάνια είναι στο σχήμα των μορίων τους, που έχει ως συνέπεια μια διαφορά και στις

κρίσιμες ιδιότητές τους ( $P_c$ ,  $T_c$  και  $V_c$ ). Οι τιμές όμως των κρίσιμων ιδιοτήτων των ισομερών αυτών είναι εξαιρετικά κοντά (ενδεικτικά αναφέρεται πως η κρίσιμη πίεση του n-βουτανίου και του i-βουτανίου είναι 551 και 529 psia αντίστοιχα), κι έτσι μια ενδεχόμενη επίδραση της πρόσθεσης των ισομερών αυτών στο GOR ή στο API κρίνεται αμελητέα.

2. Αποκλεισμός των δεδομένων με περιεκτικότητα σε μεθάνιο ( $CH_4$ ) μικρότερη του 5% καθώς τέτοιες περιπτώσεις αποτελούν εξαιρέσεις και ενδεχομένως θα απειλούσαν τη γενίκευση του μοντέλου.
3. Αποκλεισμός των δεδομένων με σχετική πυκνότητα υγρής φάσης  $\frac{d_{sto}}{1000} > 1$ , καθώς ένα τέτοιο δεδομένο πιθανότατα οφείλεται σε λάθος μέτρηση.
4. Αποκλεισμός των δειγμάτων της βάσης δεδομένων με περιεκτικότητα σε άζωτο ( $N_2$ ) και υδρόθειο ( $H_2S$ ) μεγαλύτερη του 5%.

Τόσο το άζωτο όσο και το υδρόθειο δεν μετρώνται από τον οπτικό αναλυτή ρευστών OFA, η βάση δεδομένων όμως, έχει την πληροφορία για το επί τοις εκατό κατά mole ποσοστό τους στη συνολική σύσταση των ρευστών. Έτσι για να παραχθεί ο πληθυσμός εκπαίδευσης των νευρωνικών δικτύων που εφαρμόζονται στα επόμενα στάδια, επιλέγεται το 5% ως ανώτατη τιμή ανοχής του ποσοστού του αζώτου ( $N_2$ ) και ως ανώτατη τιμή ανοχής του υδρόθειου ( $H_2S$ ) στη συνολική σύσταση του κάθε ρευστού. Επισημαίνεται στο σημείο αυτό πως μετά και από αυτό το βήμα της αρχικής επεξεργασίας της βάσης δεδομένων, το σύνολο των δειγμάτων των ρευστών είναι πλέον 660, από 752 που ήταν αρχικώς.

5. Αναγωγή επί τοις εκατό επί του συνόλου της κατά mole περιεκτικότητας των ρευστών μετά την αφαίρεση σε άζωτο ( $N_2$ ) και υδρόθειο ( $H_2S$ ).

Το βήμα αυτό αποσκοπεί στον υπολογισμό συστάσεων αντίστοιχων με τις συστάσεις που παρέχει το OFA. Έτσι, τα γραμμομοριακά κλάσματα των συστατικών στα καινούρια σύνολα θα είναι :

$$n'_i = \frac{n_i * (100 - n_{H_2S} - n_{N_2})}{100} \text{ mole \%}$$

Όπου :

$n'_i$  = η επί τοις εκατό κατά mole περιεκτικότητα του ρευστού στο συστατικό  $i$  επί του συνόλου  $(100 - n_{H_2S} - n_{N_2})$ .

$n_i$  = επί τοις εκατό κατά mole περιεκτικότητα του ρευστού στο συστατικό  $i$  πριν την αφαίρεση των  $n_{H_2S}$  και  $n_{N_2}$  από τη συνολική σύσταση του ρευστού.

6. Μετατροπή των επί τοις εκατό κατά mole συστάσεων των ρευστών, σε επί τις εκατό κατά μάζα συστάσεις.

Οι μετρήσεις του OFA αναφέρονται στην επί τοις εκατό κατά μάζα περιεκτικότητα του κάθε ρευστού στα συστατικά και κλάσματα  $CO_2, C_1, C_2, C_{3-5}$ , και  $C_{6+}$ , ενώ η βάση δεδομένων δίνει πληροφορία για την γραμμομοριακή σύσταση των επιμέρους συστατικών. Για την μετατροπή λοιπόν των επί τοις εκατό κατά mole συστάσεων σε επί τοις εκατό κατά μάζα συστάσεις, ακολουθούν οι εξής υπολογισμοί :

Εφαρμόζοντας τη σχέση (4.1) υπολογίζουμε τη μάζα των συστατικών  $CO_2, C_1, C_2, C_3, C_4, C_5, C_6, C_7, C_8, C_9, C_{10}, C_{11}$  στο δείγμα του κάθε ρευστού :

$$m_i = n_i \times Mw_i \quad (4.1)$$

Όπου :

$m_i$  = η μάζα του συστατικού  $i$  σε gr.

$n_i$  = το γραμμομοριακό κλάσμα του συστατικού  $i$ .

$Mw_i$  = το μοριακό βάρος του συστατικού  $i$  σε  $\frac{gr}{mole}$ .

Για τις τιμές των μοριακών βαρών των  $CO_2$ ,  $C_1$ ,  $C_2$ ,  $C_3$ ,  $C_4$ ,  $C_5$ ,  $C_6$ ,  $C_7$ ,  $C_8$ ,  $C_9$ ,  $C_{10}$ ,  $C_{11}$  χρησιμοποιήθηκαν τα παρακάτω μοριακά βάρη (κατά Katz [\*ref\*] για τα ψευδοσυστατικά) σε  $\frac{gr}{mole}$  :

$$MR_{CO_2} = 44$$

$$MR_{C_1} = 16$$

$$MR_{C_2} = 30$$

$$MR_{C_3} = 44$$

$$MR_{C_4} = 58.12$$

$$MR_{C_5} = 72.15$$

$$MR_{C_6} = 84$$

$$MR_{C_7} = 96$$

$$MR_{C_8} = 107$$

$$MR_{C_9} = 121$$

$$MR_{C_{10}} = 134$$

$$MR_{C_{11}} = 147$$

Καθώς δεν υπάρχει άμεση πληροφορία για την τιμή του μέσου μοριακού βάρους του κλάσματος  $C_{12^+}$ , από την βάση δεδομένων, ο υπολογισμός της μάζας του κλάσματος  $C_{12^+}$  έγινε από την εφαρμογή του ισοζυγίου μάζας επί του συνόλου του ρευστού ως εξής :

$$MW*100 = m_{CO_2} + m_{C_1} + m_{C_2} + m_{C_3} + m_{C_4} + m_{C_5} + m_{C_6} + m_{C_7} + m_{C_8} + m_{C_9} + m_{C_{10}} + m_{C_{11}} + m_{C_{12^+}} \quad (4.2)$$

Όπου :

$m_{C_{12^+}}$  = η μάζα του κλάσματος  $C_{12^+}$  που περιέχεται σε 100 moles ρευστού.

$MW$  = το μέσο μοριακό βάρος κάθε ρευστού σε  $\frac{gr}{mole}$ .

Λύνοντας την (4.2) ως προς  $m_{C_{12^+}}$ , υπολογίζουμε και τη μάζα του κλάσματος  $C_{12^+}$  που περιέχεται σε 100 moles ρευστού. Μετά εφαρμόζοντας τη σχέση (4.3), υπολογίζουμε την επί τοις εκατό κατά μάζα περιεκτικότητα του κάθε συστατικού :

$$M_i = \frac{m_i}{\sum_{i=1}^N m_i} \frac{gr}{gr} \quad (4.3)$$

Όπου :

$M_i$  = η επί τοις εκατό κατά μάζα περιεκτικότητα του συστατικού  $i$  στο ρευστού που απαρτίζεται από τα συστατικά  $N$ .

$m_i$  = η μάζα του συστατικού  $i$  σε  $gr$  .

$\sum_{i=1}^N m_i$  = η συνολική μάζα όλων των συστατικών  $N$  του ρευστού σε  $gr$  .

Έτσι υπολογίστηκαν οι επί της εκατό κατά μάζα περιεκτικότητες ( $M_i$ ) των συστατικών κάθε ρευστού, που είναι και οι ζητούμενες με στόχο την αντιστοίχιση των δεδομένων της βάσης με τα δεδομένα που παρέχει το OFA.

## 7. Δημιουργία των εισόδων των νευρωνικών δικτύων.

Τελικός στόχος των βημάτων της επεξεργασίας είναι οι μετρήσεις του OFA, που είναι οι επί τοις εκατό κατά μάζα περιεκτικότητες του κάθε ρευστού σε  $CO_2, C_1, C_2, C_{3-5}$ , και  $C_{6+}$ , και τα δεδομένα πεδίου, που είναι ο δείκτης πτητικότητας πετρελαίων GOR, ο δείκτης πυκνότητας πετρελαίων API, και η σχετική πυκνότητα της αερίου φάσεως του ρευστού  $S_g$  να αποτελούν τις εισόδους των δικτύων που εφαρμόζονται στο επόμενο στάδιο. Έτσι λοιπόν, το διάνυσμα εισόδου των νευρωνικών δικτύων που εφαρμόζονται στη

συνέχεια είναι το  $X(X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8)$ , όπου η καθεμία από τις 8 συνιστώσες του είναι αντίστοιχα :

$$X_1 = M_{CO_2}$$

$$X_2 = M_{C_1}$$

$$X_3 = M_{C_2}$$

$$X_4 = (M_{C_3} + M_{C_4} + M_{C_5}) = M_{C_{3-5}}$$

$$X_5 = (M_{C_6} + M_{C_7} + M_{C_8} + M_{C_9} + M_{C_{10}} + M_{C_{11}} + M_{C_{12}}) = M_{C_{6+}}$$

$$X_6 = \log(GOR)$$

$$X_7 = \frac{141.5}{\left(\frac{d_{STO}}{1000}\right)} - 131.5 = API$$

$$X_8 = S_g$$

Όπου :

$M_{CO_2}$  = η επί τοις εκατό κατά μάζα περιεκτικότητα του ρευστού σε  $CO_2$ .

$M_{C_1}$  = η επί τοις εκατό κατά μάζα περιεκτικότητα του ρευστού σε  $C_1$ .

$M_{C_2}$  = η επί τοις εκατό κατά μάζα περιεκτικότητα του ρευστού σε  $C_2$ .

$M_{C_{3-5}}$  = η επί τοις εκατό περιεκτικότητα του ρευστού στο κλάσμα  $C_{3-5}$ .

$M_{C_{6+}}$  = η επί τοις εκατό περιεκτικότητα του ρευστού στο κλάσμα  $C_{6+}$ .

$d_{STO}$  = η πυκνότητα της υγρής φάσης του ρευστού σε  $\frac{gr}{cm^3}$ .

Λόγω του μεγάλου εύρους διακύμανσης της τιμής του GOR (συγκεκριμένα από 4.9621 μέχρι  $9.0345 \cdot 10^5$  scf/STB, σαν λόγος αερίου-πετρελαίου ορίζεται το κλάσμα που έχει αριθμητή τον όγκο του παραχθέντος αερίου μετρημένο σε κανονικές συνθήκες και παρονομαστή τον όγκο του παραχθέντος πετρελαίου μετρημένο και αυτόν σε κανονικές συνθήκες) επιλέγεται να μετασχηματιστεί το GOR ως  $\log(GOR)$  (νεπέριος λογάριθμος του GOR) πριν μπει ως πληροφορία στις εισόδους. Η επιλογή αυτή αποσκοπεί στην διευκόλυνση της διαδικασίας εκπαίδευσης των δικτύων, καθώς ‘συμπιέζοντας’ το εύρος των τιμών του GOR μέσω του λογαρίθμου (συγκεκριμένα από 1.6018 μέχρι 13.7140 scf/STB), πυκνώνει η κατανομή των τιμών της εισόδου  $X_6$ , με αποτέλεσμα να περιορίζεται η αβεβαιότητα του δικτύου στα διαστήματα όπου δεν υπάρχει πληροφορία για το GOR. Έτσι η έκτη συνιστώσα του διανύσματος εισόδου είναι η  $X_6 = \log(GOR)$ .

## 4.2 ΕΦΑΡΜΟΓΗ ΤΩΝ ΤΕΧΝΗΤΩΝ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ.

Για την ανάπτυξη του παρόντος μοντέλου αναπτύχθηκαν 5 διαφορετικά νευρωνικά δίκτυα, τα οποία έχουν ως κοινό το διάνυσμα εισόδου  $X(X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8)$ , αλλά εκπαιδεύονται για διαφορετικά διανύσματα εξόδου το καθένα, ώστε να επιτευχθεί ο τελικός στόχος του μοντέλου. Πιο συγκεκριμένα, μετά το πέρας της εκπαίδευσης των νευρωνικών δικτύων η πληροφορία που θα υπάρχει διαθέσιμη ως προς τη σύσταση των ρευστών θα είναι η επί τοις εκατό κατά μάζα περιεκτικότητα του κάθε ρευστού σε  $CO_2$ ,  $C_1$ ,  $C_2$ ,  $C_3$ ,  $C_4$ ,  $C_5$ ,  $C_{6-8}$ ,  $C_{9-11}$  και  $C_{12+}$ .

- **1° Νευρωνικό Δίκτυο :**

Ως έξοδος του πρώτου νευρωνικού δικτύου που αναπτύσσεται ορίζεται το διάνυσμα  $Y_1$  :



$$Y_1 = \left( \frac{M_{C_3}}{M_{C_{3-5}}} \right)$$

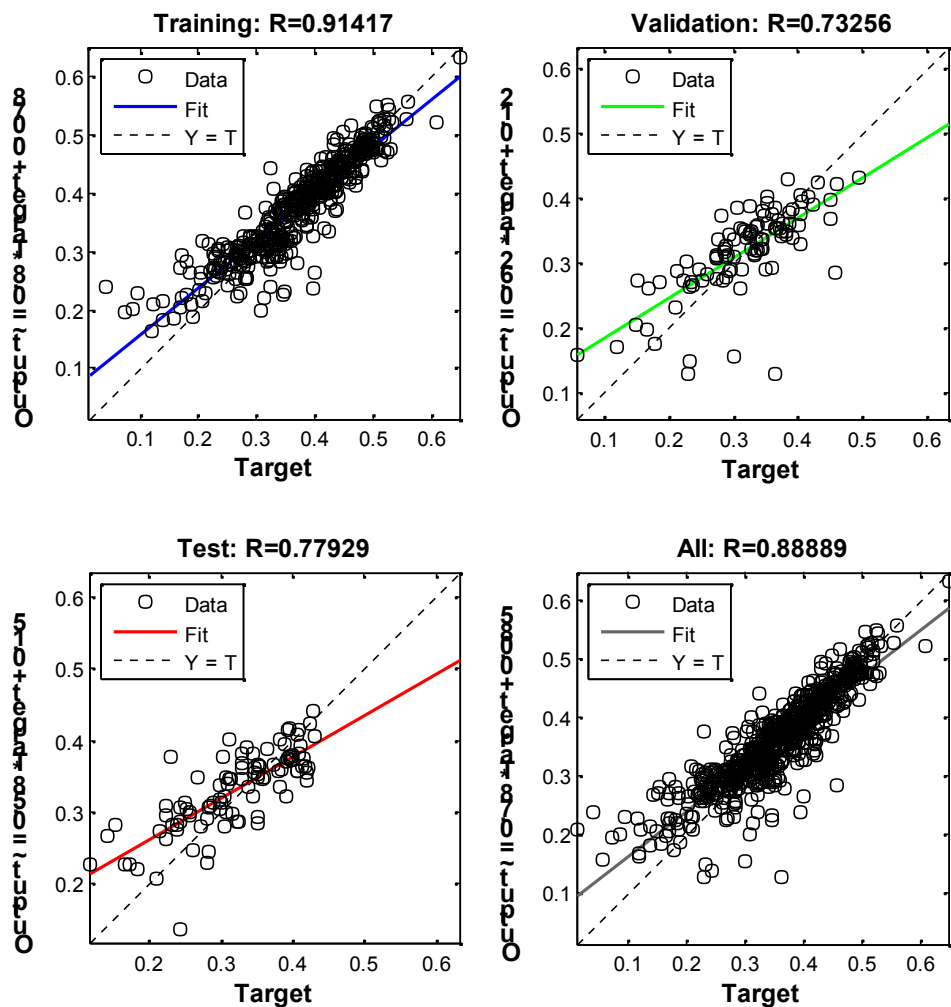
Κατά αυτόν τον τρόπο το πρώτο νευρωνικό δίκτυο εκπαιδεύεται για να μάθει να υπολογίζει τον λόγο του  $C_3$ , στο κλάσμα  $C_{3-5}$ . Μετά το πέρας της εκπαίδευσής του το δίκτυο θα δίνει ως έξοδο το  $Y_1'$ , που θα είναι η βέλτιστη προσέγγιση του  $Y_1$ . Έτσι η πρόβλεψη του μοντέλου για το  $M_{C_3}'$ , που είναι η τιμή της επί τοις εκατό κατά μάζα περιεκτικότητας του ρευστού σε  $C_3$ , θα προκύπτει ως εξής :

$$M_{C_3}' = Y_1' * M_{C_{3-5}} \quad (4.4)$$

Επισημαίνεται στο σημείο αυτό πως ο λόγος για τον οποίον δεν επιλέγεται απευθείας ως έξοδος του δικτύου η  $M_{C_3}$ , που είναι η επί τοις εκατό κατά μάζα περιεκτικότητα του ρευστού σε  $C_3$ , είναι για να αποφευχθεί το ενδεχόμενο απόκλισης της τιμής του  $M_{C_3}'$  από το  $M_{C_{3-5}}$ .

Το παρόν νευρωνικό δίκτυο, είναι δίκτυο πρόσθιας τροφοδότησης ενός κρυφού επιπέδου και εκτέλεσε την εκπαίδευσή του με τον αλγόριθμο μάθησης Levenberg-Marquardt. Υπενθυμίζεται πως το σύνολο χαρακτηρισμένων παραδειγμάτων επί του οποίου εκπαιδεύτηκε το παρόν δίκτυο, είναι το  $\mathfrak{N}_1 \{X(X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8), Y_1\}$  που αποτελείται από 660 δείγματα ρευστών. Από το σύνολο των 660 δειγμάτων, 462 δείγματα χρησιμοποιήθηκαν ως σημεία εκπαίδευσης (training data), 99 δείγματα χρησιμοποιήθηκαν ως σημεία επιβεβαίωσης (validation data) και ως σημεία ελέγχου (test data) χρησιμοποιήθηκαν τα υπόλοιπα 99. Το κρυφό επίπεδο του δικτύου αποτελείται από 10 νευρώνες.

Αποτελέσματα 1<sup>ου</sup> Νευρωνικού Δικτύου :



**Εικόνα 4.1 :** Χαρακτηριστικά εκπαίδευσης του πρώτου νευρωνικού δικτύου.

Όπως φαίνεται στην Εικόνα 4.1, τα αποτελέσματα της εκπαίδευσης του πρώτου δικτύου κρίθηκαν ικανοποιητικά. Στον άξονα των  $x$  των

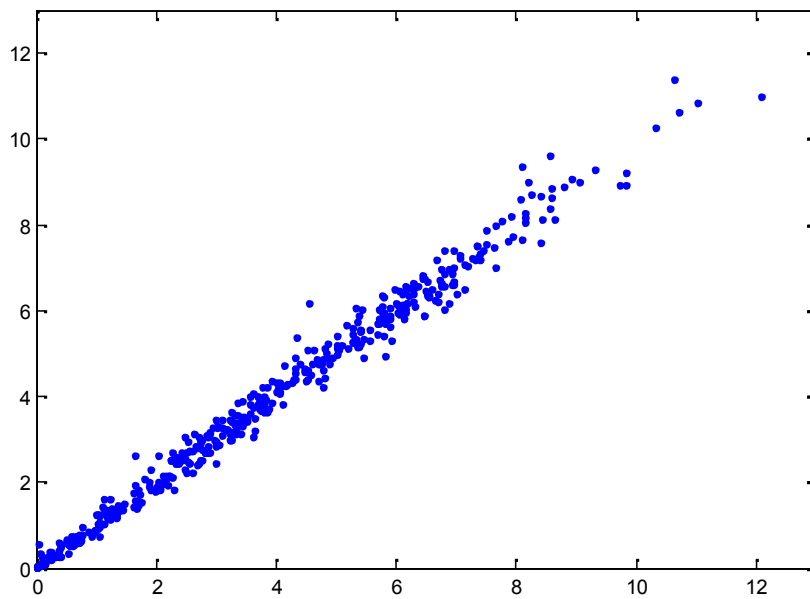
διαγραμμάτων βρίσκονται οι τιμές του  $Y_1$  που αποτέλεσαν τον μαθησιακό στόχο του δικτύου, ενώ στον άξονα των  $\psi$  βρίσκονται οι προβλέψεις  $Y_1'$  που παράγει το δίκτυο μετά την εκπαίδευσή του. Το πάνω αριστερά διάγραμμα αφορά τα σημεία εκπαίδευσης του δικτύου (training), το πάνω δεξιά τα σημεία επιβεβαίωσης του δικτύου (validation), το κάτω αριστερά τα σημεία ελέγχου (τεστ), και το κάτω δεξιά στο σύνολο των σημείων (training+validation+test).

Βασικό κριτήριο για την αξιολόγηση της εκπαίδευσης, αποτελεί το γεγονός πως το μέσο σχετικό σφάλμα των προβλέψεων του δικτύου για τα σημεία της εκπαίδευσης (training), είναι κοντά με αυτό των σημείων επιβεβαίωσης (validation). Συγκεκριμένα  $\bar{R}_{tr} = \frac{Y_{1_{tr}} - Y_{1_{tr}}'}{Y_{1_{tr}}} = 6.02\%$ , και  $\bar{R}_{val} = 5.69\%$ . Αυτό

σημαίνει πως το δίκτυο έχει την τάση να γενικεύει καλά, γεγονός που αποδεικνύεται κι από τις προβλέψεις του σε σημεία επί των οποίων δεν έχει εκπαιδευτεί (test).

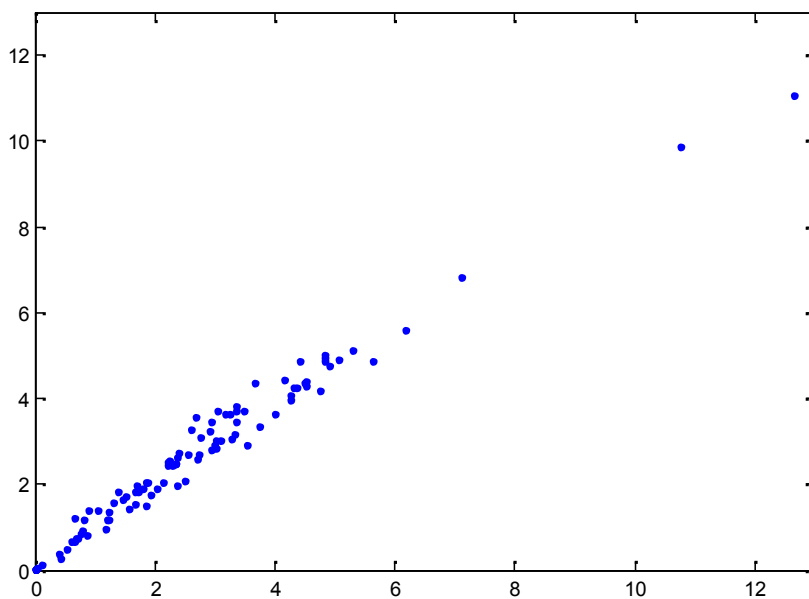
Παρακάτω παρατίθενται τα ιδίου τύπου διαγράμματα για τις προβλέψεις της επί τοις εκατό κατά μάζα περιεκτικότητας των ρευστών σε  $C_3$  που υπολογίστηκαν από τη σχέση (4.4) ως εξής :

$$M_{C_3}' = Y_1' * M_{C_{3-5}}$$

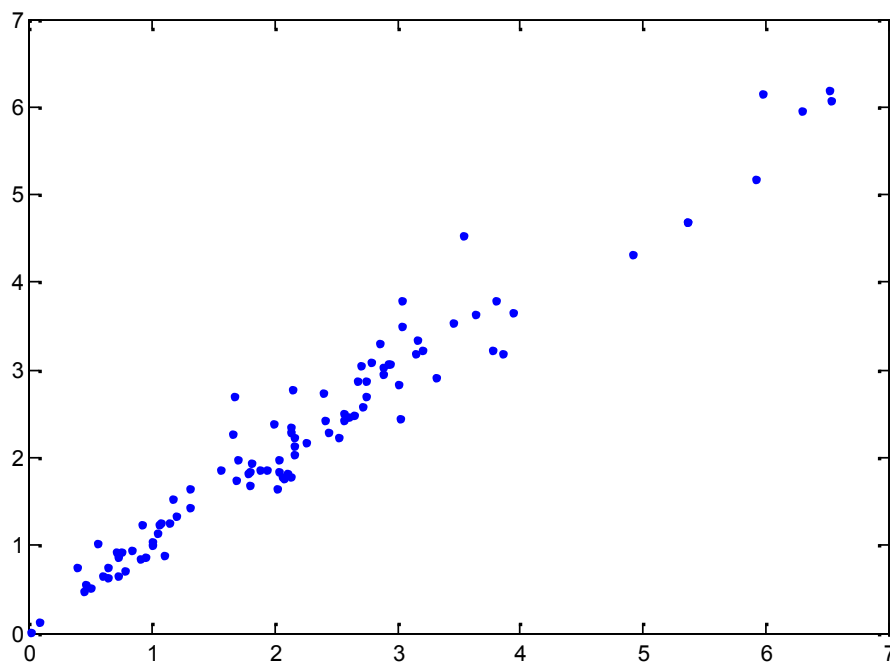


**Εικόνα 4.2 :** Σημεία εκπαίδευσης (training). Στον άξονα των  $\chi$  βρίσκονται οι τιμές  $M_{C_3}$  και στον άξονα των  $\psi$  οι τιμές των  $M_{C_3}'$ .

Το διάγραμμα της εικόνας 4.2 αναφέρεται στα σημεία εκπαίδευσης (training). Κάθε σημείο (μπλε κουκίδα) με συντεταγμένες  $(M_{C_3}, M_{C_3}')$  αντιπροσωπεύει και ένα ρευστό. Η συγκέντρωση των σημείων κοντά στην διαγώνιο (που είναι η  $\psi=\chi$ ), υποδεικνύει την πολύ μικρή απόκλιση των  $M_{C_3}$  από τα  $M_{C_3}'$ , που είναι και το ζητούμενο μετά το πέρας της εκπαίδευσης του πρώτου δικτύου.



**Εικόνα 4.3 :** Σημεία επιβεβαίωσης (validation). Στον άξονα των  $\chi$  βρίσκονται οι τιμές  $M_{C_3}$  και στον άξονα των  $\psi$  οι τιμές των  $M_{C_3}'$ . Το μέσο σχετικό σφάλμα είναι  $\bar{R} = 4.62\%$ .



**Εικόνα 4.4 :** Σημεία ελέγχου (test). Στον άξονα των  $\chi$  βρίσκονται οι τιμές  $M_{C_3}$  και στον άξονα των  $\psi$  οι τιμές των  $M_{C_3}'$ . Το μέσο σχετικό σφάλμα είναι  $\bar{R} = 6.30\%$ .

Το μέσο σχετικό σφάλμα επί του συνόλου των σημείων για τον υπολογισμό της επί τοις εκατό κατά μάζα περιεκτικότητας των ρευστών σε  $C_3$ , είναι  $\bar{R}_{M_{C_3}} = 6.12\%$ .

- **2<sup>ο</sup> Νευρωνικό Δίκτυο :**

Ως έξοδος του δεύτερου κατά σειρά νευρωνικού δικτύου που εφαρμόζεται ορίζεται το διάνυσμα  $Y_2$  :

$$Y_2 = \left( \frac{M_{C_4}}{M_{C_{4-5}}} \right)$$

Όπου :

$$M_{C_{4-5}} = M_{C_4} + M_{C_5}$$

Κατά αυτόν τον τρόπο το δεύτερο νευρωνικό δίκτυο εκπαιδεύεται για να μάθει να υπολογίζει τον λόγο του  $C_4$ , στο κλάσμα  $C_{4-5}$ . Μετά το πέρας της εκπαίδευσής του το δίκτυο θα δίνει ως έξοδο το  $Y_2'$ , που θα είναι η βέλτιστη προσέγγιση του  $Y_2$ . Έτσι η πρόβλεψη του μοντέλου για το  $M_{C_4}'$ , που είναι η τιμή της επί τοις εκατό κατά μάζα περιεκτικότητας του ρευστού σε  $C_4$ , θα προκύπτει ως εξής :

$$M_{C_4}' = Y_2' * M_{C_{4-5}}' \quad (4.5)$$

Όπου :

$$M_{C_{4-5}}' = M_{C_{3-5}} - M_{C_3}'$$

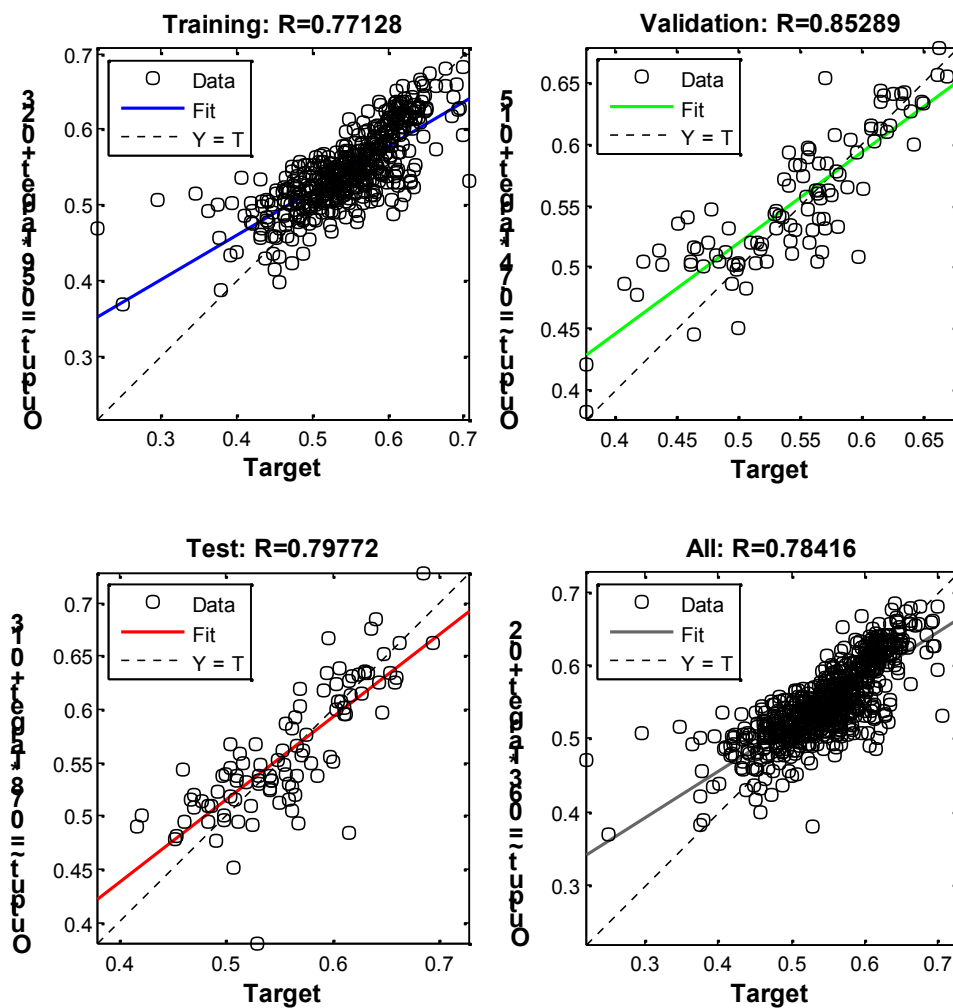
Μετά το πέρας της εκπαίδευσης του πρώτου και του δεύτερου νευρωνικού δικτύου το  $M_{C_5}'$  υπολογίζεται ως εξής :

$$M_{C_5}' = M_{C_{3-5}} - M_{C_3}' - M_{C_4}' \quad (4.6)$$

Επισημαίνεται στο σημείο αυτό πως επιλέγεται ως διάνυσμα εξόδου ο λόγος  $\frac{M_{C_4}}{M_{C_{3-5}}}$  αντί του λόγου  $\frac{M_{C_4}}{M_{C_{3-5}}}$  για να αποφευχθεί κατά τον υπολογισμό του  $M_{C_5}'$ , που είναι η πρόβλεψη του μοντέλου για την επί τοις εκατό κατά μάζα περιεκτικότητα του ρευστού σε  $C_5$ , το ενδεχόμενο της αρνητικής τιμής.

Το παρόν Νευρωνικό δίκτυο είναι κι αυτό δίκτυο πρόσθιας τροφοδότησης ενός κρυφού επιπέδου και εκτέλεσε την εκπαίδευσή του με τον αλγόριθμο μάθησης Levenberg-Marquardt. Υπενθυμίζεται πως το σύνολο χαρακτηρισμένων παραδειγμάτων επί του οποίου εκπαιδεύτηκε το παρόν δίκτυο, είναι το  $\mathfrak{Z}_2 \{X(X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8), Y_2\}$  που αποτελείται από 660 δείγματα ρευστών. Από το σύνολο των 660 δειγμάτων, 462 δείγματα χρησιμοποιήθηκαν ως σημεία εκπαίδευσης (training data), 99 δείγματα χρησιμοποιήθηκαν ως σημεία επιβεβαίωσης (validation data) και ως τεστ σημεία (test data) χρησιμοποιήθηκαν τα υπόλοιπα 99. Το κρυφό επίπεδο του δικτύου αποτελείται από 10 νευρώνες.

## Αποτελέσματα 2<sup>ου</sup> Νευρωνικού Δικτύου :

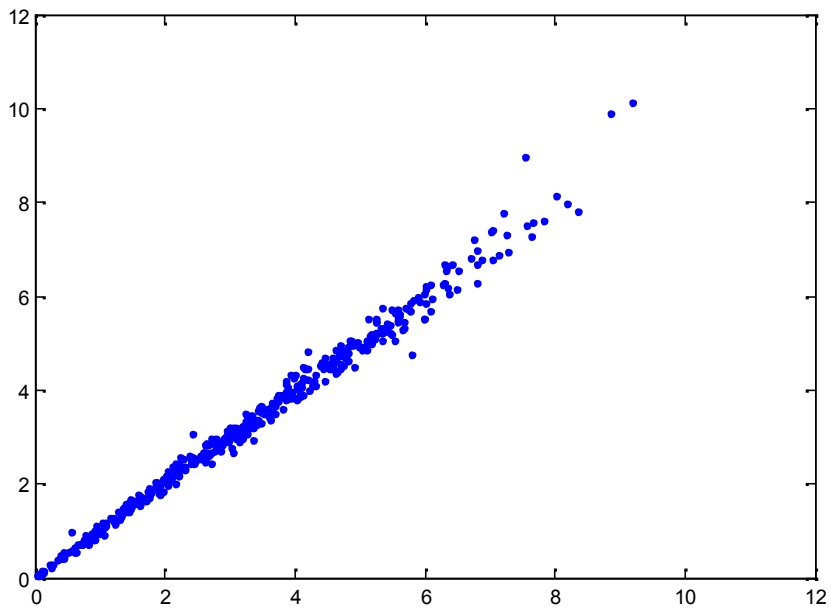


**Εικόνα 4.5 :** Χαρακτηριστικά εκπαίδευσης 2<sup>ου</sup> νευρωνικού δικτύου.

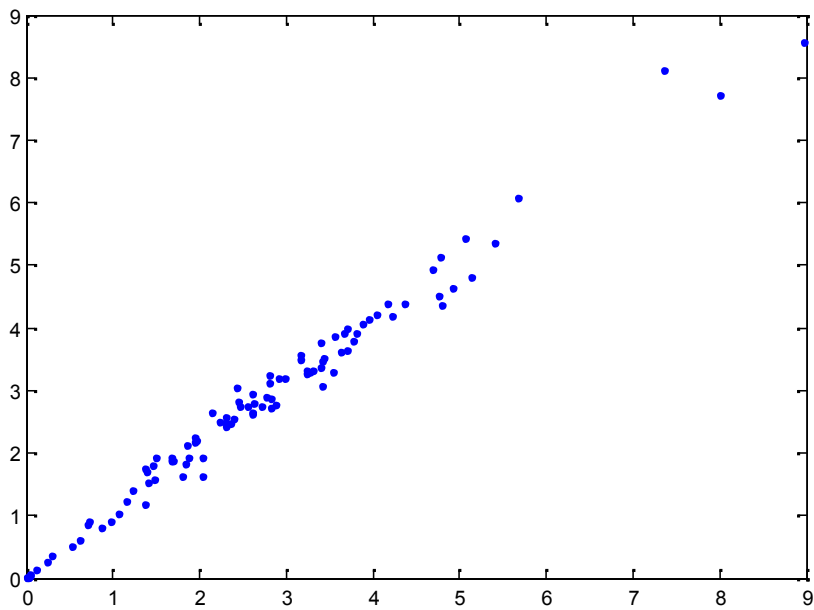
Όπως φαίνεται από την εικόνα 4.5 το δίκτυο φαίνεται να εκπαιδεύτηκε ικανοποιητικά. Ακολουθούν τα αντίστοιχα διαγράμματα για τον υπολογισμό της επί τοις εκατό κατά μάζα περιεκτικότητας των ρευστών σε  $C_4$ , όπως προκύπτει από την χρήση των προβλέψεων του 2<sup>ου</sup> δικτύου από τη σχέση (4.5) :

$$M_{C_4}' = Y_1' * M_{C_{4-5}}'$$

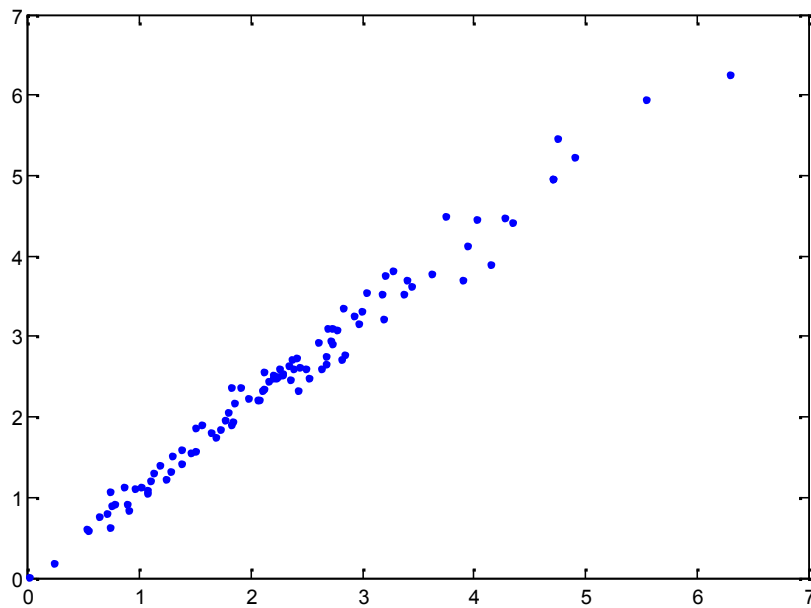




**Εικόνα 4.6 :** Σημεία εκπαίδευσης (training). Στον άξονα των  $\chi$  βρίσκονται οι τιμές  $M_{C_4}$  και στον άξονα των  $\psi$  οι τιμές των  $M_{C_4}'$ . Το μέσο σχετικό σφάλμα είναι  $\bar{R} = 2.04\%$ .



**Εικόνα 4.7 :** Σημεία επιβεβαίωσης (validation). Στον άξονα των  $\chi$  βρίσκονται οι τιμές  $M_{C_4}$  και στον άξονα των  $\psi$  οι τιμές των  $M_{C_4}'$ . Το μέσο σχετικό σφάλμα είναι  $\bar{R} = 5.87\%$ .

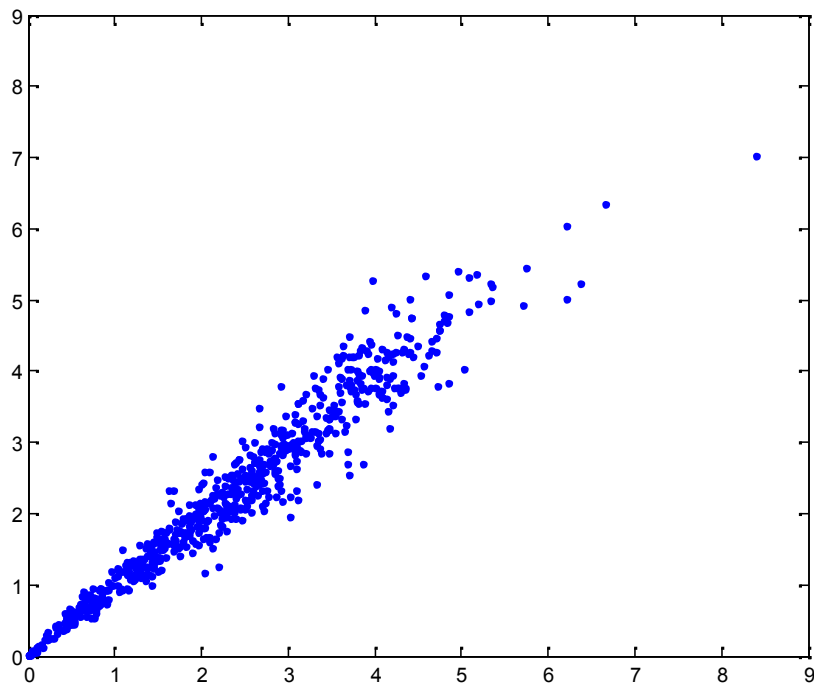


**Εικόνα 4.8 :** Σημεία ελέγχου (test). Στον άξονα των  $x$  βρίσκονται οι τιμές  $M_{C_4}$  και στον άξονα των  $y$  οι τιμές των  $M_{C_4}'$ . Το μέσο σχετικό σφάλμα είναι  $\bar{R} = 6.04\%$ .

Το μέσο σχετικό σφάλμα της περιεκτικότητας των ρευστών σε  $C_4$  επί του συνόλου των σημείων είναι  $\bar{R} = 3.15\%$ .

Υπενθυμίζεται πως οι προβλέψεις του μοντέλου για την επί τοις εκατό κατά μάζα περιεκτικότητα των ρευστών σε  $C_5$  υπολογίζονται από τη σχέση (4.6) ως εξής :

$$M_{C_5}' = M_{C_{3-5}} - M_{C_3}' - M_{C_4}'$$



**Εικόνα 4.9 :** Σύνολο σημείων. Στον άξονα των  $x$  βρίσκονται οι τιμές  $M_{C_5}$  και στον άξονα των  $y$  οι τιμές των  $M_{C_5}'$ . Το μέσο σχετικό σφάλμα είναι  $\bar{R} = 2.37\%$ .

- **3<sup>ο</sup> νευρωνικό δίκτυο :**

Ως έξοδος του τρίτου νευρωνικού δικτύου που εφαρμόζεται ορίζεται το διάνυσμα  $Y_3$  :

$$Y_3 = \left( \frac{M_{C_{6-8}}}{M_{C_{6^+}}} \right)$$

Κατά αυτόν τον τρόπο το τρίτο νευρωνικό δίκτυο εκπαιδεύεται για να μάθει να υπολογίζει τον λόγο του κλάσματος  $C_{6-8}$ , στο κλάσμα  $C_{6^+}$ . Μετά το πέρας της εκπαίδευσής του το δίκτυο θα δίνει ως έξοδο το  $Y_3'$ , που θα είναι η βέλτιστη προσέγγιση του  $Y_3$ . Έτσι η πρόβλεψη του μοντέλου για το  $M_{C_{6-8}}'$ , που είναι η

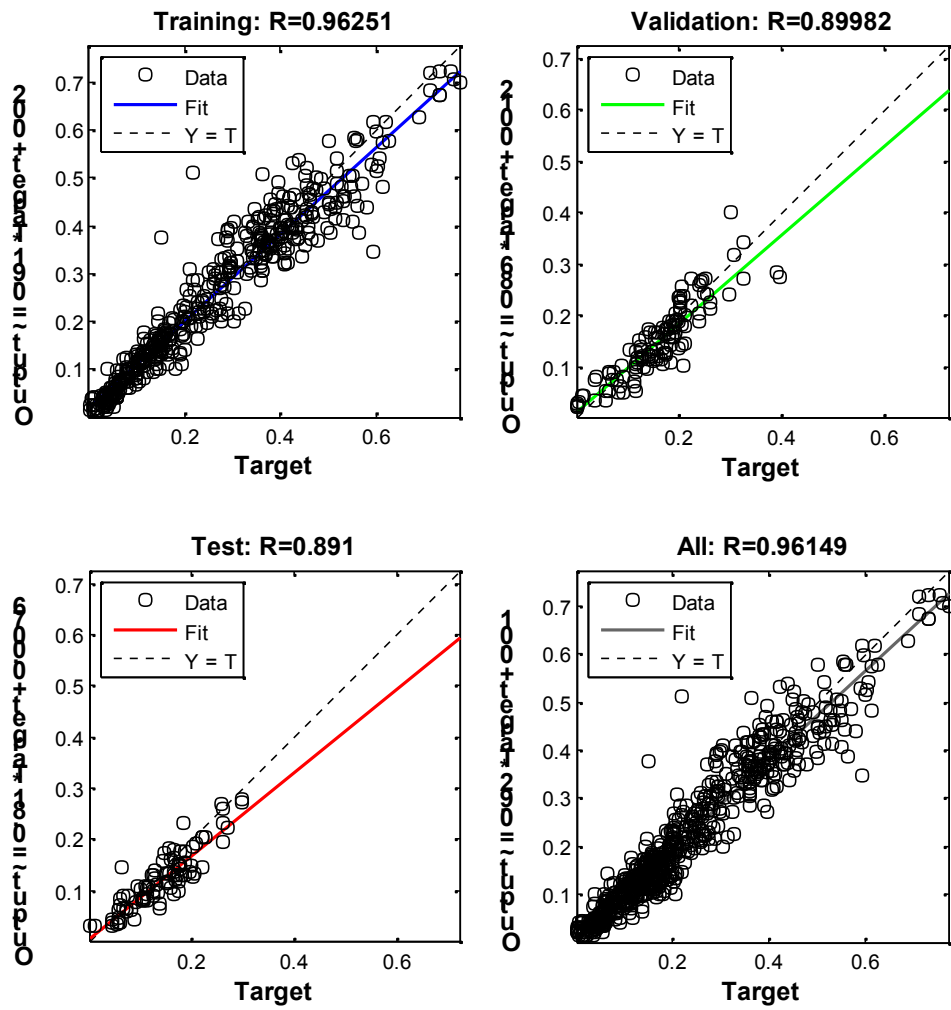
τιμή της επί τοις εκατό κατά μάζα περιεκτικότητας του ρευστού στο κλάσμα  $C_{6-8}$ , θα προκύπτει ως εξής :

$$M_{C_{6-8}}' = Y_3' * M_{C_{6+}} \quad (4.7)$$

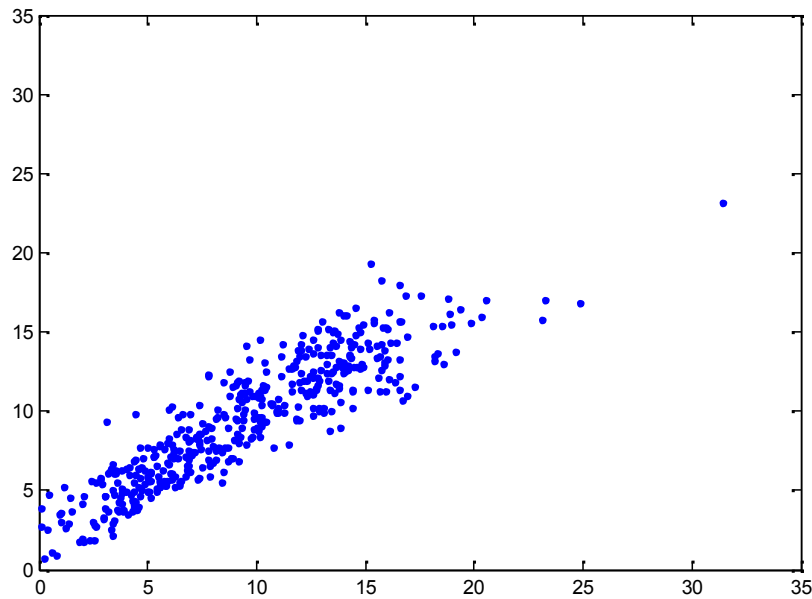
Επισημαίνεται στο σημείο αυτό πως ο λόγος για τον οποίον δεν επιλέγεται απευθείας ως έξοδος του δικτύου η  $M_{C_{6-8}}$ , που είναι η επί τοις εκατό κατά μάζα περιεκτικότητα του ρευστού σε  $C_{6-8}$ , είναι για να αποφευχθεί το ενδεχόμενο απόκλισης της τιμής του  $M_{C_{6+}}'$  από το  $M_{C_{6+}}$ .

Το παρόν Νευρωνικό δίκτυο πρόσθιας τροφοδότησης ενός κρυφού επιπέδου, εκτέλεσε την εκπαίδευσή του με τον αλγόριθμο μάθησης Levenberg-Marquardt. Υπενθυμίζεται πως το σύνολο χαρακτηρισμένων παραδειγμάτων επί του οποίου εκπαιδεύτηκε το παρόν δίκτυο, είναι το  $\mathfrak{S}_3\{X(X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8), Y_3\}$  που αποτελείται από 660 δείγματα ρευστών. Από το σύνολο των 660 δειγμάτων, 450 δείγματα χρησιμοποιήθηκαν ως σημεία εκπαίδευσης (training data), 120 δείγματα χρησιμοποιήθηκαν ως σημεία επιβεβαίωσης (validation data) και ως τεστ σημεία (test data) χρησιμοποιήθηκαν τα υπόλοιπα 90. Το κρυφό επίπεδο του δικτύου αποτελείται από 18 νευρώνες.

Αποτελέσματα 3<sup>ου</sup> Νευρωνικού Δικτύου :

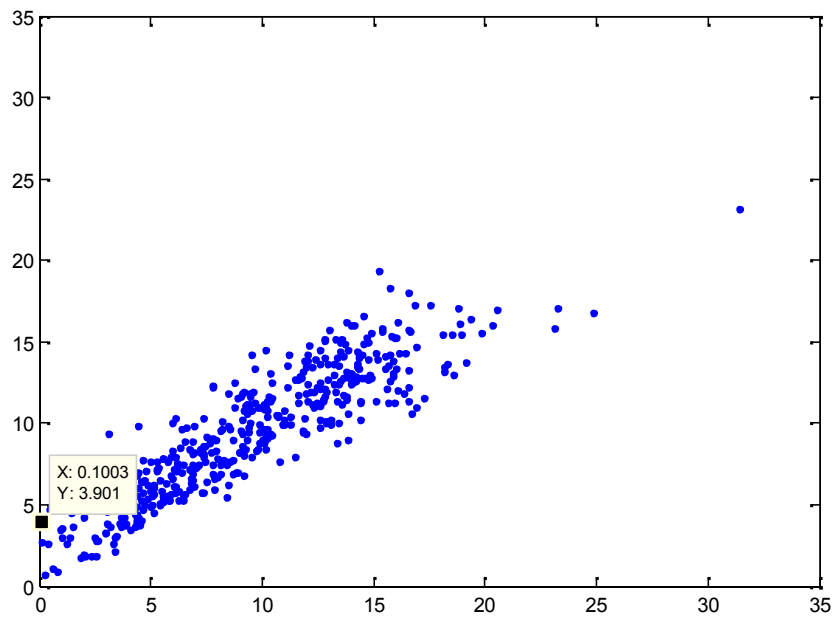


Εικόνα 4.10 : Χαρακτηριστικά εκπαίδευσης του τρίτου νευρωνικού δικτύου.



**Εικόνα 4.11 :** Σημεία εκπαίδευσης (training). Στον άξονα των  $\chi$  βρίσκονται οι τιμές  $M_{C_{6-8}}$  και στον άξονα των  $\psi$  οι τιμές των  $M_{C_{6-8}}'$ .

Όπως φαίνεται στη εικόνα 4.1 για τα δεδομένα εκπαίδευσης, οι προβλέψεις του δικτύου για την επί τοις εκατό κατά μάζα περιεκτικότητα των ρευστών στο κλάσμα  $C_{6-8}$  παρουσιάζουν μεγάλη σχετική απόκλιση από τις πραγματικές, στις πολύ χαμηλές τιμές (0-1% στον άξονα των  $\chi$ ). Υπάρχουν δηλαδή κάποια σημεία τα οποία το δίκτυο δυσκολεύεται να μάθει. Αυτό σημαίνει πως τα σημεία αυτά αποκλίνουν σημαντικά από μια γενική τάση του συνόλου των παραδειγμάτων. Πράγματι όπως προκύπτει από τη βάση δεδομένων τα σημεία αυτά ανήκουν σε μια ομάδα ρευστών των οποίων η χημική σύσταση παρουσιάζει την εξής ιδιαιτερότητα : συγκεντρώνει το μεγαλύτερο μέρος (περί το 80%) στα ακραία συστατικά  $C_1$  και  $C_{12+}$ . Μια τέτοιου τύπου γραμμομοριακή σύσταση παραπέμπει σε ρευστά που είναι προϊόντα βιοαποδόμησης (biodegradation), ρευστά δηλαδή των οποίων τα ενδιάμεσα συστατικά έχουν καταναλωθεί από μικροοργανισμούς. Ακολουθεί παράδειγμα :



**Εικόνα 4.12 :** Επανάληψη της εικόνας 4.11. Σημεία εκπαίδευσης (training). Στον άξονα των  $x$  βρίσκονται οι τιμές  $M_{C_{6-8}}$  και στον άξονα των  $y$  οι τιμές των  $M_{C_{6-8}}'$ . Το μαύρο τετράγωνο αποτελεί σημείο ιδιαίτερου ενδιαφέροντος.

Το σημείο που έχει επιλεγεί στο διάγραμμα της εικόνας 4.12 αντιστοιχεί στο παράδειγμα 312 της βάσης δεδομένων, και όπως φαίνεται από την αναλυτική του σύσταση πρόκειται για προϊόν βιοαποδόμησης :

$$CO_2 = 3.6831 \%$$

$$C_1 = 24.0562 \% \quad \leftarrow$$

$$C_2 = 0.2182 \%$$

$$C_3 = 0.0821 \%$$

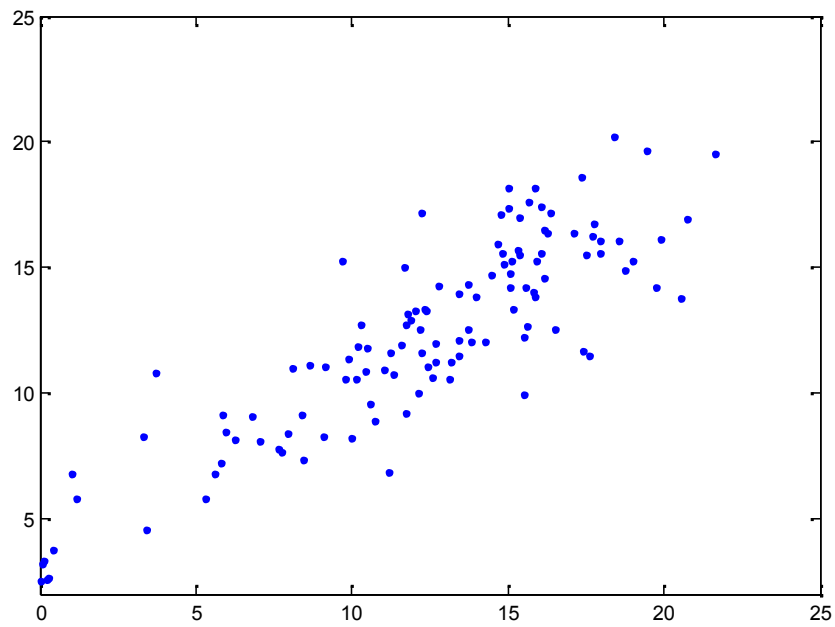
$$C_4 = 0.1621 \%$$

$$C_5 = 0.1141 \%$$

$$C_{6-8} = 0.2042 \%$$

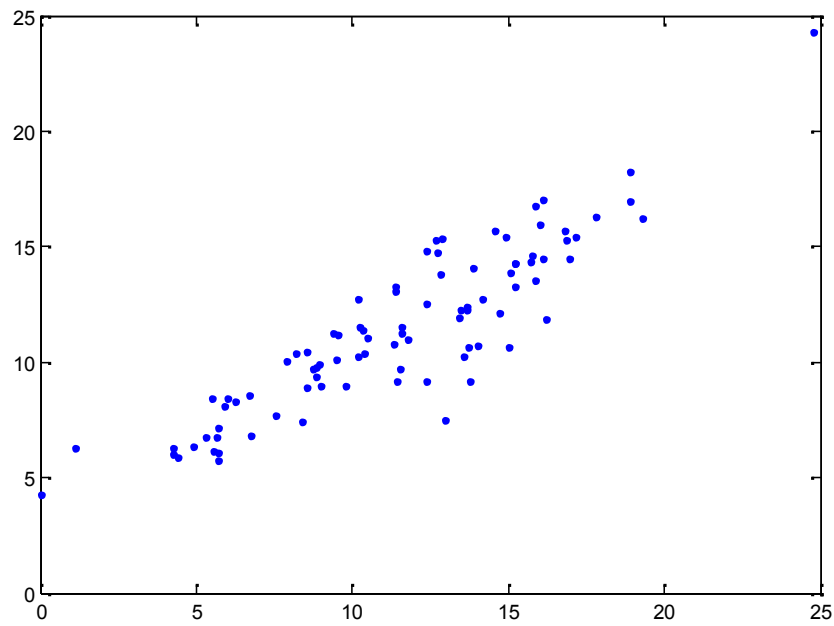
$$C_{9-11} = 2.3840 \%$$

$$C_{12^+} = 69.0960 \% \quad \leftarrow$$

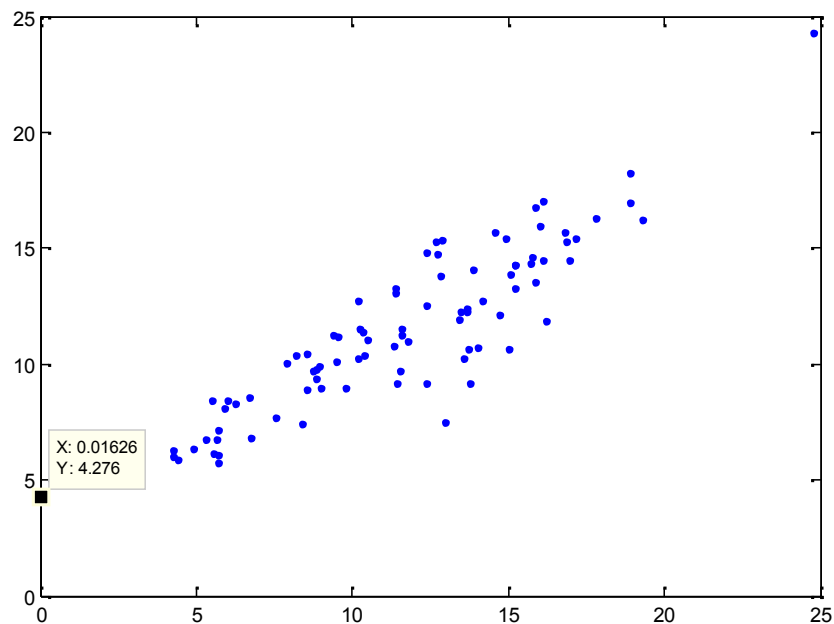


**Εικόνα 4.13 :** Σημεία επιβεβαίωσης (validation). Στον άξονα των  $\chi$  βρίσκονται οι τιμές  $M_{C_{6-8}}$  και στον άξονα των  $\psi$  οι τιμές των  $M_{C_{6-8}}'$ .





**Εικόνα 4.14 :** Σημεία ελέγχου (test). Στον άξονα των  $\chi$  βρίσκονται οι τιμές  $M_{C_{6-8}}$  και στον άξονα των  $\psi$  οι τιμές των  $M_{C_{6-8}}$ .



**Εικόνα 4.15 :** Επανάληψη εικόνας 4.14. Σημεία ελέγχου (test). Στον άξονα των  $\chi$  βρίσκονται οι τιμές  $M_{C_{6-8}}$  και στον άξονα των  $\psi$  οι τιμές των  $M_{C_{6-8}}$ . Με μαύρο τετράγωνο αναπαρίστανται το παράδειγμα 579.

Αναλυτική σύσταση του παραδείγματος 579 από τη βάση δεδομένων :

$$CO_2 = 0.1303 \%$$

$$C_1 = 24.7354 \% \quad \leftarrow$$

$$C_2 = 0.0481 \%$$

$$C_3 = 0.0220 \%$$

$$C_4 = 0.0321 \%$$

$$C_5 = 0.0160 \%$$

$$C_{6-8} = 0.0401 \%$$

$$C_{9-11} = 1.9041 \%$$

$$C_{12^+} = 73.0718 \% \quad \leftarrow$$

- **4<sup>ο</sup> νευρωνικό δίκτυο :**

Ως έξοδος του τέταρτου νευρωνικού δικτύου που εφαρμόζεται ορίζεται το διάνυσμα  $Y_4$  :

$$Y_4 = \left( \frac{M_{C_{9-11}}}{M_{C_{9^+}}} \right)$$

Όπου :

$$M_{C_{9^+}} = M_{C_9} + M_{C_{10}} + M_{C_{11}} + M_{C_{12^+}}$$

Κατά αυτόν τον τρόπο το τέταρτο νευρωνικό δίκτυο εκπαιδεύεται για να μάθει να υπολογίζει τον λόγο του κλάσματος  $C_{9-11}$ , στο κλάσμα  $C_{9+}$ . Μετά το πέρας της εκπαίδευσής του το δίκτυο θα δίνει ως έξοδο το  $Y_3'$ , που θα είναι η βέλτιστη προσέγγιση του  $Y_3$ . Έτσι ο υπολογισμός του  $M_{C_{9-11}}'$ , που είναι η τιμή της επί τοις εκατό κατά μάζα περιεκτικότητας του ρευστού στο κλάσμα  $C_{9-11}$ , θα προκύπτει ως εξής :

$$M_{C_{9-11}}' = Y_3' * M_{C_{9+}}' \quad (4.8)$$

Όπου :

$$M_{C_{9+}}' = M_{C_{6+}} - M_{C_{6-8}}'$$

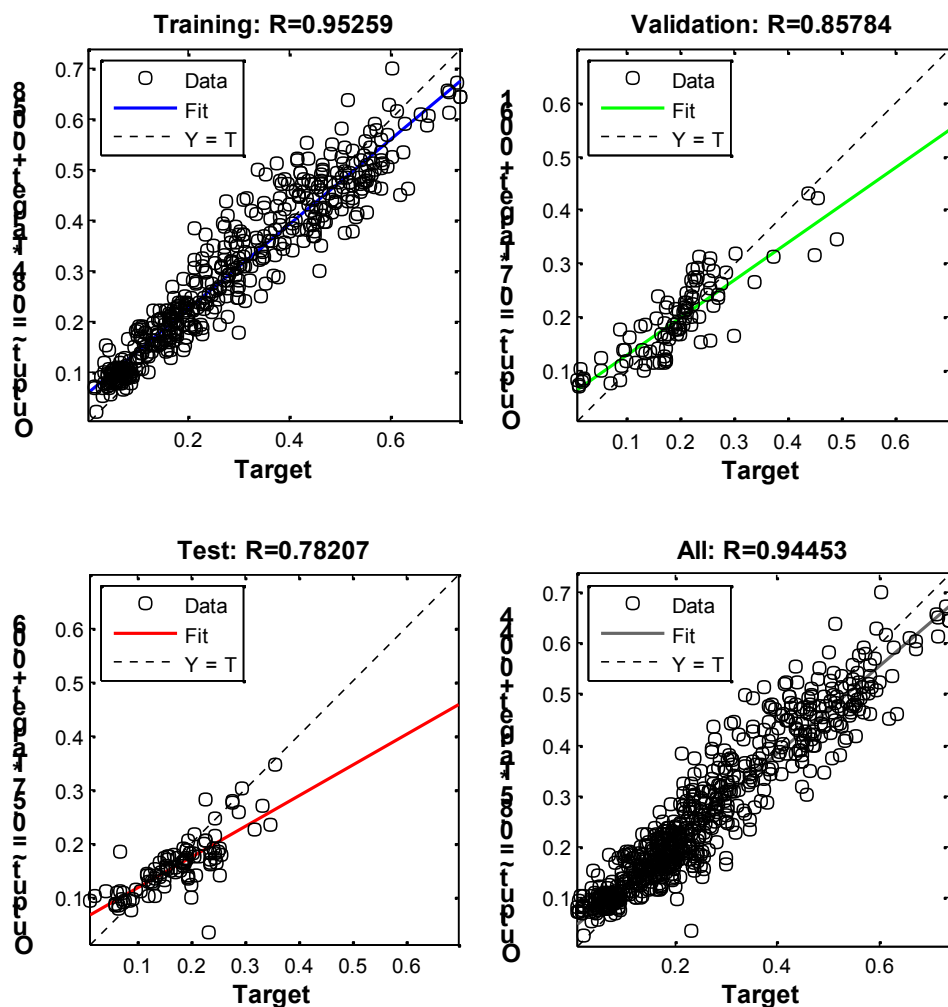
Επισημαίνεται στο σημείο αυτό πως επιλέγεται ως διάνυσμα εξόδου ο λόγος  $\frac{M_{C_{9-11}}}{M_{C_{9+}}}$  αντί του λόγου  $\frac{M_{C_{9-11}}}{M_{C_{6+}}}$  για να αποφευχθεί κατά τον υπολογισμό του  $M_{C_{12+}}'$ , που είναι η πρόβλεψη του μοντέλου για την επί τοις εκατό κατά μάζα περιεκτικότητα του ρευστού σε  $C_{12+}$ , το ενδεχόμενο της αρνητικής τιμής. Μετά το πέρας της εκπαίδευσης του τρίτου και του τέταρτου νευρωνικού δικτύου το  $M_{C_{12+}}'$  υπολογίζεται ως εξής :

$$M_{C_{12+}}' = M_{C_{6+}} - M_{C_{6-8}}' - M_{C_{9-11}}' \quad (4.9)$$

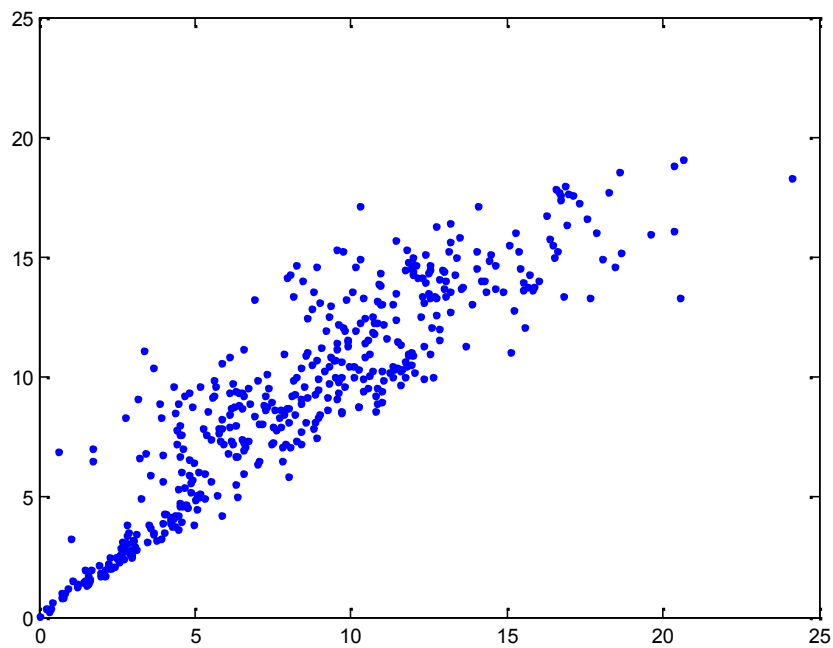
Το παρόν Νευρωνικό είναι κι αυτό δίκτυο πρόσθιας τροφοδότησης ενός κρυφού επιπέδου και εκτέλεσε την εκπαίδευσή του με τον αλγόριθμο μάθησης Levenberg-Marquardt. Υπενθυμίζεται πως το σύνολο χαρακτηρισμένων παραδειγμάτων επί του οποίου εκπαιδεύτηκε το παρόν δίκτυο, είναι το  $\mathfrak{S}_4 \{X(X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8), Y_4\}$  που αποτελείται από 660 δείγματα

ρευστών. Από το σύνολο των 660 δειγμάτων, 462 δείγματα χρησιμοποιήθηκαν ως σημεία εκπαίδευσης (training data), 99 δείγματα χρησιμοποιήθηκαν ως σημεία επιβεβαίωσης (validation data) και ως τεστ σημεία (test data) χρησιμοποιήθηκαν τα υπόλοιπα 99. Το κρυφό επίπεδο του δικτύου αποτελείται από 20 νευρώνες.

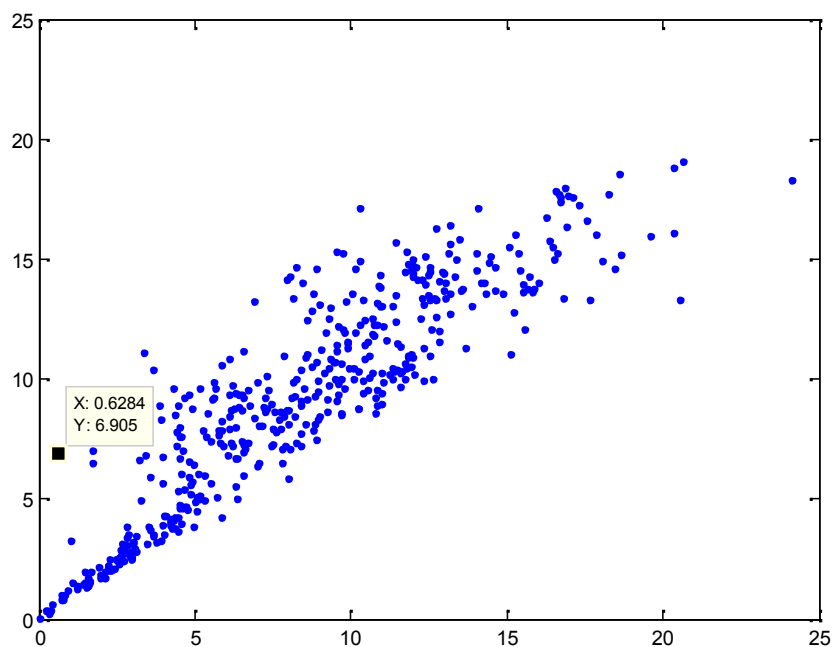
Αποτελέσματα 4<sup>ου</sup> Νευρωνικού Δικτύου :



Εικόνα 4.16 : Χαρακτηριστικά εκπαίδευσης του τέταρτου νευρωνικού δικτύου.



**Εικόνα 4.17 :** Σημεία εκπαίδευσης (test). Στον άξονα των  $\chi$  βρίσκονται οι τιμές  $M_{C_{9-11}}$  και στον άξονα των  $\psi$  οι τιμές των  $M_{C_{9-11}}$ . Το μέσο σχετικό σφάλμα είναι  $\bar{R} = 18.23\%$ .



**Εικόνα 4.18 :** Επανάληψη εικόνας 4.17. Σημεία εκπαίδευσης (training). Στον άξονα των  $\chi$  βρίσκονται οι τιμές  $M_{C_{9-11}}$  και στον άξονα των  $\psi$  οι τιμές των  $M_{C_{9-11}}$ . Με μαύρο τετράγωνο αναπαρίστανται το παράδειγμα 347.

Αναλυτική σύσταση του παραδείγματος 347 της βάσης δεδομένων που αποδεικνύει πως το παράδειγμα αυτό είναι προϊόν βιοαποδόμησης :

$$CO_2 = 0.8411 \%$$

$$C_1 = 24.8708 \% \quad \leftarrow$$

$$C_2 = 0.1302 \%$$

$$C_3 = 0.0581 \%$$

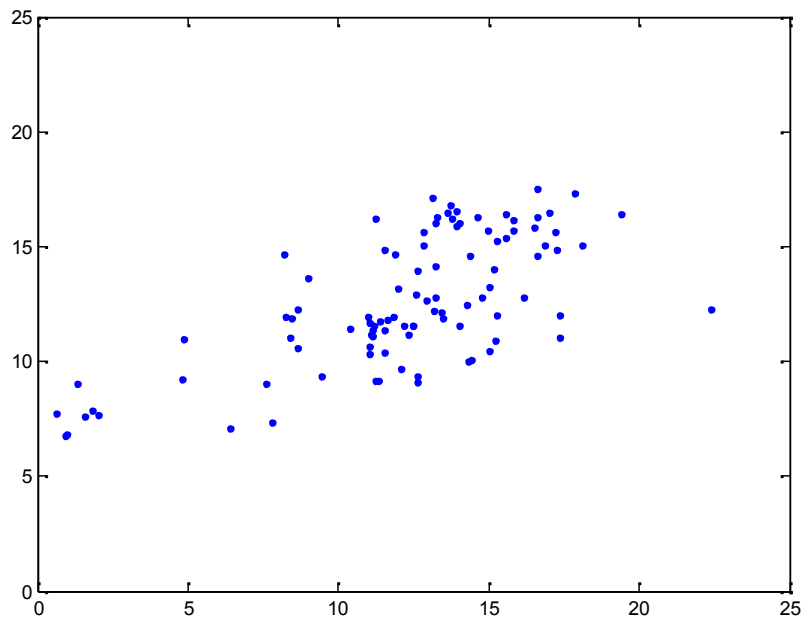
$$C_4 = 0.1282 \%$$

$$C_5 = 0.1582 \%$$

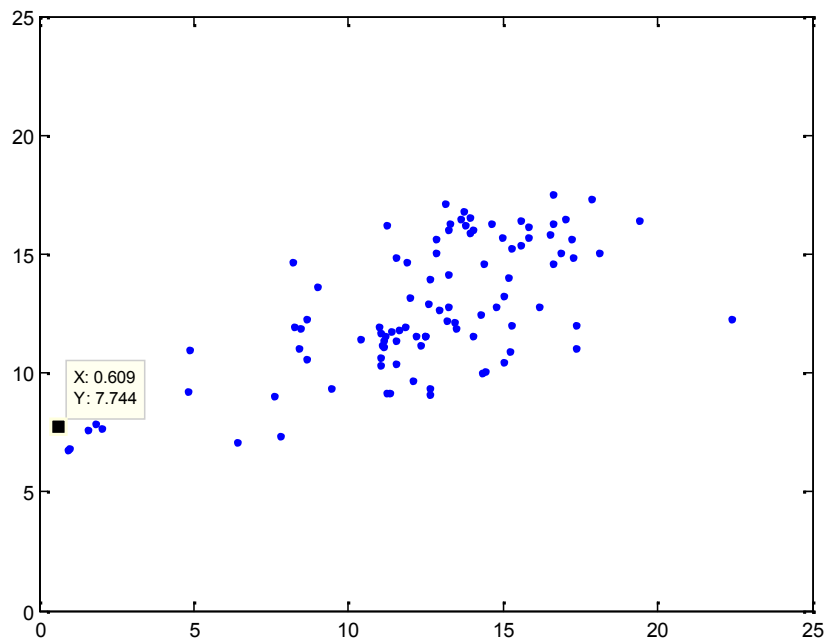
$$C_{6-8} = 0.2323 \%$$

$$C_{9-11} = 1.0734 \%$$

$$C_{12^+} = 72.5077 \% \quad \leftarrow$$



**Εικόνα 4.19 :** Σημεία επιβεβαίωσης (validation). Στον άξονα των  $\chi$  βρίσκονται οι τιμές  $M_{C_{9-11}}$  και στον άξονα των  $\psi$  οι τιμές των  $M_{C_{9-11}}$ .



**Εικόνα 4.20 :** Σημεία επιβεβαίωσης (validation). Στον άξονα των  $\chi$  βρίσκονται οι τιμές  $M_{C_{9-11}}$  και στον άξονα των  $\psi$  οι τιμές των  $M_{C_{9-11}}$ . Το μαύρο τετράγωνο εκπροσωπεί το παράδειγμα 471.

Αναλυτική σύσταση του παραδείγματος 471 :

$$CO_2 = 0.0460 \%$$

$$C_1 = 25.5760 \% \quad \leftarrow$$

$$C_2 = 0.1840 \%$$

$$C_3 = 0.0620 \%$$

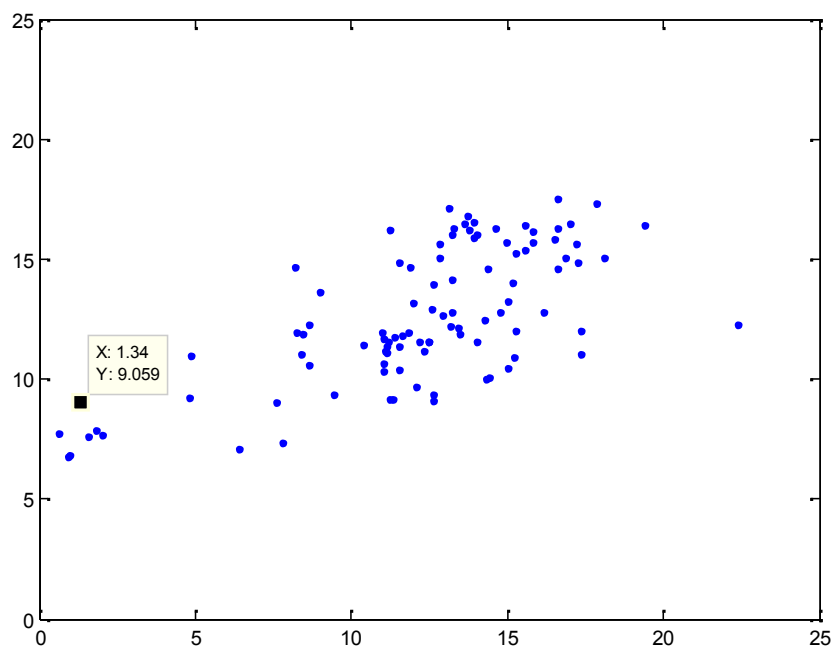
$$C_4 = 0.0880 \%$$

$$C_5 = 0.0540 \%$$

$$C_{6-8} = 0.3100 \%$$

$$C_{9-11} = 1.0780 \%$$

$$C_{12^+} = 72.6020 \% \quad \leftarrow$$



**Εικόνα 4.21 :** Σημεία επιβεβαίωσης (validation). Στον άξονα των  $x$  βρίσκονται οι τιμές  $M_{C_{9-11}}$



και στον άξονα των  $\psi$  οι τιμές των  $M_{C_{9-11}}$ . Το μαύρο τετράγωνο εκπροσωπεί το παράδειγμα 543.

Αναλυτική σύσταση του παραδείγματος 543 :

$$CO_2 = 0.0460 \%$$

$$C_1 = 25.5760 \% \quad \leftarrow$$

$$C_2 = 0.1840 \%$$

$$C_3 = 0.0620 \%$$

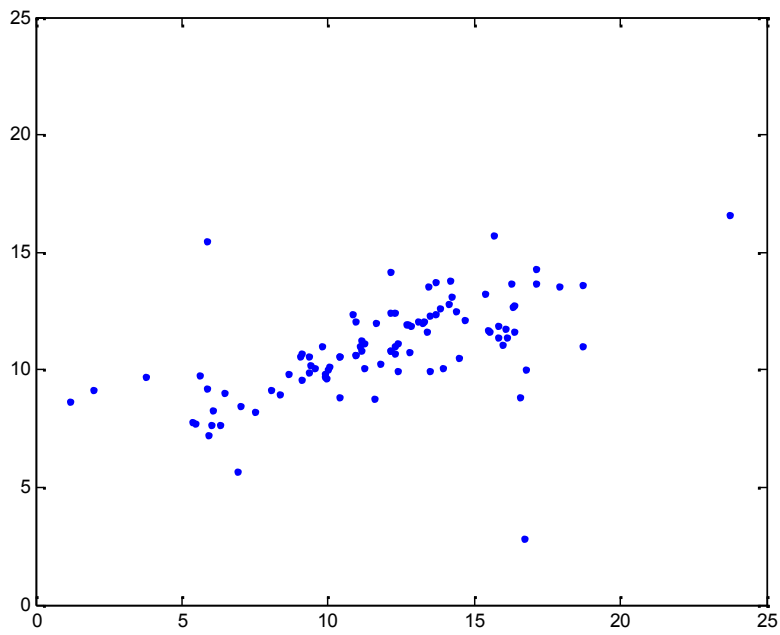
$$C_4 = 0.0880 \%$$

$$C_5 = 0.0540 \%$$

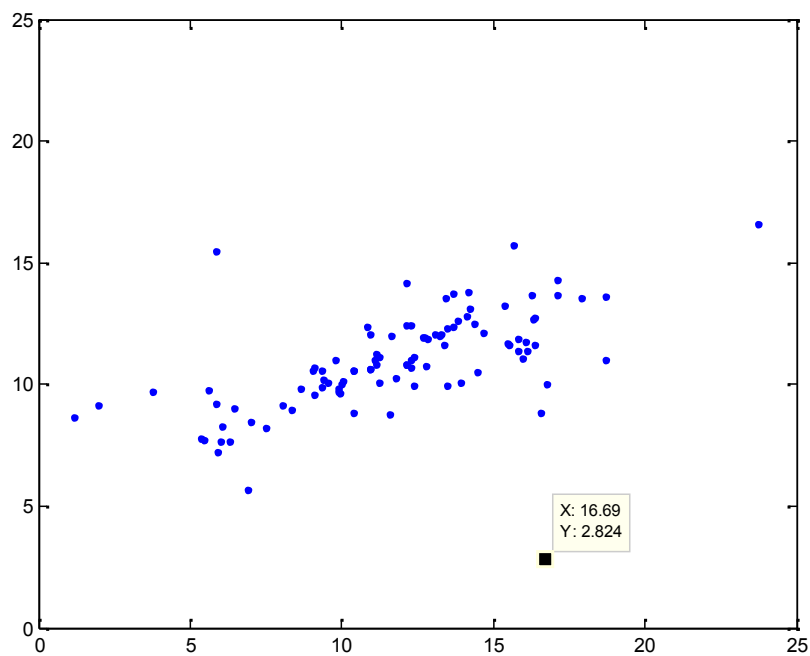
$$C_{6-8} = 0.3100 \%$$

$$C_{9-11} = 1.0780 \%$$

$$C_{12^+} = 72.6020 \% \quad \leftarrow$$



**Εικόνα 4.22 :** Σημεία ελέγχου (test). Στον άξονα των  $\chi$  βρίσκονται οι τιμές  $M_{C_{9-11}}$  και στον άξονα των  $\psi$  οι τιμές των  $M_{C_{9-11}}$ .



**Εικόνα 4.23 :** Σημεία ελέγχου (test). Στον άξονα των  $\chi$  βρίσκονται οι τιμές  $M_{C_{9-11}}$  και στον άξονα των  $\psi$  οι τιμές των  $M_{C_{9-11}}$ . Το μαύρο τετράγωνο εκπροσωπεί το παράδειγμα 543.

Η αναλυτική σύσταση του παραδείγματος 543 είναι η εξής :

$$CO_2 = 20.2397 \% \quad \leftarrow$$

$$C_1 = 6.1484 \%$$

$$C_2 = 0.9900 \%$$

$$C_3 = 1.2169 \%$$

$$C_4 = 1.7346 \%$$

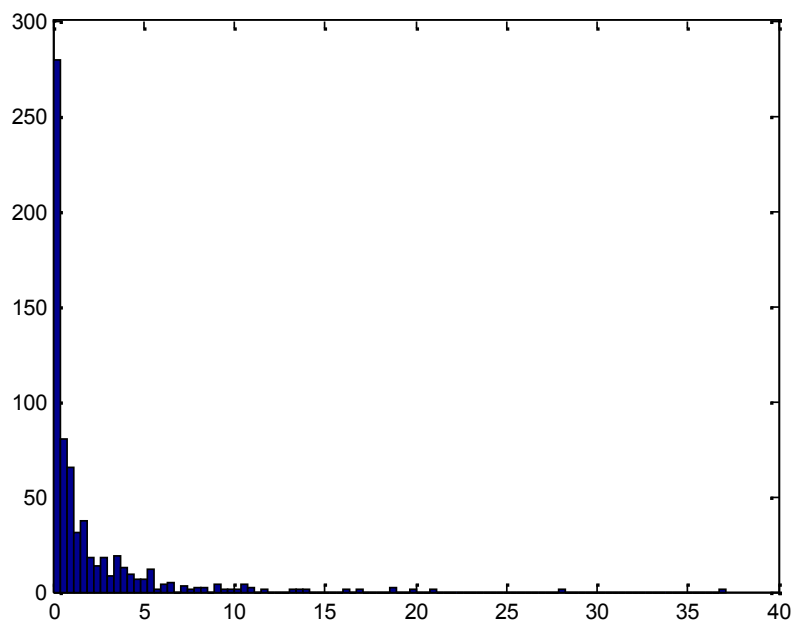
$$C_5 = 4.1952 \%$$

$$C_{6-8} = 24.9051 \%$$

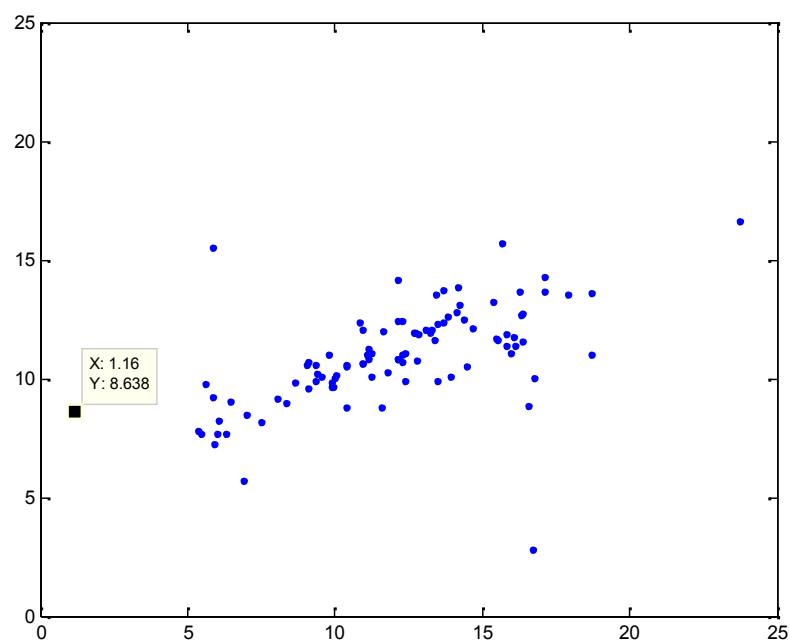
$$C_{9-11} = 17.5254 \%$$

$$C_{12^+} = 23.044 \%$$

Όπως φαίνεται από την αναλυτική σύσταση το παράδειγμα 542 έχει πολύ υψηλή περιεκτικότητα σε  $CO_2$ , και όπως αποδεικνύεται από το κάτωθι ιστόγραμμα των επί τοις εκατό κατά mole περιεκτικοτήτων των ρευστών σε  $CO_2$ , αποτελεί εξαίρεση, και άρα το δίκτυο δεν είναι δυνατόν να κάνει αξιόπιστη πρόβλεψη.



**Εικόνα 4.24 :** Ιστογράμμο. Στον άξονα των  $\chi$  : η επί τοις εκατό κατά τοιλε περιεκτικότητα του κάθε ρευστού σε  $CO_2$ . Στον άξονα των  $\psi$  : το πλήθος των παραδειγμάτων.



**Εικόνα 4.25 :** Σημεία ελέγχου (test). Στον άξονα των  $\chi$  βρίσκονται οι τιμές  $M_{C_{9-11}}$  και στον άξονα των  $\psi$  οι τιμές των  $M_{C_{9-11}}'$ . Το μαύρο τετράγωνο εκπροσωπεί το παράδειγμα 579.

Η αναλυτική σύσταση του παραδείγματος 579 υποδεικνύει πως πρόκειται για προϊόν βιοαποδόμησης :

$$CO_2 = 0.1303 \%$$

$$C_1 = 24.7354 \% \quad \leftarrow$$

$$C_2 = 0.0481 \%$$

$$C_3 = 0.0220 \%$$

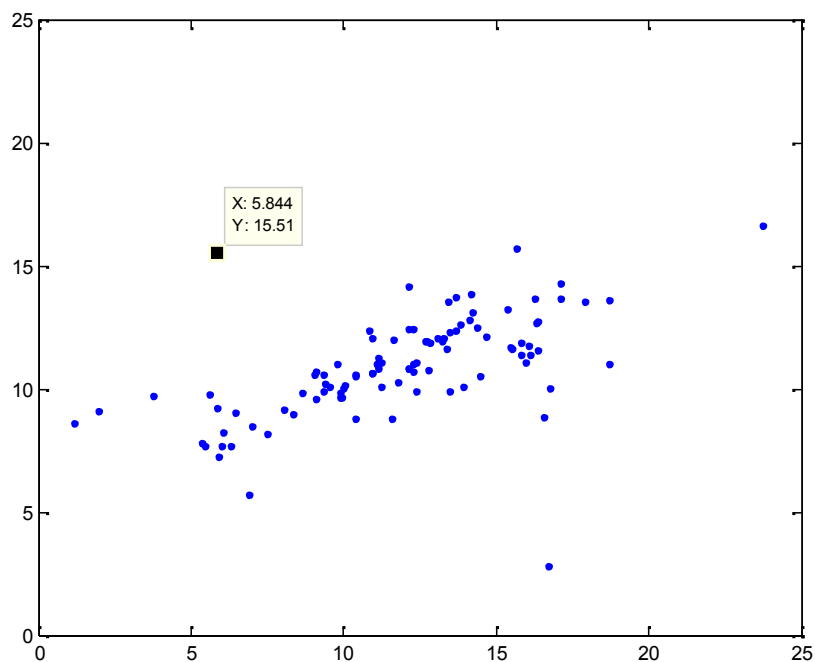
$$C_4 = 0.0321 \%$$

$$C_5 = 0.0160 \%$$

$$C_{6-8} = 0.0401 \%$$

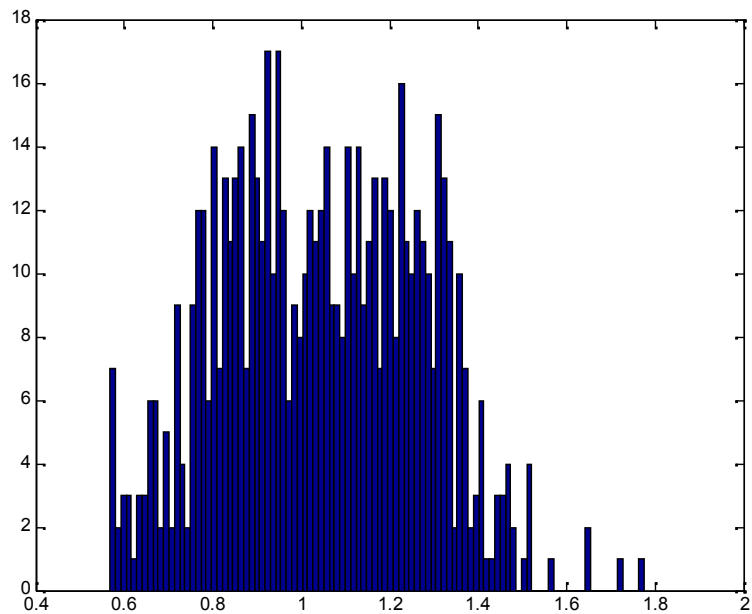
$$C_{9-11} = 0.0401 \%$$

$$C_{12^+} = 73.0718 \% \quad \leftarrow$$



**Εικόνα 4.26 :** Σημεία ελέγχου (test). Στον άξονα των  $\chi$  βρίσκονται οι τιμές  $M_{C_{9-11}}$  και στον άξονα των  $\psi$  οι τιμές των  $M_{C_{9-11}}$ . Το μαύρο τετράγωνο εκπροσωπεί το παράδειγμα 623.

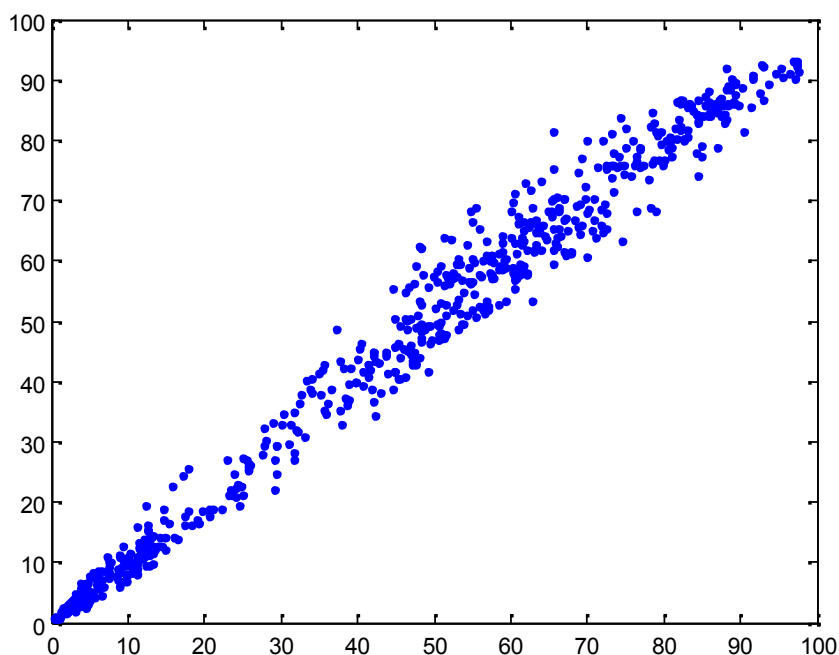
Η ιδιαιτερότητα που παρουσιάζει το παράδειγμα 623 είναι η τιμή της σχετικής πυκνότητας της αερίου φάσεως. Συγκεκριμένα  $S_g^{(623)} = 1.77$ , και όπως φαίνεται από το παρακάτω ιστόγραμμα των σχετικών πυκνοτήτων της αερίου φάσεως των ρευστών, το δείγμα 623 αποτελεί εξαίρεση καθώς εμφανίζει τη μέγιστη τιμή.



**Εικόνα 4.27 :** Ιστογράμμο. Στον άξονα των  $\chi$  :  $S_g$ . Στον άξονα των  $\psi$  : το πλήθος των δειγμάτων της βάσης δεδομένων.

Υπενθυμίζεται πως οι προβλέψεις του μοντέλου για την επί τοις εκατό κατά μάζα περιεκτικότητα των ρευστών σε  $C_{12^+}$  υπολογίζονται από τη σχέση (4.9) ως εξής :

$$M_{C_{12^+}} = M_{C_{6^+}} - M_{C_{6-8}} - M_{C_{9-11}}$$



**Εικόνα 4.28 :** Σύνολο σημείων. Στον άξονα των  $\chi$  βρίσκονται οι τιμές  $M_{C_{12^+}}$  και στον άξονα των  $\psi$  οι τιμές των  $M_{C_{12^+}}$ . Το μέσο σχετικό σφάλμα είναι  $\bar{R} = 1.14\%$ .

Συνοψίζοντας τις εργασίες που εκτελέστηκαν έως αυτό το σημείο, υπενθυμίζεται πως μετά κι από την εφαρμογή των πρώτων τεσσάρων νευρωνικών δικτύων οι πληροφορίες που θα διατίθενται πλέον είναι οι επί τοις εκατό κατά μάζα περιεκτικότητες των ρευστών σε  $CO_2$ ,  $C_1$ ,  $C_2$ ,  $C_3$ ,  $C_4$ ,  $C_5$ ,  $C_{6-8}$ ,  $C_{9-11}$ , και  $C_{12^+}$ , και μένει ως τελευταίο βήμα για την διεκπεραίωση του τελικού στόχου της εργασίας, η μετατροπή των περιεκτικότητων αυτών, σε περιεκτικότητες κατά mole. Για την μετατροπή αυτή, λείπει η πληροφορία για το μέσο μοριακό βάρος του  $C_{12^+}$ , έτσι το επόμενο νευρωνικό δίκτυο που αναπτύχθηκε στόχο έχει να παράσχει τη πληροφορία αυτή.



## 5<sup>ο</sup> Νευρωνικό Δίκτυο :

Το πέμπτο και τελευταίο νευρωνικό δίκτυο εκπαιδεύεται επί του συνόλου χαρακτηρισμένων παραδειγμάτων  $\mathfrak{S}_5 \{X(X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8), Y_5\}$ , όπου :

$$Y_5 = MR_{C_{6^+}} \frac{gr}{mole}$$

Όπου το  $MR_{C_{6^+}}$  είναι το μέσο μοριακό βάρος του κλάσματος  $C_{6^+}$  και υπολογίστηκε από την αρχική βάση δεδομένων ως εξής :

$$MR_{C_{6^+}} = \sum_{i=6}^{12^+} \left( \frac{N_{C_i}}{N_{C_{6^+}}} * MR_{C_i} \right) \quad (4.10)$$

Όπου :

$N_{C_i}$  = το γραμμομοριακό κλάσμα του  $C_i$  στο ρευστό.

$N_{C_{6^+}}$  = το γραμμομοριακό κλάσμα του κλάσματος  $C_{6^+}$  στο ρευστό.

$MR_{C_i}$  = το μοριακό βάρος του συστατικού  $C_i$  για τιμές του  $i = 6 \sim 11$ , ενώ όταν  $i = 12^+$ , το μέσο μοριακό βάρος του κλάσματος  $C_{12^+}$ .

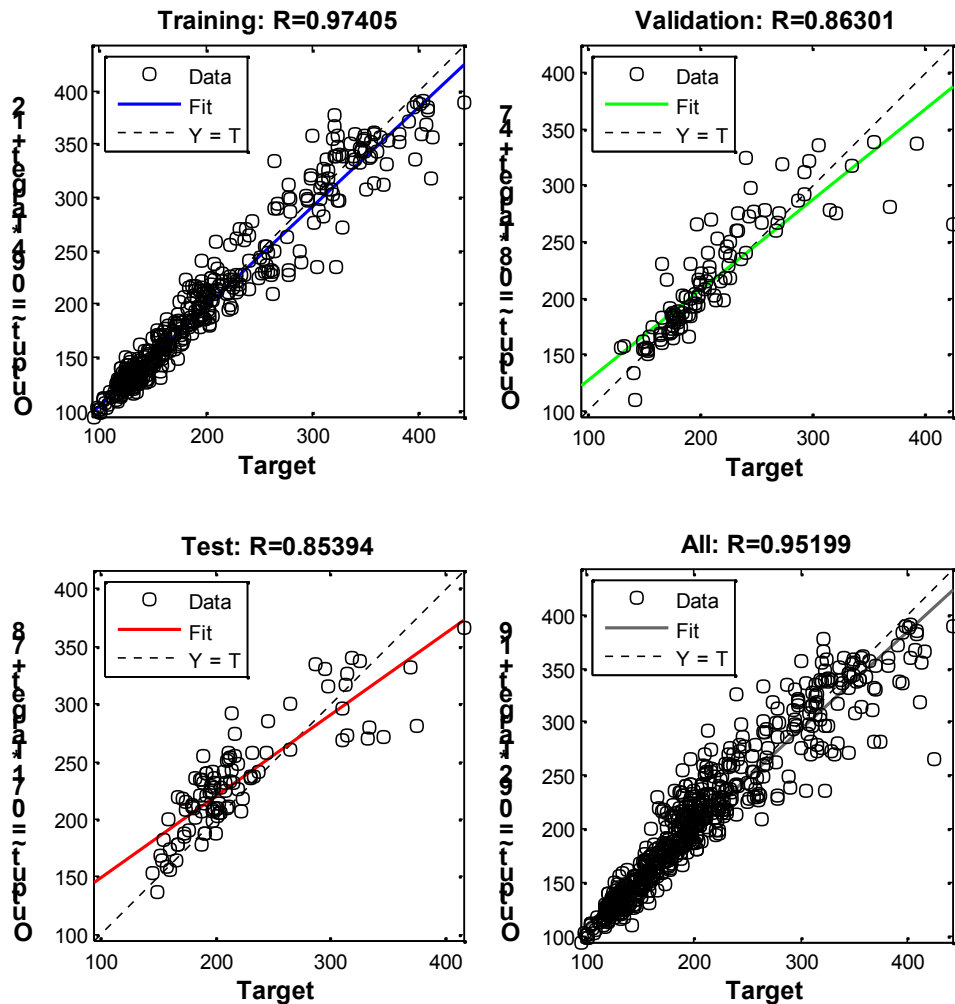
Ο λόγος που επιλέγεται ως μαθησιακός στόχος του δικτύου το μέσο μοριακό βάρος του κλάσματος  $C_{6^+}$  έναντι αυτού του κλάσματος  $C_{12^+}$ , είναι το γεγονός

πως κατά τον υπολογισμό του μέσου μοριακού βάρους του  $C_{12^+}$  από την εφαρμογή του ισοζυγίου μάζας επί του συνόλου του κάθε ρευστού της αρχικής βάσης δεδομένων, προέκυπταν αποτελέσματα χωρίς φυσική σημασία (π.χ. μέσο μοριακό βάρος του  $C_{12^+}$  :  $19000 \frac{gr}{mole}$ ). Αυτό συνέβαινε σε περιπτώσεις όπου το γραμμομοριακό κλάσμα του  $C_{12^+}$  ήταν της τάξης του 0.02%. Μία τέτοιας ποιότητας πληροφορία λοιπόν θα δυσχέραινε την εκπαίδευση του δικτύου. Έτσι κρίθηκε ως ασφαλέστερη οδός η εκπαίδευση δικτύου για την πρόβλεψη του μέσου μοριακού βάρους του  $C_{6^+}$ , ώστε με την εφαρμογή του ισοζυγίου μάζας στις προβλέψεις του δικτύου αυτού, να υπολογίζεται αμέσως ο αριθμός των moles του  $C_{12^+}$  που περιέχονται σε 100 γραμμάρια ρευστού.

Στόχος του παρόντος δικτύου είναι η βέλτιστη προσέγγιση του μέσου μοριακού βάρους του κλάσματος  $C_{6^+}$ , το  $MR_{C_{6^+}}$ . Με την εφαρμογή του ισοζυγίου μάζας στο προβλεπόμενο από το δίκτυο  $MR_{C_{6^+}}$ , θα υπολογιστούν τα moles του κλάσματος  $C_{12^+}$  ώστε να ολοκληρωθεί το τελικό ζητούμενο της παρούσης εργασίας που είναι η γραμμομοριακή σύσταση των πετρελαϊκών ρευστών σε  $CO_2$ ,  $C_1$ ,  $C_2$ ,  $C_3$ ,  $C_4$ ,  $C_5$ ,  $C_{6-8}$ ,  $C_{9-11}$  και  $C_{12^+}$ .

Όπως και όλα τα προηγούμενα δίκτυα που εφαρμόστηκαν, έτσι και το παρόν δίκτυο είναι πρόσθιας τροφοδότησης ενός κρυφού επιπέδου και εκτέλεσε την εκπαίδευσή του με τον αλγόριθμο μάθησης Levenberg-Marquardt. Υπενθυμίζεται πως το σύνολο χαρακτηρισμένων παραδειγμάτων επί του οποίου εκπαιδεύτηκε το παρόν δίκτυο, είναι το  $\mathfrak{S}_5\{X(X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8), Y_5\}$  που αποτελείται από 660 δείγματα ρευστών. Από το σύνολο των 660 δειγμάτων, 462 δείγματα χρησιμοποιήθηκαν ως σημεία εκπαίδευσης (training data), 99 δείγματα χρησιμοποιήθηκαν ως σημεία επιβεβαίωσης (validation data) και ως τεστ σημεία (test data) χρησιμοποιήθηκαν τα υπόλοιπα 99. Το κρυφό επίπεδο του δικτύου αποτελείται από 20 νευρώνες.

Αποτελέσματα 5<sup>ου</sup> Νευρωνικού Δικτύου :



Εικόνα 4.29 : Χαρακτηριστικά εκπαίδευσης του πέμπτου νευρωνικού δικτύου.

Με μέσο σχετικό σφάλμα  $\bar{R}_{Y_5} = \frac{(Y_5 - Y_5')}{Y_5}$ , τα αποτελέσματα για την πρόβλεψη του μέσου μοριακού βάρους του κλάσματος  $C_{6+}$  είχαν ως εξής :

Για τα δεδομένα εκπαίδευσης (training) :  $\bar{R} = 0.84\%$

Για τα δεδομένα επιβεβαίωσης (validation) :  $\bar{R} = 3.45\%$

Για τα δεδομένα ελέγχου (test) :  $\bar{R} = 7.30\%$

#### 4.3 ΥΠΟΛΟΓΙΣΜΟΣ ΓΡΑΜΜΟΜΟΡΙΑΚΩΝ ΚΛΑΣΜΑΤΩΝ.

Δεδομένου πως μετά την εφαρμογή των προαναφερθέντων νευρωνικών δικτύων, είναι διαθέσιμη η επί τοις εκατό κατά μάζα περιεκτικότητα του κάθε ρευστού σε  $CO_2$ ,  $C_1$ ,  $C_2$ ,  $C_3$ ,  $C_4$ ,  $C_5$ ,  $C_{6-8}$ ,  $C_{9-11}$  και  $C_{12+}$ , η μετατροπή των περιεκτικοτήτων αυτών, πλην της περιεκτικότητας στο κλάσμα  $C_{12+}$ , σε κατά mole περιεκτικότητες γίνεται με την εφαρμογή της σχέσης (4.1) ως εξής :

$$N_i' = \frac{M_i'}{MR_i} \text{ moles}$$

Όπου :

$N_i'$  = τα moles του συστατικού ή κλάσματος συστατικών  $i$  που προβλέπεται από το μοντέλο πως περιέχονται σε 100 γραμμάρια ρευστού.

$M_i'$  = η επί τοις εκατό κατά μάζα περιεκτικότητα του ρευστού στο συστατικό ή κλάσμα συστατικών  $i$  όπως προβλέπεται από το μοντέλο.

$MR_i$  = το μοριακό βάρος του συστατικού  $i$ , ή το μέσο μοριακό βάρος του κλάσματος συστατικών  $i$ .

Στο στάδιο αυτό χρησιμοποιούνται ξανά τα μοριακά βάρη του κάθε συστατικού κατά Katz. Τα μέσα μοριακά βάρη των κλασμάτων  $C_{6-8}$  και  $C_{9-11}$  υπολογίζονται ως εξής :

$$MR_{C_{6-8}} = \bar{n}_{C_6} * MR_{C_6} + \bar{n}_{C_7} * MR_{C_7} + \bar{n}_{C_8} * MR_{C_8} \quad (4.11)$$

$$MR_{C_{9-11}} = \bar{n}_{C_9} * MR_{C_9} + \bar{n}_{C_{10}} * MR_{C_{10}} + \bar{n}_{C_{11}} * MR_{C_{11}} \quad (4.12)$$

Όπου :

$\bar{n}_{C_6}$  = η μέση τιμή του γραμμομοριακού ποσοστού του  $C_6$  στο κλάσμα  $C_{6-8}$  .

$\bar{n}_{C_7}$  = η μέση τιμή του γραμμομοριακού ποσοστού του  $C_7$  στο κλάσμα  $C_{6-8}$  .

$\bar{n}_{C_8}$  = η μέση τιμή του γραμμομοριακού ποσοστού του  $C_8$  στο κλάσμα  $C_{6-8}$  .

$\bar{n}_{C_9}$  = η μέση τιμή του γραμμομοριακού ποσοστού του  $C_9$  στο κλάσμα  $C_{9-11}$  .

$\bar{n}_{C_{10}}$  = η μέση τιμή του γραμμομοριακού ποσοστού του  $C_{10}$  στο κλάσμα  $C_{9-11}$  .

$\bar{n}_{C_{11}}$  = η μέση τιμή του γραμμομοριακού ποσοστού του  $C_{11}$  στο κλάσμα  $C_{9-11}$  .

Τα moles του κλάσματος του  $C_{12^+}$  υπολογίζονται από την εφαρμογή της σχέσης (4.13) :

$$N_{C_{6^+}} = N_{C_{6-8}} + N_{C_{9-11}} + N_{C_{12^+}} \Leftrightarrow \frac{M_{C_{6^+}}}{MR_{C_{6^+}}} = \frac{M_{C_{6-8}}}{MR_{C_{6-8}}} + \frac{M_{C_{9-11}}}{MR_{C_{9-11}}} + N_{C_{12^+}} \quad (4.13)$$

Λύνοντας την (4.13) ως προς  $N_{C_{12^+}}$  ολοκληρώνεται η πληροφορία για το σύνολο των moles του ρευστού, κι έτσι με την εφαρμογή της σχέσης (4.14),

υπολογίζεται και το τελικό ζητούμενο της εν λόγω εργασίας που είναι η γραμμομοριακή σύσταση του ρευστού σε  $CO_2$ ,  $C_1$ ,  $C_2$ ,  $C_3$ ,  $C_4$ ,  $C_5$ ,  $C_{6-8}$ ,  $C_{9-11}$  και  $C_{12+}$

$$n_i' = \frac{N_i}{\sum N_i} \quad (4.14)$$

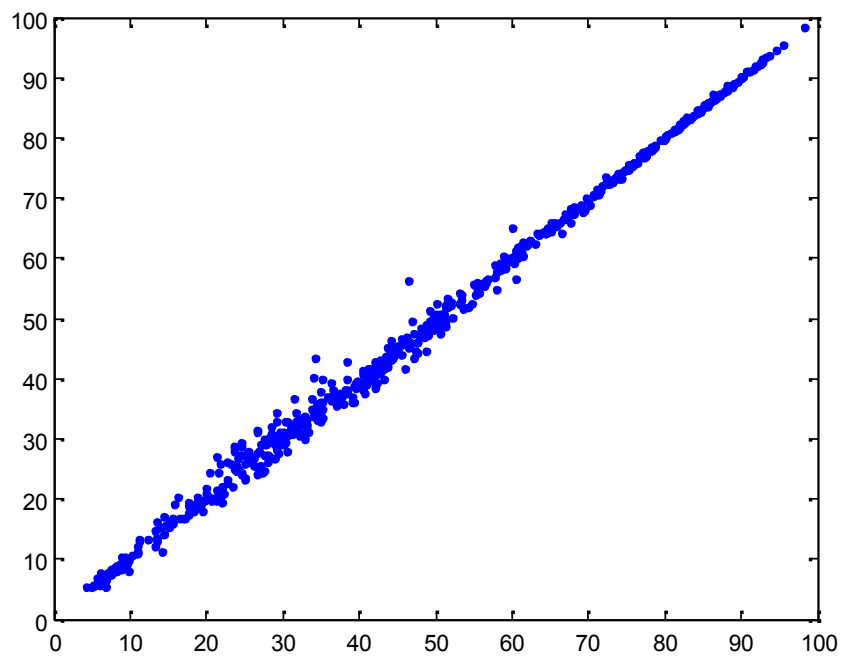
Όπου :

$n_i'$  = η πρόβλεψη του μοντέλου για την επί τοις εκατό κατά mole περιεκτικότητα του ρευστού στο συστατικό ή κλάσμα συστατικών  $i$ .

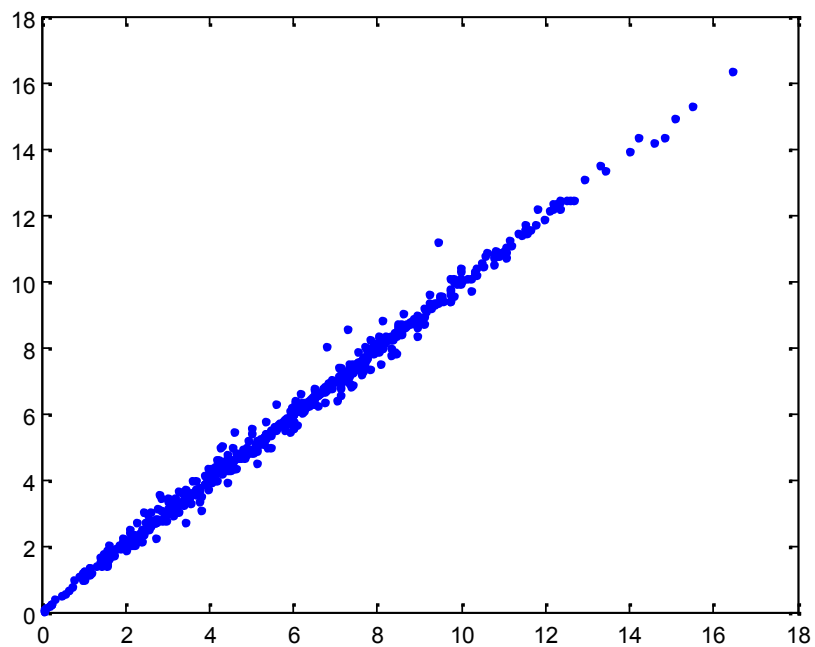
$N_i$  = τα moles του συστατικού  $i$  όπως υπολογίζονται από τις σχέσεις (4.1) ή (4.13).

$\sum N_i$  = το σύνολο των moles όλων των συστατικών που περιέχονται σε 100 γραμμάρια ρευστού.

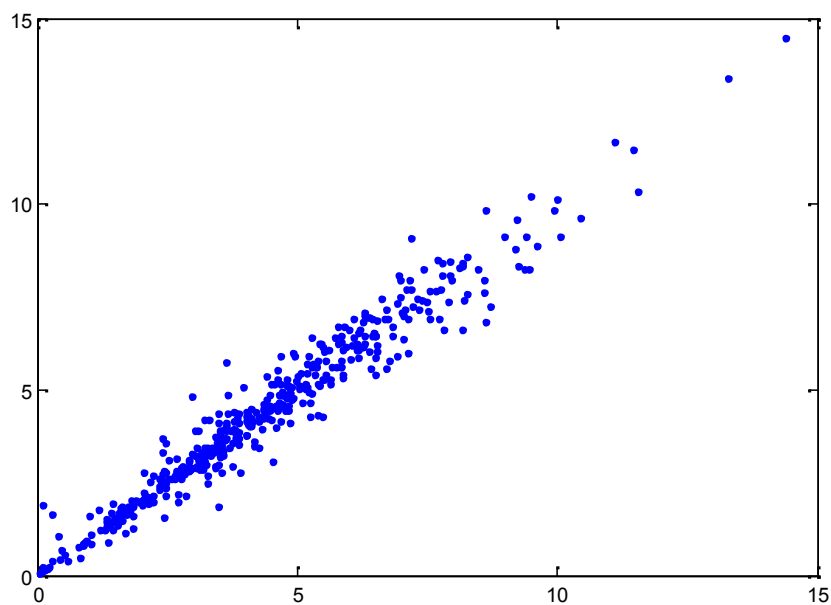
Τα αποτελέσματα των προβλέψεων του μοντέλου για τα γραμμομοριακά κλάσματα των συστατικών  $CO_2$ ,  $C_1$ ,  $C_2$ ,  $C_3$ ,  $C_4$ ,  $C_5$ ,  $C_{6-8}$ ,  $C_{9-11}$  και  $C_{12+}$ , αναπαρίστανται στα παρακάτω διαγράμματα από τους άξονες των  $\chi$ , ενώ στους άξονες των  $\psi$  αναπαρίστανται οι αντίστοιχες τιμές για τα ίδια γραμμομοριακά κλάσματα όπως προέκυψαν από τα εργαστηριακά δεδομένα. Κατά αυτόν τον τρόπο όσο το πλήθος των σημείων προσεγγίζει τη διαγώνιο, τόσο οι προβλέψεις του μοντέλου κρίνονται ικανοποιητικές.



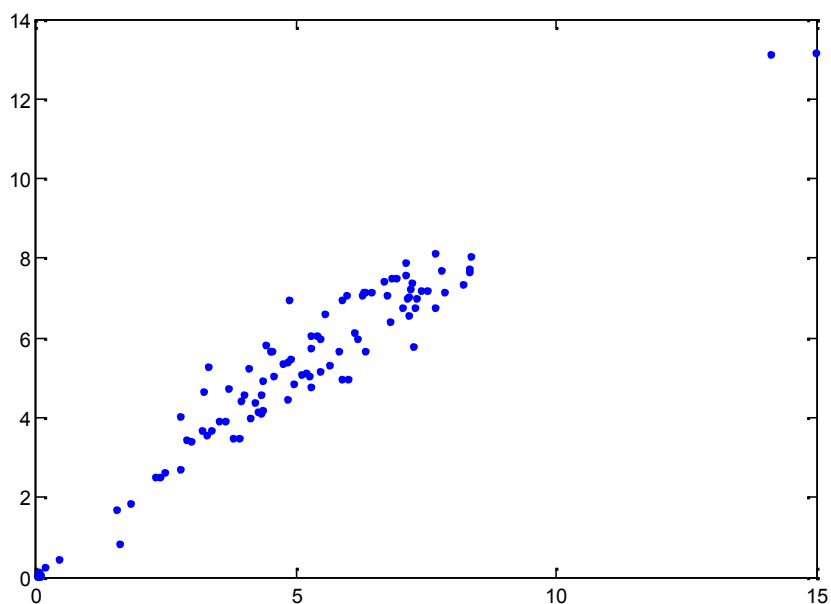
**Εικόνα 4.30 :** Σύνολο σημείων. Άξονας των  $\chi : n_{C_1}$ . Άξονας των  $\psi : n_{C_1}$ . Μέσο σχετικό σφάλμα  $\bar{R}_{C_1} = 2.5809\%$ .



**Εικόνα 4.31 :** Σύνολο σημείων. Άξονας των  $\chi : n_{C_2}$ . Άξονας των  $\psi : n_{C_2}$ . Μέσο σχετικό σφάλμα  $\bar{R}_{C_2} = 2.6733\%$ .

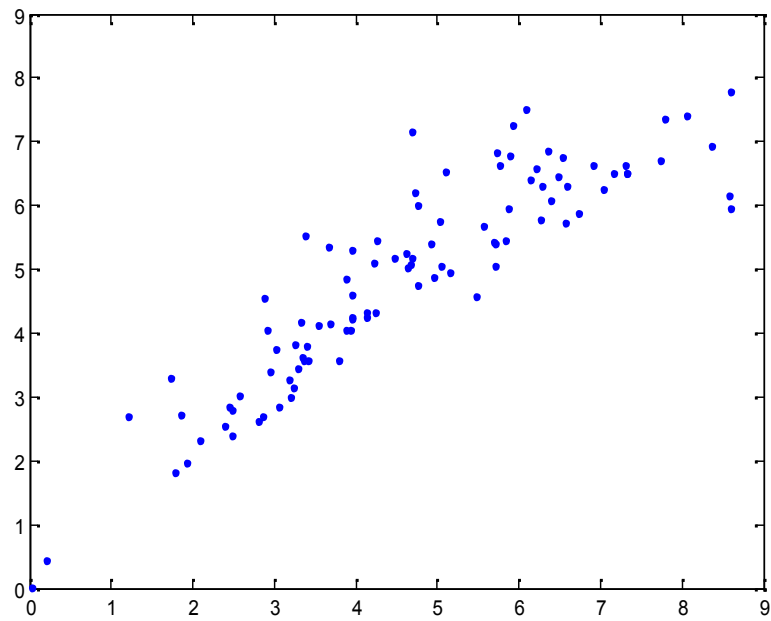


**Εικόνα 4.32 :** Σημεία εκπαίδευσης (training). Άξονας των  $\chi : n_{C_3}'$ . Άξονας των  $\psi : n_{C_3}$ . Μέσο σχετικό σφάλμα  $\bar{R} = 6.78\%$ .

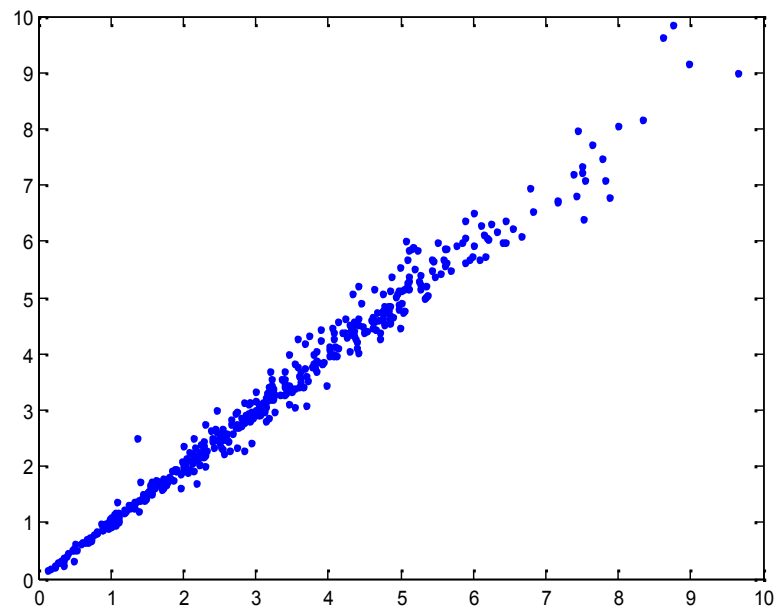


**Εικόνα 4.33 :** Σημεία επιβεβαίωσης (validation). Άξονας των  $\chi : n_{C_3}'$ . Άξονας των  $\psi : n_{C_3}$ . Μέσο σχετικό σφάλμα  $\bar{R} = 7.34\%$ .

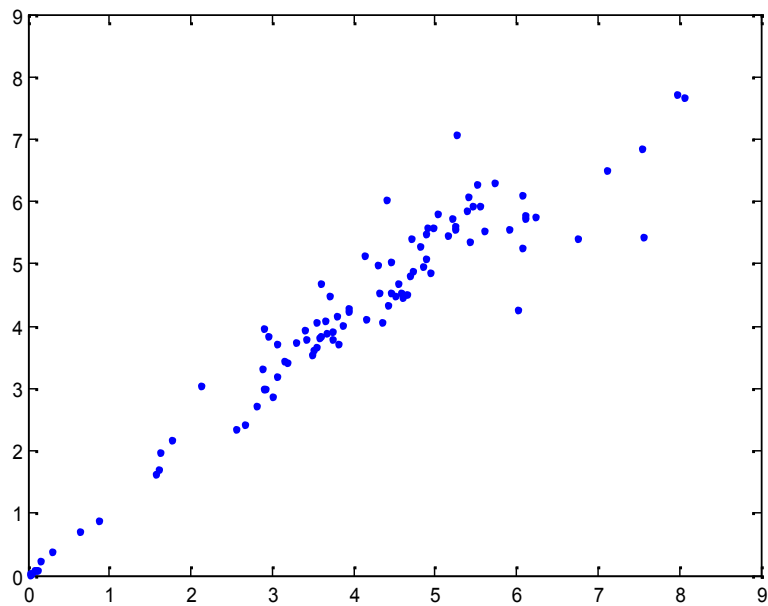




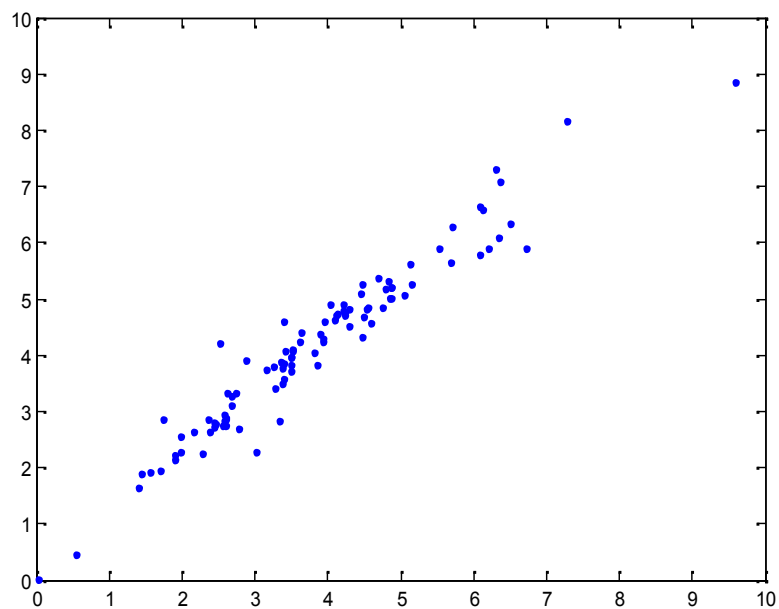
**Εικόνα 4.34 :** Σημεία ελέγχου (test). Αξονας των  $\chi : n_{C_3}$  '. Αξονας των  $\psi : n_{C_3}$  . Μέσο σχετικό σφάλμα  $\bar{R} = 8.15\%$  .



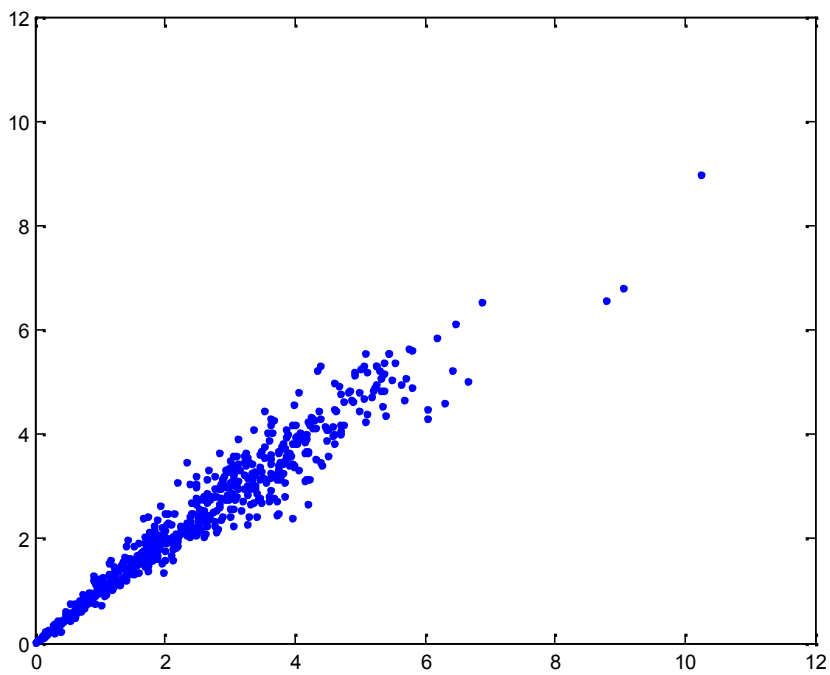
**Εικόνα 4.35 :** Σημεία εκπαίδευσης (training). Αξονας των  $\chi : n_{C_4}$  '. Αξονας των  $\psi : n_{C_4}$  . Μέσο σχετικό σφάλμα  $\bar{R} = 1.08\%$  .



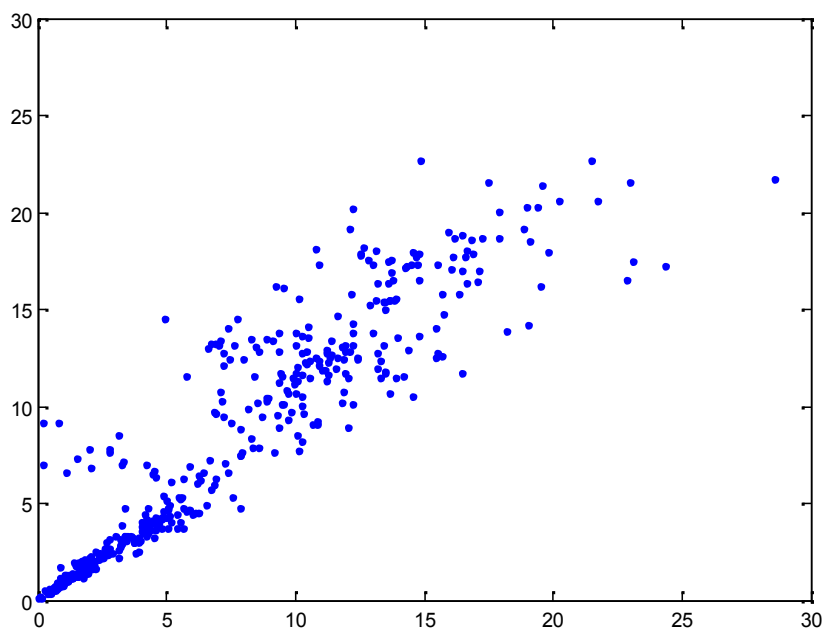
**Εικόνα 4.36 :** Σημεία επιβεβαίωσης (validation). Άξονας των  $\chi : n_{C_4}$ . Άξονας των  $\psi : n_{C_4}$ . Μέσο σχετικό σφάλμα  $\bar{R} = 7.46\%$ .



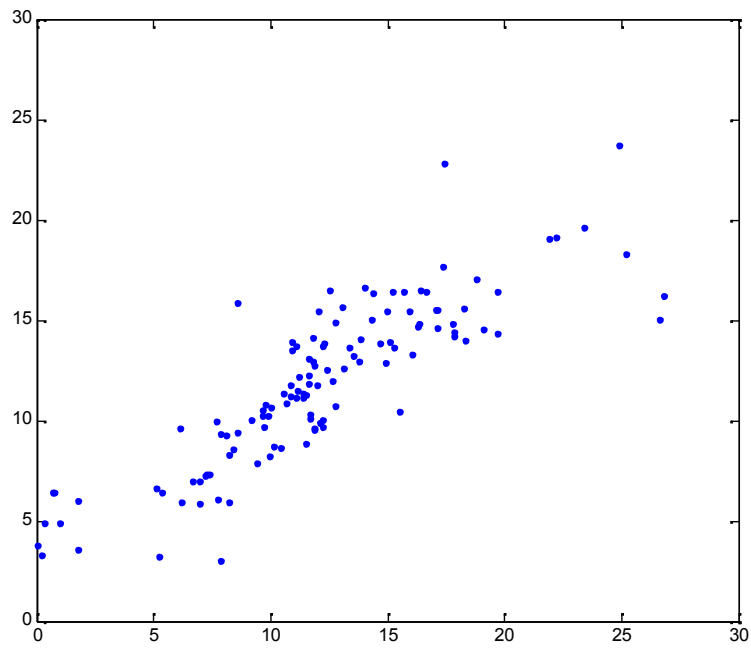
**Εικόνα 4.37 :** Σημεία ελέγχου (test). Άξονας των  $\chi : n_{C_4}$ . Άξονας των  $\psi : n_{C_4}$ . Μέσο σχετικό σφάλμα  $\bar{R} = 7.91\%$ .



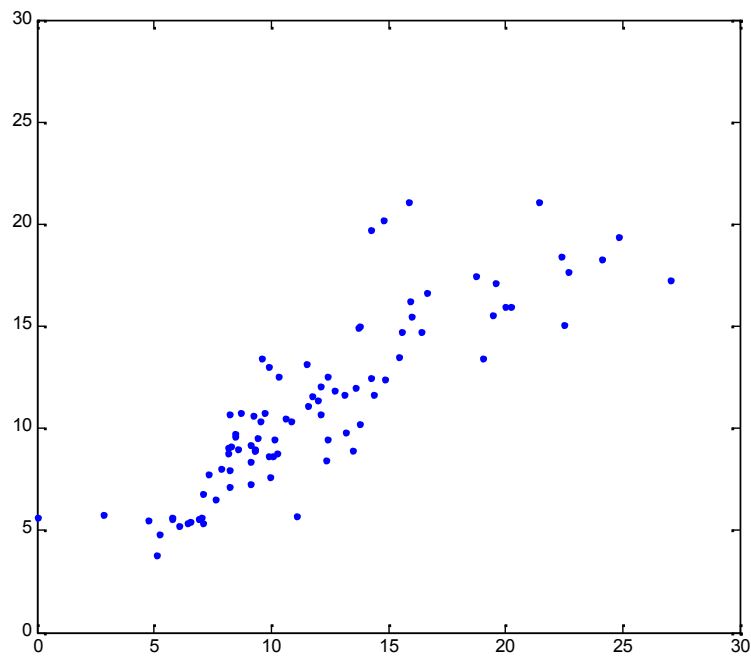
**Εικόνα 4.38 :** Σύνολο σημείων. Άξονας των  $\chi : n_{C_5}$ . Άξονας των  $\psi : n_{C_5}$ . Μέσο σχετικό σφάλμα  $\bar{R} = 1.04\%$ .



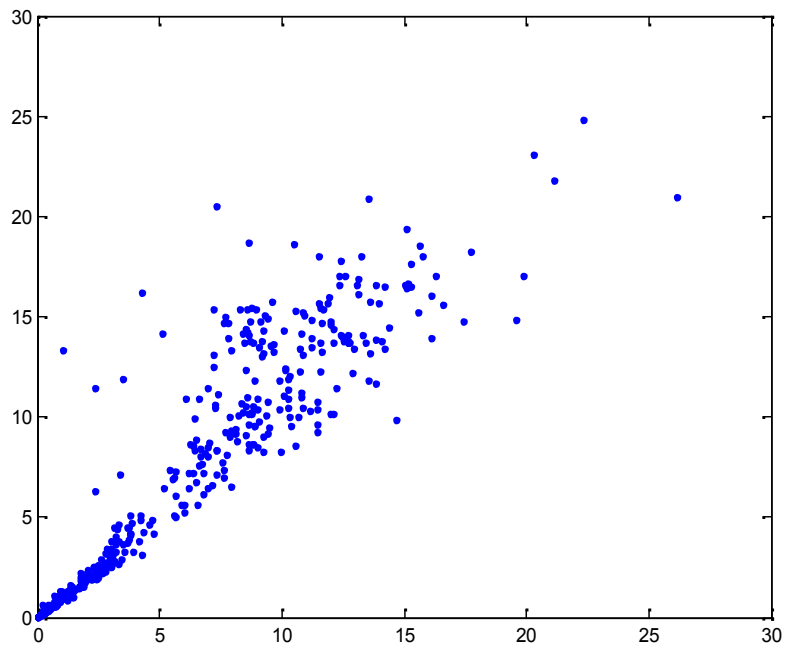
**Εικόνα 4.39 :** Σημεία εκπαίδευσης (training). Άξονας των  $\chi : n_{C_{6-8}}$ . Άξονας των  $\psi : n_{C_{6-8}}$ . Μέσο σχετικό σφάλμα  $\bar{R} = 22.09\%$ .



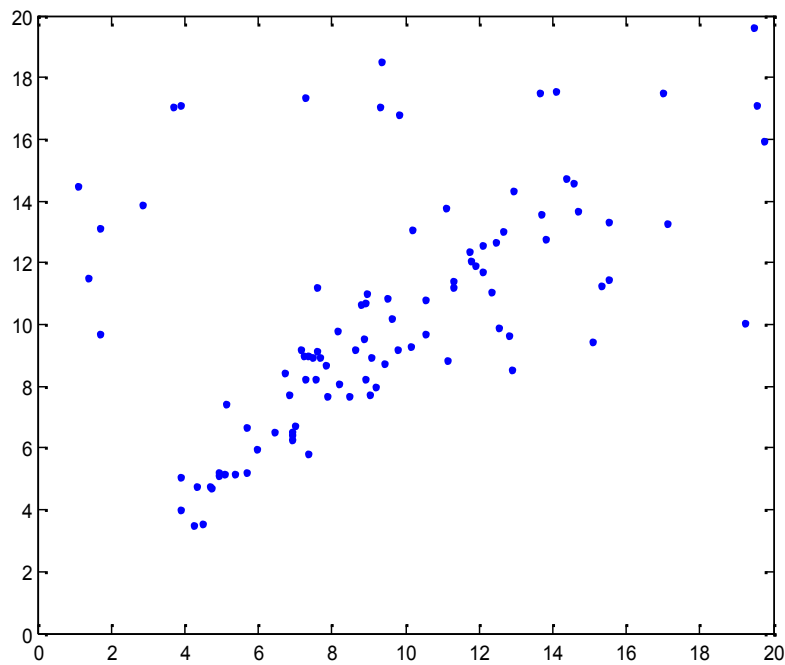
**Εικόνα 4.40 :** Σημεία επιβεβαίωσης (validation). Άξονας των  $\chi : n_{C_{6_8}}$ . Άξονας των  $\psi : n_{C_{6_8}}$ . Μέσο σχετικό σφάλμα  $\bar{R} = 164.91\%$ .



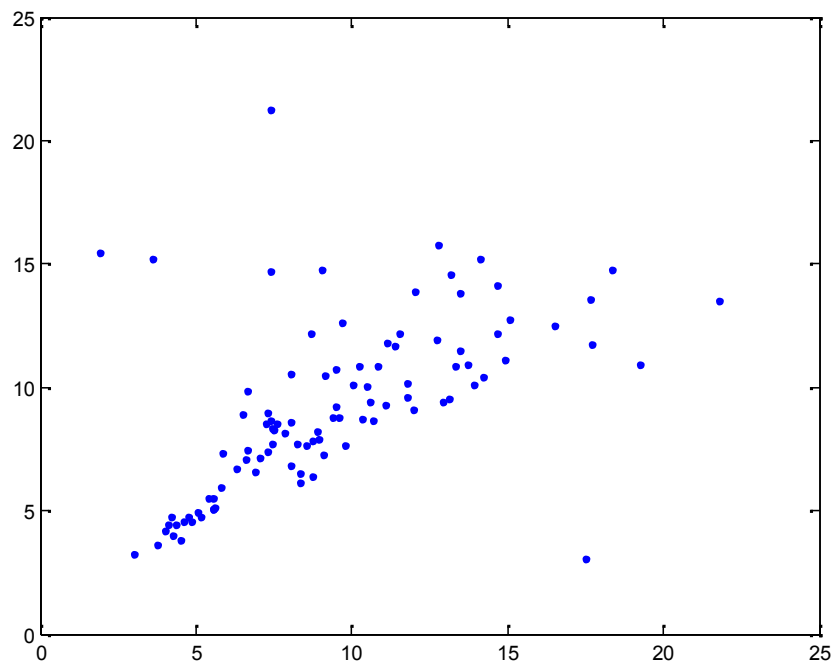
**Εικόνα 4.41 :** Σημεία ελέγχου (test). Άξονας των  $\chi : n_{C_{6_8}}$ . Άξονας των  $\psi : n_{C_{6_8}}$ . Μέσο σχετικό σφάλμα  $\bar{R} = 132.91\%$ .



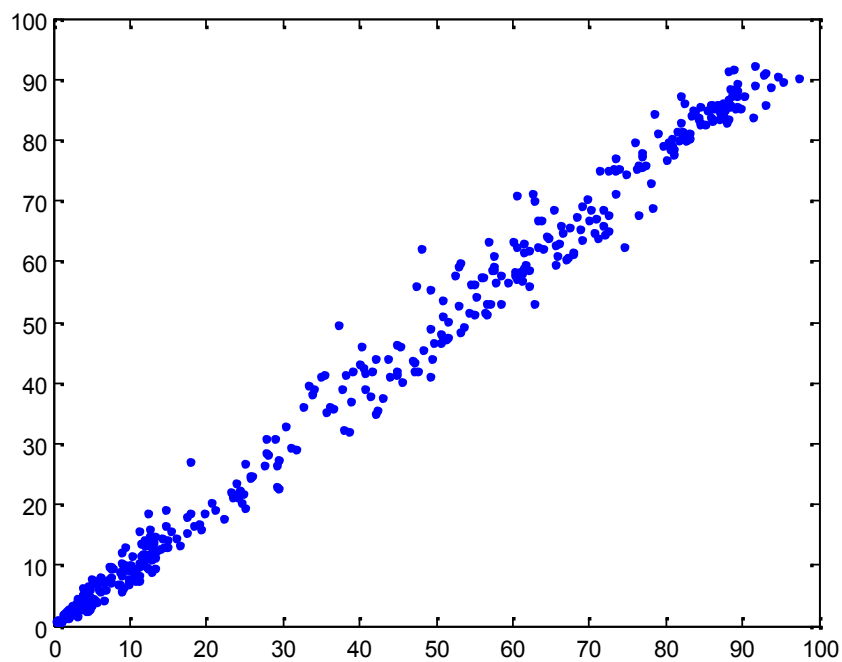
**Εικόνα 4.42 :** Σημεία εκπαίδευσης (training). Άξονας των  $\chi : n_{C_{9\_11}}$ . Άξονας των  $\psi : n_{C_{9\_11}}$ . Μέσο σχετικό σφάλμα  $\bar{R} = 34.91\%$ .



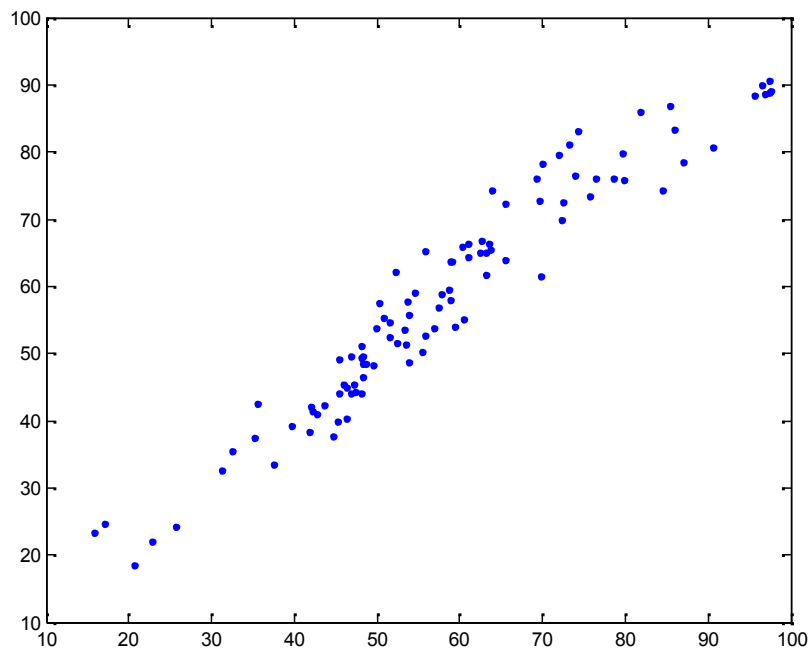
**Εικόνα 4.43 :** Σημεία επιβεβαίωσης (validation). Άξονας των  $\chi : n_{C_{9\_11}}$ . Άξονας των  $\psi : n_{C_{9\_11}}$ . Μέσο σχετικό σφάλμα  $\bar{R} = 54.12\%$ .



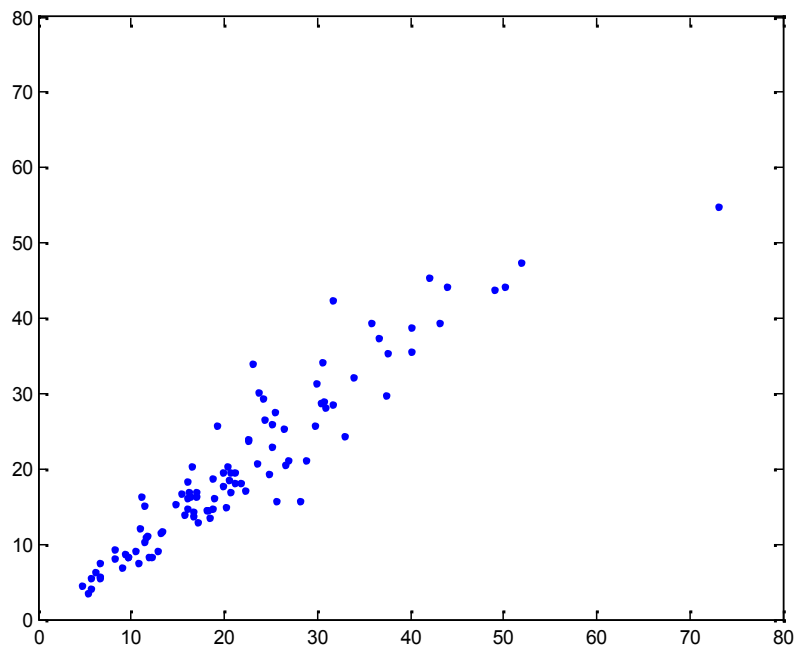
**Εικόνα 4.44 :** Σημεία ελέγχου (test). Άξονας των  $\chi : n_{C_9_{-11}}$ . Άξονας των  $\psi : n_{C_9_{-11}}$ . Μέσο σχετικό σφάλμα  $\bar{R} = 32.04\%$



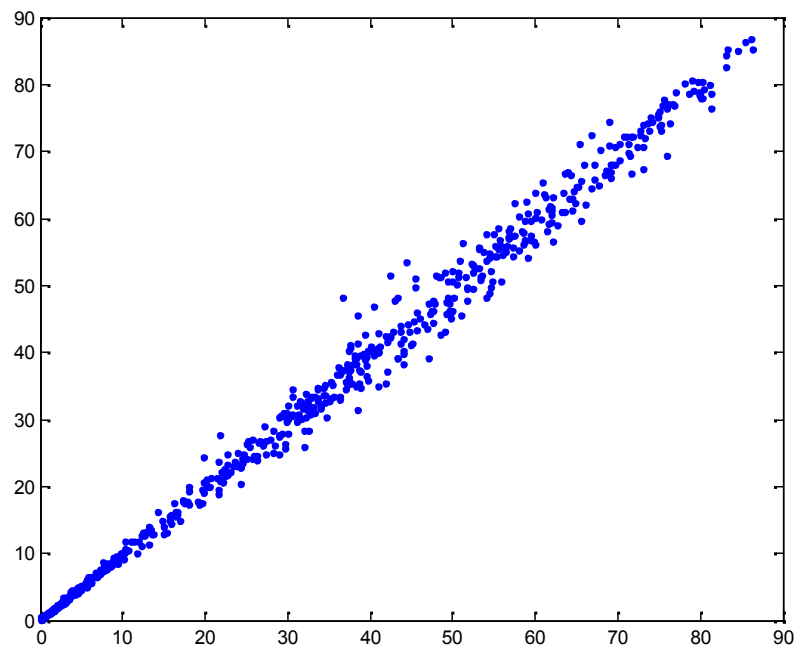
**Εικόνα 4.45 :** Σημεία εκπαίδευσης (training) του πέμπτου νευρωνικού δικτύου. Άξονας των  $\chi : n_{C_{12^+}}$ . Άξονας των  $\psi : n_{C_{12^+}}$ . Μέσο σχετικό σφάλμα  $\bar{R} = 0.65\%$ .



**Εικόνα 4.46 :** Σημεία επιβεβαίωσης (validation) του πέμπτου νευρωνικού δικτύου. Άξονας των  $\chi : n_{C_{12}^+}$ . Άξονας των  $\psi : n_{C_{12}^+}$ . Μέσο σχετικό σφάλμα  $\bar{R} = 1.06\%$ .



**Εικόνα 4.47 :** Σημεία ελέγχου (test) του πέμπτου νευρωνικού δικτύου. Άξονας των  $\chi : n_{C_{12}^+}$ . Άξονας των  $\psi : n_{C_{12}^+}$ . Μέσο σχετικό σφάλμα  $\bar{R} = 3.47\%$ .



**Εικόνα 4.48:** Στον άξονα των  $\psi$  :  $n_{C_{6^+}}$  ενώ στον άξονα των  $\chi$  :  $n_{C_{6^+}}'$ .

Όπως φαίνεται και στα παραπάνω διαγράμματα, το μοντέλο προβλέπει τα γραμμομοριακά κλάσματα των συστατικών  $C_1$ ,  $C_2$ ,  $C_3$ ,  $C_4$  και  $C_5$  με εξαιρετική επιτυχία, και αρκετά ικανοποιητικά τα γραμμομοριακά κλάσματα των  $C_{6-8}$ ,  $C_{9-11}$ , και  $C_{12^+}$ . Αυτή η λιγότερο καλή απόδοση του μοντέλου στις προβλέψεις των επί τοις εκατό κατά mole περιεκτικοτήτων των ρευστών στα κλάσματα  $C_{6-8}$ ,  $C_{9-11}$  οφείλεται στον παράγοντα της ποικιλομορφίας των ενώσεων όσο ο αριθμός των ατόμων του άνθρακα σε μια ένωση υδρογονανθράκων αυξάνεται, αυτό έχει ως συνέπεια την μικρή ακρίβεια στα μέσα μοριακά βάρη των κλασμάτων αυτών. Αντιθέτως η πρόβλεψη του γραμμομοριακού κλάσματος του  $C_{12^+}$  δεν έχει το πρόβλημα αυτό, καθώς τα moles του υπολογίστηκαν από τη σχέση (4.13) :

$$\frac{M_{C_{6^+}}}{MR_{C_{6^+}}}' = \frac{M_{C_{6-8}}}{MR_{C_{6-8}}} + \frac{M_{C_{9-11}}}{MR_{C_{9-11}}} + N_{C_{12^+}}.$$



Ως προς τα διαστήματα εμπιστοσύνης, οι προβλέψεις του μοντέλου παρουσιάζουν τα εξής χαρακτηριστικά με απόλυτο σφάλμα  $E_i = n_i - n_i'$  και μέσο σχετικό σφάλμα  $\bar{R}_i = abs\left(\frac{E_i}{n_i}\right)$ :

Η τιμή του απολύτου σφάλματος για μία πρόβλεψη της τιμής του γραμμομοριακού κλάσματος του  $C_1$ , έχει 68.27% πιθανότητα να βρεθεί στο διάστημα (-1.2550 1.3626), 95.45% πιθανότητα να βρεθεί στο διάστημα (-2.6097 2.5059) και 99.73% πιθανότητα να βρεθεί στο διάστημα (-3.8727 3.7802). Επίσης το μέσο σχετικό σφάλμα  $\bar{R}_{C_1} = 2.5809\%$ .

Η τιμή του απολύτου σφάλματος για μία πρόβλεψη της τιμής του γραμμομοριακού κλάσματος του  $C_2$ , έχει 68.27% πιθανότητα να βρεθεί στο διάστημα (-0.2433 0.1932), 95.45% πιθανότητα να βρεθεί στο διάστημα (-0.4616 0.4114) και 99.73% πιθανότητα να βρεθεί στο διάστημα (-0.6799 0.6297). Το μέσο σχετικό σφάλμα είναι  $\bar{R}_{C_2} = 2.6733\%$ .

Η τιμή του απολύτου σφάλματος για μία πρόβλεψη της τιμής του γραμμομοριακού κλάσματος του  $C_3$ , έχει 68.27% πιθανότητα να βρεθεί στο διάστημα (-0.5890 0.7138), 95.45% πιθανότητα να βρεθεί στο διάστημα (-1.2404 1.3652) και 99.73% πιθανότητα να βρεθεί στο διάστημα (-1.8918 2.0166). Το μέσο σχετικό σφάλμα κατά απόλυτη τιμή είναι  $\bar{R}_{C_3} = 16.1272\%$ .

Η τιμή του απολύτου σφάλματος για μία πρόβλεψη της τιμής του γραμμομοριακού κλάσματος του  $C_4$ , έχει 68.27% πιθανότητα να βρεθεί στο διάστημα (-0.2991 0.4639), 95.45% πιθανότητα να βρεθεί στο διάστημα (-0.6806 0.8454) και 99.73% πιθανότητα να βρεθεί στο διάστημα (-1.0620 1.2268). Το μέσο σχετικό σφάλμα κατά απόλυτη τιμή είναι  $\bar{R}_{C_4} = 7.2981\%$ .

Η τιμή του απολύτου σφάλματος για μία πρόβλεψη της τιμής του γραμμομοριακού κλάσματος του  $C_5$ , έχει 68.27% πιθανότητα να βρεθεί στο διάστημα (-0.4181 0.3901), 95.45% πιθανότητα να βρεθεί στο διάστημα (-0.7458 0.7177) και 99.73% πιθανότητα να βρεθεί στο διάστημα (-1.0735 1.0454). Το μέσο σχετικό σφάλμα κατά απόλυτη τιμή είναι  $\overline{R}_{C_5} = 10.2382\%$ .

Η τιμή του απολύτου σφάλματος για μία πρόβλεψη της τιμής του γραμμομοριακού κλάσματος του  $C_{6-8}$ , έχει 68.27% πιθανότητα να βρεθεί στο διάστημα (-2.6298 1.8198), 95.45% πιθανότητα να βρεθεί στο διάστημα (-4.8546 4.0446) και 99.73% πιθανότητα να βρεθεί στο διάστημα (-7.0794 6.2693). Το μέσο σχετικό σφάλμα κατά απόλυτη τιμή είναι  $\overline{R}_{C_{6-8}} = 71.6989\%$ .

Η τιμή του απολύτου σφάλματος για μία πρόβλεψη της τιμής του γραμμομοριακού κλάσματος του  $C_{9-11}$ , έχει 68.27% πιθανότητα να βρεθεί στο διάστημα (-1.4374 4.7532), 95.45% πιθανότητα να βρεθεί στο διάστημα (-4.5327 7.8486) και 99.73% πιθανότητα να βρεθεί στο διάστημα (-7.6281 10.9439). Το μέσο σχετικό σφάλμα κατά απόλυτη τιμή είναι  $\overline{R}_{C_{9-11}} = 21.04\%$ .

Η τιμή του απολύτου σφάλματος για μία πρόβλεψη της τιμής του γραμμομοριακού κλάσματος του  $C_{12+}$ , έχει 68.27% πιθανότητα να βρεθεί στο διάστημα (-3.8761 3.3700), 95.45% πιθανότητα να βρεθεί στο διάστημα (-7.4991 6.9930) και 99.73% πιθανότητα να βρεθεί στο διάστημα (-11.1222 10.6161). Το μέσο σχετικό σφάλμα κατά απόλυτη τιμή είναι  $\overline{R}_{C_{12+}} = 21.7331\%$ .

Η τιμή του απολύτου σφάλματος για μία πρόβλεψη της τιμής του γραμμομοριακού κλάσματος του  $C_{6+}$ , έχει 68.27% πιθανότητα να βρεθεί στο διάστημα (-2.0891 1.9390), 95.45% πιθανότητα να βρεθεί στο διάστημα (-4.1031 3.9530) και 99.73% πιθανότητα να βρεθεί στο διάστημα (-6.1171 5.9671). Το μέσο σχετικό σφάλμα κατά απόλυτη τιμή είναι  $\overline{R}_{C_{6+}} = 5.3558\%$ .

#### 4.4 ΣΥΜΠΕΡΑΣΜΑΤΑ.

Αποτέλεσμα της παρούσης εργασίας είναι πως για πρώτη φορά, δεδομένα αναλυτικής σύστασης ρευστού υδρογονανθράκων υπό μορφή μοριακής περιεκτικότητας μπορούν να καταστούν διαθέσιμα στο πολύ πρώιμο στάδιο των διαγραφιών παραγωγής. Το μοντέλο που αναπτύχθηκε, κατάφερε με την εφαρμογή των νευρωνικών δικτύων, να διευρύνει και να μετατρέψει τις μετρήσεις του οπτικού αναλυτή ρευστών OFA από επί τοις εκατό κατά μάζα περιεκτικότητα των ρευστών σε  $CO_2$ ,  $C_1$ ,  $C_2$ ,  $C_{3-5}$ , και  $C_{6+}$ , σε επί τοις εκατό κατά mol περιεκτικότητα των ρευστών σε  $CO_2$ ,  $C_1$ ,  $C_2$ ,  $C_3$ ,  $C_4$ ,  $C_5$ ,  $C_{6-8}$ ,  $C_{9-11}$ , και  $C_{12+}$ . Για να γίνει πιο αντιληπτή η σημασία μίας τέτοιας πληροφορίας σε ένα τόσο πρώιμο ερευνητικό στάδιο, αρκεί να αναφερθεί πως τα δεδομένα που διατίθενται από τα αποτελέσματα του μοντέλου δύνανται να εξαχθούν από το πεδίο, κατά τη διάρκεια της διεξαγωγής των μετρήσεων του OFA, ενώ μέχρι τώρα η απόκτηση αναλυτικής σύστασης του ρευστού του ταμιευτήρα προϋπέθετε την χρονοβόρα αποστολή του δείγματος για εργαστηριακή ανάλυση. Αυτό βέβαια δεν σημαίνει πως το παρόν μοντέλο μπορεί να υποκαταστήσει μία εργαστηριακή ανάλυση, μπορεί όμως να συντελέσει εξαιρετικά στην ακρίβεια των μελετών που εκπονούνται από τους μηχανικούς πετρελαίου κατά την αξιολόγηση μίας γεώτρησης.

Τα τεχνητά νευρωνικά δίκτυα από τα οποία αποτελείται το μοντέλο συμπεριφέρθηκαν ικανοποιητικά στην εφαρμογή αυτή, παρόλο που το σύνολο των χαρακτηρισμένων παραδειγμάτων με το οποίο εκπαιδεύτηκαν, περιείχε σχετικά περιορισμένη πληροφορία αναλογικά με εύρος του μαθησιακού στόχου. Το γεγονός πως τα νευρωνικά δίκτυα κατάφεραν να γενικεύουν καλά, αποδεικνύει την ύπαρξη σχέσης ανάμεσα στο διάνυσμα εισόδου και τα διανύσματα εξόδου που χρησιμοποιήθηκαν κάθε φορά.

Το γεγονός της λιγότερο ικανοποιητικής γενίκευσης των δικτύων που

εκπαιδεύτηκαν για να προβλέπουν τους λόγους  $Y_3 = \left( \frac{M_{C_{6-8}}}{M_{C_{6+}}} \right)$  και  $Y_4 = \left( \frac{M_{C_{9-11}}}{M_{C_{9+}}} \right)$

ήταν ουσιαστικά αναμενόμενο, καθώς ο παράγοντας της ποικιλομορφίας των

ενώσεων των κλασμάτων  $C_{6-8}$  και  $C_{9-11}$  δεν ήταν δυνατόν να συνυπολογιστεί, αφού έλειπε η πληροφορία για τα ακριβή μέσα μοριακά βάρη των κλασμάτων αυτών. Την γενίκευση των προαναφερθέντων δικτύων φαίνεται να επηρέασε επίσης η παρουσία στη βάση δεδομένων που χρησιμοποιήθηκε, πετρελαϊκών ρευστών που είναι προϊόντα βιοαποδόμησης, τα οποία όντας λίγα αποτέλεσαν εξαιρέσεις ενός γενικού κανόνα. Συμπεραίνεται λοιπόν πως όσο τα δεδομένα με τα οποία τροφοδοτούνται τα νευρωνικά δίκτυα αποκλίνουν του γενικού κανόνα ή της γενικής τάσης του συνόλου των παραδειγμάτων εκπαίδευσης, τόσο η δυνατότητά τους να γενικεύουν καλά κινδυνεύει.

Μία πρόταση για περαιτέρω έρευνα επί του αντικειμένου της παρούσης εργασίας, θα μπορούσε να είναι η αφαίρεση των προϊόντων βιοαποδόμησης από την βάση εκπαίδευσης με στόχο την βελτίωση της γενίκευσης του τρίτου και του τέταρτου νευρωνικού δικτύου, καθώς όπως φάνηκε από την μελέτη των αποτελεσμάτων, τα παραδείγματα αυτά αποτέλεσαν εξαιρέσεις και άρα ήταν αδύνατο από τα δίκτυα να τα μάθουν.

## **5. ΒΙΒΛΙΟΓΡΑΦΙΑ :**

1. Ν. Βαρότσης. Μηχανική Ταμιευτήρων. Πολυτεχνείο Κρήτης.

2. Darwin V. Ellis & Julian M. Singer. Well Logging for Earth Scientists.
3. site : [www.rigzone.com](http://www.rigzone.com).
4. Ross Crain. Petrofysical Handbook.
5. site : [petrowiki.org](http://petrowiki.org).
6. site : [www.slb.com](http://www.slb.com).
7. Rob Badry, Derrel Fincher, Oliver Mullins, Bob Schroeder, Tony Smits.  
Downhole Optical Analysis of Formation Fluids.
8. Tim Ireland, Jeffrey Josheph, Nick Colley, Patrick Reignier, Simon Richardson. MDT: A Wireline Testing Breakthrough.
9. C.Dong, SPE, and M. O' Keefe, SPE, Schlumberger; H. Elshahawi, SPE, Shell; S. Williams, SPE, StatoilHydro; D. Stensland, Eni Norge; P. Hegeman, SPE, R. Vasques, T. Terabayashi, O. Mullins, SPE, and E. Donzier , Schlumberger. New Downhole Fluid Analysis Tool for Improved Reservoir Characterization. (SPE 108566)
10. Α. Π. Βαλαβανίδης. Φασματοσκοπία Οργανικών Ενώσεων, Αθήνα 2006.
11. Ε. Ι. ΚΑΜΙΤΣΟΣ & Γ. Δ. ΧΡΥΣΙΚΟΣ. Στοιχεία Φασματοσκοπίας Υπερύθρου Ινστιτούτο Θεωρητικής & Φυσικής Χημείας. Εθνικό Ίδρυμα Ερευνών.
12. New downhole fluid analysis tool for improved reservoir characterization.
13. Schlumberger. In Situ Fluid Analyzer brochure.
14. Beer-Lambert law, site : [life.nthu.edu.tw/~labcjw/.../beerslaw.htm](http://life.nthu.edu.tw/~labcjw/.../beerslaw.htm)
15. Peter S.Hegeman (Schlumberger), Chengli Dong (Schlumberger), Nikos Varotsis (Tech. U of Crete), Vassilis Gaganis (Tech. U of Crete). Paper : Application of Artificial Neural Networks to Downhole Fluid Analysis. 123423 – PA SPE Journal Paper 2009.
16. Albert Tarantola. Inverse Problem Theory and methods for parameter estimation. Institute de Physique du Globe de Paris, Universite de Paris 6, Paris, France.
17. Διονύσιος Θ. Χρηστόπουλος. Σημειώσεις του Προπτυχιακού Μαθήματος 'Εισαγωγή στη Γεωστατιστική'. Χανιά Απρίλιος 2008.
18. Simon Haykin. Νευρωνικά Δίκτυα και Μηχανική Μάθηση. Τρίτη έκδοση.
19. Computer Science. site : [cs.stanford.edu](http://cs.stanford.edu).

- 20.** Ι. Βλαχάβας, Π. Κεφαλάς, Ν. Βασιλειάδης, Φ. Κόκκορας, Η. Σακελλαρίου. Τεχνητή Νοημοσύνη. Β' έκδοση.
- 21.** 'ARTIFICIAL INTELLIGENCE Comes of Age in Oil and Gas'. JPT (Journal of Petroleum Technology) January 2013.
- 22.** Ali Abedini, Siavash Ashoori, Yaser Saki. 'Application of Neural Model for Prediction of Asphaltene Precipitation'. 132760 –MS SPE Conference Paper 2010.
- 23.** Varotsis N., Gaganis V., Nighswander J., Guieze P., " Novel Non-Iterative Method for Predicting the PVT Behavior of ReservoirFluids", SPE 56745, presented at the 1999 SPE Annual TechnicalConference, Oct 3-6, 1999, Houston, USA.
- 24.** Varotsis N. Schlumberger, Gaganis V. Consultant, Nighswander J. Schlumberger. 'Quality Assurance Tool for PVT Simulator Predictions'. 81751 PA-SPE Journal Paper – 2002.
- 25.** Adeyemi B.J., Sulaimon A.A., 'Predicting Wax Formation Using Artificial Neural Networks'. 163026 –MS SPE Conference Paper 2012.
- 26.** Katz, D.L., and Firoozabadi, A., 1978. Predicting phase behavior of condensate/crude oil systems using methane interaction coefficients: JPT: 1649-1655