



ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ

Τμήμα Μηχανικών Ορυκτών Πόρων

Εργαστήριο Ανάλυσης Ρευστών
& Πυρήνων Υπόγειων Ταμιευτήρων

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Μελέτη της Επίδρασης της Μίξης Κλασμάτων στην Διαμόρφωση του Αριθμού Οκτανίων της Βενζίνης

ΑΝΔΡΙΑΝΑ ΜΙΝΤΖΑ

ΕΞΕΤΑΣΤΙΚΗ ΕΠΙΤΡΟΠΗ:

Πασαδάκης Ν.: *Επίκουρος Καθηγητής (Επιβλέπων)*

Γαλετάκης Μ.: *Επίκουρος Καθηγητής*

Γαγάνης Β.: *Διδάκτωρ Μηχανολόγος Μηχανικός*

ΧΑΝΙΑ

ΔΕΚΕΜΒΡΙΟΣ 2006

αφιερώνεται στον Φρίξο

Περίληψη

Ο αριθμός οκτανίων (Α.Ο.) αποτελεί χαρακτηριστική ιδιότητα των βενζινών και καθορίζει σε μεγάλο βαθμό την εμπορική τους αξία ως καυσίμων σε μηχανές εσωτερικής καύσης. Η ιδιότητα αυτή διαμορφώνεται ανάλογα με το είδος και την αναλογία μίξης των επιμέρους κλασμάτων πετρελαίου που είναι διαθέσιμα για την παραγωγή της βενζίνης. Στην βιομηχανική πρακτική χρησιμοποιούνται συνήθως ημι-εμπειρικές σχέσεις πρόβλεψης του Α.Ο. σε μίγματα κλασμάτων. Δεδομένου ότι ο Α.Ο. δεν αποτελεί προσθετική ιδιότητα τα κλάσματα που περιέχονται στο τελικό προϊόν εμφανίζουν διαφορετικό βαθμό επίδρασης στην διαμόρφωση του Α.Ο. του μίγματος.

Στην παρούσα εργασία διερευνήθηκε συστηματικά η επίδραση στον Α.Ο. της τελικής βενζίνης, της συγκέντρωσης και του Α.Ο. των κλασμάτων που χρησιμοποιούνται για την παραγωγή της στο διυλιστήριο της Motor Oil Hellas (MOH). Αναπτύχθηκαν αλγόριθμοι πρόβλεψης με βάση τα δεδομένα παραγωγής, με χρήση τεχνητών νευρωνικών δικτύων (ΤΝΔ), ενώ παράλληλα έγινε μία προσπάθεια διερεύνησης της επίδρασης μίξης (blend effect) των επιμέρους κλασμάτων βενζίνης στον τελικό Α.Ο.. Με χρήση της χημειομετρικής τεχνικής παλινδρόμησης κυρίων συνιστωσών (Principal Component Regression, PCR) αναπτύχθηκε ένα γραμμικό μοντέλο ως προς τις μεταβλητές σύστασης και ως προς τις αλληλεπιδράσεις τους ανά δύο, οι συντελεστές του οποίου αντιπροσωπεύουν την επίδραση του αντίστοιχου κλάσματος στην διαμόρφωση του τελικού Α.Ο.. Επιπλέον, με εφαρμογή της τεχνικής του στατιστικού πειραματικού σχεδιασμού παρασκευάστηκαν εργαστηριακά μίγματα βενζίνης, με συγκεντρώσεις κλασμάτων εντός των συνηθών ορίων τους στην βενζίνη, αλλά μεγαλύτερη διακύμανση τελικών Α.Ο., καλύπτοντας έτσι σχετικά ομοιόμορφα τον χώρο των δυνατών Α.Ο.. Με βάση τους Α.Ο. και τις συστάσεις των επιμέρους κλασμάτων, αναπτύχθηκαν με ΤΝΔ αντίστοιχοι αλγόριθμοι πρόβλεψης Α.Ο..

Οι αλγόριθμοι προβλέπουν αξιόπιστα τον αριθμό οκτανίων της τελικής βενζίνης, χρησιμοποιώντας ως δεδομένα εισαγωγής μόνο την περιεκτικότητα και τον Α.Ο. των επιμέρους κλασμάτων και μπορούν να χρησιμοποιηθούν τόσο για την πρόβλεψη του αριθμού οκτανίων των τελικών βενζινών του διυλιστηρίου, όσο και ως εργαλείο επιλογής της βέλτιστης σύστασης της βενζίνης, εισάγοντας επιπλέον κριτήρια κόστους και διαθεσιμότητας των κλασμάτων.

Ευχαριστίες

Ευχαριστίες εκφράζονται στα μέλη της τριμελούς εξεταστικής επιτροπής, τον επίκουρο καθηγητή κ.κ. Ν. Πασαδάκη, επιβλέποντα της παρούσας εργασίας, τον επίσης επίκουρο καθηγητή του Τμήματος Μηχανικών Ορυκτών Πόρων κ.κ. Μ. Γαλετάκη και τον διδάκτορα Μηχανολόγο Μηχανικό κ.κ. Β. Γαγάνη για την πολύτιμη καθοδήγησή τους κατά την εκπόνηση και αξιολόγηση της παρούσας εργασίας.

Ευχαριστίες εκφράζονται επίσης στην Motor Oil Hellas (ΜΟΗ) και ιδιαίτερα στα στελέχη της κ.κ. Δ. Βουγιουκλάκη, κ.κ. Χ. Φωτεινόπουλο και κ.κ. Π. Κοτσόκολο για την παροχή δεδομένων και την καθοδήγησή τους κατά τον χειρισμό των τελευταίων.

Ιδιαίτερα θα πρέπει να ευχαριστήσω τον αδερφό μου και γραφίστα Κ. Μίντζα για την σημαντική βοήθεια στην απόδοση των σχημάτων που χρησιμοποιήθηκαν και γενικότερα την επιμέλεια της εικονογράφησης του κειμένου, καθώς και την φίλη Μηχανικό Ορυκτών Πόρων Κ. Αναστασιάδου τόσο για την ηθική υποστήριξη, όσο για τις σημαντικές παρατηρήσεις κατά την συγγραφή του κειμένου.

Τέλος, θα ήθελα να εκφράσω την ευγνωμοσύνη μου στους γονείς μου και τους φίλους για την κατανόηση και την συνεχή ενθάρρυνση, χωρίς την οποία η περάτωση του όλου πονήματος δεν θα ήταν δυνατή.

Πίνακας Περιεχομένων

Περίληψη	i
Ευχαριστίες	iii
Πίνακας Περιεχομένων	v
Περιγραφή Αντικειμένου της Εργασίας	1
1 Βενζίνη και Αριθμός Οκτανίων	5
1.1 Αριθμός Οκτανίων	5
Κύκλος Λειτουργίας Μηχανών Εσωτερικής Καύσης	5
Προσδιορισμός Αριθμού Οκτανίων Βενζίνης	9
1.2 Βενζίνη	11
Συστατικά της βενζίνης	13
Προδιαγραφές Βενζίνης	17
Διεργασίες παραγωγής βενζίνης	19
1.3 Καύση βενζίνης	23
Ανώμαλες Μορφές Καύσης	23
Παράγοντες που επηρεάζουν τη αυτανάφλεξη	25
Αριθμός Οκτανίων και Χημική Σύσταση Βενζίνης	27
2 Πολυπαραμετρική Ανάλυση Δεδομένων	35
2.1 Εισαγωγή στην Χημειομετρία	35
2.2 Διερευνητική Ανάλυση	37
Βασικές Έννοιες Στατιστικής	38
Προεπεξεργασία Δεδομένων	41
Συσχέτιση Μεταβλητών	43
2.3 Μέθοδοι Παλινδρόμησης Πολυπαραμετρικών Δεδομένων	48
Πολλαπλή Γραμμική Παλινδρόμηση	48
Παλινδρόμηση Κυρίων Συνιστωσών	49
Παλινδρόμηση Μερικών Ελαχίστων Τετραγώνων	49
3 Νευρωνικά Δίκτυα	51
3.1 Φυσικοί και τεχνητοί νευρώνες	52
Φυσικοί νευρώνες	52
Τεχνητοί νευρώνες –ο κόμβος TLU	54
3.2 Τεχνητά νευρωνικά δίκτυα	59
Αρχιτεκτονική νευρωνικών δικτύων	59
Λειτουργία νευρωνικού δικτύου ευθείας προώθησης	61
Εκπαίδευση νευρωνικών δικτύων	65
3.3 Συναρτήσεις Μεταφοράς και Στοιχεία Τοπολογίας	71
Είδη Συναρτήσεων Μεταφοράς	72
Αρχιτεκτονική και Προσέγγιση Συνάρτησης	74
3.4 Γενίκευση	76
Μέγεθος Συνόλου Εκπαίδευσης και Δικτύου	77
Σύνολο επαλήθευσης	78
4 Σχεδιασμός Πειράματος	81
4.1 Είδη πειραμάτων	82
Τύποι Παραγοντικών Πειράματων	82
Εναλλακτικές Τακτικές Πειραματισμού	84
4.2 Στρατηγική Υλοποίησης Πειραματικού Σχεδιασμού	85
Μεταβλητές Διεργασίας	86
Μοντελοποίηση διαδικασίας	90
4.3 Τυπικές Τεχνικές Σχεδιασμού Πειράματος	97
Πλήρης Παραγοντικός Σχεδιασμός Πειράματος	97
Κλασματικός Παραγοντικός Σχεδιασμός Πειράματος	101
D-optimal πειραματικός σχεδιασμός	104

4.4	Πειραματικός Σχεδιασμός σε Μίγματα Συστατικών	105
	Μοντέλα σε μίγματα συστατικών	106
	Τεχνικές Σχεδιασμού σε Μίγματα Συστατικών	108
5	Πρόβλεψη Αριθμού Οκτανίων σε Βενζίνες και Διερεύνηση Επίδρασης Μίξης	111
5.1	Μοντέλα Πρόβλεψης Αριθμού Οκτανίων από Δεδομένα Παραγωγής	111
	Δεδομένα Παραγωγής	111
	Υφιστάμενοι Αλγόριθμοι Πρόβλεψης Αριθμού Οκτανίων	113
	Ανάπτυξη Μοντέλων Πρόβλεψης Α.Ο. με Τεχνητά Νευρωνικά Δίκτυα	116
5.2	Διερεύνηση Επίδρασης της Μίξης στον Α.Ο.	125
5.3	Μοντέλα Πρόβλεψης Αριθμού Οκτανίων από Πειραματικά Δεδομένα	128
	Πειραματικός Σχεδιασμός	129
	Πειραματικά Δεδομένα	130
	Ανάπτυξη Μοντέλων Πρόβλεψης Α.Ο Εργαστηριακών Μιγμάτων	132
	Συμπεράσματα- Προτάσεις	137
	Βιβλιογραφία	139
	Παράρτημα	143

Περιγραφή Αντικειμένου της Εργασίας

Η βενζίνη αποτελεί την βασική πηγή κέρδους για την βιομηχανία πετρελαίου, δεδομένου ότι σε ένα εκσυγχρονισμένο διυλιστήριο μέχρι και 70% του ακατέργαστου αργού πετρελαίου που οδηγείται αρχικά σε κλασματική απόσταξη μετατρέπεται σε βενζίνη. Η βενζίνη παράγεται με μίξη διαφόρων κλασμάτων-προϊόντων μονάδων δευτερογενών διεργασιών εξευγενισμού των αρχικών κλασμάτων απόσταξης του πετρελαίου. Τυπικά κλάσματα ανάμιξης είναι τα προϊόντα καταλυτικής αναμόρφωσης και πυρόλυσης, ισομερισμού, αλκυλίωσης, καθώς ελαφρά κλάσματα κορυφής άλλων μονάδων (βουτάνια), τα οποία μαζί με άλλα βελτιωτικά πρόσθετα συμμετέχουν σε συνταγές μίξης, τέτοιες ώστε το τελικό προϊόν να πληρεί τις υπάρχουσες από την νομοθεσία προδιαγραφές ιδιοτήτων καθώς και τα περιβαλλοντικά πρότυπα. Χαρακτηριστικές ιδιότητες της βενζίνης είναι η πίεση ατμού Reid (RVP), τα σημεία απόσταξης κατά ASTM, το flash point, η περιεκτικότητα σε αρωματικές ενώσεις και θείο κ.α., με τον αριθμό οκτανίων (A.O.) να αποτελεί έναν από τους σημαντικότερους ποιοτικούς δείκτες της. Ο A.O. χαρακτηρίζει το καύσιμο ως προς την αντικροτική του συμπεριφορά, δηλαδή την ικανότητά του να μη προκαλεί πρόωρη ανάφλεξη στις μηχανές εσωτερικής καύσης. Στην παρούσα εργασία ως A.O. αναφέρονται, τόσο ο ερευνητικός αριθμός οκτανίων (Research Octane Number, RON), όσο και ο αριθμός οκτανίων κινητήρα (Motor Octane Number, MON) κατά τα πρότυπα της ASTM. Στο παρελθόν η χρήση αλκυλιωμένων παραγώγων του μολύβδου ως πρόσθετων στη βενζίνη εξασφάλιζε τους απαιτούμενους υψηλούς A.O.. Με την απαγόρευση όμως της χρήσης των προσθέτων αυτών για περιβαλλοντικούς λόγους (από τις αρχές του 2002 και στην Ελλάδα) η έρευνα εστίασε σε νέους τρόπους παραγωγής βενζινών υψηλού A.O.. Έτσι ο τεχνολογικός σχεδιασμός των διυλιστηρίων αναμορφώθηκε για να καλυφθεί η διαφορά των 3-6 A.O. (ανάλογα με την προδιαγραφή), η οποία προέκυψε από τον αποκλεισμό των προσθέτων μολύβδου.

Μία σημαντική εφεδρεία των διυλιστηρίων για την εξασφάλιση των επιδιωκόμενων τιμών-στόχων του A.O. είναι η βελτιστοποίηση της διαδικασίας μίξης των επιμέρους κλασμάτων που χρησιμοποιούνται για την παραγωγή της τελικής βενζίνης. Τόσο το κόστος, όσο και η διαθεσιμότητα των κλασμάτων αυτών ποικίλει όχι μόνο μεταξύ διυλιστηρίων, αλλά ακόμη και στο ίδιο διυλιστήριο σε διαφορετικές χρονικές περιόδους, ενώ οι τιμές του A.O. εξαρτώνται μη γραμμικά από την σύνθεση του μίγματος –εξάρτηση που αναφέρεται στην βιβλιογραφία ως επίδραση μίξης (blend effect). Η φύση του φαινομένου μίξης δεν έχει πλήρως διερευνηθεί με συνέπεια την απουσία αναλυτικών σχέσεων υπολογισμού του A.O. από τις συγκεντρώσεις των συστατικών του τελικού προϊόντος. Η συνήθης πρακτική μίξης στηρίζεται σε ημι-εμπειρικά γραμμικά κυρίως μοντέλα πρόβλεψης του A.O. του μίγματος. Σήμερα είναι διαθέσιμα στην αγορά αυτοματοποιημένα συστήματα μίξης, τα οποία προσδιορίζουν με υπέρυθρους αναλυτές

τον αριθμό οκτανίων των διαθέσιμων για μίξη κλασμάτων και καθορίζουν την αναλογία μίξης για την εξασφάλιση του Α.Ο.-στόχου, ικανοποιώντας ταυτόχρονα μια συνάρτηση ελαχιστοποίησης του κόστους. Τα συστήματα αυτά έχουν υψηλό κόστος εγκατάστασης και προσαρμογής σε κάθε διυλιστήριο. Ως εκ τούτου υπάρχει έντονο ενδιαφέρον από τις διυλιστηριακές μονάδες για ανάπτυξη αλγορίθμων πρόβλεψης του Α.Ο. που να χαρακτηρίζονται από την απαιτούμενη ακρίβεια και ευρωστία (διατήρηση της ακρίβειας πρόβλεψης κατά τις συνήθεις μεταβολές στην τροφοδοσία της μονάδας ανάμιξης) και μπορούν να χρησιμοποιηθούν είτε εντός (συνεχής κύκλος λειτουργίας) ή εκτός της ροής παραγωγής στην μονάδα ανάμιξης.

Μέχρι σήμερα έχουν αναπτυχθεί διάφορα μοντέλα πρόβλεψης του Α.Ο. σε μίγματα βενζινών, τα οποία ενσωματώνουν την μη γραμμική εξάρτηση του Α.Ο. μοντελοποιώντας την με συναρτήσεις που περιλαμβάνουν ένα γραμμικό τμήμα και έναν μη γραμμικό όρο διόρθωσης. Η γνωστή στην πετρελαϊκή βιομηχανία μέθοδος Ethyl RT-70 (Healy, κ.α. (1959)) εφαρμόζει μη γραμμική εξάρτηση της ευαισθησίας του καυσίμου (RON-MON) συναρτήσει της περιεκτικότητας σε ολεφίνες και αρωματικές ενώσεις των επιμέρους συστατικών του. Στην ίδια κατεύθυνση κινείται και το μοντέλο που παρουσιάστηκε από τον Stewart (1959), και το οποίο εκφράζει την μη γραμμικότητα μέσω της περιεκτικότητας σε ολεφίνες των κλασμάτων της βενζίνης. Ακολουθώντας την φιλοσοφία εισαγωγής μη γραμμικότητας σε ένα βασικό γραμμικό μοντέλο, οι Gary και Handwerl (1994) ανέπτυξαν την μέθοδο των αριθμών οκτανίων μίγματος (Blend Octane Numbers, BONs). Τα κλάσματα της βενζίνης αναμιγνύονται προσθετικά ανάλογα με τις ογκομετρικές τους συστάσεις ως προς τους αριθμούς BON τους, οι οποίοι αντικαθιστούν τους πραγματικούς αριθμούς RON και MON. Οι αριθμοί BON προσδιορίζονται με ανάλυση παλινδρόμησης συναρτήσει των RON και MON εισάγοντας την απαραίτητη μη γραμμικότητα στο μοντέλο. Μια παρόμοια μέθοδος που μετασχηματίζει τους μη γραμμικούς αριθμούς οκτανίων μίγματος σε γραμμικά εξαρτώμενα μεγέθη μίγματος παρουσιάστηκε από τους Rusin, κ.α. (1981). Μία άλλη προσέγγιση αποτελεί η ανάπτυξη γραμμικών μοντέλων που περιλαμβάνουν και δευτεροβάθμιους όρους των αλληλεπιδράσεων μεταξύ των συστατικών, οπότε η επίδραση μίξης στην εκτίμηση του αριθμού οκτανίων εκφράζεται από το σύνολο των συντελεστών των αλληλεπιδράσεων αυτών (Modgometry (1991), Zahed, κ.α. (1993)).

Αν και η απόκλιση του Α.Ο. μίγματος από την αυστηρά προσθετική συμπεριφορά είναι σχετικά μικρή, η χρήση γραμμικού μοντέλου ως προς τις συστάσεις των κλασμάτων οδηγεί σε ανεπιθύμητα σφάλματα υπολογισμού μερικών βαθμών Α.Ο.. Επιπρόσθετα, η σχετικά μικρή πρόοδος στην μελέτη της χημικής δομής σε σχέση με τον Α.Ο. δεν έχει βοηθήσει ουσιαστικά στην ανάπτυξη τροποποιημένων γραμμικών μοντέλων που να ενσωματώνουν με μη γραμμικούς όρους τα φυσικά αίτια αυτής της απόκλισης. Τα τεχνητά νευρωνικά δίκτυα λόγω της εξαιρετικής ευελιξίας που εμφανίζουν στην προσέγγιση συνάρτησης μπορούν να αποδειχθούν

πολύτιμα εργαλεία για την ανάπτυξη μη γραμμικών μοντέλων πρόβλεψης Α.Ο.. Σε πρόσφατη μελέτη των Murty και Rao (2004) αναπτύχθηκε ένα μοντέλο τεχνητού νευρωνικού δικτύου για την πρόβλεψη του αριθμού RON μίγματος βενζίνης συναρτήσει των ογκομετρικών συγκεντρώσεων πέντε κλασμάτων μίξης και ορισμένων ιδιοτήτων της βενζίνης, όπως η πυκνότητα, το RVP, τα σημεία απόσταξης και η περιεκτικότητα σε θείο. Οι ιδιότητες όμως που χρησιμοποιήθηκαν ως μεταβλητές εισόδου του μοντέλου υπολογίζονται προσθετικά από τις συστάσεις των επιμέρους κλασμάτων μη προσδίδοντας έτσι επιπλέον πληροφορία. Ένα αρκετά αξιόπιστο μοντέλο πρόβλεψης του αριθμού RON που αναπτύχθηκε με νευρωνικά δίκτυα και ενσωματώνει τις τυχόν ποιοτικές μεταβολές στα κλάσματα τροφοδοσίας παρουσιάστηκε από τους Πασαδάκη, κ.α. (2006). Ως μεταβλητές εισόδου χρησιμοποιούνται, εκτός από τις ογκομετρικές συγκεντρώσεις των κλασμάτων στο μίγμα, και οι Α.Ο. των επιμέρους κλασμάτων ζυγισμένοι με τις αντίστοιχες συγκεντρώσεις.

Στην παρούσα εργασία, ακολουθώντας την ίδια προσέγγιση, αναπτύχθηκαν με χρήση νευρωνικών δικτύων μοντέλα πρόβλεψης και των Α.Ο. (RON και MON) μιγμάτων βενζινών με μεταβλητές εισόδου τους αντίστοιχους Α.Ο. οκτώ κυρίων κλασμάτων μίξης ζυγισμένους με τις ογκομετρικές περιεκτικότητες αυτών. Τα κλάσματα που χρησιμοποιήθηκαν περιελάμβαναν τα προϊόντα από τις μονάδες καταλυτικής πυρόλυσης ρευστοποιημένης κλίνης, καταλυτικής αναμόρφωσης, ισομερισμού, αλκυλίωσης και διμερισμού, μαζί με τα κλάσματα βουτανίων και πεντανίων και οξυγονούχα πρόσθετα αύξηση του Α.Ο.. Τα δίκτυα εκπαιδεύτηκαν σε ένα μεγάλο πληθυσμό δειγμάτων της τυπικής παραγωγής του διυλιστηρίου της Motor Oil Hellas (MOH) και προβλέπουν αξιόπιστα τους αριθμούς οκτανίων με σφάλμα της τάξης των 0.2 βαθμών Α.Ο., μικρότερο από την επαναληψιμότητα των πρότυπων εργαστηριακών μεθόδων προσδιορισμού τους κατά ASTM. Επιπλέον έγινε μια προσπάθεια διερεύνησης της επίδρασης των παραπάνω κλασμάτων μίξης τόσο μεμονωμένα, όσο και των αλληλεπιδράσεων τους ανά δύο στην διαμόρφωση του Α.Ο. του τελικού μίγματος, με την ανάπτυξη γραμμικού ως προς τους συντελεστές του μοντέλου που περιλαμβάνει και δευτεροβάθμιους όρους αλληλεπιδράσεων μεταξύ των συστατικών.

Όλες οι παραπάνω μελέτες πρόβλεψης του Α.Ο. βασίζονται σε δεδομένα παραγωγής, ενσωματώνοντας προηγούμενη εμπειρία στην διαδικασία ανάμιξης. Το εύρος των συστάσεων των επιμέρους συστατικών μίγματος βενζίνης καλύπτει τα συνήθη όρια υπό όρους διαθεσιμότητας και κόστους των κλασμάτων τροφοδοσίας σε μία τυπική παραγωγή. Οι περιοχές όμως του χώρου δράσης του φαινομένου μίξης που καλύπτουν είναι περιορισμένες, αφού πρόκειται για μίγματα με βάση εμπειρικές “συνταγές” που εστιάζουν στις επιθυμητές αποκρίσεις ανάλογα με τις προδιαγραφές σε Α.Ο. του τελικού εμπορικού προϊόντος, με συνέπεια να είναι αμφισβητήσιμη η ικανότητα πρόβλεψης των μοντέλων σε μίγματα διαφοροποιημένης σύστασης.

Κατά την παρούσα μελέτη επιχειρήθηκε η ανάπτυξη γενικότερων μοντέλων πρόβλεψης των τιμών των αριθμών οκτανίων (RON και MON) σε βενζίνες με βάση εργαστηριακά παρασκευασμένα μίγματα που παρουσιάζουν μεγαλύτερη ποικιλία Α.Ο.. Με χρήση νευρωνικών δικτύων αναπτύχθηκαν αλγόριθμοι πρόβλεψης των Α.Ο. που εμφανίζουν ικανοποιητική ακρίβεια παρά το μικρό σχετικά πλήθος δεδομένων εκπαίδευσης. Μία πιο εκτεταμένη έρευνα στην κατεύθυνση αυτή μπορεί να οδηγήσει σε καθορισμό της γενικότερης επιφάνειας απόκρισης που παρουσιάζει το φαινόμενο, συμβάλλοντας έτσι στην κατανόηση της επίδρασης μίξης αλλά και τον επαναπροσδιορισμό των περιοχών βέλτιστης σύστασης της βενζίνης.

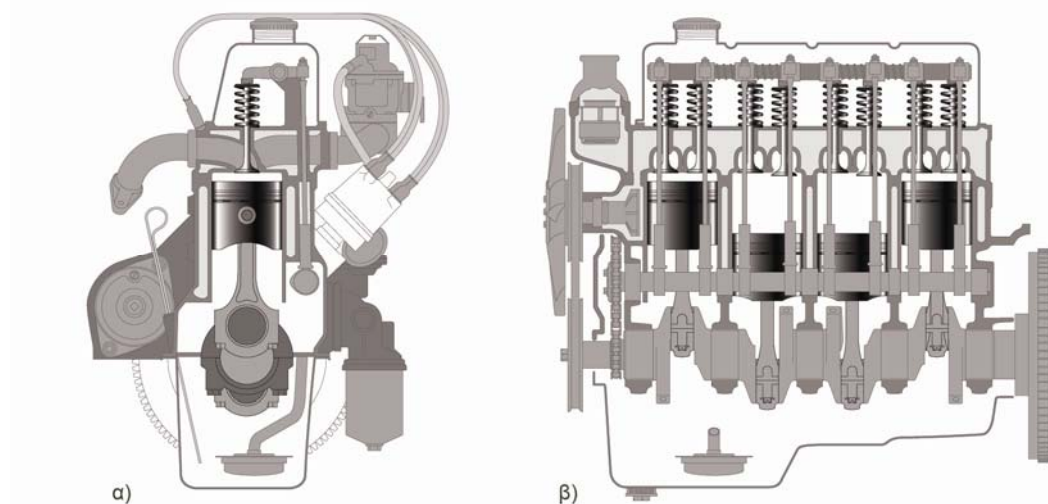
Τέλος, στην εργασία αυτή παρουσιάζεται μία αναλυτική περιγραφή των ιδιοτήτων της βενζίνης και ιδιαίτερα της αντικροτικής συμπεριφοράς της. Γίνεται επίσης συνοπτική αναφορά των θεωρητικών “εργαλείων” –χημειομετρικές μέθοδοι ανάλυσης πολυπαραμετρικών δεδομένων, τεχνητά νευρωνικά δίκτυα και στατιστικός πειραματικός σχεδιασμός– που χρησιμοποιήθηκαν.

1 Βενζίνη και Αριθμός Οκτανίων

1.1 Αριθμός Οκτανίων

Ο **αριθμός οκτανίων**, A.O. (octane rating or number) αποτελεί μία από τις πιο ενδεικτικές για την ποιότητά της ιδιότητα της βενζίνης και χαρακτηρίζει το καύσιμο ως προς την ικανότητα του να αποτρέπει την πρόωρη έκρηξη λόγω αυτανάφλεξης στις μηχανές εσωτερικής καύσης, ή αλλιώς ως προς την **αντικροτική συμπεριφορά** του (antiknock property). Η ενεργειακή απόδοση του κινητήρα εξαρτάται άμεσα από τον αριθμό οκτανίων του καυσίμου, αφού η χρήση βενζινών υψηλού αριθμού οκτανίων επιτρέπει την αύξηση της επιτρεπόμενης αναλογίας συμπίεσης. Ως αριθμός οκτανίων βενζίνης ορίζεται η % κ.ο. σε ισο-οκτάνιο περιεκτικότητα δυαδικού μίγματος κανονικού επτανίου-ισοοκτανίου, το οποίο εμφανίζει την ίδια κροτική συμπεριφορά με την εξεταζόμενη βενζίνη σε πρότυπο κινητήρα.

Κύκλος Λειτουργίας Μηχανών Εσωτερικής Καύσης

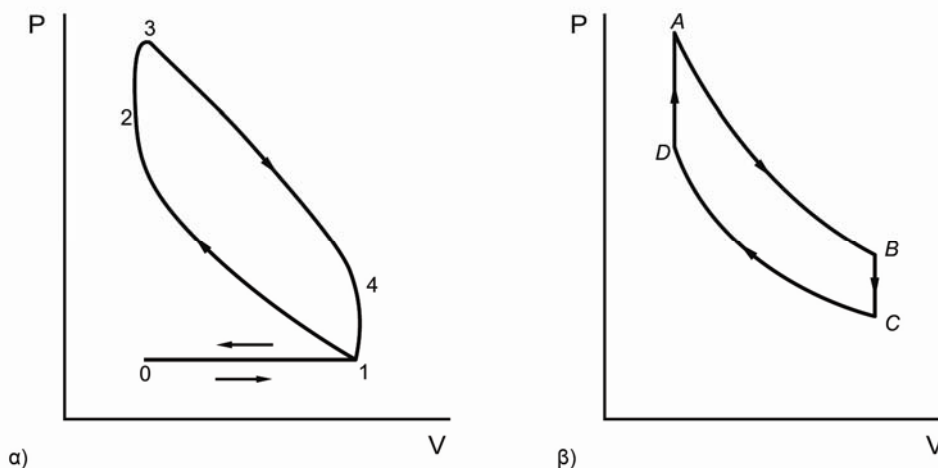


Σχήμα 1-1 Τετρακύλινδρος κινητήρας εσωτερικής καύσης. α) Μετωπική και β) πλευρική τομή

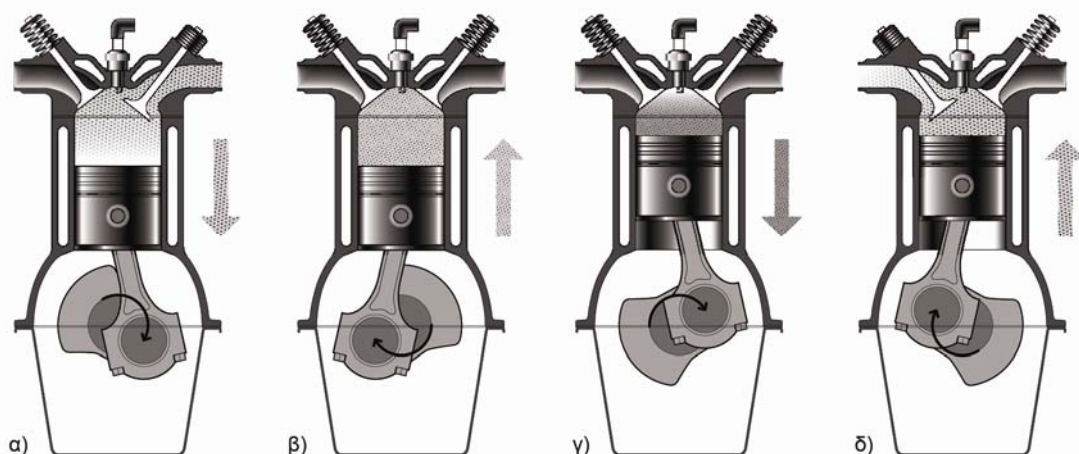
Στις εμβολοφόρες μηχανές εσωτερικής καύσης, Μ.Ε.Κ., η ενέργεια κίνησης παράγεται στον κινητήρα (Σχ. 1-1) με μετατροπή της θερμότητας, που εκλύεται κατά την καύση του καυσίμου (μίγμα βενζίνης και αέρα) εντός κλειστών κυλίνδρων, σε μηχανική ενέργεια. Μετά την ανάμιξη του στον εξαερωτήρα (καρμπυρατέρ), το καύσιμο αναρροφάται από τον θάλαμο καύσης στο άνω τμήμα κάθε κυλίνδρου, στο εσωτερικό του οποίου συμπιέζεται και στην συνέχεια αναφλέγεται από τον σπινθηριστή (μπουζί). Η αύξηση του όγκου του μίγματος από την δημιουργία των προϊόντων της καύσης, ωθεί το έμβολο σε προς τα κάτω κίνηση. Η

παλινδρομική κίνηση του εμβόλου μετατρέπεται μέσω του διωστήρα (μπιέλα) σε περιστροφική του στροφαλοφόρου άξονα.

Ο κύκλος λειτουργίας στους βενζινοκινητήρες είναι ο τετράχρονος ή κύκλος του Otto (Σχ. 1-2α)) που αποτελείται από τέσσερα στάδια που αντιστοιχούν σε τέσσερις διαδρομές του εμβόλου (Σχ. 1-3). Στο πρώτο στάδιο, όπου η βαλβίδα εισαγωγής του καυσίμου είναι ανοιχτή και η βαλβίδα εξαγωγής κλειστή, το έμβολο κινείται προς τα κάτω, υπό σταθερή πίεση, αναρροφώντας το μίγμα αέρα-καυσίμου στο εσωτερικό του κυλίνδρου. Η κίνηση αυτή αντιπροσωπεύεται από την μεταβολή $0 \rightarrow 1$ του θερμοδυναμικού διαγράμματος. Κατά το δεύτερο στάδιο (μεταβολή $1 \rightarrow 2 \rightarrow 3$), όπου και οι δύο βαλβίδες είναι κλειστές, το μίγμα με άνοδο του εμβόλου συμπιέζεται σχεδόν αδιαβατικά κατά μήκος της γραμμής $1 \rightarrow 2$ και εξαερώνεται λόγω της αύξησης της θερμοκρασίας. Κατόπιν το συμπιεσμένο μίγμα αναφλέγεται, με έναυση από το σπινθηριστή, με ταχεία καύση που δεν αφήνει περιθώρια μεταβολής του όγκου, ενώ η πίεση εξακολουθεί να αυξάνεται (μεταβολή $2 \rightarrow 3$). Στο τρίτο στάδιο, που είναι και το μοναδικό του κύκλου με παραγωγή έργου, τα προϊόντα της καύσης εκτονώνονται σχεδόν αδιαβατικά κατά μήκος της γραμμής $3 \rightarrow 4$, ωθώντας το έμβολο προς τα έξω. Η βαλβίδα εξόδου ανοίγει και η πίεση ελαττώνεται ταχύτατα (μεταβολή $4 \rightarrow 1$). Στο τελευταίο στάδιο το έμβολο ανέρχεται για την απαγωγή των εναπομεινάντων καυσαερίων δια της ανοιχτής βαλβίδας εξαγωγής κατά την ισοβαρή $1 \rightarrow 0$.



Σχήμα 1-2 Θερμοδυναμικά διαγράμματα α) διεργασιών της μηχανής Otto και β) πρότυπου κύκλου Otto

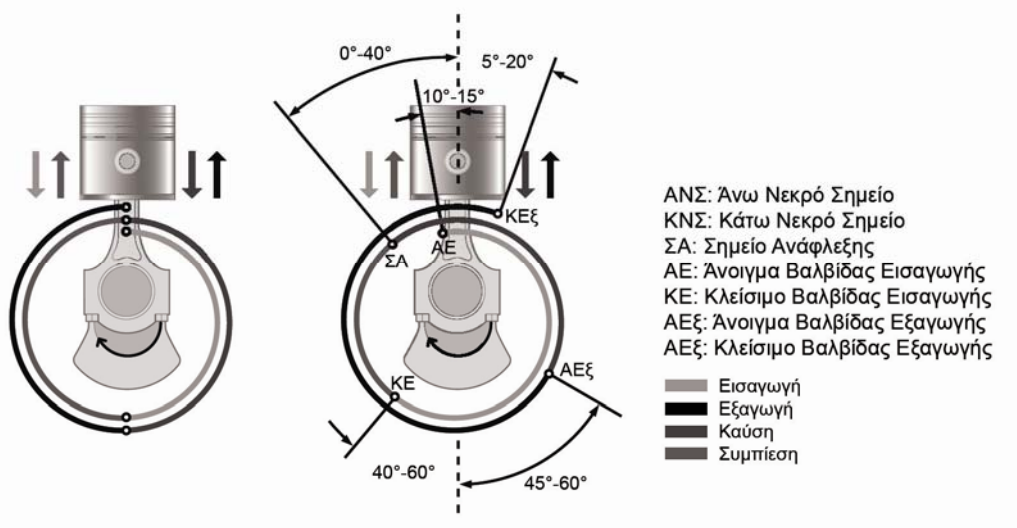


Σχήμα 1-3 Τετράχρονος κύκλος λειτουργίας. α) Εισαγωγή, β) συμπίεση, γ) εκτόνωση και δ) εξαγωγή

Για έναν πλήρη κύκλο λειτουργίας (συμπίεση, καύση, εκτόνωση και εναλλαγή των αερίων), απαιτούνται δύο πλήρεις περιστροφές του στροφαλοφόρου άξονα (720° περιστροφής) ή αλλιώς δύο πλήρεις κινήσεις του συστήματος εμβόλου-διωστήρα με αρχή την ανώτερη θέση του εμβόλου που καλείται **Άνω Νεκρό Σημείο** (Α.Ν.Σ.) της διαδρομής του. Ο συνολικός κύκλος μπορεί αποδοθεί με κυκλικό διάγραμμα χρόνων¹ των τεσσάρων φάσεων λειτουργίας, όπως αυτό του Σχ. 1-4), όπου απεικονίζονται οι χρονισμοί ανοίγματος και κλεισίματος των βαλβίδων εισαγωγής-εξαγωγής που αποτελούν τα όρια κάθε φάσης. Είναι εμφανές ότι στην πραγματικότητα υπάρχει επικάλυψη στους χρονισμούς των βαλβίδων, ώστε το άνοιγμα και κλείσιμο να μην συμπίπτει ακριβώς με την αρχή και το τέλος των φάσεων λειτουργίας, όπως αυτές ορίζονται θεωρητικά (Σχ. 1-4) από την παλινδρομική κίνηση του εμβόλου. Η βαλβίδα εξαγωγής ανοίγει πριν (45° - 60°) το έμβολο φτάσει στο **Κάτω Νεκρό Σημείο** (Κ.Ν.Σ.) και κλείνει λίγο μετά το Α.Ν.Σ. (5° - 15°), ενώ η βαλβίδα εισαγωγής ανοίγει πριν το Α.Ν.Σ. (10° - 15°) και κλείνει μετά το Κ.Ν.Σ. (40° - 60°), ώστε να για ένα μικρό χρονικό διάστημα να παραμένουν ταυτόχρονα ανοιχτές, επιτρέποντας ροή αερίων προς και από τον κύλινδρο που διευκολύνει την πλήρωσή του με μίγμα καυσίμου και απαγωγή των καυσαερίων. Επιπλέον, δεδομένου ότι στην πραγματικότητα απαιτείται πεπερασμένο χρονικό διάστημα για την ολοκλήρωση της καύσης όλο του καυσίμου, η έναυση του σπινθηριστή ρυθμίζεται λίγο πριν (0° - 40°) το Α.Ν.Σ., ώστε η μέγιστη πίεση να παρουσιαστεί μετά το Α.Ν.Σ. (10° - 20°) για την εφαρμογή στο έμβολο της μέγιστης δυνατής ώθησης κατά την εκτόνωση.

¹ Πρόκειται στην πραγματικότητα για διάγραμμα μεταβολής της γωνίας περιστροφής του στροφάλου από το Α.Ν.Σ., η οποία αντιστοιχεί σε χρονική στιγμή (για $t_0=0\text{sec}$ ο στρόφαλος βρίσκεται στο Α.Ν.Σ.). Για γνωστή ταχύτητα περιστροφής του κινητήρα n (rpm), η γωνιακή ταχύτητα ω (rad/sec) δίνεται από την σχέση:

$$\omega = 2\pi n = \frac{2\pi}{60} n = \frac{\delta\phi}{\delta t} \Rightarrow \delta\phi \text{ (rad)} = \frac{2\pi}{60} n \delta t \text{ (sec)} \quad \text{ή} \quad \delta\phi \text{ (}^\circ\text{)} = \frac{2\pi}{60} \frac{360}{2\pi} n \delta t = 6 n \delta t \text{ (sec)}$$



Σχήμα 1-4 Διάγραμμα χρόνων φάσεων ενός τετράχρονου κινητήρα

Η απόδοση του κινητήρα εξαρτάται από την ποσότητα της θερμικής ενέργειας που μετατρέπεται σε έργο. Για το υπολογισμό του *συντελεστή θερμικής απόδοσης* γίνεται υπόθεση κυκλικής μηχανής (Σχ. 1-2β)) που λειτουργεί με *ιδανικό αέριο* (αέρα) και έχει την ίδια απόδοση με την Μ.Ε.Κ. του Otto. Η απόδοση δίνεται από την σχέση:

$$\eta = \frac{-W_s}{Q_{DA}} \quad (1.1)$$

όπου W_s είναι το αξονικό έργο που ανταλλάσσει το σύστημα με το περιβάλλον μέσω του εκκεντροφόρου άξονα και Q_{DA} η θερμότητα που απορροφά το σύστημα από το περιβάλλον κατά την θεωρητική ισοβαρή μεταβολή $D \rightarrow A$, που αντιστοιχεί στην ταχεία καύση του μίγματος καυσίμου εντός του κυλίνδρου.

Η εξίσωση (1.1) μετασχηματίζεται με την χρήση των *καταστατικών εξισώσεων* ιδανικού αερίου στην ακόλουθη μορφή:

$$\eta = 1 - \frac{1}{r^{\gamma-1}} \quad (1.2)$$

όπου $r = V_c/V_D$ η **αναλογία συμπίεσης** (compression ratio) που αντιστοιχεί στον λόγο των όγκων του μίγματος πριν και μετά το στάδιο της συμπίεσης και $\gamma = C_p/C_v$ ο λόγος των *ειδικών θερμοτήτων* του ρευστού υπό σταθερή πίεση και όγκο αντίστοιχα.

Η τελευταία σχέση καταδεικνύει την άμεση εξάρτηση της απόδοσης του θεωρητικού κύκλου για κλειστό σύστημα, και επομένως του παραγόμενου έργου της Μ.Ε.Κ. ανά μονάδα μάζας του καυσίμου, από την αναλογία συμπίεσης. Έλεγχοι πάνω σε πραγματικές μηχανές Otto επιβεβαιώνουν ότι για χαμηλές τιμές της αναλογίας συμπίεσης, ο συντελεστής θερμικής απόδοσης αυξάνεται ταχύτατα με το r , ενώ για μεγαλύτερες τιμές της ο ρυθμός αύξησης της

απόδοσης ελαττώνεται. Τυπική τιμή αναλογίας συμπίεσης είναι 9:1, ενώ η μέγιστη απόδοση μίας Μ.Ε.Κ. επιτυγχάνεται για αναλογία συμπίεσης 17:1.

Προσδιορισμός Αριθμού Οκτανίων Βενζίνης

Μέχρι το 1912 η αναλογία συμπίεσης στις μηχανές εσωτερικής καύσης ήταν περιορισμένη λόγω του ανεπιθύμητου “κτύπου” που οδηγούσε σε γρήγορη φθορά της μηχανής. Αρχικά υπεύθυνο για την ανώμαλη καύση θεωρήθηκε το ηλεκτρικό σύστημα ανάφλεξης που προστέθηκε στα αυτοκίνητα παράλληλα με την μίζα (αυτόματη ηλεκτρική εκκίνηση). Στην προσπάθεια προσδιορισμού των αιτιών του “κτύπου”, διεξήχθησαν μελέτες της μεταβολής της πίεσης κατά το στάδιο της καύσης σε πρότυπο κινητήρα, που έδειξαν ότι το φαινόμενο δεν οφειλόταν σε προανάφλεξη, όπως θεωρήθηκε αρχικά, αλλά σε απότομη άνοδο της πίεσης μετά την ανάφλεξη. Η χρήση μιας φωτογραφικής μηχανής υψηλής ταχύτητας (Midgley & Boyd) παράλληλα με την εισαγωγή ενός δείκτη “αναπήδησης βελόνας” για τον ποσοτικό χαρακτηρισμό του “κτύπου”, συνέβαλλε σημαντικά στην ερμηνεία του φαινομένου. Αργότερα εισήχθει (Ricardo,) η έννοια της **μέγιστης ωφέλιμης αναλογίας συμπίεσης** (Highest useful compression Ratio, HUCR), για τον προσδιορισμό της οποίας χρησιμοποιήθηκε μια μηχανή μεταβλητής συμπίεσης.

Η έννοια του αριθμού οκτανίων εισήχθει από τον Graham Edgar (1927), που πρότεινε μία κλίμακα μέτρησης της αντικροτικής συμπεριφοράς της βενζίνης σε σχέση με αυτή που εμφανίζει δυαδικό μίγμα υδρογονανθράκων σε διάφορες αναλογίες. Για το πρότυπο αυτό μίγμα προτάθηκαν δύο υδρογονάνθρακες που μπορούσαν να παραχθούν σε ικανοποιητική καθαρότητα και ποσότητα, το “κανονικό επτάνιο”, που παράγεται με απόσταξη ελαίου πεύκων Jeffrey, και ένα συνθετικό οκτάνιο, το “ισοοκτάνιο” ή “2, 4, 4-τριμεθυλοπεντάνιο”. Οι υδρογονάνθρακες αυτοί θεωρήθηκαν ιδανικοί λόγω της τεράστιας διαφοράς στην αντικροτική συμπεριφορά τους (δόθηκαν οι αυθαίρετες τιμές 0 και 100 για τα καθαρά τους μίγματα αντίστοιχα), αλλά και των παρόμοιων πτητικών ιδιοτήτων τους (Πιν. 1-1), που διασφαλίζουν την σταθερή πτητική συμπεριφορά των μιγμάτων τους στις διάφορες αναλογίες από 0:100 ως 100:0.

Πίνακας 1-1 Πτητικές ιδιότητες κανονικού επτανίου και ισοοκτανίου

	Σημείο τήξης (°C)	Σημείο βρασμού (°C)	Πυκνότητα (g/ml)	Θερμότητα ατμοποίησης (MJ/kg σε 25°C)
κανονικό επτάνιο	-90.7	98.4	0.684	0.365
ισοοκτάνιο	-107.45	99.3	0.6919	0.308

Με πρότυπο το παραπάνω μίγμα αναφοράς αναπτύχθηκε μία ολόκληρη σειρά μηχανών και συνθηκών δοκιμής, αλλά σήμερα χρησιμοποιούνται σχεδόν αποκλειστικά οι μέθοδοι προσδιορισμού του **ερευνητικού αριθμού οκτανίων** (Research Octane Number, RON), και του **αριθμού οκτανίων κινητήρα** (Motor Octane Number, MON). Η ανάγκη διεθνούς τυποποίησης για τον χαρακτηρισμό των καυσίμων καθιέρωσε για αρκετές δεκαετίες την πρότυπη μέθοδο ερευνητικής δοκιμής (Research Test Method) κατά ASTM (American Society for Testing and Materials). Οι συνθήκες λειτουργίας της δοκιμής (Πιν. 1-2) ήταν αντιπροσωπευτικές για τις απαιτήσεις των οχημάτων σε καύσιμα μέχρι τα μέσα του '60, οπότε η ανάπτυξη της τεχνολογίας των κινητήρων (μεγαλύτερες ταχύτητες και επιταχύνσεις, αυξημένο φορτίο) οδήγησε σε αποκλίσεις της συμπεριφοράς των καυσίμων από τον προδιαγραφόμενο ερευνητικό αριθμό οκτανίων τους. Για τον λόγο αυτό αναπτύχθηκε η πρότυπη μέθοδος δοκιμής κινητήρα (Motor Test Method) που χρησιμοποιεί εντονότερες συνθήκες κατά την δοκιμή καυσίμου σε πρότυπο κινητήρα.

Πίνακας 1-2 Συνθήκες δοκιμής των μεθόδων προσδιορισμού αριθμού οκτανίων

Συνθήκες Λειτουργίας	RON	MON
Μέθοδος δοκιμής	ASTM D2699-92	ASTM D2700-92
Μηχανή	Cooperative Fuels Research (CFR)	Cooperative Fuels Research (CFR)
Ταχύτητα Μηχανής	600 RPM	900 RPM
Θερμοκρασία Αέρα Εισαγωγής	ποικίλλει με την βαρομετρική πίεση 88kPa σε 19.4°C, 101.6kPa σε 52.2°C	38°C
Υγρασία Αέρα Εισαγωγής	3.56-7.12g H ₂ O/ kg ξηρού αέρα	3.56-7.12g H ₂ O/ kg ξηρού αέρα
Θερμοκρασία Μίγματος Εισαγωγής	δεν καθορίζεται	149°C
Θερμοκρασία Ψύξης	100°C	100°C
Θερμοκρασία Λιπαντικού	57°C	57°C
Έναυση σπινθηριστή	13° BTDC*	ποικίλλει με την αναλογία συμπίεσης 14°-26° BTDC
Χοάνη εξαερωτήρα	ποικίλλει με το υψόμετρο 14.3mm σε 0-500m, 15.1mm σε 500-1000m	14.3 mm

*Before Top Dead Cycle

Και οι δύο μέθοδοι χρησιμοποιούν μία ειδική μονοκύλινδρο μηχανή μεταβλητής συμπίεσης (4:1-18:1), γνωστή ως μηχανή CFR (Co-operative Fuel Research), με κύλινδρο διαμέτρου 82.5mm και διαδρομή εμβόλου 114.3mm. Η κεφαλή και ο κύλινδρος αποτελούν ενιαίο τμήμα που έχει την δυνατότητα να μετακινείται για ρύθμιση της αναλογίας συμπίεσης, ενώ ο εξαερωτήρας είναι

τεσσάρων χοάνων με δυνατότητα ρύθμισης διαφορετικής αναλογίας αέρα-καυσίμου σε καθεμία απ' αυτές. Ένας αισθητήρας στο θάλαμο συμπίεσης μετρά τις ταχείες μεταβολές της πίεσης που προκαλεί ο κτύπος και το ενισχυμένο σήμα μετράται σε μετρητή κτύπου με κλίμακα 0-100. Για τον προσδιορισμό του Α.Ο. ενός καυσίμου η μηχανή ρυθμίζεται σε αναλογία συμπίεσης, τέτοια ώστε, για αναλογία αέρα-καυσίμου στον εξαερωτήρα που να παράγει το μέγιστο δυνατό κτύπο, να λαμβάνεται για δείγμα του καυσίμου τιμή 50 ± 2 στο μετρητή κτύπου (knockmeter). Παρασκευάζονται δύο μίγματα επτανίου-ισοοκτανίου με τιμές Α.Ο. μικρότερη και μεγαλύτερη αντίστοιχα από την αναμενόμενη τιμή του Α.Ο. του εξεταζόμενου δείγματος και τοποθετούνται σε διαφορετικές χοάνες του εξαερωτήρα, οι οποίες επίσης ρυθμίζονται σε αναλογίες αέρα-καυσίμου μέγιστου κτύπου. Ο Α.Ο. του καυσίμου υπολογίζεται με γραμμική παρεμβολή των μετρήσεων του μετρητή κτύπου για το εξεταζόμενο δείγμα και για τα δείγματα αναφοράς (περίπου 60-70 για αυτό με μικρότερο Α.Ο. και 30-40 για εκείνο με μεγαλύτερο Α.Ο.).

Ο σχεδιασμός της μηχανής και του οχήματος επηρεάζει τις απαιτήσεις τόσο σε RON, όσο και σε MON του καυσίμου, με τις περισσότερες μηχανές σήμερα να εμφανίζουν τον λεγόμενο περιορισμό του ενός βαθμού οκτανίων, που σημαίνει μία μεταβολή ενός βαθμού RON είναι ασφαλής αν συνοδεύεται από ίση και αντίθετη μεταβολή του MON. Για τον πλήρη χαρακτηρισμό μίας βενζίνης, επομένως, δεν αρκεί ο προσδιορισμός ενός μόνο από τους δύο αριθμούς και χρησιμοποιείται εκτός από τον δείκτη αντικροτικής συμπεριφοράς $(RON+MON)/2$, και η **ευαισθησία καυσίμου** (fuel sensitivity), που υπολογίζεται ως διαφορά των δύο αριθμών οκτανίων του, $RON-MON$, που για τις συνήθεις βενζίνες έχει τιμή περίπου 10. Πρόκειται για δείκτη συμπεριφοράς ενός καυσίμου σε μεταβαλλόμενες συνθήκες λειτουργίας κινητήρα, αφού οι δύο μέθοδοι προσδιορισμού του Α.Ο. διαφέρουν ουσιαστικά στην θερμοκρασία εισαγωγής του μίγματος και την ταχύτητα της μηχανής.

1.2 Βενζίνη

Η βενζίνη είναι προϊόν επεξεργασίας του αργού πετρελαίου. Το μίγμα υδρογονανθράκων που περιλαμβάνει το τελευταίο διαχωρίζεται με κλασματική απόσταξη σε διάφορα κλάσματα, μίγματα υδρογονανθράκων επίσης, η σύσταση των οποίων καθορίζεται από τα όρια της θερμοκρασίας απόσταξης. (Πιν. 1-3). Το αρχικό ακατέργαστο προϊόν της πρωτογενούς αυτής φυσικής διεργασίας που προορίζεται για παραγωγή καυσίμων αυτοκίνησης ανέρχεται σε 30-40% του αργού πετρελαίου που αποστάζεται και αποτελεί την πρώτη ύλη, η οποία με περαιτέρω κατεργασία και ανάμιξη με άλλα συστατικά, δίνει το τελικό προϊόν που αναφέρεται ως **βενζίνη** (gasoline).

Πίνακας 1-3 Προϊόντα κλασματικής απόσταξης πετρελαίου

Προϊόν απόσταξης	Σημείο Βρασμού (°C)	Χρήσεις
αέρια (gases)	<40	• καύσιμο διυλιστηρίου • υγραέριο LPG (Liquified petroleum gas) • πετροχημικά
ελαφριά νάφθα (light naphtha)	25-130	• βενζίνη • διαλύτες
βαριά νάφθα (heavy naphtha)	80-200	• βενζίνη • διαλύτες • καύσιμα αεροπορίας
κηροζίνη (kerosene)	150-250	• καύσιμα αεροπορίας • διαλύτες
ελαφρύ gasoil (light gasoil)	200-320	• πετρέλαιο κίνησης (diesel oil) • πετρέλαιο θέρμανσης
βαρύ gasoil (heavy gasoil)	260-400	• πετρέλαιο κίνησης (diesel oil) • πετρέλαιο θέρμανσης
gasoil απόσταξης υπό κενό (vacuum gasoil)	400-600	• τροφοδοσία μονάδων πυρόλυσης • λιπαντικά
υπόλειμμα ατμοσφαιρικής απόσταξης (residue)	>400	• μαζούτ (fuel oil) • τροφοδοσία απόσταξης υπό κενό
υπόλειμμα απόσταξης υπό κενό (vacuum residue)	>600	• μαζούτ (fuel oil) • ασφαλτος

Οι βενζίνες είναι προϊόντα μίξης κλασμάτων που προκύπτουν από εξειδικευμένες κατεργασίες διύλισης. Λόγω του ό,τι υπήρξαν πάντοτε καθοριστικός παράγοντας για την ανάπτυξη της αυτοκινητοβιομηχανίας, η έρευνα στον τομέα των καυσίμων μετέβαλε συνεχώς την παραγωγή της βενζίνης. Στις αρχές του 20ού η βενζίνη παραγόταν αποκλειστικά ως απλό απόσταγμα του πετρελαίου, οι βελτιώσεις όμως στις μηχανές εσωτερικής καύσης (M.E.K.) επηρέασαν γρήγορα και τις απαιτήσεις στην ποιότητα των καυσίμων. Στην δεκαετία του '10 δοκιμάστηκαν μηχανές εσωτερικής καύσης που χρησιμοποιούσαν ως καύσιμο την κηροζίνη. Το κλάσμα αυτό οδηγούσε σε έντονο "κτύπημα" στον κινητήρα καθιστώντας τις ακατάλληλες. Κατέδειξε όμως την σημαντικότητα του φαινομένου και κατεύθυνε την έρευνα στην αναζήτηση ικανοποιητικών αντικροτικών παραγόντων. Σημαντική καινοτομία υπήρξε για την βελτίωση της αντικροτικής συμπεριφοράς των καυσίμων η προσθήκη τετρααιθυλιούχου μολύβδου.

Για μερικές δεκαετίες η έρευνα αφορούσε σχεδόν αποκλειστικά τις μολυβδωμένες βενζίνες. Μειώνεται η περιεκτικότητα σε θείο που επιδρά αρνητικά στην ικανότητα αύξησης αριθμού οκτανίων του αλκυλικού μολύβδου, ενώ παράλληλα εντοπίζεται η αρνητική επίδραση και των μεγαλύτερων μορίων υδρογονανθράκων και μέχρι τη δεκαετία του '40 εισάγεται η καταλυτική πυρόλυση. Στα μέσα του 20ού αιώνα αυξάνεται η αναλογία συμπίεσης στους κινητήρες των

Μ.Ε.Κ. και επομένως οι απαιτήσεις για καύσιμα υψηλότερων αριθμών οκτανίων. Αυξάνονται οι περιεκτικότητες σε τετρααιθυλιούχο μόλυβδο και μειώνονται η περιεκτικότητα σε θείο και το ποσοστό των ολεφινών, ενώ εισάγονται νέες διαδικασίες διύλισης (υδρογονοπυρόλυση) με σκοπό την παραγωγή κλασμάτων υδρογονανθράκων υψηλού αριθμού οκτανίων και καλύτερη απόκριση στην προσθήκη μολύβδου. Δεν σημειώνονται σημαντικές αλλαγές μέχρι το '70, οπότε για περιβαλλοντικούς λόγους εισήχθησαν οι αμόλυβδες βενζίνες με αυξημένο ποσοστό αρωματικών συστατικών, ενώ η περιεκτικότητα σε θείο και ολεφίνες παραμένει σε σταθερά επίπεδα. Από το 1990, καταργούνται σταδιακά τα πρόσθετα μολύβδου και αντικαθίστανται από οξυγονούχα πρόσθετα σε μια προσπάθεια αύξησης του αριθμού οκτανίων με περιορισμό των ιδιαίτερα τοξικών αρωματικών ενώσεων.

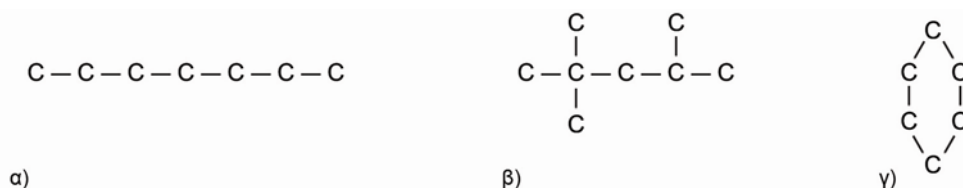
Συστατικά της βενζίνης

Η βενζίνη αποτελεί κατά το μεγαλύτερο μέρος της μίγμα υδρογονανθράκων, ενώ εμπλουτίζεται και με διάφορα βελτιωτικά πρόσθετα που διαμορφώνουν τις τελικές φυσικές και χημικές ιδιότητες του καυσίμου, όπως τα οξυγονούχα πρόσθετα και οι αλκυλικές ενώσεις μολύβδου για την αύξηση του αριθμού οκτανίων. Η βενζίνη περιέχει γύρω στους 500 υδρογονάνθρακες με αλυσίδες τεσσάρων ως δώδεκα ατόμων άνθρακα, με εύρος σημείων βρασμού 30°C-220°C σε ατμοσφαιρική πίεση και πυκνότητες, σχετική 0.71-0.79 g/ml και A.P.I (American Petroleum Institute) 69-48.

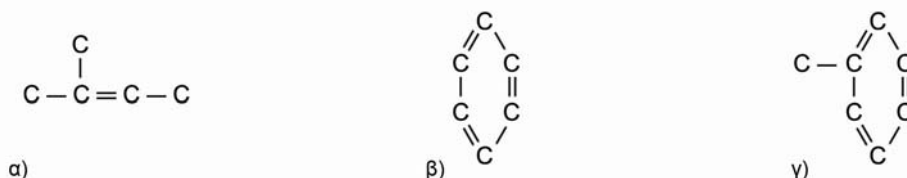
Οι υδρογονάνθρακες της βενζίνης διακρίνονται σε κορεσμένους και ακόρεστους, ανάλογα με το είδος των δεσμών μεταξύ των ατόμων άνθρακα: απλοί και διπλοί ή τριπλοί δεσμοί αντίστοιχα. Οι κορεσμένοι υδρογονάνθρακες (αλκάνια ή παραφίνες) αποτελούν το κύριο συστατικό των μολυβδωμένων βενζινών, καίγονται στον αέρα με λευκή φλόγα και ο αριθμός οκτανίων τους εξαρτάται από το είδος της αλυσίδας και τον αριθμό ατόμων άνθρακα. Τα κανονικά αλκάνια (ευθείες αλυσίδες) με μοριακό τύπο C_nH_{2n+2} εμφανίζουν χαμηλό αριθμό οκτανίων που μειώνεται καθώς αυξάνεται ο αριθμός ατόμων άνθρακα, σε αντίθεση με τα αλκάνια με διακλαδισμένη αλυσίδα (ισο-αλκάνια), τα οποία εμφανίζουν μεγαλύτερους αριθμούς οκτανίων που αυξάνονται με την αύξηση του μεγέθους του μορίου. Τα κυκλικά αλκάνια ή ναφθένια (C_nH_{2n}), όπως το κυκλοεξάνιο, εμφανίζουν επίσης υψηλούς αριθμούς οκτανίων.

Οι ακόρεστοι υδρογονάνθρακες της βενζίνης είναι περισσότερο ασταθείς, λόγω της ύπαρξης διπλών (αλκένια) και τριπλών (αλκίνια) δεσμών ατόμων άνθρακα στο μόριό τους, παρουσιάζουν αυξημένη τοξικότητα και κατά την καύση τους δημιουργούν αιθαλομίχλη. Τα αλκένια ή ολεφίνες, αν και εμφανίζουν τους επιθυμητούς αριθμούς οκτανίων, είναι χημικά ενεργά και τοξικά ώστε περιορίζονται σε μικρό ποσοστό, ενώ τα αλκίνια (ακετυλένια), εμφανίζονται σε ίχνη στα καύσιμα

αυτοκινήτων και μόνο σε φτωχά ραφιναρισμένες βενζίνες. Τα αρένια ή αρωματικές ενώσεις (βενζόλιο, τολουόλιο) αποτελούν πάνω από 40% της βενζίνης, αν και τα τελευταία χρόνια σταδιακά μειώνονται κάτω από 20%. Αν και τοξικά, έχουν υψηλούς αριθμούς οκτανίων και χρησιμοποιούνται σε μεγάλο ποσοστό σε αμόλυβδες βενζίνες υψηλής ποιότητας αντικαθιστώντας τα βελτιωτικά μολύβδου. Τέλος, τα πολυκυκλικά αρωματικά (Polynuclear Aromatics, PNAs), όπως η ναφθαλίνη, υπάρχουν σε πολύ μικρές συγκεντρώσεις στην βενζίνη, ενώ η χρήση των μεγαλύτερων μελών της σειράς είναι απαγορευτική λόγω της υψηλής τοξικότητάς τους. Στα Σχ. 1-5 και Σχ. 1-6 παρουσιάζονται οι αλυσίδες των κυριότερων εκπροσώπων των παραπάνω ομάδων υδρογονανθράκων.



Σχήμα 1-5 Παραφίνες α) κανονικό επτάνιο, β) ισοοκτάνιο (2, 2, 4 τριμεθυλοπεντάνιο) και γ) κυκλοεξάνιο



Σχήμα 1-6 Ολεφίνες και αρωματικές ενώσεις. α) 2-μεθυλοβουτένιο-2, β) βενζόλιο και γ) τολουόλιο

Μία από τις σημαντικότερες κατηγορίες προσθέτων στα καύσιμα είναι αυτή των ουσιών με αντικροτική δράση. Από τις αρχές του 20ου αιώνα δοκιμάστηκαν διάφορες ουσίες, όπως το ιώδιο και η ανιλίνη, με πολύ καλά αποτελέσματα, δεν είχαν όμως πρακτική εφαρμογή και χρησιμοποιήθηκαν μόνο ως πρότυπα αναφοράς. Το 1922 διαπιστώθηκε η ισχυρή αντικροτική δράση των τετρααιθυλιούχου TEL (tetra ethyl lead) και τετραμεθυλιούχου μολύβδου TML (tetra methyl lead), η προσθήκη των οποίων στην βενζίνη έγινε γρήγορα η πιο οικονομική μέθοδος αύξησης του αριθμού οκτανίων. Λόγω των ανεπιθύμητων επικαθίσεων μολύβδου στα τμήματα της μηχανής, η χρήση αλκυλικών ενώσεων μολύβδου είναι δυνατή μόνο με παράλληλη προσθήκη οργανικών αλογονούχων ενώσεων, όπως το διβρωμίδιο αιθυλενίου και το διχλωρίδιο αιθυλενίου, τα άτομα αλογόνων των οποίων αντιδρούν με τον μολύβδο σχηματίζοντας πτητικά άλατα αλογονούχου μολύβδου. Η τοξικότητα όμως των καυσαερίων καθώς και η αδυναμία να κατασκευαστούν ανθεκτικοί στον μολύβδο καταλύτες εξάτμισης, έχουν οδηγήσει τα τελευταία χρόνια σε σταδιακή κατάργηση των προσθέτων μολύβδου, ενώ ακόμα και στις χώρες που διατηρούνται οι μολυβδωμένες βενζίνες έχουν μειωθεί τα επιτρεπόμενα επίπεδα μολύβδου. Άλλες οργανομεταλλικές ενώσεις, που κατά καιρούς έχουν προταθεί για αύξηση του αριθμού

οκτανίων, χωρίς όμως να ξεπεραστούν προβλήματα, όπως η τοξικότητα, το κόστος, τα αυξημένα ποσοστά επικαθίσεων, κ.α., είναι το μέθυλο-κυκλοπενταδιένυλο-τρικαρβονυλιούχο μαγγάνιο MMT, το καρβονύλιο του νικελίου, το πεντακαρβονύλιο του σιδήρου ($\text{Fe}(\text{CO})_5$), το φερροκένιο ή δικυκλοπενταδιενυλιούχος σίδηρος, και το οξείδιο του σιδήρου (Fe_3O_4).

Για την διατήρηση στις αμόλυβδες βενζίνες των υψηλών αριθμών οκτανίων που επιτεύχθηκαν με την χρήση μολύβδου, σήμερα, εκτός από την αύξηση του ποσοστού των αρωματικών, χρησιμοποιούνται οξυγονούχα πρόσθετα. Πρόκειται για οργανικές ενώσεις που περιέχουν στο μόριό τους άτομα οξυγόνου, κυρίως αλκοόλες ($\text{C}_x\text{-O-H}$) ή αιθέρες ($\text{C}_x\text{-O-C}_y$) μέχρι έξι ατόμων άνθρακα (Σχ. 1-7). Η παρουσία οξυγόνου δεν αυξάνει το ενεργειακό δυναμικό του καυσίμου, αλλά η δομή τους παρέχει μια σημαντική τιμή του δείκτη αντικροτικής συμπεριφοράς, έτσι ώστε να καθίστανται ικανά υποκατάστατα των αρωματικών. Παράγονται από ορυκτά καύσιμα (μεθανόλη, μεθυλικός τριτοταγής βουτυλικός αιθέρας MTBE, τεταρτοταγής αμυλομεθυλαιθέρας TAME) και από βιομάζα (βιοαιθανόλη, αιθυλικός τριτοταγής βουτυλικός αιθέρας (ETBE)). Οι φυσικές ιδιότητες των ενώσεων αυτών είναι σημαντικά διαφορετικές από εκείνες των υδρογονανθράκων της βενζίνης και επομένως τα όρια τους διατηρούνται σε χαμηλά επίπεδα, ενώ το κλάσμα των υδρογονανθράκων πρέπει να τροποποιηθεί κατάλληλα ώστε να αξιοποιηθούν οι αντικροτικές και πτητικές ιδιότητες των οξυγονούχων προσθέτων. Σήμερα χρησιμοποιούνται σε καύσιμα με σημαντικά τροποποιημένο κλάσμα υδρογονανθράκων, που αναφέρονται ως **ανασχηματισμένες βενζίνες** (reformulated gasolines, RFGs), ώστε παράλληλα με την αύξηση του αριθμού οκτανίων να είναι δυνατή η μείωση των τοξικών ακόρεστων υδρογονανθράκων και του βενζολίου.



Σχήμα 1-7 Οξυγονούχες ενώσεις α) αιθανόλη και β) μεθυλικός τριτοταγής βουτυλικός αιθέρας

Οι βενζίνες μπορούν να περιέχουν και άλλα βελτιωτικά πρόσθετα, όπως μικρές ποσότητες άλλων ενισχυτικών του αριθμού οκτανίων, αντιοξειδωτικά και μεταλλικούς απενεργοποιητές για αύξηση της σταθερότητας της βενζίνης, τροποποιητές ανεπιθύμητων επικαθίσεων στο σύστημα ανάφλεξης (μπουζί) και προκαλούν προανάφλεξη, επιφανειακά ενεργά συστατικά που λειτουργούν ως αντιψυκτικά, αποτρέπουν τον σχηματισμό επικαθίσεων, βελτιώνουν την εξατμισμό και μειώνουν τις εκπομπές NO_x , αντιδιαβρωτικά, και τέλος χρωστικές ουσίες και ιχνηθέτες για εύκολο διαχωρισμό των διαφόρων ποιοτήτων και για λόγους ασφάλειας.

Πίνακας 1-4 Υδρογονάνθρακες της βενζίνης και οξυγονούχα πρόσθετα

	Οργανική ένωση	RON	MON	Σ.Β. (°C)	d (g/ml σε 15°C)	AIT (°C)
παραφίνες	βουτάνιο	113	114	-0.5	αέριο	370
	πεντάνιο	62	66	35	0.626	260
	εξάνιο	19	22	69	0.659	225
	επτάνιο	0	0	98	0.684	225
	οκτάνιο	-18	-16	126	0.703	220
ισοπαραφίνες	2 μεθυλοπροπάνιο	122	120	-12	αέριο	460
	2 μεθυλοβουτάνιο	100	104	28	0.620	420
	2 μεθυλοπεντάνιο	82	78	62	0.653	306
	3 μεθυλοπεντάνιο	86	80	64	0.664	
	2 μεθυλοεξάνιο	40	42	90	0.679	
	3 μεθυλοεξάνιο	56	57	91	0.687	
	2, 2 διμεθυλοπεντάνιο	89	93	79	0.674	
	2,2,3 τριμεθυλοβουτάνιο	112	112	81	0.690	420
	2,2,4 τριμεθυλοπεντάνιο	100	100	98	0.692	415
κυκλοαλκάνια	κυκλοπεντάνιο	141	141	50	0.751	380
	μεθυλοκυκλοπεντάνιο	107	99	72	0.749	
	κυκλοεξάνιο	110	97	81	0.779	245
	μεθυλοκυκλοεξάνιο	104	84	101	0.770	250
αρωματικές ενώσεις	βενζόλιο	98	91	80	0.874	560
	τολουόλιο	124	112	111	0.867	480
	αιθυλοβενζόλιο	124	107	136	0.867	430
	m-ξυλόλιο	162	124	138	0.868	463
	p-ξυλόλιο	155	126	138	0.866	530
	ο-ξυλόλιο	126	102	144	0.870	530
	3-αιθυλοτολουόλιο	162	138	158	0.865	
	1,3,5-τριμεθυλοβενζόλιο	170	136	163	0.864	
	1,2,4-τριμεθυλοβενζόλιο	148	124	168	0.889	
ολεφίνες	2-πεντένιο	154	138	37	0.649	
	2-μεθυλοβουτένιο-2	176	140	36	0.662	
	2-μεθυλοπεντένιο	159	148	67	0.690	
	κυκλοπεντένιο	171	126	44	0.774	
οξυγονούχα	μεθανόλη	133	105	65	0.796	385
	αιθανόλη	129	102	78	0.794	365
	ισοπροπυλική αλκοόλη	118	98	82	0.790	399
	MTBE	116	103	55	0.745	
	ETBE	118	102	72	0.745	
	TAME	111	98	86	0.776	

(http://www.faqs.org/faqs/autos/gasoline-faq/part2/)

Οι βενζίνες αποτελούν λοιπόν προϊόντα μίξης κλασμάτων-προϊόντων εξειδικευμένων κατεργασιών διύλισης, ώστε να πληρούν τις εκάστοτε ποιοτικές προδιαγραφές και τους ισχύοντες περιβαλλοντικούς κανονισμούς. Διαχωρίζονται σε ποιότητες ανάλογα με τον αριθμό οκτανίων, που όμως δεν αντιστοιχούν απαραίτητα στην ίδια χημική σύσταση και συνεπώς το ενεργειακό περιεχόμενο, την ταχύτητα φλόγας, την θερμοκρασία ανάφλεξης και τα υπόλοιπα ποιοτικά χαρακτηριστικά του καυσίμου. Μια τυπική σύσταση βενζίνης περιλαμβάνει 15% παραφίνες, 30% ισοπαραφίνες, 12% κυκλοαλκάνια, 35% αρωματικές ενώσεις και 8% ολεφίνες, ενώ τα νεώτερα οξυγονωμένα καύσιμα περιέχουν μέχρι και 12-15% οξυγονούχες ενώσεις που αντικαθιστούν ποσοστό των αρωματικών ενώσεων και ολεφινών, μειώνοντας το εύρος του σημείου βρασμού. Σήμερα στις υψηλής ποιότητας αμόλυβδες βενζίνες το ποσοστό των αρωματικών τείνει να είναι κάτω του 25% για περιβαλλοντολογικούς λόγους. Οι κυριότερες ενώσεις των παραπάνω ομάδων με τους αριθμούς οκτανίων (RON, MON), το σημείο βρασμού, την πυκνότητα και την ελάχιστη θερμοκρασία αυτανάφλεξης (minimum autoignition temperature, AIT) που εμφανίζουν παρουσιάζονται στον Πιν. 1-4.

Προδιαγραφές Βενζίνης

Οι βενζίνες παράγονται ώστε να ικανοποιούν καθορισμένες προδιαγραφές και πρότυπα ποιότητας. Οι προδιαγραφές αυτές αφορούν φυσικά και χημικά χαρακτηριστικά της βενζίνης που επηρεάζουν την ορθή και ασφαλή λειτουργία των μηχανών εσωτερικής καύσης και εξασφαλίζουν συμμόρφωση με τους περιβαλλοντικούς κανονισμούς. Κατά τον ποιοτικό έλεγχο εξετάζονται η **πτητικότητα** (volatility), τα **χαρακτηριστικά καύσης** (combustion characteristics), η **οξειδωτική σταθερότητα** (stability), η **ανθεκτικότητα σε διάβρωση** (corrosiveness) και το περιεχόμενο σε θείο, μόλυβδο, φώσφορο και μαγγάνιο και υδρογονάνθρακες.

Πτητικότητα

Η πτητικότητα του καυσίμου έχει επίδραση στις εκπομπές καυσαερίων, καθώς και στην ομαλή κίνηση των οχημάτων υπό ορισμένες συνθήκες λειτουργίας. Είναι ιδιότητα που εξαρτάται τόσο από την ατμοσφαιρική πίεση και θερμοκρασία, όσο και από τις συνθήκες λειτουργίας του κινητήρα, με αποτέλεσμα οι προδιαγραφές να περιλαμβάνουν διαφορετικές κατηγορίες ορίων ανάλογα με την τοποθεσία και την εποχή χρήσης του καυσίμου. Πτητικές βενζίνες εμφανίζουν μεγάλες απώλειες και ανωμαλίες στην λειτουργία της μηχανής εξαιτίας της δημιουργίας φυσαλίδων αερίων στην γραμμή ροής του καυσίμου (φράγμα ατμών). Η παρουσία, εξάλλου, κλασμάτων με υψηλό σημείο βρασμού αυξάνει σημαντικά τα επίπεδα εκπομπής ανεπιθύμητων υδρογονανθράκων (βενζολίου, βουταδιενίου) και αλδεϋδών (φορμαλδεϋδης, ακεταλδεϋδης). Βενζίνες που εξατμίζονται δύσκολα δυσχεραίνουν την εκκίνηση της μηχανής και μπορεί να

οδηγήσουν σε πάγωμα του εξαερωτήρα στις χαμηλές θερμοκρασίες, χαμηλή επιτάχυνση, κτύπο στον κινητήρα λόγω μη ομοιόμορφης κατανομής καυσίμου στους κυλίνδρους, διακοπή ροής του καυσίμου λόγω ύπαρξης ατμών σε περιόδους υψηλών θερμοκρασιών και επιμόλυνση του λιπαντικού στον στροφαλοθάλαμο.

Η πτητικότητα του καυσίμου καθορίζεται από τις συνθήκες απόσταξης και την τάση ατμών. Οι προδιαγραφές ορίζουν τις θερμοκρασίες εξάτμισης για διάφορα ποσοστά του καυσίμου. Τα όρια απόσταξης αναφέρονται στην μέγιστη θερμοκρασία εξάτμισης του 10% (50-70°C), του 50% (110-121°C) και του 90% (185-190°C) του καυσίμου, το τελικό σημείο βρασμού (225°C) της βενζίνης, την ελάχιστη θερμοκρασία εξάτμισης του 50% του καυσίμου (77°C) και το μέγιστο ποσοστό υπολείμματος (2%) μετά από την απόσταξη. Καθορίζονται επίσης τα όρια τάσης ατμών (54-103 kPa) για κάθε κατηγορία και για τις κατηγορίες περιόδων υψηλών θερμοκρασιών το μέγιστο (1050-1250) του δείκτη ατμόφραξης (Vapor Lock Index, VLI), που συνυπολογίζει την τάση ατμών και το ποσοστό του ελαφρού κλάσματος της βενζίνης για τον προσδιορισμό του λόγου υγρού-ατμών.

Χαρακτηριστικά καύσης

Δεδομένου ότι οι βενζίνες περιέχουν κυρίως υδρογονάνθρακες με παρόμοιες φυσικοχημικές ιδιότητες, η μόνη σημαντική μεταβλητή μεταξύ των διάφορων ποιοτήτων είναι ο αριθμός οκτανίων του καυσίμου. Η θερμοκρασία καύσης δεν εμφανίζει σημαντικές διακυμάνσεις (γύρω στους 2000°C για την ισοβαρή αδιαβατική καύση), ενώ εξαρτάται κυρίως από το φορτίο (load) και τον σχεδιασμό των μηχανών. Τα επίπεδα οξυγόνου στα καύσιμα ρυθμίζονται σύμφωνα με τα πρότυπα ατμοσφαιρικής ποιότητας, καθώς η προσθήκη οξυγονούχων τροποποιεί τις διαδικασίες αντιδράσεων προ-έναυση αλλά παράλληλα μειώνει το ενεργειακό περιεχόμενο του καυσίμου. Ο αριθμός οκτανίων χαρακτηρίζει το καύσιμο ως προς την τάση εμφάνισης κτύπου στον κινητήρα, καθορίζοντας έτσι την επιτρεπόμενη αναλογία συμπίεσης και επομένως την ενεργειακή απόδοση του κινητήρα. Ο **δείκτης αντικροτικής συμπεριφοράς** (antiknock index) του καυσίμου $(RON+MON)/2$ δεν εντάσσεται σε αυστηρά καθορισμένα όρια αλλά εξαρτάται από τις απαιτήσεις των Μ.Ε.Κ., ενώ τυπικά θεωρούνται καύσιμα με δείκτη αντικροτικής συμπεριφοράς 87, 89, 91 (για αμόλυβδες βενζίνες), και 88 (για μολυβδωμένες βενζίνες).

Οξειδωτική Σταθερότητα

Με την πάροδο του χρόνου τα ακόρεστα συστατικά της βενζίνης οξειδώνονται με την παρουσία αέρα και πολυμερίζονται σε κομμωδείς ουσίες, που ονομάζονται **κόλλες** ή **κομμίες** (gums) και οι οποίες κατακρημνίζονται και επικάθονται στις επιφάνειες του εξαερωτήρα, της κυλινδροκεφαλής και των βαλβίδων, επηρεάζοντας την ορθή λειτουργία των Μ.Ε.Κ.. Οι

αντιδράσεις οξείδωσης καταλύονται από μέταλλα, όπως ο χαλκός, και χρησιμοποιούνται αντιοξειδωτικά πρόσθετα και μεταλλικοί απενεργοποιητές για την σταθεροποίηση του καυσίμου. Οι δοκιμές που πραγματοποιούνται είναι η μέτρηση της υπάρχουσας κόλλας (Maximum Solvent Washed Gum-Existent Gum) που προσδιορίζει την ποσότητα κολλών σε 5mg/ 100ml και η σταθερότητα οξείδωσης (Minimum Oxidation Stability) που μετρά το χρόνο κατακρήμνισης κόλλας παρουσία οξυγόνου πίεσης 100 psi σε θερμοκρασία 100°C (με ελάχιστο όριο 360 min). Τέλος, μία βενζίνη παραγωγής δεν πρέπει να περιέχει νερό ή ίζημα –πρέπει να είναι καθαρή και διαυγής σε θερμοκρασία 21°C (70°F).

Ανθεκτικότητα σε διάβρωση

Η ύπαρξη θείου ή θειούχων ενώσεων (μερκαπτάνες, σουλφίδια και δισουλφίδια) στις βενζίνες έχει σαν αποτέλεσμα την παραγωγή κατά την καύση τους καυσαερίων (SO₂) με διαβρωτική επίδραση στα μηχανικά μέρη και το σύστημα εξάτμισης και επιβάρυνση του περιβάλλοντος. Το θείο επίσης μειώνει την απόκριση των βενζινών στην αύξηση οκτανίων με προσθήκη τετρααιθυλίουχου μολύβδου (alkyl lead octane response), και έχει καταστροφική επίδραση στους καταλύτες εξάτμισης. Η ανθεκτικότητα σε διάβρωση των βενζινών καθορίζεται με την δοκιμή διάβρωσης ελάσματος χαλκού (copper strip corrosion) που μετρά το ενεργό θείο και τον προσδιορισμό περιεκτικότητας σε θείο (maximum sulfur content), το όρια του οποίου είναι 0.15% κ.β. για μη αμόλυβδες βενζίνες και 0.10% κ.β. για αμόλυβδες με τυπικές τιμές μεταξύ 0.03% κ.β -0.04% κ.β..

Άλλοι περιβαλλοντικοί περιορισμοί που έχουν θεσπιστεί για την παραγωγή καυσαερίων θέτουν όρια πυκνότητας (720-775kg/m³ στους 15°C), μέγιστου περιεχόμενου μολύβδου 1.1g Pb/l για τις μολυβδωμένες και 0.005g Pb/l για τις αμόλυβδες βενζίνες, ενώ το περιεχόμενο σε φώσφορο λόγω της αρνητικής επίδρασής του στους καταλύτες εξάτμισης έχει περιοριστεί για όλες τις βενζίνες σε 0.0013g P/l. Για την μείωση των τοξικών εκπομπών μέγιστο όριο για το βενζόλιο έχει τεθεί το 1%, που περιορίζει το συνολικό ποσοστό των αρωματικών γύρω στο 27%, ενώ υπάρχουν επιπλέον περιορισμοί για αέριες εκπομπές υδρογονανθράκων, CO και οξειδίων του αζώτου NO_x.

Διεργασίες παραγωγής βενζίνης

Τα διυλιστήρια σήμερα για να καλύψουν τις αυξημένες απαιτήσεις για βενζίνες υψηλού αριθμού οκτανίων, χρησιμοποιούν, εκτός της αρχικής ατμοσφαιρικής απόσταξης του αργού πετρελαίου, **κατεργασίες αναμόρφωσης** (refining processes) των κλασμάτων απόσταξης του αργού πετρελαίου. Η ανάπτυξη τέτοιων μεθόδων επεξεργασίας στοχεύει στην βελτίωση των ιδιοτήτων και την αύξηση της παραγωγής των κλασμάτων που θα αποτελέσουν την πρώτη ύλη για την

σύνθεση του καυσίμου, καθώς και την οικονομική διαχείριση της συνολικής παραγωγής του διυλιστηρίου που περιλαμβάνει, εκτός της βενζίνης, αέρια καύσιμα, πετρέλαιο κίνησης και θέρμανσης, λιπαντικά και άλλα υποπροϊόντα. Οι διεργασίες που πραγματοποιούνται στις συνήθεις διυλιστηριακές εγκαταστάσεις κατατάσσονται σε τρεις κατηγορίες: πρωτογενείς, δευτερογενείς και δευτερεύουσες.

Πίνακας 1-5 Πρωτογενείς διεργασίες διυλιστηρίου

Φυσικές Διεργασίες Διαχωρισμού	Διεργασίες Χημικής Μετατροπής
απόσταξη • ατμοσφαιρική • υπό κενό κρυστάλλωση • αποπαραφίνωση • απελαίωση παραφίνης εκχύλιση απορρόφηση απογύμνωση προσρόφηση θερμική διάχυση	πυρολυτικές • θερμική πυρόλυση • καταλυτική πυρόλυση • υδρογονοπυρόλυση επικοδομητικές • πολυμερισμός (διμερισμός) • αλκυλίωση μετατρεπτικές • θερμική αναμόρφωση • καταλυτική αναμόρφωση • ισομερισμός • κυκλοποίηση • αφυδρογόνωση • αποθείωση

Ο όρος πρωτογενείς περιλαμβάνει όλες τις βασικές διεργασίες (Πιν. 1-5), τόσο τις φυσικές, όπως ο διαχωρισμός του αργού πετρελαίου σε κλάσματα (απόσταξη), όσο και τις χημικές κατά τις οποίες παράγονται διάφορα συστατικά από τα κλάσματα του πετρελαίου με αλλαγή της μοριακής δομής των υδρογονανθράκων τους. Στο στάδιο των δευτερογενών διεργασιών πραγματοποιείται εξευγενισμός των προϊόντων των πρωτογενών, με τις ίδιες ουσιαστικά μεθόδους, προσαρμοσμένες έτσι ώστε, το τελικό προϊόν να πληρεί τις απαιτούμενες προδιαγραφές, ενώ οι δευτερεύουσες αναφέρονται σε διεργασίες παρασκευής του έτοιμου τελικού προϊόντος, όπως η μίξη κλασμάτων βενζινών και η προσθήκη των απαραίτητων προσθέτων. Η βελτίωση των ιδιοτήτων των παραγόμενων σήμερα βενζινών οφείλεται κυρίως στην ανάπτυξη των χημικών μεθόδων κατεργασίας και τον εκσυγχρονισμό των εν λόγω μονάδων, οι κυριότερες από τις οποίες παρουσιάζονται παρακάτω. Πριν από τις κύριες αυτές μονάδες παραγωγής των κλασμάτων μίξης της βενζίνης χρησιμοποιούνται μονάδες **αποθείωσης** και **υδρογονοκατεργασίας νάφθας**, όπου πραγματοποιείται προκατεργασία της τροφοδοσίας τους για την απομάκρυνση ανεπιθύμητων προσμίξεων.

Καταλυτική Αναμόρφωση

Η **αναμόρφωση** (reforming) εντάσσεται στις χημικές διεργασίες μετατροπής που στο σύνολό τους στοχεύουν στην αναδιάταξη των μορίων των υδρογονανθράκων των κλασμάτων κυρίως ελαφριάς και βαριάς νάφθας για την παραγωγή προϊόντων-κλασμάτων μίξης με καλύτερη αντικροτική συμπεριφορά. Πρόκειται για μία διεργασία κατά την οποία τα κορεσμένα συστατικά (παραφίνες και ναφθένια) χαμηλού αριθμού οκτανίων της βαριάς νάφθας από την ατμοσφαιρική απόσταξη, μετατρέπονται σε προϊόντα υψηλού Α.Ο. (100-104 RON) που περιέχουν περίπου 60% αρωματικά συστατικά. Διακρίνονται δύο τύποι αναμόρφωσης, η θερμική και η καταλυτική, με την τελευταία να κυριαρχεί λόγω της αυξημένης απόδοσης και καλύτερης ποιότητας του προϊόντος. Ο καταλύτης που χρησιμοποιείται είναι Pt/Al_2O_3 , ενώ για την υψηλή απόδοση απαιτούνται θερμοκρασίες της τάξης $500^{\circ}C$ σε 10-20atm. Η μονάδα καταλυτικής αναμόρφωσης αποτελεί επίσης και την βασική πηγή υδρογόνου για τις υπόλοιπες διυλιστηριακές διεργασίες.

Οι κυριότερες αντιδράσεις που λαμβάνουν χώρα στην μονάδα αναμόρφωσης είναι: αφυδρογόνωση ναφθενίων προς αρωματικά και υδρογόνο

1. ισομερισμός παραφινών και ολεφινών προς ισο-αλκάνια και ναφθενίων προς αρωματικά
2. αφυδρογονωτική κυκλοποίηση παραφινών προς αρωματικά και υδρογόνο
3. υδρογονοπυρόλυση (αντιδράσεις διάσπασης με προσθήκη H_2) παραφινών προς παραφίνες με μικρότερη ανθρακική αλυσίδα.
4. αποθείωση

Ισομερισμός

Ο **ισομερισμός** (isomerization) αποτελεί καταλυτική διεργασία χημικής μετατροπής των παραφινών ευθείας αλυσίδας σε ισο-παραφίνες με σκοπό την αύξηση του Α.Ο. (80-88 RON) του κλάσματος της ελαφριάς νάφθας που δεν αναμορφώνεται. Ο καταλύτης που χρησιμοποιείται είναι $AlCl_3/Al_2O_3$, ενώ η θερμοκρασία κυμαίνεται από $180^{\circ}C$ (βουτάνιο) και $400^{\circ}C$ (πεντάνιο, εξάνιο) σε πίεση περίπου 20atm.

Πυρόλυση

Για την αξιοποίηση των βαρέων κλασμάτων της απόσταξης του αργού πετρελαίου και της μετέπειτα απόσταξης υπό κενό του υπολείμματος της πρώτης, αναπτύχθηκε η διεργασία της πυρόλυσής τους προς ελαφρύτερα υψηλότερης αξίας προϊόντα (βενζίνη υψηλού Α.Ο., diesel και υγραέρια). Αποτελεί μία από τις βασικότερες μεθόδους παραγωγής κλασμάτων μίξης βενζινών και περιλαμβάνει ένα μεγάλο εύρος διεργασιών που κατηγοριοποιούνται σε:

- **θερμική πυρόλυση** (thermal cracking), κατά την οποία πραγματοποιείται θερμική διάσπαση ($425-650^{\circ}C$ σε 15atm) μορίων ανώτερων υδρογονανθράκων, απουσία αέρος, προς

μικρότερα μόρια αλκανίων και αλκενίων, που οδηγεί σε παραγωγή μεγάλων ποσοτήτων αερίων και βενζίνης χαμηλού αριθμού οκτανίων. Κατά την θερμική πυρόλυση παράγεται χαμηλής ποιότητας προϊόν λόγω της υψηλής περιεκτικότητάς του σε ασταθείς και εύκολα πολυμεριζόμενες ολεφίνες, γι' αυτό και σταδιακά αντικαθίσταται από μονάδες καταλυτικής πυρόλυσης. Οι κυριότερες αντιδράσεις είναι:

1. διάσπαση παραφινών προς μικρότερα μόρια παραφινών και ολεφινών
 2. αφυδρογόνωση παραφινών προς ολεφίνες και ναφθενίων προς αρωματικά
 3. ισομερισμός ολεφινών
 4. πολυμερισμός ολεφινών
 5. περαιτέρω αφυδρογόνωση ολεφινών προς διολεφίνες
- **καταλυτική πυρόλυση** (catalytic cracking), με αντιδράσεις καταλυτικής διάσπασης βαρέων κλασμάτων πετρελαίου που περιέχουν C15⁺ υδρογονάνθρακες με υψηλό σημείο βρασμού, π.χ. αερίλαιο, προς κλάσματα βενζίνης που περιέχουν 30% αρωματικά και 20-30% ολεφίνες σε αντιδραστήρες καταλυτικής κλίνης. Η διεργασία αυτή, που υλοποιείται σε ρευστοποιημένη (fluidized, FCC) ή σταθερή κλίνη, χρησιμοποιεί ως καταλύτη μίγμα αργιλοπυριτικών ενώσεων Al₂O₃, SiO₂ σε όξινο περιβάλλον. Η καταλυτική πυρόλυση δεν διαφέρει ουσιαστικά από την θερμική, επιτελείται όμως σε μεγαλύτερες θερμοκρασίες (650-750°C σε 2-3atm) και περιλαμβάνει επιπλέον αντιδράσεις υδρογόνωσης των παραγόμενων ολεφινών και αφυδρογονωτικής κυκλίσωσης παραφινών προς σταθερότερα και υψηλότερης αντικροτικής αξίας προϊόντα μίξης βενζινών (92-94RON).
 - **υδροπυρόλυση** (hydrocracking) κατά την οποία πραγματοποιείται διάσπαση βαρέων κλασμάτων πετρελαίου παρουσία υδρογόνου προς κορεσμένους και αρωματικούς υδρογονάνθρακες.

Αλκυλίωση

Ο όρος **αλκυλίωση** (alkylation) αναφέρεται στην αντίδραση αντικατάστασης ενός υδρογόνου μιας οργανικής ένωσης από ένα αλκύλιο και χαρακτηρίζει την διεργασία κατά την οποία ελαφρές αέριες ολεφίνες αντιδρούν με ισοβουτάνιο σε έντονα όξινες καταλυτικές συνθήκες (H₂SO₄, HF) για την παραγωγή προϊόντος που περιέχει μεγαλύτερα μόρια ισοπαραφινών. Η μονάδα χρησιμοποιείται για την αξιοποίηση ελαφρών ακόρεστων υδρογονανθράκων από την μονάδα FCC για την παραγωγή συστατικού με υψηλές αντικροτικές ιδιότητες (93-95 RON).

Διμερισμός

Ο **διμερισμός** (dimate) αποτελεί διεργασία μέσω της οποίας αυξάνεται η παραγωγή βενζίνης από ελαφρές αέριες ολεφίνες. Εφαρμόζεται κυρίως σε μονάδες διμερισμού των προπυλενίων της καταλυτικής πυρόλυσης για την παραγωγή μίγματος εξανίων με Α.Ο. 95-97 RON. Το

προπυλένιο αντιδρά εύκολα προς προϊόντα διμερισμού ως πολυμερισμού παρουσία καταλυτών, που βασίζονται σε οργανομεταλλικές ενώσεις του Al, σε θερμοκρασίες 27°-55°C. Μαζί με την αλκυλίωση αποτελούν εποικοδομητικές διεργασίες, κατά τις οποίες το μεγαλύτερο μέρος των ελαφρών υδρογονανθράκων C₃ και C₄, που προέρχονται από την μονάδα FCC, μετατρέπεται σε συστατικά βενζινών.

1.3 Καύση βενζίνης

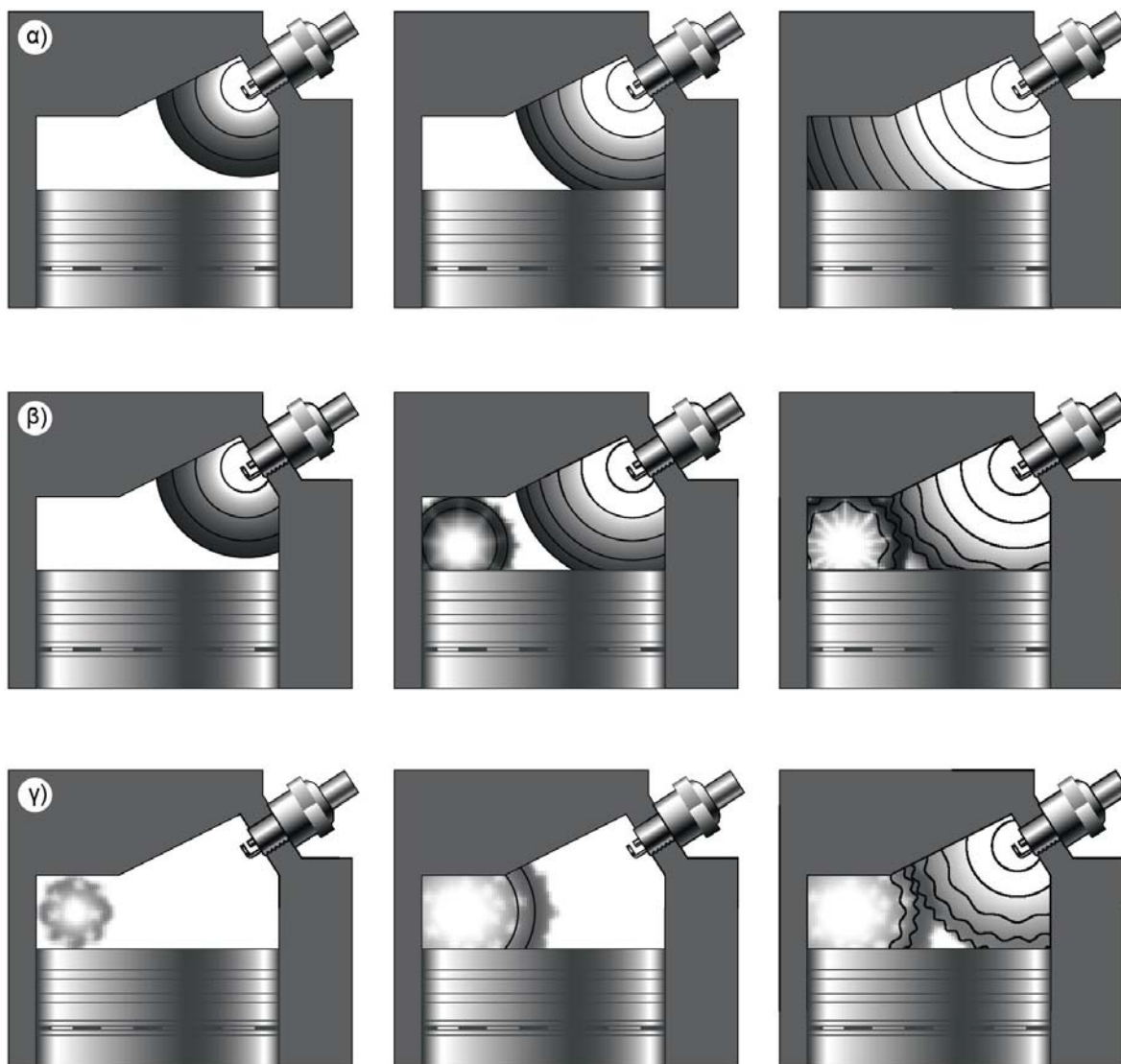
Η καύση στον κινητήρα είναι ένα πολύπλοκο φαινόμενο, που η ομαλή εξέλιξή του επηρεάζει σημαντικά την απόδοση του κινητήρα, την κατανάλωση καυσίμου, το χρόνο ζωής της μηχανής και τις δυνατότητες του οχήματος κατά την οδήγηση. Οι μεταβολές της πίεσης κατά την διάρκεια της καύσης εξαρτώνται από πολλούς παράγοντες, που αφορούν το σχεδιασμό της μηχανής (διάταξη κινητήρα, μορφή θαλάμου καύσης, θέση και αρχή λειτουργίας σπινθηριστή, χρονισμός έναυσης, αναλογία συμπίεσης, αναλογία μίγματος αέρα-καυσίμου), τις συνθήκες λειτουργίας (θερμοκρασία, φορτίο, ταχύτητα περιστροφής) και προφανώς την ποιότητα του καυσίμου (πτητικότητα, αριθμός οκτανίων, χημική δομή των συστατικών του).

Ανώμαλες Μορφές Καύσης

Κατά την ανάφλεξη του σπινθηριστή, σχηματίζεται ένας πυρήνας φλόγας που διογκώνεται προοδευτικά διαδίδοντας την φλόγα δια του ομογενούς εξαερωμένου μίγματος καυσίμου-αέρα. Για να ληφθεί η μέγιστη ενέργεια από τη βενζίνη, το συμπιεσμένο μίγμα πρέπει να καεί ομοιόμορφα μέσα στον θάλαμο καύσης, έως ότου καταναλωθούν όλα τα καύσιμα. Σε κανονική λειτουργία του κινητήρα, η καύση είναι **ομαλή** με στρωτό μέτωπο φλόγας που προωθείται ομαλά κατά μήκος του θαλάμου καύσης (Σχ. 1-8α)), προκαλώντας γρήγορη διαστολή του καυσίμου που εφαρμόζει ομαλή ώθηση στο έμβολο.

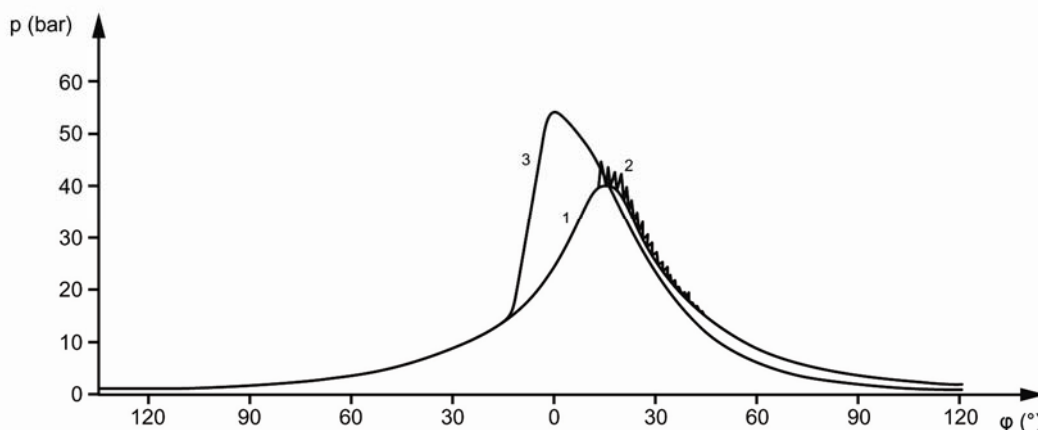
Με την έναρξη του σταδίου της καύσης, παράλληλα με την διάδοση του μετώπου φλόγας, πραγματοποιείται στον θάλαμο καύσης μια σειρά αντιδράσεων προέναυσης στα άκαυτα “υπόλοιπα αερίων” εγγύς του μετώπου της φλόγας. Πρόκειται για αλυσίδα αντιδράσεων που πυροδοτείται και κλιμακώνεται με την ραγδαία αύξηση πίεσης και θερμοκρασίας, λόγω τόσο της μηχανικής εισπίεσης του εμβόλου, όσο και του ωστικού κύματος καύσης. Κάποια από τα προϊόντα είναι ασταθή στις συνθήκες αυτές και τείνουν να αυτοαναφλέγονται δημιουργώντας τοπικά μεμονωμένους πυρήνες φλογών (Σχ. 1-8β)) εντός της περιοχής των άκαυστων αερίων. Η εκρηκτική αυτή καύση, που συμβαίνει αυθόρμητα και σποραδικά, συμπαρασύρει σε αυτανάφλεξη όλο και μεγαλύτερο μέρος των συστατικών του άκαυστου υπολοίπου, σχηματίζοντας ένα δεύτερο τυρβώδες μέτωπο κύματος που υπερτίθεται του κυρίου κύματος,

παράγοντας ένα πριονωτό διάγραμμα πίεσης (Σχ. 1-9) που αντικατοπτρίζει τις απότομες ταλαντώσεις του φορτίου στο θάλαμο καύσης· ταλαντώσεις που συνοδεύονται από χαρακτηριστικό “κτύπο” στην μηχανή.



Σχήμα 1-8 Μέτωπο φλόγας κανονικής καύσης, προανάφλεξης και αυτανάφλεξης

Εκτός από την **αυτανάφλεξη** (autoignition), άλλη ανώμαλη μορφή καύσης αποτελεί η **προανάφλεξη** (preignition), κατά την οποία συμβαίνει πρόωρη ανάφλεξη (πριν την έναυση σπινθήρα) του μίγματος αέρα-καυσίμου σε περιοχές του θαλάμου καύσης κοντά στα τοιχώματα, λόγω των πυρακτωμένων επικαθίσεων άνθρακα στις επιφάνειες του εμβόλου, των βαλβίδων ή του σπινθηριστή. Η εξέλιξη του φαινομένου (Σχ. 1-8γ)) μοιάζει με αυτή της αυτανάφλεξης και ανάλογα με την έντασή του επηρεάζει αρνητικά την μετατροπή της θερμότητας καύσης σε μηχανική ενέργεια, που απάγεται ταχέια από τα τοιχώματα οδηγώντας σε υπερθέρμανση και αστοχία των εμβόλων.



(Internal Combustion Engines, Ferguson (1986))

Σχήμα 1-9 Δυναμοδεικτικό διάγραμμα καύσης τετράχρονου βενζινοκινητήρα. 1) Ομαλή καύση, 2) κρουστική καύση, λόγω αυτανάφλεξης του καυσίμου και 3) πυρανάφλεξη λόγω πρόωρης ανάφλεξης υπέρθερμου μέρους του μίγματος

Παρόλο που δεν υπάρχει άμεση σύνδεση μεταξύ της κρουστικής συμπεριφοράς (αυτανάφλεξης) ενός καυσίμου και της τάσης για προανάφλεξη, και οι δύο ευνοούνται από και οδηγούν σε συνθήκες υψηλής πίεσης και θερμοκρασίας, έτσι ώστε οι δύο καταστάσεις να ενισχύουν η μία την άλλη. Μη ομαλή καύση του μίγματος καυσίμου-αέρα συνήθως αναφέρεται σε συνδυασμό των δύο φαινομένων, αφού η κρουστική καύση προκαλεί υπερθέρμανση των τοιχωμάτων του θαλάμου καύσης που επιφέρει την προανάφλεξη, ενώ αντίστροφα η προανάφλεξη έχει σαν αποτέλεσμα την πρόωρη θερμοκρασιακή άνοδο εξαναγκάζοντας τα ασταθή συστατικά της βενζίνης να παραμένουν για περισσότερο χρόνο σε θερμοκρασίες κοντά στο όριο αυτανάφλεξής τους.

Παράγοντες που επηρεάζουν τη αυτανάφλεξη

Ο κυριότερος παράγοντας επίδρασης στην κρουστική συμπεριφορά της βενζίνης είναι η χημική δομή των υδρογονανθράκων και η απόκρισή τους στα πρόσθετα αύξησης του Α.Ο.. Η εσωτερική δομή των συστατικών του καυσίμου, αλλά και η αλληλεπίδραση μεταξύ τους είναι αυτές που καθορίζουν την χημική διεργασία της καύσης και τις εξωτερικές συνθήκες που αποτελούν το όριο αυτανάφλεξής, και επομένως τον αριθμό οκτανίων του εν λόγω καυσίμου. Παρ' όλο που το είδος και η αναλογία μίξης των επιμέρους κλασμάτων στην βενζίνη διαμορφώνουν καθοριστικά τον Α.Ο. της, η επίδραση τους δεν είναι ανάλογη των επιμέρους Α.Ο. τους, δηλαδή της αντικροτικής συμπεριφοράς που εμφανίζουν αυτά μεμονωμένα. Το φαινόμενο αναφέρεται ως **επίδραση μίξης** (blend effect) και η φύση του εξαρτάται σε μεγάλο βαθμό από τον πολύπλοκο μηχανισμό οξείδωσης των υδρογονανθράκων.

Σημαντικό παράγοντα για την ανθεκτικότητα του καυσίμου έναντι της τάσης αυτανάφλεξης αποτελούν επίσης και οι πτητικές του ιδιότητες. Η μείωση του τελικού σημείου βρασμού αυξάνει τον αριθμό οκτανίων. Δεδομένου ότι οι παραφίνες αποτελούν ένα από τα κύρια συστατικά της βενζίνης, και ο αριθμός οκτανίων τους μειώνεται με την αύξηση του μήκους της αλυσίδας ατόμων άνθρακα ή το μέγεθος του δακτυλίου, ενώ αυξάνεται με το ποσοστό διακλάδωσης αλυσίδας, η παρουσία βαρέων υδρογονανθράκων μειώνει το ποσοστό των πτητικών συστατικών με παράλληλη μείωση του Α.Ο.. Παρόμοια είναι και η επίδραση της απώλειας του πτητικού κλάσματος της βενζίνης λόγω παρατεταμένης ή ακατάλληλης αποθήκευσης. Αντίθετα χαρακτηριστικά καύσης του καυσίμου, όπως το ενεργειακό περιεχόμενο και οι ταχύτητα φλογών, δεν επηρεάζουν την αντικροτική συμπεριφορά του καυσίμου. Τα οξυγονούχα πρόσθετα διαμορφώνουν υψηλούς Α.Ο., αν και το ενεργειακό τους περιεχόμενο είναι σημαντικά χαμηλότερο από των υδρογονανθράκων, και η ταχύτητα φλογών δεν αποτελεί καθοριστικό παράγοντα, αφού δεν διαφοροποιείται μεταξύ των διαφόρων συστατικών για να έχει σημαντική επίδραση στον Α.Ο..

Οι απαιτήσεις σε οκτάνια ενός καυσίμου καθορίζονται από τις κατασκευαστικές ιδιαιτερότητες της μηχανής εσωτερικής καύσης. Είναι προφανές από την εξίσωση (1.2) ότι η απόδοση της μηχανής αυξάνεται με την αναλογία συμπίεσης, όμως η αύξησή της ισοδυναμεί με περισσότερο τεταμένες συνθήκες πίεσης και θερμοκρασίας που εντείνουν το φαινόμενο της αυτανάφλεξης. Για την χρήση επομένως Μ.Ε.Κ. μεγαλύτερης απόδοσης, απαραίτητη προϋπόθεση είναι η χρήση ανθεκτικότερων καυσίμων, δηλαδή καυσίμων μεγαλύτερου αριθμού οκτανίων. Στον Πιν. 1-6 παρουσιάζονται οι απαιτούμενοι αριθμοί οκτανίων για διάφορες αναλογίες συμπίεσης.

Πίνακας 1-6 Αναλογίες συμπίεσης και Α.Ο. που απαιτούν

Αναλογία συμπίεσης	Αριθμός οκτανίων
5:1	72
6:1	81
7:1	87
8:1	92
9:1	96
10:1	100
11:1	104
12:1	108

(<http://www.faqs.org/faqs/autos/gasoline-faq/part2/>)

Μία άλλη παράμετρος λειτουργίας, που επηρεάζει την αυτανάφλεξη, είναι η αναλογία αέρα-καυσίμου που χρησιμοποιείται και η οποία σχετίζεται με την στοιχειομετρία της αντίδρασης καύσης. Ένα εμπλουτισμένο σε βενζίνη μίγμα ελαττώνει το φαινόμενο του “κτύπου”, ενώ η

χρήση φτωχότερου μίγματος μειώνει την ταχύτητα του μετώπου φλόγας και συνεπώς μικρότερο ποσοστό θερμότητας μετατρέπεται σε μηχανική ενέργεια, αυξάνοντας έτσι την θερμοκρασία των τοιχωμάτων του κυλίνδρου και ενισχύοντας την αυτανάφλεξη. Είναι ενδεικτικό ότι για αύξηση κατά μία μονάδα του λόγου αέρα-καυσίμου η απαίτηση του αριθμού οκτανίων αυξάνεται κατά δύο μονάδες του δείκτη αντικροτικής συμπεριφοράς.

Η τάση για αυτανάφλεξη εντείνεται με την πρόωρη έναυση του σπινθηριστή, η οποία πυροδοτεί και την έναυση των αντιδράσεων προφλογών, παρέχοντας τους περισσότερο χρόνο για σχηματισμό των αυταναφλεγόμενων ειδών πριν το έμβολο φτάσει στην κατάλληλη θέση για εφαρμογή της μέγιστης δυνατής ώθησης (10°-25° μετά το Α.Ν.Σ.). Στον Πιν. 1-7 δίδονται ενδεικτικές τιμές του απαιτούμενου αριθμού οκτανίων για διαφορετικούς χρονισμούς έναυσης του σπινθηριστή σε σχέση με τον προτεινόμενο για κάθε όχημα (μερικές μοίρες πριν το Α.Ν.Σ.).

Πίνακας 1-7 Χρονισμός έναυσης σπινθήρα

Όχημα	Κανονικός Χρονισμός	Έναυση 5° μετά	Έναυση 5° πριν
A	88	91	93
B	86	90.5	94.5
C	85.5	88	90
D	84	87.5	91
E	82.5	87	90

(<http://www.faqs.org/faqs/autos/gasoline-faq/part2/>)

Άλλοι παράγοντες που επηρεάζουν την αυτανάφλεξη και επομένως τις απαιτήσεις σε αριθμό οκτανίων είναι η θερμοκρασία του μίγματος και το φορτίο, η αύξηση των οποίων οδηγεί σε αύξηση της θερμοκρασίας της μηχανής και της πίεσης των άκαυστων υπολοίπων εντείνοντας την πρόωρη έκρηξη στον κινητήρα. Αρνητική επίδραση έχουν επίσης οι οργανικές και ανόργανες επικαθίσεις στα μηχανικά μέρη, που αυξάνουν την αναλογία συμπίεσης περιορίζοντας τον όγκο στον θάλαμο καύσης, αυξάνουν την θερμική αγωγιμότητα οδηγώντας σε υπερθέρμανση και λειτουργούν καταλυτικά στις αντιδράσεις σχηματισμού αυταναφλεγόμενων ειδών. Αντίθετα η ταχύτητα περιστροφής του κινητήρα μειώνει με την αύξησή της τον διαθέσιμο χρόνο εξέλιξης των αντιδράσεων.

Αριθμός Οκτανίων και Χημική Σύσταση Βενζίνης

Ο αριθμός οκτανίων αντικατοπτρίζει την ικανότητα του καυσίμου να αντισταθεί στην εν δυνάμει τάση του για αυθόρμητη ανάφλεξη υπό ορισμένες συνθήκες δοκιμής. Δεδομένου ότι η διαδικασία της καύσης είναι δυναμική μεταβολή, όχι μόνο του μέρους του καυσίμου που μετέχει κάθε στιγμή στην καύση, αλλά και του άκαυστου υπολοίπου αερίων πέραν του μετώπου

φλόγας, η σύσταση του οποίου μεταβάλλεται συνεχώς κατά την εξέλιξη του φαινομένου ως την πλήρη κατανάλωσή του, ο αριθμός οκτανίων εξαρτάται από την δυνατότητα του καυσίμου να σχηματίζει σε συνθήκες προανάφλεξης προϊόντα με όριο αυτανάφλεξης εκτός των δεδομένων συνθηκών. Η σημασία της χημικής δομής ήταν εμφανής ήδη από τις αρχικές μελέτες πάνω στην κρουστική συμπεριφορά που παρουσιάζουν κατά την καύση τους οι καθαροί υδρογονάνθρακες (Lovell (1948)). Τα κύρια συμπεράσματα ήταν ότι για τα μεν αλκάνια η κρίσιμη αναλογία συμπίεσης –αναλογία συμπίεσης στην οποία παρατηρείται ο “κτύπος”– μειώνεται με την αύξηση του μήκους της ευθείας αλυσίδας ατόμων άνθρακα και αυξάνεται με την αύξηση του ποσοστού των διακλαδώσεων στο μόριο, ενώ για τα αλκένια αυξάνεται όσο πιο κεντρική είναι η θέση του διπλού δεσμού ατόμων άνθρακα. Το φαινόμενο της αυτανάφλεξης εξαρτάται, επομένως, άμεσα από την δομή των υδρογονανθράκων και η κατανόηση της χημείας οξείδωσης των διαφόρων ομάδων τους και των αλληλεπιδράσεων μεταξύ τους, παράλληλα με τον προσδιορισμό των μεταβολών πίεσης και θερμοκρασίας στο περιβάλλον καύσης, είναι ουσιαστική στην αντιμετώπιση των ανεπιθύμητων, αυθόρμητων εκρήξεων που περιορίζουν την απόδοση των Μ.Ε.Κ..

Χημεία οξείδωσης υδρογονανθράκων

Η οξείδωση της αέριας φάσης των υδρογονανθράκων είχε μελετηθεί ήδη από τις αρχές του 1900¹. Το φαινόμενο των *ψυχρών φλογών* (cool flames) (Perkin (1882)) ήταν ήδη γνωστό, ενώ μέχρι την δεκαετία του '30 είχε παρατηρηθεί (Pease (1929)) και η συμπεριφορά του **αρνητικού συντελεστή θερμοκρασίας** (negative temperature coefficient, NTC)². Η υπόθεση της μη αλυσιδωτής διαδοχής των αντιδράσεων, που επικρατούσε μέχρι τότε, δεν εξηγούσε την συμπεριφορά των ψυχρών φλογών, δηλαδή την ανάφλεξη και την μεγάλη ταχύτητα της οξείδωσης υδρογονανθράκων στις χαμηλές θερμοκρασίες, αλλά και την αρνητική επίδραση της θερμοκρασίας στον ρυθμό της (NTC). Σήμερα ο μηχανισμός της οξείδωσης υδρογονανθράκων αναλύεται με την θεωρία των αλυσιδωτών αντιδράσεων που βασίζεται στην υπόθεση ότι η διάσπαση των αντιδρώντων μορίων ελέγχεται από ενεργά κέντρα, όπως ελεύθερες ρίζες και άτομα (Semenov (1929)).

Ένα τέτοιο διακλαδιζόμενο σύστημα χημικών αντιδράσεων –παρόμοιο με την διαδικασία πυρηνικής διάσπασης– εμφανίζει αυτοεπιταχυνόμενη συμπεριφορά, ενώ ο ρυθμός της συνολικής αντίδρασης μπορεί να ελεγχθεί επηρεάζοντας το σύνολο των ενεργών κέντρων. Πρόκειται για ένα φαινόμενο χημικής επαγωγής που κυριαρχείται από **εκφυλισμένη**

¹ <http://dspace.library.drexel.edu/retrieve/1860/384>

² <http://dspace.library.drexel.edu/handle/1860/557>

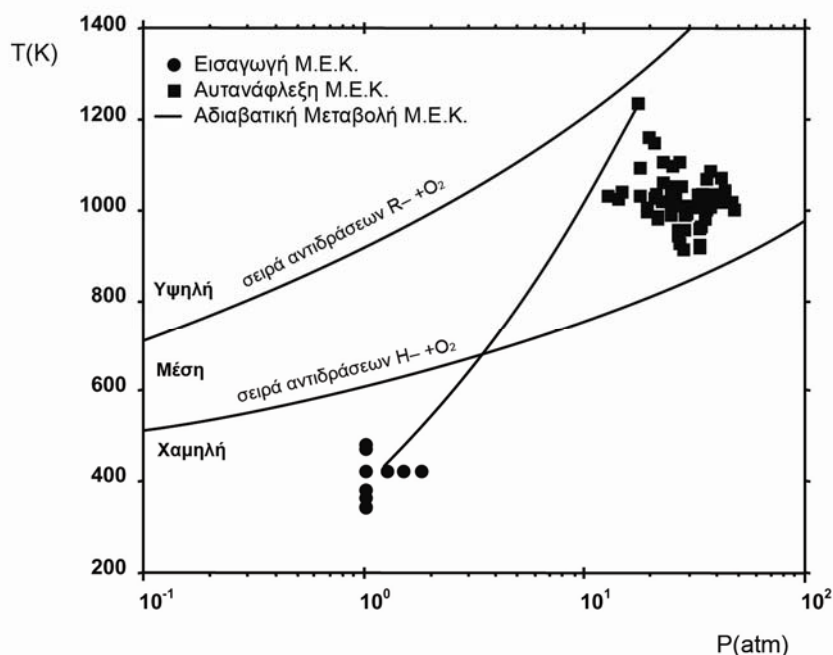
διακλάδωση (degenerate branching) στην εξέλιξη των χημικών αντιδράσεων, κατά την οποία τα σταθερά ενδιάμεσα ακολουθούν επιλεκτικά διαφορετική πορεία αντίδρασης ανάλογα με τις εξωτερικές συνθήκες (πίεση, θερμοκρασία, απελευθέρωση θερμότητας, κ.α.)· είτε μη αλυσιδωτή για να σχηματίσουν νέα σταθερά μόρια, είτε κλάδο αλυσιδωτής για τον σχηματισμό έντονα ενεργών ελεύθερων ριζών. Οι αντιδράσεις οξειδωσης μπορούν να ταξινομηθούν σε διάφορα είδη κατά την πορεία του φαινομένου (Pilling (1997)):

- πρωτογενούς έναρξης (primary initiation) με σχηματισμό ριζών από τα μητρικά μόρια του καυσίμου
- δευτερογενούς έναρξης (secondary initiation) με σχηματισμό ριζών από άλλα “σταθερά” ενδιάμεσα
- διάδοσης αλυσίδας (chain propagation) κατά τις οποίες ο αριθμός των ριζών παραμένει σταθερός
- διακλάδωσης αλυσίδας (chain branching) που αυξάνουν τον αριθμό των ελεύθερων ριζών
- περάτωσης του φαινομένου (termination) με απαλοιφή των ελεύθερων ριζών από την αντίδραση

Μελέτες της κρουστικής συμπεριφοράς στις Μ.Ε.Κ. με έναυση σπινθήρα έδειξαν ότι η θερμοκρασία των άκαυστων υπολοίπων αερίου μίγματος που πρόκεινται του προελαύνοντος μετώπου φλόγας αυξάνεται σταδιακά, λόγω σημαντικής απελευθέρωσης θερμότητας από το σώμα του άκαυστου αερίου και συνεχούς συμπίεσής του από το μέτωπο φλόγας, μέχρι την θερμικά και χημικά τεταμένη εκείνη κατάσταση που πυροδοτεί την αυτανάφλεξη. Η θερμοκρασία του αερίου μίγματος και οι χημικές και φυσικές συνθήκες που την ελέγχουν, συσχετίζουν την αυτανάφλεξη στο θάλαμο συμπίεσης με την οξειδωση των υδρογονανθράκων, αφού η τελευταία μπορεί να χωριστεί σε τρία θερμοκρασιακά στάδια, τα όρια των οποίων εξαρτώνται από την επικρατούσα κατά την αντίδραση πίεση. Σε πίεση $P = 1\text{atm}$, οι αντιδράσεις οξειδωσης χωρίζονται σε:

1. Χαμηλής Θερμοκρασίας ($T < 650\text{K}$)
2. Μέσης Θερμοκρασίας ($650\text{K} < T < 1000\text{K}$)
3. Υψηλής Θερμοκρασίας ($T > 1000\text{K}$)

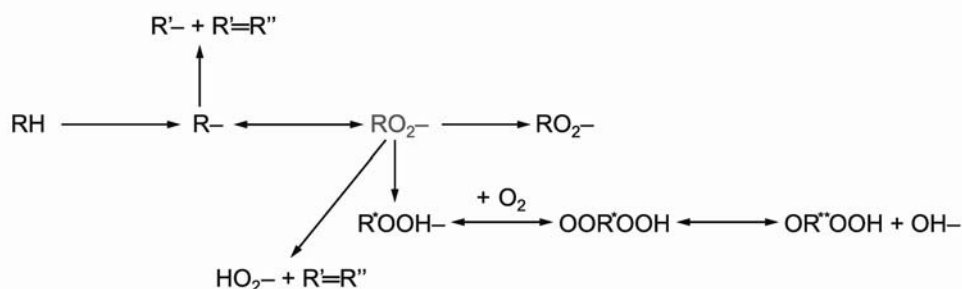
Τα θερμοκρασιακά όρια των παραπάνω σταδίων μετατοπίζονται συνεχώς με την αύξηση της πίεσης καθώς εξελίσσεται η φάση της συμπίεσης (Σχ. 1-10). Αν και η αντίδραση οξειδωσης εξαρτάται και από τις συνθήκες λειτουργίας (αρχική θερμοκρασία μίγματος αέρα-καυσίμου, αναλογία συμπίεσης, χρονισμός έναυσης σπινθήρα), οι πειραματικές μελέτες συγκλίνουν στο ότι τα υπόλοιπα αερίου αυτοαναφλέγονται στην περιοχή των μέσων θερμοκρασιών.



(<http://dspace.library.drexel.edu/retrieve/1860/384>)

Σχήμα 1-10 Μεταβολή πίεσης και θερμοκρασίας του άκαυστου υπολοίπου αερίων (Wang (1999))

Δεδομένου ότι το καύσιμο παραμένει για πεπερασμένο χρονικό διάστημα σε συνθήκες εντός της περιοχής των χαμηλών θερμοκρασιών, η τελευταία αναδεικνύεται εξίσου σημαντική για την πορεία των αντιδράσεων οξείδωσης που θα οδηγήσουν στην ανεπιθύμητη αυτανάφλεξη. Χαρακτηρίζεται από αντιδράσεις σχηματισμού αλκυλοϋπεροξειδικών ριζών $RO_2\cdot$, για τα μικρότερα μόρια υδρογονανθράκων, ή ισομερών τους $R^*OOH\cdot$, για μεγαλύτερα, και σταθερών μορίων οξυγονούχων υδρογονανθράκων. Κατά το στάδιο των ενδιάμεσων θερμοκρασιών κυριαρχούν οι αντιδράσεις του υδροϋπεροξειδίου $HO_2\cdot$ και τα χαρακτηριστικότερα σταθερά προϊόντα είναι αλκένια, οξυγονούχες ενώσεις και μεθάνιο, ενώ με την είσοδο στην περιοχή των υψηλών θερμοκρασιών οι κυριότερες ενεργές ρίζες είναι το υδροξύλιο $OH\cdot$, οξυγόνο $O\cdot$ και υδρογόνο $H\cdot$ και η διαδικασία ολοκληρώνεται με την αποσύνθεση των αλκυλικών ριζών. Στο Σχ. 1-11 παρουσιάζεται ένα σχηματικό διάγραμμα που περιγράφει εν συντομία τους κυριότερους κλάδους των αντιδράσεων που λαμβάνουν χώρα στην περιοχή των χαμηλών και ενδιάμεσων θερμοκρασιών.



Σχήμα 1-11 Διάγραμμα αντιδράσεων χαμηλών και ενδιάμεσων θερμοκρασιών οξείδωσης υδρογονανθράκων.

Αρνητικός Συντελεστής Θερμοκρασίας

Αν και η ανάφλεξη τοποθετείται στην περιοχή υψηλών θερμοκρασιών, φαινόμενα που σχετίζονται με την καύση, όπως αυτό των ψυχρών φλογών, η συμπεριφορά αρνητικού συντελεστή θερμοκρασίας, καθώς και η αυτανάφλεξη παρουσιάζονται νωρίτερα εντός των περιοχών χαμηλών και ενδιάμεσων θερμοκρασιών. Γύρω στους 420K πυροδοτείται ο μηχανισμός οξείδωσης με σταδιακή αφαίρεση ατόμων υδρογόνου των υδρογονανθράκων από μοριακό οξυγόνο.



Η ενδοθερμική αυτή αντίδραση είναι εξαιρετικά αργή¹ και εξαρτάται από την ενέργεια του δεσμού του αποσπώμενου υδρογόνου. Η ενέργεια του δεσμού C–H είναι συνάρτηση της θέσης του υδρογόνου στο μόριο και εξαρτάται από τον αριθμό ατόμων άνθρακα με τα οποία συνδέεται ο άνθρακας του δεσμού, με ασθενέστερους τους τριτοταγείς δεσμούς C–H, όπου ο C συνδέεται με τρία άτομα άνθρακα, που ακολουθούνται από τους δευτεροταγείς και πρωτοταγείς, για δύο και ένα άτομα άνθρακα αντίστοιχα. Λόγω της απαιτούμενης ενέργειας ενεργοποίησης, η διαδικασία απόσπασης είναι πολύ εκλεκτική στο άτομο υδρογόνου και το μόριο του υδρογονάνθρακα που θα συμμετάσχει, προτιμώντας, για παράδειγμα, τα κανονικά αλκάνια (κανονικό επτάνιο), οι ευθείες αλυσίδες των οποίων έχουν πληθώρα δευτεροταγών δεσμών C–H, έναντι των ισοαλκανίων (ισοοκτάνιο) με διακλαδισμένες αλυσίδες που περιλαμβάνουν περισσότερους πρωτοταγείς δεσμούς.

Η ταχύτητα της αντίδρασης αυξάνεται με την αύξηση της θερμοκρασίας, ενώ παράλληλα οι παραγόμενες αλκυλικές ρίζες αρχίζουν να αντιδρούν με μοριακό οξυγόνο προς σχηματισμό $\text{RO}_2\cdot$.

¹ Για να εξασφαλίσει την απαιτούμενη ενέργεια, η αντίδραση απαιτεί την πρόσκρουση με ένα τρίτο συστατικό (M) του μίγματος (ελεύθερη ρίζα ή σταθερό μόριο).



Κατά την μετάβαση, όμως, στην περιοχή των ενδιάμεσων θερμοκρασιών η αντίδραση ακολουθεί αρχίζει να ακολουθεί και έναν δεύτερο δρόμο σχηματισμού υδροϋπεροξειδίου HO_2- και ολεφινών.



Οι αντιδράσεις αυτές είναι ανταγωνιστικές, αφού η αυξανόμενη με την θερμοκρασία παραγωγή HO_2- , αυξάνει παράλληλα την συγκέντρωση των προϊόντων της (A.1), γεγονός που έχει αρνητική επίδραση στην ταχύτητα της και οδηγεί σε μικρότερους ρυθμούς παραγωγής RO_2- . Στο στάδιο αυτό παρατηρείται εξισορρόπηση του μηχανισμού οξειδωσης με σταθεροποίηση των συγκεντρώσεων των ενεργών κέντρων δραστηριοποίησής του. Καθώς όμως η θερμοκρασία αυξάνει η αποσύνθεση του HO_2- προς $OH-$ υπερισχύει και ο μηχανισμός οξειδωσης αρχίζει ξανά να επιταχύνει.

Το φαινόμενο αυτό προσωρινής ανακοπής της εξέλιξης των αλυσιδωτών αντιδράσεων οξειδωσης με την αύξηση της θερμοκρασίας, λόγω αποκατάστασης χημικής ισορροπίας κατά των σχηματισμό ενδιάμεσων συστατικών, ονομάζεται αρνητικός συντελεστής θερμοκρασίας, NTC, και παρατηρήθηκε αρχικά σε πειράματα οξειδωσης προπανίου. Αποτελεί χαρακτηριστική συμπεριφορά των παραφινών, για πολλές από τις οποίες έχει παρουσιάζεται σε εύρος θερμοκρασιών 600-800K, και αποτελεί εξήγηση της μικρής ευαισθησίας που παρουσιάζουν έναντι των άλλων υδρογονανθράκων, αφού στις εντονότερες συνθήκες ελέγχου του MON η δράση του NTC εντείνεται, επίσης, διατηρώντας την ανθεκτικότητα του καυσίμου. Αρνητικός συντελεστής θερμοκρασίας έχει παρατηρηθεί και σε κάποιες ολεφίνες και αρωματικούς υδρογονάνθρακες, το φαινόμενο όμως δεν είναι τόσο δραστικό για τον συνολικό μηχανισμό όπως συμβαίνει κατά την οξειδωση των αλκανίων.

Οξειδωση Βενζίνης

Ο μηχανισμός οξειδωσης περιλαμβάνει ένα τεράστιο πλήθος αντιδράσεων, η γνώση των οποίων είναι ακόμη ελλιπής. Είναι ενδεικτικό ότι για υδρογονάνθρακες C_7 οι αντιδράσεις που συμμετέχουν ανέρχονται σε χιλιάδες με εκατοντάδες ενδιάμεσα συστατικά. Επίσης, παρ' όλο, που έχουν πραγματοποιηθεί εκτεταμένες μελέτες της διεργασίας καύσης των καθαρών υδρογονανθράκων ή δυαδικών μιγμάτων τους, η έρευνα, όσον αφορά την αλληλεπίδραση των κυριότερων ομάδων τους, πόσο μάλλον σε μίγματα με σύσταση που να προσεγγίζει αυτή των πραγματικών βενζινών, βρίσκεται ακόμη σε αρχικά στάδια. Μελέτες της συμπεριφοράς μιγμάτων ισοοκτανίου και κανονικού επτανίου με αλκένια και αρωματικούς υδρογονάνθρακες (Leppard (1992), Khan (1998)), έδειξαν ότι οι μηχανισμοί αυτανάφλεξης των ακόρεστων υδρογονανθράκων αλληλεπιδρούν έτσι ώστε να επιβραδύνουν την συνολική εξέλιξη της

οξειδωσης του μίγματος. Αν και τα αρωματικά δεν εμφανίζουν έντονη δραστηριότητα στην περιοχή των χαμηλών θερμοκρασιών, επηρεάζουν σημαντικά τις αλυσιδωτές αντιδράσεις επιδρώντας στο συνολικό απόθεμα ελευθέρων ριζών. Επίσης για μίγμα αλκανίου-αλκενίου προσδιορίστηκαν οι αναλογίες των ενδιάμεσων ειδών καύσης για κάθε συστατικό, απ' όπου προέκυψε ότι το αλκένιο επιδρά ελάχιστα στον μηχανισμό οξειδωσης του αλκανίου, ενώ αντίθετα η παρουσία αλκανίου μεταβάλλει σημαντικά του δρόμους αντίδρασης του αλκενίου.

Η παρουσία προσθέτων αύξησης του αριθμού οκτανίων στα μίγματα βενζινών λειτουργεί, επίσης, στην κατεύθυνση επιβράδυνσης του συνολικού μηχανισμού οξειδωσης. Αν και, τόσο οι αλκυλικές ενώσεις μολύβδου, όσο και τα οξυγονούχα πρόσθετα, είναι εξίσου αποτελεσματικά στην καθυστέρηση της αυτανάφλεξης και την καταστολή του "κτύπου", οι χημικοί μηχανισμοί δράσης τους είναι εξ' ολοκλήρου διαφορετικοί. Οι οξυγονούχες ενώσεις επιβραδύνουν σημαντικά την εξέλιξη των αντιδράσεων προφλογών χαμηλών θερμοκρασιών, καταναλώνοντας τις ενεργές ρίζες (κυρίως $\text{OH}\cdot$) προς σχηματισμό ισοβουτενίου. Το ισοβουτένιο αντιδρά με επιπλέον $\text{OH}\cdot$ με προϊόντα ανενεργές ρίζες, όπως αλλύλιο ($\text{CH}_2=\text{CHCH}_2\cdot$) και μέθυαλλύλιο, και σταθερά είδη, όπως διένια (αλλένια), ανθεκτικά σε περαιτέρω οξειδωση. Οι αλκυλικές ενώσεις του μολύβδου, αντίθετα, παρεμβαίνουν στο στάδιο των ενδιάμεσων θερμοκρασιών, απενεργοποιώντας την σημαντικότερη ανεπιθύμητη ακολουθία αλυσιδωτών αντιδράσεων, με κατανάλωση της κυρίαρχης ενεργής ρίζας $\text{HO}_2\cdot$ από οξείδια μολύβδου.

Αν και δεν είναι πλήρως κατανοητές οι διαδικασίες αλληλεπίδρασης των συστατικών της βενζίνης στην εξέλιξη του φαινομένου της αυτανάφλεξης των άκαυστων υπολοίπων αερίων, είναι προφανές ότι η συμπεριφορά των μιγμάτων διαφοροποιείται σημαντικά από αυτή των καθαρών συστατικών τους και σε καμία περίπτωση δεν είναι ανάλογη της συγκέντρωσή τους στο καύσιμο. Η περαιτέρω έρευνα πάνω στην οξειδωση των υδρογονανθράκων και η διερεύνηση των αλληλεπιδράσεων των επιμέρους κλασμάτων της βενζίνης είναι καθοριστική για την εξήγηση του φαινομένου μίξης και την καλύτερη πρόβλεψη του αριθμού οκτανίων των παραγόμενων βενζινών.

2 Πολυπαραμετρική Ανάλυση Δεδομένων

2.1 Εισαγωγή στην Χημειομετρία

Ο όρος **χημειομετρία** (chemometrics) εισήχθει από τους Svante Wold και Bruce R. Kowalski στις αρχές της δεκαετίας του '70, για να περιγράψει μεθόδους και τεχνικές που σχετίζονται με την μαθηματική επεξεργασία και ερμηνεία των χημικών δεδομένων. Πρόκειται για μεθόδους της Άλγεβρας, Αναλυτικής Γεωμετρίας και Στατιστικής, που ήδη έβρισκαν εφαρμογή στην *βιομετρία* (biometrics) και *οικονομετρία* (econometrics), προσαρμοσμένες όμως στα χημικά δεδομένα. Πριν από την περίοδο αυτή οι χημικοί περιορίζονταν κυρίως στην μονοπαραμετρική ανάλυση δεδομένων, με προσδιορισμό στατιστικών παραμέτρων τους, όπως η μέση τιμή, η σταθερή απόκλιση, το επίπεδο εμπιστοσύνης και την συσχέτιση των μεταβλητών. Από τις αρχές της δεκαετίας του '70 όμως, δόθηκε τεράστια ώθηση στην ανάπτυξη της χημειομετρίας, που αντιμετωπίζεται πλέον ως ξεχωριστός κλάδος της Χημείας, κυρίως λόγω της αυτοματοποίησης και εξέλιξης των αναλυτικών οργάνων που έφερε τους χημικούς και ιδιαίτερα τους αναλυτικούς χημικούς αντιμέτωπους με τις δυσκολίες στην ανάκτηση πληροφορίας από τεράστιο όγκο δεδομένων, που μέχρι τότε απασχολούσαν αποκλειστικά τους κοινωνικούς επιστήμονες και οικονομολόγους. Σημαντικά συνέβαλε και η πρόοδος που σημειώθηκε στην τεχνολογία της μικροηλεκτρονικής αυξάνοντας τις υπολογιστικές δυνατότητες ώστε να καλύψει τις αυξημένες απαιτήσεις των χημειομετρικών μεθόδων.

Πέρα από την επεξεργασία δεδομένων, η χημειομετρία περιλαμβάνει ένα σύνολο μαθηματικών εργαλείων που χρησιμοποιούνται κατά την εφαρμογή μιας αναλυτικής μεθόδου στο σύνολο της, από την επιλογή της μεθόδου ως την τελική ερμηνεία και αξιοποίηση της πληροφορίας των δεδομένων σε επιστημονικούς, αλλά και βιομηχανικούς τομείς. Κατά την εφαρμογή την εργαστηριακής αναλυτικής μεθόδου χρησιμοποιούνται χημειομετρικά κριτήρια αξιολόγησής της που αφορούν την ακρίβεια, την ευαισθησία και το όριο ανίχνευσης, καθώς και τεχνικές βελτιστοποίησης των παραμέτρων της, όπως η πολλαπλή παλινδρόμηση. Για την οργάνωση της διεξαγωγής μετρήσεων, η χημειομετρία παρέχει πολύτιμα εργαλεία υπό τον γενικό τίτλο του πειραματικού σχεδιασμού, όπως τον πλήρη και μερικό πειραματικό σχεδιασμό, τον σχεδιασμό D-optimal, κ.α., που περιορίζουν τον αριθμό των απαιτούμενων μετρήσεων ανάλογα με τον στόχο της έρευνας, μειώνοντας έτσι τον όγκο δεδομένων που περιέχει χρήσιμη πληροφορία. Θέματα της χημειομετρίας αποτελούν επίσης η ανάκτηση ψηφιακών δεδομένων (θεωρία δειγματοληψίας, όριο ανίχνευσης), η προεπεξεργασία τους (κανονικοποίηση και μετασχηματισμοί), η επιλογή και ανάδειξη των ιδιομορφιών τους μέσω στατιστικών παραμέτρων τους (πίνακας διακύμανσης-συνδιακύμανσης) και γραμμικών συνδυασμών των μετρούμενων

ιδιοτήτων-παραμέτρων (μέθοδος κυρίων συνιστωσών, ανάλυση παραγόντων) και η ενίσχυση σήματος (μείωση θορύβου, χαρακτηρισμός και μοντελοποίηση του σήματος) με την χρήση μαθηματικών φίλτρων εξομάλυνσης και υπέρθεσης.

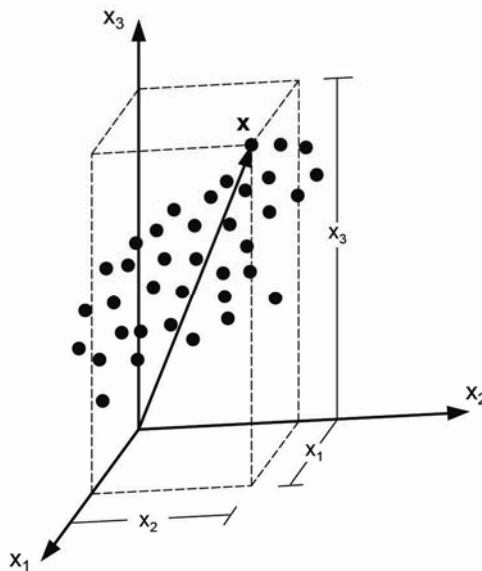
Τέλος, η χημειομετρία περιλαμβάνει ένα σύνολο τεχνικών πολυπαραμετρικής ανάλυσης που εφαρμόζονται για την ερμηνεία και αξιοποίηση της πληροφορίας των χημικών δεδομένων. Ανάλογα με τον στόχο της έρευνας, οι μέθοδοι ανάλυσης πολυμεταβλητών δεδομένων μπορούν να διαχωριστούν σε μεθόδους **ταξινόμησης** (classification), **πρόβλεψης** (prediction) και **βελτιστοποίησης** (optimization). Η ταξινόμηση αφορά τον διαχωρισμό των αντικειμένων-δειγμάτων και την καταχώρισή τους σε δύο ή περισσότερες ομάδες ή τάξεις. Η **αναγνώριση μορφών** της δομής ενός δειγματικού χώρου (pattern recognition) περιλαμβάνει μεθόδους **μη επιβλεπόμενης** ανάλυσης (unsupervised analysis), οι οποίες στοχεύουν στην ανάδειξη της διαφορετικότητας των δειγμάτων και αφορούν κυρίως ανάλυση σημνών για διαχωρισμό των δειγμάτων σε τάξεις χωρίς εκ των προτέρων πληροφορία για την σωστή ταξινόμησή τους, και μεθόδους **επιβλεπόμενης** ή **εποπτευόμενης ανάλυσης** (supervised analysis) για ανάπτυξη αλγορίθμων ταξινόμησης των δειγμάτων, οι οποίες λαμβάνουν υπ' όψιν τις αναμενόμενες αποκρίσεις-στόχους τους.

Οι μέθοδοι *πολυπαραμετρικής βαθμονόμησης και παλινδρόμησης*, που χρησιμοποιούνται για προσαρμογή συνάρτησης μεταξύ των ιδιοτήτων-παραμέτρων του δειγματικού χώρου, είναι ιδιαίτερα δημοφιλείς τα τελευταία χρόνια λόγω της ευρείας χρήσης τους και σε βιομηχανικές εφαρμογές. Οι αλγόριθμοι πρόβλεψης και βελτιστοποίησης που προκύπτουν από τεχνικές όπως η πολυωνυμική προσέγγιση, η πολλαπλή γραμμική παλινδρόμηση, η παλινδρόμηση κυρίων συνιστωσών, και η μέθοδος μερικών ελαχίστων τετραγώνων βρίσκουν πρακτική εφαρμογή σε συστήματα παραγωγής για τον έμμεσο προσδιορισμό ιδιοτήτων προϊόντων από άλλα χαρακτηριστικά τους χωρίς την παρεμβολή χρονοβόρων ποιοτικών ελέγχων στην συνεχή λειτουργία.

Οι παραπάνω μέθοδοι προϋποθέτουν σε μεγάλο βαθμό την ύπαρξη γραμμικών σχέσεων έστω και υπό λανθάνουσα μορφή στον δειγματικό χώρο που να επιτρέπουν την υποκατάσταση των μεταβλητών από μικρότερο αριθμό νέων μεταβλητών-παραγόντων. Η πρόοδος, όμως, που συντελείται τα τελευταία χρόνια στην επιστήμη των πληροφοριών και τα εφαρμοσμένα μαθηματικά έχουν εξοπλίσει την χημειομετρία με νέα εργαλεία: τεχνικές όπως τα νευρωνικά δίκτυα, τους γενετικούς αλγόριθμους (genetic algorithms) και την προσομοιούμενη ανόπτηση (simulated annealing), που έρχονται να αντιπαρέλθουν τις αδυναμίες των πολυπαραμετρικών μεθόδων.

2.2 Διερευνητική Ανάλυση

Οι πειραματικές μετρήσεις για την μελέτη φυσικοχημικών διεργασιών, η συγκέντρωση των οποίων γίνεται βάση πειραματικού σχεδιασμού που εξετάζεται σε ξεχωριστό κεφάλαιο, συγκροτούνται για τις ανάγκες των χημειομετρικών εφαρμογών σε πίνακα δεδομένων $\mathbf{X}$. Στις περισσότερες περιπτώσεις τα n δείγματα του πληθυσμού αναλύονται ως προς περισσότερες από μία ιδιότητές τους και ο πίνακας των πολυπαραμετρικών δεδομένων $\mathbf{X}_{(n \times m)} = [x_{ij}]$, με $i=1,2,\dots,n$ και $j=1,2,\dots,m$, έχει στήλες τις m μεταβλητές $X_1, X_2, \dots, X_j, \dots, X_m$ και γραμμές τα n διανύσματα των τιμών τους $\mathbf{x}_i = \{x_{ij}\} = \{x_{i1}, x_{i2}, \dots, x_{im}\}$ ανά δείγμα. Στον πολυδιάστατο χώρο που ορίζουν οι μεταβλητές κάθε δείγμα αντιστοιχεί σε ένα σημείο με συντεταγμένες $(x_1, x_2, \dots, x_m)$ τις προβολές του στους άξονες των μεταβλητών. Στο σχήμα Σχ. 2-1 απεικονίζεται ένας πληθυσμός δειγμάτων προσδιορισμένων ως προς τρεις ιδιότητες τους.



Σχήμα 2-1 Γραφική απεικόνιση πολυπαραμετρικών δεδομένων στον 3-διάστατο χώρο

Στις περισσότερες εφαρμογές, το μεγάλο πλήθος δεδομένων καθιστά εξαιρετικά δυσχερή την μελέτη των αριθμητικών στοιχείων του πίνακα $\mathbf{X}$ και την εξαγωγή συμπερασμάτων για τις τάσεις διαμόρφωσης κάθε μεταβλητής και των εξαρτήσεων μεταξύ τους, ενώ η γενική εμποπτεία των δεδομένων με χρήση γραφικής απεικόνισης περιορίζεται αυστηρά στις τρεις διαστάσεις. Ο σκοπός της **διερευνητικής ανάλυσης** (exploratory analysis) είναι να αποκτηθεί, ήδη από τα αρχικά στάδια και πριν την εφαρμογή οποιασδήποτε μεθόδου ταξινόμησης ή προσέγγισης συναρτησιακής εξάρτησης μεταξύ των μεταβλητών, όσο το δυνατόν καλύτερη άποψη της εσωτερικής δομής των δεδομένων και να αναδειχθεί το μεγαλύτερο ποσοστό της εσωτερικής τους μεταβλητότητας, εκείνης δηλαδή που εμπεριέχει χρήσιμη πληροφορία. Κατάλληλη

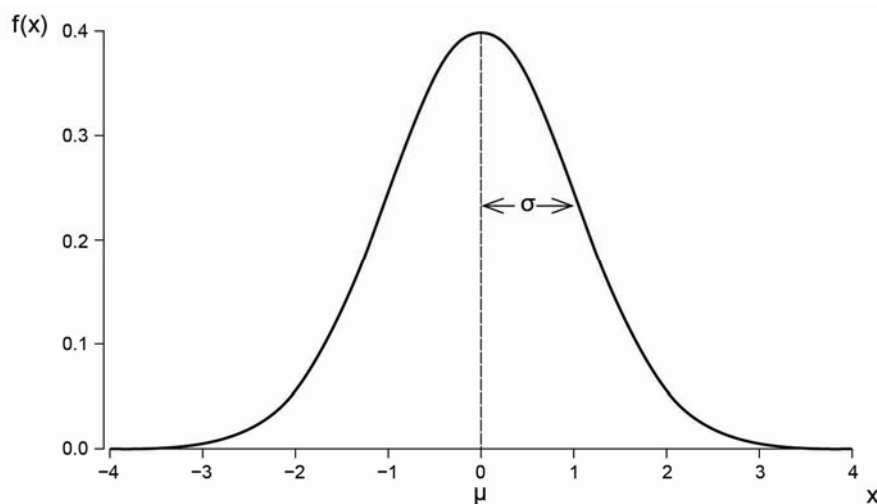
μαθηματική επεξεργασία και προσδιορισμός στατιστικών μεγεθών είναι δυνατόν να απομονώσουν το μεγαλύτερο ποσοστό του πειραματικού θορύβου και να αναδείξουν τις πραγματικές τάσεις που διέπουν την διασπορά των δεδομένων στον πειραματικό χώρο που ορίζουν.

Βασικές Έννοιες Στατιστικής

Αν και μία ιδανική κατανομή των πειραματικών δεδομένων είναι στην πράξη ανέφικτη, λόγω την ύπαρξης συστηματικών και τυχαίων σφαλμάτων, η συνήθης υπόθεση για πειραματικά δεδομένα φυσικών διεργασιών είναι ότι οι τιμές μίας μεταβλητής ακολουθούν *κανονική* ή *Gaussian κατανομή*. Η ιδανική κανονική κατανομή για μεμονωμένη τυχαία μεταβλητή X εκφράζει την πιθανότητα εμφάνισης μίας τιμής x της μεταβλητής για ένα μεγάλο πληθυσμό εντός ορισμένου εύρους τιμών. Η μορφή της κατανομής παρουσιάζεται στο Σχ. 2-2 και περιγράφεται μαθηματικά από την σχέση:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (2.1)$$

Η καμπύλη είναι συμμετρική γύρω από την μέση τιμή μ και το εύρος της ως προς την μέση τιμή δίνεται από την διασπορά σ^2 .



Σχήμα 2-2 Κανονική ή Gaussian κατανομή

Μέτρα θέσης και διασποράς

Για την περιγραφή της κατανομής συχνοτήτων μιας μεταβλητής χρησιμοποιούνται τα λεγόμενα μέτρα θέσης και διασποράς, που εκφράζουν την θέση του “κέντρου” των τιμών της μεταβλητής και την διασπορά τους γύρω από το “κέντρο” αντίστοιχα. Έστω η τυχαία μεταβλητή X_i ως προς την οποία εξετάζεται ένα σύνολο πεπερασμένου αριθμού n δειγμάτων του πληθυσμού. Μία

από τις κυριότερες παραμέτρους της κατανομής είναι η **μέση τιμή** ή **αριθμητικός μέσος** (average ή arithmetic mean) που αποτελεί στατιστική εκτίμηση της πραγματικής μέσης τιμής του συνολικού πληθυσμού. Η μέση τιμή $\bar{x}_j$ της μεταβλητής X_j , που αντιπροσωπεύεται από το διάνυσμα-στήλη με τιμές $\mathbf{x}_j = \{x_{ij}\} = \{x_{1j}, x_{2j}, \dots, x_{nj}\}$, με στοιχεία x_{ij} τις τιμές της μεταβλητής, ορίζεται μαθηματικά από την σχέση:

$$\bar{x}_j = \frac{1}{n} \sum_{i=1}^n x_{ij} \quad (2.2)$$

Πρόκειται για μέγεθος που μαζί με την **διάμεσο τιμή** (median), η οποία αντιπροσωπεύει την τιμή κάτω από την οποία βρίσκεται το 50% του συνόλου των εξεταζόμενων δειγμάτων, προσδιορίζουν την θέση των πειραματικών δεδομένων στον μονοδιάστατο χώρο της εν λόγω μεταβλητής. Επίσης η απόκλιση των τιμών των δύο αυτών μεγεθών, που είναι μηδενική για ιδανική κατανομή, εκφράζει την παραμόρφωση της καμπύλης κατανομής της εξεταζόμενης μεταβλητής σε σχέση με αυτή της ιδανικής.

Το απλούστερο από τα μέτρα διασποράς είναι το **εύρος** (range) που υπολογίζεται ως διαφορά της ελάχιστης από την μέγιστη τιμή της μεταβλητής.

$$R = x_{j_{\max}} - x_{j_{\min}} \quad (2.3)$$

Πιο εξεζητημένο μέτρο της διασποράς που εκφράζει την απόκλιση των τιμών από την μέση τιμή αποτελεί η **διακύμανση** ή **διασπορά** (variance) –στατιστική εκτίμηση της πραγματικής διακύμανσης του συνολικού πληθυσμού–, που ορίζεται από την σχέση:

$$\text{VAR}_{x_j} = s_j^2 = \frac{1}{n} \sum_{i=1}^n (x_{ij} - \bar{x}_j)^2 \quad (2.4)$$

$$\text{ή με χρήση πινάκων} \quad \text{VAR}_{x_j} = s_j^2 = \frac{1}{n} \mathbf{x}_{d_j}^T \cdot \mathbf{x}_{d_j} \quad (2.5)$$

όπου $\mathbf{x}_{d_j}$ το διάνυσμα-στήλη των διορθωμένων ως προς την μέση τιμή τιμών της μεταβλητής $\mathbf{x}_{d_j} = \mathbf{x}_j - \bar{x}_j$.

Η τετραγωνική ρίζα της διακύμανσης δίνει ένα άλλο μέτρο διασποράς, την **τυπική απόκλιση** (standard deviation), που πλεονεκτεί αυτής στο ότι εκφράζεται στις ίδιες μονάδες μέτρησης με την εξεταζόμενη μεταβλητή.

$$\text{STDEV}_{x_j} = s_j = \sqrt{s_j^2} = \sqrt{\text{VAR}_{x_j}} \quad (2.6)$$

Τέλος, ένα μέτρο *σχετικής διασποράς* των τιμών των δειγμάτων, ανεξάρτητο μονάδων και γι' αυτό εξαιρετικά χρήσιμο για την σύγκριση κατανομών διαφορετικών μεταβλητών, αποτελεί ο **συντελεστής μεταβλητότητας** (coefficient of variation), που εκφράζει την τυπική απόκλιση μίας κατανομής ως προς την μέση τιμή της.

$$CV_{X_j} = \frac{s_j}{\bar{X}_j} \quad (2.7)$$

Συνδιακύμανση και συσχέτιση

Τα παραπάνω μεγέθη αποτελούν περιγραφικούς στατιστικούς δείκτες της διασποράς που παρουσιάζει μία μεμονωμένη μεταβλητή του συνόλου των πολυπαραμετρικών δεδομένων. Αν και ο προσδιορισμός τους αποτελεί ένα σημαντικό τμήμα της ανάλυσης, η μονοπαραμετρική αυτή εξέταση δεν μπορεί να αποδώσει την εσωτερική μεταβλητότητα του συνολικού πίνακα των πολυπαραμετρικών δεδομένων. Στα πλαίσια της **πολυπαραμετρικής ανάλυσης** (multivariate analysis), που αποτελεί και το κύριο θέμα της χημειομετρικής επεξεργασίας, ενδιαφέρει ιδιαίτερα ο προσδιορισμός μεγεθών που αποδίδουν τις αλληλεπιδράσεις μεταξύ των μεταβλητών, αφού στατιστική ανεξαρτησία τους σπάνια εμφανίζεται σε πραγματικά προβλήματα. Ο βαθμός της αλληλεπίδρασης εκφράζεται από την **συνδιακύμανση** ή **συνδιασπορά** (covariance) των μεταβλητών ανά δύο. Για δύο μεταβλητές X_j και X_ℓ του πίνακα δεδομένων, με $\mathbf{x}_j = \{x_{ij}\} = \{x_{1j}, x_{2j}, \dots, x_{nj}\}$ και $\mathbf{x}_\ell = \{x_{i\ell}\} = \{x_{1\ell}, x_{2\ell}, \dots, x_{n\ell}\}$, η συνδιακύμανση δίνεται από την σχέση

$$COV_{X_j X_\ell} = s_{j\ell} = \frac{1}{n} \sum_{i=1}^n (x_{ij} - \bar{x}_j)(x_{i\ell} - \bar{x}_\ell) \quad (2.8)$$

$$\text{ή} \quad COV_{X_j X_\ell} = s_{j\ell} = \frac{1}{n} \mathbf{x}_{dj}^T \cdot \mathbf{x}_{d\ell} = \frac{1}{n} \mathbf{x}_{d\ell}^T \cdot \mathbf{x}_{dj} = s_{\ell j} \quad (2.9)$$

Κατά την επεξεργασία μεγάλου όγκου πολυπαραμετρικών δεδομένων είναι συνήθης η απόδοση των στατιστικών παραμέτρων τους υπό μορφή πινάκων. Η διασπορά των μεταβλητών του δειγματικού χώρου και βαθμός αλληλεπίδρασης μεταξύ τους ανά δύο εκφράζεται με τον πίνακα διακύμανσης-συνδιακύμανσης:

$$COV_x = \frac{1}{n} \mathbf{X}_d^T \cdot \mathbf{X}_d \quad (2.10)$$

$$\text{με στοιχεία} \quad COV_{x_{j\ell}} = \frac{1}{n} \sum_{i=1}^n (x_{ij} - \bar{x}_j)(x_{i\ell} - \bar{x}_\ell) \quad (2.11)$$

όπου $i=1,2,\dots,n$ και $j, \ell=1,2,\dots,m$

Ο πίνακας διακύμανσης-συνδιακύμανσης είναι *τετραγωνικός* $m \times m$ και *συμμετρικός*, αφού για την συνδιακύμανση ισχύει $s_{ij} = s_{ji}$, με τα διαγώνια στοιχεία του $s_{ij} = s_i^2$, για $i=j$ να αντιστοιχούν στις διακυμάνσεις των μεταβλητών.

¹ Όταν η συνδιακύμανση, όπως και η διακύμανση, δεν υπολογίζονται από τον συνολικό πληθυσμό έρευνας, αλλά από ένα μικρό σχετικά δείγμα του ($n < 30$) τότε ο συντελεστής του αθροίσματος στις σχέσεις είναι $1/(n-1)$

Για τον προσδιορισμό του βαθμού συσχέτισης μεταξύ των μεταβλητών, απαλλαγμένου της επίδρασης των μονάδων μέτρησης χρησιμοποιείται ο **συντελεστής συσχέτισης** (correlation coefficient).

$$r_{j\ell} = \frac{\text{COV}_{X_j, X_\ell}}{s_j s_\ell} = \frac{s_{j\ell}}{s_j s_\ell} \quad (2.12)$$

Ο συντελεστής συσχέτισης είναι καθαρός αριθμός που λαμβάνει τιμές μεταξύ -1 και 1 , όπου τιμές του κοντά στην μονάδα υποδεικνύουν ισχυρή θετική αλληλεξάρτηση των μεταβλητών, κοντά στο -1 ισχυρή αρνητική αλληλεξάρτηση, ενώ μηδενικές τιμές του εμφανίζουν μεταβλητές γραμμικά ανεξάρτητες.

Προεπεξεργασία Δεδομένων

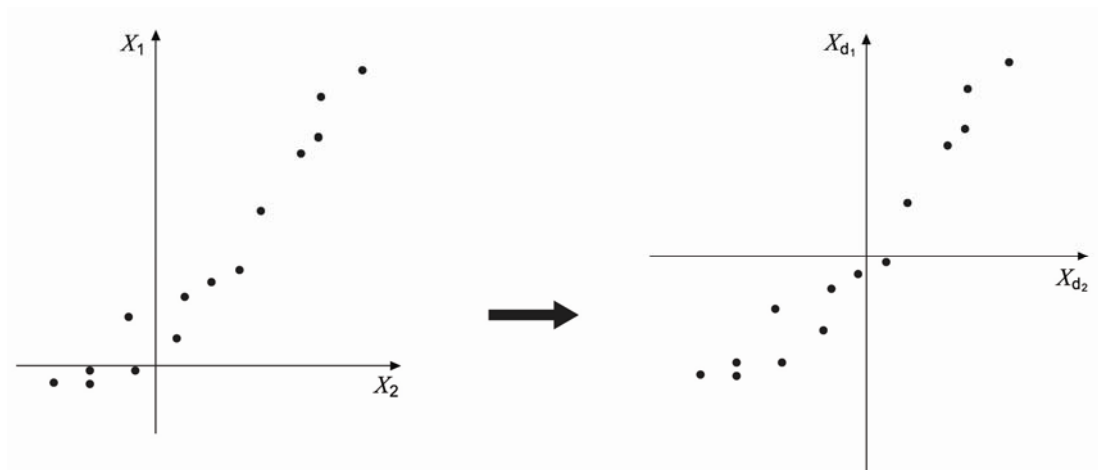
Οι ακατέργαστες μετρήσεις που προκύπτουν κατά την διεξαγωγή ενός πειράματος, είτε πρόκειται για αριθμητικές τιμές, είτε είναι φασματικής μορφής, αφού ψηφιοποιηθούν και ανάλογα με την φύση των μετρήσεων και με το είδος της ανάλυσης που πρόκειται να ακολουθήσει (ταξινόμηση, πρόβλεψη ή βελτιστοποίηση) είναι απαραίτητο να υποστούν και την ανάλογη **προεπεξεργασία** (preprocessing). Τα ανεπεξέργαστα δεδομένα συνήθως εμφανίζουν κατανομή ακατάλληλη για ανάλυση. Οι επιδράσεις περιβάλλοντος, τα συστηματικά σφάλματα των οργάνων, οι διαφορετικές μονάδες μέτρησης των παραμέτρων, η διαφορετική διασπορά των μετρούμενων ιδιοτήτων, κ.α., δυσχεραίνουν την ικανότητα των χημειομετρικών μεθόδων για εξαγωγή πληροφορίας με νόημα. Η προεπεξεργασία των δεδομένων μειώνει σημαντικά τον “θόρυβο” από τέτοιες επιδράσεις. Η κυρίως προεπεξεργασία αφορά κυρίως την στάθμιση και κανονικοποίηση των δεδομένων, ενώ μπορεί να επεκταθεί και (αν κρίνεται απαραίτητο) σε χρήση άλλων μετασχηματισμών του πίνακά τους.

Μεταφορά αξόνων

Κατά το πρώτο στάδιο της προεπεξεργασίας η μέση τιμή κάθε μεταβλητής αφαιρείται από όλες τις τιμές της. Η διαδικασία αυτή ονομάζεται **μεσοστάθμιση** ή **διόρθωση ως προς την μέση τιμή** (mean-centering) και μεταφέρει το σύστημα αξόνων του m -διάστατου χώρου στο κέντρο του σμήνους των σημείων των δειγμάτων (κέντρο του δειγματικού χώρου). Οι διορθωμένες τιμές της j -οστής μεταβλητής του μεσοσταθμισμένου πίνακα $\mathbf{X}_d$ δίνονται από την σχέση:

$$x_{dij} = x_{ij} - \bar{x}_j \quad (2.13)$$

Η μεσοστάθμιση των μεταβλητών διασφαλίζει ότι όλα τα αποτελέσματα θα είναι ερμηνεύσιμα υπό όρους απόστασης από την μέση τιμή. Στο Σχ. 2-3 δίνεται η γεωμετρική απεικόνιση της μεσοστάθμισης των τιμών σε διδιάστατο δειγματικό χώρο.



Σχήμα 2-3 Απεικόνιση σε δύο διαστάσεις της διόρθωσης ως προς την μέση τιμή των στοιχείων του πίνακα δεδομένων

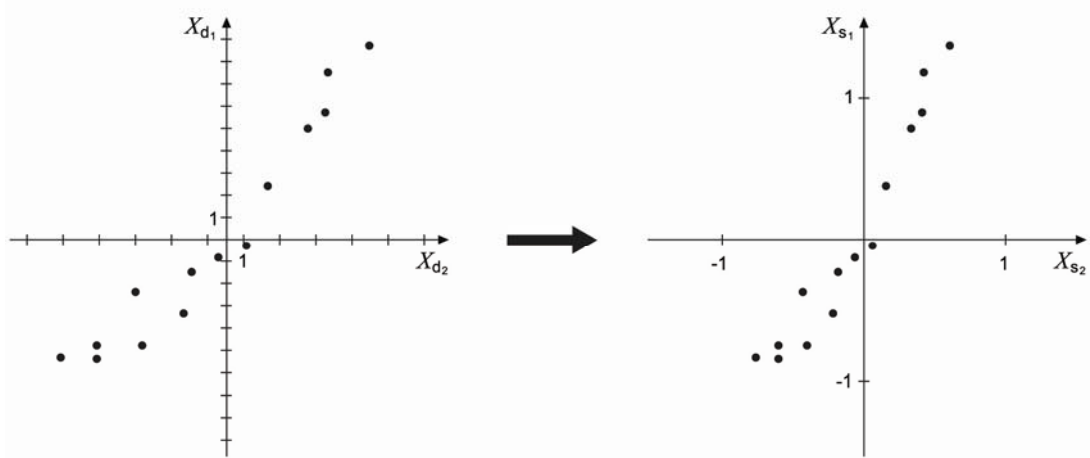
Στάθμιση Δεδομένων

Μέθοδοι προβολής σημείων του δειγματικού χώρου σε άξονες, όπως η ανάλυση κυρίων συνιστωσών (PCA), η παλινδρόμηση κυρίων συνιστωσών (PCR) και η παλινδρόμηση μερικών ελαχίστων τετραγώνων (PLS) στοχεύουν στον προσδιορισμό διευθύνσεων μέγιστης διασποράς του σμήνους των σημείων που αντιστοιχούν στον πληθυσμό των δεδομένων. Επομένως εξαρτώνται από την σχετική διακύμανση των μεταβλητών. Αναλόγως με το είδος της πληροφορίας που αναμένεται να εξαχθεί από τα δεδομένα, χρησιμοποιούνται βάρη στάθμισης των δεδομένων ανά μεταβλητή. Συνήθως διαιρούνται με την τυπική απόκλιση, που εκφράζει την διασπορά τους στις ίδιες μονάδες με την αρχική μεταβλητή. Η διαδικασία αναφέρεται ως **στάθμιση** (scaling) του πίνακα δεδομένων.

Η επεξεργασία του πίνακα δεδομένων με διόρθωσή τους ως προς την μέση τιμή και στάθμιση ως προς την διακύμανση κάθε μεταβλητής αναφέρεται ως **τυποποίηση δεδομένων** (data standardization). Ο τυποποιημένος πίνακας $\mathbf{X}_s$ έχει στοιχεία που δίνονται από την σχέση:

$$x_{sij} = \frac{x_{ij} - \bar{x}_j}{s_j} \quad (2.14)$$

Η τυποποίηση των δεδομένων μεταβάλλει την ενδοσημειακή τους απόσταση, έτσι ώστε όλες οι μεταβλητές να εμφανίζουν την ίδια διακύμανση ίση με την μονάδα. Χρησιμοποιείται για πολυπαραμετρική επεξεργασία μεταβλητών διαφορετικού είδους ή μεταβλητές με διαφορετική μονάδα μέτρησης ή και διαφορετικό εύρος τιμών, ενέχει όμως τον κίνδυνο να οδηγήσει σε μείωση της αξιοπιστίας του μοντέλου αν στην ανάλυση συμπεριλαμβάνονται και μεταβλητές με μικρή τυπική απόκλιση που περιέχουν θόρυβο, οπότε η επίδρασή τους μεγεθύνεται. Στο Σχ. 2-4 παρουσιάζεται η γεωμετρική απεικόνιση διπαραμετρικών τυποποιημένων δεδομένων.



Σχήμα 2-4 Απεικονίζει σε δύο διαστάσεις της τυποποίησης του πίνακα δεδομένων

Στην ίδια βάση διατήρησης της συσχέτισης των δεδομένων και ανάλογα με τις ανάγκες της ανάλυσης μπορούν να εφαρμοστούν και άλλοι τύποι προεπεξεργασίας, όπως η **κανονικοποίηση** (normalization) των δεδομένων των δειγμάτων:

ως προς το άθροισμα των μεταβλητών:
$$x_{nij} = x_{ij} / \sum_{j=1}^m x_{ij} \quad (2.15)$$

ως προς το εύρος κάθε μεταβλητής:
$$x_{nij} = (x_{ij} - x_{jmin}) / (x_{jmax} - x_{jmin}) \quad (2.16)$$

ως προς το μέγιστο κάθε μεταβλητής:
$$x_{nij} = x_{ij} / x_{jmax} \quad (2.17)$$

ως κάποια ιδιότητα αναφοράς:
$$x_{nij} = x_{ij} / x_{is} \quad (2.18)$$

Τέλος, αν κρίνεται αναγκαίο γίνεται χρήση **μετασχηματισμών** (transformations) του πίνακα δεδομένων. Χαρακτηριστικό παράδειγμα είναι η λογαρίθμηση του πίνακα δεδομένων στις περιπτώσεις όπου διαπιστώνεται εκθετική εξάρτηση μεταξύ των παραμέτρων του δειγματικού χώρου:

$$x_{trij} = \log(x_{ij}) \quad (2.19)$$

Συσχέτιση Μεταβλητών

Πιο εξελιγμένες μέθοδοι διερευνητικής ανάλυσης αποτελούν οι τεχνικές πολυπαραμετρικής επεξεργασίας δεδομένων, οι οποίες αποσκοπούν στην ανάδειξη της σημαντικότητας κάθε μεταβλητής και των συσχετίσεων μεταξύ τους, με απώτερο στόχο την μείωση του αριθμού των στοιχείων, διατηρώντας παράλληλα το μεγαλύτερο ποσοστό της χρήσιμης πληροφορίας που περιέχει ο αρχικός πίνακας δεδομένων. Η μελέτη πολλών πολυμεταβλητών προβλημάτων μπορεί να απλοποιηθεί με εξέταση όχι μόνο των αρχικών μεταβλητών αλλά και γραμμικών συνδυασμών τους. Συγκροτείται δηλαδή, με χρήση μαθηματικών τεχνικών, ένα νέο σύνολο μεταβλητών κάθε μία από τις οποίες αποτελεί άθροισμα των αρχικών μεταβλητών

σταθμισμένων κατάλληλα. Οποιαδήποτε κι αν είναι η χρησιμοποιούμενη μέθοδος, ο στόχος είναι να μειωθεί ο αριθμός μεταβλητών που υπεισέρχονται στην κυρίως ανάλυση, μετασχηματίζοντας των αρχικό πίνακα σε ένα νέο μικρότερων διαστάσεων, τα στοιχεία του οποίου περιέχουν μία βελτιωμένη ανακατανομή της αρχικής διασποράς των δεδομένων.

Γραμμικοί Συνδυασμοί Μεταβλητών

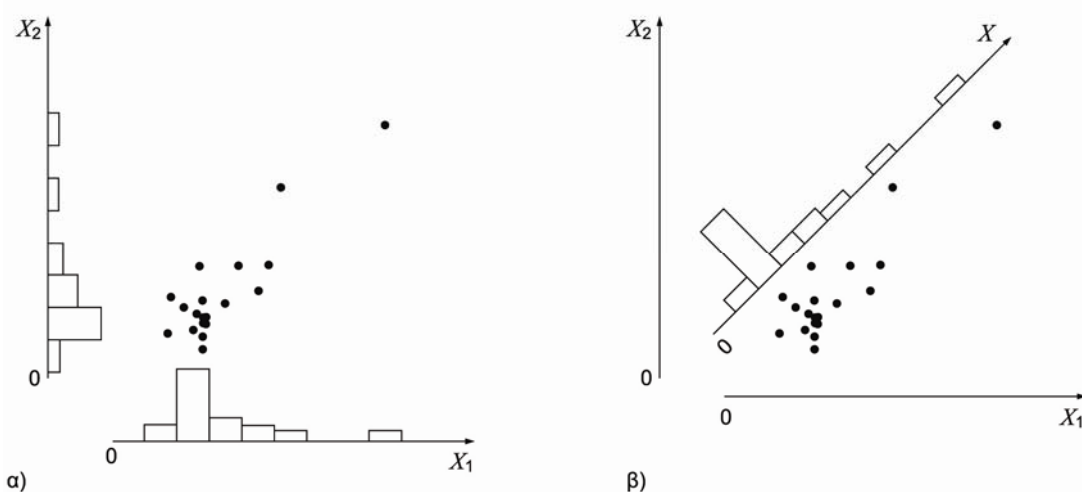
Η εξέταση του πίνακα διακύμανσης-συνδιακύμανσης συνήθως υποδεικνύει την ύπαρξη ισχυρών συσχετίσεων μεταξύ κάποιων εκ των μεταβλητών. Στο Σχ. 2-5α) παρουσιάζεται το γράφημα διασποράς των δεδομένων ως προς δύο μεταβλητές X_1 και X_2 που παρουσιάζουν ισχυρή συσχέτιση μεταξύ τους. Στο ίδιο διάγραμμα απεικονίζονται και οι μονοδιάστατες κατανομές συχνοτήτων των μεταβλητών, που προκύπτουν από τις προβολές των σημείων στους αντίστοιχους άξονες. Η ισχυρή συσχέτιση μεταξύ των μεταβλητών δίνει την δυνατότητα αντικατάστασης των αξόνων τους από κάποιο μεμονωμένο άξονα νέας μεταβλητής X , πάνω στον οποίο οι προβολές των δειγμάτων να παρουσιάζουν μεγαλύτερη διασπορά (Σχ. 2-5β)) και επομένως να περιέχεται σ' αυτόν μεγαλύτερη ποσότητα πληροφορίας σε σχέση με εκείνη που περιέχεται σε κάθε έναν από τους αρχικούς. Η μεταβλητή που αντιστοιχεί στον άξονα αυτό μπορεί να αποτελεί γραμμικό συνδυασμό των αρχικών μεταβλητών X_1 και X_2 :

$$X = aX_1 + bX_2 \quad (2.20)$$

υπό συνθήκη κανονικοποίησης
για τους συντελεστές:

$$a^2 + b^2 = 1 \quad (2.21)$$

ώστε οι τιμές της να σταθμίζονται στο εύρος τιμών που ορίζονται οι αρχικές μεταβλητές.



Σχήμα 2-5 Γράφημα διασποράς και κατανομές δεδομένων α) ως προς δύο υψηλής συσχέτισης μεταβλητές X_1 και X_2 και β) ως προς τον γραμμικό συνδυασμό τους $X = aX_1 + bX_2$

Ο προηγούμενος γραμμικός συνδυασμός στην πραγματικότητα αποτελεί στροφή των αρχικών αξόνων-μεταβλητών κατά γωνία φ , αφού οι συντελεστές a και b μπορούν να γραφούν ως τριγωνομετρικοί αριθμοί της φ :

$$\begin{aligned} a &= \cos \varphi \\ b &= \sin \varphi \end{aligned} \quad (2.22)$$

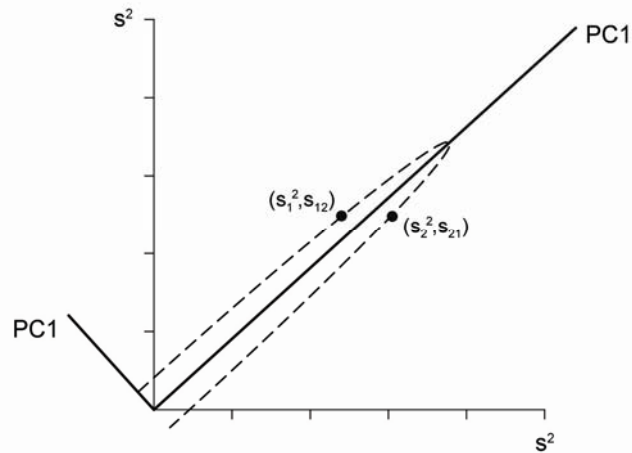
Σωστή επιλογή των a και b ή φ μπορεί να οδηγήσει σε μείωση των διαστάσεων του αρχικού πίνακα δεδομένων. Αν το ποσοστό της διακύμανσης στον δεύτερο άξονα X' του περιστραμμένου συστήματος είναι εξαιρετικά μικρό και μπορεί να θεωρηθεί ότι το μεγαλύτερο ποσοστό της χρήσιμης πληροφορίας περικλείεται στον πρώτο γραμμικό συνδυασμό των X_1 και X_2 , ο X' απαλείφεται και η μεταβλητή X αντικαθιστά τις αρχικές X_1 και X_2 στον πίνακα δεδομένων.

Ανάλυση Κυρίων Συνιστωσών

Η αναδιοργάνωση και καταμερισμός της μεταβλητότητας που αναλαμβάνουν οι αρχικές μεταβλητές σε νέες, καθώς και ο ως ένα βαθμό διαχωρισμός της χρήσιμης πληροφορίας που αυτή περιέχει από τον πειραματικό θόρυβο, αποτελούν τους κύριους στόχους της **ανάλυσης κυρίων συνιστωσών** (principal component analysis, PCA). Πρόκειται για χημειομετρική επεξεργασία πολυπαραμετρικών δεδομένων που εφαρμόζει περιστροφή και μετασχηματισμό του αρχικού m -διάστατου συστήματος αξόνων, έτσι ώστε οι άξονες του νέου συστήματος (κύριες συνιστώσες, PCs) να βρίσκονται πάνω στις κατευθύνσεις μέγιστης διασποράς του σμήνους δεδομένων. Ο μετασχηματισμός αυτός εφαρμόζεται υπό τον περιορισμό *ορθογωνιότητας* του συστήματος, έτσι ώστε οι νέες μεταβλητές που αντιπροσωπεύουν οι κύριες συνιστώσες να είναι *ασυσχέτιστες*. Με την ανάλυση κυρίων συνιστωσών επιχειρείται επομένως ο προσδιορισμός των τυχόν υφιστάμενων γραμμικών συσχετίσεων-τάσεων ενός μεγάλου αριθμού μεταβλητών, με δύο κυρίως στόχους: την συμπύκνωση της πληροφορίας που περιέχεται στις αρχικές μεταβλητές σε ένα μικρότερο αριθμό νέων μεταβλητών, με όσο το δυνατόν μικρότερη απώλεια πληροφορίας, καθώς επίσης και την διερεύνηση της εσωτερικής δομής των δεδομένων που ορίζουν οι συνδυασμοί αυτοί μεταβλητών και την φυσική ερμηνεία της.

Ο προσδιορισμός των κυρίων συνιστωσών γίνεται με βάση τις *ιδιοτιμές* (eigenvalues) και τα *ιδιοδιανύσματα* (eigenvectors) του πίνακα $\mathbf{C}_{(m \times m)}$ διακύμανσης-συνδιακύμανσης του πίνακα δεδομένων. Για διδιάστατο χώρο δύο μεταβλητών X_1 και X_2 , ο πίνακας $\mathbf{C}_{(2 \times 2)}$ μπορεί να παρασταθεί γραφικά σε σύστημα αξόνων διασποράς από δύο σημεία με συντεταγμένες (s_1^2, s_{12}) και (s_2^2, s_{21}) αντίστοιχα (Σχ. 2-6). Τα σημεία αυτά βρίσκονται στην περιφέρεια έλλειψης με κέντρο το κέντρο του συστήματος αξόνων και άξονες τις κύριες συνιστώσες. Η κλίση του μεγάλου

άξονα ή άξονα πρώτης κύριας συνιστώσας είναι ίση με το αντίστοιχο ιδιοδιάνυσμα του πίνακα **C** και το μήκος του ισούται με την αντίστοιχη ιδιοτιμή. Η δεύτερη κύρια συνιστώσα ορίζεται με βάση το δεύτερο ιδιοδιάνυσμα ή ιδιοτιμή.



Σχήμα 2-6 Γραφική απεικόνιση του πίνακα διακύμανσης-συνδιακύμανσης δύο μεταβλητών και των κυρίων συνιστωσών τους

Η μαθηματική επεξεργασία του πίνακα δεδομένων για τον προσδιορισμό των κυρίων συνιστωσών περιλαμβάνει τον μετασχηματισμό του πίνακα διακύμανσης-συνδιακύμανσης $C_{(m \times m)}$ των αρχικών μεταβλητών σε γινόμενο δύο πινάκων:

$$C = S \cdot L \quad (2.23)$$

όπου **S** ο πίνακας των scores factors με γραμμές τις m κύριες συνιστώσες και στήλες τα n δείγματα του πολυπαραμετρικού πληθυσμού

L ο πίνακας των loadings factors με γραμμές τις m κύριες συνιστώσες και στήλες τα m συνημίτονα κατεύθυνσης

Για τον προσδιορισμό των **L** και **S**, αρχικά υπολογίζεται ο *συμμετρικός, τετραγωνικός* πίνακας διακύμανσης-συνδιακύμανσης $C_{(m \times m)}$ του τυποποιημένου πίνακα δεδομένων:

$$C = X_d^T \cdot X_d \quad (2.24)$$

Ο πίνακας **C** διαγωνοποιείται με εύρεση πίνακα **Q** τέτοιου ώστε:

$$Q^{-1} \cdot C \cdot Q = [\lambda_j \delta_{jk}] = \Lambda \quad (2.25)$$

όπου $\delta_{jk} = \begin{cases} 0, & j \neq k \\ 1, & j = k \end{cases}$ το δέλτα του Kronecker

και λ_j μία ιδιοτιμή της χαρακτηριστικής εξίσωσης $C \cdot q_j = \lambda_j q_j$ με q_j την j-οστή του *ορθογώνιου* πίνακα **Q** ($Q^{-1} = Q^T$) που αναφέρεται ως ιδιοδιάνυσμα του **C**

Το πρώτο μέλος της (2.25) γράφεται:

$$Q^{-1} \cdot C \cdot Q = Q^{-1} \cdot (X_d^T \cdot X_d) \cdot Q = Q^T \cdot (X_d^T \cdot X_d) \cdot Q = (Q^T \cdot X_d^T) \cdot X_d \cdot Q = S^T \cdot S \quad (2.26)$$

Είναι επομένως:

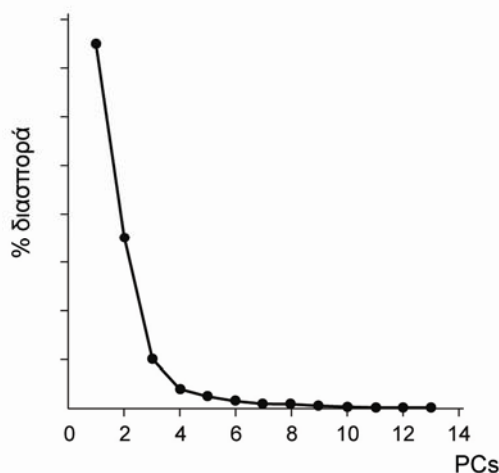
$$\mathbf{X}_d \cdot \mathbf{Q} = \mathbf{S} \text{ ή } \mathbf{X}_d = \mathbf{S} \cdot \mathbf{Q}^{-1} = \mathbf{S} \cdot \mathbf{Q}^T \quad (2.27)$$

Οι ζητούμενοι επομένως πίνακες των loadings και scores factors δίνονται από τις σχέσεις:

$$\mathbf{S} = \mathbf{X}_d \cdot \mathbf{Q} \quad \text{και} \quad \mathbf{L} = \mathbf{Q}^T \quad (2.28)$$

Ο αναλυτικός υπολογισμός των ιδιοτιμών και ιδιοδιανυσμάτων είναι αριθμητικά επιβαρυνμένος και γι' αυτό οι πίνακες loadings και scores factors στην πράξη προσδιορίζονται με επαναληπτικές μεθόδους αριθμητικής ανάλυσης που υλοποιούνται σε ανάλογα λογισμικά υπολογιστών.

Όπως αναφέρθηκε παραπάνω το μέγεθος της ιδιοτιμής υποδεικνύει την σημαντικότητα του αντίστοιχου ιδιοδιανύσματος και πόση από την αρχική διασπορά του πίνακα δεδομένων εξηγείται από την αντίστοιχη κύρια συνιστώσα. Στο Σχ. 2-7 δίνεται ένα παράδειγμα διαγράμματος της διασποράς ως προς τις κύριες συνιστώσες της ανάλυσης. Είναι εμφανές ότι οι πέντε πρώτες κύριες συνιστώσες περιέχουν αθροιστικά πάνω από 95% της συνολικής διασποράς. Γενικά γίνεται επιλογή των σημαντικότερων κυρίων συνιστωσών από το σύνολο των m PCs, εκείνων δηλαδή που περιέχουν το μεγαλύτερο ποσοστό πληροφορίας και μπορούν να περιγράψουν κατά το καλύτερο δυνατό τρόπο το σύνολο των δειγμάτων, με όσο το δυνατό λιγότερη απώλεια πληροφορίας. Ο αριθμός των κυρίων συνιστωσών που θα παραμείνουν, και ο οποίος ορίζει μετασχηματισμένο πίνακα δεδομένων μικρότερων διαστάσεων, εξαρτάται από τις ανάγκες της περαιτέρω ανάλυσης— τις απαιτήσεις σε ακρίβεια και την διαθεσιμότητα υπολογιστικού χρόνου.



Σχήμα 2-7 Διάγραμμα του % ποσοστού της εξηγήσιμης ανά κύρια συνιστώσα διασποράς

2.3 Μέθοδοι Παλινδρόμησης Πολυπαραμετρικών Δεδομένων

Η προσέγγιση συναρτησιακής σχέσης μεταξύ ανεξάρτητων και εξαρτημένων μεταβλητών αποτελεί έναν από τους σημαντικότερους στόχους στην ποσοτική χημική ανάλυση, είτε για πρόβλεψη ιδιοτήτων από άλλες εύκολα μετρούμενες, είτε για βελτιστοποίηση των παραμέτρων μιας αναλυτικής διαδικασίας ή των ποιοτικών χαρακτηριστικών βιομηχανικού προϊόντος. Οι τεχνικές που χρησιμοποιούνται αναφέρονται ως μέθοδοι **ανάλυσης πολυπαραμετρικής βαθμονόμησης και παλινδρόμησης** (multivariate calibration & regression analysis) και χρησιμοποιούνται για την ανάπτυξη κατάλληλου μαθηματικού μοντέλου συσχέτισης των ανεξάρτητων και εξαρτημένων μεταβλητών. Αποτελούν μεθόδους εποπτευόμενης ανάλυσης για την ανάπτυξη γραμμικών κυρίως μοντέλων με χρήση των μετρούμενων αποκρίσεων-στόχων των εξαρτημένων μεταβλητών ως προς τις οποίες εξετάζεται το πολυπαραμετρικό σύνολο δεδομένων.

Πολλαπλή Γραμμική Παλινδρόμηση

Η μέθοδος της **πολλαπλής γραμμικής παλινδρόμησης** (Multiple Linear Regression, MLR)) αποτελεί μία από τις απλούστερες και συνηθέστερα εφαρμοζόμενες πολυπαραμετρικές μεθόδους στατιστικής επεξεργασίας για ανάπτυξη γραμμικών μοντέλων πρόβλεψης ιδιοτήτων. Αποτελεί στην πραγματικότητα επέκταση –για περισσότερες της μίας ανεξάρτητες μεταβλητές– της μονοπαραμετρικής πολυωνυμικής παλινδρόμησης των ελαχίστων τετραγώνων. Για πολυπαραμετρικό πληθυσμό n δειγμάτων που εξετάζονται ως προς m ανεξάρτητες μεταβλητές X_j με $j=1,2,\dots,m$ και μία εξαρτημένη μεταβλητή Y , το γραμμικό μοντέλο της πολλαπλής γραμμικής παλινδρόμησης δίνεται από την σχέση:

$$\mathbf{y} = \mathbf{X} \cdot \boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (2.29)$$

όπου $\mathbf{y}$ το διάνυσμα της μεμονωμένης εξαρτημένης μεταβλητής Y και με στοιχεία τις μετρούμενες τιμές-στόχους.

$\boldsymbol{\beta}$ το διάνυσμα των συντελεστών των ανεξάρτητων μεταβλητών του πολυωνύμου προσέγγισης

και $\boldsymbol{\varepsilon}$ το διάνυσμα του υπολειπόμενου σφάλματος του πολυωνύμου προσέγγισης

Η εξαρτημένη μεταβλητή αποτελεί γραμμικό συνδυασμό των m ανεξάρτητων μεταβλητών με συντελεστές πολυωνύμου που προσδιορίζονται από την σχέση:

$$\boldsymbol{\beta} = (\mathbf{X}^T \cdot \mathbf{X})^{-1} \cdot \mathbf{X}^T \cdot \mathbf{y} \quad (2.30)$$

Παλινδρόμηση Κυρίων Συνιστωσών

Η ανάλυση πολλαπλής γραμμικής παλινδρόμησης περιλαμβάνει συχνά μεγάλο αριθμό ανεξάρτητων μεταβλητών που εμφανίζουν ισχυρή συσχέτιση μεταξύ τους, γεγονός που εισάγει πλεονασμό στο μοντέλο αφού μπορεί να χρησιμοποιούνται περισσότερες μεταβλητές από τις απαραίτητες για ικανοποιητική πρόβλεψη της εξαρτημένης μεταβλητής. Η χρήση τεχνικών παλινδρόμησης που εφαρμόζουν γραμμικό μετασχηματισμό του πίνακα δεδομένων με παράλληλη μείωση του αριθμού των παραμέτρων μπορεί να οδηγήσει σε αποκοπή του μεγαλύτερου ποσοστού του περιεχόμενου θορύβου και την ανάπτυξη ακριβέστερων μοντέλων πρόβλεψης. Η φιλοσοφία των μεθόδων αυτών στηρίζεται στην ανάλυση κυρίων συνιστωσών με πιο δημοφιλή ανάμεσα τους την **παλινδρόμηση κυρίων συνιστωσών** (Principal Component Regression, PCR).

Η χημειομετρική αυτή τεχνική συνίσταται στον προσδιορισμό των κυρίων συνιστωσών (PCs) του αρχικού συνόλου των ανεξάρτητων μεταβλητών, οι οποίες στη συνέχεια χρησιμοποιούνται ως δεδομένα εισαγωγής σε ένα αλγόριθμο πολλαπλής παλινδρόμησης για την πρόβλεψη της εξαρτημένης μεταβλητής. Αφορά μία επεξεργασία δύο σταδίων: ανάλυση κυρίων συνιστωσών του πίνακα $\mathbf{X}$ των ανεξάρτητων μεταβλητών και κατόπιν εφαρμογή πολλαπλής γραμμικής παλινδρόμησης χρησιμοποιώντας τις κύριες συνιστώσες που προέκυψαν από τα αρχικά δεδομένα ως ανεξάρτητες μεταβλητές για την πρόβλεψη της εξαρτημένης μεταβλητής.

$$\mathbf{X} = \mathbf{S} \cdot \mathbf{L} + \mathbf{E} \quad (2.31)$$

$$\mathbf{y} = \mathbf{S} \cdot \boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (2.32)$$

Η εξαρτημένη μεταβλητή προσεγγίζεται με πολυωνυμική συνάρτηση πρώτου βαθμού ως προς τις $p < m$ κύριες συνιστώσες με συντελεστές που υπολογίζονται από τον πίνακα των scores factors με βάση την σχέση:

$$\boldsymbol{\beta} = (\mathbf{S}^T \cdot \mathbf{S})^{-1} \cdot \mathbf{S}^T \cdot \mathbf{y} \quad (2.33)$$

Παλινδρόμηση Μερικών Ελαχίστων Τετραγώνων

Μια σχετικά νέα τεχνική με παρόμοια φιλοσοφία με αυτή της PCR είναι η **παλινδρόμηση μερικών ελαχίστων τετραγώνων** (Partial Least Squares, PLS). Η μέθοδος περιλαμβάνει συμπίκνωση δεδομένων και αποσύνθεση του πίνακα, όχι μόνο των ανεξάρτητων, $\mathbf{X}$, αλλά και των εξαρτημένων μεταβλητών, $\mathbf{Y}$, για την ανάπτυξη γραμμικού μοντέλου συσχέτισης μεταξύ τους της μορφής:

$$\mathbf{Y} = \mathbf{X} \cdot \mathbf{B} + \mathbf{E} \quad (2.34)$$

Οι συντελεστές του πίνακα $\mathbf{B}$ υπολογίζονται έτσι ώστε να μεγιστοποιούν κατά το δυνατόν την συνδιακύμανση μεταξύ των πινάκων των εξαρτημένων και των ανεξαρτητών μεταβλητών, και

επομένως την χρήσιμη πληροφορία συσχέτισης των εξαρτημένων και ανεξαρτήτων μεταβλητών.

Για τον σκοπό αυτό εφαρμόζεται ανάλυση παρόμοια με αυτή των κυρίων συνιστωσών, για προσδιορισμό των PLS-συνιστωσών, τόσο στον πίνακα $\mathbf{X}$, όσο και στον πίνακα $\mathbf{Y}$ των εξαρτημένων μεταβλητών.

Αρχικά παράγεται ένας πίνακας συντελεστών βάρους $\mathbf{W}$ τέτοιος, ώστε για τον πίνακα των ανεξαρτήτων μεταβλητών $\mathbf{X}$ να ισχύει:

$$\mathbf{S}_x = \mathbf{W} \cdot \mathbf{X} \quad (2.35)$$

όπου $\mathbf{S}_x$ ο πίνακας των scores factors του $\mathbf{X}$

Έπειτα ο πίνακας $\mathbf{Y}$ των εξαρτημένων μεταβλητών γράφεται ως γινόμενο των πινάκων:

$$\mathbf{Y} = \mathbf{S}_x \cdot \mathbf{L}_y \quad (2.36)$$

όπου $\mathbf{L}_y$ ο πίνακας των loadings factors του $\mathbf{Y}$

Ο πίνακας των εξαρτημένων μεταβλητών γράφεται:

$$\mathbf{Y} = \mathbf{X} \cdot \mathbf{B} \quad (2.37)$$

όπου $\mathbf{B} = \mathbf{W} \cdot \mathbf{L}_y$ οι συντελεστές του γραμμικού μοντέλου παλινδρόμησης

Η PLS είναι πιο γρήγορη από την PCA στην επεξεργασία και την αποσύνθεση των πινάκων, ενώ τα μοντέλα πρόβλεψης με PLS παρουσιάζονται συγκρίσιμα με εκείνα της PCR.

3 Νευρωνικά Δίκτυα

Τα νευρωνικά δίκτυα αποτελούν ένα ιδιαίτερο κεφάλαιο στην χημειομετρική επεξεργασία δεδομένων. Πρόκειται για μία σχετικά νέα μέθοδο προσέγγισης τέτοιων προβλημάτων, ραγδαία αναπτυσσόμενη τα τελευταία χρόνια, που καλείται να καλύψει τις αδυναμίες των άλλων στατιστικών μεθόδων. Επίσης θεωρούνται εναλλακτική άλλων μη γραμμικών στατιστικών μεθόδων πολυπαραμετρικής ανάλυσης δεδομένων, όπως η μη γραμμική παλινδρόμηση και η ανάλυση σμηνών. Επομένως τα νευρωνικά δίκτυα καλούνται να επιλύσουν προβλήματα ταξινόμησης ή πρόβλεψης, όπως είναι η αναγνώριση χαρακτήρων κειμένου, η επεξεργασία εικόνας και ήχου, και πεδία ανθρώπινης ειδίκευσης, όπως η ιατρική διάγνωση, η γεωφυσική ανίχνευση πετρελαίου, η πρόβλεψη οικονομικών δεικτών αγοράς κ.α..

Τέτοιου είδους προβλήματα εμπίπτουν στην περιοχή της κλασσικής **τεχνητής νοημοσύνης** (artificial intelligence, AI), όρος που χρησιμοποιείται για να υπογραμμίσει την ικανότητα ενός συστήματος να προσομοιώνει την ανθρώπινη εγκεφαλική επεξεργασία ερεθισμάτων και λήψη αποφάσεων. Ένα τέτοιο σύστημα, όπως και το αντίστοιχό του βιολογικό, αν και μειονεκτεί στην ταχύτητα εκτέλεσης πράξεων σε σχέση με τους συμβατικούς υπολογιστές, μπορεί να επεξεργαστεί γρήγορα πολύπλοκα δεδομένα κάνοντας χρήση **παράλληλου κατανεμημένου υπολογισμού**¹ (parallel distributed computing), που επιτρέπει λειτουργίες, όπως ανοχή στα σφάλματα και τον θόρυβο, συνειρμική αναγνώριση και ταυτόχρονη επεξεργασία μεγάλου όγκου πληροφορίας. Πρόκειται για μη γραμμικά συστήματα, που βρίσκουν εφαρμογές στην στατιστική μηχανική και τον αυτοματισμό, παρέχοντας εφαρμόσιμες λύσεις μεγάλης ακρίβειας σε ψηφιακά δεδομένα. Χαρακτηρίζονται επίσης και ως **συνδεσμικά** συστήματα (connectionist)², λόγω της αρχιτεκτονικής τους και του είδους της διαδικασίας –με ιδιαίτερη έμφαση στη σημασία των ενδονευρωνικών συνδέσεων –που εφαρμόζουν κατά την μοντελοποίηση του ανθρώπινου εγκεφάλου.

Ένα τεχνητό νευρωνικό δίκτυο (ΤΝΔ) είναι μια αλληλοσυνδεόμενη συνάθροιση από απλά στοιχεία επεξεργασίας, **μονάδες ή κόμβους**, η αρχή λειτουργίας των οποίων βασίζεται στους φυσικούς νευρώνες. Τα νευρικά κύτταρα του ανθρώπινου εγκεφάλου επικοινωνούν μέσω

¹ Παραλληλισμός είναι η ταυτόχρονη και συγκριτική εκτέλεση ενός συνόλου ακολουθιακών βημάτων και διαδικασιών. Κάθε μονάδα επεξεργασίας του συστήματος λειτουργεί ανεξάρτητα και ταυτόχρονα (παράλληλα με) τις άλλες, και η “γνώση” στο σύστημα κατανέμεται μεταξύ του συνόλου των ρυθμιστικών συνδέσεων των μονάδων, αντί να συγκεντρώνεται σε μερικές περιοχές μνήμης όπως σε ένα συμβατικό υπολογιστή.

² Η θεωρία του συνδετισμού εντάσσει τα νευρωνικά δίκτυα στο ευρύτερο πλαίσιο των μοντέλων που εμπνέονται από τη φυσιολογία της ανθρώπινης γνωστικής λειτουργίας.

ηλεκτρικών παλμικών σημάτων μικρή διάρκειας που ευαισθητοποιούν το κυτταρικό τους τοίχωμα, αφού πρώτα υποστούν ρυθμιστική επεξεργασία στους μεταξύ τους συνδετικούς κόμβους, τέτοια, ώστε να έχουν διεγερτική ή αποτρεπτική επίδραση στην παραγωγή επόμενου παλμού. Με όμοιο τρόπο, στους τεχνητούς νευρώνες, η ικανότητα προώθησης του δικτύου αποθηκεύεται στις μεταξύ των μονάδων δυνάμεις σύνδεσης, ή **βάρη**, που ρυθμίζονται με μια επαναληπτική διαδικασία προσαρμογής (αλγόριθμος εκμάθησης) σε ερεθίσματα από ένα σύνολο πρότυπων δειγμάτων, όπως ακριβώς ο ανθρώπινος εγκέφαλος “μαθαίνει” να αντιδρά σωστά σε εξωτερικά ερεθίσματα συλλέγοντας εμπειρία από προηγούμενα ανάλογα προβλήματα. Βασικές ιδιότητες επομένως ενός νευρωνικού δικτύου είναι *η καλή γνώση του χώρου του προβλήματος και η ικανότητα γενίκευσης*, δηλαδή η ικανότητά του να ταξινομεί ή να προβλέπει σωστά τόσο τα δείγματα για τα οποία έχει εκπαιδευτεί, όσο και κάθε νεοεισαγόμενο δείγμα.

Τα τεχνητά νευρωνικά δίκτυα που χρησιμοποιούνται στις συνήθεις εφαρμογές είναι αρκετά πολύπλοκα, επιβαρυνμένα υπολογιστικά συστήματα, που απαιτούν για την υλοποίησή τους εξελιγμένους ηλεκτρονικούς υπολογιστές (μεγάλη μνήμη και ταχύτητα, βελτιωμένα λογισμικά κλπ.), και επομένως η έρευνα και εξέλιξη τους συμβαδίζει με την ανάπτυξη της επιστήμης της πληροφορικής.

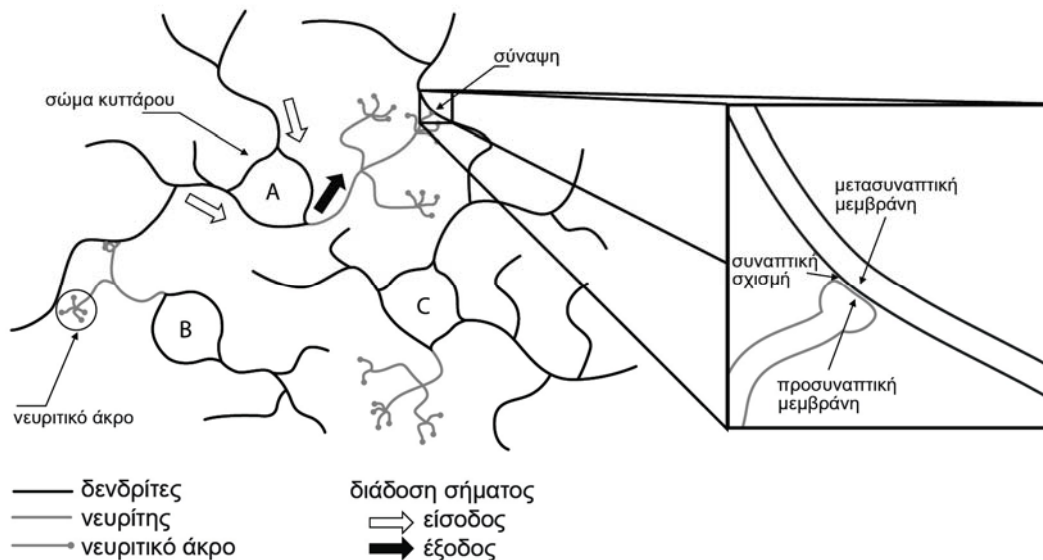
3.1 Φυσικοί και τεχνητοί νευρώνες

Οι δομικές μονάδες των τεχνητών νευρωνικών δικτύων είναι οι τεχνητοί νευρώνες, που διαμορφώνονται με στόχο να ενσωματώσουν την βασική δυνατότητα –επεξεργασίας και προώθησης πληροφοριών –των πραγματικών, βιολογικών νευρώνων. Παρόλο που οι φυσικοί νευρώνες, εκτός από πολυσύνθετοι, παρουσιάζουν μεταξύ τους μεγάλες διαφορές στις λεπτομέρειες δομής και λειτουργίας τους, στην πλειοψηφία τους εμφανίζουν κάποιες κοινές τυπικές ιδιότητες που, με βάση τη συνήθη υπόθεση του **συνδετισμού** (connectionism), είναι εκείνες που παρέχουν και τον κύριο όγκο της υπολογιστικής τους δυνατότητας και προσπαθούν να προσομοιώσουν τα τεχνητά αντίστοιχά τους.

Φυσικοί νευρώνες

Ένας τυπικός νευρώνας αποτελείται (Σχ. 3-1) από το κυρίως τμήμα ή **σώμα** (soma) που περιέχει τα συνήθη σε όλα τα κύτταρα οργανίδια (πυρήνας, μιτοχόνδρια, σώμα Golgi, κλπ.). Η κύρια μορφολογική του διαφορά είναι η πληθώρα ινών που καταλήγουν στο σώμα του κυττάρου. Μία από αυτές, ο **νευρίτης** (axon), είναι αρμόδια για τη διαβίβαση των σημάτων προς τους άλλους νευρώνες και επομένως μπορεί να θεωρηθεί ως η έξοδος του νευρώνα. Όλες

οι άλλες, οι **δενδρίτες** (dendrites), μεταφέρουν σήματα από άλλους νευρώνες στο σώμα του κύτταρου, οπότε λειτουργούν ως νευρικές είσοδοι. Έχουν συνήθως διακλαδιζόμενη δομή σχηματίζοντας δενδριτικούς **άξονες** (arbors). Ο νευρίτης απομακρυνόμενος από το σώμα του νευρώνα χωρίζεται σε **κλάδους** (collaterals) που επιτρέπουν την επικοινωνία με πολλούς άλλους νευρώνες. Για έναν ορισμένο νευρώνα A, οι νευρώνες, όπως ο B, που παρέχουν την είσοδό του ονομάζονται **προσαγωγοί** (afferent), ενώ τα άλλα κύτταρα (C) που λαμβάνουν σήματα από τον A, αναφέρονται ως **απαγωγοί** (efferent). Οι νευρίτες των προσαγωγών **εννευρώνουν** (innervate) τον νευρώνα συνδέοντάς τον με τους δενδρίτες σε σημεία συμβολής που καλούνται **συνάψεις** (synapses). Στην σύναψη (Σχ. 3-1)), την **νευριτική απόληξη** (axon terminal) βρίσκεται σχεδόν σε επαφή¹ με ένα μικρό τμήμα της δενδριτικής επιφάνειας –την **μετασυναπτική μεμβράνη** (postsynaptic membrane).



Σχήμα 3-1 Σχηματική δομή φυσικών νευρώνων και απεικόνιση της σύναψης στην νευριτική απόληξη

Αυτό που διαφοροποιεί τους νευρώνες από τα άλλα κύτταρα, επιτρέποντάς τους να λειτουργήσουν ως μηχανισμός προώθησης σήματος, οφείλεται κατά ένα μεγάλο μέρος στις ιδιότητες του επιφανειακού καλύμματος του νευρώνα, τη **μεμβράνη** (membrane), που υποστηρίζει μια ευρεία ποικιλία ηλεκτροχημικών διεργασιών. Σε ισορροπία, η νευρική μεμβράνη διατηρεί εκατέρωθεν της σταθερή διαφορά δυναμικού (70mV), με το εσωτερικό της αρνητικά πολωμένο. Για να υποστηρίξει όμως η νευρική μεμβράνη ηλεκτρικά σήματα πρέπει η

¹ Υπάρχει συναπτική σχισμή (synaptic cleft), μεταξύ της προσυναπτικής (presynaptic) μεμβράνης της νευριτικής απόληξης και της αντίστοιχης μετασυναπτικής του, της τάξεως των 20nm (2×10^{-8} m).

κατάσταση πόλωσης της ή **δυναμικό μεμβράνης** (membrane potential) να είναι δυναμικά μεταβαλλόμενη. Ένα σήμα ή **δυναμικό δράσης** (action potential), που πρόκειται να διαβιβαστεί διαμέσου ενός κεντρικού νευρώνα, παράγεται σ' έναν από τους προσαγωγούς νευρώνες, έχοντας πάντα την ίδια χαρακτηριστική μορφή ακίδας παλμού, διαδίδεται κατά μήκος του νευρίτη και προσκρούει σε μία νευρωνική είσοδο-σύναψη, φθάνοντας σε ένα από τα νευριτικά του άκρα (προσυναπτικό τμήμα της σύναψης). Στην νευριτική απόληξη, σε έναν μεγάλο αριθμό μικρών **κυστιδίων** (vesicles), περιέχεται μια χημική ουσία, ο **νευροδιαβιβαστής** (neurotransmitter).

Με την άφιξη του δυναμικού δράσης τα κυστίδια μεταναστεύουν στην προσυναπτική μεμβράνη και απελευθερώνουν το νευροδιαβιβαστή στην συναπτική σχισμή, που δεσμεύεται κατόπιν χημικά από ακροδέκτες στην μετασυναπτική μεμβράνη, εκκινώντας μια ηλεκτροχημική διαδικασία που μεταβάλλει την κατάσταση πόλωσης της μεμβράνης τοπικά στη σύναψη. Η μεταβολή αυτή χαρακτηρίζεται ως **μετασυναπτικό δυναμικό** (Postsynaptic Potential, PSP) και μπορεί είτε να **αποπολώσει** (depolarize) τη μεμβράνη μειώνοντας το αρνητικό δυναμικό βάσης της, είτε να την **υπερπολώσει** (hyperpolarize) σε ένα ακόμα μεγαλύτερο αρνητικό δυναμικό. Τα θετικά PSPs είναι ενισχυτικά (Excitatory Postsynaptic Potentials, EPSPs), υποβοηθώντας την νευρική παραγωγή σημάτων, ενώ εκείνα που υπερπολώνουν την μεμβράνη είναι αποτρεπτικά (Inhibitory Postsynaptic Potentials, IPSPs). Ενώ τα δυναμικά δράσης έχουν όλα την ίδια μορφή σημάτων και ίδια μέγιστη τιμή, επιδεικνύοντας έναν χαρακτήρα “όλα ή τίποτα”, τα PSPs μπορούν να πάρουν μία συνεχή σειρά τιμών ανάλογα με την αποδοτικότητα της σύναψης. Το προκύπτον PSP διασπείρεται έξω από τη σύναψη, διαβιβάζεται κατά μήκος του αντίστοιχου δενδρίτη προς το σώμα του κυττάρου και καταλήγει στον **εκφυτικό κώνο** (axon hillock) –το εξειδικευμένο τμήμα του κυτταρικού σώματος του νευρώνα από την οποία εκφύεται ο νευρίτης.

Ταυτόχρονα με αυτό, συμβαίνουν χιλιάδες άλλα συναπτικά γεγονότα που έχουν ως αποτέλεσμα την άφιξη ενός πλήθους PSPs στο εκφυτικό κώνο, όπου και αθροίζονται για να παραχθεί το προκύπτον δυναμικό μεμβράνης. Αν το αυτό υπερβεί έναν ορισμένο κατώτατο όριο, το λεγόμενο **κατώφλιο** (threshold), παράγεται ή όχι από το σώμα του κυττάρου ένα δυναμικό δράσης, που κατόπιν διαδίδεται στον νευρίτη, διαμέσου κάθε κλάδου, φθάνοντας τελικά στην απόληξη του νευρίτη για να διαβιβαστεί με την ίδια διαδικασία στον επόμενο νευρώνα.

Τεχνητοί νευρώνες –ο κόμβος TLU

Η απλούστερη, και ιστορικά η πρώτη (McCulloch & Pitts (1943)), μονάδα επεξεργασίας τεχνητού νευρωνικού δικτύου είναι η **λογική μονάδα κατωφλίου** ή **λήψης αποφάσεων** (threshold logic unit, TLU). Αυτή η μονάδα νευρώνα υιοθετεί την δομή φυσικού νευρώνα ενσωματώνοντας μόνο τα βασικά χαρακτηριστικά προώθησης σήματος, και, υποθέτοντας την

εισαγωγή ενός συνόλου σημάτων, λαμβάνει απόφαση αποδοχή τους ή μη προκειμένου να παράγει έξοδο-νέο σήμα. Ο “όλα ή τίποτα” χαρακτήρας του δυναμικού δράσης, ως σήμα εισόδου ή εξόδου, αποδίδεται με ένα δυαδικό (Boolean) μέγεθος που παίρνει συμβατικά τις τιμές 0 και 1, ή εναλλακτικά $-1, 1$, ενώ η ρυθμιστική επίδραση της σύναψης αντιστοιχίζεται με το γινόμενο του σήματος εισόδου επί μία τιμή βάρους, θετική ή αρνητική, ώστε να υλοποιούνται οι ενισχυτικές ή αποτρεπτικές δράσεις της αντίστοιχα. Επομένως για m δυαδικά σήματα εισόδου $x_1, x_2, \dots, x_m$ σ’ έναν κόμβο ορίζονται m βάρη $w_1, w_2, \dots, w_m$, οπότε προκύπτουν τα γινόμενα $w_1x_1, w_2x_2, \dots, w_mx_m$ (ανάλογα των PSPs), που εισέρχονται στον τεχνητό νευρώνα TLU όπου αθροίζονται· διαδικασία που προσομοιώνει αυτή που πραγματοποιείται στον εκφυτικό κώνο. Το άθροισμα τους παράγει την **ενεργοποίηση** (activation) α του κόμβου (αντιστοιχεί στο δυναμικό μεμβράνης του εκφυτικού κώνου) και δίνεται από την σχέση:

$$\alpha = \sum_{j=1}^m w_j x_j = w_1 x_1 + w_2 x_2 + \dots + w_m x_m \quad (3.1)$$

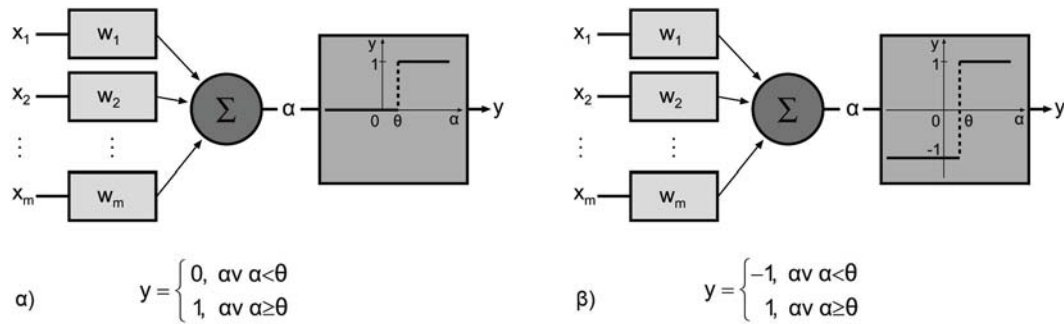
ή με χρήση διανυσμάτων για την είσοδο και τα βάρη:

$$\alpha = \mathbf{w}^T \cdot \mathbf{x} \quad (3.2)$$

όπου $\mathbf{w} = (w_1, w_2, \dots, w_m)$ και $\mathbf{x} = (x_1, x_2, \dots, x_m)$. Προκειμένου να παραχθεί το δυναμικό δράσης του νευρώνα απαιτείται η εφαρμογή ενός τελεστή απόφασης για την τιμή που θα λάβει το δυαδικό σήμα εξόδου, που υλοποιείται με την λεγόμενη **συνάρτηση μεταφοράς** ή **ενεργοποίησης** (transfer ή activation function). Η απλούστερη μορφή που μπορεί να χρησιμοποιηθεί είναι η **συνάρτηση βήματος** (step function) ή αλλιώς **αυστηρά περιοριστική** (hard-limiter), η οποία καθορίζει ότι αν η ενεργοποίηση α υπερβαίνει ένα κατώφλιο-όριο απόφασης θ τότε η έξοδος y του κόμβου λαμβάνει την τιμή 1, ενώ, αν είναι μικρότερη, την τιμή 0 (ή -1 αντίστοιχα). Ο τύπος αυτής της συνάρτησης μεταφοράς δίνεται από τη σχέση:

$$y = \begin{cases} 0, & \text{αν } \alpha < \theta \\ 1, & \text{αν } \alpha \geq \theta \end{cases} \quad \text{ή αντίστοιχα} \quad y = \begin{cases} -1, & \text{αν } \alpha < \theta \\ 1, & \text{αν } \alpha \geq \theta \end{cases} \quad (3.3)$$

Το σχηματισμένο διάγραμμα της συνολικής δομής και λειτουργίας ενός κόμβου TLU παρουσιάζεται στο Σχ. 3-2.



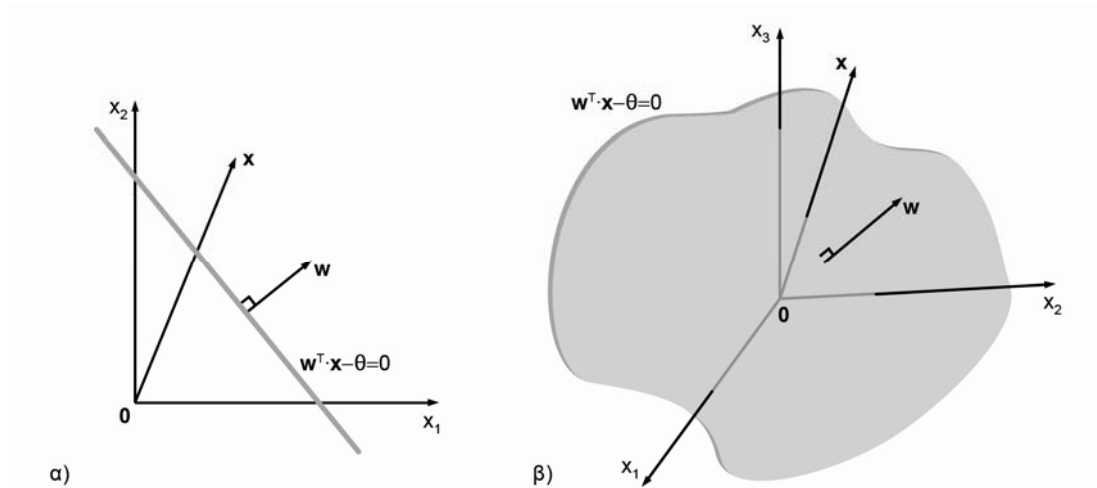
Σχήμα 3-2 Νευρώνας TLU που χρησιμοποιεί δυαδικές τιμές εισόδου -εξόδου α) 0, 1 και β) -1, 1.

Ο κόμβος TLU παρουσιάζει σταθερή συμπεριφορά στην υποβάθμιση των σημάτων εισόδου και την αστοχία του ηλεκτρομηχανολογικού εξοπλισμού υλοποίησής του. Λόγω της μη γραμμικής εξάρτησης της εξόδου y από την είσοδο $\mathbf{x}$, οι τυχόν ακούσιες μικρές μεταβολές στις παραμέτρους x_i του δείγματος ή τα w_i του συστήματος, αλλάζουν μεν την ενεργοποίηση α , εφ' όσον όμως αυτή δεν διασχίζει το όριο απόφασης θ , δεν επηρεάζουν την απόκριση της μονάδας. Αυτή η συμπεριφορά είναι χαρακτηριστική των *μη γραμμικών* συστημάτων, που δεν υπακούν στον περιορισμό αναλογίας εισόδου-εξόδου των γραμμικών συστημάτων, έτσι ώστε το μέγεθος της μεταβολής στην έξοδο να μην απεικονίζει απαραίτητα αυτό της εισόδου. Η *αντίσταση στον θόρυβο και τα συστηματικά σφάλματα* είναι ιδιότητα που διατηρούν και οι πιο εξελιγμένοι νευρώνες σε όλες τις αρχιτεκτονικές νευρωνικών δικτύων και η οποία τους καθιστά ιδιαίτερα ελκυστικούς στις εφαρμογές πολυπαραμετρικής ανάλυσης δεδομένων.

Η έννοια του πολυδιάστατου χώρου εισάγεται, ήδη από το απλοποιημένο αυτό μοντέλο, ταυτόχρονα με την ικανότητα γραμμικού διαχωρισμού που χαρακτηρίζει τους τεχνητούς νευρώνες. Η μονάδα TLU ταξινομεί τα m -παραμετρικά δείγματα εισόδου που ορίζουν ένα χώρο m διαστάσεων σε δύο τάξεις ανάλογα με τη δυαδική απόκρισή της (0 ή 1) σε κάθε ένα απ' αυτά. Οι δύο τάξεις αντιστοιχούν σε περιοχές του δειγματικού χώρου που ορίζονται από *υπερεπίπεδο* –διαστατικό ισοδύναμο ευθείας (σε δύο διαστάσεις, 2D) ή *επιπέδου* (σε τρεις διαστάσεις, 3D). Η εξίσωση αυτού του υπερεπιπέδου απόφασης δίνεται από την σχέση:

$$\mathbf{w}^T \cdot \mathbf{x} - \theta = 0 \quad (3.4)$$

που για διδιάστατο δείγμα εισόδου (x_1, x_2) αντιστοιχεί σε εξίσωση ευθείας. Το διάνυσμα βαρών είναι πάντα κάθετο στο υπερεπίπεδο απόφασης, καθορίζοντας τον προσανατολισμό του, ενώ η τιμή του κατωφλίου προσδιορίζει την θέση του στον χώρο (για $\theta = 0$ το υπερεπίπεδο απόφασης διέρχεται από την αρχή των m αξόνων του χώρου). Στο Σχ. 3-3 παρουσιάζονται τα όρια απόφασης, ευθεία και επίπεδο, ενός κόμβου TLU δύο και τριών εισόδων αντίστοιχα. Οι τρεις διαστάσεις, δηλαδή τρεις είσοδοι, αποτελούν και το όριο γεωμετρικής απεικόνισης.



Σχήμα 3-3 Γεωμετρική απεικόνιση του ορίου απόφασης κόμβου TLU α) με δύο και β) τρεις εισόδους

Επειδή, τόσο τα βάρη, όσο και το κατώφλιο αποτελούν τα βασικά στοιχεία επεξεργασίας και προώθησης σήματος του νευρώνα, είναι δυνατό να ενσωματωθεί το τελευταίο στο διάνυσμα των βαρών της μονάδας. Από την εξίσωση (3.4) έχουμε:

$$\mathbf{w}^T \cdot \mathbf{x} + (-1)\theta = \mathbf{w}^T \cdot \mathbf{x} + 1(-\theta) = \mathbf{w}^T \cdot \mathbf{x} + b = 0 \quad (3.5)$$

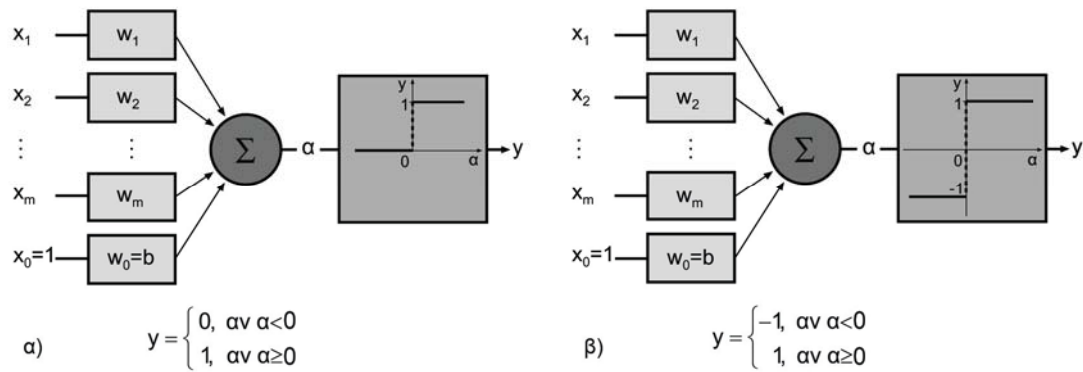
$$\mathbf{w}'^T \cdot \mathbf{x}' = 0 \quad (3.6)$$

όπου $\mathbf{w}' = (-\theta, w_1, w_2, \dots, w_m) = (w_0, w_1, w_2, \dots, w_m)$ και $\mathbf{x}' = (1, x_1, x_2, \dots, x_m)$ τα επαυξημένα διανύσματα των βαρών και των εισόδων. Η αρνητική τιμή του ορίου απόφασης $w_0 = b = -\theta$ αναφέρεται συνήθως ως **πόλωση** (bias ή offset) του νευρώνα.

Η τροποποίηση αυτή του νευρώνα έχει ως αποτέλεσμα την ενσωμάτωση του ορίου απόφασης στην συνάρτηση μεταφοράς και επιτρέπει χρήση συνάρτηση μεταφοράς συμμετρικής ως προς την αρχή των αξόνων και χειρισμό συνολικά όλων των παραμέτρων λειτουργίας της μονάδας. Για τον νευρώνα που εξετάστηκε η τροποποίηση αυτή φαίνεται σχηματικά στο Σχ. 3-4, ενώ η ενεργοποίηση και η έξοδος του θα είναι:

$$\alpha = \mathbf{w}'^T \cdot \mathbf{x}' = \sum_{j=0}^m w_j x_j = w_0 x_0 + w_1 x_1 + w_2 x_2 + \dots + w_m x_m \quad (3.7)$$

$$y = \begin{cases} 0, & \text{αν } \alpha < 0 \\ 1, & \text{αν } \alpha \geq 0 \end{cases} \quad \text{ή αντίστοιχα} \quad y = \begin{cases} -1, & \text{αν } \alpha < 0 \\ 1, & \text{αν } \alpha \geq 0 \end{cases} \quad (3.8)$$



Σχήμα 3-4 Νευρώνας με ενσωματωμένο το όριο απόφασης.

Τέλος, η απλοποίηση, που υιοθετήθηκε με τον περιορισμό σε δυαδικές τιμές του σήματος εισόδου και εξόδου, μπορεί να διευρυνθεί και για πληροφορία συνεχών τιμών. Εξάλλου, στην πραγματικότητα, οι φυσικοί νευρώνες κωδικοποιούν τις τιμές σημάτων τους σε συνεχείς συναρτήσεις έναυσης δυναμικού δράσης, που αντιπροσωπεύουν την συχνότητα του σήματος, αντί να χρησιμοποιούν την παρουσία ή μη μεμονωμένων παλμών. Για την είσοδο του νευρώνα αρκεί απλά να επεκταθεί το εύρος τιμών των x_i σε συνεχή σειρά τιμών στο διάστημα $(-1, 1)$, αντί των διακριτών $-1, 1$. Όμως, η χρήση της συνάρτησης βήματος καθιστά δυαδική την έξοδο, αποκλείοντας την δυνατότητα παραγωγής συνεχούς καμπύλης σημάτων για τα δίκτυα των TLU. Αυτό μπορεί να υπερκεραστεί θέτοντας ως συνάρτηση μεταφοράς της μονάδας μία συνεχή, άνω και κάτω φραγμένη, συμμετρική ως προς την αρχή των αξόνων συνάρτηση $g(\alpha)$, όπως η **εκθετική σιγμοειδής** (logistic sigmoid) (Σχ. 3-5), έτσι ώστε η έξοδος y να εξαρτάται ομαλά από την ενεργοποίηση α και επομένως από το διάνυσμα $\mathbf{x}$ που περιγράφει το δείγμα εισόδου.

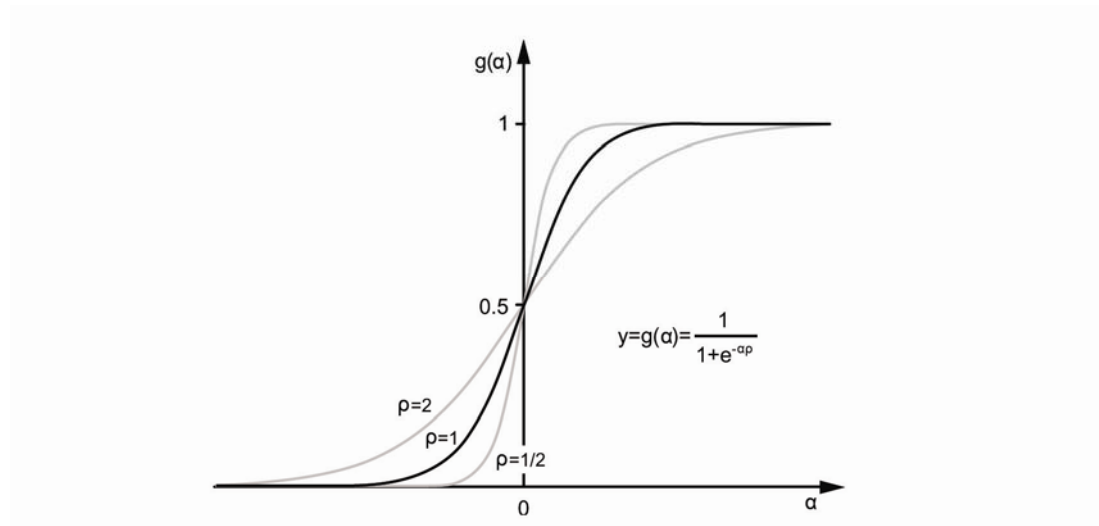
$$y = g(\alpha) = g(x_1, x_2, \dots, x_m) \quad (3.9)$$

Με την εφαρμογή συνεχούς συνάρτησης για την παραγωγή της εξόδου γίνεται εμφανής η δυνατότητα των δικτύων να χρησιμοποιηθούν, εκτός από προβλήματα ταξινόμησης δειγμάτων, και για την πρόβλεψη της εξάρτησης μιας ιδιότητας (έξοδος y) από άλλες παραμέτρους (είσοδος $\mathbf{x}$) του συνόλου δειγμάτων, διαμορφώνοντας μία συναρτησιακή σχέση μεταξύ αυτών των μεγεθών, που αποτελεί ταυτόχρονα και την μαθηματική περιγραφή λειτουργίας του νευρώνα:

$$y = g\left(\sum_{j=0}^m w_j x_j\right) \quad \text{ή} \quad y = g\left(\sum_{j=1}^m w_j x_j + b\right) \quad (3.10)$$

ή με χρήση διανυσμάτων για την είσοδο και τα βάρη:

$$y = g(\mathbf{w}^T \cdot \mathbf{x}') \quad \text{ή} \quad y = g(\mathbf{w}^T \cdot \mathbf{x} + b) \quad (3.11)$$



Σχήμα 3-5 Η εκθετική σιγμοειδής για διάφορες τιμές της παραμέτρου ρ .

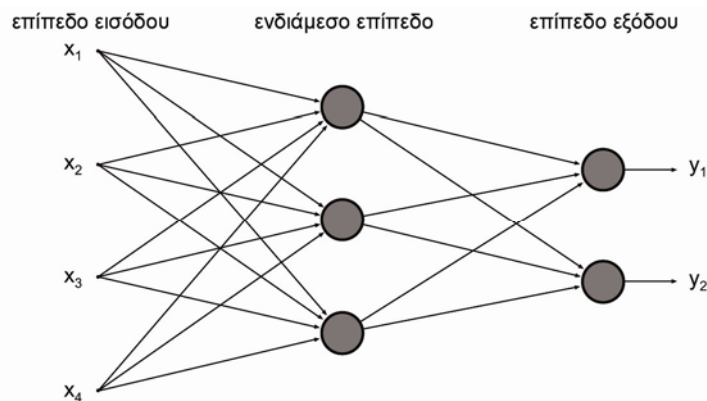
3.2 Τεχνητά νευρωνικά δίκτυα

Αρχιτεκτονική νευρωνικών δικτύων

Ένα τεχνητό νευρωνικό δίκτυο αποτελείται από ένα σύνολο μονάδων επεξεργασίας (τεχνητοί νευρώνες) οργανωμένων συνήθως σε ένα ή περισσότερα **επίπεδα** (layers). Κάθε μεμονωμένος νευρώνας (κόμβος) λειτουργεί ανεξάρτητα ως στοιχείο επεξεργασίας των εισερχόμενων σ' αυτόν τιμών εφαρμόζοντας βάρη στάθμισής τους για την διαμόρφωση της ενεργοποίησης και παράγοντας, μέσω της συνάρτησης μεταφοράς, μοναδική έξοδο. Λειτουργεί έτσι ως ενδιάμεσο στοιχείο στην προώθηση σήματος του συνολικού συστήματος, τόσο ως απαγωγός για τους προηγούμενους κόμβους λαμβάνοντας ως εισόδους τις εξόδους αυτών, όσο και ως προσαγωγός για τους επόμενους παρέχοντας την έξοδο του ως μία εκ των εισόδων τους. Η οργάνωση επομένως του συνολικού δικτύου καθορίζεται, εκτός από το πλήθος των μονάδων επεξεργασίας, από τις υπάρχουσες συνδέσεις μεταξύ τους.

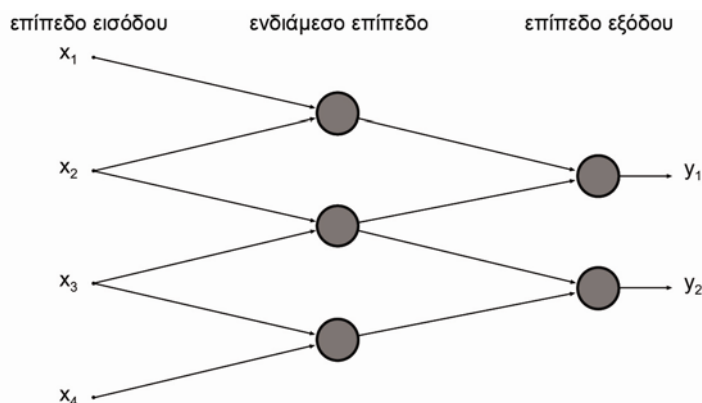
Η απλούστερη απεικόνιση της λειτουργίας αυτής είναι η οργάνωση των κόμβων σε συνεχή επίπεδα, όπου για ένα δίκτυο πλήρους διασύνδεσης, όλοι οι κόμβοι ενός επιπέδου λαμβάνουν εισόδους από τους κόμβους του προηγούμενου επιπέδου και αποστέλλουν τις εξόδους τους στους κόμβους του επόμενου επιπέδου. Τα επίπεδα του δικτύου διακρίνονται στο **επίπεδο εισόδου** (input layer), τα **ενδιάμεσα επίπεδα** (hidden layers) και το **επίπεδο εξόδου** (output layer) και η συνολική αρχιτεκτονική χαρακτηρίζεται από το σύνολο των νευρώνων ανά επίπεδο. Ένα δίκτυο αρχιτεκτονικής $4 \times 3 \times 2$ αναπτύσσεται σε ένα επίπεδο τεσσάρων εισόδων, ένα

ενδιάμεσο επίπεδο τριών κόμβων και επίπεδο εξόδου δύο κόμβων (Σχ. 3-6). Στα επόμενα σχήματα, για λόγους απλοποίησης της απεικόνισης τα βάρη έχουν ενσωματωθεί στις συνδέσεις μεταξύ των κόμβων.



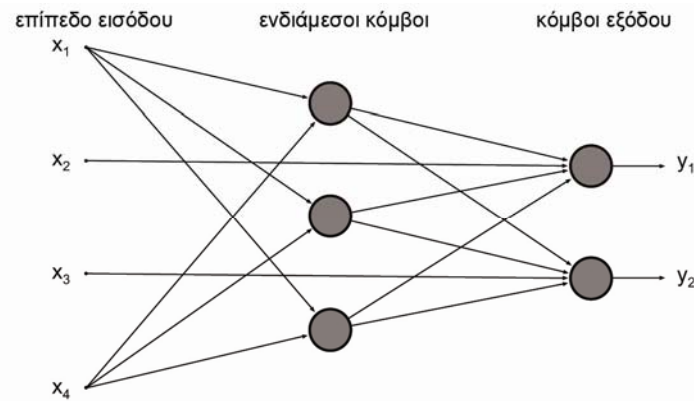
Σχήμα 3-6 Αρχιτεκτονική 4×3×2 δικτύου ευθείας προώθησης πλήρους διασύνδεσης

Βέβαια σε πολλές περιπτώσεις δεν είναι απαραίτητη η πλήρης διασύνδεση μεταξύ των κόμβων (Σχ. 3-7), ενώ επιπλέον κάποιες από τις εισόδους μπορεί να συνδέονται απ' ευθείας με την έξοδο¹, σε μη διαβαθμισμένη σε επίπεδα οργάνωση του δικτύου (Σχ. 3-8).



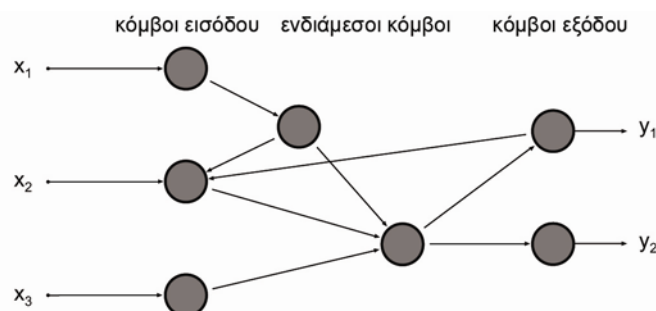
Σχήμα 3-7 Αρχιτεκτονική 4×3×2 δικτύου ευθείας προώθησης μη πλήρους διασύνδεσης

¹ Στην περίπτωση προσέγγισης συναρτησιακής σχέσης μεταξύ των ανεξάρτητων και των εξαρτημένων μεταβλητών, κάτι τέτοιο εφαρμόζεται για αναγνώριση από το σύστημα της γραμμικής σχέσης κάποιων εκ των ανεξάρτητων μεταβλητών με την έξοδο (με χρήση γραμμικών κόμβων εξόδου) παράλληλα με την μη γραμμική εξάρτηση της τελευταίας από τις υπόλοιπες.



Σχήμα 3-8 Αρχιτεκτονική δικτύου ευθείας προώθησης με μη διαβαθμισμένη σε επίπεδα οργάνωση

Όλες οι προηγούμενες αρχιτεκτονικές δομές περιγράφουν **δίκτυα ευθείας προώθησης** (feedforward networks) που επιτρέπουν στα σήματα να διαδίδονται μόνο προς μία κατεύθυνση, από την είσοδο προς την έξοδο. Πρόκειται για δίκτυα που χρησιμοποιούνται κυρίως για ταξινόμηση ή προσέγγιση συνάρτησης, στα οποία τα σήματα υφίστανται επεξεργασία από μία πεπερασμένη σειρά κόμβων πριν συμβάλλουν σε μία τελική απόκριση, χωρίς να παρεμβάλλεται κανένα σύστημα ανατροφοδότησης πληροφοριών από την έξοδο στην είσοδο. Αντίθετα στα **δίκτυα ανατροφοδότησης ή επαναλαμβανόμενα** (feedback ή recurrent networks) η διάδοση γίνεται εντός κλειστών βρόγχων, με τις εξόδους κάποιων εκ των κόμβων να επανεισάγονται σε προηγούμενους κόμβους. (Σχ. 3-9). Πρόκειται για δυναμικά συστήματα, τα οποία για δεδομένες εισόδους παράγουν μεταβαλλόμενες με τον χρόνο αποκρίσεις και χρησιμοποιούνται για εφαρμογή διαφορικών εξισώσεων.



Σχήμα 3-9 Τυπική αρχιτεκτονική δικτύου ανατροφοδότησης

Λειτουργία νευρωνικού δικτύου ευθείας προώθησης

Η λειτουργία ενός νευρωνικού δικτύου ευθείας προώθησης πλήρους διασύνδεσης που δέχεται m εισόδους και αναπτύσσεται σε δύο ενδιάμεσα επίπεδα και το επίπεδο εξόδου με k , p και s μονάδες επεξεργασίας αντίστοιχα, δηλαδή με αρχιτεκτονική $m \times k \times p \times s$, παρουσιάζεται αναλυτικά

στο Σχ. 3-10 και, κατ' αντιστοιχία με αυτή του μεμονωμένου νευρώνα (εξισώσεις (3.10) και (3.11)), περιγράφεται μαθηματικά από την σχέση:

$$y_t^3 = g^3 \left(\sum_{r=1}^p w_{tr}^{3,2} g^2 \left(\sum_{\ell=1}^k w_{r\ell}^{2,1} g^1 \left(\sum_{j=1}^m w_{\ell j}^{1,1} x_j + b_{\ell}^1 \right) + b_r^2 \right) + b_t^3 \right) \quad (3.12)$$

που για δίκτυο με u επίπεδα, το v -οστό ενδιάμεσο επίπεδο ($v=1, \dots, u$) του οποίου περιέχει e κόμβους, είναι:

x_j η j -οστή είσοδος ($j=1, \dots, m$) του συστήματος,

$w_{fd}^{v,v-1}$ το βάρος ζύγισης της εξόδου του d -οστού κόμβου ($d=1, \dots, c$) του $(v-1)$ -οστού ενδιάμεσου επιπέδου με c κόμβους που εισάγεται στον f -οστό κόμβο ($f=1, \dots, e$) του v -οστού ενδιάμεσου επιπέδου

και ειδικά για το επίπεδο εισόδου, $w_{\ell j}^{1,1}$ το βάρος ζύγισης της j -οστής εισόδου x_j του συστήματος που εισάγεται στον ℓ -οστό κόμβο ($\ell=1, \dots, k$) του 1ου ενδιάμεσου επιπέδου,

b_f^v η πόλωση του f -οστού κόμβου κόμβο του v -οστού ενδιάμεσου επιπέδου,

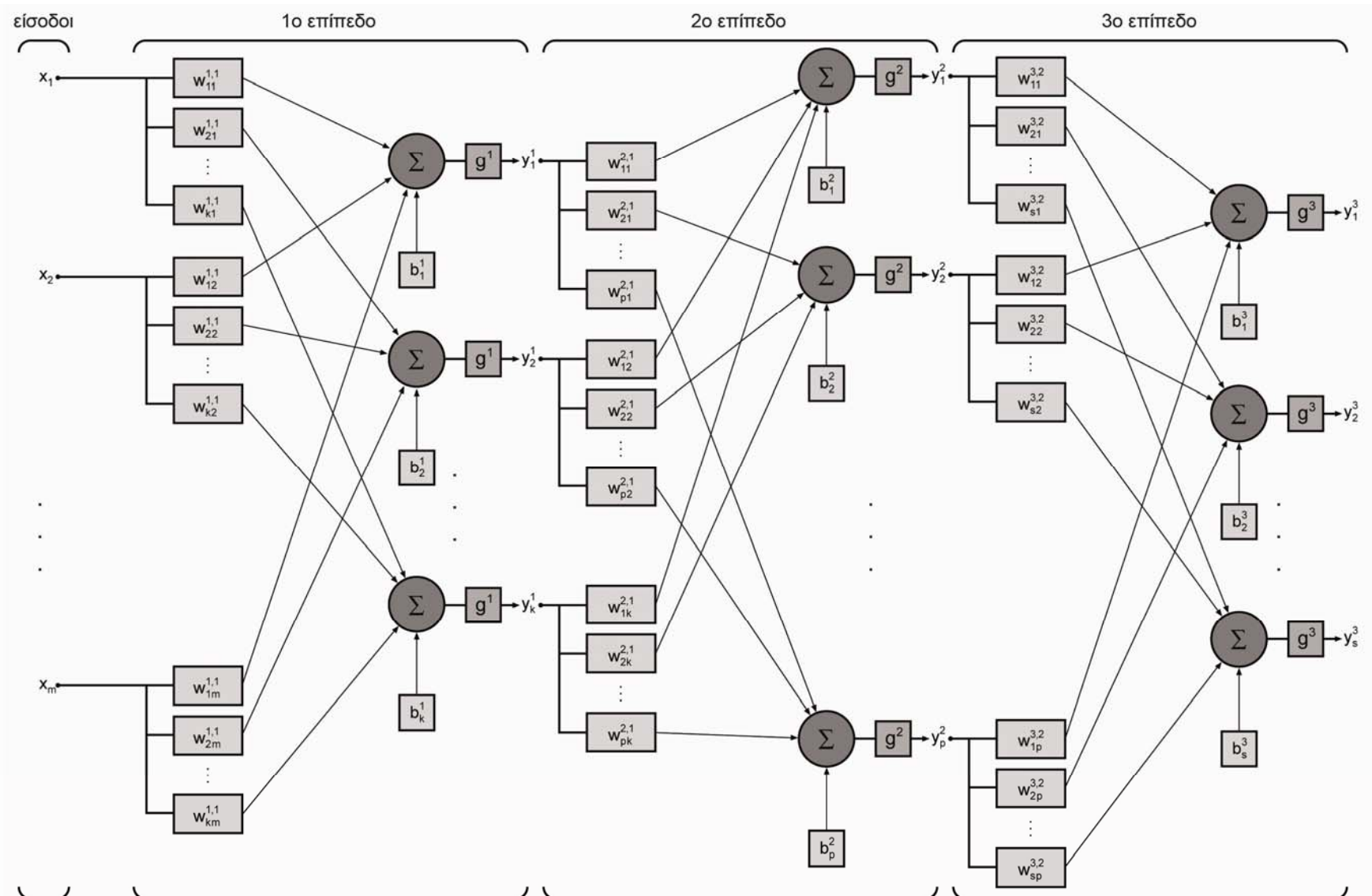
g^v η συνάρτηση μεταφοράς του v -οστού ενδιάμεσου επιπέδου¹,

y_f^v η έξοδος του f -οστού κόμβου κόμβο του v -οστού ενδιάμεσου επιπέδου.

Η ενεργοποίηση και η έξοδος του f -οστού κόμβου κόμβο του v -οστού ενδιάμεσου επιπέδου δίνονται από τις εξισώσεις:

$$\alpha_f^v = \sum_{d=1}^c w_{fd}^{v,v-1} y_d^{v-1} + b_f^v \quad \text{και} \quad y_f^v = g^v(\alpha_f^v) = g^v \left(\sum_{d=1}^c w_{fd}^{v,v-1} y_d^{v-1} + b_f^v \right) \quad (3.13)$$

¹ Αξίζει να σημειωθεί ότι όλοι οι κόμβοι κάθε επιπέδου χρησιμοποιούν την ίδια συνάρτηση μεταφοράς. Η χρήση διαφορετικών συναρτήσεων ανά κόμβο καθιστά εξαιρετικά πολύπλοκη την λειτουργία του δικτύου, χωρίς να βελτιώνει ουσιαστικά την αποτελεσματικότητά του.



Σχήμα 3-10 Αναλυτική απεικόνιση λειτουργίας νευρωνικού δικτύου αρχιτεκτονικής $m \times k \times p \times s$

Η πολύπλοκη αυτή λειτουργία του νευρωνικού δικτύου μπορεί να περιγραφεί καλύτερα με χρήση πινάκων για τις λειτουργικές παραμέτρους του συστήματος. Για το παραπάνω δίκτυο η διανυσματική έξοδος θα δίνεται από την σχέση:

$$\mathbf{y}^3 = \mathbf{g}^3 \left(\mathbf{W}^{3,2} \cdot \mathbf{g}^2 \left(\mathbf{W}^{2,1} \cdot \mathbf{g}^1 \left(\mathbf{W}^{1,1} \cdot \mathbf{x} + \mathbf{b}^1 \right) + \mathbf{b}^2 \right) + \mathbf{b}^3 \right) \quad (3.14)$$

όπου $\mathbf{x}_{m \times 1}$ το διάνυσμα των m εισόδων του συστήματος

$\mathbf{W}_{e \times c}^{v,v-1}$ ο πίνακας βαρών της εξόδου του $(v-1)$ -οστού ενδιάμεσου επιπέδου που εισάγεται στο v -οστό ενδιάμεσο επίπεδο, και ειδικά για το επίπεδο εισόδου,

$\mathbf{W}_{k \times m}^{1,1}$ ο πίνακας βαρών της διανυσματικής εισόδου $\mathbf{x}_{m \times 1}$ του συστήματος που εισάγεται στο 1ο ενδιάμεσο επίπεδο,

$\mathbf{b}_{e \times 1}^v$ το διάνυσμα των πολώσεων των e κόμβων του v -οστού ενδιάμεσου επιπέδου,

$\mathbf{y}_{e \times 1}^v$ η διανυσματική έξοδος του v -οστού ενδιάμεσου επιπέδου, και

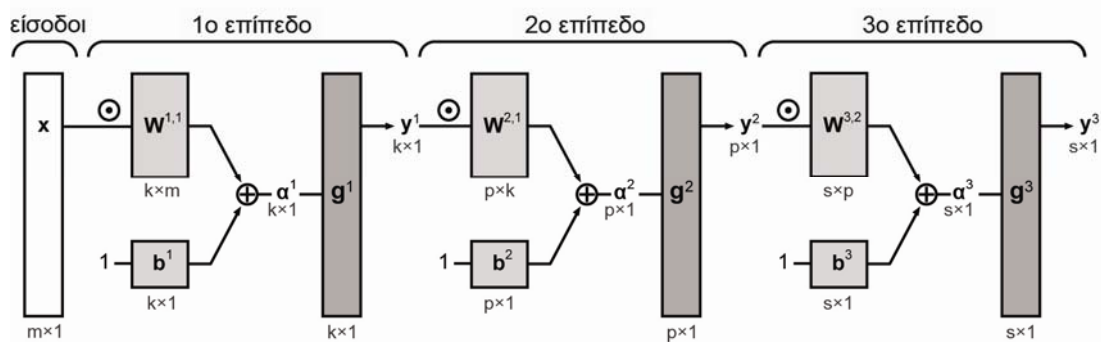
$\mathbf{g}^v$ το διάνυσμα των συναρτήσεων μεταφοράς του v -οστού ενδιάμεσου επιπέδου, με:

$$\mathbf{g}(\mathbf{z}) = \begin{bmatrix} g(z_1) \\ g(z_2) \\ \vdots \\ g(z_n) \end{bmatrix} \quad (3.15)$$

Τα διανύσματα της ενεργοποίησης και της εξόδου του v -οστού ενδιάμεσου επιπέδου δίνονται από τις εξισώσεις:

$$\mathbf{a}_{e \times 1}^v = \mathbf{W}_{e \times c}^{v,v-1} \cdot \mathbf{y}_{c \times 1}^{v-1} + \mathbf{b}_{e \times 1}^v \quad \text{και} \quad \mathbf{y}_{e \times 1}^v = \mathbf{g}_{e \times 1}^v(\mathbf{a}_{e \times 1}^v) = \mathbf{g}_{e \times 1}^v \left(\mathbf{W}_{e \times c}^{v,v-1} \cdot \mathbf{y}_{c \times 1}^{v-1} + \mathbf{b}_{e \times 1}^v \right) \quad (3.16)$$

Σχηματικά το νευρωνικό δίκτυο του Σχ. 3-10 μπορεί να αποδοθεί από την συμπυκνωμένη απεικόνιση του Σχ. 3-11.



Σχήμα 3-11 Απεικόνιση λειτουργίας νευρωνικού δικτύου αρχιτεκτονικής $m \times k \times p \times s$ με χρήση πινάκων για τις εισόδους, τα βάρη, τις πολώσεις και τις εξόδους

Εκπαίδευση νευρωνικών δικτύων

Πέρα από την λειτουργία προώθησης σήματος δοσμένης της επιφάνειας απόφασης ή της συναρτησιακής εξάρτησης των εξόδων από τις ανεξάρτητες μεταβλητές, βασική προϋπόθεση είναι η σωστή συμπεριφορά του δικτύου στη λήψη ορισμένου ερεθίσματος, δηλαδή η αναγνώριση του σήματος εισόδου και η κατάλληλη απόκριση της μονάδας. Το δίκτυο θα πρέπει να ανταποκριθεί στην εύρεση της ζητούμενης (ή κατάλληλης) επιφάνειας απόφασης στην περίπτωση ταξινόμησης δειγμάτων, ή –με όρους προσέγγισης συνάρτησης– στην πρόβλεψη της ζητούμενης (ή κατάλληλης) συναρτησιακής σχέσης εισόδου-εξόδου. Απαιτείται λοιπόν μια διαδικασία που να καθιστά την μονάδα, και κατ' επέκταση το συνολικό δίκτυο, κατάλληλα στον εκάστοτε δειγματικό χώρο, μια διαδικασία δηλαδή που να προσαρμόζει τα στοιχεία επεξεργασίας του συστήματος, βάρη και πολώσεις, στις παραμέτρους $\mathbf{x}$, $\mathbf{y}$ ενός πληθυσμού δειγμάτων εντός του χώρου αυτού.

Η διαδικασία ρύθμισης των λειτουργικών παραμέτρων του νευρωνικού δικτύου αναφέρεται ως **εκπαίδευση** (training) του δικτύου. Κατά την εκπαίδευσή του ένα δίκτυο τροποποιεί κατάλληλα τα στοιχεία επεξεργασίας του, ώστε να ταξινομεί σωστά τα δείγματα ή να προβλέπει με ακρίβεια τις τιμές των εξαρτημένων μεταβλητών τους, χρησιμοποιώντας ως πρότυπο ένα σύνολο “γνωστών” δοκιμαστικών δειγμάτων, το **σύνολο εκπαίδευσης** (training set), που καλύπτει όσο το δυνατόν πληρέστερα των ζητούμενο χώρο εφαρμογής του. Η διαδικασία εκπαίδευσης διακρίνεται σε επιβλεπόμενη ή εποπτευόμενη και μη επιβλεπόμενη ή αυτοοργάνωση. Στην πρώτη περίπτωση παρέχεται στο δίκτυο πληροφορία για την αναμενόμενη “ορθή” λειτουργία με την μορφή τιμών-στόχων για κάθε δείγμα εκπαίδευσης και το δίκτυο καλείται να προσαρμοστεί έτσι ώστε να προβλέπει όσο το δυνατό καλύτερα κάθε δείγμα, δηλαδή οι έξοδοι του να προσεγγίζουν όσο το δυνατόν τους στόχους. Στην μη εποπτευόμενη εκπαίδευση, αντίθετα, παρέχεται στο δίκτυο μόνο ο πίνακας $\mathbf{X}$ των δεδομένων του συνόλου εκπαίδευσης χωρίς καμία πληροφορία για τις αναμενόμενες αποκρίσεις και καλείται το ίδιο να ανιχνεύσει και να προσαρμοστεί στα κοινά χαρακτηριστικά των δεδομένων εισαγωγής του. Οι μέθοδοι μη εποπτευόμενης εκπαίδευσης χρησιμοποιούνται κυρίως για ταξινόμηση.

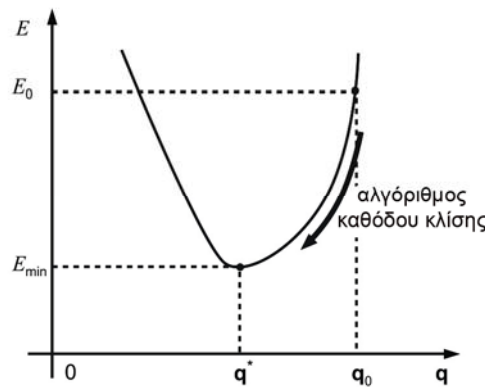
Η εσωτερική λειτουργία του δικτύου κατά την οποία αυτό τροποποιεί τα στοιχεία επεξεργασίας του ανάλογα με την πληροφορία των δεδομένων εισαγωγής αναφέρεται ως **εκμάθηση** (learning) του συνόλου εκπαίδευσης από το δίκτυο. Πρόκειται για μια επαναληπτική διαδικασία παρουσίας στο δίκτυο του συνόλου εκπαίδευσης, σε κάθε κύκλο της οποίας, πραγματοποιούνται μικρές αλλαγές στα βάρη και τις πολώσεις έτσι ώστε η απόκριση του συστήματος να προσεγγίζει καλύτερα την έξοδο-στόχο. Η εντολή για την αλλαγή των βαρών σε κάθε βήμα, ή αλλιώς **εποχή** (epoch) εκπαίδευσης, αποτελεί τον **κανόνα εκμάθησης** (learning

rule). Αποτελεί στην ουσία έναν αλγόριθμο εύρεσης του βέλτιστου πίνακα βαρών και πολώσεων του συστήματος, ή αλλιώς έναν αλγόριθμο βελτιστοποίησης μίας συνάρτησης που εκφράζει την απόδοση του δικτύου συναρτήσει των στοιχείων επεξεργασίας του. Στην περίπτωση επιβλεπόμενης εκπαίδευσης, ως μέτρο της απόδοσης του δικτύου μπορεί να χρησιμοποιηθεί μία συνάρτηση σφάλματος που εκφράζει την απόκλιση των εξόδων του δικτύου από τις επιθυμητές τιμές τους

$$E = E(\mathbf{q}, \mathbf{X}) \quad (3.17)$$

όπου $\mathbf{q}$ το διάνυσμα παραμέτρων του δικτύου, που περιλαμβάνει το σύνολο των στοιχείων επεξεργασίας του δικτύου (βάρη και πολώσεις) και $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_i, \dots, \mathbf{x}_n)$ ο πίνακας των εισόδων για όλα τα δείγματα του συνόλου εκπαίδευσης. Η διαδικασία εκπαίδευσης (βελτιστοποίηση της E) συνίσταται στον προσδιορισμό των στοιχείων επεξεργασίας, αρχίζοντας από κάποιες τυχαίες αυθαίρετες τιμές τους, για τα οποία το συνολικό σφάλμα ελαχιστοποιείται (Σχ. 3-12).

$$\mathbf{q} \rightarrow \mathbf{q}^*, \text{ ώστε } E \rightarrow E_{\min}$$



Σχήμα 3-12 Σχηματική λειτουργία του αλγορίθμου καθόδου κλίσης

Αλγόριθμος καθόδου κλίσης

Για την εκπαίδευση των νευρωνικών δικτύων μπορεί να χρησιμοποιηθεί ένα πλήθος αλγορίθμων βελτιστοποίησης, μεταξύ των οποίων οι BFGS, Trust Region, Levenberg-Marquardt, Newton. Ένας ιδιαίτερα δημοφιλής αλγόριθμος βελτιστοποίησης είναι αυτός της **καθόδου κλίσης** (gradient descent), σύμφωνα με τον οποίο σε κάθε επανάληψη της διαδικασίας η μεταβολή της κάθε παραμέτρου επεξεργασίας θα δίνεται:

$$\Delta w = -\alpha \frac{\partial E}{\partial w} \quad (3.18)$$

όπου α , $0 < \alpha < 1$ ένας συντελεστής που καθορίζει το βήμα της μεταβολής σε κάθε επανάληψη και αναφέρεται ως **ρυθμός εκμάθησης** (learning rate).

Ας θεωρήσουμε την απλή περίπτωση ενός κόμβου με διάνυσμα εισόδου $\mathbf{x}' = (1, x_1, x_2, \dots, x_m)$, διάνυσμα βαρών $\mathbf{w}' = (b, w_1, w_2, \dots, w_m) = (w_0, w_1, w_2, \dots, w_m)$ και έξοδο y . Η χρήση των επαυξημένων πινάκων για τις εισόδους και τα βάρη επιτρέπει στον κανόνα εκμάθησης να χειρίζεται ενιαία όλες τις λειτουργικές παραμέτρους της μονάδας, χρησιμοποιώντας τις ίδιες εξισώσεις τόσο για τα βάρη, όσο και για την πόλωση. Για ένα σύνολο εκπαίδευσης n δειγμάτων δίνονται, για κάθε δείγμα i , οι τιμές $\{x_i, t_i\}$ των ανεξάρτητων μεταβλητών και του στόχου της εξαρτημένης μεταβλητής αντίστοιχα. Ως συνάρτηση απόδοσης της μονάδας μπορεί να χρησιμοποιηθεί το συνολικό σφάλμα, δηλαδή η μέση τιμή του σφάλματος υπολογισμού¹ του y για κάθε δείγμα:

$$E = \frac{1}{n} \sum_{i=1}^n e_i \quad (3.19)$$

$$e_i = \frac{1}{2} (t_i - y_i)^2 \quad (3.20)$$

Σε κάθε επανάληψη του αλγορίθμου εκπαίδευσης της μονάδας τα βάρη μεταβάλλονται κατά:

$$\Delta w_j = -\alpha \frac{\partial E}{\partial w_j} = -\frac{\alpha}{n} \left(\sum_{i=1}^n \frac{\partial e_i}{\partial w_j} \right) \quad (3.21)$$

$$\text{όπου } \frac{\partial e_i}{\partial w_j} = -(t_i - y_i) \frac{\partial y_i}{\partial w_j} = -(t_i - y_i) \frac{dg(\alpha)}{d\alpha} \frac{\partial \alpha}{\partial w_j} = -(t_i - y_i) \frac{dg(\alpha)}{d\alpha} \frac{\partial \left(\sum_{j=0}^m w_j x_j \right)}{\partial w_j} = -g'(\alpha) (t_i - y_i) x_j$$

Ο κανόνας εκμάθησης που περιγράφεται από την (3.20) χρησιμοποιήθηκε αρχικά από τους Widrow & Hoff (1960) για εκπαίδευση γραμμικών μονάδων ADALINEs (ADaptive LINear Elements)² και είναι γνωστός ως **κανόνας δέλτα** (δ -rule), λόγω του όρου διαφοράς $\delta = (t_i - y_i) = (t_i - \alpha)$ που υπεισέρχεται στην μεταβολή κάθε βάρους.

Ο κανόνας- δ μπορεί να χρησιμοποιηθεί ανάλογα και στην περίπτωση εκπαίδευσης νευρωνικού δικτύου ενός επιπέδου k κόμβων, όπου το σφάλμα για κάθε δείγμα i του συνόλου εκπαίδευσης είναι το άθροισμα των σφαλμάτων σε κάθε κόμβο, αφού οι κόμβοι δρουν ανεξάρτητα.

$$e_i = \frac{1}{2} \sum_{\ell=1}^k (t_{i\ell} - y_{i\ell})^2 \quad (3.22)$$

¹ Η μορφή του σφάλματος της εξίσωσης (3.20) αντικαθιστά το σφάλμα προσέγγισης $e_i = t_i - y_i$ για λόγους απλοποίησης των υπολογισμών κατά τις παραγωγίσεις

² Πρόκειται για μονάδες επεξεργασίας, αντίστοιχες του TLU, που εφαρμόζουν γραμμική συνάρτηση μεταφοράς στην ενεργοποίησή τους $y = g(\alpha) = \alpha$, με $g'(\alpha) = 1$

Ο κανόνας εκμάθησης θα περιγραφεί από την σχέση:

$$\Delta w_{\ell j} = -\alpha \frac{\partial E}{\partial w_{\ell j}} = -\frac{\alpha}{n} \left(\sum_{i=1}^n \frac{\partial e_i}{\partial w_{\ell j}} \right) \quad (3.23)$$

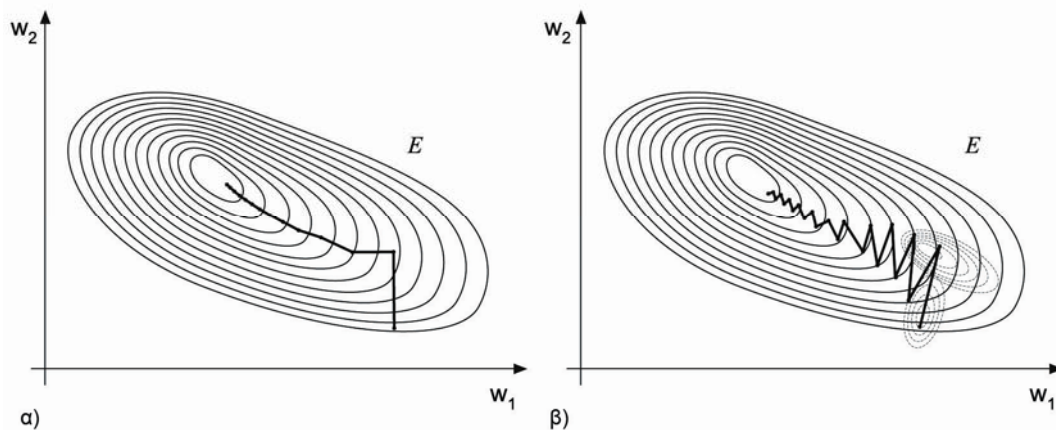
όπου η κλίση του σφάλματος για το j -οστό βάρος του ℓ -οστού κόμβου ($\ell=1, \dots, k$) θα δίνεται από την εξίσωση $\frac{\partial e_i}{\partial w_{\ell j}} = -g'(\alpha_\ell) (t_i - y_i) x_{ij}$, αφού οι υπόλοιπες έξοδοι του δικτύου, εκτός της y_ℓ , είναι ανεξάρτητες του βάρους $w_{\ell j}$.

Ήδη, από την απλή περίπτωση του μονοεπίπεδου δικτύου, είναι φανερό ότι η συνάρτηση σφάλματος E , και επομένως και οι μερικοί παράγωγοί της, εξαρτώνται από το σύνολο των δειγμάτων εκπαίδευσης και συνεπώς για την εκτίμηση των κλίσεων $\partial E / \partial w_{\ell j}$, και άρα και των διορθωτικών μεταβολών $\Delta w_{\ell j}$ που θα εφαρμοστούν στα βάρη σε κάθε εποχή εκπαίδευσης, θα πρέπει προηγουμένως να έχουν παρουσιαστεί στο δίκτυο όλα τα δείγματα. Αυτή η **ομαδική εκπαίδευση** (batch training) συνολικού χειρισμού του πίνακα δεδομένων $\mathbf{X}$ των δειγμάτων εκπαίδευσης οδηγεί στην πραγματική κάθοδο κλίσης, ενώ ένας πιο ταχύς τύπος εκπαίδευσης είναι η **ανά δείγμα εκπαίδευση** (incremental training), κατά την οποία η προσαρμογή των βαρών και πολώσεων πραγματοποιείται μετά την παρουσίαση κάθε δείγματος, δηλαδή με βάση τις κλίσεις $\partial e_i / \partial w_{\ell j}$ που αποτελούν μερική εκτίμηση της πραγματικής κλίσης της συνάρτησης σφάλματος.

Στην περίπτωση αυτή, η μεταβολή των βαρών σε κάθε εποχή, που συμπίπτει με την εφαρμογή στο δίκτυο του μεμονωμένου κάθε φορά δείγματος, θα δίνεται από την σχέση:

$$\Delta w_{\ell j} = -\alpha \frac{\partial e_i}{\partial w_{\ell j}} = \alpha g'(\alpha_\ell) (t_i - y_i) x_{ij} \quad (3.24)$$

Επειδή χρησιμοποιούνται εκτιμήσεις της πραγματική κλίση, η πρόοδος κατά την ελαχιστοποίηση της E δεν είναι ομαλή και επιτρέπει μεταβολές των βαρών που οδηγούν σε τοπική αύξηση της E . Όμως συνολικά το αποτέλεσμα είναι μια συστηματική μείωση του σφάλματος. Στο Σχ. 3-13 παρουσιάζεται σχηματικά η λειτουργία των δύο μεθόδων εκπαίδευσης σε χάρτη ισοϋψών της συνάρτησης σφάλματος ως προς δύο λειτουργικές παραμέτρους δικτύου w_1 και w_2 .



Σχήμα 3-13 Σχηματική εξέλιξη α) ομαδικής και β) ανά δείγμα εκπαίδευσης

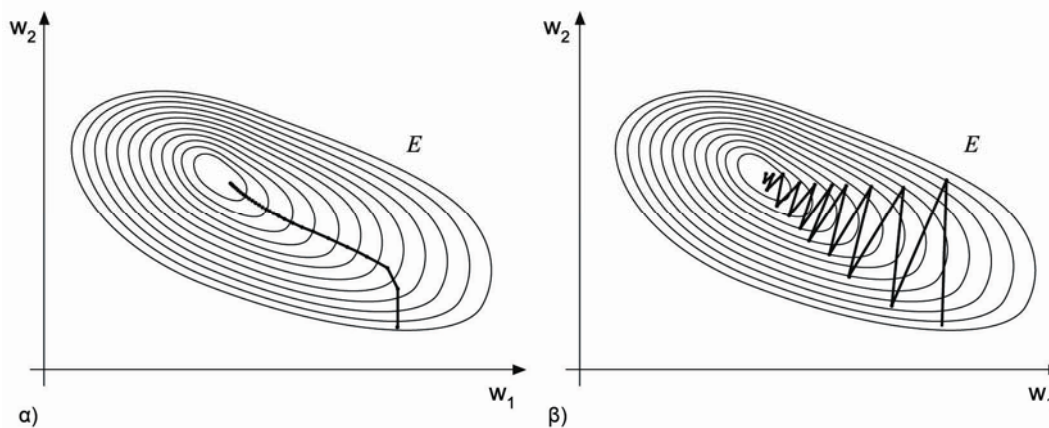
Αλγόριθμος ανάστροφης διάδοσης

Για την εκπαίδευση πολυεπίπεδων νευρωνικών δικτύων ευθείας προώθησης, ο αλγόριθμος εύρεσης των βέλτιστων τιμών των στοιχείων επεξεργασίας που χρησιμοποιείται συνήθως, αποτελεί γενίκευση του κανόνα-δ, ώστε να συνυπολογίζονται στην συνάρτηση σφάλματος τόσο τα σφάλματα των κόμβων εξόδου, όσο και τα σφάλματα εκείνων των ενδιάμεσων επιπέδων. Είναι προφανές ότι τα τελευταία δεν μπορούν να υπολογιστούν άμεσα, παρά μόνο αφού υπολογιστούν τα τελικά σφάλματα (αποκλίσεις αποκρίσεων δικτύου από τους στόχους) των κόμβων εξόδου. Για κάθε ενδιάμεσο κόμβο, το σφάλμα του υπολογίζεται ως το άθροισμα των σφαλμάτων των επόμενων κόμβων σταθμισμένων με τα βάρη των αντίστοιχων μεταξύ τους συνάψεων, που αντικατοπτρίζει την συμμετοχή του εν λόγω κόμβου στα τελικά σφάλματα του επιπέδου εξόδου. Ένας τέτοιος αλγόριθμος εκπαίδευσης, που βασίζεται στην διαβίβαση των σφαλμάτων κατά αντίθετη φορά από αυτή της διαστροφώσεως των επιπέδων του δικτύου, ονομάζεται **αλγόριθμος ανάστροφης διάδοσης** (backpropagation algorithm).

Οι διορθώσεις των βαρών με βάση τον αλγόριθμο ανάστροφης διάδοσης μπορούν να συνεχίζονται έπ' άπειρο, αν δεν οριστεί κάποιο **κριτήριο διακοπής** (stopping criterion) του επαναληπτικού αυτού αλγόριθμου. Ένα τέτοιο κατάλληλο κριτήριο θα ήταν μία ελάχιστη τιμή της συνάρτησης σφάλματος, η οποία να καλύπτει τις απαιτήσεις σε ακρίβεια της εφαρμογής. Στην πράξη χρησιμοποιείται ένα συνδυασμός συνθηκών για την παύση της εκπαίδευσης που περιλαμβάνουν τον μέγιστο αριθμό εποχών εκπαίδευσης, την ελάχιστη τιμή της συνάρτησης σφάλματος, τον μέγιστο επιτρεπόμενο χρόνο εκπαίδευσης, κ.α..

Η χρήση ενός αλγόριθμου βελτιστοποίησης, όπως αυτός της καθόδου κλίσης, που βασίζεται στον υπολογισμό των μερικών παραγώγων της συνάρτησης απόδοσης του δικτύου, είναι αρκετά ικανοποιητική για εκπαίδευση τυπικών νευρωνικών δικτύων που χρησιμοποιούνται

συνήθως για προσέγγιση συναρτησιακής εξάρτησης μεταξύ μεταβλητών. Η εφαρμογή ισχυρότερων μεθόδων βελτιστοποίησης, όπως αυτή του Newton¹, δεν συνίσταται, παρά την ταχύτερη σύγκλιση που επιτυγχάνεται, λόγω της υπολογιστικής επιβάρυνσης. Διάφορα μειονεκτήματα της καθόδου κλίσης, όπως ο εγκλωβισμός σε τοπικά ελάχιστα της συνάρτησης σφάλματος και η μικρή ταχύτητα εκπαίδευσης, μπορούν να υπερκεραστούν με τροποποιήσεις στον βασικό αλγόριθμο.



Σχήμα 3-14 Σχηματική εξέλιξη εκπαίδευσης με χρήση α) χαμηλού και β) υψηλού ρυθμού εκμάθησης

Η απόδοση του αλγόριθμου καθόδου κλίσης εξαρτάται άμεσα από τον εφαρμοζόμενο ρυθμό εκμάθησης α . Ένας υψηλός ρυθμός εκμάθησης, οδηγεί σε ταλάντωση, αντίστοιχη με εκείνη της ανά δείγμα εκπαίδευσης (Σχ. 3-14 β)), που ενέχει τον κίνδυνο ο αλγόριθμος να καταστεί ασταθής, παράλληλα όμως επιτρέπει στον αλγόριθμο να υπερπηδήσει τυχόν κοντινά κοίλα της συνάρτησης ώστε να διαφύγει από τοπικά ελάχιστα της. Αντίθετα μικροί ρυθμοί εκμάθησης καθυστερούν σημαντικά την σύγκλιση του αλγορίθμου και οδηγούν συχνά σε εγκλωβισμό του σε τοπικά-μη ολικά ελάχιστα (Σχ. 3-14 α)). Η επιλογή κατάλληλου ρυθμού εκμάθησης δεν είναι εφικτή πριν την εκπαίδευση, πόσο μάλλον αφού ο βέλτιστος ρυθμός εκμάθησης μεταβάλλεται κατά την διάρκεια της τελευταίας. Μία σημαντική τροποποίηση βελτίωσης της απόδοσης του αλγόριθμου καθόδου κλίσης είναι η χρήση **μεταβλητού ρυθμού εκμάθησης** (adaptive learning rate) που να ανταποκρίνεται στην τοπική πολυπλοκότητα της επιφάνειας σφάλματος. Η τεχνική αυτή επιλέγει σε κάθε εποχή εκπαίδευσης μέγεθος βήματος όσο το δυνατόν μεγαλύτερο, ελέγχοντας παράλληλα την σταθερότητα του αλγορίθμου. Η διόρθωση των βαρών σε κάθε εποχή γίνεται με βάση τον τρέχοντα ρυθμό εκμάθησης, το σφάλμα που προκύπτει συγκρίνεται με το προηγούμενο και επιλέγεται ο βέλτιστος για την εποχή ρυθμός εκμάθησης (μικρότερος ή

¹ Ο αλγόριθμος Newton χρησιμοποιεί δεύτερης τάξης μερικές παραγώγους της συνάρτησης σφάλματος

μεγαλύτερος του τρέχοντος), ανάλογα με το αν το νέο σφάλμα είναι μεγαλύτερο ή μικρότερο του προηγούμενου.

Στην ίδια λογική της προσαρμοσίμης στην επιφάνεια σφάλματος εκπαίδευσης (adaptive training) κινείται και η προσθήκη στον βασικό αλγόριθμο καθόδου κλίσης του **όρου ορμής** (momentum term). Ο αλγόριθμος τροποποιείται έτσι ώστε να διατηρεί “μνήμη” των διορθωτικών αλλαγών των βαρών που πραγματοποιήθηκαν σε προηγούμενες εποχές εκπαίδευσης και κατά κάποιο τρόπο εισάγεται η –δανεισμένη από την Φυσική –έννοια της αδράνειας για το σύστημα. Στην μεταβολή ενός βάρους του δικτύου κατά την v -οστή εποχή εκπαίδευσης προστίθεται ο επιπλέον όρος ορμής, που περιλαμβάνει την διόρθωση του εν λόγω βάρους κατά την προηγούμενη ($v-1$) εποχή. Η νέα διόρθωση θα δίνεται από μία σχέση της μορφής:

$$(\Delta w_{fd}^{v,v-1})^{(v)} = -\alpha \left(\frac{\partial E}{\partial w_{fd}^{v,v-1}} \right)^{(v)} + \lambda (\Delta w_{fd}^{v,v-1})^{(v-1)} \quad (3.25)$$

όπου λ , $0 < \lambda < 1$ η σταθερά ορμής.

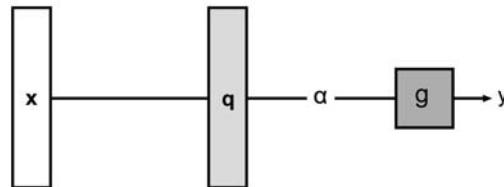
Για μηδενική τιμή της σταθεράς ορμής, η μεταβολή του βάρους ορίζεται αποκλειστικά βάση της τρέχουσας κλίσης της συνάρτησης σφάλματος¹, ενώ μεγαλύτερες τιμές της αναγκάζουν τον αλγόριθμο να συνυπολογίσει για την παραγωγή του $(\Delta w_{fd}^{v,v-1})^{(v)}$ τις κλίσεις στο τρέχον και το προηγούμενο βήμα εκπαίδευσης, επιτρέποντας του να προσαρμοστεί όχι μόνο στην τοπική κλίση, αλλά και στις ευρύτερες τάσεις της επιφάνειας σφάλματος. Ενεργώντας λοιπόν αντίστοιχα με ένα χαμηλής διέλευσης φίλτρο, η ορμή επιτρέπει στο δίκτυο για να αγνοήσει τις τοπικές ανωμαλίες στην επιφάνεια λάθους, επιταχύνοντας τον αλγόριθμο όπου οι κλίσεις είναι πρακτικά σταθερές ή ανακόπτοντας τον ρυθμό εξέλιξης του όταν οι διαδοχικές κλίσεις μεταβάλλονται σημαντικά.

3.3 Συναρτήσεις Μεταφοράς και Στοιχεία Τοπολογίας

Η λειτουργική συμπεριφορά των περισσότερων τεχνητών νευρώνων βασίζεται στο τυπικό μοντέλο ενεργοποίησης-εξόδου, που φαίνεται σχηματικά στο Σχ. 3-15, όπου $\mathbf{x} = \{x_i\}$ το διάνυσμα των m εισόδων και $\mathbf{q} = \{q_i\}$ το διάνυσμα των L λειτουργικών παραμέτρων της μονάδας. Η ενεργοποίηση υπολογίζεται ως συνάρτηση των εισόδων και των παραμέτρων της

¹ Παραλλαγές της επαναληπτικής σχέσης (3.25) μπορούν να περιλαμβάνουν για τον πρώτο όρο έναν επιπλέον συντελεστή, εκτός του ρυθμού εκμάθησης, που εξαρτάται από την σταθερά ορμής ($1-\lambda$, λ , κ.α.), ώστε να ενισχύεται ή μετριάζεται η επίδραση της τρέχουσας κλίσης της επιφάνειας σφάλματος στις διορθώσεις που πραγματοποιούνται σε κάθε εποχή εκπαίδευσης.

μονάδας, $\alpha = \alpha(q_i, x_j)$ και η έξοδος y προκύπτει ως συνάρτηση της ενεργοποίησης, $y = y(\alpha)$. Στους συνήθεις νευρώνες (TLU, perceptrons, κ.α.) η ενεργοποίηση δίνεται σχεδόν αποκλειστικά από το σταθμισμένο άθροισμα¹ των εισόδων με $m+1$ παραμέτρους λειτουργίας της μονάδας: τα βάρη w_j και την πόλωση b . Επομένως ο βασικός διαχωρισμός των μονάδων των νευρωνικών δικτύων που χρησιμοποιούνται για προσέγγιση συνάρτησης γίνεται με βάση την συνάρτηση μεταφοράς g , που εφαρμόζεται για παραγωγή της εξόδου συναρτήσει της ενεργοποίησης α .



Σχήμα 3-15 Μοντέλο ενεργοποίησης-εξόδου των τεχνητών νευρώνων.

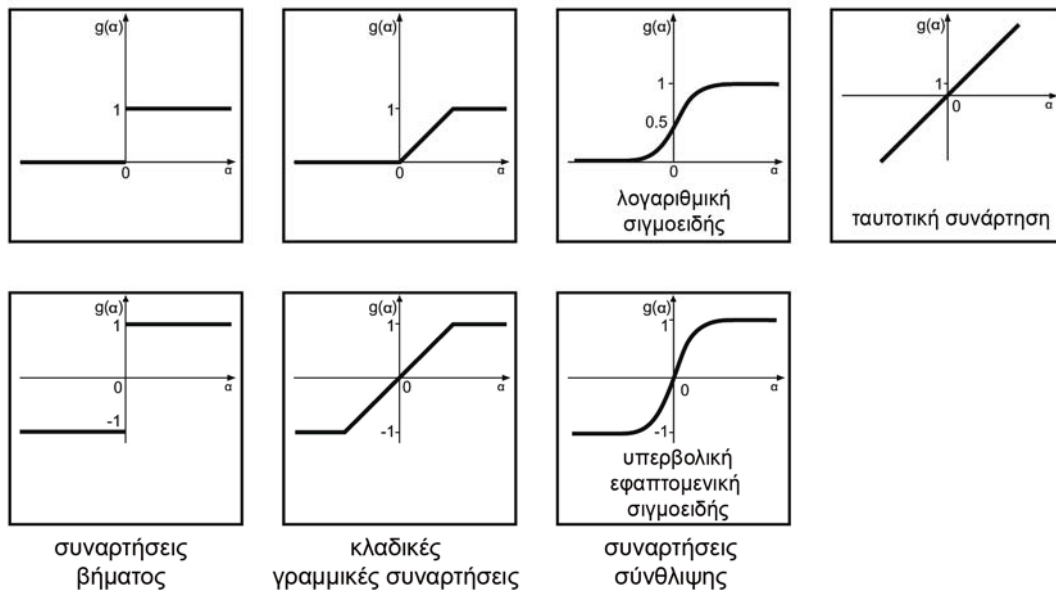
Είδη Συναρτήσεων Μεταφοράς

Η συνάρτηση μεταφοράς που χρησιμοποιούν οι μονάδες του δικτύου εξαρτάται από τις ανάγκες της εκάστοτε εφαρμογής. Οι κόμβοι TLU χρησιμοποιούν, όπως ήδη συζητήθηκε, την συνάρτηση βήματος (ή αυστηρά περιοριστική), είτε φραγμένη στο διάστημα $(0, 1)$, είτε στο $(-1, 1)$, που δίνεται από τις εξισώσεις (3.8). Ο “όλα ή τίποτα” χαρακτήρας που αποδίδει στην έξοδο αυτή η συνάρτηση ενεργοποίησης, μπορεί να εξομαλυνθεί με χρήση **γραμμικών μονάδων** (linear units) που εφαρμόζουν γραμμικές συναρτήσεις μεταφοράς με εξίσωση:

$$y = g(\alpha) = \begin{cases} 0, & \text{αν } \alpha < 0 \\ \alpha, & \text{αν } 0 \leq \alpha < 1 \\ 1, & \text{αν } \alpha \geq 1 \end{cases} \quad \text{ή αντίστοιχα} \quad y = g(\alpha) = \begin{cases} -1, & \text{αν } \alpha < -1 \\ \alpha, & \text{αν } -1 \leq \alpha < 1 \\ 1, & \text{αν } \alpha \geq 1 \end{cases} \quad (3.26)$$

Η χρήση μόνο γραμμικών κόμβων, είναι φανερό, ότι έχει σαν αποτέλεσμα την ανάπτυξη αποκλειστικά γραμμικής συναρτησιακής σχέσης μεταξύ των ανεξάρτητων μεταβλητών–εισόδων και των εξαρτημένων–εξόδων του συστήματος

¹ Αυτή η μορφή ενεργοποίησης αποτελεί στην πραγματικότητα μία γραμμική σχέση μεταξύ των εισόδων του νευρώνα. Αφού η συνάρτηση ενεργοποίησης είναι ανεξάρτητη από την μορφή της ενεργοποίησης, θα μπορούσε να χρησιμοποιηθεί μία οποιαδήποτε διαφορετική σχέση μεταξύ των παραμέτρων του συστήματος και των εισόδων, όπως η πολυγραμμική μορφή $\alpha = \sum_{k=1}^{2^n} w_k \prod_{j \in J_k} x_j$, που εφαρμόζουν οι λεγόμενες sigma-pi μονάδες ή η μορφή απόστασης ($\alpha = \|\mathbf{w} - \mathbf{x}\|$) που χρησιμοποιούν οι μονάδες RBF (Radial Basis Functions). Η εφαρμογή, όμως, του σταθμισμένου αθροίσματος των εισόδων για την παραγωγή της ενεργοποίησης, εκτός από την απλούστευση των υπολογισμών διασφαλίζει την σύγκλιση του συστήματος σε οποιαδήποτε συνάρτηση μεταξύ των εξόδων y (εξαρτημένες μεταβλητές) και των εισόδων x (ανεξάρτητες μεταβλητές).



Σχήμα 3-16 Τύποι συναρτήσεων μεταφοράς

Η εφαρμογή, όμως, των περισσότερων αλγορίθμων βελτιστοποίησης (με εξαίρεση την μέθοδο Simplex) που χρησιμοποιούνται στην εκπαίδευση των νευρωνικών δικτύων απαιτεί τον υπολογισμό των παραγώγων της g σε όλο το εύρος τιμών της ενεργοποίησης, προϋπόθεση που δεν πληρούν οι TLU και οι γραμμικοί κόμβοι λόγω της κλαδικής μορφής της συνάρτησης μεταφοράς που χρησιμοποιούν. Αυτό, καθώς και η ανάγκη για προσέγγιση της μη γραμμικής εξάρτησης μεταξύ των μεταβλητών οδήγησε στην χρήση των **συναρτήσεων σύνθλιψης** (squashing functions) –άνω και κάτω φραγμένων, συνεχών και παραγωγίσιμων– για την παραγωγή της εξόδου, συνηθέστερες από τις οποίες είναι οι λεγόμενες –λόγω της μορφής τους– σιγμοειδείς (λογαριθμική και υπερβολική εφασπτομενική) (Σχ. 3-16). Οι δύο αυτές συναρτήσεις μεταφοράς παράγουν έξοδο στο διάστημα $(0, 1)$ και $(-1, 1)$ αντίστοιχα, που εξαρτάται ομαλά, μη γραμμικά από την ενεργοποίηση και επομένως από τις εισόδους της μονάδας, και περιγράφονται από τις εξισώσεις:

$$y = g(a) = \frac{1}{1 + e^{-a}} \quad \text{και} \quad y = g(a) = \frac{2}{1 + e^{-2a}} - 1 \quad (3.27)$$

Οι μονάδες που χρησιμοποιούν σιγμοειδή συνάρτηση μεταφοράς αναφέρονται ως ημιγραμμικού τύπου, λόγω της δυνατότητας να προσεγγιστεί η σιγμοειδής από μια συνεχή, τμηματικά-γραμμική συνάρτηση. Οι σιγμοειδείς συναρτήσεις εκτός των άλλων πλεονεκτημάτων που παρουσιάζουν, είναι εύκολα παραγωγίσιμες, με πρώτες παραγώγους $g'(a) = g(a)(1 - g(a))$ και $g'(a) = 2g(a)(1 - g(a))$ αντίστοιχα, γεγονός που απλοποιεί εξαιρετικά τους υπολογισμούς του αλγορίθμου εκπαίδευσης.

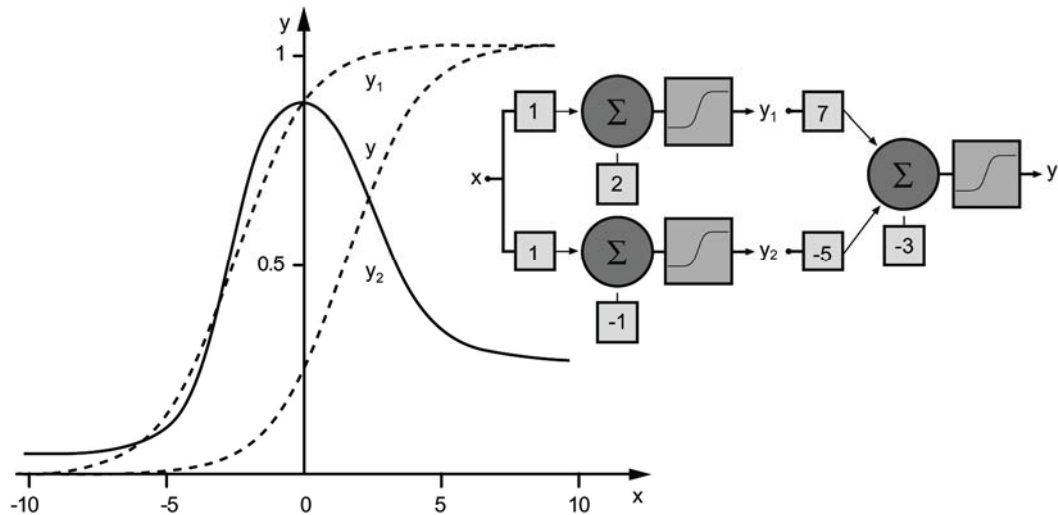
Η απαίτηση για άνω και κάτω φράγμα της συνάρτησης μεταφοράς αφορά το εύρος τιμών της εξόδου κάθε μονάδας επεξεργασίας, έτσι ώστε να βρίσκονται εντός κάποιου διαστήματος κανονικοποίησης $((0, 1)$ ή $(-1, 1)$) και να μπορούν να αποτελέσουν άμεσα εισόδους για τους απαγωγούς της μονάδας νευρώνες. Οι μονάδες του επιπέδου εξόδου, όμως, είτε για να μην εισάγουν επιπλέον μη γραμμικότητα στην συνάρτηση που καλείται να παρεμβάλει το δίκτυο στις τιμές των μεταβλητών X και Y , είτε για να παράγουν απ' ευθείας μη κανονικοποιημένες αποκρίσεις-εξόδους του συστήματος, χρησιμοποιούν συνήθως ως συνάρτηση μεταφοράς την **ταυτοτική** συνάρτηση (identical function).

$$y = g(a) = a \quad (3.28)$$

Τέλος, όπως φαίνεται και στο Σχ. 3-16, οι συναρτήσεις μεταφοράς που εφαρμόζουν οι τεχνητοί νευρώνες είναι, εκτός ελαχίστων εξαιρέσεων (π.χ. Gaussian συνάρτηση μεταφοράς σε μονάδες RBF), μονότονες και μάλιστα αύξουσες. Αποδεικνύεται ότι η χρήση αύξουσας συνάρτησης ενεργοποίησης έχει σαν αποτέλεσμα ομαλότερη συνάρτηση βελτιστοποίησης χωρίς πολλά τοπικά ελάχιστα που μπορούν να εγκλωβίσουν τον αλγόριθμο εκπαίδευσης μακριά από το ζητούμενο ολικό ελάχιστο.

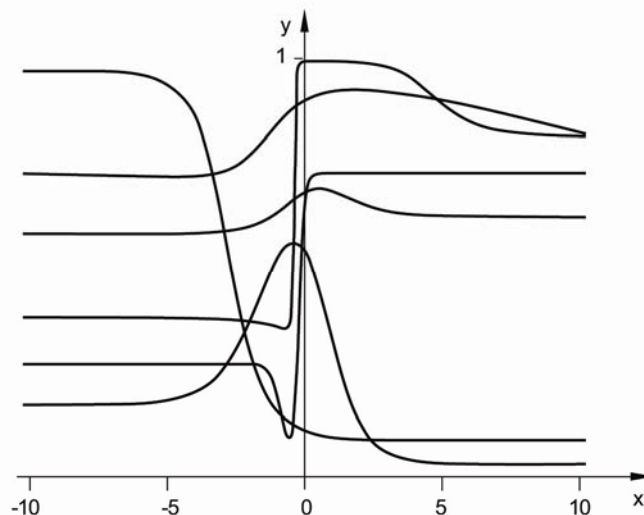
Αρχιτεκτονική και Προσέγγιση Συνάρτησης

Όπως έχει ήδη αναφερθεί, μία από τις πιο ενδιαφέρουσες εφαρμογές των νευρωνικών δικτύων είναι η προσέγγιση της πραγματικής συνάρτησης που διέπει τις μεταβλητές ενός φυσικού φαινομένου. Η λειτουργία του δικτύου εφαρμόζει μία μαθηματική σχέση με βάση την εξίσωση (3.12) μεταξύ των μεταβλητών εισόδου και εξόδου του συστήματος, που για προσέγγισης συνάρτησης μίας ανεξάρτητης μεταβλητής μπορεί να περιγραφεί με το παράδειγμα του απλού δικτύου δύο επιπέδων που παρουσιάζεται στο Σχ. 3-17. Το δίκτυο δέχεται μεμονωμένη είσοδο x σε ενδιάμεσο επίπεδο με δύο κόμβους με βάρη 1, 1 και πολώσεις 2, -1, αντίστοιχα, που παρέχουν είσοδο σε έναν κόμβο εξόδου με βάρη 7, -5 και πόλωση -3 για την παραγωγή της εξόδου y . Στο διάγραμμα του σχήματος δίνεται η γραφική παράσταση της συνάρτησης $y = g(x)$ που παράγεται από το σύστημα, όταν και οι τρεις κόμβοι εφαρμόζουν λογαριθμική σιγμοειδή συνάρτηση μεταφοράς. Οι διακεκομμένες καμπύλες αντιστοιχούν στις εξόδους των κόμβων του ενδιάμεσου επιπέδου και διατηρούν την μορφή της σιγμοειδούς, ενώ η σιγμοειδής του κόμβου εξόδου λειτουργεί για να συμπιέσει την ενεργοποίησή του στο διάστημα $(0, 1)$, διατηρώντας όμως για την έξοδο y την ίδια βασική μορφή με την τελευταία.



Σχήμα 3-17 Δίκτυο $1 \times 2 \times 1$ για απόδοση προσέγγισης συνάρτησης μίας μεταβλητής.

Διαφοροποίηση των βαρών των κόμβων του ενδιαμέσου επιπέδου αλλάζει την μορφή ως προς την κλίση της σιγμοειδούς εξόδου τους (Σχ. 3-5), ενώ αύξηση της απόστασης των πολώσεων τους απομακρύνει τις καμπύλες εξόδου τους y_1 και y_2 διευρύνοντας την καμπάνα της καμπύλης της τελικής εξόδου y . Διαφοροποίηση, εξάλλου των βαρών του κόμβου εξόδου, μεταβάλλει την απόσταση μεταξύ των επιπέδων των άκρων της καμπύλης εξόδου y , ενώ η αλλαγή της πόλωσης του μεταβάλλει τον βαθμό συμπίεσης της ενεργοποίησης της από την σιγμοειδή εξόδου.



Σχήμα 3-18 Καμπύλες απόκρισης του δικτύου $1 \times 2 \times 1$ για διάφορες τιμές των λειτουργικών παραμέτρων

Είναι δυνατό για διάφορους συνδυασμούς των λειτουργικών παραμέτρων του ίδιου δικτύου (μεταβάλλοντας τις τιμές τους και ως προς τα πρόσημα) να προκύψει ένα πλήθος συναρτήσεων

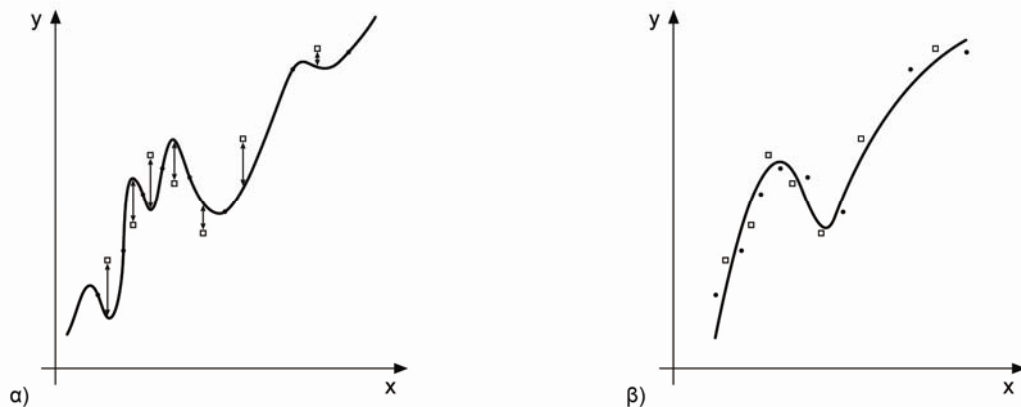
(Σχ. 3-18) που δεν περιορίζεται σε καμπύλες δύο κοίλων (αντίστοιχα με ένα πολυώνυμο 3ου βαθμού), καθιστώντας εμφανή, ήδη από το απλό αυτό δίκτυο, την ευελιξία των δικτύων στην προσέγγιση συνάρτησης μεταξύ μεταβλητών X και Y . Το όριο πολυπλοκότητας της καμπύλης εξόδου εξαρτάται από το πλήθος των λειτουργικών παραμέτρων, που αποτελούν και τους βαθμούς ελευθερίας του συστήματος. Με την προσθήκη ενός επιπλέον κόμβου στο ενδιάμεσο επίπεδο, δηλαδή με αύξηση των παραμέτρων κατά δύο, χωρίς όμως να τροποποιηθεί η βασική μορφή εξίσωσης της συνάρτησης εξόδου (3.12), η καμπύλη μπορεί να καταστεί ακόμη πιο σύνθετη. Επιπλέον ενδιάμεσα επίπεδα αυξάνουν ραγδαία την ευελιξία του συστήματος, διαφοροποιώντας όμως δραστικά και την συνάρτηση απόκρισης με προσθήκη επιπλέον σιγμοειδούς, ενώ η αύξηση της πολυπλοκότητας του δικτύου επηρεάζει και την δυνατότητα γενίκευσής του. Αποδεικνύεται (Funahashi, 1989, Hornik, 1989) ότι η χρήση ενός μεμονωμένου ενδιάμεσου επιπέδου, είναι επαρκής για την προσέγγιση οποιασδήποτε συνεχούς συνάρτησης μεταξύ των X και Y .

3.4 Γενίκευση

Έχει γίνει ήδη αναφορά στην προϋπόθεση των “καλά” εκπαιδευμένων δικτύων για καλή ικανότητα **γενίκευσης** (generalization) στον χώρο του προβλήματος, δηλαδή στην ιδιότητά τους να ταξινομούν ή να προβλέπουν σωστά, εκτός από τα δείγματα στα οποία έχουν εκπαιδευτεί, και κάθε νέο δείγμα εντός του χώρου εκπαίδευσης. Το κυριότερο πρόβλημα που επηρεάζει αρνητικά την ικανότητα γενίκευσης του δικτύου είναι η **υπερεκπαίδευσή** (overtraining) του, που αντιστοιχεί σε υπερπροσαρμογή του συστήματος στα δείγματα του συνόλου εκπαίδευσης. Το δίκτυο προσπαθώντας να υιοθετήσει όλη την διαθέσιμη πληροφορία προσαρμόζει τις παραμέτρους του ώστε να προβλέπει σωστά όλα τα δείγματα εκπαίδευσης σε τέτοιο βαθμό που κινδυνεύει να καταστεί ανίκανο να διερευνήσει τις πραγματικές τάσεις που διέπουν το φαινόμενο.

Στο Σχ. 3-19 παρουσιάζεται ένα χαρακτηριστικό παράδειγμα υπερεκπαίδευσης-ανεπαρκούς ικανότητας γενίκευσης δικτύου που εκπαιδεύτηκε για προσέγγιση συνάρτησης μίας ανεξάρτητης μεταβλητής. Με κύκλους σημειώνονται τα δείγματα εκπαίδευσης, ενώ τα τετράγωνα αντιστοιχούν σε άγνωστα για το δίκτυο δείγματα που αποτελούν το **σύνολο ελέγχου** (test set), τις εξαρτημένες μεταβλητές των οποίων το σύστημα καλείται να προβλέψει μετά την ολοκλήρωση της εκπαίδευσής του. Προκειμένου να μειώσει στο ελάχιστο το σφάλμα, το δίκτυο τείνει να εφαρμόσει μία εξαιρετικά πολύπλοκη καμπύλη (Σχ. 3-19α)) που διέρχεται από όλα τα σημεία του συνόλου εκπαίδευσης, με αποτέλεσμα όμως να εμφανίζει ανεπαρκή ικανότητα γενίκευσης, αφού προβλέπει τα άγνωστα δείγματα με μεγάλο σφάλμα. Αν το δίκτυο δεν είχε εξαναγκαστεί να προσαρμοστεί τόσο καλά στα δείγματα εκπαίδευσης, ώστε να εισάγει πλαστές

διακυμάνσεις στην συνάρτηση παραβλέποντας τις πραγματικές τάσεις στην εξάρτηση εισόδου-εξόδου, θα υλοποιούσε μία απλούστερη καμπύλη (Σχ. 3-19β)). Το υπολειπόμενο σφάλμα πρόβλεψης θα ήταν λίγο μεγαλύτερο για τα δείγματα εκπαίδευσης, η ικανότητα γενίκευσής του όμως θα αυξανόταν, αφού θα διατηρούσε την ακρίβεια πρόβλεψής του και για τα δείγματα ελέγχου.



Σχήμα 3-19 α) Γενίκευση και β) υπερεκπαίδευση στο πλαίσιο προσέγγισης συνάρτησης

Μέγεθος Συνόλου Εκπαίδευσης και Δικτύου

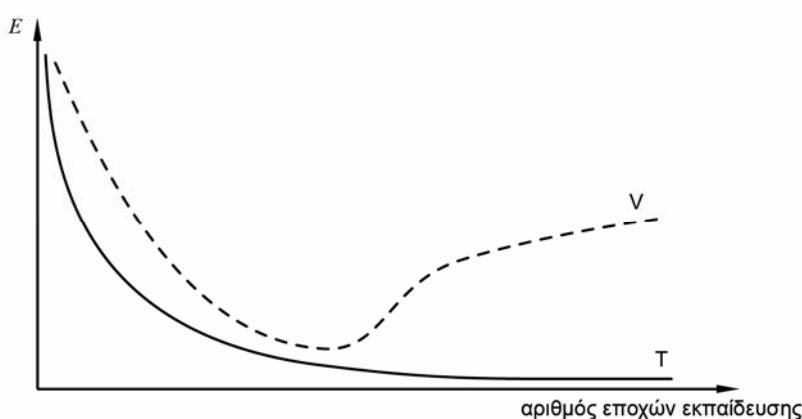
Όπως υπονοήθηκε από το προηγούμενο παράδειγμα, η ικανότητα γενίκευσης επηρεάζεται τόσο από το μέγεθος του συνόλου εκπαίδευσης, όσο και από την πολυπλοκότητα –και επομένως τους βαθμούς ελευθερίας– του δικτύου. Ένα μικρό σύνολο εκπαίδευσης, καθώς και η ποιότητα των δεδομένων του, *περιορίζει* δραστικά το δίκτυο, αφού, εκπαιδευόμενο ώστε συλλάβει όλη την μεταβλητότητα του πίνακα δεδομένων συμπεριλαμβανομένου και του περιεχόμενου θορύβου, οδηγείται σε ψευδή απόδοση της εξάρτησης μεταξύ εισόδων-εξόδων. Ένας αυξημένος αριθμός σημείων εκπαίδευσης στο προηγούμενο παράδειγμα δεν θα επέτρεπε στην καμπύλη προσέγγισης τόσες πολλές κυμάνσεις, διατηρώντας ικανοποιητική γενίκευση του συστήματος. Στο ίδιο αποτέλεσμα θα μπορούσε να οδηγηθεί το δίκτυο με περιορισμό των βαθμών ελευθερίας του, ή αλλιώς περικοπή περιττών ενδιάμεσων κόμβων επεξεργασίας, αναγκάζοντας τους τελευταίους να λάβουν βάρη σύνδεσης κοντά στο μηδέν.

Είναι σαφές επομένως, ότι είναι απαραίτητο για την διασφάλιση της ικανότητας γενίκευσης να μπορεί να προσδιοριστεί ο επαρκής αριθμός δεδομένων εκπαίδευσης σε συνδυασμό με την βέλτιστη αρχιτεκτονική του νευρωνικού δικτύου που χρησιμοποιείται σε κάθε περίπτωση. Επειδή στην πράξη ο αριθμός των δειγμάτων είναι δεδομένος και συχνά περιορισμένος, έχουν αναπτυχθεί διάφορες τεχνικές σχεδιασμού της τοπολογίας του δικτύου είτε με χρήση στατιστικών μεθόδων πολυπαραμετρικής ανάλυσης πριν την εκπαίδευσή του, είτε δυναμικά

κατά την διάρκεια αυτής με προσθήκη ή αφαίρεση ενδιάμεσων μονάδων. Στην ίδια λογική ελέγχου της ελευθερίας του συστήματος κινούνται και οι μέθοδοι περικοπής του δικτύου, που αντί να μεταβάλλουν την αρχιτεκτονική του θέτουν περιορισμούς για τις παραμέτρους λειτουργίας του.

Σύνολο επαλήθευσης

Μέθοδοι όπως αυτές που αναφέρθηκαν είναι συχνά πολύπλοκες, επιβαρημένες υπολογιστικά και επομένως χρονοβόρες για τις συνήθεις εφαρμογές. Στην πραγματικότητα δεν είναι πάντα απαραίτητο να καθοριστεί ο βέλτιστος αριθμός ενδιάμεσων κρυφών κόμβων, αλλά η διασφάλιση καλής γενίκευσης μπορεί να πραγματοποιηθεί με την χρήση ενός συνόλου γνωστών δειγμάτων, που δεν συμμετέχουν άμεσα στην εκπαίδευση του δικτύου, αλλά χρησιμοποιούνται παράλληλα με αυτήν για την αξιολόγηση της λειτουργίας του. Το σύνολο αυτό αναφέρεται ως **σύνολο επαλήθευσης** (validation set) και χρησιμοποιείται αντίστοιχα με ένα σύνολο ελέγχου κατά την διάρκεια όμως της εκπαίδευσης, με υπολογισμό του σφάλματος πρόβλεψής του σε κάθε εποχή.



Σχήμα 3-20 Σφάλμα πρόβλεψης των συνόλων εκπαίδευσης και επαλήθευσης κατά την διάρκεια της εκπαίδευσης

Σε αρχικά και επόμενα στάδια της εκπαίδευσης η λειτουργία του δικτύου αντιστοιχεί σε αυτή του Σχ. 3-19β) και μόνο μετά από παρατεταμένη εκπαίδευσης το δίκτυο διαμορφώνει συμπεριφορά ανάλογη του Σχ. 3-19α). Είναι δυνατό λοιπόν να διασφαλιστεί επαρκής ικανότητα γενίκευσης αν διακοπεί η εκπαίδευση του δικτύου σε κατάλληλο στάδιο, εφαρμόζοντας την μέθοδο *πρώωρης διακοπής*. Τα διαθέσιμα δεδομένα χωρίζονται σε σύνολο εκπαίδευσης T και σύνολο επαλήθευσης V (με αναλογία περίπου 10:1), το δίκτυο εκπαιδεύεται στο σύνολο T , ενώ προσδιορίζεται το σφάλμα και των δύο συνόλων κατά την διάρκεια της εκπαίδευσης. Μία χαρακτηριστική συμπεριφορά δικτύου παρουσιάζεται στο Σχ. 3-20. Κριτήριο για παύση της

εκπαίδευσης είναι η ελάχιστη απόκλιση των δύο σφαλμάτων, στο στάδιο δηλαδή όπου η γενίκευση δικτύου ως προς τα άγνωστα για το δίκτυο δείγματα του V είναι βέλτιστη. Το σφάλμα εκπαίδευσης του δικτύου πέραν του σημείου αυτού συνεχίζει να μειώνεται ακολουθώντας τον αλγόριθμο καθόδου κλίσης, εις βάρος όμως της ικανότητας γενίκευσης του δικτύου.

4 Σχεδιασμός Πειράματος

Η επεξεργασία δεδομένων με χημειομετρικές μεθόδους μπορεί να αναδείξει σημαντική ποσότητα χρήσιμης πληροφορίας, απαλλαγμένη από το μεγαλύτερο ποσοστό του πειραματικού θορύβου και να αποκαλύψει τις σε πρώτη εξέταση μη προφανείς αλληλεξαρτήσεις που υφίστανται μεταξύ των μεταβλητών. Βασική προϋπόθεση είναι η ύπαρξη πληροφορίας στα διαθέσιμα δεδομένα, που, εκτός από τις περιπτώσεις συγκέντρωσής τους από μεγάλους πληθυσμούς καταγραφών παραγωγής ή διαδικασίας, συλλέγονται ακολουθώντας ένα πειραματικό πλάνο που έχει καθορισμένους στόχους και προσδοκά στην εξαγωγή έγκυρων και αντικειμενικών συμπερασμάτων. Η επιβεβαίωση μίας θεωρίας ή η αναζήτηση των αρχών που διέπουν μία φυσική διεργασία και ο καθορισμός των σημαντικότερων παραμέτρων της, η μοντελοποίηση μέσω αναλυτικών σχέσεων της εν λόγω διεργασίας, και τέλος ο προσδιορισμός των τιμών των παραμέτρων που βελτιώνουν την συμπεριφορά ενός συστήματος, απαιτούν για την αποτελεσματικότητα της ανάλυσης την συστηματική συλλογή δεδομένων. Ο **στατιστικός πειραματικός σχεδιασμός** ή **σχεδιασμός πειράματος** (statistical experimental design or design of experiment, DoE) αποτελεί την μεθοδολογία κατάστρωσης πλάνου διεξαγωγής πειραμάτων με στόχο την εξαγωγή της μέγιστης δυνατής πληροφορίας με εφαρμογή του ελάχιστου αριθμού πειραματικών δοκιμών. Η έννοια του σχεδιασμού πειράματος εισήχθει από τον Fisher στην δεκαετία του '20 και περιλαμβάνει ένα σύνολο μεθόδων για την εκ των προτέρων διαχείριση και επιλογή των συνθηκών διεξαγωγής των πειραματικών δοκιμών, ανάλογα με τις απαιτήσεις της ανάλυσης που πρόκειται να εφαρμοστεί στα δεδομένα.

Ένα στατιστικά σχεδιασμένο πείραμα αποτελείται από ένα σύνολο πειραματικών δοκιμών που είναι οι βέλτιστες δυνατές για την κάλυψη των αναγκών της ζητούμενης ανάλυσης. Αφού καθοριστούν οι αντικειμενικοί στόχοι και οι μεταβλητές που επηρεάζουν τις ποιοτικές παραμέτρους της διεργασίας, επιλέγεται το είδος του **μοντέλου διεργασίας** (process model) που θα χρησιμοποιηθεί, δηλαδή η μορφή της μαθηματικής σχέσης για την περιγραφή των αλληλεξαρτήσεων που υφίστανται μεταξύ δύο τουλάχιστον παραμέτρων, εκ των οποίων η μία ορίζεται ως ανεξάρτητη και η δεύτερη ως εξαρτημένη. Στον πειραματικό σχεδιασμό οι ανεξάρτητες μεταβλητές –ποιοτικές ή ποσοτικές–, αναφέρονται ως **παράγοντες** (factors) και οι μεταβολές τους καθορίζονται από τον ερευνητή, ενώ το αποτέλεσμα κάθε μεταβολής τους, που αντικατοπτρίζει την μεταβολή των εξαρτημένων μεταβλητών, αναφέρεται ως **απόκριση** (response) του συστήματος. Κατά την διάρκεια μίας σειράς πειραματικών δοκιμών οι ανεξάρτητες μεταβλητές παίρνουν τιμές εντός κάποιων ορίων ανάλογα με την φύση της μεταβλητής και τον στόχο του πειράματος, οι οποίες αναφέρονται ως **επίπεδα** ή **στάθμες** (levels). Η διαφορά μεταξύ δύο αποκρίσεων που προκύπτει από δύο διαφορετικά επίπεδα ενός

παράγοντα αναφέρεται ως **αποτέλεσμα** (effect) της συγκεκριμένης μεταβολής. Τέλος, ως **πειραματική μονάδα** (experimental unit or run) ορίζεται κάθε πειραματική δοκιμή που περιλαμβάνει μία και μοναδική επέμβαση, διαφορετική κάθε φορά, στις τιμές των μεταβλητών που υπεισέρχονται στο σύστημα μελέτης. Μία από τις βασικότερες επιδιώξεις του στατιστικού πειραματικού σχεδιασμού είναι η μείωση του αριθμού των παραγόντων, περιορίζοντάς τους σε αυτούς που έχουν άμεση και σημαντική επίδραση στην εξέλιξη της διεργασίας, και ο προσδιορισμός των βέλτιστων επιπέδων τους για την παραγωγή δεδομένων με την μέγιστη δυνατή χρήσιμη πληροφορία.

4.1 Είδη πειραμάτων

Σε κάθε εφαρμοσμένο πρόβλημα όπου απαιτείται η συγκέντρωση δεδομένων από πειραματικές μετρήσεις, είναι απαραίτητος ο καθορισμός των στόχων και η επιλογή του τρόπου προσέγγισης. Ανάλογα με την ζητούμενη λύση του προβλήματος, τα πειράματα χωρίζονται γενικά σε δύο κατηγορίες: τα **συγκριτικά** (comparative) και τα **παραγοντικά** (factorial). Τα συγκριτικά πειράματα χρησιμοποιούνται για την σύγκριση των αποτελεσμάτων διαφορετικών δοκιμών, με μεταβολή μόνο κάποιων εκ των παραμέτρων (συνήθως του ίδιου δείγματος), στην απόκριση του συστήματος. Ο στόχος περιορίζεται σε επιλογή μεταξύ ορισμένων διαθέσιμων εναλλακτικών λύσεων και υλοποιείται με εκλεκτικές δοκιμές για την αξιολόγηση των αποτελεσμάτων ενός ή περισσότερων παραγόντων. Δηλαδή εξετάζεται η απόκριση είτε κατά την μεταβολή μεταξύ δύο ή περισσότερων επιπέδων ενός μεμονωμένου παράγοντα (π.χ. μία χημική αντίδραση εξελίσσεται πιο γρήγορα σε χαμηλές, μέσες ή υψηλές θερμοκρασίες), είτε κατά την αντικατάσταση ενός παράγοντα από έναν άλλο (π.χ. μία ένωση παρουσιάζει ισχυρότερη καταλυτική δράση από μία άλλη για δεδομένη αντίδραση). Αντίθετα στα παραγοντικά πειράματα, στόχος είναι η διερεύνηση της συμπεριφοράς του συστήματος συνολικά, δηλαδή οι μεταβολές της απόκρισής του κατά την ταυτόχρονη μεταβολή διαφορετικών παραγόντων.

Τύποι Παραγοντικών Πειραμάτων

Τα πεδία χρήσης των παραγοντικών πειραμάτων περιλαμβάνουν μεταξύ άλλων την διερεύνηση της δομής και των παραμέτρων που καθορίζουν την συμπεριφορά νέων άγνωστων συστημάτων, την προσομοίωση, μέσω κάποιου μοντέλου παλινδρόμησης, μίας φυσικοχημικής διεργασίας για ακριβή ποσοτική πρόβλεψη της απόκρισης ή την ανάπτυξη ενός μοντέλου προσέγγισης της ευρύτερης επιφάνειας απόκρισης. Ο στόχος στην μελέτη της επιφάνειας απόκρισης του συστήματος είναι η ρύθμιση των επιπέδων των παραγόντων επιρροής του συστήματος για την επίτευξη απόκρισης-στόχου (π.χ. ρύθμιση παραμέτρων αναλυτικού οργάνου κατά την βαθμονόμησή του με βάση πρότυπο δείγμα, καθορισμός της αναλογίας των

συστατικών ενός μίγματος), για την ελαχιστοποίηση ή μεγιστοποίηση της απόκρισης (π.χ. μεγιστοποίηση της παραγωγής, ελαχιστοποίηση της ποσότητας των υποπροϊόντων ή του κόστους παραγωγής), για περιορισμό της εσωτερικής διασποράς της απόκρισης που οφείλεται σε δύσκολα ελεγχόμενες παραμέτρους (π.χ. μη σταθερές ποσότητες ή ιδιότητες των πρώτων υλών τροφοδοσίας μίας μονάδας παραγωγής), για ανάπτυξη της σταθερότητας της απόκρισης στις μεταβολές των παραμέτρων (π.χ. μία μηχανή εσωτερικής καύσης θα πρέπει να ανταποκρίνεται επιτυχώς στις δοκιμές αντοχής υπό τεταμένες μεταβαλλόμενες συνθήκες), ή τέλος για την συνδυαστική κάλυψη των επιθυμητών απαιτήσεων, που συνήθως αφορούν περισσότερες από μία ιδιότητες-στόχους. Τα τελευταία αποτελούν προβλήματα βελτιστοποίησης ή αριστοποίησης μίας εργαστηριακής ή βιομηχανικής διεργασίας.

Για την επίλυση τέτοιου είδους προβλημάτων χρησιμοποιούνται **τεχνικές σχεδιασμού πειραμάτων βελτιστοποίησης** (optimizing experimental designs), που προϋποθέτουν όμως την προηγούμενη ποιοτική και ποσοτική περιγραφή του συστήματος. Είναι φανερό λοιπόν, ότι η έννοια του παραγοντικού σχεδιασμού πειράματος εμπεριέχει, εκτός από την ταυτόχρονη και συστηματική μεταβολή όλων των παραγόντων, και την **διαδοχική ή επαναληπτική προσέγγιση** (sequential or iterative approach) της πειραματικής διαδικασίας, κατά την οποία σχεδιάζεται μία σειρά πειραμάτων, που διεξάγονται σταδιακά και ακολουθώντας ιεραρχική πορεία ως προς τους στόχους τους, και όπου η αναγκαιότητα, το είδος και η πορεία του επόμενου πειράματος καθορίζεται από τα αποτελέσματα της ανάλυσης του προηγούμενου.

Τα **διερευνητικά πειράματα** (screening experimental designs), που ο άμεσος στόχος τους, στα πλαίσια της κατανόησης της διεργασίας, είναι να απομονώσουν από το σύνολο των πιθανών μεταβλητών τις λιγότερο σημαντικές (δεν περιέχουν ή περιέχουν μη σημαντική πληροφορία) και να ιεραρχήσουν τους κύριους παράγοντες επιρροής διεξάγοντας ένα μικρό αριθμό πειραματικών δοκιμών, πραγματοποιούνται συνήθως στα πρώτα στάδια της έρευνας. Η ανάλυση και ερμηνεία των δεδομένων τέτοιων διερευνητικών πειραμάτων οδηγούν σε μία προσεγγιστική ποσοτική εκτίμηση της επίδρασης των παραγόντων, επιτρέποντας, κατά τον πειραματικό σχεδιασμό της επόμενης φάσης μελέτης, μία ακόμη πιο αυστηρή επιλογή μεταξύ των ισχυρότερων ως προς το αποτέλεσμα, με περιορισμό παράλληλα του εύρους του πειραματικού χώρου. Αξιοποιώντας την χρήσιμη αυτή πληροφορία, ο σχεδιασμός των **πειραμάτων μοντελοποίησης** (modeling experimental designs), που ακολουθούν, προσβλέπει στην ανάπτυξη ενός λειτουργικού μοντέλου της διαδικασίας με την μορφή μίας υψηλής προβλεψιμότητας της απόκρισης μαθηματικής συνάρτησης, οι συντελεστές της οποίας προσδιορίζονται με μεγάλη ακρίβεια. Ένα τέτοιο μοντέλο είναι αναγκαίο για την πλήρη διερεύνηση του φαινομένου που μελετάται και ικανό να αποδώσει την επίδραση των μεταβλητών που υπεισέρχονται στο πρόβλημα, αλλά και τις αλληλεπιδράσεις μεταξύ τους,

εξυπηρετώντας έτσι στον ακόλουθο σχεδιασμό του πλάνου όλων των απαραίτητων δοκιμών για τον καθορισμό του βέλτιστου συνδυασμού των επιπέδων των παραγόντων, κατά το στάδιο των πειραμάτων βελτιστοποίησης.

Εναλλακτικές Τακτικές Πειραματισμού

Βέβαια εκτός από τον στατιστικό πειραματικό σχεδιασμό, οι αρχές και τεχνικές του οποίου δεν διαφοροποιούνται σημαντικά για τις παραπάνω κατηγορίες, εξαιρετικά διαδεδομένες είναι και οι απλούστερες τακτικές πειραματισμού της **ad hoc προσέγγισης** και της **μεταβολής ενός μόνο παράγοντα κάθε φορά**. Στην πρώτη περίπτωση εφαρμόζεται μη συστηματική μεταβολή των παραγόντων, με πειραματικές δοκιμές που δεν έχουν εκ των προτέρων προγραμματιστεί, αλλά ακολουθούν μία πορεία, που βασιζόμενη στην διαθέσιμη γνώση και εμπειρία, αναπροσαρμόζεται κάθε φορά ανάλογα με τα προκύπτοντα αποτελέσματα. Στην ίδια λογική απλοποίησης κινείται και η πιο μεθοδική τακτική της μεταβολής των επιπέδων ενός μόνο παράγοντα, διατηρώντας κατά την διάρκεια του πειράματος σταθερές όλες τις υπόλοιπες μεταβλητές και εξαντλώντας την μελέτη επιρροής ενός μεμονωμένου παράγοντα, προτού προχωρήσει σε δοκιμές διερεύνησης της επίδρασης των υπολοίπων. Η βηματική αυτή προσέγγιση είναι αρκετά ελκυστική λόγω της απλότητας του πλάνου των πειραματικών δοκιμών, έναντι της σύγχυσης που δημιουργεί η πολυπλοκότητα του στατιστικού σχεδιασμού, και της αντικατάστασης του πολυπαραμετρικού προβλήματος από πλήθος μονοπαραμετρικών, που πλεονεκτούν στην ευκολία ερμηνείας των αποτελεσμάτων.

Η θεώρηση όμως μονοδιάστατων χώρων, αποκλείει *a priori* την πιθανότητα αλληλεπιδράσεων μεταξύ των παραγόντων, που αποτελούν μάλλον κανόνα για τα πραγματικά συστήματα, αδυνατώντας έτσι να διερευνήσει το συνολικό πειραματικό χώρο, καθώς και να προβλέψει την συμπεριφορά του συστήματος για τιμές των παραγόντων εκτός των ορίων που χρησιμοποιήθηκαν στο συγκεκριμένο πείραμα. Το μειονέκτημα αυτό, που εντείνεται ακόμα περισσότερο στον τυχαίο *ad hoc* πειραματισμό, παράλληλα με την αδυναμία τέτοιων τακτικών να προσδιορίσουν εκ των προτέρων τον αριθμό των απαιτούμενων δοκιμών, με κίνδυνο να γίνει απαγορευτικά υψηλός, καθιστούν τις τελευταίες αναποτελεσματικές στην επίτευξη των στόχων που προαναφέρθηκαν. Μη συστηματικοί σχεδιασμοί πειράματος μπορούν ίσως να καλύψουν τις απαιτήσεις μίας συγκριτικής μελέτης, αλλά η χρήση τους είναι ακατάλληλη σε περισσότερο απαιτητικά προβλήματα και ειδικότερα σε προβλήματα βελτιστοποίησης, όπου τα αποτελέσματά τους, εκτός από ανεπαρκή, μπορούν συχνά να οδηγήσουν σε λανθασμένα συμπεράσματα.

4.2 Στρατηγική Υλοποίησης Πειραματικού Σχεδιασμού

Η πολυπαραμετρική ανάλυση στοχεύει στην εξαγωγή της μέγιστης δυνατής πληροφορίας από τον πίνακα δεδομένων, η επεξεργασία των οποίων καθίσταται πολύ αποδοτικότερη, αν ο εν λόγω πίνακας είναι στατιστικά σχεδιασμένος, έτσι ώστε να εξυπηρετεί τον κύριο στόχο της μελέτης. Οι βασικές επιδιώξεις του πειραματικού σχεδιασμού είναι η αύξηση της *αποδοτικότητας* της πειραματικής διεργασίας και η *εστίασή* της στην ανάκτηση αποκλειστικά χρήσιμης πληροφορίας. Τόσο η μείωση των απαραίτητων δοκιμών, όσο και ο περιορισμός τους εντός της περιοχής ενδιαφέροντος, έχουν σαν αποτέλεσμα την παραγωγή ενός “δομημένου” πίνακα δεδομένων, που περιέχει δηλαδή σημαντικό ποσοστό εσωτερικής μεταβλητότητας, η οποία και θα αποτελέσει την βάση για ένα σταθερό και ακριβές πολυπαραμετρικό μοντέλο, ενώ παράλληλα οδηγούν σε μείωση των δαυστάσεων του, καθιστώντας τον πιο εύχρηστο για στατιστική επεξεργασία. Η διασφάλιση της καταλληλότητας για την μετέπειτα ανάλυση του πίνακα μετρήσεων προϋποθέτει τον εκ των προτέρων προσδιορισμό των αντικειμενικών στόχων του πειραματισμού, ορισμό του πεδίου μελέτης και καθορισμό των κατευθυντήριων χειρισμών κατά την κατάσχεση του πλάνου δοκιμών.

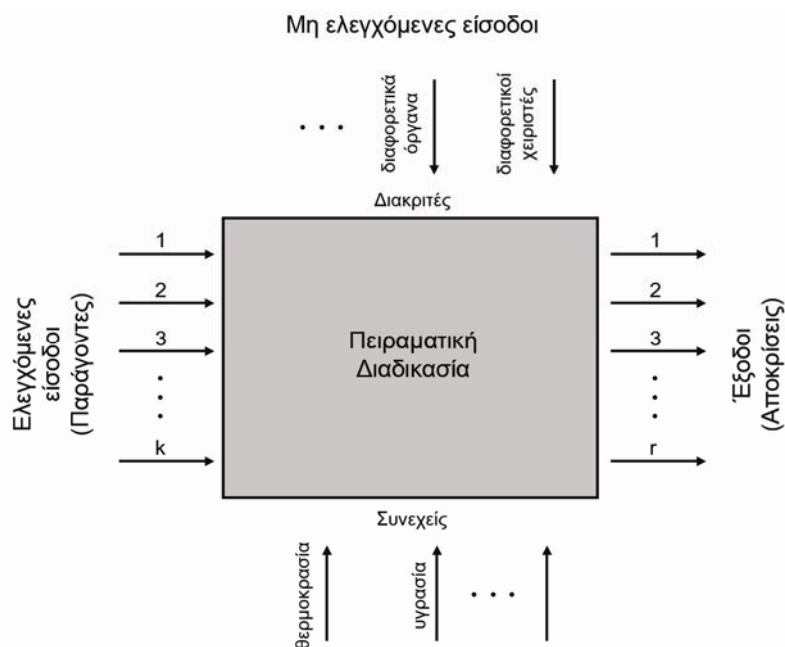
Για την κάλυψη των ζητούμενων σε ένα πείραμα και αφού έχει ολοκληρωθεί το στάδιο μίας πρώτης εξοικείωσης με την εξεταζόμενη διεργασία, που περιλαμβάνει παρατήρηση, συνήθως με την διεξαγωγή κάποιων αναγνωριστικών δοκιμών, συγκέντρωση της προηγούμενης εμπειρικής ή αναλυτικής γνώσης και επισημάνση των υποσχόμενων δυνατοτήτων της έρευνας-σχεδιασμού, ακολουθείται ένα διάγραμμα ροής των απαραίτητων βημάτων σχεδιασμού, εκτέλεσης και ανάλυσης του παραγοντικού πειράματος. Αν και είναι αρκετά ελαστικό σε πιθανές τροποποιήσεις, που εξαρτώνται από το δεδομένο πρόβλημα, σε γενικές γραμμές περιλαμβάνει τα εξής στάδια:

1. Προσδιορισμός των αντικειμενικών στόχων του πειράματος
2. Επιλογή των παραγόντων που θα αποτελέσουν τις ανεξάρτητες μεταβλητές και των πεδίων τιμών τους, παράλληλα με τον καθορισμό των ανεξέλεγκτων μεταβλητών-συμπαράγοντων που πιθανώς επιδρούν στην διεργασία
3. Επιλογή των εξαρτημένων μεταβλητών που θα αντιπροσωπεύουν την απόκριση του συστήματος
4. Επιλογή του μοντέλου για την μαθηματική περιγραφή της διεργασίας που ερευνάται
5. Επιλογή εκείνης της τεχνικής σχεδιασμού που παρουσιάζει μεγαλύτερη συμβατότητα με τον στόχο του πειράματος και καθορισμός του αριθμού των παραγόντων και της απαιτούμενης ακρίβειας των μετρήσεων.

6. Επιλογή των πειραματικών σημείων για την όσο το δυνατόν πλήρη και συμμετρική κάλυψη του πειραματικού χώρου και έλεγχος της επαρκούς κάλυψής του για τα σημεία που επιλέχθηκαν.
7. Εκτέλεση πειραμάτων
8. Ανάλυση των αποτελεσμάτων, έλεγχος της συνέπειας των δεδομένων με τις υποθέσεις του πειράματος και ανάπτυξη του μοντέλου
9. Οπτικοποίηση και ερμηνεία των αποτελεσμάτων του μοντέλου
10. Χρήση του μοντέλου (εφαρμογή ή σχεδιασμός νέου πειράματος)

Μεταβλητές Διεργασίας

Αφού προσδιοριστεί ο στόχος και τα περιγραφικά χαρακτηριστικά του πειραματισμού, που απαντούν σε θέματα, όπως ποια η διαθέσιμη γνώση για το σύστημα μελέτης, τα κύρια ζητούμενα της διερεύνησης, οι μεταβλητές που σχετίζονται με το πρόβλημα, καθώς και ποιες από αυτές θα αντιπροσωπεύονται στην μετρούμενη απόκριση, διαμορφώνεται ένα λογικό διάγραμμα μοντελοποίησης της διαδικασίας, που αποτελεί και την βάση για τον σχεδιασμό του πειράματος (Σχ.°4-1). Ο σκοπός του είναι να σχηματοποιηθεί η λειτουργία του αναμενόμενου μοντέλου περιγραφής του συστήματος, το οποίο, προτού αποφασιστεί το είδος της μαθηματικής συνάρτησης, βοηθά στον καθορισμό, τόσο των μεταβλητών που υπεισέρχονται στην διεργασία με την μορφή εισόδων, όσο και εκείνων που θα αποτελέσουν την παραγόμενη έξοδο-απόκριση.



Σχήμα 4-1 Λογικό διάγραμμα σχεδιασμού για την εποπτική παρουσίαση του μοντέλου της πειραματικής διεργασίας

Η εκτέλεση ενός στατιστικού πειράματος βασίζεται στον έλεγχο των μεταβολών των μεταβλητών εκείνων, των οποίων τα αποτελέσματα ενδιαφέρουν και θεωρείται ότι έχουν άμεση συσχέτιση με την απόκριση. Η επιλογή τους γίνεται κατά την κρίση του ερευνητή με βάση την υφιστάμενη γνώση, ενώ η γενική πρακτική είναι αρχικά να εισάγεται ένας ικανός αριθμός τους (ίσως περισσότερες από τις απαραίτητες) κατά τα διερευνητικά πειράματα, που μειώνεται κατά τα επόμενα στάδια των σχεδιασμών πρόβλεψης και βελτιστοποίησης. Αποτελούν τις ελεγχόμενες εισόδους του συστήματος, αλλιώς **μεταβλητές σχεδιασμού** (*design variables*) ή παράγοντες, και μπορούν να είναι είτε *ποσοτικές* –συνεχείς ή διακριτές (π.χ. θερμοκρασία, πίεση, συγκέντρωση % κ.ο. συστατικών μίγματος, pH, κ.α.), είτε *κατηγορικές* (π.χ. χρώμα, είδος καταλύτη χημικής αντίδρασης, χρήση ή μη ενός βελτιωτικού προσθέτου, κ.α.). Το εύρος διακύμανσης κάθε παράγοντα υφίσταται τόσο φυσικούς περιορισμούς, εγγενείς της φυσικοχημικής διεργασίας, όσο και εξωγενείς που επιβάλλονται για τεχνολογικούς ή οικονομικούς λόγους, ενώ τα επίπεδα των τιμών τους καθορίζονται ανάλογα με την επιθυμητή πυκνότητα κάλυψης του πειραματικού χώρου. Η συνήθης πρακτική στον πειραματικό σχεδιασμό είναι η χρήση δύο επιπέδων τιμών, μέγιστης και ελάχιστης, για κάθε παράγοντα, που έχει το πλεονέκτημα της διερεύνησης της συμπεριφοράς του συστήματος για όλο το εύρος του πεδίου τιμών των μεταβλητών του με μειωμένο αριθμό δοκιμών, καλύπτει επαρκώς τις απαιτήσεις των περισσότερων προβλημάτων και προσαρμόζεται σε όλες τις τεχνικές σχεδιασμού. Ο λεγόμενος σχεδιασμός δύο επιπέδων επιτρέπει τον αυτόματο υπολογισμό των υπολοίπων απαραίτητων, για κάθε τεχνική, επιπέδων τιμών κατά την πιο λεπτομερειακή διερεύνηση του χώρου, όπως του ενδιάμεσου επιπέδου για κεντρικά δείγματα και επιπέδων εκτός του προκαθορισμένου εύρους για δείγματα αστέρα σε σχεδιασμούς βελτιστοποίησης.

Για λόγους ευκρίνειας του πειραματικού πλάνου και απλοποίησης των υπολογισμών εφαρμόζεται κατάλληλη **κωδικοποίηση** (*coding*) των επιπέδων κάθε παράγοντα που αντικαθιστά τις τιμές του στα διάφορα επίπεδα με τις ανηγμένες τους στο διάστημα $(-1, 1)$. Για ποσοτική μεταβλητή X , με τιμές $\mathbf{x} = \{x_i\} = \{x_1, x_2, \dots, x_n\}$ και $x_i \in [x_a, x_b] = [m-d, m+d]$, όπου $m = (x_a + x_b)/2$, $d = (x_b - x_a)/2$, η αντίστοιχη ανηγμένη ορίζεται από την σχέση:

$$x_{ic} = (x_i - m)/d \Rightarrow x_{ic} = \frac{2x_i - (x_a + x_b)}{x_b - x_a} \quad (4.1)$$

Αν πρόκειται για σχεδιασμό δύο επιπέδων τότε κάθε παράγοντας αντιπροσωπεύεται από δυαδική ανηγμένη μεταβλητή $\{x_{ic}\} = \{-1, +1\}$, ή πιο απλά τα δύο επίπεδα τιμών του συμβολίζονται με $(-, +)$. Ανάλογα κωδικοποιούνται και οι κατηγορικές μεταβλητές, που για τις μεν *ordinal*, με διαβάθμιση δηλαδή στις τιμές τους, η “καλύτερη” και “χειρότερη” ποιότητα συμβολίζονται με $-$ και $+$ αντίστοιχα, ενώ για τις *nominal*, που χρησιμοποιούν τιμές-“ετικέτες”

(π.χ. ηλεκτρικό ρεύμα AC/DC, τεχνητό/φυσικό γλυκαντικό πρόσθετο, κ.α.), ο συμβολισμός (–, +) δίνεται αυθαίρετα. Σε μια τέτοια κωδικοποίηση, το **κύριο αποτέλεσμα** (main effect) κάθε παράγοντα, που αντικατοπτρίζει την μεταβολή της απόκρισης του συστήματος για μεταβολή του παράγοντα από το χαμηλότερο στο υψηλότερο επίπεδό του και ο προσδιορισμός του οποίου λαμβάνει υπ’ όψιν το σύνολο των πειραματικών μετρήσεων αξιοποιώντας *όλη την διαθέσιμη πληροφορία*, υπολογίζεται ως διαφορά των μέσων όρων των αποκρίσεων στο υψηλό (+) και το χαμηλό επίπεδο (–) του παράγοντα:

$$\text{main effect}_x = \bar{y}_{\text{για } x_{ic}=+} - \bar{y}_{\text{για } x_{ic}=-} \quad (4.2)$$

Όλες οι υπόλοιπες μεταβλητές που υπεισέρχονται σε ένα πρόβλημα πειραματικού σχεδιασμού χαρακτηρίζονται ως **μεταβλητές εκτός σχεδιασμού** (non-design variables) και περιλαμβάνουν, εκτός από τις μετρούμενες αποκρίσεις του συστήματος που αντικατοπτρίζουν το αποτέλεσμα των μεταβολών των παραγόντων, και άλλους αστάθμητους παράγοντες που επεμβαίνουν στην διεργασία. Πρόκειται για μεταβλητές που επιδρούν παράλληλα με τους παράγοντες σχεδιασμού στην απόκριση του συστήματος, αλλά τα αποτελέσματά τους εμφανίζονται ως σφάλμα στην έξοδο, αφού δεν έχουν συμπεριληφθεί εξ’ αρχής στον πειραματικό σχεδιασμό ως ανεξάρτητες μεταβλητές της διεργασίας. Επομένως είναι ουσιαστικό να ληφθούν υπόψη οι εν δυνάμει επιδράσεις τους και να περιοριστούν στο ελάχιστο, διαφορετικά αυξάνεται σημαντικά ο περιεχόμενος στα πειραματικά δεδομένα θόρυβος και αναιρείται η βασική επιδίωξη του στατιστικού σχεδιασμού πειράματος για *εστίαση* σε χρήσιμη πληροφορία. Οι μη ελεγχόμενες εισαγωγές στην πειραματική διεργασία είναι είτε μεταβλητές που θεωρήθηκαν επουσιώδεις (δεν συμπεριλαμβάνονται στους κύριους παράγοντες επίδρασης) κατά την επιλογή των παραγόντων, είτε ανεξέλεγκτες παράμετροι του πειράματος που επηρεάζουν κατά συστηματικό τρόπο την διαδικασία.

Ακόμη και σε πειράματα που εκτελούνται υπό εργαστηριακές συνθήκες, γεγονός που επιτρέπει καλύτερο έλεγχο επί της διαδικασίας, είναι πολλές φορές αδύνατη η απαλοιφή μεταβλητών που δεν εισάγονται στον σχεδιασμό. Για να περιοριστεί η επίδραση τους στην απόκριση του συστήματος, υπεισέρχονται στον σχεδιασμό ως **σταθερές συνθήκες** (constant conditions) του πειράματος. Η διατήρηση σταθερών επιπέδων μεταβλητών, όπως σταθερή θερμοκρασία και υγρασία περιβάλλοντος, σταθερές ρυθμίσεις αναλυτικού οργάνου, τυποποίηση διαδικασίας και εκτέλεση των δοκιμών από το ίδιο άτομο, δεν εξαλείφει μεν το αποτέλεσμά τους στην απόκριση, αλλά μεταβάλλει τις μετρήσεις πάντα κατά τον ίδιο τρόπο (κατά σταθερή τιμή), ώστε να εμφανίζουν μηδενική επίδραση στην διακύμανση της μεταβλητής εξόδου. Ανάλογος χειρισμός υιοθετείται και για μεταβλητές που σχετίζονται άμεσα με το φαινόμενο, αλλά δεν

συμπεριλαμβάνονται στο σύνολο των παραγόντων επίδρασης, είτε γιατί το αποτέλεσμα τους βρίσκεται στα όρια του πειραματικού σφάλματος¹, είτε γιατί θεωρήθηκαν επουσιώδεις για τους στόχους της δεδομένης διερεύνησης. Μία από τις χρήσεις των προαναφερθέντων συγκριτικών πειραμάτων, είναι και η επιλογή του κατάλληλου σταθερού επιπέδου μίας τέτοιας μεταβλητής, ώστε να διασφαλίζεται η ομαλή πορεία του φαινομένου (π.χ. μέση θερμοκρασία που να μην τροποποιεί δραματικά τον μηχανισμό χημικής αντίδρασης) ή να βελτιώνεται η παραγωγή της εξόδου του συστήματος (π.χ. επιλογή διαλύτη πλύσης κυψελίδας φασματοφωτόμετρου με την μικρότερη αλληλεπίδραση με τα δραστικά συστατικά μελετούμενου μίγματος βενζίνης). Τέλος, αν η σταθεροποίηση μίας μεταβλητής με υπολογίσιμη επίδραση είναι ανέφικτη, όπως συμβαίνει συχνά σε βιομηχανικές διεργασίες, είναι απαραίτητο να επανεξεταστεί η εισαγωγή της στους παράγοντες σχεδιασμού.

Μία άλλη τεχνική ελέγχου, τόσο των μη σημαντικών γνωστών παραγόντων, όσο και των άγνωστων εισόδων-παραμέτρων του πειράματος, που μπορεί να παράγουν συστηματικά, ανάλογα με τον χρόνο (π.χ. συγκέντρωση ρύπων στον ανιχνευτή αναλυτικού οργάνου στην διάρκεια μίας σειράς πειραματικών δοκιμών, σταδιακή εξοικείωση του ερευνητή κατά την διεξαγωγή του πειράματος), ή τυχαία σφάλματα (π.χ. ανεπιθύμητες προσμίξεις σε παρτίδες πρώτων υλών παραγωγής ενός βιομηχανικού προϊόντος), είναι η **τυχαιοποίηση** (randomization) της σειράς διεξαγωγής των προκαθορισμένων δοκιμών. Το πειραματικό σφάλμα που παράγεται από τις μη ελεγχόμενες παραμέτρους του συστήματος πρέπει να είναι τυχαίο και ανεξάρτητο. Η εκτέλεση των δοκιμών με τυχαία σειρά, αντί της κανονικής ταξινόμησης (σύμφωνα με τα επίπεδα τιμών των παραγόντων), με την οποία σχεδιάζονται αρχικά, εξαλείφει την υπέρθεση της επίδρασης της μεταβολής με τον χρόνο επί του πραγματικού αποτελέσματος μιας άγνωστης παραμέτρου, που έχει σαν συνέπεια παρερμηνεία κατά την ανάλυση των δεδομένων. Αντί της πλήρους τυχαιοποίησης σε όλο το σύνολο των δοκιμών, όταν η εφαρμογή της καθίσταται αδύνατη ως προς συγκεκριμένες παραμέτρους (π.χ. πολύπλοκες ρυθμίσεις αναλυτικού οργάνου για διαφορετικά οργανικά δείγματα, χρονοβόρα παρασκευή δείγματος, παράλληλη χρήση περισσότερων μηχανών δοκιμής, αδύνατη εναλλαγή ερευνητών όταν πρόκειται για περισσότερους του ενός, κ.α.), χρησιμοποιείται η τεχνική της

¹ Οι επαναληπτικές δοκιμές σε κεντρικό σημείο ή σημείο αναφοράς του πειραματικού χώρου (και όχι απλά επαναληπτικές μετρήσεις στο ίδιο δείγμα για τον έλεγχο επαναληπτικότητας του αναλυτικού οργάνου, του χειριστή, κ.τ.λ.), που εμπεριέχονται σε έναν στατιστικό σχεδιασμό για την εκτίμηση του *συνολικού πειραματικού σφάλματος*, επιτρέπουν την σύγκριση της εξηγήσιμης διακύμανσης της απόκρισης, που οφείλεται στις ελεγχόμενες μεταβλητές, με την τυχαία διακύμανση των μη ελεγχόμενων παραμέτρων. Αν οι τιμές των τελευταίων προκύπτουν της ίδιας τάξης μεγέθους, τότε οι μεταβολές των μεταβλητών εξόδου δεν μπορούν να συσχετιστούν με την εξεταζόμενη μεταβλητή.

ομαδοποίησης (blocking) των δοκιμών με βάση τα επίπεδα μίας ή περισσότερων τέτοιων δύσκολα ρυθμιζόμενων παραμέτρων. Η μεταβλητή αυτή λειτουργεί ως **παράγοντας ομαδοποίησης** (blocking factor), αν κάθε επίπεδο του κύριου παράγοντα σχεδιασμού εμφανίζεται για ίδιο αριθμό δοκιμών σε κάθε ομάδα-επίπεδο της εν λόγω μεταβλητής. Τα δεδομένα κάθε τέτοιας ομογενούς ομάδας δοκιμών επεξεργάζονται χωριστά για την ανάδειξη και απαλοιφή του συστηματικού σφάλματος που οφείλεται στον παράγοντα ομαδοποίησης, ενώ για τον περιορισμό των παραποιοητικών επιδράσεων των υπολοίπων ανεξέλεγκτων παραμέτρων εφαρμόζεται τυχαία σειρά εκτέλεσης των διαφορετικών ομάδων, καθώς και των δοκιμών εντός της ίδιας ομάδας.

Μοντελοποίηση διαδικασίας

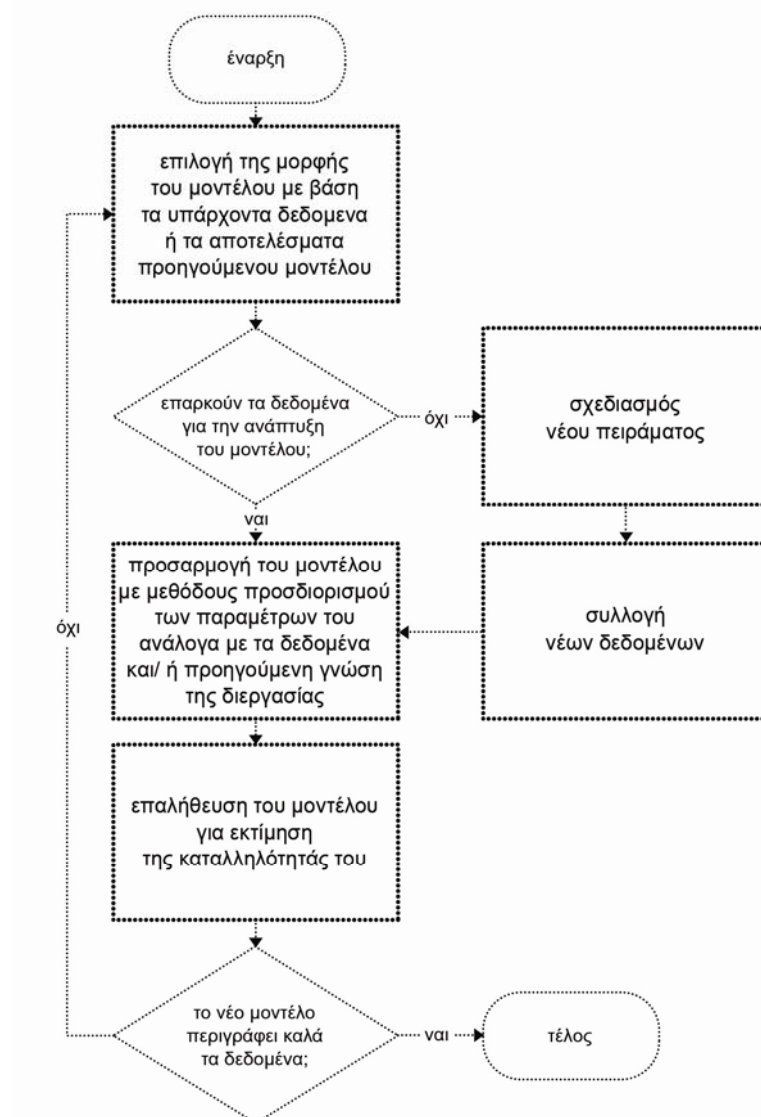
Η πειραματική διερεύνηση συστημάτων είναι κοινή πρακτική τόσο στην εργαστηριακή έρευνα, όσο και στην διαχείριση βιομηχανικών διεργασιών, όπου ανεξάρτητα με το είδος της εφαρμογής και τις ιδιαίτερες επιδιώξεις του προβλήματος, η διενέργεια πειραματικών μετρήσεων είναι απαραίτητη για την ανάπτυξη ενός μαθηματικού μοντέλου περιγραφής του φαινομένου. Ο όρος *μοντελοποίηση της διεργασίας* αναφέρεται στην συνοπτική περιγραφή της συνολικής μεταβλητότητας από μία μεταβλητή Y , που αποτελείται από έναν αιτιοκρατικό όρο με την μορφή μαθηματικής συνάρτησης μίας ή περισσότερων μεταβλητών και έναν τυχαίο όρο που δίδεται από ορισμένη κατανομή πιθανότητας.

$$Y = f(\mathbf{X}, \boldsymbol{\beta}) + \varepsilon \quad (4.3)$$

όπου $\mathbf{X} = \{X_j\} = \{X_1, X_2, \dots, X_k\}$ το διάνυσμα των k παραγόντων πρόβλεψης του σχεδιασμού, $\boldsymbol{\beta} = \{\beta_0, \beta_1, \dots\}$ το διάνυσμα του σταθερού όρου και των παραμέτρων της συνάρτησης, και ε το τυχαίο σφάλμα που απεικονίζει την απόκλιση των δεδομένων από την θεωρούμενη μαθηματική σχέση.

Η ανάπτυξη ενός μοντέλου περιγραφής του διερευνούμενου φαινομένου αποτελεί μία επαναληπτική διαδικασία, που προσαρμόζεται ανάλογα στις ιδιαιτερότητες της ανάλυσης που θα εφαρμοστεί, αλλά ακολουθεί σε γενικές γραμμές τα τρία στάδια **επιλογής** (selection) του μοντέλου, **προσαρμογής** (fitting) του στα δεδομένα και **επαλήθευσης** (validation) του μοντέλου (Σχ. 4-2). Η μορφή της συνάρτησης f σχετίζεται με την φύση της διεργασίας και τις απαιτήσεις του πειραματικού σχεδιασμού, αλλά συνήθως χρησιμοποιούνται πολυωνυμικές συναρτήσεις πρώτου και δευτέρου βαθμού. Είναι απαραίτητη μία αρχική θεώρηση του είδους του μοντέλου που αναμένεται να προκύψει κατά το διενεργούμενο πείραμα, το οποίο μαθηματικοποιεί κατά κάποιον τρόπο τους επιδιωκόμενους στόχους. Ο ελάχιστος απαιτούμενος αριθμός των σημείων του πειράματος (παράγοντες και επίπεδα τιμών τους), που

αντιπροσωπεύεται από τον πίνακα $X_{n \times k}$ (η πειραματικές δοκιμές και k παράγοντες σχεδιασμού) της εξίσωσης (4.3) και επιλέγεται κατά το στάδιο της εφαρμογής της κατάλληλης τεχνικής σχεδιασμού, καθορίζεται από τον αριθμό των αγνώστων παραμέτρων της (4.3). Εκτός από τις βασικές χρήσεις του μοντέλου μιας διεργασίας για πρόβλεψη, βαθμονόμηση και βελτιστοποίηση, είναι επίσης δυνατό να χρησιμοποιηθεί και για μία αρχική **εκτίμηση** (estimation) της συνάρτησης παλινδρόμησης πριν την εκτέλεση των πειραματικών δοκιμών, δίνοντας τυχαίες, κοντά στην πραγματικότητα, τιμές στις αποκρίσεις Y του συστήματος, ώστε να γίνει ένας πρώτος έλεγχος της επαρκούς κάλυψης του πειραματικού χώρου και της λειτουργικότητας του μοντέλου και να καταστεί δυνατή μία πρόωρη, άρα και οικονομικότερη, επανεξέταση των παραμέτρων του σχεδιασμού, αν αυτό κριθεί απαραίτητο.



Σχήμα 4-2 Διάγραμμα ροής ανάπτυξης του μοντέλου της διεργασίας

Η πολυπαραμετρική ανάλυση των μετρήσεων-δεδομένων, που προκύπτουν κατά την εκτέλεση του πειράματος, στοχεύει στον προσδιορισμό των άγνωστων συντελεστών της (4.3) και την εκτίμηση της κατανομής του σφάλματος ε . Το τελευταίο περιλαμβάνει εκτός από το σφάλμα των μετρήσεων και το σφάλμα του μοντέλου, αφού το θεωρούμενο πολυωνυμικό μοντέλο αποτελεί μόνο προσέγγιση της πραγματικής συναρτησιακής σχέσης που διέπει το φαινόμενο και συνδέει την απόκριση με τις ανεξάρτητες μεταβλητές του συστήματος. Γενικά, η εκτίμηση των συντελεστών πραγματοποιείται με αριθμητικές μεθόδους (ελαχίστων τετραγώνων, μέγιστης πιθανοφάνειας), όπου η συνάρτηση βελτιστοποίησης είναι μία αλγοριθμική σχέση των αποκρίσεων και του συναρτησιακού όρου της (4.3), συνήθως κάποια από τις νόρμες του διανύσματος-σφάλματος της προσέγγισης, π.χ.:

$$\|\varepsilon\|_1 = \sum_{i=1}^n |\varepsilon_i|, \|\varepsilon\|_2 = \sqrt{\sum_{i=1}^n |\varepsilon_i|^2}, \|\varepsilon\|_\infty = \max_i |\varepsilon_i| \quad (4.4)$$

όπου $\varepsilon_i = y_i - f(x_{ij}, \hat{\beta})$ το σφάλμα της i -οστής μέτρησης y_i , με $\hat{\beta}$ το διάνυσμα των εκτιμώμενων συντελεστών κατά τις σταδιακές επαναλήψεις του αλγορίθμου μέχρι την εύρεση του άριστου πολυωνύμου προσέγγιση.

Στο αρχικό στάδιο της διεξαγωγής διερευνητικών πειραμάτων χρησιμοποιείται ένα γραμμικό μοντέλο ως προς τους παράγοντες, της μορφής:

$$y_i = \beta_0 + \sum_{j=1}^k \beta_j x_{ijc} + \varepsilon_i \quad (4.5)$$

όπου y_i η μετρούμενη απόκριση της i -στης δοκιμής

x_{ijc} το κωδικοποιημένο επίπεδο του j παράγοντα κατά την i -στη δοκιμή

β_0 ο σταθερός όρος

β_j ο συντελεστής του j παράγοντα του σχεδιασμού που αντικατοπτρίζει το ημιαποτέλεσμα του παράγοντα στην απόκριση

ε_i το σφάλμα μέτρησης της i -στης δοκιμής

Το γενικό μοντέλο της (4.3) γράφεται:

$$\mathbf{y} = \mathbf{X} \cdot \boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (4.6)$$

όπου

$$\mathbf{X} = \begin{bmatrix} 1 & \mathbf{A} \\ \vdots & \\ 1 & \end{bmatrix} \quad (4.7)$$

ο πίνακας σχεδιασμού, $\mathbf{A} = \{\mathbf{x}_{jc}\} = \{\mathbf{x}_{1c}, \mathbf{x}_{2c}, \dots, \mathbf{x}_{kc}\}$ ο πίνακας του πλάνου των μετρήσεων, με

$\mathbf{x}_{jc} = (x_{1jc}, x_{2jc}, \dots, x_{njc})^T$, $\mathbf{y} = (y_1, y_2, \dots, y_n)^T$, $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_k)^T$ και $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)^T$.

Στον διερευνητικό σχεδιασμό ο πίνακας των δοκιμών είναι *ισοσταθμισμένος* (balanced), δηλαδή τα δύο επίπεδα τιμών $\{x_{jc}\} = \{-1, +1\}$ του j παράγοντα εμφανίζονται ίσες φορές το καθένα στο σύνολο n των δοκιμών, γεγονός που έχει ως συνέπεια κάθε παράγοντας να έχει μέση τιμή μηδέν, ενώ το πλάνο των μετρήσεων καταστρώνεται ώστε κάθε στήλη του $\mathbf{A}$ είναι *ορθογώνια* ως προς τις υπόλοιπες, δηλ.:

$$\mathbf{x}_{jc}^T \cdot \mathbf{x}_{\ell c} = 0 \text{ ή } \sum_{i=1}^n x_{ijc} x_{i\ell c} = 0, \text{ για } j = \ell \quad \text{και} \quad \bar{x}_{jc} = \frac{1}{n} \sum_{i=1}^n x_{ijc} = 0 \quad (4.8)$$

Οι παραπάνω συνθήκες, που συμπυκνώνονται στην σχέση:

$$\mathbf{X}^T \cdot \mathbf{X} = n\mathbf{I}_n \quad (4.9)$$

διασφαλίζουν ότι οι παράγοντες του σχεδιασμού είναι ασυσχέτιστοι ανά δύο, γεγονός που απλοποιεί σημαντικά το μοντέλο, αφού ο αλγόριθμος των ελαχίστων τετραγώνων θα έχει απλή μορφή και οι συντελεστές του γραμμικού μοντέλου, οπότε και τα αποτελέσματα κάθε παράγοντα, μπορούν να εκτιμηθούν ανεξάρτητα. Από την (4.6) έχουμε:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \cdot \mathbf{X})^{-1} \cdot \mathbf{X}^T \cdot \mathbf{y} = \frac{1}{n} \mathbf{X}^T \cdot \mathbf{y} \quad (4.10)$$

$$\hat{\beta}_j = \frac{1}{n} \left(\sum_{i \text{ με } x_{ijc}=+1} y_i - \sum_{i \text{ με } x_{ijc}=-1} y_i \right) \quad \hat{\beta}_0 = \frac{1}{n} \sum_{i=1}^n y_i = \bar{y} \quad (4.11)$$

Δεδομένου ότι οι αποκρίσεις για το μέγιστο επίπεδο κάθε παράγοντα έχουν πλήθος $n/2$ (αντίστοιχα και για το ελάχιστο), δηλ.:

$$\bar{y} = \frac{2}{n} \sum_{i \text{ με } x_{ijc}=+1} y_i \quad (4.12)$$

ο συντελεστής κάθε παράγοντα ισοδυναμεί με το ημιαποτέλεσμα (half effect) του παράγοντα:

$$\text{main effect}_{x_j} = 2\hat{\beta}_j \quad (4.13)$$

Η γραμμική συνάρτηση μπορεί να μην αποδειχθεί ιδιαίτερα ακριβής στην περιγραφή της διεργασίας, αλλά εξυπηρετεί τις επιδιώξεις της αρχικής διερεύνησης, αφού ο στόχος εδώ είναι να προσδιοριστεί αν ένας παράγοντας επιδρά σημαντικά στην απόκριση και όχι να προσδιοριστεί με ακρίβεια η επίδρασή του. Η ανάγκη να συμπεριληφθούν όσο το δυνατόν περισσότεροι εν δυνάμει παράγοντες στον διερευνητικό σχεδιασμό, προκειμένου να μην υπάρξουν λανθασμένες αποφάσεις κατά την επιλογή των σημαντικότερων εξ' αυτών, σε συνδυασμό με την απαίτηση για κατ' ελάχιστον $k+1$ δοκιμές για τον προσδιορισμό των συντελεστών β_0 και β_j του γραμμικού μοντέλου, περιορίζει τον όγκο του πειράματος που θα αυξανόταν δραματικά με την προσθήκη δευτεροβάθμιων όρων ως προς τους παράγοντες.

Σε επόμενο στάδιο πειραματισμού, με περιορισμένο πλέον τον αριθμό των ανεξάρτητων μεταβλητών, το ζητούμενο είναι η περαιτέρω μείωση των παραγόντων με κριτήριο την περιεχόμενη πληροφορία, ποσοτικοποιώντας τόσο την επίδραση που παρουσιάζουν αυτοί

μεμονωμένα στην διαμόρφωση της απόκρισης του συστήματος, όσο και τις αλληλεπιδράσεις μεταξύ τους (ανταγωνιστικές ή συνέργιας). Η πειραματική διερεύνηση εξαντλείται σε λεπτομερή περιγραφή του πειραματικού χώρου, που ορίζουν οι διαφαινόμενοι κυρίαρχοι παράγοντες, με στόχο την ανάπτυξη αλγορίθμου με καλή ακρίβεια πρόβλεψης των αποκρίσεων, τόσο για τα υπάρχοντα δεδομένα, όσο και για νέα άγνωστα δείγματα εντός του χώρου μελέτης (γενίκευση του μοντέλου).

Για λόγους απλότητας το μοντέλο εξακολουθεί να θεωρείται γραμμικό ως προς τους παράγοντες και ως προς τις αλληλεπιδράσεις τους, που εκλαμβάνονται ως επιπλέον παράγοντες του σχεδιασμού. Αυτό, εκτός από την απλοποίηση των υπολογισμών, δίνει την δυνατότητα στον ερευνητή να διερευνήσει τα αποτελέσματα των αλληλεπιδράσεων μεταξύ των παραγόντων ανεξάρτητα από τα αποτελέσματα που παρουσιάζουν αυτοί μεμονωμένα.

$$y_i = \beta_0 + \sum_{j=1}^k \beta_j x_{ijc} + \sum_{j=1}^{k-1} \sum_{\substack{\ell=2 \\ \ell > j}}^k \beta_{j,\ell} x_{ijc} x_{i\ell c} + \varepsilon_i \quad (4.14)$$

όπου $x_{ijc} x_{i\ell c}$ το γινόμενο των κωδικοποιημένων επιπέδων των παραγόντων j και ℓ κατά την i -στη δοκιμή

$\beta_{j,\ell}$ ο συντελεστής της αλληλεπίδρασης μεταξύ των παραγόντων j και ℓ του σχεδιασμού που αντικατοπτρίζει το ημιαποτέλεσμα της αλληλεπίδρασής τους

Στην εξίσωση (4.6), ο πίνακας σχεδιασμού είναι επανυξημένος σε σχέση με αυτόν του διερευνητικού μοντέλου ως προς τον πίνακα των αλληλεπιδράσεων των παραγόντων ανά δύο

$\mathbf{C} = \{\mathbf{x}_{j,\ell c}\} = \{\mathbf{x}_{1,2c}, \dots, \mathbf{x}_{1,kc}, \mathbf{x}_{2,3c}, \dots, \mathbf{x}_{2,kc}, \dots, \mathbf{x}_{k-1,kc}\}$, με $\mathbf{x}_{j,\ell c} = (x_{1jc} x_{1\ell c}, x_{2jc} x_{2\ell c}, \dots, x_{njc} x_{n\ell c})^T$:

$$\mathbf{X} = \begin{bmatrix} 1 \\ \vdots \\ \mathbf{A} \quad \mathbf{C} \\ 1 \end{bmatrix} \quad (4.15)$$

ενώ το διάνυσμα των συντελεστών και του σταθερού όρου είναι:

$$\boldsymbol{\beta} = (\beta_0, (\beta_j)_{j=1, \dots, k-1 \text{ και } \ell=1, \dots, k \text{ με } j < \ell})^T = (\beta_0, \beta_1, \dots, \beta_k, \beta_{1,2}, \dots, \beta_{1,k}, \beta_{2,3}, \dots, \beta_{2,k}, \dots, \beta_{k-1,k})^T.$$

Το μοντέλο διατηρεί τις ίδιες συνθήκες με το απλό γραμμικό, όσον αφορά τις ανεξάρτητες μεταβλητές, τόσο για τις πραγματικές του συστήματος των k παραγόντων, όσο και για τις $k(k-1)/2$ των αλληλεπιδράσεων των παραγόντων ανά δύο, ενώ ο ελάχιστος απαιτούμενος αριθμός δοκιμών για την ύπαρξη της λύσης του $\boldsymbol{\beta}$ είναι $k + k(k-1)/2 + 1$. Οι συντελεστές του μοντέλου υπολογίζονται απ' ευθείας συναρτήσει των αποκρίσεων, σύμφωνα με την (4.11) για τους παράγοντες και τον σταθερό όρο, ενώ για τις αλληλεπιδράσεις είναι:

$$\hat{\beta}_{j,\ell} = \frac{1}{n} \left(\sum_{i \text{ με } x_{ijc} x_{i\ell c} = +1} y_i - \sum_{i \text{ με } x_{ijc} x_{i\ell c} = -1} y_i \right) \quad (4.16)$$

Η (4.14) μπορεί να περιλαμβάνει και όρους αλληλεπιδράσεων μεταξύ τριών ή περισσότερων παραγόντων, που όμως επιβαρύνουν σημαντικά τους υπολογισμούς χωρίς να προσφέρουν ουσιαστικά σε ακρίβεια για τις συνήθεις εφαρμογές, ενώ για τον ίδιο λόγο οι αλληλεπιδράσεις ανά δύο κάποιων εκ των παραγόντων μπορούν να τεθούν εξ' αρχής μηδενικές.

Στα περισσότερα εφαρμοσμένα προβλήματα, ο απώτερος στόχος είναι η βελτιστοποίηση του συστήματος μελέτης, που αφορά την διαδικασία ρύθμισης των παραγόντων ώστε να εξασφαλίζεται η καλύτερη δυνατή απόκριση του συστήματος. Όλες οι τεχνικές πειραματικού σχεδιασμού αριστοποίησης στοχεύουν στην ανάπτυξη ενός μοντέλου που να αποδίδει όσο το δυνατόν καλύτερα την επιφάνεια απόκρισης, δηλαδή την επιφάνεια που δημιουργεί στον χώρο η απόκριση του συστήματος, προσδιορισμένη σε ολόκληρο το πεδίο τιμών του καθενός από τους παράγοντες επιρροής, προκειμένου να διερευνηθεί αυτή ως προς τα υπάρχοντα μέγιστα ή ελάχιστα σημεία της. Η προηγούμενη πειραματική διερεύνηση, με την ανάπτυξη γραμμικών μοντέλων ως προς τους εν δυνάμει παράγοντες και τις αλληλεπιδράσεις ανταγωνισμού και συνέργιας μεταξύ τους, είναι προαπαιτούμενη, για την οριοθέτηση της βέλτιστης περιοχής του ευρύτερου πειραματικού χώρου και τον περιορισμό του αριθμού των κύριων παραγόντων επιρροής. Δίδεται έτσι η δυνατότητα για οικονομικά και χρονικά εφικτό σχεδιασμό πειράματος με αυξημένο τοπικά αριθμό δοκιμών και χρήση περισσότερο πολύπλοκων συναρτήσεων προσέγγισης, ώστε να αναδειχθούν όλες οι κυμάνσεις της επιφάνειας απόκρισης που ενδέχεται να αποκρύψουν το βέλτιστο σημείο.

Στην περιοχή ενός βέλτιστου η απόκριση δεν μπορεί να προσεγγιστεί ικανοποιητικά από το γραμμικό μοντέλο ακόμη κι αν αυτό περιλαμβάνει όλες τις δυνατές αλληλεπιδράσεις των παραγόντων, οπότε χρησιμοποιούνται δευτεροβάθμια πολυώνυμα προσέγγισης. Στο μοντέλο της (4.14) προστίθενται και οι όροι των τετραγώνων των παραγόντων.

$$y_i = \beta_0 + \sum_{j=1}^k \beta_j x_{ijc} + \sum_{j=1}^{k-1} \sum_{\substack{\ell=2 \\ \ell > j}}^k \beta_{j,\ell} x_{ijc} x_{i\ell c} + \sum_{j=1}^k \beta_{jj} x_{ijc}^2 + \varepsilon_i \quad (4.17)$$

Η εξίσωση που περιγράφει το μοντέλο είναι η (4.6), με πίνακα σχεδιασμού τον

$$\mathbf{X} = \begin{bmatrix} 1 \\ \vdots \\ \mathbf{A} \quad \mathbf{C} \quad \mathbf{Q} \\ 1 \end{bmatrix} \quad (4.18)$$

όπου $\mathbf{Q} = \{\mathbf{x}_{jc}^2\} = \{\mathbf{x}_{1c}^2, \mathbf{x}_{2c}^2, \dots, \mathbf{x}_{kc}^2\}$ ο πίνακας των τετραγώνων των παραγόντων, με

$\mathbf{x}_{jc}^2 = (x_{1jc}^2, x_{2jc}^2, \dots, x_{njc}^2)^T$, ενώ το διάνυσμα των συντελεστών και του σταθερού όρου είναι:

$$\boldsymbol{\beta} = (\beta_0, (\beta_j)_{j=1, \dots, k-1 \text{ και } \ell=1, \dots, k \text{ με } j < \ell}, (\beta_{jj}))^T = (\beta_0, \beta_1, \dots, \beta_k, \beta_{1,2}, \dots, \beta_{1,k}, \beta_{2,3}, \dots, \beta_{2,k}, \dots, \beta_{k-1,k}, \beta_{1,1}, \dots, \beta_{k,k})^T$$

Ο ελάχιστος επομένως απαιτούμενος αριθμός δοκιμών για την ύπαρξη της λύσης του $\boldsymbol{\beta}$ είναι $k + k(k+1)/2 + 1$, αν και στα πειράματα βελτιστοποίησης συνήθως εκτελείται πλεονάζων αριθμός

δοκιμών για την εύρεση της άριστης λύσης, αφού αν ο αριθμός των βαθμών ελευθερίας είναι πολύ μικρός υπάρχει κίνδυνος *υπερπροσαρμογής* του μοντέλου στα δεδομένα, με χαμηλή όμως προβλεψιμότητα σε νέα άγνωστα δείγματα και ιδιαίτερα στον προσδιορισμό του βέλτιστου.

Στο μοντέλο βελτιστοποίησης δεν υπάρχει απαίτηση για ασυσχέτιστες μεταβλητές, ενώ τα επίπεδα των παραγόντων πρέπει να είναι τουλάχιστον τρία, αφού η χρήση μόνο των τιμών $-1, +1$ θα είχε ως αποτέλεσμα οι τετραγωνικοί παράγοντες να είναι σταθερά μονάδες και να μην διαχωρίζονται από τον σταθερό όρο. Συνήθως χρησιμοποιείται η κωδικοποίηση τριών τιμών επιπέδων για κάθε παράγοντα $\{x_{ic}\} = \{-1, 0, +1\}$.

Πίνακας 4-1 Είδη παραγοντικών σχεδιασμών και χαρακτηριστικά τους

	Στατιστικός πειραματικός σχεδιασμός		
	Διερευνητικός	Πρόβλεψης	Βελτιστοποίησης
Αριθμός παραγόντων	$k \geq 8$	$4 \leq k < 8$	$k < 4$
Μαθηματικό μοντέλο	γραμμικό ως προς τους παράγοντες	παράγοντες διεργασίας γραμμικό ως προς τους παράγοντες και τις αλληλεπιδράσεις παράγοντες μιγμάτων	δευτεροβάθμιο
Αντικειμενικός στόχος	γραμμικό Cox μείωση παραγόντων	τετραγωνικό Cox μείωση παραγόντων	τετραγωνικό Cox βελτιστοποίηση απόκρισης
Αριθμός δοκιμών	ελάχιστος	ελάχιστος	πλεονάζων
Τυπικές τεχνικές σχεδιασμού	σχεδιασμός Plackett-Burman • αξονικός σχεδιασμός • σχεδιασμός $\{k, 1\}$ Simplex-Lattice	παράγοντες διεργασίας κλασματικός παραγοντικός σχεδιασμός παράγοντες μιγμάτων • σχεδιασμός Simplex-Centroid • σχεδιασμός $\{k, m\}$ Simplex-Lattice • σχεδιασμός D-optimal	• σύνθετος κεντρικός σχεδιασμός • σχεδιασμός D-optimal • σχεδιασμός Simplex-Centroid • σχεδιασμός $\{k, m\}$ Simplex-Lattice • σχεδιασμός D-optimal

Είναι φανερό ότι ο στατιστικός πειραματικός σχεδιασμός, πολύ πριν την εφαρμογή της κατάλληλης τεχνικής κατάστρωσης του πλάνου των μετρήσεων και την διεξαγωγή τους, απαιτεί την πλήρη εξοικείωση του ερευνητή με το διερευνούμενο φαινόμενο και την συγκέντρωση της διαθέσιμης γνώσης και εμπειρίας, προκειμένου να ληφθούν οι σωστές αποφάσεις επιλογής των παραμέτρων που αφορούν την υπό σχεδιασμό πορεία της πειραματική διαδικασία. Ανάλογα με τους στόχους και το στάδιο της διερεύνησης του συστήματος, επιλέγονται οι μεταβλητές που

υπεισέρχονται στο σύστημα, το μοντέλο περιγραφής της συμπεριφοράς τους, και η τεχνική επιλογής των πειραματικών σημείων, όπως παρουσιάζονται σε γενικές γραμμές στον Πιν. 4-1.

4.3 Τυπικές Τεχνικές Σχεδιασμού Πειράματος

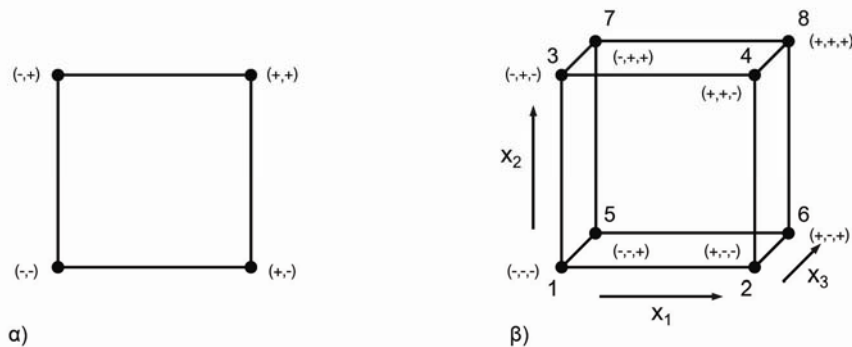
Ακόμη και στην περίπτωση των συγκριτικών πειραμάτων, αλλά κυρίως στα παραγοντικά πειράματα είναι απαραίτητος ο σχεδιασμός του πίνακα των δοκιμών κατά τρόπο συστηματικό για τον εκ των προτέρων προσδιορισμό του αριθμού των απαιτούμενων δοκιμών, τόσο για την οικονομική διαχείριση της πειραματικής μελέτης, όσο και για αποφυγή της σύγχυσης που δημιουργεί το πλήθος των πειραματικών σημείων, ενώ διαφορετική πρακτική ενέχει τον κίνδυνο διεξαγωγής περιττών δοκιμών ή παράλειψης καθοριστικών για την ανάλυση σημείων. Η λογική του παραγοντικού πειράματος έγκειται στην *ταυτόχρονη* μεταβολή περισσοτέρων του ενός παραγόντων, έτσι ώστε να εξασφαλίζεται η διερεύνηση της συμπεριφοράς του συστήματος σε όλο το εύρος των τιμών που μπορούν να λάβουν οι παράγοντες του (υποχώρος μελέτης) και οι τεχνικές σχεδιασμού προσδιορίζουν όλους τους δυνατούς συνδυασμούς των επιπέδων των παραγόντων που απαιτούνται για την ζητούμενη σε κάθε περίπτωση κάλυψη του πειραματικού χώρου. Οι τεχνικές σχεδιασμού πειράματος, που αφορούν στατιστικές μεθόδους κατάστρωσης του πλάνου των δοκιμών, μπορούν να ταξινομηθούν σε κατηγορίες ανάλογα με το είδος του πειράματος που απαιτείται για την διεξαγόμενη κάθε φορά έρευνα (Πιν. 4-1), και εφαρμόζουν σχεδόν στο σύνολό τους την λογική του **πλήρους παραγοντικού πειραματικού σχεδιασμού** (*full factorial experimental design*), που αναπτύσσεται παρακάτω, με διαφοροποιήσεις που αφορούν κυρίως την πυκνότητα κάλυψης του πειραματικού χώρου και την πυκνότητα διακριτοποίησης. Τέλος, οι τεχνικές που χρησιμοποιούνται σε συγκριτικές πειραματικές μελέτες υιοθετούν επίσης την λογική του συστηματικού προσδιορισμού του αριθμού των πειραματικών μονάδων και αφορούν τεχνικές τυχαιοποίησης στην περίπτωση ενός παράγοντα ή ομαδοποίησης για δύο ή περισσότερους¹.

Πλήρης Παραγοντικός Σχεδιασμός Πειράματος

Στον πλήρη παραγοντικό σχεδιασμό πειράματος ο αριθμός των N δοκιμών περιορίζεται σε όλους τους δυνατούς συνδυασμούς μεταξύ των παραγόντων ενός συστήματος και για όλα τα επίπεδα των τιμών τους. Αν $L_1, L_2, \dots, L_k$ τα επίπεδα μεταβολής της τιμής των k παραγόντων, τότε το σύνολο των $L_1 \times L_2 \times \dots \times L_k$ πειραματικών δοκιμών συνιστούν ένα πλήρες παραγοντικό

¹ Σχεδιασμοί Latin Square για δύο παράγοντες ομαδοποίησης, Graeco-Latin Square για τρεις και Hyper-Graeco-Latin Square για τέσσερις

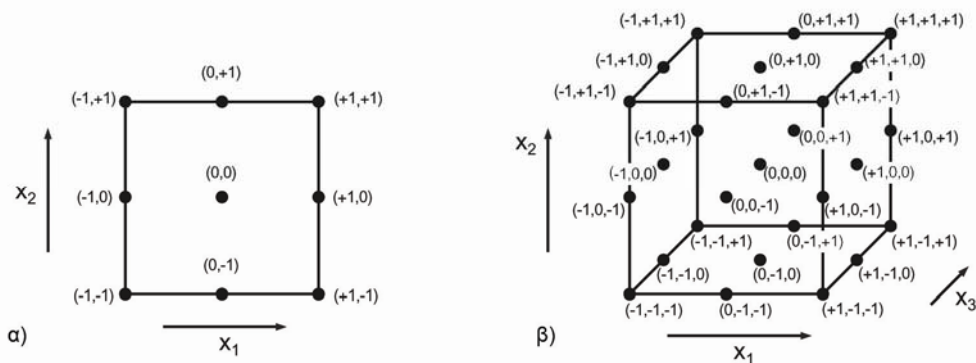
πρόβλημα για το δεδομένο σύστημα. Στην ειδική περίπτωση που οι παράγοντες έχουν τον ίδιο αριθμό επιπέδων $L_1=L_2=\dots=L_k=L$, το σύνολο των πειραματικών δοκιμών είναι $N=L^k$. Η χρήση όλων των δυνατών συνδυασμών έχει σαν αποτέλεσμα έναν πλήρως ισοσταθμισμένο πίνακα δοκιμών, ενώ περιλαμβάνεται ικανός αριθμός δοκιμών για την ανάπτυξη μοντέλου με ανεξάρτητες μεταβλητές, εκτός των παραγόντων και των αλληλεπιδράσεων μεταξύ τους ανά δύο, και όλες τις άλλες δυνατές αλληλεπιδράσεις τους. Είναι φανερό, όμως, ότι ο όγκος του πειράματος καθίσταται απαγορευτικός για περισσότερους των τεσσάρων παραγόντων, αφού για σχεδιασμούς μόνο δύο επιπέδων ο αριθμός των δοκιμών αυξάνεται εκθετικά ($2^2=4$, $2^3=8$, $2^4=16$, $2^5=32$, κ.τ.λ.), πόσο μάλλον για χρήση περισσότερων επιπέδων τιμών ($3^2=9$, $3^3=27$, $3^4=81$, $3^5=243$, κ.τ.λ.). Στα διερευνητικά πειράματα, αλλά και τα πειράματα μοντελοποίησης ο πλήρης παραγοντικός σχεδιασμός παρέχει έναν υψηλό αριθμό βαθμών ελευθερίας, που αυξάνει όμως το κόστος του πειράματος χωρίς να προσδίδει επιπλέον χρήσιμη πληροφορία, αφού η ενσωμάτωση στο μοντέλο διεργασίας αλληλεπιδράσεων άνω των δύο παραγόντων δεν αυξάνει σημαντικά την ακρίβεια του μοντέλου.



Σχήμα 4-3 Πλήρης παραγοντικός σχεδιασμός α) 2^2 β) 2^3

Η απλούστερη περίπτωση παραγοντικού πειράματος είναι αυτή των παραγόντων με δύο επίπεδα τιμών κωδικοποιημένων ως -1 , $+1$ ή απλά $-$, $+$, γνωστός ως πλήρης παραγοντικός σχεδιασμός 2^k , που για πειραματικούς χώρους δύο και τριών διαστάσεων είναι δυνατό να παρασταθεί γραφικά (Σχ. 4-3). Ο περιορισμός των επιπέδων των παραγόντων σε δύο δεν είναι εσφαλμένος, αφού, ειδικά σε μία βιομηχανική διεργασία, το εύρος μεταβολής των παραμέτρων είναι δεδομένο και συνήθως μικρό λόγω των περιορισμών της παραγωγής. Το τετράγωνο και ο κύβος απεικονίζουν τον διδιάστατο και τρισδιάστατο πειραματικό χώρο, ενώ κάθε εσωτερικό σημείο τους αντιστοιχεί σε κάθε θεωρητικά δυνατό πείραμα. Οι κορυφές των σχημάτων αντιστοιχούν στα πειράματα που πρόκειται να εκτελεστούν και αντιπροσωπεύουν όλους τους δυνατούς συνδυασμούς για τις ακραίες τιμές όλων των παραγόντων. Το βασικό πλεονέκτημα του σχεδιασμού αυτού είναι ότι απαιτεί σχετικά μικρό αριθμό πειραμάτων, ενώ ταυτόχρονα

διερευνά την συμπεριφορά-απόκριση του συστήματος για όλο το εύρος του πεδίου τιμών των μεταβλητών του. Προφανώς η χρήση περισσότερων επιπέδων οδηγεί σε πιο “πυκνό” πειραματικό σχεδιασμό με ραγδαία αύξηση όμως του αριθμού των πειραματικών δοκιμών. Το Σχ. 4-4 απεικονίζει τις πειραματικές δοκιμές πλήρους παραγοντικού σχεδιασμού αν οριστούν τρία επίπεδα τιμών– τις οριακές και τον αριθμητικό τους μέσο– για κάθε παράγοντα $(-1, 0, +1)$.



Σχήμα 4-4 Πλήρης παραγοντικός σχεδιασμός α) 3^2 β) 3^3

Ο πίνακας του πλάνου των δοκιμών που δημιουργείται περιλαμβάνει τα επίπεδα των παραγόντων, εκείνα όλων των δυνατών αλληλεπιδράσεων μεταξύ τους, καθώς και τις τιμές των αποκρίσεων του συστήματος για κάθε μία από τις πειραματικές μονάδες του σχεδιασμού. Για σχεδιασμό 2^3 , που αναφέρεται σε πλήρη παραγοντικό σχεδιασμό δύο επιπέδων $-$, $+$ τριών παραγόντων x_1 , x_2 , x_3 , το γενικό πειραματικό πλάνο παρουσιάζεται στον Πιν. 4-2, όπου οι στήλες x_1 , x_2 , x_3 αποτελούν το πλάνο των δοκιμών (τις συντεταγμένες των ακραίων σημείων του τρισδιάστατου πειραματικού χώρου, όπου θα διεξαχθούν πειραματικές μετρήσεις), οι στήλες $x_1 x_2$, $x_1 x_3$, $x_2 x_3$ και $x_1 x_2 x_3$ τις αλληλεπιδράσεις των μεταβλητών εισόδου ανά δύο και ανά τρεις αντίστοιχα, ενώ η απόκριση του συστήματος σε κάθε πειραματική δοκιμή y_i με $i=1, \dots, 2^3$ αναγράφεται στην τελευταία στήλη. Στην πρώτη στήλη αναγράφεται ο αριθμός της κάθε πειραματικής μονάδας σύμφωνα με την αρχική κανονική σειρά του σχεδιασμού (πριν την τυχαιοποίηση) των δοκιμών. Ο γενικός κανόνας σε ένα σχεδιασμό 2^k για την κατασκευή του πλάνου των δοκιμών σε κανονική σειρά είναι: Η i -οστή στήλη (x_i) αρχίζει με 2^{i-1} επαναλήψεις του $-$ που εναλλάσσονται με 2^{i-1} επαναλήψεις του $+$ για το σύνολο των 2^k δοκιμών.

Πίνακας 4-2 Πλήρες πλάνο παραγοντικού σχεδιασμού τριών παραγόντων δύο επιπέδων συμπεριλαμβανομένου και των παραγόντων αλληλεπίδρασης μεταξύ των μεταβλητών

Αριθμός δοκιμής	x_1	x_2	x_3	$x_1 \cdot x_2$	$x_1 \cdot x_3$	$x_2 \cdot x_3$	$x_1 \cdot x_2 \cdot x_3$	y_i
1	-	-	-	+	+	+	-	y_1
2	+	-	-	-	-	+	+	y_2
3	-	+	-	-	+	-	+	y_3
4	+	+	-	+	-	-	-	y_4
5	-	-	+	+	-	-	+	y_5
6	+	-	+	-	+	-	-	y_6
7	-	+	+	-	-	+	-	y_7
8	+	+	+	+	+	+	+	y_8

Ο πίνακας X του πειραματικού σχεδιασμού 2^k , που χρησιμοποιείται για την ανάπτυξη του μοντέλου (3.6) της διεργασίας και είναι όμοιος με τον Πιν. 4-2 αν η πρώτη στήλη αντικατασταθεί με διάνυσμα μονάδων ή +, καλύπτει τις συνθήκες ορθογωνιότητας της (4.9) που διασφαλίζουν ότι οι παράγοντες και οι αλληλεπιδράσεις μεταξύ τους είναι ασυσχέτιστες. Επομένως ο πλήρης παραγοντικός σχεδιασμός 2^3 , και κατ' επέκταση κάθε σχεδιασμός 2^k , δίνει την δυνατότητα για ανάπτυξη του πλήρους μοντέλου

$$\hat{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{23} x_2 x_3 + \beta_{13} x_1 x_3 + \beta_{123} x_1 x_2 x_3 \quad (4.19)$$

όπου οι συντελεστές, και άρα τα κύρια αποτελέσματα (εξίσωση (4.13)) των παραγόντων και των αλληλεπιδράσεων οποιασδήποτε τάξης μεταξύ τους υπολογίζονται ανεξάρτητα, αποδίδοντας έτσι την “φυσική” σημασία της επίδρασης τους στην πορεία της μελετούμενης διεργασίας.

Σε ένα πραγματικό πείραμα οι δοκιμές δεν περιορίζονται σε 2^k , αλλά θα πρέπει να συνυπολογιστούν και αυτές των επαναληπτικών μετρήσεων και των κεντρικών σημείων ελέγχου ή αναφοράς. Τα **κεντρικά σημεία** (center points), στα μέσα δηλαδή επίπεδα όλων των παραγόντων, χρησιμοποιούνται για τον έλεγχο της **κυρτότητας** (curvature) της επιφάνειας απόκρισης. Η ύπαρξη ή όχι γραμμικότητας, που αποτελεί μία από τις βασικές υποθέσεις του γραμμικού μοντέλου των κύριων παραγόντων και αλληλεπιδράσεων, μπορεί να ελεγχθεί με το αν η απόκριση στο κεντρικό σημείο προσεγγίζει ή όχι την μέση τιμή των αποκρίσεων των σημείων του παραγοντικού σχεδιασμού. Η εμφάνιση υψηλής κυρτότητας υποδεικνύει την ανάγκη αντικατάστασης του γραμμικού μοντέλου από τετραγωνικό για την περιγραφή της διεργασίας. Στις περισσότερες των περιπτώσεων κεντρικό σημείο είναι και κάποιο από τα **σημεία αναφοράς** (reference samples), τα οποία εισάγονται ανεξάρτητα στον σχεδιασμό και μπορούν επίσης να περιλαμβάνουν σημεία για τυπικές (πιο συχνά εμφανιζόμενες στην διεργασία) τιμές των παραγόντων με γνωστή απόκριση (από προηγούμενα πειράματα ή

εμπειρία). Τέλος, για την εκτίμηση του σφάλματος μέτρησης της απόκρισης, είναι απαραίτητο να συμπεριληφθούν στο πλάνο δοκιμών και **επαναληπτικές μετρήσεις** (replicates). Η συνηθέστερη και οικονομικότερη πρακτική είναι η διεξαγωγή επαναληπτικών πειραμάτων στα κεντρικά σημεία ή τα τυχόν διαφορετικά σημεία αναφοράς και η εκτίμηση του σφάλματος του πειράματος με βάση τις μετρήσεις στα σημεία αυτά, αλλά μπορούν να εφαρμοστούν επαναληπτικές μετρήσεις στο σύνολο των πειραματικών δοκιμών για παράλληλο έλεγχο της ομοιομορφίας της διασποράς της απόκρισης σε όλο τον πειραματικό χώρο¹.

Κλασματικός Παραγοντικός Σχεδιασμός Πειράματος

Ο σχεδιασμός και η διεξαγωγή ενός πλήρους παραγοντικού πειράματος είναι ο ασφαλέστερος τρόπος για να διερευνηθεί λεπτομερώς ένας πειραματικός χώρος. Ακόμα όμως και στην περίπτωση που ο αριθμός των παραγόντων ενός πειράματος είναι σχετικά μικρός, ο αριθμός των δοκιμών, μαζί με τις επιπρόσθετες μετρήσεις στα κεντρικά σημεία και τις απαιτούμενες επαναληπτικές, καθιστά απαγορευτική την χρήση του, λόγω του υψηλού κόστους και του μεγάλου χρόνου που απαιτεί, ενώ ειδικά για τα διερευνητικά πειράματα ($k > 8$) είναι βασικό να εφαρμοστούν οικονομικότεροι σχεδιασμοί, που επιτρέπουν εξίσου καλή διερεύνηση του χώρου με ουσιαστικά μικρότερο αριθμό δοκιμών από τον αντίστοιχο πλήρη παραγοντικό σχεδιασμό. Το παραπάνω πρόβλημα αντιμετωπίζεται συχνά με την εφαρμογή του **κλασματικού παραγοντικού πειραματικού σχεδιασμού** (fractional factorial experimental design), κατά τον οποίο επιλέγεται να εκτελεστεί μόνο ένα κατάλληλο κλάσμα του συνόλου των συνδυασμών των επιπέδων των παραγόντων του πλήρους παραγοντικού σχεδιασμού, δηλαδή ένα τμήμα των μετρήσεων που να διατηρεί τις ιδιότητες του συνολικού πίνακα δοκιμών και την μέγιστη δυνατή πληροφορία για τα αποτελέσματα των παραγόντων και των αλληλεπιδράσεων μεταξύ τους.

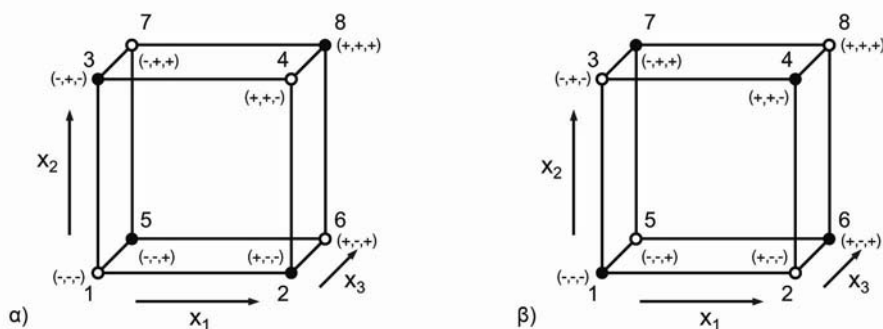
Στον κλασματικό παραγοντικό σχεδιασμό πειράματος L^{k-p} , το πλάνο των μετρήσεων περιέχει το κλάσμα $1/L^p$ του συνόλου των μετρήσεων του πλήρους παραγοντικού σχεδιασμού (ο σχεδιασμός 2^{k-1} περιέχει το $1/2$ των μετρήσεων του 2^k , ο σχεδιασμός 2^{k-2} το $1/4$, κ.ο.κ). Για την κατασκευή του πίνακα μετρήσεων ενός κλασματικού παραγοντικού σχεδιασμού 2^{k-p} , στην πραγματικότητα ο πίνακας του αντίστοιχου πλήρους ομαδοποιείται ως προς p παράγοντές του. Για την διατήρηση της μέγιστης δυνατής πληροφορίας εντός των κλασμάτων αυτών, επιλέγονται ως παράγοντες ομαδοποίησης αυτοί των μεγαλύτερου βαθμού αλληλεπιδράσεων, δεδομένου ότι θεωρούνται λιγότερο σημαντικές από τις υπόλοιπες και τους κύριους παράγοντες. Έτσι, σε

¹ Βασική υπόθεση της ανάλυσης αποτελεί η ύπαρξη *ομοιογένειας της διακύμανσης* της απόκρισης στον πειραματικό χώρο.

έναν 2^{3-1} κλασματικό σχεδιασμό, παράγοντας ομαδοποίησης λαμβάνεται αυτός των αλληλεπιδράσεων τρίτου βαθμού, ενώ κριτήριο ομαδοποίησης είναι το επίπεδο τιμών του (- ή +), ώστε για $x_1x_2x_3=1$ προκύπτει ο σχεδιασμός τεσσάρων μετρήσεων που δίνεται στον Πιν. 4-3, ενώ για $x_1x_2x_3=-1$ ο πίνακας περιλαμβάνει τις υπόλοιπες τιμές του Πιν. 4-2.

Πίνακας 4-3 Κλασματικός παραγοντικός σχεδιασμός 2^{3-1}

Αριθμός δοκιμής	x_1	x_2	x_3	$x_1 \cdot x_2$	$x_1 \cdot x_3$	$x_2 \cdot x_3$	$x_1 \cdot x_2 \cdot x_3$	y_i
1	+	-	-	-	-	+	+	y_1
2	-	+	-	-	+	-	+	y_2
3	-	-	+	+	-	-	+	y_3
4	+	+	+	+	+	+	+	y_4



Σχήμα 4-5 Κλασματικός παραγοντικός σχεδιασμός 2^{3-1} με κριτήριο ομαδοποίησης α) την $x_1x_2x_3=1$ και β) την $x_1x_2x_3=-1$.

Διεξάγοντας ένα μόνο κλάσμα του αριθμού των μετρήσεων του πλήρους παραγοντικού σχεδιασμού, δεν υπάρχει η δυνατότητα ανάπτυξης μοντέλου που να περιλαμβάνει τους κύριους παράγοντες και όλες τις δυνατές αλληλεπιδράσεις μεταξύ τους και επιπλέον στο δημιουργούμενο μοντέλο οι άγνωστοι συντελεστές αντικατοπτρίζουν μόνο κατά προσέγγιση το αποτέλεσμα των αντίστοιχων παραγόντων, γιατί εμφανίζεται το φαινόμενο της “σύγχυσης” (confounding) μεταξύ των αποτελεσμάτων. Το κριτήριο ομαδοποίησης του σχεδιασμού μπορεί να χρησιμοποιηθεί για τον προσδιορισμό αυτής της σύγχυσης, με συνεχή πολλαπλασιασμό του με κάθε έναν από τους παράγοντες. Για τον παραπάνω 2^{3-1} κλασματικό σχεδιασμό θα είναι: $x_1x_2x_3=1 \Rightarrow x_1x_2x_3^2=x_3 \Rightarrow x_1x_2 \cdot 1=x_3 \Rightarrow x_1x_2=x_3$. Με όμοιο τρόπο λαμβάνονται και οι $x_1x_3=x_2$ και $x_2x_3=x_1$. Αυτό σημαίνει ότι στο δημιουργούμενο μοντέλο οι συντελεστές των κυρίων παραγόντων δεν θα δίνουν αποκλειστικά το αποτέλεσμα των παραγόντων αυτών, αφού το τελευταίο δεν διακρίνεται από εκείνο των αλληλεπιδράσεων ανά δύο των παραγόντων.

Η μείωση του αριθμού των δοκιμών έχει επομένως σαν αποτέλεσμα την μείωση της **διακριτικής ικανότητας** (resolution) του σχεδιασμού. Η διακριτική ικανότητα περιγράφει τον βαθμό στο οποίο συγχέονται τα αποτελέσματα των κυρίων παραγόντων με εκείνα των αλληλεπιδράσεων ανά δύο, ανά τρεις, κ.ο.κ. και ισούται με τον μικρότερο βαθμό των αλληλεπιδράσεων που χρησιμοποιούνται ως παράγοντας ομαδοποίησης. Έτσι, σε σχεδιασμούς με διακριτική ικανότητα III, όπως ο 2^{3-1} που παρουσιάστηκε, αποτελέσματα κυρίων παραγόντων συγχέονται με αυτά των αλληλεπιδράσεων μεταξύ δύο παραγόντων, αλλά όχι μεταξύ τους¹, σε σχεδιασμούς με διακριτική ικανότητα IV αποτελέσματα κυρίων παραγόντων δεν συγχέονται με αυτά των αλληλεπιδράσεων μεταξύ δύο παραγόντων, αλλά τα τελευταία συγχέονται μεταξύ τους και σε σχεδιασμούς με διακριτική ικανότητα V, ούτε τα αποτελέσματα των κυρίων παραγόντων, ούτε των αλληλεπιδράσεων ανά δύο συγχέονται μεταξύ τους, αλλά τα αποτελέσματα των αλληλεπιδράσεων ανά δύο συγχέονται με αυτά των αλληλεπιδράσεων ανά τρεις, κ.ο.κ..

Μία ιδιαίτερα οικονομική τεχνική κλασματικού παραγοντικού σχεδιασμού είναι ο **σχεδιασμός Plackett-Burman** (Plackett-Burman, 1946) με την οποία καταστρώνονται πειράματα με αριθμό δοκιμών $N=k+1$ πολλαπλάσιο του τέσσερα αντί της δύναμης του δύο. Επειδή στους σχεδιασμούς αυτούς τα αποτελέσματα των κυρίων παραγόντων συγχέονται έντονα με αυτά των αλληλεπιδράσεων ανά δύο, η τεχνική χρησιμοποιείται αποκλειστικά σε διερευνητικά πειράματα. Η κατασκευή του πλάνου των μετρήσεων πραγματοποιείται ως εξής: Χρησιμοποιώντας ως βάση την *αρχική γραμμή* του πίνακα (Πιν. 4-4), κάθε επόμενη γραμμή αποτελεί μετατόπιση προς τα δεξιά της προηγούμενης ενώ το πρώτο στοιχείο αυτής λαμβάνει την τελευταία τιμή της προηγούμενης μέχρις ότου συμπληρωθούν οι k γραμμές του πίνακα του σχεδιασμού $k+1$ Plackett-Burman. Ο πίνακας δοκιμών συμπληρώνεται με μία ακόμη γραμμή που περιέχει αποκλειστικά “-”. Στον Πιν. 4-5 δίνεται παράδειγμα σχεδιασμού Plackett-Burman με 8 δοκιμές που μπορεί να χρησιμοποιηθεί για το πολύ 7 παράγοντες.

¹ Αν ως παράγοντας ομαδοποίησης επιλεγόταν ο $x_1 x_2$, τότε θα συγχέονταν μεταξύ τους και τα αποτελέσματα των κυρίων παραγόντων.

Πίνακας 4-4 Αρχική γραμμή για σχεδιασμό Plackett-Burman $k+1=N$ δοκιμών

N																				
4	-	+	+																	
8	-	-	+	-	+	+	+													
12	+	-	-	-	+	-	+	+	-	+	+									
16	-	-	-	+	-	-	+	+	-	+	-	+	+	+	+	+				
20	-	-	-	-	+	-	+	-	+	+	+	+	+	-	-	+	+	-	+	+

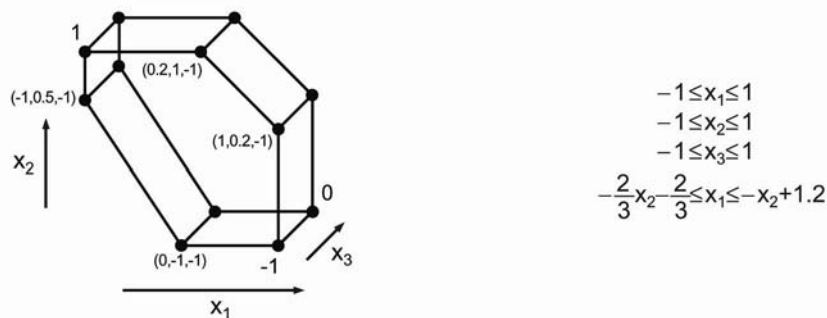
Πίνακας 4-5 Πλάνο σημείων μετρήσεων σχεδιασμού Plackett-Burman $7+1=8$ δοκιμών

Αριθμός δοκιμής	x_1	x_2	x_3	x_4	x_5	x_6	x_7
1	-	-	+	-	+	+	+
2	+	-	-	+	-	+	+
3	+	+	-	-	+	-	+
4	+	+	+	-	-	+	-
5	-	+	+	+	-	-	+
6	+	-	+	+	+	-	-
7	-	+	-	+	+	+	-
8	-	-	-	-	-	-	-

D-optimal πειραματικός σχεδιασμός

Οι κυριότερες τεχνικές πειραματικού σχεδιασμού που παρουσιάστηκαν προϋποθέτουν για την εφαρμογή τους ορισμένο βαθμό ακρίβειας, ορθογωνιότητα και ισοσταθμισμένο πίνακα μετρήσεων. Υπάρχουν όμως περιπτώσεις, όπου οι κλασικές τεχνικές δεν καλύπτουν τις ιδιαιτερότητες μίας ορισμένης πειραματικής έρευνας, όπως όταν η ομαδοποιημένη δομή ή το μέγεθος του πειράματος δεν μπορεί να προσαρμοστεί σε τυπικούς σχεδιασμούς ομαδοποίησης, αν δεν είναι εφικτοί όλοι συνδυασμοί των επιπέδων των παραγόντων (συστατικά μίγματος) ή ο πειραματικός χώρος είναι μη κανονικός (Σχ. 4-6) λόγω *περιορισμών* (constraints) στα επίπεδα ενός ή περισσότερων παραγόντων, ή όταν απαιτείται η ανάπτυξη μη γραμμικού μοντέλου, και στις οποίες χρησιμοποιούνται ειδικές τεχνικές σχεδιασμού πειράματος που βασίζονται σε αλγόριθμους που υλοποιούνται με υπολογιστικές εφαρμογές. Μία τέτοια τεχνική αποτελεί και ο **D-optimal πειραματικός σχεδιασμός**, με τον οποίο προσδιορίζεται μία υποομάδα του συνόλου δοκιμών του πλήρους παραγοντικού σχεδιασμού, που δεν καλύπτει απαραίτητα τις συνθήκες ορθογωνιότητας των πλήρους και κλασματικού παραγοντικών σχεδιασμών, και η

οποία εξασφαλίζει την καλύτερη δυνατή κάλυψη του πειραματικού χώρου, ανάλογα με τον επιθυμητό αριθμό δοκιμών και το μοντέλο που εφαρμοστεί.



Σχήμα 4-6 Απεικόνιση πειραματικού τρισδιάστατου πειραματικού χώρου με περιορισμούς στα επίπεδα των παραγόντων

Τέτοιες μέθοδοι χρησιμοποιούν κάποιο κριτήριο για τον προσδιορισμό της βέλτιστης ομάδας δοκιμών, που στην περίπτωση του D-optimal σχεδιασμού είναι η μεγιστοποίηση της ορίζουσας $D = |\mathbf{X}^T \cdot \mathbf{X}|$ του πίνακα διακύμανσης-συνδιακύμανσης. Ο αλγόριθμος καλείται δηλαδή να επιλέξει από δοσμένο σύνολο δοκιμών, που αποτελεί συνήθως όλους τους δυνατούς επιτρεπτούς συνδυασμούς μεταξύ των όλων των επιπέδων τιμών των παραγόντων, εκείνο το σύνολο δεδομένων με την μέγιστη μεταβλητότητα και την ελάχιστη συμμεταβλητότητα.

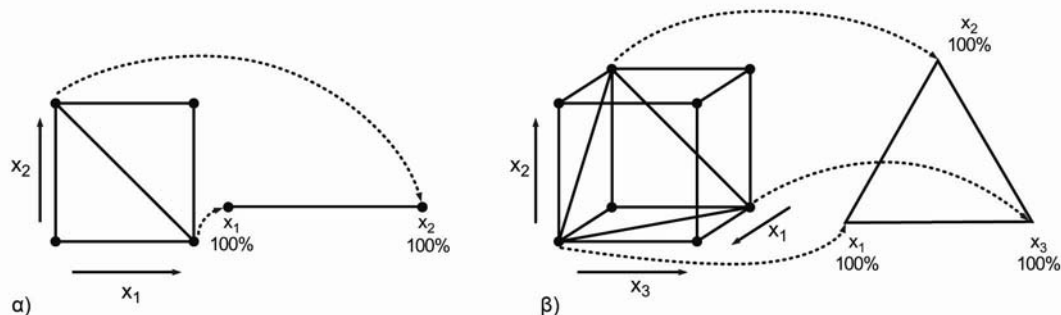
4.4 Πειραματικός Σχεδιασμός σε Μίγματα Συστατικών

Στην ειδική περίπτωση που η πειραματική έρευνα αφορά μίγματα συστατικών, οι παράγοντες σχεδιασμού αποτελούν συγκεντρώσεις των συστατικών του μίγματος, ενώ απόκριση του συστήματος είναι συνήθως κάποια ιδιότητά του. Οι παράγοντες στα πειράματα αυτά, ανεξάρτητα από επιπλέον εξωγενείς περιορισμούς, υπόκεινται στην παρακάτω *αθροιστική συνθήκη*:

$$x_1 + x_2 + \dots + x_k = \sum_{j=1}^k x_j = 1 \quad x_j \geq 0, j=1, \dots, k \quad (4.20)$$

με την οποία οι ανεξάρτητες μεταβλητές του συστήματος μειώνονται σε $k-1$. Ο πειραματικός χώρος **σχεδιασμού σε μίγματα** (mixture design) δεν χρησιμοποιεί ορθοκανονικές συντεταγμένες απεικόνισης, αλλά τις λεγόμενες *βαρυκεντρικές συντεταγμένες* (barycentric coordinates), στις οποίες το εύρος των k παραγόντων μίγματος αποδίδεται με ένα κανονικό

*simplex*¹ $k-1$ διαστάσεων με k κορυφές (Σχ. 4-7). Οι κορυφές των σχημάτων αντιπροσωπεύουν τα σημεία στα οποία ο αντίστοιχος παράγοντας έχει τιμή 1 ή 100% με μηδενική συμμετοχή των υπολοίπων συστατικών, ενώ τα ευθύγραμμα τμήματα που ορίζονται από δύο κορυφές (πλευρές ή ακμές) αντιπροσωπεύουν δυαδικά μίγματα των δύο αυτών συστατικών σε διάφορες αναλογίες ανάλογα με την απόσταση από τις κορυφές.



Σχήμα 4-7 Πειραματικός χώρος α) μίας διάστασης (ευθεία simplex) για δύο συστατικά και β) δύο διαστάσεων (ισοσκελές τρίγωνο simplex) για τρία συστατικά μίγματος

Μοντέλα σε μίγματα συστατικών

Η αθροιστική συνθήκη (4.20) έχει σαν αποτέλεσμα την μείωση των συντελεστών της συναρτησιακής σχέσης μεταξύ των συμμετοχών των συστατικών του μίγματος και της απόκρισης του συστήματος, που πρέπει να προσδιοριστούν κατά την ανάπτυξη του μοντέλου. Έτσι στην περίπτωση των αντίστοιχων διερευνητικών πειραμάτων, για το γραμμικό μοντέλο, θα είναι:

$$y_i = \beta_0 + \sum_{j=1}^k \beta_j x_{ij} + \varepsilon_i \quad (4.21)$$

$$y_i = \beta_0 \sum_{j=1}^k x_{ij} + \sum_{j=1}^k \beta_j x_{ij} + \varepsilon_i \quad (4.22)$$

$$y_i = \sum_{j=1}^k \beta'_j x_{ij} + \varepsilon_i, \text{ με } \beta'_j = \beta_0 + \beta_j \quad (4.23)$$

Επειδή όμως το παραπάνω μοντέλο αποδεικνύεται δύσκολο ως προς την ερμηνεία των αποτελεσμάτων των παραγόντων, στα μίγματα εφαρμόζονται κυρίως τα λεγόμενα **Cox-μοντέλα** (Cox-Models) (Cox, 1971), στα οποία οι άγνωστοι συντελεστές των παραγόντων, και κατ' επέκταση των αλληλεπιδράσεων μεταξύ τους και των τετραγώνων τους, χρησιμοποιούνται

¹ Ως **κανονικό πολύτοπο** (regular simplex) ορίζεται ένα γεωμετρικό σχήμα που έχει $n+1$ κορυφές σε χώρο n διαστάσεων. Έτσι στο χώρο μίας διάστασης είναι ευθύγραμμο τμήμα, των δύο διαστάσεων κανονικό τρίγωνο και των τριών κανονικό τετράεδρο.

για την ερμηνεία των αποτελεσμάτων των παραγόντων στην απόκριση του συστήματος σε σχέση με ένα γνωστό μίγμα αναφοράς (Cox-αποτελέσματα) ή σε σχέση με το κέντρο βάρους του πειραματικού χώρου.

Στα διερευνητικά πειράματα χρησιμοποιείται το γραμμικό Cox-μοντέλο, που έχει την μορφή:

$$y_i = \beta_0 + \sum_{j=1}^k \beta_j x_{ij} + \varepsilon_i, \text{ με την συνθήκη } \sum_{j=1}^k \beta_j r_j = 0 \quad (4.24)$$

όπου r_j η συμμετοχή του συστατικού j στο μίγμα αναφοράς

β_0 η αναμενόμενη απόκριση του μίγματος αναφοράς

β_j ο συντελεστής του j παράγοντα του σχεδιασμού, που αναφέρεται ως Cox-αποτέλεσμα, δηλαδή η αναμενόμενη μεταβολή της απόκρισης λόγω μεταβολής της συμμετοχής του συστατικού j κατά μήκος του Cox-άξονα (Σχ. 4-8)

Για την ανάπτυξη μοντέλου σε διερευνητικά πειράματα απαιτούνται $(k+1)-1=k$ δοκιμές.

Στα πειράματα τόσο μοντελοποίησης, όσο και βελτιστοποίησης¹ χρησιμοποιείται το τετραγωνικό Cox-μοντέλο, που έχει την μορφή:

$$y_i = \beta_0 + \sum_{j=1}^k \beta_j x_{ij} + \sum_{j=1}^{k-1} \sum_{\ell=j}^k \beta_{j,\ell} x_{ij} x_{i\ell} + \sum_{j=1}^k \beta_{jj} x_{ij}^2 + \varepsilon_i \quad (4.25)$$

$$\text{με τις συνθήκες } \sum_{j=1}^k \beta_j r_j = 0, \sum_{\ell=1}^k \beta_{j,\ell} r_j r_\ell = 0, \text{ με } \beta_{j,\ell} = \beta_{\ell,j}, j < \ell$$

Σε μορφή πινάκων ένα Cox-μοντέλο γράφεται:

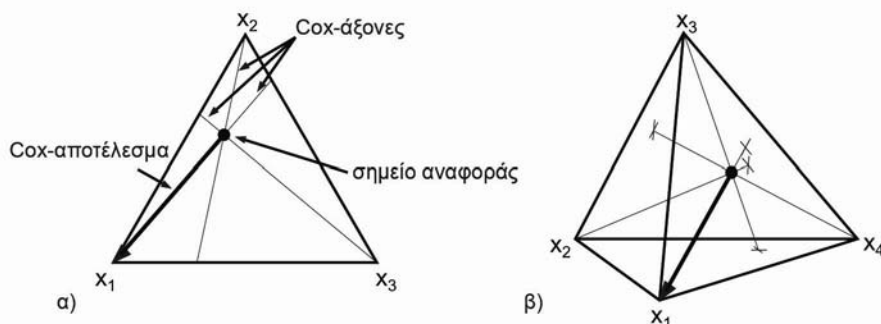
$$\mathbf{y} = \mathbf{X} \cdot \boldsymbol{\beta} + \boldsymbol{\varepsilon}, \text{ με την συνθήκη } \mathbf{R} \cdot \boldsymbol{\beta} = \mathbf{0} \quad (4.26)$$

Ο πίνακας σχεδιασμού $\mathbf{X}$ για το γραμμικό Cox-μοντέλο δίνεται από την (4.7), με διάνυσμα συνθήκης $\mathbf{R} = (0, r_1, \dots, r_k)$, ενώ για το τετραγωνικό Cox-μοντέλο από την (4.18), με πίνακα

$$\text{συνθήκης } \mathbf{R} = \begin{bmatrix} 0 & r_1 & \dots & r_k & 0 & \dots & \dots & 0 & 0 & \dots & 0 & 0 & 0 & \dots & \dots & \dots & 0 \\ 0 & 0 & \dots & 0 & r_1 r_2 & r_1 r_3 & \dots & r_1 r_k & 0 & \dots & 0 & \vdots & r_1^2 & 0 & \dots & \dots & 0 \\ \vdots & \vdots & & \vdots & r_1 r_2 & 0 & \dots & 0 & r_2 r_3 & \dots & r_2 r_k & \vdots & 0 & r_2^2 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & 0 & r_1 r_3 & & \vdots & r_2 r_3 & & 0 & \vdots & 0 & \ddots & & \vdots & \\ \vdots & \vdots & & \vdots & \vdots & & \ddots & \vdots & \ddots & & \vdots & 0 & \vdots & \vdots & & \vdots & 0 \\ 0 & 0 & \dots & 0 & 0 & \dots & \dots & r_1 r_k & 0 & \dots & r_2 r_k & r_{k-1} r_k & 0 & 0 & \dots & 0 & r_k^2 \end{bmatrix}$$

Για τον προσδιορισμό των $k+k(k+1)/2+1$ αγνώστων συντελεστών υπό $(k+1)$ επιπλέον συνθήκες, απαιτούνται $k+k(k+1)/2+1-(k+1)=k(k+1)/2$ δοκιμές.

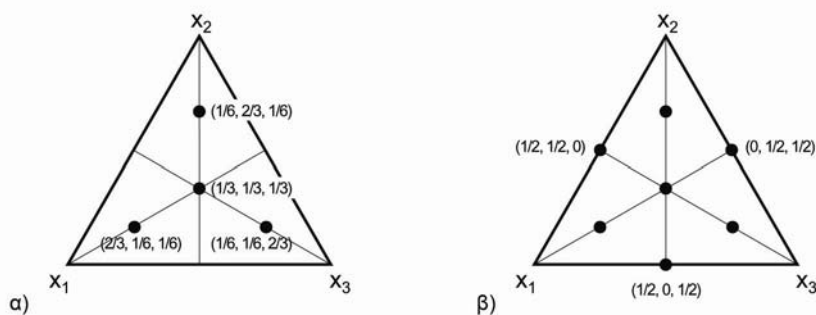
¹ Εκτός του γραμμικού Cox-μοντέλου, στον σχεδιασμό σε μίγματα συστατικών δεν μπορούν να εφαρμοστούν μοντέλα χωρίς τετραγωνικούς όρους, αφού οι οποιοσδήποτε αλληλεπιδράσεις δεν είναι ανεξάρτητες των τετραγώνων των παραγόντων εξαιτίας των συνθηκών για το δείγμα αναφοράς.



Σχήμα 4-8 Γραφική απεικόνιση του Cox-αποτελέσματος κανονικού πειραματικού χώρου α) για τρία και β) για τέσσερα συστατικά μίγματος

Τεχνικές Σχεδιασμού σε Μίγματα Συστατικών

Μια από τις πιο διαδεδομένες τεχνικές κατάστρωσης του πλάνου των δοκιμών στην περίπτωση συστατικών μίγματος είναι ο **αξονικός σχεδιασμός** (axial design), στον οποίο τα σημεία μετρήσεων λαμβάνονται πάνω στους άξονες (Cox-άξονες, Σχ. 4-8) που διέρχονται από τις κορυφές του simplex πειραματικού χώρου¹ και ένα σημείο αναφοράς εσωτερικό του χώρου, σε ίσες αποστάσεις από το τελευταίο, και προαιρετικά στα σημεία που οι άξονες τέμνουν τα όρια του χώρου. Ο αξονικός σχεδιασμός χαρακτηρίζεται ως *n*-τάξης αξονικός σχεδιασμός (*n*-side axial design) αν περιλαμβάνει *n* σημεία ανά άξονα (Σχ. 4-9) και επομένως *nk+1* δοκιμές και χρησιμοποιείται αποκλειστικά σε διερευνητικά πειράματα για ανάπτυξη γραμμικών μοντέλων.

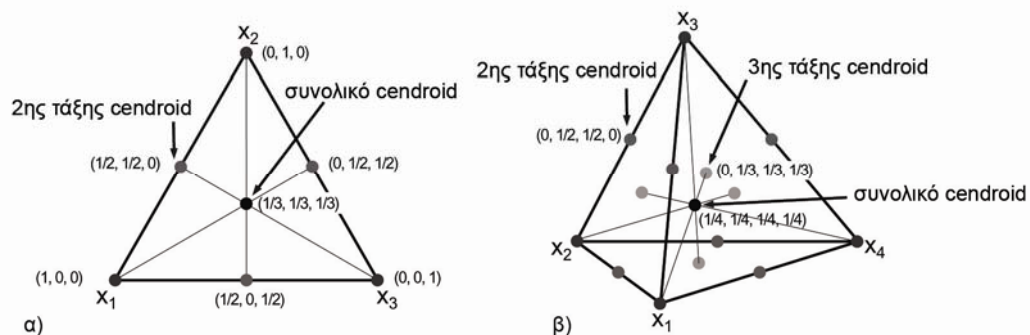


Σχήμα 4-9 Αξονικοί σχεδιασμοί α) one-side και β) two-side για μίγμα τριών συστατικών με το κέντρο βάρους ως σημείο αναφοράς

Στην περίπτωση, που το σημείο αναφοράς συμπίπτει με το κέντρο βάρους του simplex, και τα σημεία μετρήσεων είναι οι κορυφές, τα άκρα των αξόνων και προαιρετικά σημεία πάνω

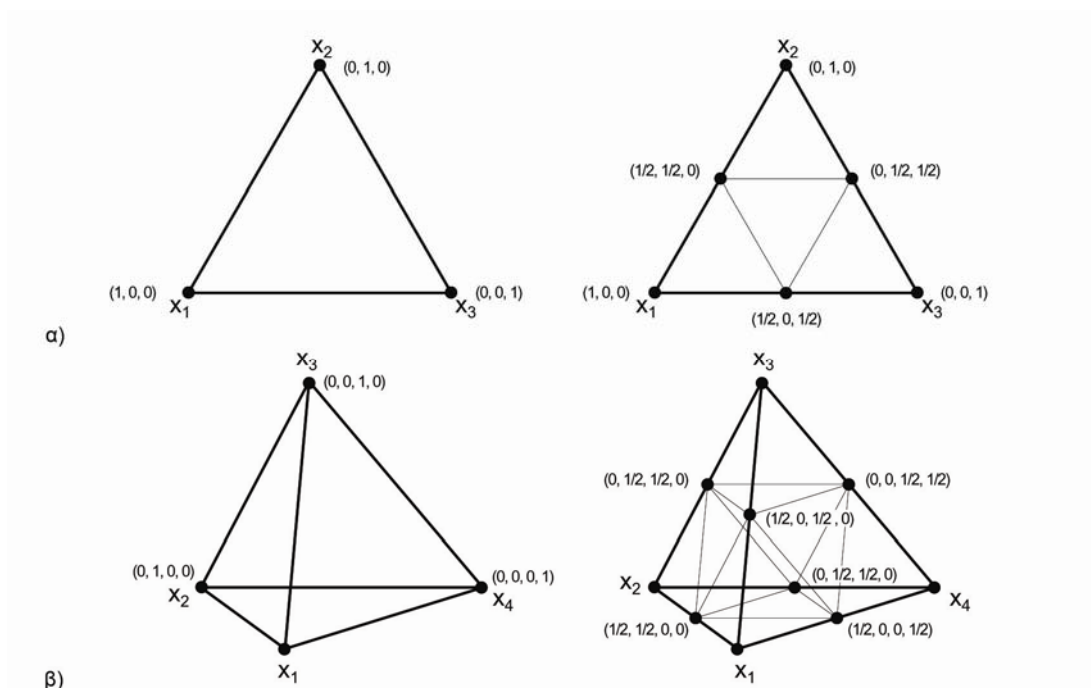
¹ Ισχύει και στην περίπτωση που ο χώρος δεν είναι κανονικό simplex, αλλά τμήμα του, λόγω περιορισμών στο εύρος των συμμετοχών των συστατικών στο μίγμα.

σ' αυτούς σε ίσες αποστάσεις από το κέντρο βάρους, ο σχεδιασμός ονομάζεται **simplex-centroid**. Ως *centroid* σημείο (βαρύκεντρο) μίας γραμμής, ενός επιπέδου και γενικότερα ενός χώρου ορίζεται η μέση τιμή των ακραίων σημείων του χώρου αυτού. Οι κορυφές του simplex (βαρύκεντρα σημείων) αντιστοιχούν σε καθαρά συστατικά, τα centroid δεύτερης τάξης (βαρύκεντρα ακμών του simplex) σε δυαδικά μίγματα ίσης συμμετοχής των δύο αντίστοιχων συστατικών, τα centroid τρίτης τάξης (βαρύκεντρα εδρών του simplex) σε τριαδικά μίγματα ίσης συμμετοχής των τριών αντίστοιχων συστατικών, κ.ο.κ. (Σχ. 4-10). Οι δοκιμές ενός simplex-centroid σχεδιασμού είναι $2^k + 1$, περισσότερες δηλαδή από τις ελάχιστες απαιτούμενες για την κατασκευή του τετραγωνικού μοντέλου, με αποτέλεσμα η τεχνική αυτή να χρησιμοποιείται κυρίως σε προβλήματα βελτιστοποίησης και σε πειράματα όπου το θεωρούμενο μοντέλο περιλαμβάνει και πολυωνυμικούς όρους μεγαλύτερου βαθμού.



Σχήμα 4-10 Σχεδιασμοί simplex-centroid για α) για μίγμα τριών και β) τεσσάρων συστατικών.

Μία άλλη τεχνική σχεδιασμού πειράματος σε μίγματα συστατικών είναι ο **σχεδιασμός {k,m} simplex-lattice**, στον οποίο τα k συστατικά λαμβάνουν $m+1$ ισαπέχοντα επίπεδα τιμών στο διάστημα $(0, 1)$, $x_i = 0, 1/m, 2/m, \dots, 1$ (Σχ. 4-11). Ο αριθμός των δοκιμών σε έναν $\{k,m\}$ simplex-lattice σχεδιασμό είναι $(k+m-1)!/m!(k-1)!$, που για $m=1$ δίνει k δοκιμές, αρκετές δηλαδή για ανάπτυξη γραμμικού μοντέλου, ενώ για $m=2$ δίνει $k(k+1)/2$ που είναι οι ελάχιστες απαιτούμενες για ανάπτυξη τετραγωνικού μοντέλου. Σε αντίθεση με τον αξονικό σχεδιασμό, τόσο ο σχεδιασμός simplex-lattice, όσο και ο simplex-centroid, εκτός του συνολικού centroid σημείου, περιλαμβάνουν μετρήσεις στα όρια του simplex πειραματικού χώρου.



Σχήμα 4-11 Σχεδιασμοί simplex-lattice με $m=1$ και $m=2$ για α) για μίγμα τριών και β) τεσσάρων συστατικών.

Σε ειδικές περιπτώσεις, όπως όταν, εκτός των παραγόντων μίγματος, υπεισέρχονται στην μελέτη και άλλοι παράγοντες διεργασίας (π.χ. θερμοκρασία, πίεση) ή όταν οι περιορισμοί στα επίπεδα των παραγόντων ορίζουν πειραματικούς χώρους στους οποίους είναι αδύνατο να εφαρμοστούν οι παραπάνω τεχνικές, χρησιμοποιούνται συνδυασμοί των παραπάνω τεχνικών, όπως συνδυασμός κλασματικού παραγοντικού με simplex-lattice, ή τεχνικές που υλοποιούνται από αλγοριθμικές υπολογιστικές εφαρμογές, όπως ο **D-optimal πειραματικός σχεδιασμός μιγμάτων**.

5 Πρόβλεψη Αριθμού Οκτανίων σε Βενζίνες και Διερεύνηση Επίδρασης Μίξης

5.1 Μοντέλα Πρόβλεψης Αριθμού Οκτανίων από Δεδομένα Παραγωγής

Δεδομένα Παραγωγής

Για την δημιουργία των αλγορίθμων πρόβλεψης του Α.Ο. που αναπτύχθηκαν στην παρούσα εργασία χρησιμοποιήθηκαν 156 δείγματα βενζινών που αφορούν δεδομένα παραγωγής της MotorOilHellas (ΜΟΗ) στην διάρκεια ενός έτους (Ιανουάριος-Οκτώβριος 2002), ώστε να διασφαλίζεται η απαραίτητη διακύμανση στην σύνθεση των κλασμάτων μίξης λόγω διαφορετικής τροφοδοσίας αργού πετρελαίου ή διαφοροποιήσεων στις λειτουργικές παραμέτρους των επιμέρους μονάδων. Για την ανάπτυξη των αλγορίθμων πρόβλεψης του RON χρησιμοποιήθηκε το σύνολο των καταγεγραμμένων δειγμάτων, ενώ, λόγω έλλειψης δεδομένων, για πρόβλεψη του MON χρησιμοποιήθηκαν μόνο 122 δείγματα. Οι καταγραφές περιελάμβαναν την % κ.ο. περιεκτικότητα των επιμέρους κλασμάτων στο τελικό προϊόν μίξης και τους Α.Ο., τόσο των κλασμάτων όσο και της τελικής βενζίνης. Ο πειραματικός προσδιορισμός των Α.Ο. στα κλάσματα και στις τελικές βενζίνες έγινε στο εργαστήριο της ΜΟΗ, σύμφωνα με τις πρότυπες μεθόδους ASTM D2699-92 και ASTM D2700-92 για τους RON και MON αντίστοιχα. Η μονάδα ανάμιξης-παραγωγής βενζίνης χρησιμοποιεί ως τροφοδοσία προϊόντα από κατεργασίες αναμόρφωσης των κλασμάτων απόσταξης του αργού πετρελαίου και συγκεκριμένα τα προϊόντα καταλυτικής πυρόλυσης ρευστοποιημένης κλίνης (FCC), καταλυτικής αναμόρφωσης (REF), ισομερισμού (ISO), αλκυλίωσης (ALK) και διμερισμού (DIM), μαζί με κλάσματα βουτανίων (C₄) και ισοπεντανίων (C₅) και οξυγονούχα πρόσθετα (MTBE). Εκτός των οκτώ αυτών κύριων κλασμάτων οι βενζίνες παραγωγής περιέχουν συμπληρωματικά ελαφριά νάφθα, αερίελαια (gasoil) και αμόλυβδη βενζίνη 95 RON σε μικρά ποσοστά (0-9.88% κ.ο.).

Τα δείγματα που χρησιμοποιήθηκαν αντιστοιχούν σε προϊόντα τυπικής παραγωγής του διυλιστηρίου και καλύπτουν ένα ευρύ φάσμα ποιοτήτων τελικού προϊόντος οι οποίες χαρακτηρίζονται με βάση τον Α.Ο. τους. Αναλυτικά το σύνολο των δειγμάτων περιελάμβανε:

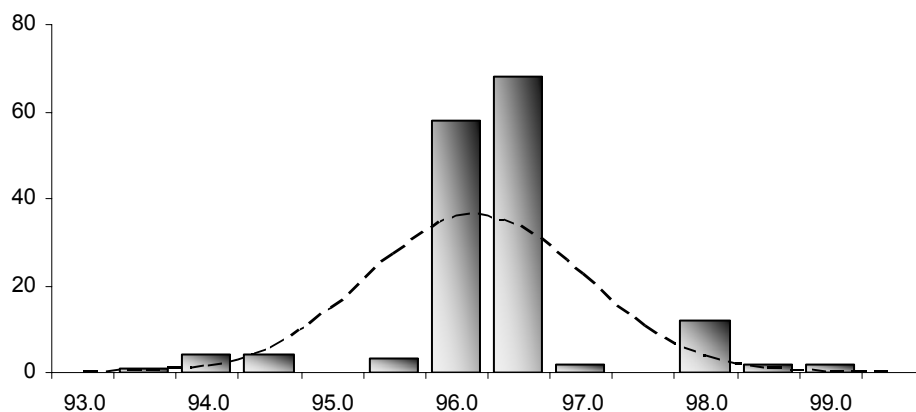
- 51 δείγματα αμόλυβδη βενζίνη 95 RON
- 70 δείγματα αμόλυβδη βενζίνη (LRP) 96RON
- 9 δείγματα αμόλυβδη βενζίνη BP 96RON
- 14 δείγματα αμόλυβδη βενζίνη 98RON (SUPERUNL)
- 2 δείγματα αμόλυβδη βενζίνη με 99RON (SUPERUNL)
- 4 δείγματα BENZEXPO (gasoline export)

- 6 δείγματα GASOLINE (tanks-gasoline from units)

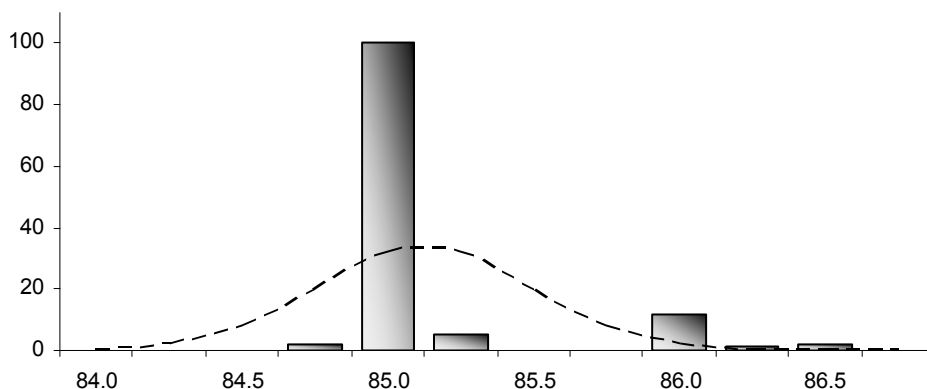
Αντίστοιχα τα δείγματα που χρησιμοποιήθηκαν για την πρόβλεψη του MON περιελάμβαναν:

- 44 δείγματα αμόλυβδη βενζίνη 95 RON
- 56 δείγματα αμόλυβδη βενζίνη (LRP) 96RON
- 6 δείγματα αμόλυβδη βενζίνη BP 96RON
- 13 δείγματα αμόλυβδη βενζίνη 98RON (SUPERUNL)
- 2 δείγματα αμόλυβδη βενζίνη με 99RON (SUPERUNL)
- 1 δείγμα BENZEXPO (gasoline export)

Οι κατανομές των RON και MON στον συνολικό αντίστοιχο πληθυσμό δίνονται υπό μορφή ιστογραμμάτων στα Σχ. 5-1 και Σχ. 5-2. Είναι εμφανές ότι, για το μεγαλύτερο ποσοστό των δειγμάτων, οι Α.Ο. είναι περιορισμένοι σε μικρό εύρος τιμών (95.5-96.5 RON και 98 RON, 85 MON).



Σχήμα 5-1 Ιστογράμμα του αριθμού RON του τελικού προϊόντος (156 δείγματα)



Σχήμα 5-2 Ιστογράμμα του αριθμού MON του τελικού προϊόντος (122 δείγματα)

Για την πρόβλεψη του RON αναπτύχθηκαν δύο μοντέλα νευρωνικών δικτύων. Στο πρώτο ως ανεξάρτητες μεταβλητές χρησιμοποιήθηκαν απλά οι % κ.ο. περιεκτικότητες των 8 κύριων κλασμάτων μίξης, ενώ στο δεύτερο τα γινόμενα αυτών επί τον RON των επιμέρους κλασμάτων, έτσι ώστε το μοντέλο να διατηρεί την ακρίβεια πρόβλεψης, ακόμα και όταν μεταβάλλεται η ποιότητα της τροφοδοσίας στο χρόνο. Παρόλο που τα δύο μοντέλα δεν παρουσίασαν σημαντικές διαφορές στην ακρίβεια πρόβλεψης, ακριβώς με στόχο την ευελιξία αυτή των μοντέλων, υιοθετήθηκε γενικότερα σε όλες τις επόμενες προβλέψεις η δεύτερη προσέγγιση. Για την πρόβλεψη του MON το μοντέλο που αναπτύχθηκε χρησιμοποιεί επίσης ως εισόδους τους αριθμούς MON των επιμέρους κλασμάτων σταθμισμένους με τις % κ.ο. περιεκτικότητές τους.

Τα κλάσματα που χρησιμοποιούνται για παραγωγή της τελικής βενζίνης και τα όρια διακύμανσης (min-max) των Α.Ο. τους στον πληθυσμό μελέτης, δίδονται στον Πιν. 5-1.

Πίνακας 5-1 Κλάσματα βενζίνης και όρια διακύμανσης του Α.Ο. τους

Κλάσμα		Αριθμός Οκτανίων (RON)		Αριθμός Οκτανίων (MON)	
		min	max	min	max
Βενζίνη καταλυτικής πυρόλυσης	FCC	92.4	94.0	80.1	81.2
Βενζίνη καταλυτικής αναμόρφωσης	REF	100.1	103.3	89.4	92.1
Βενζίνη ισομερισμού	ISO	77.5	84.9	79.7	82.6
Βενζίνη αλκυλίωσης	ALK	93.3	94.6	91.2	92.1
Βενζίνη διμερισμού	DIM	93.8	97.6	79.7	81.8
Οξυγονούχα πρόσθετα	MTBE	115.0		95.0	
Βουτάνια	C ₄	92.0		82.0	
Ισοπεντάνια	C ₅	101.7		90.3	

Υφιστάμενοι Αλγόριθμοι Πρόβλεψης Αριθμού Οκτανίων

Όπως έχει ήδη συζητηθεί παραπάνω ο αριθμός οκτανίου δεν είναι προσθετική ιδιότητα της βενζίνης. Αυτό μπορεί να φανεί με την σύγκριση των μετρούμενων αριθμών οκτανίων (RON και MON) με τους αριθμούς οκτανίων (RON_{ad} και MON_{ad}), που προκύπτουν αθροιστικά από τους αριθμούς οκτανίων των επιμέρους κλασμάτων ζυγισμένους με τις %κ.ο. συγκεντρώσεις τους στην τελική βενζίνη. Οι διαφορές των προσθετικά υπολογιζόμενων τιμών και των αντίστοιχων πειραματικών κυμαίνονται από -2.3 έως 1.6 για τον RON και -1.2 έως 2.1 για τον MON.

Παρόλο που δεν υπάρχει ασφαλής μέθοδος προσδιορισμού της συγκέντρωσης των επιμέρους κλασμάτων που να διασφαλίζει τον επιθυμητό Α.Ο. της παραγόμενης βενζίνης, η ανάμιξη στην μονάδα της MOH γίνεται με την χρήση ημι-εμπειρικών εξισώσεων που χρησιμοποιούν για τον

προσδιορισμό του Α.Ο. της τελικής βενζίνης τους *συντελεστές ανάμιξης* του κάθε κλάσματος μεμονωμένα και του μίγματος. Το μοντέλο αυτό, παρόμοια με την μέθοδο των αριθμών οκτανίων μίγματος (BONs) (Gary και Handwerl (1994)), αποτελεί στην πραγματικότητα μία εξέλιξη του απλού προσθετικού μοντέλο που διατηρεί την φιλοσοφία του τελευταίου για γραμμική συσχέτιση των ογκομετρικών περιεκτικοτήτων με τους επιμέρους Α.Ο. των κλασμάτων ενώ ταυτόχρονα εισάγει την υφιστάμενη μη γραμμικότητα στο μοντέλο.

Στην βάση του προσθετικού μοντέλου, το οποίο υπολογίζει τον Α.Ο. του μίγματος R ως άθροισμα των Α.Ο. R_i των κλασμάτων μίξης, σταθμισμένο με τις ογκομετρικές περιεκτικότητες τους y_i στο μίγμα,

$$R = \sum y_i R_i \quad (5.1)$$

το παραπάνω μοντέλο υπολογίζει τον τροποποιημένο Α.Ο. του μίγματος F (συντελεστής ανάμιξης μίγματος) ως άθροισμα των τροποποιημένων Α.Ο. των κλασμάτων μίξης F_i (συντελεστής ανάμιξης συστατικού), σταθμισμένο με τις ογκομετρικές περιεκτικότητες τους y_i στο μίγμα:

$$F = \sum y_i F_i \quad (5.2)$$

Οι πραγματικοί Α.Ο. R και R_i εξαρτώνται μη γραμμικά από τους συντελεστές ανάμιξης F και F_i αντίστοιχα:

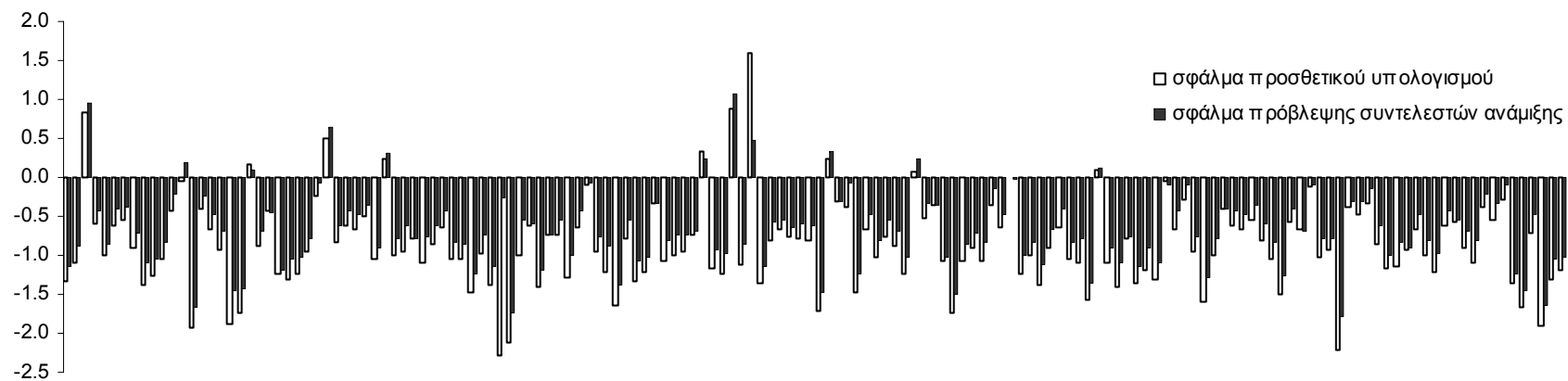
$$R = f(F) \quad (5.3)$$

$$F_i = h(R_i) \quad (5.4)$$

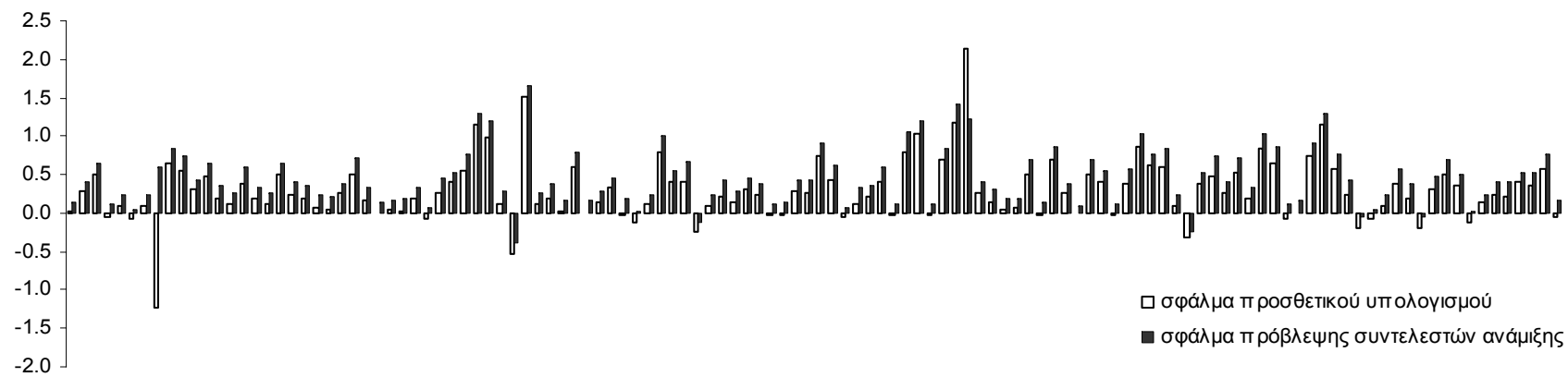
όπου f και h κλαδικές πολυωνυμικές συναρτήσεις με σημείο αλλαγής τύπου συνάρτησης κάποια οριακή τιμή του F και του R_i αντίστοιχα.

Οι εξισώσεις που εφαρμόζει το μοντέλο αυτό για τον υπολογισμό του Α.Ο. του μίγματος παρατίθενται στον Παράρτημα (§Π.2).

Το σφάλμα προσδιορισμού του Α.Ο. υπολογιζόμενου προσθετικά, καθώς και με χρήση των ημι-εμπειρικών αυτών εξισώσεων για τα δείγματα που εξετάστηκαν, απεικονίζεται στα Σχ. 5-3 και Σχ. 5-4. Το σφάλμα των υπολογισμών έχει μέση τιμή -0.86 και -0.66 RON και 0.29 και 0.46 MON αντίστοιχα. Όπως παρατηρείται, τόσο ο προσθετικός υπολογισμός, όσο και η πρόβλεψη των παραπάνω εξισώσεων της MOH, τείνει να υπερεκτιμά τον πραγματικό αριθμό MON του τελικού μίγματος, ενώ αντίστοιχα υποεκτιμά τον RON.



Σχήμα 5-3 Σφάλμα υπολογισμού του RON προσθετικά και με χρήση εξισώσεων συντελεστών ανάμιξης



Σχήμα 5-4 Σφάλμα υπολογισμού του MON προσθετικά και με χρήση εξισώσεων συντελεστών ανάμιξης

Ανάπτυξη Μοντέλων Πρόβλεψης A.O. με Τεχνητά Νευρωνικά Δίκτυα

Προεπεξεργασία δεδομένων

Για τις ανάγκες ανάπτυξης των μοντέλων, από το σύνολο των δειγμάτων δημιουργήθηκαν σε κάθε περίπτωση δύο υποσύνολα: το σύνολο εκπαίδευσης του τεχνητού νευρωνικού δικτύου (ΤΝΔ), και το σύνολο ελέγχου. Η επιλογή αυτή προέκυψε μετά από τυχαία διάταξη των δειγμάτων, αλλά με παράλληλο έλεγχο των μεγίστων –ελαχίστων τιμών των εξαρτημένων και ανεξάρτητων μεταβλητών της ανάλυσης, ώστε το σύνολο ελέγχου να βρίσκεται εντός του δειγματικού χώρου που ορίζει το σύνολο μελέτης, δηλαδή οι τιμές όλων των δειγμάτων ελέγχου να βρίσκονται εντός του εύρους που ορίζει για κάθε μεταβλητή το σύνολο μελέτης (Πιν. 5-2). Το πλήθος των δειγμάτων ελέγχου είναι περίπου 10% των συνολικών διαθέσιμων δειγμάτων.

Η προεπεξεργασία των δεδομένων έγινε εσωτερικά στον κώδικα εκπαίδευσης του δικτύου που αναπτύχθηκε σε περιβάλλον Matlab. Για το σύνολο εκπαίδευσης, χρησιμοποιήθηκε η συνάρτηση *premnmx* του Matlab που κανονικοποιεί τα δεδομένα εισαγωγής και τους στόχους στο διάστημα [-1, 1] με βάση την εξίσωση

$$x_{n_{ij}} = 2(x_{ij} - x_{jmin}) / (x_{jmax} - x_{jmin}) - 1 \quad (5.5)$$

ενώ για τα δεδομένα του συνόλου ελέγχου χρησιμοποιήθηκε η *tramnmx* που εφαρμόζει αντίστοιχη κανονικοποίηση ως προς τα όρια του συνόλου εκπαίδευσης.

Μοντέλα πρόβλεψης RON

Το πρώτο μοντέλο πρόβλεψης του RON εκπαιδεύτηκε με εισόδους τις % κ.ο περιεκτικότητες των επιμέρους κλασμάτων μίξης, θεωρώντας ότι οι A.O. των τελευταίων είναι σταθεροί, το οποίο εμφανίζει αρκετά ικανοποιητικά αποτελέσματα. Παρόλο όμως που οι A.O. των επιμέρους κλασμάτων δεν μεταβάλλονται σημαντικά (Πίν. 5-1), είναι γεγονός ότι οι διαφορές τους επιδρούν καθοριστικά στην διαμόρφωση του A.O. του τελικού μίγματος. Για την αξιοποίηση, επομένως, όλης της δυνατής πληροφορίας των διαθέσιμων δεδομένων επιλέχθηκε στα επόμενα μοντέλα να χρησιμοποιηθούν ως είσοδοι οι A.O. των επιμέρους κλασμάτων ζυγισμένοι με τις αντίστοιχες συστάσεις τους στο μίγμα.

Τα όρια διακύμανσης των συγκεντρώσεων (% κ.ο.) των επιμέρους κλασμάτων στο τελικό μίγμα, καθώς και τα όρια του A.O. (RON) των τελικών βενζινών για τον πληθυσμό των 156 δειγμάτων, που χρησιμοποιήθηκαν ως ανεξάρτητες και εξαρτημένες αντίστοιχα μεταβλητές στην ανάλυση, παρουσιάζονται στον Πίν. 5-2. Στον ίδιο πίνακα δίνεται και η σύσταση μίας τυπικής βενζίνης παραγωγής του διυλιστηρίου.

Πίνακας 5-2 Όρια διακύμανσης του A.O. (RON) και των συγκεντρώσεων κλασμάτων (% κ.ο.) για τις βενζίνες παραγωγής της MOH

	FCC	REF	ISO	ALK	MTBE	DIM	C ₄	C ₅	RON
min	6.60	0.73	0.00	0.00	0.00	0.00	0.00	0.00	93.5
max	97.14	43.37	32.02	12.32	12.87	14.15	7.42	7.31	99.0
τυπικές τιμές	36	29	15	7	4	4	3	2	

Και στα δύο μοντέλα πρόβλεψης του RON, η αρχιτεκτονική του ΤΝΔ ευθείας προώθησης που υλοποιήθηκε σε περιβάλλον Matlab ήταν 8×6×5×1, με συναρτήσεις μεταφοράς τις *tansig*, *tansig* και *purelin* του Matlab για κάθε επίπεδο αντίστοιχα. Η *tansig* εφαρμόζει στην ενεργοποίηση του αντίστοιχου επιπέδου υπερβολική εφαπτομενική σιγμοειδή συνάρτηση μεταφοράς για την διαμόρφωση των ημιγραμμικών κόμβων του συστήματος, εισάγοντας έτσι μη γραμμικότητα στο μοντέλο.

$$g(\alpha) = \text{tansig}(\alpha) = \frac{2}{(1 + e^{-2\alpha}) - 1} \quad (5.6)$$

ενώ η *purelin* είναι η ταυτοτική συνάρτηση για την υλοποίηση γραμμικών κόμβων εξόδου του δικτύου:

$$g(\alpha) = \text{purelin}(\alpha) = \alpha \quad (5.7)$$

Τα ΤΝΔ εκπαιδεύτηκαν με χρήση του αλγορίθμου *traingdx* του προγράμματος. Πρόκειται για αλγόριθμο ανάδρομης διάδοσης που εφαρμόζει τον γενικευμένο κανόνα εκμάθησης καθόδου κλίσης με μεταβλητό ρυθμό εκμάθησης και όρο ορμής σε ομαδική εκπαίδευση. Ο έλεγχος καλής γενίκευσης του δικτύου έγινε με παράλληλη παρακολούθηση των σφαλμάτων για τα σύνολα εκπαίδευση και ελέγχου κατά την πορεία της εκπαίδευσης και έλεγχο της κανονικότητας της κατανομής των σφαλμάτων. Τα σύνολα εκπαίδευσης και ελέγχου ήταν τα ίδια και για τα δύο μοντέλα, με πλήθος δειγμάτων 140 και 16 αντίστοιχα.

Κατά την διεξαγωγή της παρούσας εργασίας εκπαιδεύτηκαν αρχικά διάφορα ΤΝΔ για την επιλογή της κατάλληλης τοπολογίας με δοκιμές διαφορετικών αρχιτεκτονικών και σύγκριση των αποδόσεων των δικτύων κάθε φορά σε σχέση με το σφάλμα του συνόλου ελέγχου, της κανονικότητας της κατανομής των σφαλμάτων και του μεγέθους του διανύσματος των βαρών. Σε επόμενο στάδιο ακολούθησε η εκπαίδευση ενός μεγάλου αριθμού ΤΝΔ της ίδιας τοπολογίας προκειμένου να εξασφαλιστεί ότι τα μεμονωμένα δίκτυα περιγράφουν εξίσου καλά την εξάρτηση μεταξύ του RON και των ανεξάρτητων μεταβλητών που χρησιμοποιήθηκαν κάθε φορά ως εισαγωγές. Τα στατιστικά αποτελέσματα των “καλύτερης απόδοσης” μοντέλων πρόβλεψης του RON, που περιλαμβάνουν το μέσο απόλυτο σφάλμα (MAEP), τη τυπική απόκλιση (SEP), το μέσο τετραγωνικό σφάλμα (RMSEP) και τα ελάχιστο και μέγιστο απόλυτα σφάλματα της

πρόβλεψης (minAEP και maxAEP), δίνονται στους πίνακες Πιν. 5-3 και Πιν. 5-4. Η ακρίβεια των μοντέλων κρίνεται ικανοποιητική δεδομένου ότι η επαναληψιμότητα μέτρησης της πρότυπης μεθόδου ASTM είναι 0.2-0.3 RON.

Πίνακας 5-3 Σφάλματα του μοντέλου πρόβλεψης RON με εισόδους τις % κ.ο. περιεκτικότητες των επιμέρους κλασμάτων μίξης για τα δείγματα του συνόλου εκπαίδευσης και ελέγχου

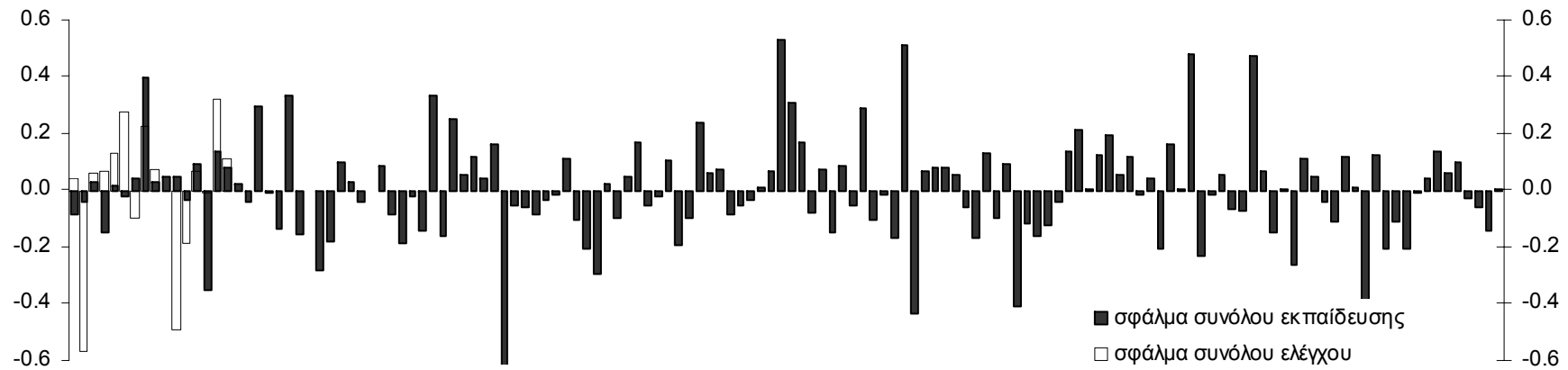
Σύνολο δειγμάτων	Αριθμός δειγμάτων	MAEP	SEP	RMSEP	minAEP	maxAEP
εκπαίδευσης	140	0.1	0.2	0.0	0.0	0.6
ελέγχου	16	0.2	0.2	0.1	0.0	0.6

Πίνακας 5-4 Σφάλματα του μοντέλου πρόβλεψης RON με εισόδους τους αριθμούς RON των επιμέρους κλασμάτων μίξης ζυγισμένους με % κ.ο. περιεκτικότητες για τα δείγματα του συνόλου εκπαίδευσης και ελέγχου

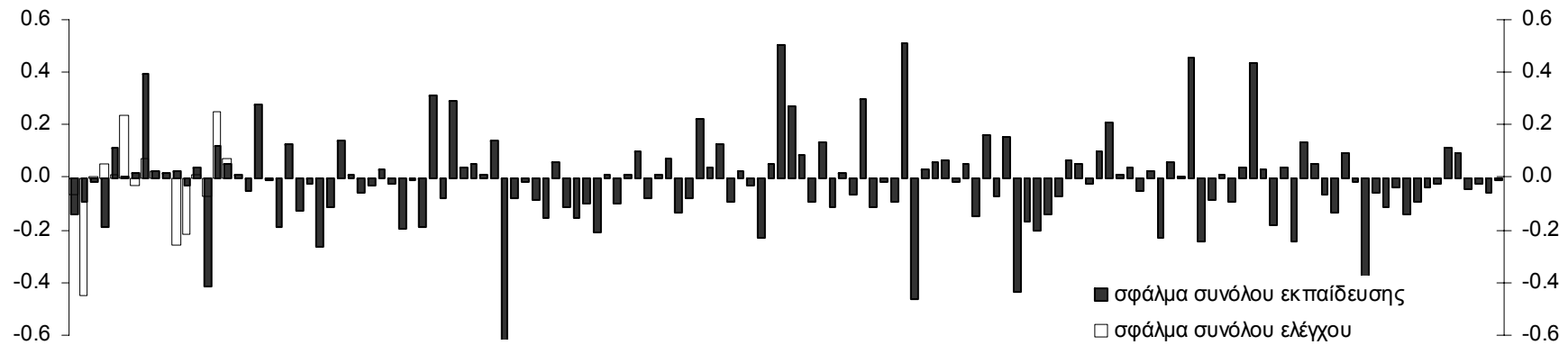
Σύνολο δειγμάτων	Αριθμός δειγμάτων	MAEP	SEP	RMSEP	minAEP	maxAEP
εκπαίδευσης	140	0.1	0.2	0.0	0.0	0.6
ελέγχου	16	0.1	0.2	0.0	0.0	0.4

Οι τιμές των σφαλμάτων πρόβλεψης για τα δύο μοντέλα που αναπτύχθηκαν δίνονται στα διαγράμματα των Σχ. 5-5 και Σχ. 5-6.

.

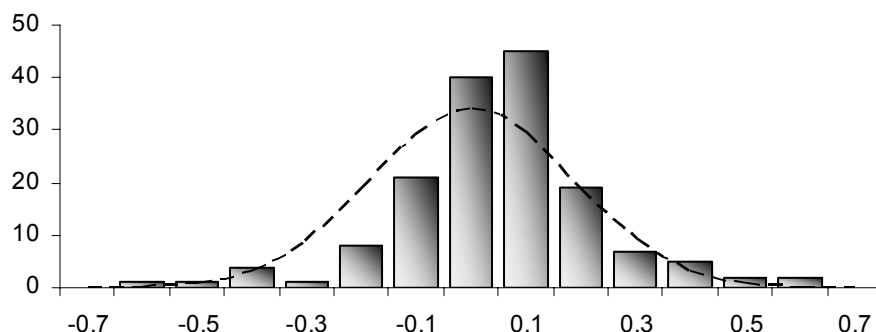


Σχήμα 5-5 Διάγραμμα σφαλμάτων πρόβλεψης του RON του μοντέλου με εισόδους τις % κ.ο. περιεκτικότητες των επιμέρους κλασμάτων μίξης για τα δείγματα του συνόλου εκπαίδευσης και ελέγχου.

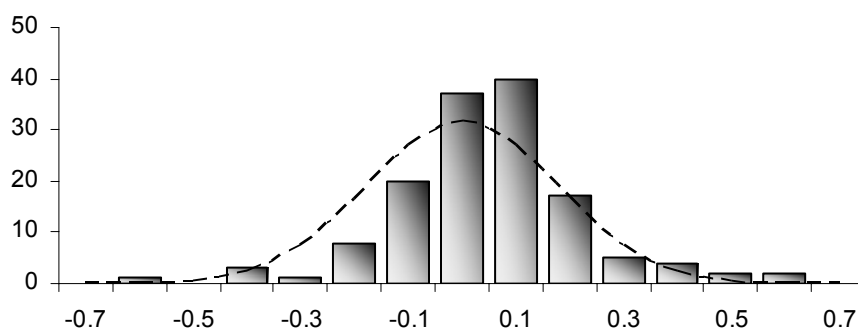


Σχήμα 5-6 Διάγραμμα σφαλμάτων πρόβλεψης του RON του μοντέλου με εισόδους τους αριθμούς RON των επιμέρους κλασμάτων μίξης ζυγισμένους με % κ.ο. περιεκτικότητες για τα δείγματα του συνόλου εκπαίδευσης και ελέγχου.

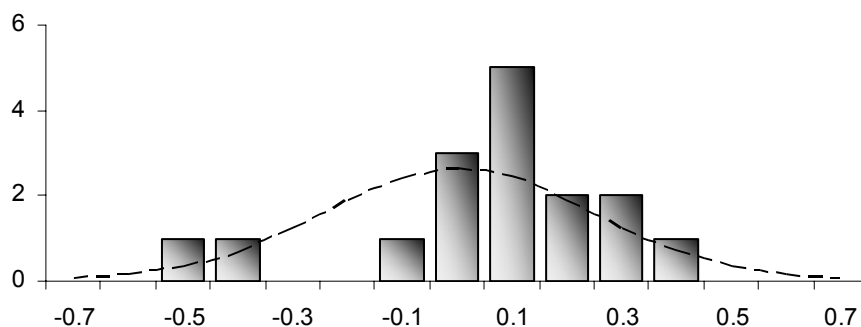
Στα επόμενα διαγράμματα (Σχ. 5-7 ως Σχ. 5-12) δίνονται υπό μορφή ιστογραμμάτων οι κατανομές των σφαλμάτων πρόβλεψης των τελικών μοντέλων πρόβλεψης του RON για το σύνολο των δειγμάτων και για τα δείγματα του συνόλου εκπαίδευσης και ελέγχου.



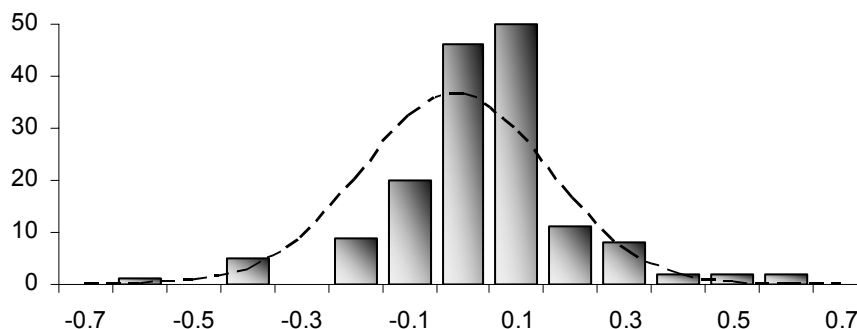
Σχήμα 5-7 Κατανομή των σφαλμάτων πρόβλεψης του RON για το σύνολο των δειγμάτων για το μοντέλο με εισόδους τις % κ.ο. περιεκτικότητες των επιμέρους κλασμάτων μίξης



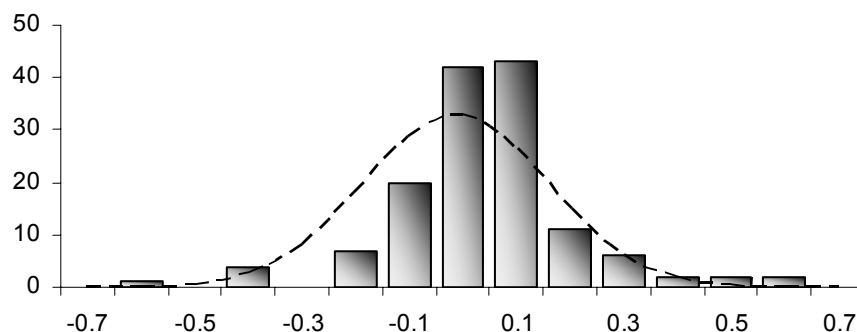
Σχήμα 5-8 Κατανομή των σφαλμάτων πρόβλεψης του RON για τα δείγματα του συνόλου εκπαίδευσης για το μοντέλο με εισόδους τις % κ.ο. περιεκτικότητες των επιμέρους κλασμάτων μίξης



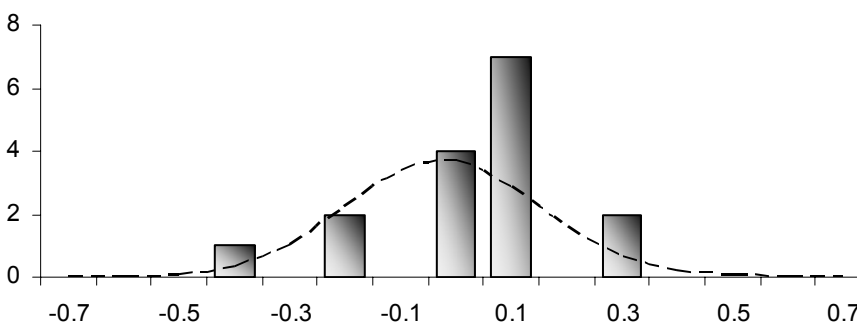
Σχήμα 5-9 Κατανομή των σφαλμάτων πρόβλεψης του RON για τα δείγματα του συνόλου ελέγχου για το μοντέλο με εισόδους τις % κ.ο. περιεκτικότητες των επιμέρους κλασμάτων μίξης



Σχήμα 5-10 Κατανομή των σφαλμάτων πρόβλεψης του RON για το σύνολο των δειγμάτων για το μοντέλο με εισόδους τους αριθμούς RON των επιμέρους κλασμάτων μίξης ζυγισμένους με % κ.ο. περιεκτικότητες



Σχήμα 5-11 Κατανομή των σφαλμάτων πρόβλεψης του RON για τα δείγματα του συνόλου εκπαίδευσης για το μοντέλο με εισόδους τους αριθμούς RON των επιμέρους κλασμάτων μίξης ζυγισμένους με % κ.ο. περιεκτικότητες



Σχήμα 5-12 Κατανομή των σφαλμάτων πρόβλεψης του RON για τα δείγματα του συνόλου ελέγχου για το μοντέλο με εισόδους τους αριθμούς RON των επιμέρους κλασμάτων μίξης ζυγισμένους με % κ.ο. περιεκτικότητες

Μοντέλο πρόβλεψης MON

Τα μοντέλα πρόβλεψης του MON που αναπτύχθηκαν χρησιμοποιούν ως εισόδους τους αριθμούς MON των επιμέρους κλασμάτων ζυγισμένους με τις αντίστοιχες συστάσεις τους στο μίγμα. Η αρχιτεκτονική των ΤΝΔ που εκπαιδεύτηκαν ήταν $8 \times 5 \times 4 \times 1$, με συναρτήσεις μεταφοράς τις *tansin*, *tansin* και *purelin* του Matlab για κάθε επίπεδο αντίστοιχα. Το πλήθος των δειγμάτων για τα σύνολα εκπαίδευσης και ελέγχου ήταν 109 και 13 αντίστοιχα.

Τα όρια διακύμανσης των συγκεντρώσεων (% κ.ο.) των επιμέρους κλασμάτων καθώς και του Α.Ο. (MON) των τελικών βενζινών, που χρησιμοποιήθηκαν ως εξαρτημένες και ανεξάρτητες αντίστοιχα μεταβλητές στην ανάλυση, δίνονται στον Πίν. 5-5.

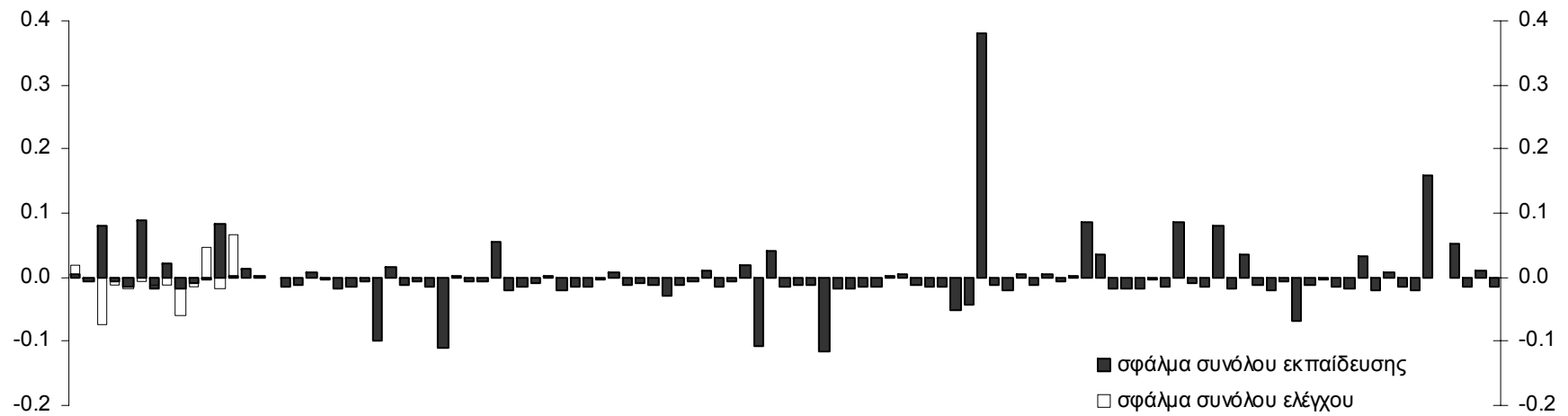
Πίνακας 5-5 Όρια διακύμανσης του Α.Ο. (MON) και των συγκεντρώσεων κλασμάτων (% κ.ο.) για τις βενζίνες παραγωγής της MOH

	FCC	REF	ISO	ALK	MTBE	DIM	C ₄	C ₅	RON
min	6.60	21.40	0.00	0.00	0.00	0.00	0.00	0.00	84.6
max	41.13	43.37	32.02	15.99	12.87	14.15	7.42	2.95	86.5

Στον Πιν. 5-6 δίνονται τα στατιστικά αποτελέσματα του μοντέλου πρόβλεψης του MON που παρουσίασε την καλύτερη απόδοση, ενώ στο διάγραμμα του Σχ. 5-13 παρουσιάζονται οι τιμές των σφαλμάτων για τα δείγματα που χρησιμοποιήθηκαν. Η ακρίβεια του μοντέλου κρίνεται ικανοποιητική δεδομένου ότι η επαναληψιμότητα μέτρησης της πρότυπης μεθόδου ASTM είναι 0.3-0.5 MON.

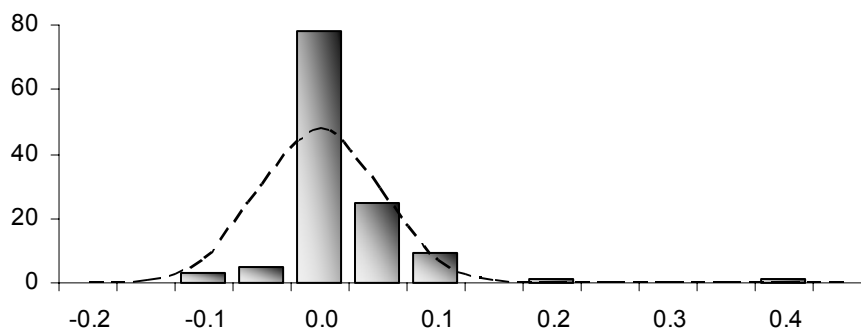
Πίνακας 5-6 Σφάλματα του μοντέλου πρόβλεψης του αριθμού MON με εισόδους τους MON των επιμέρους κλασμάτων μίξης ζυγισμένους με % κ.ο. περιεκτικότητες για τα δείγματα του συνόλου εκπαίδευσης και ελέγχου

Σύνολο δειγμάτων	Αριθμός δειγμάτων	MAEP	SEP	RMSEP	minAEP	maxAEP
εκπαίδευσης	109	0.0	0.1	0.0	0.0	0.4
ελέγχου	13	0.0	0.0	0.0	0.0	0.1

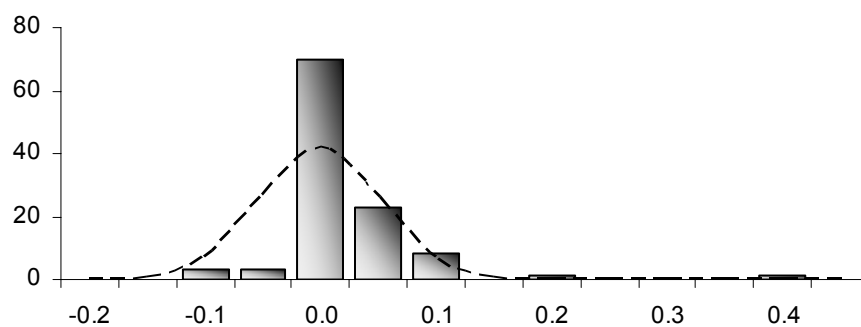


Σχήμα 5-13 Διάγραμμα σφαλμάτων πρόβλεψης του MON του μοντέλου με εισόδους τους αριθμούς MON των επιμέρους κλασμάτων μίξης ζυγισμένους με % κ.ο. περιεκτικότητες για τα δείγματα του συνόλου εκπαίδευσης και ελέγχου.

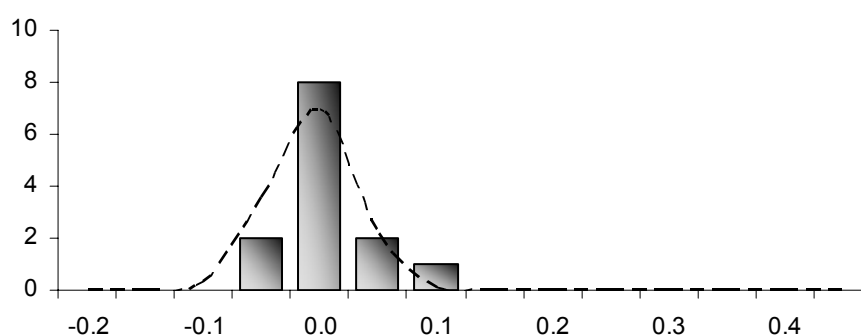
Στα επόμενα διαγράμματα (Σχ. 5-14 ως Σχ. 5-16) δίνονται υπό μορφή ιστογραμμάτων οι κατανομές των σφαλμάτων πρόβλεψης του τελικού μοντέλου πρόβλεψης του αριθμού MON για το σύνολο των δειγμάτων και για τα δείγματα του συνόλου εκπαίδευσης και ελέγχου.



Σχήμα 5-14 Κατανομή των σφαλμάτων πρόβλεψης του MON για το σύνολο των δειγμάτων



Σχήμα 5-15 Κατανομή των σφαλμάτων πρόβλεψης του MON για τα δείγματα του συνόλου εκπαίδευσης

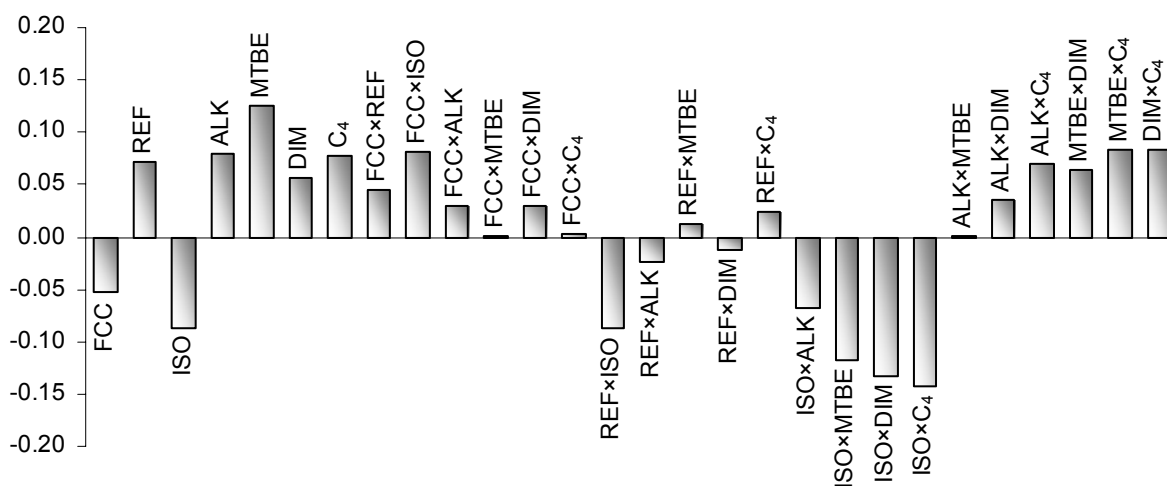


Σχήμα 5-16 Κατανομή των σφαλμάτων πρόβλεψης του MON για τα δείγματα του συνόλου ελέγχου

5.2 Διερεύνηση Επίδρασης της Μίξης στον Α.Ο.

Με στόχο την διερεύνηση των αλληλεπιδράσεων μεταξύ των κλασμάτων κατά την μίξη τους και την επίδραση που έχουν στον τελικό Α.Ο., τα ίδια δεδομένα παραγωγής που χρησιμοποιήθηκαν για την ανάπτυξη των αλγορίθμων πρόβλεψης του RON με νευρωνικά δίκτυα, υπεβλήθησαν και σε ανάλυση παλινδρόμησης κυρίων συνιστωσών (Principal Component Regression PCR). Στην μοντελοποίηση του Α.Ο. χρησιμοποιήθηκαν ως δεδομένα εισαγωγής μόνο οι τρεις πρώτες κύριες συνιστώσες των αρχικών μεταβλητών σύστασης οι οποίες αποδίδουν τις κύριες τάσεις αποκόπτοντας έτσι το μεγαλύτερο ποσοστό του “πειραματικού θορύβου”. Επιπλέον των οκτώ αρχικών μεταβλητών σύστασης χρησιμοποιήθηκαν ως δεδομένα εισαγωγής και τα γινόμενά τους ανά δύο, ώστε να διαπιστωθεί το μέγεθος της αλληλεπίδρασής τους.

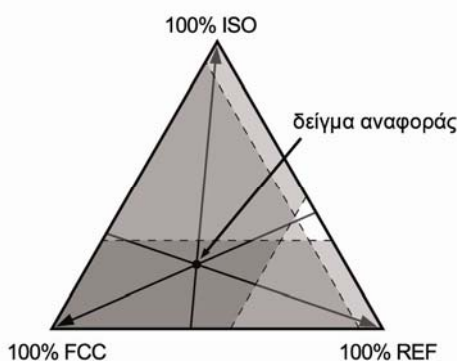
Η ανάλυση πραγματοποιήθηκε χρησιμοποιώντας το χημειομετρικό λογισμικό Unscrambler 7.5 της Camo ASA. Για την πρόβλεψη του τελικού Α.Ο. (RON) χρησιμοποιήθηκαν οι τρεις πρώτες κύριες συνιστώσες (PCs), οι οποίες αποδίδουν το 70% της συνολικής διασποράς του αρχικού πίνακα δεδομένων. Οι συντελεστές του γραμμικού μοντέλου που αναπτύχθηκε παρουσιάζονται στο διάγραμμα του Σχ. 5-17. Το μέγεθος του κάθε συντελεστή είναι χαρακτηριστικό του βαθμού σημαντικότητας του αντίστοιχου κλάσματος στην διαμόρφωση του τελικού αριθμού οκτανίων, ενώ το πρόσημο υποδηλώνει την θετική ή αρνητική επίδρασή του αντίστοιχα. Οι συντελεστές των γινομένων των μεταβλητών χαρακτηρίζουν ποιοτικά (αλληλεπίδραση συνέργειας ή ανταγωνισμού) και ποσοτικά την αλληλεπίδραση που υφίσταται κατά την μίξη των αντίστοιχων κλασμάτων.



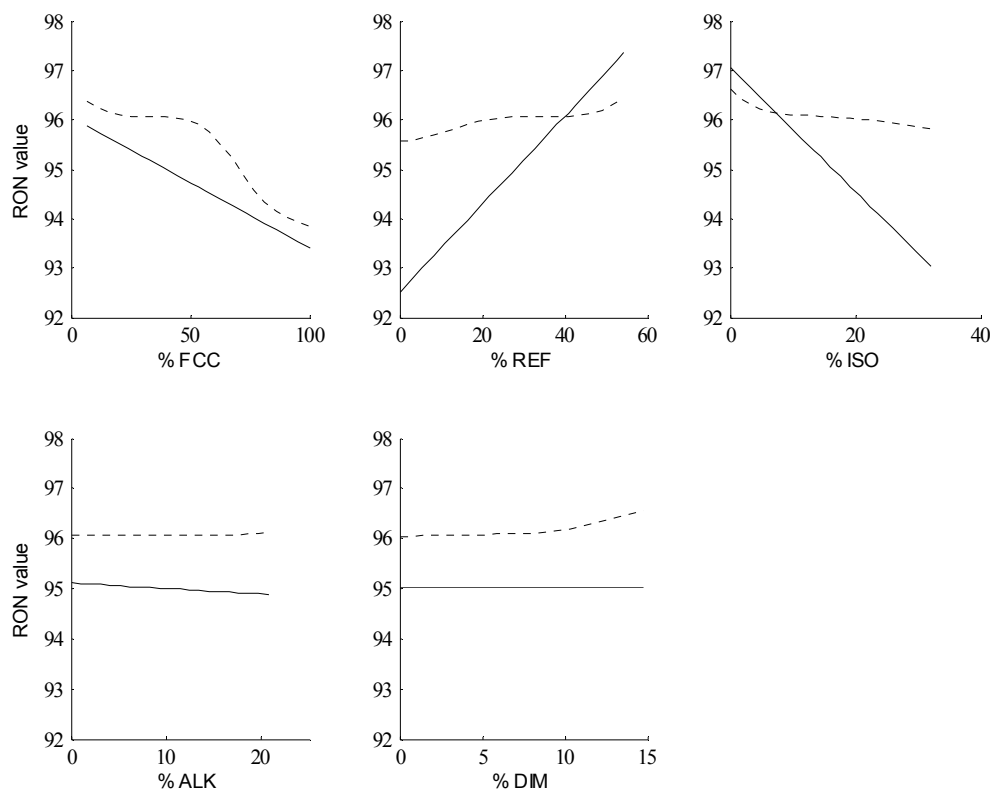
Σχήμα 5-17 Συντελεστές του γραμμικού ως προς τα επιμέρους κλάσματα και τις αλληλεπιδράσεις τους μοντέλου πρόβλεψης του αριθμού RON της βενζίνης που αναπτύχθηκε με εφαρμογή της μεθόδου PCR

Τα κλάσματα FCC και ISO επιδρούν αρνητικά στην διαμόρφωση του τελικού Α.Ο., ενώ αντίθετα όλα τα υπόλοιπα κλάσματα έχουν θετική επίδραση. Θετική επίσης είναι και η αλληλεπίδραση ανάμεσα στο κλάσμα FCC και στα υπόλοιπα κλάσματα της βενζίνης. Το μικρό μέγεθος των συντελεστών αλληλεπίδρασης FCC×MTBE και FCC×C₄ πρέπει να αποδοθεί σε διαφορετικούς μηχανισμούς ανάφλεξης που αναπτύσσονται ανάμεσα σε μίγματα (όπως είναι το FCC) και σε καθαρά συστατικά, όπως είναι το MTBE και τα βουτάνια. Αντίθετα η μίξη του REF, με τα υπόλοιπα κλάσματα έχει αρνητική επίδραση στον Α.Ο., με εξαίρεση την μίξη του με MTBE και C₄. Όμοια η μίξη της βενζίνης ισομερισμού ISO με όλα τα υπόλοιπα κλάσματα εμφανίζεται να έχει αρνητικό πρόσημο. Τέλος, η μίξη της βενζίνης αλκυλίωσης (ALK) με βενζίνη διμερισμού (DIM) και C₄ έχει θετικό αποτέλεσμα στον αριθμό οκτανίων του μίγματος.

Μία διαφορετική προσέγγιση στην μελέτη της διερεύνησης της επίδρασης των επιμέρους κλασμάτων στην διαμόρφωση του Α.Ο. του μίγματος παρουσιάστηκε από τους Πασαδάκη, κ.α. (2006), με εφαρμογή σε “συνθετικά” μίγματα του αλγορίθμου ΤΝΔ πρόβλεψης του RON μίγματος που εκπαιδεύτηκε σε βιομηχανικά μίγματα βενζινών επτά συστατικών (αντίστοιχα με αυτά που χρησιμοποιήθηκαν στην παρούσα εργασία). Με βάση την σύσταση μίας τυπικής βενζίνης (δείγμα αναφοράς) τα “συνθετικά” αυτά μίγματα λήφθηκαν ως σημεία πειραματικού αξονικού σχεδιασμού πάνω στους Cox-άξονες, κάθε ένας από τους οποίους διέρχεται από το σημείο αναφοράς και μία από τις κορυφές του χώρου simplex επτά συστατικών που ορίζει το σύνολο δειγμάτων εκπαίδευσης του ΤΝΔ. Στο Σχ. 5-18 απεικονίζεται ένα σχηματικό παράδειγμα, για ανάμιξη τριών συστατικών, των αξόνων πάνω στους οποίους βρίσκονται τέτοια συνθετικά μίγματα εντός του μη κανονικού χώρου simplex που ορίζουν τα επιτρεπόμενα εύρη συστάσεων των αντίστοιχων κλασμάτων. Στην πραγματικότητα, τα δείγματα που παράγονται με αυτόν τον τρόπο αντιπροσωπεύουν μίγματα της τυπικής βενζίνης και αντίστοιχου εξεταζόμενου κάθε φορά κλάσματος για διάφορες % κ.ο. περιεκτικότητες του τελευταίου.



Σχήμα 5-18 Cox-άξονες σε μη κανονικό simplex (έντονα διαγραμμισμένη περιοχή) τριών συστατικών



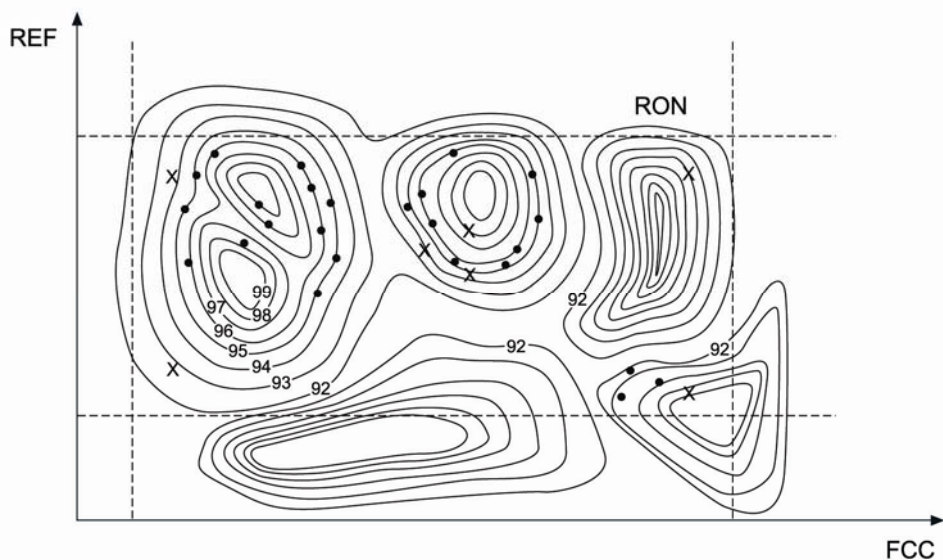
Σχήμα 5-19 Διαγράμματα εξάρτησης του αριθμού RON από τα επιμέρους συστατικά που χρησιμοποιήθηκαν για τα συνθετικά δείγματα (διακεκομμένη: πρόβλεψη ΤΝΔ, συνεχής γραμμή: προσθετικός υπολογισμός)

Οι τιμές RON των "συνθετικών" μιγμάτων υπολογίστηκαν από το εκπαιδευμένο σε βιομηχανικά δείγματα μοντέλο πρόβλεψης ΤΝΔ. Στο Σχ. 5-19 δίδονται τα διαγράμματα των εκτιμώμενων από το μοντέλο τιμών RON ως προς την σύσταση πέντε βασικών κλασμάτων μίξης, ενώ για λόγους σύγκρισης παρουσιάζεται και ο αντίστοιχος προσθετικός υπολογισμός τους. Είναι εμφανής η απόκλιση από την γραμμική συμπεριφορά, ενώ για κάθε κλάσμα η διακεκομμένη καμπύλη αντιπροσωπεύει τον βαθμό αυτό της απόκλισης που εξαρτάται σαφώς από την %κ.ο. περιεκτικότητα του αντίστοιχου κλάσματος. Αν και σε γενικές γραμμές τα αποτελέσματα ως προς την επίδραση του κάθε μεμονωμένου συστατικού είναι σύμφωνα με αυτά που προέκυψαν από την ανάλυση παλινδρόμησης κυρίων συνιστωσών (αρνητική επίδραση των FCC και ISO, θετική επίδραση των υπολοίπων κλασμάτων), η προσέγγιση αυτή αποδίδει καλύτερα την φύση της εξάρτησης του Α.Ο., αφού παρέχει την δυνατότητα διερεύνησης για όλο το εύρος συστάσεων των κλασμάτων και στηρίζεται σε πιο ευέλικτο στην ενσωμάτωση της μη γραμμικότητας του φαινομένου μοντέλο ΤΝΔ. Βέβαια πρέπει να σημειωθεί ότι τα αποτελέσματα αυτά παρήχθησαν με βάση ορισμένη σύσταση βενζίνης και δεν διασφαλίζουν ότι η

συμπεριφορά αυτή των μεμονωμένων κλασμάτων διατηρείται ίδια σε σημαντικά διαφορετικές συγκεντρώσεις των υπολοίπων συστατικών του μίγματος.

5.3 Μοντέλα Πρόβλεψης Αριθμού Οκτανίων από Πειραματικά Δεδομένα

Η ανάπτυξη των μοντέλων πρόβλεψης των Α.Ο. με χρήση ΤΝΔ επέδειξε την δυνατότητα ανάπτυξης αλγορίθμων πρόβλεψης της ιδιαίτερης αυτής ιδιότητας των καυσίμων με ικανοποιητική ακρίβεια και ανέδειξε την ικανότητα των ΤΝΔ να προσεγγίσουν αποτελεσματικά την μη γραμμική της εξάρτηση από τις συγκεντρώσεις των επιμέρους κλασμάτων. Το μικρό πλήθος όμως των διαθέσιμων δεδομένων παραγωγής και μάλιστα για το περιορισμένο χρονικό διάστημα του ενός έτους, καθώς και το γεγονός ότι, για το μεγαλύτερο ποσοστό των δειγμάτων, οι Α.Ο., και ιδιαίτερα ο αριθμός MON, υπήρξαν περιορισμένοι σε εξαιρετικά μικρό εύρος τιμών (95.5-96.5, 98 RON και 85 MON), οδήγησαν στην ανάγκη συλλογής δεδομένων που καλύπτουν ευρύτερα την δοσμένη περιοχή ενδιαφέροντος του χώρου και με μεγαλύτερη διακύμανση των μεταβλητών απόκρισης. Δεδομένου ότι ένα τέτοιο εγχείρημα, υπό την προϋπόθεση χρήσης δεδομένων παραγωγής, απαιτεί την χρονοβόρα και επομένως απαγορευτική συγκρότηση μίας εξαιρετικά μεγάλης βάσης δεδομένων, επιλέχθηκε να διεξαχθεί ο σχεδιασμός της πειραματικής διεξαγωγής ενός μικρού πλήθους εργαστηριακών μετρήσεων, ο δομημένος πίνακας δεδομένων των οποίων περιέχει όσο το δυνατόν μεγαλύτερη πληροφορία.



Σχήμα 5-20 Σχηματική απόδοση της επιφάνειας απόκρισης του αριθμού RON συναρτήσει δύο συστατικών του μίγματος

Τα εργαστηριακά δεδομένα αντανakλούν σε μεγάλο βαθμό την αποκτηθείσα εμπειρία του διυλιστηρίου και περιοριζόμενα απ' αυτή μπορεί να συγκεντρώνονται μόνο σε ορισμένες

περιοχές του χώρου μίξης που εμφανίζουν τις επιθυμητές αποκρίσεις-τιμές των απαιτούμενων Α.Ο. ανά ποιότητα βενζίνης. Στο Σχ. 5–20 απεικονίζεται ένα σχηματικό παράδειγμα της πιθανής επιφάνειας απόκρισης του Α.Ο. μίγματος βενζίνης ως προς δύο συστατικά του. Με κύκλους σημειώνονται τα δείγματα παραγωγής και με X τα εργαστηριακά πειραματικά δείγματα. Αν και η επιφάνεια απόκρισης του Α.Ο. μπορεί να παρουσιάζει διάφορους “λόφους” και “κοιλιάδες”, το μεγαλύτερο ποσοστό των εργοστασιακών δειγμάτων βρίσκονται πάνω σε ισουψείς των 95 και 96 RON και μάλιστα εκείνες που τοποθετούνται πάνω και αριστερά του χώρου που ορίζεται από τα όρια συμμετοχής των συστατικών του μίγματος. Ο περιορισμένος αριθμός δειγμάτων με 98 RON ή 93 RON βρίσκεται σε ακόμα πιο στενές περιοχές και αντιστοιχεί μόνο σε κάποιες από τις δυνατές “συνταγές” μίξης για τις αντίστοιχες ποιότητες βενζίνης. Η ανάπτυξη αλγορίθμων με βάση τέτοια δεδομένα μπορεί να δώσει σταθερά μοντέλα ικανοποιητικής ακρίβειας για τις περιοχές του χώρου που ορίζουν οι αντίστοιχες ισουψείς, οι οποίοι όμως εμφανίζουν αδυναμία πρόβλεψης του Α.Ο. σε μίγματα με σημαντικά διαφοροποιημένες συστάσεις. Η συλλογή δεδομένων από στατιστικά σχεδιασμένα πειράματα, εκτός από το πλεονέκτημα μείωσης των συστηματικών σφαλμάτων που διασφαλίζει η εργαστηριακή παρασκευή των μιγμάτων έναντι της βιομηχανικής παραγωγής, συμβάλλει στην καλύτερη κάλυψη της δοσμένης περιοχής ενδιαφέροντος. Δεδομένα που αντιστοιχούν σε περιφερειακή κάλυψη του πειραματικού χώρου μπορούν να δώσουν πιο γενικά μοντέλα πρόβλεψης, ενώ κατάλληλη σταδιακή πύκνωση των μετρήσεων είναι δυνατό να αποδώσει την πραγματική επιφάνεια απόκρισης του Α.Ο., συμβάλλοντας έτσι στην διερεύνηση της επίδρασης μίξης, αλλά και την βελτιστοποίησης της διαδικασίας ανάμιξης στην λειτουργία παραγωγής του διυλιστηρίου.

Πειραματικός Σχεδιασμός

Με βάση την εμπειρία των προηγούμενων μοντέλων επιλέχθηκαν να παρασκευαστούν μίγματα βενζινών αποκλειστικά των κυριότερων συστατικών χωρίς επιπλέον προσθήκες από άλλες μονάδες. Ο σχεδιασμός του πειράματος της εργαστηριακής παρασκευής μιγμάτων βενζίνης υλοποιήθηκε με χρήση του λογισμικού πακέτου χημειομετρικής πολυπαραμετρικής επεξεργασίας δεδομένων The Unscrambler 7.5 (CAMO ASA), σε δύο στάδια 15 και 25 μετρήσεων, και κάλυπτε τον καθορισμένο πειραματικό χώρο με δείγματα στα όρια του τελευταίου, καθώς και επαναληπτικά δείγματα στα κεντρικά του σημεία. Η τεχνική που χρησιμοποιήθηκε ήταν αυτή του D-optimal πειραματικού σχεδιασμού σε μίγματα με παράγοντες σχεδιασμού τις % κ.ο. συγκεντρώσεις επτά επιμέρους κλασμάτων και απόκριση του συστήματος τον Α.Ο. του μίγματος. Οι μόνοι περιορισμοί (constraints) του σχεδιασμού αφορούσαν τα όρια διακύμανσης των ογκομετρικών συγκεντρώσεων των κλασμάτων μίξης, όπως δίνονται στον Πιν. 5-7., και καθορίστηκαν λαμβάνοντας υπόψη την δυναμικότητα των επιμέρους μονάδων τροφοδοσίας της MOH και το κόστος του τελικού προϊόντος.

Πίνακας 5-7 Περιορισμοί του πειραματικού σχεδιασμού για τα όρια διακύμανσης των ογκομετρικών συγκεντρώσεων των κλασμάτων μίξης

	FCC	REF	ISO	ALK	MTBE	DIM	C₄
min	0	20	0	0	0	0	0
max	80	35	25	18	12	10	6

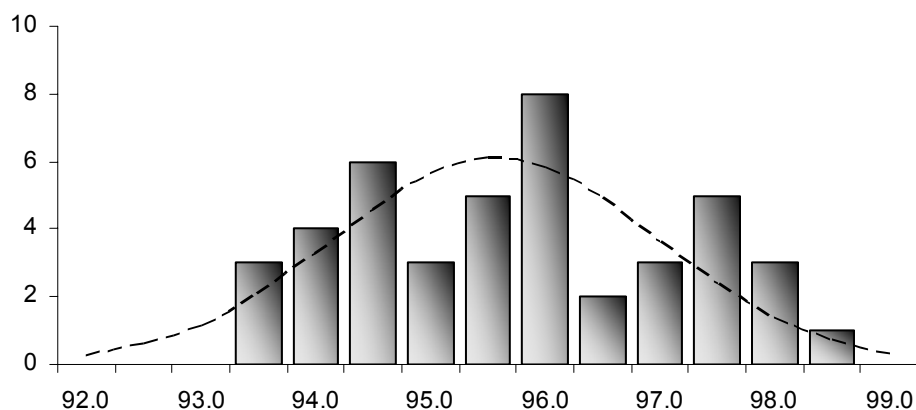
Πειραματικά Δεδομένα

Η διαδικασία παρασκευής των μιγμάτων πραγματοποιήθηκε στο εργαστήριο της ΜΟΗ, όπου και μετρήθηκαν οι Α.Ο. των επιμέρους κλασμάτων και των τελικών μιγμάτων. Η χρήση του κλάσματος βουτανίων δεν ήταν δυνατή, λόγω της πτητικότητας του σε ανοιχτό σύστημα, το ποσοστό συμμετοχής του οποίου αποδόθηκε στα κλάσματα αναμόρφωσης και ισομερισμού. Εκτός των σχεδιασμένων μιγμάτων παρασκευάστηκαν και κάποια επιπλέον δείγματα αναφοράς με σύσταση αντίστοιχη των τυπικών βενζινών παραγωγής.

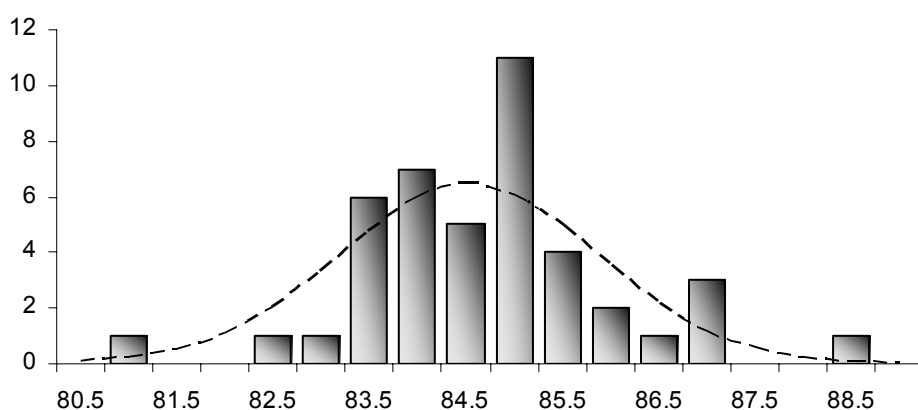
Συγκεντρώθηκαν τελικά 43 δείγματα εργαστηριακά παρασκευασμένων βενζινών, ο πίνακας τιμών των οποίων παρατίθεται στο Παράρτημα. Τα όρια διακύμανσης των Α.Ο. (RON, MON) των μιγμάτων και των % κ.ο. συγκεντρώσεων των επιμέρους κλασμάτων μίξης παρουσιάζονται στον Πίν. 5-8. Ο χώρος του μη κανονικού simplex που ορίζουν (αντίστοιχος αυτού του Σχ. 5-18) καλύπτει την περιοχή με τις πλέον συνηθισμένες συστάσεις της βενζίνης δεδομένου ότι τα όρια των δεδομένων παραγωγής τον επικαλύπτουν (Πίν. 5-2 και Πίν. 5-5), όμως οι μετρούμενοι Α.Ο. τους εμφανίζουν καλύτερη ισοκατανομή κατανομή στο εύρος τιμών τους, όπως φαίνεται στα Σχ. 5-21 και Σχ. 5-22 σε σύγκριση με τα Σχ. 5-1 και Σχ. 5-2.

Πίνακας 5-8 Όρια διακύμανσης των Α.Ο. (RON, MON) και των % κ.ο.συγκεντρώσεων κλασμάτων για τα εργαστηριακά μίγματα βενζινών

	FCC	REF	ISO	ALK	MTBE	DIM	RON	MON
min	16.00	7.40	0.00	0.00	0.00	0.00	84.8	80.6
max	80.00	41.00	27.70	18.00	10.00	10.00	98.2	88.5

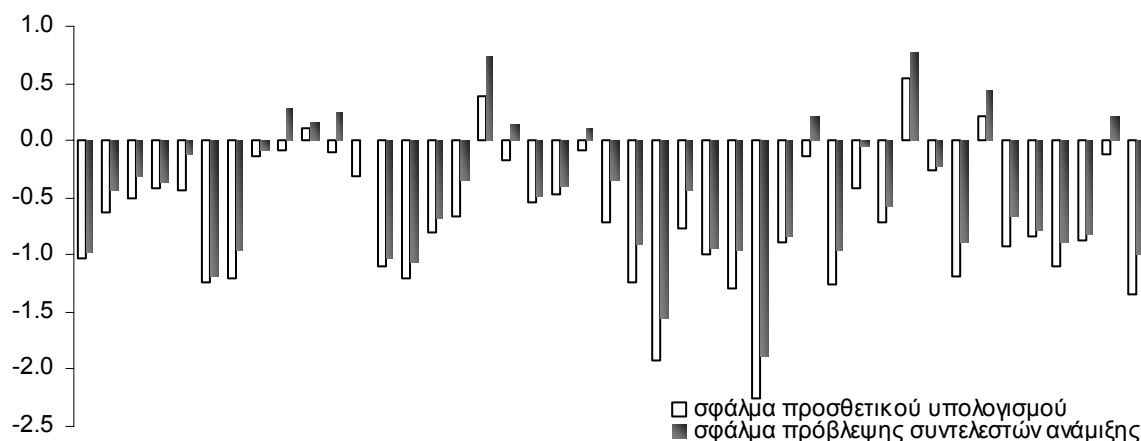


Σχήμα 5-21 Ιστόγραμμα του αριθμού RON του εργαστηριακών μιγμάτων (43 δείγματα)

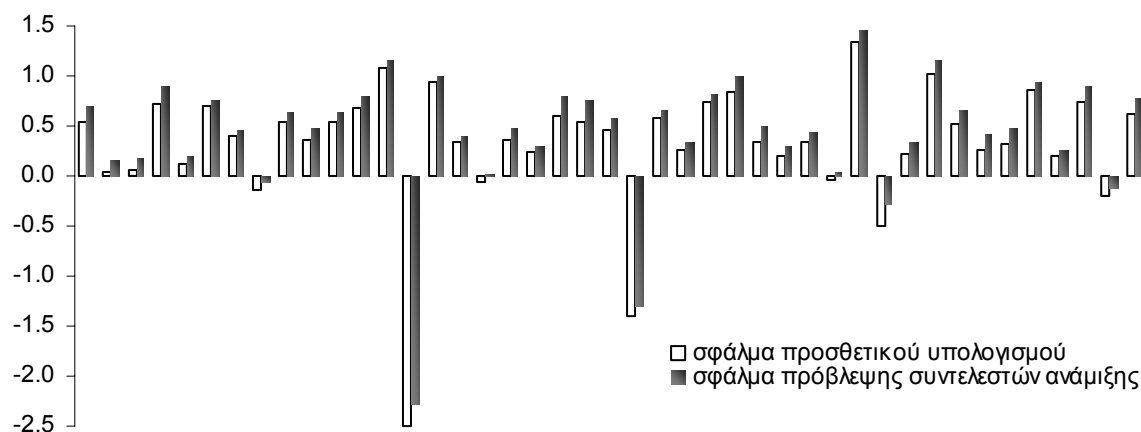


Σχήμα 5-22 Ιστόγραμμα του αριθμού MON του εργαστηριακών μιγμάτων (43 δείγματα)

Αξίζει να σημειωθεί ότι και στα εργαστηριακά παρασκευασμένα μίγματα βενζινών ο προσθετικός υπολογισμός των Α.Ο., όπως και η πρόβλεψη του μοντέλου συντελεστών ανάμιξης της ΜΟΗ εμφανίζουν την τάση “να σφάλουν κατά τον ίδιο τρόπο”, υποεκτιμώντας τον αριθμό RON και υπερεκτιμώντας τον MON του τελικού μίγματος, παρουσιάζοντας αντίστοιχες τιμές σφαλμάτων πρόβλεψης (Σχ. 5-23, Σχ. 5-24).



Σχήμα 5-23 Σφάλμα υπολογισμού του RON προσθετικά και με χρήση εξισώσεων συντελεστών ανάμιξης για τα εργαστηριακά παρασκευασμένα μίγματα βενζινών



Σχήμα 5-24 Σφάλμα υπολογισμού του MON προσθετικά και με χρήση εξισώσεων συντελεστών ανάμιξης για τα εργαστηριακά παρασκευασμένα μίγματα βενζινών

Ανάπτυξη Μοντέλων Πρόβλεψης Α.Ο. Εργαστηριακών Μιγμάτων

Για την ανάπτυξη των μοντέλων με βάση τα εργαστηριακά παρασκευασμένα μίγματα βενζινών ακολουθήθηκε η ίδια μεθοδολογία των προηγούμενων μοντέλων. Δημιουργήθηκαν με τυχαία ανακατανομή των δειγμάτων τα σύνολα εκπαίδευσης και ελέγχου με πλήθος δειγμάτων 38 και 5 αντίστοιχα. Επιλέχθηκε, και για τις δύο περιπτώσεις (RON και MON), ως καταλληλότερη αρχιτεκτονική του ΤΝΔ αυτή του ενός ενδιάμεσου επιπέδου τριών κόμβων ($6 \times 3 \times 1$), με συνάρτηση μεταφοράς στο ενδιάμεσο και το επίπεδο εξόδου την υπερβολική εφάπτομενική σιγμοειδή. Διαπιστώθηκε ότι η χρήση ενός μόνο ενδιάμεσου επιπέδου ήταν επαρκής, χωρίς να επιβαρύνεται η ταχύτητα εκπαίδευσης του ΤΝΔ, λόγω του μικρού πλήθους δεδομένων. Επιπλέον έχει σαν αποτέλεσμα ένα μικρότερο διάνυσμα παραμέτρων (τιμές βαρών) καθιστώντας πιο εύχρηστο τον αλγόριθμο πρόβλεψης. Η μέθοδος εκπαίδευσης διατηρήθηκε η ίδια με των προηγούμενων μοντέλων.

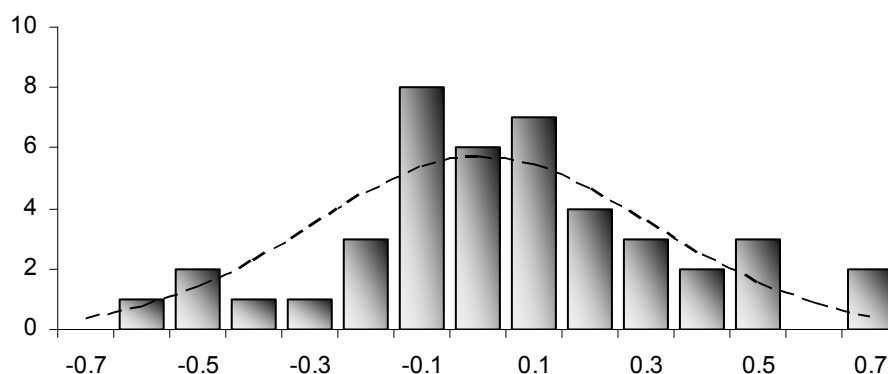
Όλα τα υποψήφια μοντέλα ΤΝΔ που εκπαιδεύτηκαν είχαν ως εισόδους τους Α.Ο. των επιμέρους κλασμάτων σταθμισμένους με τις % κ.ο. συγκεντρώσεις τους στο μίγμα, στην λογική διατήρησης της σταθερότητας του μοντέλου στις αναμενόμενες μεταβολές της ποιότητας της τροφοδοσίας στο χρόνο, και απόκριση τον αντίστοιχο Α.Ο. του μίγματος. Τα αποτελέσματα των καλύτερης απόδοσης μοντέλων παρουσιάζονται αναλυτικά στην συνέχεια.

Μοντέλο πρόβλεψης RON

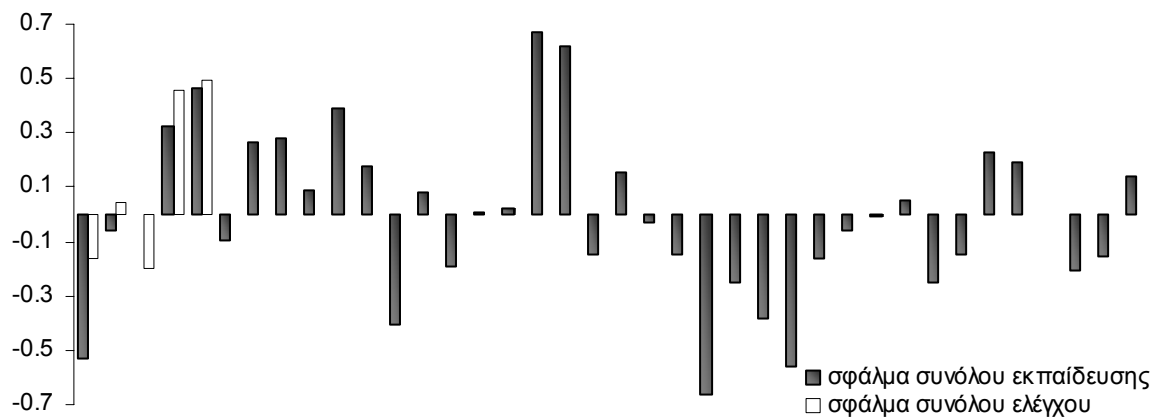
Στον Πιν. 5-9 δίνονται τα στατιστικά αποτελέσματα του μοντέλου πρόβλεψης του RON που παρουσίασε την καλύτερη απόδοση, ενώ στα διαγράμματα των Σχ. 5-25 και Σχ. 5-26 παρουσιάζονται η κατανομή των σφαλμάτων για το σύνολο των δειγμάτων και οι τιμές τους για τα δείγματα των συνόλων εκπαίδευσης και ελέγχου. Το μοντέλο εμφανίζει μικρότερη ακρίβεια σε σχέση με το αντίστοιχο που εκπαιδεύτηκε με δεδομένα παραγωγής, κυρίως λόγω του μικρού πλήθους δεδομένων, αλλά εξακολουθεί να είναι ικανοποιητική στα όρια της επαναληψιμότητας μέτρησης της πρότυπης μεθόδου ASTM (0.2-0.3 RON).

Πίνακας 5-9 Σφάλματα του μοντέλου πρόβλεψης RON για τα εργαστηριακά παρασκευασμένα δείγματα

Σύνολο δειγμάτων	Αριθμός δειγμάτων	MAEP	SEP	RMSEP	minAEP	maxAEP
εκπαίδευσης	38	0.2	0.3	0.1	0.0	0.7
ελέγχου	5	0.3	0.3	0.1	0.0	0.5



Σχήμα 5-25 Κατανομή των σφαλμάτων του μοντέλου πρόβλεψης του RON για το σύνολο των εργαστηριακά παρασκευασμένων δειγμάτων.



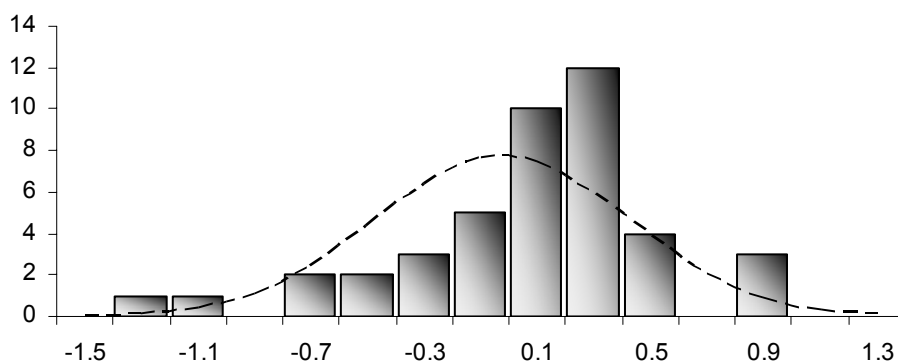
Σχήμα 5-26 Διάγραμμα σφαλμάτων του μοντέλου πρόβλεψης του RON για τα δείγματα του συνόλου εκπαίδευσης και ελέγχου.

Μοντέλο πρόβλεψης MON

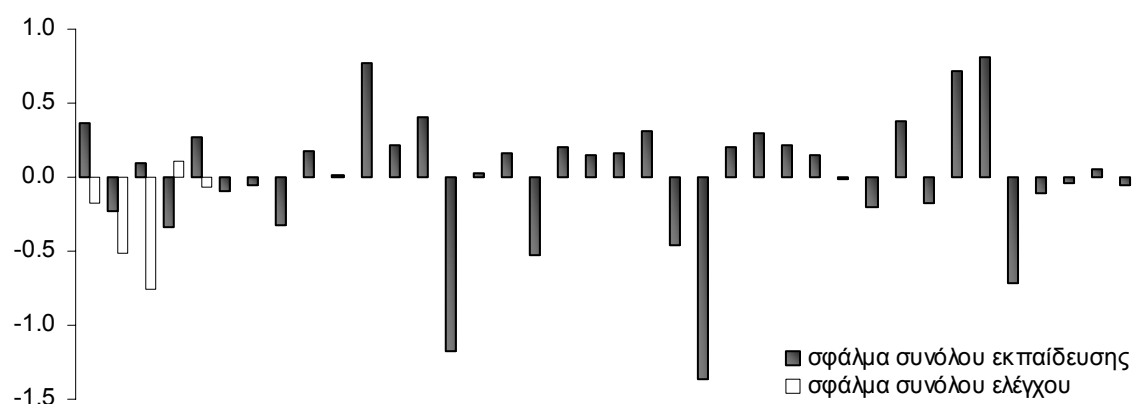
Στον Πιν. 5-10 δίνονται τα στατιστικά αποτελέσματα του μοντέλου πρόβλεψης του MON που παρουσίασε την καλύτερη απόδοση, ενώ στα διαγράμματα των Σχ. 5-27 και Σχ. 5-28 παρουσιάζονται η κατανομή των σφαλμάτων για το σύνολο των δειγμάτων και οι τιμές τους για τα δείγματα των συνόλων εκπαίδευσης και ελέγχου. Το μοντέλο δεν εμφανίζει καλή ακρίβεια σε σχέση με εκείνο που εκπαιδεύτηκε με δεδομένα παραγωγής και η κατανομή των σφαλμάτων περιέχει μεροληψία (bias) προς θετικές τιμές, αλλά όπως και στην περίπτωση του RON, το μικρό πλήθος των μετρήσεων δεν επιτρέπει περαιτέρω βελτίωση των βαρών του ΤΝΔ χωρίς τον κίνδυνο υπερεκπαίδευσης και απώλειας της ικανότητας γενίκευσης του αλγορίθμου. Επιπλέον, θα πρέπει να σημειωθεί ότι το παρόν μοντέλο σε σύγκριση με το αρχικό των δεδομένων παραγωγής δεν είναι “στραγγαλισμένο” σε μία μεμονωμένη τιμή (85 MON) και ένας εμπλουτισμός των δεδομένων και επανεκπαίδευσή του μπορεί να οδηγήσει στην ανάπτυξη ενός σταθερότερου και υψηλότερη διακριτικής ικανότητας ως προς τα δείγματα εισαγωγής μοντέλου πρόβλεψης του MON, με την απαιτούμενη ακρίβεια πρόβλεψης.

Πίνακας 5-10 Σφάλματα του μοντέλου πρόβλεψης MON για τα εργαστηριακά παρασκευασμένα δείγματα

Σύνολο δειγμάτων	Αριθμός δειγμάτων	MAEP	SEP	RMSEP	minAEP	maxAEP
εκπαίδευσης	38	0.3	0.4	0.2	0.0	1.4
ελέγχου	5	0.3	0.3	0.2	0.1	0.8



Σχήμα 5-27 Κατανομή των σφαλμάτων του μοντέλου πρόβλεψης του MON για το σύνολο των εργαστηριακά παρασκευασμένων δειγμάτων.



Σχήμα 5-28 Διάγραμμα σφαλμάτων του μοντέλου πρόβλεψης του MON για τα δείγματα του συνόλου εκπαίδευσης και ελέγχου.

Συμπεράσματα- Προτάσεις

Η χημειομετρική ανάλυση των δεδομένων παραγωγής επιβεβαίωσε την μη προσθετικότητα του Α.Ο., δεδομένου ότι τα κλάσματα μίξης συμμετέχουν στην ανάμιξη όχι με τον απόλυτο Α.Ο. τους, αλλά διαφοροποιημένο ανάλογα με την σύσταση και το είδος τους, καθώς επίσης και ανάλογα με την σύσταση και το είδος των υπολοίπων συστατικών που περιέχονται στο μίγμα βενζίνης. Παρουσιάστηκαν επίσης τα πλεονεκτήματα των ΤΝΔ, τόσο από άποψη ανάδειξης των υφιστάμενων γραμμικών εξαρτήσεων μεταξύ των παραμέτρων του φαινομένου μίξης, όσο και από άποψη ευελιξίας στην προσαρμογή νέων δεδομένων, έναντι άλλων αλγορίθμων πρόβλεψης.

Οι αλγόριθμοι ΤΝΔ που αναπτύχθηκαν προβλέπουν αξιόπιστα –προσεγγίζοντας την απαιτούμενη ακρίβεια υπολογισμού του –τον αριθμό οκτανίων της τελικής βενζίνης, χρησιμοποιώντας ως δεδομένα εισαγωγής μόνο την περιεκτικότητα και τον Α.Ο. των επιμέρους κλασμάτων και μπορούν να αποτελέσουν ένα σημαντικό και εξαιρετικά ταχύ εργαλείο στον προσδιορισμό των αναλογιών μίξης για την εξασφάλιση του απαιτούμενου Α.Ο. του τελικού προϊόντος στην μονάδα ανάμιξης ενός διυλιστηρίου, ενώ μπορούν να χρησιμοποιηθούν και ως εργαλείο επιλογής της βέλτιστης σύστασης της βενζίνης, εισάγοντας επιπλέον κριτήρια κόστους και διαθεσιμότητας των κλασμάτων.

Στην κατεύθυνση διερεύνησης της επίδρασης στον Α.Ο. της τελικής βενζίνης της συγκέντρωσης και του Α.Ο. των επιμέρους κλασμάτων μίξης έγινε προσπάθεια να προσδιοριστεί ο βαθμός επίδρασης κάθε μεμονωμένου κλάσματος και των αλληλεπιδράσεων τους ανά δύο. Τα αποτελέσματα συμφωνούν τόσο με την υφιστάμενη εμπειρική γνώση όσο και με άλλες μελέτες διαφορετικής προσέγγισης από αυτή του προσδιορισμού των συντελεστών–αποτελεσμάτων παραγόντων ενός γραμμικού μοντέλου που ακολουθήθηκε. Η εξέταση των δεδομένων παραγωγής βενζίνης της ΜΟΗ, καθώς και εκείνων των εργαστηριακών μιγμάτων που παρασκευάστηκαν κατά την παρούσα εργασία, επιβεβαιώνει γενικά τα συμπεράσματα στα οποία οδηγηθήκαμε, όμως η περαιτέρω ανάλυση και εξήγηση της επίδρασης μίξης των κλασμάτων βενζινών πρέπει να συνεχιστεί στηριζόμενη σε μετρήσεις Α.Ο. δυαδικών μιγμάτων.

Τέλος, οι αλγόριθμοι ΤΝΔ που αναπτύχθηκαν με βάση δεδομένα από εργαστηριακά παρασκευασμένα μίγματα κλασμάτων βενζινών, εμφανίζουν επίσης ικανοποιητική ακρίβεια πρόβλεψης παρά το μικρό μέγεθος του συνόλου εκπαίδευσης. Τα αποτελέσματα των τελευταίων αυτών μοντέλων επέδειξαν την δυνατότητα, με διεξαγωγή εκτενέστερου στατιστικά σχεδιασμένου πειράματος, του καθορισμού της συνολικής επιφάνειας απόκρισης του Α.Ο. συμβάλλοντας έτσι στην διερεύνηση της επίδρασης μίξης, αλλά και την βελτιστοποίησης της διαδικασίας ανάμιξης στην λειτουργία παραγωγής του διυλιστηρίου.

Βιβλιογραφία

Μερτίκας Σ. Π. (1999), *Τηλεπισκόπηση και Ψηφιακή Ανάλυση Εικόνας*, ΙΩΝ, Αθήνα

MOTOR OIL (ΕΛΛΑΣ) ΔΙΥΛΙΣΤΗΡΙΑ ΚΟΡΙΝΘΟΥ Α.Ε. (2005), *Εκπαιδευτικά Εγχειρίδια τ.2: Παραγωγή Καυσίμων*, Motor Oil, Κόρινθος

MOTOR OIL (ΕΛΛΑΣ) ΔΙΥΛΙΣΤΗΡΙΑ ΚΟΡΙΝΘΟΥ Α.Ε. (2005), *Εκπαιδευτικά Εγχειρίδια τ.3: Προϊόντα*, Motor Oil, Κόρινθος

Πασαδάκης Ν. (2003): *Εισαγωγή στην χημειομετρία* (Σημειώσεις Ακαδημαϊκών Διαλέξεων), Πολυτεχνείο Κρήτης, Χανιά

Ρακόπουλος Κ. (1987), *Αρχές Εμβολοφόρων Μηχανών Εσωτερικής Κάυσης*, Φούντας, Αθήνα

Σουρλίγκας Σ. (2004), *Πρόβλεψη Ιδιοτήτων του Ντήζελ με Χημειομετρική Επεξεργασία Υπέρυθρων Φασματοσκοπικών Δεδομένων* (Διπλωματική Εργασία), Πολυτεχνείο Κρήτης, Χανιά

Σταθερόπουλος Μ., Παππά Α. (1998), *Χημειομετρία*, ΕΜΠ, Τμήμα Χημικών Μηχανικών, Αθήνα

Στούρνας Σ., Λόης Ε., Ζαννίκος Φ. (1997), *Τεχνολογία Καυσίμων και Λιπαντικών*, ΕΜΠ, Τμήμα Χημικών Μηχανικών, Αθήνα

Τριάντος Α. Χ. (1999), *Πρόβλεψη του Αριθμού Οκτανίου σε Δείγματα Βενζινών με Νευρωνικά Δίκτυα και Φασματοσκοπία Υπέρυθρου* (Διπλωματική Εργασία), Πολυτεχνείο Κρήτης, Χανιά

Edelen D. G. B., Kydoniefs A. D. (1985), *Εφαρμοσμένη Γραμμική Άλγεβρα*, University Studio Press, Θεσσαλονίκη

Esbensen K. H., (2001) *Multivariate data analysis in practice*, Camo Process AS

Ferguson R.C.(1986), *Internal Combustion Engines, applied thermosciences*, Wiley, New York

Funahashi, K. (1989), *On the approximate realization of continuous mapping by neural networks*, Neural Networks v.2, p. 183-192

Gary J.H., Handwerl G.E. (1994), *Petroleum Refining Technology and Economics*, 3rd ed., Marcel Dekker, New York

Geladi P., Kowalski B (1986): *Partial least-squares regression: A tutorial*, Laboratory for chemometrics and center for process analytical chemistry, University of Washington, Analytica Chimica Acta V.185, p.1-17

Gurney K. (1997), *An introduction to neural networks*, University of Sheffield, UCL Press, UK

Gybenko G., (1989), Approximation by Superposition of a Sigmoidal Function, Mathematics of Control, Signal and Systems, v.2, No 4, p. 303-314

Healy W.C., Maasen C.W., Peterson J. (1959), *A new approach to blending octanes*, 24th Midyear Meeting of the American Petroleum Institute's Division of Refining, New York, p. 39.

Hornik K., Stinchcombe M., White H. (1989), *Multilayer Feedforward Neural Networks*, Transfer Performance by Neural Networks, Ind. Eng. Chem. Res. v.41, p.3436-3446

MATLAB: «Neural Network Tutorial », Matlab Help Documentation, Version 6.5 R13

Mike J. Adams (1995), *Chemometrics in Analytical Spectroscopy*, The Royal Society of Chemistry, Cambridge, UK

Modgomery D.C. (1991), *Design and Analysis of Experiments*, John Wiley and Sons, New York

Murray R. Spiegel, μετάφραση: Σωτήριος Κ. Περισίδης (1975), *Πιθανότητες και Στατιστική*, McGraw-Hill, New York, ΕΣΠΙ, Αθήνα

Murty B.S.N., Rao R.N. (2004), *Global optimization for prediction of blend composition of gasolines of desired octane number and properties*, Fuel Processing Technology V.85, p. 1595-1602.

Pasadakis N., Gaganis V., Foteinopoulos C., (2006), *Octane number prediction for gasoline blends*, Fuel Processing Technology V.85, p.505-509

Rusin M.H., Chung H.S., Marshall J.F. (1981), *A 'transformation' method for calculating the research and motor octane numbers of gasoline blends*, Industrial and Engineering Chemistry Fundamentals V.20, p. 195-204.

Smits J.R.M., Melssen W.J., Buydens L.M.C., Kateman G. (1994), Using artificial neural networks for solving chemical problems. Part I. Multi-layer feedforward networks, Chemometrics and Intelligent Laboratory Systems V. 22, p. 165-189.

Stewart W.E. (1959), Predict octanes for gasoline blends, Petroleum Refiner V. 38, p. 135-139.

Wold S., Esbensen K., Geladi P. (1987), *Principal component Analysis*, Chemometrics and Intelligent Laboratory Systems V.2, p. 37-52

Zahed A.H., Mullah S.A., Bashir M.D. (1993), Predict octane number for gasoline blends, Hydrocatbon Processing V.5, p.85-87.

<http://dasher.wustl.edu/bio5476/reading/chemometrics-aug-02.pdf>

<http://dspace.library.drexel.edu/handle/1860/557>

<http://dspace.library.drexel.edu/retrieve/1860/384>

http://media.wiley.com/product_data/excerpt/28/04712024/0471202428.pdf

<http://minf.uub.ac.be/fabi/calibration/multi/nn/download/down.html>

<http://plato.stanford.edu/entries/>

<http://rcw.raaiuniversity.edu/biotechnology/MScBioinformatics/data-mining-tech/lecture-notes/lecture-09-2.pdf>

<http://www.faqs.org/faqs/autos/gasoline-faq/part1-4>

<http://www.itl.nist.gov/div898/handbook/index.htm>

<http://www.mind.ilstu.edu/curriculum2/perception/mpneuron1.html>

<http://www.mli.gmu.edu/teaching/fall-2003/KDD2003/Papers/S04p07.pdf>

http://www.statistik.uni-dortmund.de/download/mitarbeiter/enache/VersuchsplanungGraduateSchool/DOE_Lecture_English_1-4.pdf

Παράρτημα

Π.1 Στατιστικά μεγέθη σφαλμάτων

Παρακάτω παρατίθενται εν συντομία οι συνήθεις εκτιμήσεις σφαλμάτων που χρησιμοποιούνται για την αξιολόγηση μοντέλων πρόβλεψης.

Σφάλμα Πρόβλεψης (Error of Prediction, EP)

Το σφάλμα πρόβλεψης ε_i για το i -οστό δείγμα του πληθυσμού δίνεται από την διαφορά της προβλεπόμενης από το μοντέλο τιμής $\hat{y}_i$ της ιδιότητας από την πραγματική τιμή y_i :

$$\varepsilon_i = y_i - \hat{y}_i \quad (\Pi.1)$$

Απόλυτο Σφάλμα Πρόβλεψης (Absolute Error of Prediction, AEP)

Η απόκλιση από την πραγματική τιμή y_i ή απόλυτο σφάλμα πρόβλεψης $|\varepsilon_i|$ για το i -οστό δείγμα του πληθυσμού είναι η απόλυτη τιμή του αντίστοιχου σφάλματος πρόβλεψης:

$$|\varepsilon_i| = |y_i - \hat{y}_i| \quad (\Pi.2)$$

Μέσο σφάλμα πρόβλεψης (Mean Error of Prediction, MEP)

Το μέσο σφάλμα της πρόβλεψης είναι η μέση τιμή του σφάλματος πρόβλεψης και προκύπτει από το λόγο του αθροίσματος των σφαλμάτων της πρόβλεψης ε_i προς το πλήθος των προβλέψεων n . Αναφέρεται και ως *μεροληψία* ή *συστηματικό σφάλμα* (bias) του μοντέλου πρόβλεψης και δίνεται από τον τύπο:

$$MEP = \frac{\sum_{i=1}^n \varepsilon_i}{n} = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)}{n} \quad (\Pi.3)$$

Μέσο απόλυτο σφάλμα πρόβλεψης (Mean absolute Error of Prediction, MAEP)

Το μέσο απόλυτο σφάλμα πρόβλεψης $|m|$ είναι το άθροισμα όλων των απόλυτων σφαλμάτων της πρόβλεψης $|\varepsilon_i|$ προς τον αριθμό των τιμών n που προβλέφθηκαν:

$$MAEP = \frac{\sum_{i=1}^n |\varepsilon_i|}{n} = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n} \quad (\Pi.4)$$

Τυπικό σφάλμα πρόβλεψης (Standard Error of Prediction, SEP)

Το τυπικό σφάλμα της πρόβλεψης είναι η τυπική απόκλιση του σφάλματος της πρόβλεψης και δίνεται από τον τύπο:

$$SEP = \sqrt{\frac{\sum_{i=1}^n (\varepsilon_i - MEP)^2}{n-1}} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i - MEP)^2}{n-1}} \quad (\Pi.5)$$

Μέσο σφάλμα της τετραγωνικής ρίζας της πρόβλεψης (Root-Mean Squared Error of Prediction, RMSEP)

Το μέσο σφάλμα της τετραγωνικής ρίζας της πρόβλεψης RMSEP είναι η τετραγωνική ρίζα του μέσου όρου των τετραγώνων των σφαλμάτων της πρόβλεψης και δίνεται από τον τύπο:

$$RMSEP = \sqrt{\frac{\sum_{i=1}^n \varepsilon_i^2}{n-1}} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (\Pi.6)$$

Μέγιστο απόλυτο σφάλμα πρόβλεψης (Maximum Absolute Error of Prediction, MaxAEP)

Το μέγιστο απόλυτο σφάλμα της πρόβλεψης $|\varepsilon|_{\max}$ δίνεται από τον τύπο:

$$|\varepsilon|_{\max} = \max\{|\varepsilon_i|\} \quad (\Pi.7)$$

Ελάχιστο απόλυτο σφάλμα πρόβλεψης (Minimum Absolute Error of Prediction, MinAEP)

Το ελάχιστο απόλυτο σφάλμα της πρόβλεψης $|\varepsilon|_{\min}$ δίνεται από τον τύπο:

$$|\varepsilon|_{\min} = \min\{|\varepsilon_i|\} \quad (\Pi.8)$$

Π.2 Μοντέλο Συντελεστών Ανάμιξης

Η ανάμιξη στην μονάδα της ΜΟΗ γίνεται με χρήση του μη γραμμικού μοντέλου των λεγόμενων συντελεστών ανάμιξης. Οι εξισώσεις που χρησιμοποιεί το μοντέλο αυτό για τον προσδιορισμό του Α.Ο. της βενζίνης με εισόδους τους Α.Ο. των επιμέρους κλασμάτων μίξης και τις ογκομετρικές συγκεντρώσεις τους στο τελικό μίγμα είναι οι εξής:

Αρχικά υπολογίζεται ο συντελεστής ανάμιξης (F_i) του κάθε κλάσματος συναρτήσει του αντίστοιχου Α.Ο. του από τις σχέσεις:

Για Α.Ο. (RON_i ή MON_i) < 85

$$F_i = 7.629395 \times 10^{-7} RON_i^4 - 1.521485 \times 10^{-4} RON_i^3 + 1.093527 \times 10^{-2} RON_i^2 + 0.3385549 RON_i + 19.79466$$

Για Α.Ο. (RON_i ή MON_i) > 85

$$F_i = -1.74719 \times 10^{-6} RON_i^4 + 2.837436 \times 10^{-4} RON_i^3 + 1.934477 \times 10^{-2} RON_i^2 - 4.449638 RON_i + 229.1377$$

Ο συντελεστής ανάμιξης (F_{HB}) του μίγματος υπολογίζεται ως άθροισμα των συντελεστών ανάμιξης (F_i) των επιμέρους κλασμάτων, ζυγισμένους με τις αντίστοιχες % κ.ο. συγκεντρώσεις τους.

Ο Α.Ο. του τελικού μίγματος προκύπτει από τις σχέσεις:

Αν $F_{HB} < 73.95$

$$RON_{HB} \text{ ή } MON_{HB} = -6.731819 \times 10^{-7} F_{HB}^4 - 2.422371 \times 10^{-5} F_{HB}^3 + 1.627614 \times 10^{-2} F_{HB}^2 + 0.3801322 F_{HB} - 2.192894$$

Αν $F_{HB} < 73.95$

$$RON_{HB} \text{ ή } MON_{HB} = 4.248186 \times 10^{-6} F_{HB}^4 - 7.606696 \times 10^{-4} F_{HB}^3 + 6.328804 \times 10^{-3} F_{HB}^2 + 6.028001 F_{HB} - 214.7961$$

Π.3 Πίνακες Βαρών των Μοντέλων και Κώδικας Υλοποίησης ΤΝΔ

Πίνακας 1 Βάρη και πολώσεις του μοντέλου πρόβλεψης RON με εισόδους τις % κ.ο. περιεκτικότητες των οκτώ επιμέρους κλασμάτων μίξης

1ο επίπεδο								
-0.045	0.7517	1.3778	0.5955	-0.619	0.029	-0.963	0.8068	1.2672
0.7388	0.5086	-0.948	-0.107	1.0259	0.286	0.5583	-0.045	-0.662
0.4881	0.0957	-0.378	0.4808	-0.612	0.6518	0.3489	1.1236	-0.453
0.5613	-0.267	-0.379	0.5901	0.2988	0.8144	-0.938	-0.676	0.3669
0.6972	-0.807	-0.629	-0.343	0.3543	1.0584	-0.529	-0.367	1.1721
0.6768	-0.216	-0.666	-0.416	0.3902	0.8141	-0.433	-0.777	2.0239
2ο επίπεδο								
0.1777	-0.879	0.3315	0.0259	1.2453	-0.944			-1.769
0.741	0.6156	-0.555	-0.094	-0.8	1.3459			-0.783
-0.444	1.0651	0.3113	1.1939	0.6622	0.3497			0.0987
0.2953	0.8614	0.6908	-0.018	0.6881	-0.863			1.0546
-0.268	-1.033	0.0253	-0.577	0.9808	-0.785			-1.962
3ο επίπεδο								
-0.545	0.7889	0.1263	0.6874	0.7644				0.1039

Πίνακας 2 Βάρη και πολώσεις του μοντέλου πρόβλεψης RON με εισόδους τους αριθμούς RON των οκτώ επιμέρους κλασμάτων μίξης ζυγισμένους με % κ.ο. περιεκτικότητές τους

1ο επίπεδο								
0.649	-1.013	-0.468	0.387	-0.390	-0.260	0.308	-0.559	0.649
1.276	0.908	0.078	-0.218	0.425	-0.439	0.416	-0.118	1.276
0.810	0.452	0.807	-0.527	0.638	0.907	-0.359	0.594	0.810
0.491	-0.749	0.062	-0.721	0.333	-0.356	0.653	0.829	0.491
0.802	0.157	-0.772	0.500	1.063	2.207	0.359	0.851	0.802
-0.044	-0.780	-0.771	-0.691	-0.007	-0.445	0.580	-1.044	-0.044
2ο επίπεδο								
0.778	-0.229	-1.145	0.155	1.408	-0.165			0.778
-0.355	-1.164	-0.747	0.903	-1.077	0.222			-0.355
-1.500	0.267	-0.871	0.319	0.413	-0.898			-1.500
0.006	-0.640	1.145	-0.233	-0.628	0.420			0.006
-0.471	-0.547	1.300	0.740	0.473	-0.163			-0.471
3ο επίπεδο								
-0.551	-0.701	0.430	-0.329	0.521				-0.551

Πίνακας 3 Βάρη και πολώσεις του μοντέλου πρόβλεψης MON με εισόδους τους αριθμούς MON των οκτώ επιμέρους κλασμάτων μίξης ζυγισμένους με % κ.ο. περιεκτικότητες τους

1ο επίπεδο								
0.864	0.155	0.234	-0.031	-0.933	0.205	-0.796	-0.306	0.864
-1.078	0.199	0.998	0.941	-0.061	-0.327	0.603	0.490	-1.078
-0.023	-0.649	0.884	0.653	-0.007	-0.383	-0.219	0.746	-0.023
-0.395	0.002	1.288	-0.627	-1.652	-1.291	0.469	-0.482	-0.395
0.802	-0.168	-0.519	0.720	-0.753	-0.788	-0.082	0.362	0.802
2ο επίπεδο								
0.894	0.828	-1.023	-1.961	-0.509				0.894
-1.291	-0.256	-1.192	0.266	1.010				-1.291
0.550	0.514	-0.849	-1.043	0.916				0.550
-0.290	-0.870	0.766	0.976	0.145				-0.290
3ο επίπεδο								
0.842	0.404	0.004	0.330					0.842

Πίνακας 4 Βάρη και πολώσεις του μοντέλου πρόβλεψης RON που αναπτύχθηκε με βάση τα εργαστηριακά παρασκευασμένα δείγματα

1ο επίπεδο						
0.245	-0.476	0.390	-0.212	-0.371	-0.192	-0.317
0.034	0.104	0.999	-0.434	-0.447	-0.386	0.011
-1.100	-0.188	2.657	-0.887	-0.987	-1.790	-0.354
2ο επίπεδο						
1.788	-2.210	1.544				0.466
3ο επίπεδο						
-1.014						0.011

Πίνακας 5 Βάρη και πολώσεις του μοντέλου πρόβλεψης MON που αναπτύχθηκε με βάση τα εργαστηριακά παρασκευασμένα δείγματα

1ο επίπεδο						
-2.036	1.152	1.324	0.004	1.432	-0.352	1.305
-0.515	-1.524	4.287	-2.936	-1.012	0.271	0.065
-0.167	0.288	0.772	0.489	0.222	-0.133	-1.908
2ο επίπεδο						
-0.222	0.216	-0.274				-0.250
3ο επίπεδο						
-2.220						-0.045

Ο κώδικας των τεχνητών νευρωνικών δικτύων υλοποιήθηκε σε περιβάλλον Matlab και δίνεται αναλυτικά:

```
clear all
load RON.txt

d=RON;
d=d';
r=140;

x=d(1:end-1,1:r);
y=d(end,1:r);

xtest=d(1:end-1,r+1:end);
ytest=d(end,r+1:end);

[xn,minx,maxx,yn,miny,maxy]=premnmx(x,y);

testx=tramnmx(xtest,minx,maxx);
testy=tramnmx(ytest,miny,maxy);

VV.P=testx;
VV.T=testy;
VV.Pi=[];
VV.Ai=[];

net=newff(minmax(xn),[6 5 1],{'tansig','tansig','purelin'}, 'traingdx');

net.trainParam.epochs = 200;

switch_flag=0;

while (switch_flag==0)
    % Store previously trained ANN
    net1 = net;
    % Typical training
    net = train(net,xn,yn);

    resn=sim(net,xn);
    restestn=sim(net,testx);
    train_err=(norm(resn-yn))^2;
    val_err1 = (norm(restestn-testy))^2;

    fprintf('Training : %6.6f   Validation : %6.6f\n',train_err, val_err1);

    an = input('Continue with typical training (Yes=1, No=2) ?');

    if (an==2)
        switch_flag=1;
    end
end
```

```
[net,tr] = train(net1,xn,yn,[],[],VV);

res=postmnmx(resn,miny,maxy);
retest=postmnmx(restestn,miny,maxy);

subplot(2,1,1)
plot(y-res)
subplot(2,1,2)
stem(ytest-retest)

a=input('save results(Y=1, N=2)? (1/2) ');
switch a
case 1
pred=[res retest]';
qb=cell2mat(net.b);
qi=cell2mat(net.iw);
ql1=cell2mat(net.lw(2,1));
ql2=cell2mat(net.lw(3,2));

save pred.txt pred -ascii
save qb.txt qb -ascii
save qi.txt qi -ascii
save ql1.txt ql1 -ascii
save ql2.txt ql2 -ascii
case 2
end
```

Π.4 Πίνακες Δεδομένων

Πίνακας 1 Δεδομένα ογκομετρικών συστάσεων και RON των συστατικών για τις βενζίνες παραγωγής της ΜΟΗ που χρησιμοποιήθηκαν για ανάπτυξη των μοντέλων πρόβλεψης του αριθμού RON του τελικού προϊόντος

α/α	FCC	REF	ISO	ALK	MTBE	DIM	C ₄	C ₅	RON _{FCC}	RON _{REF}	RON _{ISO}	RON _{ALK}	RON _{MTBE}	RON _{DIM}	RON _{C₄}	RON _{C₅}	RON
C001	39.52	33.84	13.17	6.44	3.38	0.00	3.63	0.00	93.1	101.3	84.8	94.6	115.0	96.8	92.0	101.7	96.3
C002	40.73	25.03	15.49	10.76	3.85	0.00	4.15	0.00	93.1	101.3	84.8	94.6	115.0	96.8	92.0	101.7	96.0
C003	39.84	35.33	16.00	6.51	2.31	0.00	0.00	0.00	93.2	101.3	84.8	94.6	115.0	96.8	92.0	101.7	96.3
C004	37.78	35.26	2.20	7.86	8.65	2.67	4.83	0.74	93.2	101.3	84.8	94.6	115.0	96.8	92.0	101.7	98.0
C005	39.65	33.99	18.08	5.79	2.49	0.00	0.00	0.00	93.2	101.3	84.8	94.6	115.0	96.8	92.0	101.7	96.3
C006	40.10	34.22	16.47	6.34	2.88	0.00	0.00	0.00	93.2	101.3	84.8	94.6	115.0	96.8	92.0	101.7	96.1
C007	36.14	34.15	20.75	4.97	2.17	0.00	0.59	1.23	93.3	101.3	84.8	94.6	115.0	96.8	92.0	101.7	95.9
C008	43.53	22.59	12.84	11.66	4.44	0.00	4.00	0.94	93.3	101.3	84.8	94.6	115.0	96.8	92.0	101.7	96.2
C009	43.05	29.16	14.73	6.46	2.54	0.00	3.07	0.99	93.0	101.3	84.8	94.0	115.0	96.8	92.0	101.7	96.1
C010	52.07	40.76	1.91	0.90	1.11	0.00	0.94	2.30	93.0	101.5	84.8	94.0	115.0	96.8	92.0	101.7	96.5
C011	42.02	27.94	13.78	6.18	3.68	0.00	5.33	1.07	93.0	101.5	84.8	94.0	115.0	96.8	92.0	101.7	96.1
C012	37.10	28.78	16.11	7.49	3.30	0.00	5.55	1.69	92.6	101.5	84.8	94.0	115.0	96.8	92.0	101.7	95.7
C013	32.99	31.86	22.14	5.30	2.19	0.00	3.74	1.78	92.6	101.5	84.8	94.0	115.0	96.8	92.0	101.7	95.8
C014	37.07	28.37	16.22	8.12	3.01	0.00	5.40	1.81	92.7	101.0	84.8	94.1	115.0	96.8	92.0	101.7	95.6
C015	37.02	22.95	14.35	9.87	5.69	3.40	5.38	1.36	92.7	101.0	84.7	94.1	115.0	96.8	92.0	101.7	96.1
C016	37.19	24.30	13.03	12.22	3.27	4.14	4.26	1.60	93.8	101.0	84.7	94.0	115.0	96.8	92.0	101.7	96.2
C017	33.29	32.26	19.49	4.22	3.04	1.16	4.47	2.06	93.6	101.2	84.8	94.0	115.0	95.8	92.0	101.7	96.1
C018	38.41	24.18	11.81	12.32	2.79	4.26	4.89	1.35	94.0	101.2	84.8	94.0	115.0	95.8	92.0	101.7	96.1
C019	40.90	23.39	13.89	6.13	5.27	4.74	4.40	1.29	93.8	101.2	84.8	94.0	115.0	95.8	92.0	101.7	96.2
C020	15.57	34.79	32.02	0.00	1.56	10.48	3.67	1.91	93.7	101.3	84.8	94.0	115.0	95.3	92.0	101.7	96.2
C022	18.97	37.99	28.16	1.79	1.85	5.85	3.27	2.12	93.8	101.3	84.8	94.0	115.0	95.3	92.0	101.7	95.8
C023	30.22	32.56	20.54	4.24	2.55	3.72	4.27	1.89	94.0	101.3	84.8	94.0	115.0	95.3	92.0	101.7	96.2
C024	37.92	34.41	12.69	5.69	1.44	0.00	5.83	2.01	93.8	101.3	84.8	94.0	115.0	95.0	92.0	101.7	96.6
C025	30.27	27.29	18.84	6.62	7.71	5.08	2.64	1.56	93.8	101.1	84.8	94.0	115.0	95.0	92.0	101.7	96.4
C026	28.28	34.26	16.70	8.14	2.67	4.07	3.91	1.97	93.8	101.1	84.6	94.0	115.0	95.0	92.0	101.7	96.1
C027	30.07	32.84	19.05	5.54	3.20	3.45	3.96	1.88	93.7	101.1	84.6	94.0	115.0	95.0	92.0	101.7	96.7
C028	24.67	36.71	21.52	6.85	2.48	2.18	3.63	1.97	93.8	101.1	84.6	94.0	115.0	95.0	92.0	101.7	96.2
C029	25.77	35.90	21.72	6.17	2.28	2.52	3.66	1.98	93.7	101.1	84.6	94.0	115.0	95.0	92.0	101.7	96.1
C030	24.50	37.11	4.98	11.23	7.16	5.55	7.42	2.05	93.8	101.1	84.6	94.0	115.0	95.0	92.0	101.7	98.1
C031	21.43	40.30	12.76	7.02	2.24	0.00	4.20	2.16	93.8	101.1	84.6	94.0	115.0	95.0	92.0	101.7	96.2
C032	30.89	31.22	18.52	6.69	3.31	2.78	4.88	1.71	93.7	100.7	84.6	94.0	115.0	94.8	92.0	101.7	96.5
C033	25.80	29.27	20.36	6.18	3.28	0.00	3.63	1.64	93.6	100.7	84.6	94.0	115.0	94.8	92.0	101.7	95.7
C034	31.06	31.63	18.19	6.63	3.34	3.35	3.97	1.82	93.5	100.7	84.6	93.9	115.0	94.8	92.0	101.7	96.5
C035	34.68	24.52	16.36	11.30	4.03	5.02	2.53	1.55	93.5	100.7	84.6	93.9	115.0	94.8	92.0	101.7	96.2
C036	25.64	28.11	22.22	5.24	2.49	4.67	3.19	1.80	93.6	100.7	84.6	93.9	115.0	94.8	92.0	101.7	95.7
C037	32.12	33.63	1.76	10.51	7.15	6.47	6.21	2.15	93.7	100.7	84.6	93.9	115.0	94.8	92.0	101.7	98.0
C038	27.94	32.97	17.85	6.90	3.41	5.50	3.96	1.46	93.8	100.7	84.6	93.9	115.0	94.8	92.0	101.7	96.2
C039	34.87	22.91	17.02	11.97	5.15	4.86	2.49	0.73	93.7	100.9	84.6	93.9	115.0	94.8	92.0	101.7	96.0
C040	13.26	41.33	13.74	5.15	2.86	2.60	2.33	2.28	93.4	100.9	84.6	93.9	115.0	94.8	92.0	101.7	95.5
C041	25.46	35.95	13.78	8.05	2.57	5.07	6.18	2.95	93.4	100.9	84.6	93.9	115.0	94.8	92.0	101.7	96.3
C042	29.15	34.61	14.96	8.01	3.40	2.49	4.95	2.43	92.6	100.9	84.9	93.6	115.0	94.7	92.0	101.7	96.3
C043	24.71	34.77	17.83	6.77	4.29	4.68	4.68	2.28	92.6	100.9	84.9	93.6	115.0	94.7	92.0	101.7	96.3
C044	27.44	32.78	17.52	6.35	2.80	2.88	4.00	2.10	93.5	100.9	84.9	93.6	115.0	94.7	92.0	101.7	96.2
C045	27.57	34.74	15.77	7.69	3.78	3.24	4.83	2.39	93.5	100.9	84.9	93.6	115.0	94.7	92.0	101.7	96.3
C046	31.62	27.83	16.71	7.92	3.53	4.56	5.34	1.71	92.9	100.9	84.9	93.6	115.0	94.7	92.0	101.7	96.2
C047	29.98	30.16	16.01	7.53	2.83	6.22	4.57	1.79	93.3	100.8	84.9	93.6	115.0	94.7	92.0	101.7	96.0
C048	31.30	25.53	17.36	9.28	3.91	6.65	3.43	1.50	93.3	100.8	84.9	93.6	115.0	94.3	92.0	101.7	96.0
C049	26.33	32.25	20.04	6.37	3.19	4.42	3.52	1.87	93.3	100.8	84.9	93.6	115.0	94.3	92.0	101.7	96.0
C051	27.36	33.90	7.03	11.12	5.99	6.76	4.88	2.10	93.2	100.8	84.9	93.7	115.0	94.3	92.0	101.7	98.0
C052	31.07	28.45	16.53	7.24	3.33	7.19	4.51	1.68	93.4	100.8	84.1	93.7	115.0	94.3	92.0	101.7	96.2
C053	29.54	26.11	21.89	7.06	4.64	4.81	3.93	2.02	93.4	100.8	84.1	93.7	115.0	94.3	92.0	101.7	96.4
C054	28.00	33.97	7.30	8.02	7.05	8.36	5.50	1.79	93.4	100.8	84.1	93.7	115.0	94.3	92.0	101.7	98.0
C056	26.71	25.05	25.01	7.33	4.01	5.01	3.62	1.25	93.3	101.3	84.1	93.7	115.0	95.0	92.0	101.7	95.5
C057	28.47	26.57	21.38	7.26	4.18	4.67	3.39	1.37	93.2	101.3	84.1	93.7	115.0	95.0	92.0	101.7	95.7
C058	27.09	35.59	4.99	6.51	7.05	6.01	6.56	1.85	93.2	101.3	84.1	93.7	115.0	95.0	92.0	101.7	98.0

α/α	FCC	REF	ISO	ALK	MTBE	DIM	C ₄	C ₅	RON _{FCC}	RON _{REF}	RON _{ISO}	RON _{ALK}	RON _{MTBE}	RON _{DIM}	RON _{C₄}	RON _{C₅}	RON
C059	28.86	29.06	11.86	4.47	5.10	6.12	5.61	1.46	93.2	101.3	84.1	93.7	115.0	95.0	92.0	101.7	96.0
C060	7.73	39.23	24.28	7.43	2.42	8.58	5.14	1.98	93.3	101.3	84.1	93.9	115.0	95.0	92.0	101.7	96.2
C061	31.74	28.39	15.36	5.49	4.37	6.14	4.66	1.44	93.3	101.3	84.1	93.9	115.0	95.0	92.0	101.7	96.1
C062	28.39	31.74	15.36	5.49	4.37	6.14	4.65	1.44	93.4	101.3	84.2	93.9	115.0	95.0	92.0	101.7	96.2
C063	28.63	29.84	19.12	6.93	2.92	6.25	4.75	1.56	93.3	101.3	84.2	93.9	115.0	95.0	92.0	101.7	96.1
C064	23.05	43.05	0.68	8.96	6.42	8.20	6.85	2.23	93.3	101.3	84.2	93.9	115.0	95.0	92.0	101.7	99.0
C065	33.39	31.53	15.40	6.47	1.12	3.53	6.81	1.76	93.4	101.3	84.2	93.9	115.0	95.0	92.0	101.7	96.1
C066	22.87	33.02	23.39	6.57	1.67	3.81	4.53	1.76	93.3	101.2	84.2	93.9	115.0	95.0	92.0	101.7	96.0
C067	31.44	30.80	18.69	6.33	2.53	3.76	4.67	1.78	93.3	101.2	84.2	93.9	115.0	95.0	92.0	101.7	96.0
C068	26.10	28.90	25.40	6.17	2.11	3.52	4.01	1.68	93.3	101.2	84.2	93.9	115.0	94.4	92.0	101.7	95.7
C069	32.95	27.24	19.15	6.34	4.34	3.76	4.73	1.49	93.3	101.2	84.2	93.9	115.0	94.4	92.0	101.7	96.0
C070	25.27	34.88	11.03	7.08	6.28	8.21	5.45	1.79	93.4	101.2	84.2	93.9	115.0	94.4	92.0	101.7	98.0
C071	27.51	31.05	21.23	6.18	3.15	2.98	4.76	1.60	93.4	101.2	84.2	93.9	115.0	94.4	92.0	101.7	96.0
C072	29.99	30.80	18.74	5.61	3.28	3.86	4.89	1.55	93.4	101.2	84.1	93.9	115.0	94.4	92.0	101.7	95.6
C073	28.12	32.81	8.61	8.23	9.59	5.51	5.45	1.67	93.4	101.2	84.1	93.9	115.0	94.4	92.0	101.7	98.0
C074	29.27	26.93	18.80	9.78	4.06	6.76	3.03	1.37	93.5	101.3	84.1	93.9	115.0	94.4	92.0	101.7	96.1
C075	27.91	30.24	14.32	8.44	3.05	3.69	5.43	1.53	93.5	101.3	84.1	93.9	115.0	94.4	92.0	101.7	96.1
C076	30.43	29.42	6.10	9.35	8.28	8.92	6.00	1.51	93.5	101.3	84.1	93.9	115.0	95.0	92.0	101.7	98.0
C078	34.14	29.76	15.94	7.19	3.71	2.32	5.37	1.56	93.5	101.3	84.1	94.5	115.0	95.0	92.0	101.7	96.1
C079	32.97	29.99	15.78	8.16	3.28	3.18	5.07	1.57	93.5	101.3	84.1	94.5	115.0	95.0	92.0	101.7	96.0
C080	34.41	29.74	16.42	6.10	3.81	2.54	5.31	1.66	93.5	101.3	84.6	94.5	115.0	95.0	92.0	101.7	96.0
C081	29.31	33.69	17.25	6.15	3.88	3.30	4.81	1.61	93.7	101.3	84.6	94.5	115.0	95.0	92.0	101.7	96.0
C083	33.75	29.36	17.20	6.15	3.88	3.28	4.77	1.62	93.7	101.3	84.6	94.5	115.0	95.0	92.0	101.7	96.0
C084	30.20	29.58	16.46	6.14	3.82	7.78	4.40	1.63	93.7	101.2	84.6	94.5	115.0	95.0	92.0	101.7	96.1
C085	32.13	32.16	14.44	8.46	2.76	4.28	4.07	1.71	93.7	101.2	84.6	94.5	115.0	95.0	92.0	101.7	96.0
C086	27.93	35.14	6.64	5.15	9.88	7.25	6.12	1.89	93.7	101.2	84.6	94.5	115.0	94.8	92.0	101.7	98.0
C087	33.19	25.25	21.14	6.56	3.44	4.94	4.15	1.34	93.6	101.2	84.6	94.5	115.0	94.8	92.0	101.7	96.2
C088	32.13	32.16	14.44	8.46	2.76	4.28	4.07	1.71	93.6	101.2	84.6	94.5	115.0	94.8	92.0	101.7	96.1
C089	32.42	25.56	19.59	7.89	4.07	5.00	4.21	1.26	93.6	101.2	84.6	94.3	115.0	94.8	92.0	101.7	96.1
C090	31.17	32.44	14.79	7.53	2.65	4.63	5.21	1.58	93.7	101.2	84.6	94.3	115.0	94.8	92.0	101.7	96.1
C091	32.13	24.61	21.16	8.44	2.59	5.58	3.92	1.58	93.7	101.2	84.8	94.3	115.0	94.8	92.0	101.7	96.1
C092	32.00	26.02	20.24	7.82	3.43	5.44	3.68	1.37	93.7	101.2	84.8	94.3	115.0	94.8	92.0	101.7	96.1
C093	22.16	26.23	24.64	10.38	3.23	6.82	4.99	1.55	93.7	101.2	84.8	94.3	115.0	95.0	92.0	101.7	96.0
C094	6.60	43.37	9.56	4.34	12.87	14.15	5.72	2.37	93.7	101.2	84.8	94.3	115.0	95.0	92.0	101.7	99.0
C095	34.14	29.91	15.00	7.49	2.75	3.12	4.63	1.55	93.9	101.2	84.8	94.3	115.0	95.0	92.0	101.7	96.0
C096	27.77	31.10	17.61	7.78	2.89	6.06	4.42	1.64	93.5	101.2	84.8	94.3	115.0	95.0	92.0	101.7	96.0
C097	31.83	39.22	0.00	11.40	2.17	6.14	7.10	2.13	93.5	101.2	84.8	93.7	115.0	95.0	92.0	101.7	98.1
C098	29.90	30.88	16.78	7.70	2.85	5.11	4.24	1.61	93.6	103.1	84.8	93.7	115.0	95.0	92.0	101.7	96.1
C099	31.33	31.08	17.18	7.21	2.58	3.94	4.06	1.69	93.5	103.1	84.8	93.7	115.0	95.0	92.0	101.7	96.1
C100	31.40	27.71	15.28	10.91	5.91	3.22	4.48	1.08	93.4	103.1	84.4	93.7	115.0	95.0	92.0	101.7	96.1
C101	33.83	27.69	21.75	4.84	2.33	3.50	4.40	1.66	93.4	103.3	84.4	93.7	115.0	95.0	92.0	101.7	96.0
C102	34.73	28.12	18.97	5.88	2.47	4.06	4.19	1.59	93.4	103.3	84.4	93.7	115.0	95.0	92.0	101.7	96.0
C103	33.68	27.15	19.37	6.01	3.46	4.04	4.72	1.58	93.4	103.3	84.4	93.7	115.0	95.0	92.0	101.7	96.2
C104	33.19	30.09	4.27	11.47	6.93	5.19	7.01	1.86	93.4	103.3	84.4	93.8	115.0	95.0	92.0	101.7	98.0
C105	37.01	23.25	18.96	6.57	3.72	4.46	4.62	1.42	93.4	101.9	84.4	93.8	115.0	95.0	92.0	101.7	96.2
C106	36.72	27.22	18.40	5.71	2.41	3.94	4.07	1.54	93.4	101.9	84.4	93.8	115.0	95.0	92.0	101.7	96.0
C107	36.63	24.27	20.18	10.21	3.68	3.53	0.00	1.49	93.4	101.9	84.4	93.8	115.0	95.0	92.0	101.7	96.4
C108	35.04	23.25	19.12	6.10	5.01	5.31	4.84	1.34	93.4	101.9	84.4	93.8	115.0	95.0	92.0	101.7	96.0
C109	34.18	31.70	18.82	2.22	2.90	3.40	4.96	1.83	93.5	101.9	84.4	93.8	115.0	95.0	92.0	101.7	96.0
C110	35.87	27.99	16.37	5.82	2.52	4.73	5.07	1.62	93.3	101.9	84.4	93.8	115.0	95.0	92.0	101.7	96.0
C111	35.10	28.45	18.52	3.72	3.76	3.44	5.39	1.63	93.1	101.9	84.4	93.8	115.0	95.0	92.0	101.7	96.1
C112	33.66	30.67	17.41	4.52	3.24	3.23	5.56	1.71	93.3	102.0	84.4	93.8	115.0	94.5	92.0	101.7	96.1
C113	34.08	28.23	17.11	6.31	3.08	4.84	4.86	1.49	93.3	102.0	84.4	93.7	115.0	94.5	92.0	101.7	96.1
C114	35.60	28.94	18.35	3.61	3.70	2.92	5.27	1.61	93.3	102.0	84.4	93.7	115.0	94.5	92.0	101.7	96.0
C115	34.68	28.53	17.58	5.62	3.29	4.27	4.49	1.53	93.3	102.0	84.7	93.7	115.0	94.5	92.0	101.7	96.2
C116	31.44	31.86	0.50	10.39	9.73	7.48	6.88	1.73	93.3	102.0	84.7	93.7	115.0	94.5	92.0	101.7	98.0
C117	34.84	25.16	18.78	6.61	4.50	3.40	5.33	1.37	93.3	102.0	84.7	93.7	115.0	94.5	92.0	101.7	96.0
C118	35.84	22.19	23.91	6.03	3.84	2.28	4.73	1.19	93.3	102.0	84.7	93.7	115.0	94.5	92.0	101.7	96.0
C119	39.54	27.37	17.79	4.14	4.64	0.00	5.06	1.47	93.3	102.0	84.7	93.7	115.0	94.6	92.0	101.7	95.9R
C120	38.96	27.80	16.89	5.07	4.72	0.00	5.04	1.53	93.3	102.0	84.7	93.7	115.0	94.6	92.0	101.7	96.0
C121	32.00	32.30	17.84	5.27	1.59	3.83	5.43	1.74	93.4	102.3	84.7	93.7	115.0	94.6	92.0	101.7	96.0

α/α	FCC	REF	ISO	ALK	MTBE	DIM	C ₄	C ₅	RON _{FCC}	RON _{REF}	RON _{ISO}	RON _{ALK}	RON _{MTBE}	RON _{DIM}	RON _{C4}	RON _{C5}	RON
C122	41.13	26.94	15.56	5.20	4.86	0.00	4.82	1.48	93.4	102.3	84.7	93.7	115.0	94.6	92.0	101.7	96.1
C123	36.91	31.49	19.24	1.80	3.87	0.00	5.06	1.64	93.4	102.3	84.7	94.1	115.0	94.6	92.0	101.7	96.1
C125	35.94	37.60	11.97	0.51	12.25	0.00	0.00	1.73	93.5	102.3	81.3	94.1	115.0	94.6	92.0	101.7	97.9R
C126	39.19	33.35	15.54	8.04	2.44	0.00	0.00	1.44	93.2	102.3	81.3	94.1	115.0	94.6	92.0	101.7	95.5
C127	36.64	33.60	16.71	6.41	3.01	2.11	0.00	1.51	93.0	101.8	81.3	94.1	115.0	94.8	92.0	101.7	96.0
C128	34.63	33.39	17.83	7.10	3.12	2.39	0.00	1.54	93.0	101.8	81.3	93.4	115.0	94.8	92.0	101.7	95.7
C129	34.87	34.52	16.19	8.15	2.65	2.15	0.00	1.46	92.9	101.8	81.4	93.4	115.0	94.8	92.0	101.7	96.1
C130	35.83	30.95	14.93	8.92	3.33	4.94	0.00	1.11	92.9	101.8	81.4	93.4	115.0	94.8	92.0	101.7	96.0
C131	33.74	33.31	16.11	7.85	2.87	5.02	0.00	1.10	92.8	101.8	77.5	93.4	115.0	94.8	92.0	101.7	96.1
C132	33.58	37.32	7.36	7.28	5.12	0.00	5.12	1.03	92.5	101.8	81.2	93.5	115.0	95.0	92.0	101.7	96.1
C133	38.05	33.41	10.59	6.70	2.79	2.80	0.00	1.12	92.4	101.8	81.8	93.5	115.0	95.0	92.0	101.7	96.0
C134	35.86	34.49	14.25	4.56	3.76	2.76	0.00	1.34	93.6	102.5	81.8	93.5	115.0	94.5	92.0	101.7	96.0
C135	37.17	29.78	14.19	7.00	3.93	4.56	0.00	0.00	93.1	102.5	84.0	94.0	115.0	94.5	92.0	101.7	96.1
C136	37.06	30.04	14.08	7.84	3.83	4.54	0.00	1.40	93.1	102.5	84.0	94.0	115.0	93.9	92.0	101.7	96.0
C137	37.25	29.91	14.66	7.60	3.75	4.48	0.00	1.42	93.3	103.0	83.6	93.4	115.0	93.9	92.0	101.7	96.1
C138	35.48	31.35	16.43	8.09	3.57	3.64	0.00	1.44	93.5	101.3	83.6	94.0	115.0	94.8	92.0	101.7	96.0
C139	36.16	33.39	15.28	7.53	3.08	2.84	0.00	1.73	93.5	101.1	84.0	93.8	115.0	94.9	92.0	101.7	96.0
C140	26.19	34.33	16.65	7.64	3.59	3.59	0.70	7.31	93.3	101.6	77.5	94.1	115.0	94.3	92.0	101.7	96.0
C141	35.26	29.14	17.30	5.87	3.86	3.65	0.54	4.38	93.1	100.3	77.5	93.6	115.0	94.1	92.0	101.7	96.0
C142	28.07	32.82	19.86	3.92	4.18	8.56	0.62	1.97	93.1	100.7	83.4	93.6	115.0	94.2	92.0	101.7	96.2
C143	30.58	32.40	16.39	6.62	4.52	6.97	0.67	1.84	93.1	100.5	83.5	93.6	115.0	94.1	92.0	101.7	96.1
C144	31.39	35.13	17.55	3.95	4.67	4.62	0.84	1.84	93.1	100.5	84.0	93.6	115.0	94.1	92.0	101.7	96.0
C145	40.95	26.78	15.60	4.52	2.35	4.12	0.52	1.35	93.1	101.6	83.3	94.5	115.0	94.3	92.0	101.7	96.0
C147	85.20	9.88	0.00	0.00	1.90	0.00	3.02	0.00	93.1	102.1	83.9	94.2	115.0	94.8	92.0	101.7	94.5
C148	92.06	4.96	0.00	0.00	1.24	0.00	1.75	0.00	93.1	101.7	84.0	93.3	115.0	93.8	92.0	101.7	94.1
C149	96.19	1.67	0.00	0.00	0.00	0.00	2.13	0.00	93.1	102.5	83.3	94.0	115.0	94.9	92.0	101.7	94.0
C150	35.02	28.16	12.59	9.49	3.78	7.43	1.21	1.13	93.1	102.2	83.4	93.7	115.0	94.0	92.0	101.7	96.0
C151	80.73	6.20	2.69	5.61	0.58	0.00	2.14	0.00	93.1	102.0	83.3	93.4	115.0	93.9	92.0	101.7	93.9
C152	42.19	25.47	14.17	7.69	3.49	4.20	0.84	1.95	93.1	102.0	83.3	93.4	115.0	93.8	92.0	101.7	96.0
C153	87.51	5.86	2.86	2.25	0.00	0.00	0.71	0.00	93.1	102.0	83.3	93.6	115.0	93.8	92.0	101.7	94.1
C154	96.49	0.73	0.00	0.00	0.00	0.00	2.78	0.00	93.1	102.3	81.4	93.6	115.0	93.8	92.0	101.7	93.7
C156	88.42	6.83	0.00	0.00	2.04	0.00	2.71	0.00	93.1	103.0	84.5	93.6	115.0	94.5	92.0	101.7	94.1
C157	97.14	0.92	0.00	0.00	0.00	0.00	1.94	0.00	93.1	103.0	84.5	93.6	115.0	94.5	92.0	101.7	93.5
C158	91.01	5.19	0.00	0.00	1.49	0.00	2.31	0.00	93.1	103.0	84.5	93.8	115.0	94.4	92.0	101.7	94.0
C160	40.88	34.85	9.78	0.00	6.61	0.00	5.25	0.00	93.1	102.8	82.1	93.8	115.0	97.6	92.0	101.7	96.0
C161	42.22	31.79	10.57	1.96	6.60	0.93	3.44	0.00	93.1	102.8	82.1	93.8	115.0	97.6	92.0	101.7	96.0
C162	40.08	30.84	9.95	11.42	5.43	2.28	0.00	0.00	93.1	100.1	78.3	93.7	115.0	97.3	92.0	101.7	96.3
C163	35.81	29.63	11.99	9.73	4.34	5.13	3.37	0.00	93.1	102.1	83.6	94.0	115.0	94.8	92.0	101.7	96.1
C164	40.35	26.19	14.09	8.74	3.71	4.79	2.12	0.00	93.1	101.3	83.6	94.0	115.0	94.8	92.0	101.7	95.9
C165	34.35	25.15	17.95	6.58	3.40	5.86	5.27	1.43	93.1	102.4	83.2	94.2	115.0	94.6	92.0	101.7	96.0

Πίνακας 2 Δεδομένα ογκομετρικών συστάσεων και MON των συστατικών για τις βενζίνες παραγωγής της MOH που χρησιμοποιήθηκαν για ανάπτυξη των μοντέλων πρόβλεψης του αριθμού MON του τελικού προϊόντος

α/α	FCC	REF	ISO	ALK	MTBE	DIM	C ₄	C ₅	MON _{FCC}	MON _{REF}	MON _{ISO}	MON _{ALK}	MON _{MTBE}	MON _{DIM}	MON _{C4}	MON _{C5}	MON
C014	37.07	28.37	16.22	8.12	3.01	0.00	5.40	1.81	81.2	89.8	82.6	91.2	95.0	81.8	82.0	90.3	85.0
C015	37.02	22.95	14.35	9.87	5.69	3.40	5.38	1.36	81.2	89.8	82.1	91.2	95.0	81.8	82.0	90.3	85.0
C016	37.19	24.30	13.03	12.22	3.27	4.14	4.26	1.60	80.8	89.8	82.1	92.0	95.0	81.8	82.0	90.3	85.0
C017	33.29	32.26	19.49	4.22	3.04	1.16	4.47	2.06	80.8	90.0	82.2	92.0	95.0	80.9	82.0	90.3	85.1
C018	38.41	24.18	11.81	12.32	2.79	4.26	4.89	1.35	81.1	90.0	82.2	92.0	95.0	80.9	82.0	90.3	85.1
C019	40.90	23.39	13.89	6.13	5.27	4.74	4.40	1.29	80.1	90.0	82.2	91.9	95.0	80.9	82.0	90.3	85.0
C020	15.57	34.79	32.02	0.00	1.56	10.48	3.67	1.91	80.1	89.8	82.2	91.9	95.0	80.5	82.0	90.3	85.0
C022	18.97	37.99	28.16	1.79	1.85	5.85	3.27	2.12	80.1	89.8	82.1	91.9	95.0	80.5	82.0	90.3	85.0
C023	30.22	32.56	20.54	4.24	2.55	3.72	4.27	1.89	80.1	89.8	82.1	91.9	95.0	80.5	82.0	90.3	84.9
C024	37.92	34.41	12.69	5.69	1.44	0.00	5.83	2.01	80.1	89.8	82.1	91.9	95.0	80.5	82.0	90.3	85.0
C025	30.27	27.29	18.84	6.62	7.71	5.08	2.64	1.56	80.1	89.8	82.1	91.9	95.0	80.5	82.0	90.3	85.1
C026	28.28	34.26	16.70	8.14	2.67	4.07	3.91	1.97	80.1	89.8	82.1	91.9	95.0	80.5	82.0	90.3	84.9
C027	30.07	32.84	19.05	5.54	3.20	3.45	3.96	1.88	80.1	89.8	82.1	91.8	95.0	80.5	82.0	90.3	85.1
C028	24.67	36.71	21.52	6.85	2.48	2.18	3.63	1.97	80.1	89.7	82.1	91.8	95.0	80.2	82.0	90.3	85.0
C029	25.77	35.90	21.72	6.17	2.28	2.52	3.66	1.98	80.1	89.7	82.1	91.8	95.0	80.2	82.0	90.3	85.0
C030	24.50	37.11	4.98	11.23	7.16	5.55	7.42	2.05	80.4	89.7	82.1	91.8	95.0	80.2	82.0	90.3	86.0
C031	21.43	40.30	12.76	7.02	2.24	0.00	4.20	2.16	80.1	89.7	82.1	91.8	95.0	80.2	82.0	90.3	85.0
C032	30.89	31.22	18.52	6.69	3.31	2.78	4.88	1.71	80.7	89.7	82.1	91.8	95.0	80.2	82.0	90.3	85.0
C033	25.80	29.27	20.36	6.18	3.28	0.00	3.63	1.64	80.7	89.7	82.1	91.8	95.0	80.2	82.0	90.3	85.0
C034	31.06	31.63	18.19	6.63	3.34	3.35	3.97	1.82	80.7	89.7	82.1	91.8	95.0	80.2	82.0	90.3	85.0
C035	34.68	24.52	16.36	11.30	4.03	5.02	2.53	1.55	80.7	89.8	82.1	91.8	95.0	80.2	82.0	90.3	85.0
C036	25.64	28.11	22.22	5.24	2.49	4.67	3.19	1.80	80.7	89.8	82.1	91.8	95.0	80.1	82.0	90.3	85.0
C037	32.12	33.63	1.76	10.51	7.15	6.47	6.21	2.15	80.7	89.8	82.1	91.8	95.0	80.1	82.0	90.3	86.0
C038	27.94	32.97	17.85	6.90	3.41	5.50	3.96	1.46	80.7	89.8	82.1	91.8	95.0	80.1	82.0	90.3	85.0
C039	34.87	22.91	17.02	11.97	5.15	4.86	2.49	0.73	80.2	89.8	82.1	91.8	95.0	80.1	82.0	90.3	85.0
C040	13.26	41.33	13.74	5.15	2.86	2.60	2.33	2.28	80.2	89.8	82.1	91.8	95.0	80.1	82.0	90.3	85.0
C041	25.46	35.95	13.78	8.05	2.57	5.07	6.18	2.95	80.2	89.8	82.1	91.6	95.0	80.1	82.0	90.3	85.0
C042	29.15	34.61	14.96	8.01	3.40	2.49	4.95	2.43	80.1	90.0	82.4	91.6	95.0	80.1	82.0	90.3	84.6
C043	24.71	34.77	17.83	6.77	4.29	4.68	4.68	2.28	80.1	90.0	82.4	91.6	95.0	80.1	82.0	90.3	85.0
C044	27.44	32.78	17.52	6.35	2.80	2.88	4.00	2.10	80.1	90.0	82.4	91.6	95.0	80.1	82.0	90.3	85.0
C045	27.57	34.74	15.77	7.69	3.78	3.24	4.83	2.39	80.1	90.0	82.4	91.6	95.0	80.1	82.0	90.3	85.0
C046	31.62	27.83	16.71	7.92	3.53	4.56	5.34	1.71	80.1	90.0	82.4	91.6	95.0	80.1	82.0	90.3	85.0
C047	29.98	30.16	16.01	7.53	2.83	6.22	4.57	1.79	80.4	90.0	82.4	91.6	95.0	80.1	82.0	90.3	85.0
C048	31.30	25.53	17.36	9.28	3.91	6.65	3.43	1.50	80.4	90.0	82.4	91.6	95.0	80.1	82.0	90.3	85.0
C049	26.33	32.25	20.04	6.37	3.19	4.42	3.52	1.87	80.4	90.0	82.4	91.6	95.0	80.1	82.0	90.3	85.0
C051	27.36	33.90	7.03	11.12	5.99	6.76	4.88	2.10	80.3	90.0	82.4	91.7	95.0	80.1	82.0	90.3	86.0
C052	31.07	28.45	16.53	7.24	3.33	7.19	4.51	1.68	80.3	90.1	82.0	91.7	95.0	80.1	82.0	90.3	85.0
C053	29.54	26.11	21.89	7.06	4.64	4.81	3.93	2.02	80.3	90.1	82.0	91.7	95.0	80.1	82.0	90.3	85.0
C054	28.00	33.97	7.30	8.02	7.05	8.36	5.50	1.79	80.3	90.1	82.0	91.7	95.0	80.1	82.0	90.3	86.0
C055	28.71	26.78	20.63	7.22	4.20	4.59	3.65	1.39	80.3	90.1	82.0	91.7	95.0	80.1	82.0	90.3	85.0
C056	26.71	25.05	25.01	7.33	4.01	5.01	3.62	1.25	80.3	90.1	82.0	91.7	95.0	80.1	82.0	90.3	85.0
C057	28.47	26.57	21.38	7.26	4.18	4.67	3.39	1.37	80.3	90.1	82.0	91.7	95.0	80.1	82.0	90.3	85.0
C058	27.09	35.59	4.99	6.51	7.05	6.01	6.56	1.85	80.3	90.1	82.0	91.7	95.0	80.1	82.0	90.3	86.0
C059	28.86	29.06	11.86	4.47	5.10	6.12	5.61	1.46	80.3	90.1	82.0	91.7	95.0	80.1	82.0	90.3	85.0
C060	7.73	39.23	24.28	7.43	2.42	8.58	5.14	1.98	80.3	90.1	82.0	91.7	95.0	80.1	82.0	90.3	85.0
C061	31.74	28.39	15.36	5.49	4.37	6.14	4.66	1.44	80.3	90.1	82.0	91.7	95.0	80.1	82.0	90.3	85.0
C062	28.39	31.74	15.36	5.49	4.37	6.14	4.65	1.44	80.6	90.2	82.0	91.8	95.0	79.8	82.0	90.3	85.0
C063	28.63	29.84	19.12	6.93	2.92	6.25	4.75	1.56	80.6	90.2	82.0	91.8	95.0	79.8	82.0	90.3	85.0
C064	23.05	43.05	0.68	8.96	6.42	8.20	6.85	2.23	80.6	90.2	82.0	91.8	95.0	79.8	82.0	90.3	86.4
C065	33.39	31.53	15.40	6.47	1.12	3.53	6.81	1.76	80.6	90.2	82.0	91.8	95.0	79.8	82.0	90.3	84.9
C066	22.87	33.02	23.39	6.57	1.67	3.81	4.53	1.76	80.6	90.2	82.0	91.8	95.0	79.8	82.0	90.3	85.0
C067	31.44	30.80	18.69	6.33	2.53	3.76	4.67	1.78	80.6	90.2	82.0	91.8	95.0	79.8	82.0	90.3	85.0
C068	26.10	28.90	25.40	6.17	2.11	3.52	4.01	1.68	80.6	90.2	82.0	91.8	95.0	79.8	82.0	90.3	85.0
C069	32.95	27.24	19.15	6.34	4.34	3.76	4.73	1.49	80.6	90.2	82.0	91.8	95.0	79.8	82.0	90.3	85.0
C070	25.27	34.88	11.03	7.08	6.28	8.21	5.45	1.79	80.6	90.2	82.0	91.8	95.0	79.8	82.0	90.3	86.0
C071	27.51	31.05	21.23	6.18	3.15	2.98	4.76	1.60	80.6	90.9	82.0	91.7	95.0	79.8	82.0	90.3	85.0
C072	29.99	30.80	18.74	5.61	3.28	3.86	4.89	1.55	80.6	90.9	82.0	91.7	95.0	79.8	82.0	90.3	84.9
C073	28.12	32.81	8.61	8.23	9.59	5.51	5.45	1.67	80.6	90.9	82.0	91.7	95.0	79.8	82.0	90.3	86.1
C074	29.27	26.93	18.80	9.78	4.06	6.76	3.03	1.37	80.6	90.9	82.0	91.7	95.0	79.8	82.0	90.3	85.0
C075	27.91	30.24	14.32	8.44	3.05	3.69	5.43	1.53	80.6	90.9	82.0	91.7	95.0	79.8	82.0	90.3	84.9
C076	30.43	29.42	6.10	9.35	8.28	8.92	6.00	1.51	80.6	90.9	82.0	91.7	95.0	79.8	82.0	90.3	86.0
C078	34.14	29.76	15.94	7.19	3.71	2.32	5.37	1.56	80.6	91.0	82.1	92.1	95.0	79.9	82.0	90.3	85.0
C079	32.97	29.99	15.78	8.16	3.28	3.18	5.07	1.57	80.6	91.0	82.1	92.1	95.0	79.9	82.0	90.3	85.0
C080	34.41	29.74	16.42	6.10	3.81	2.54	5.31	1.66	80.6	91.0	82.1	92.1	95.0	79.9	82.0	90.3	85.0

α/α	FCC	REF	ISO	ALK	MTBE	DIM	C₄	C₅	MON_{FCC}	MON_{REF}	MON_{ISO}	MON_{ALK}	MON_{MTBE}	MON_{DIM}	MON_{C₄}	MON_{C₅}	MON
C081	29.31	33.69	17.25	6.15	3.88	3.30	4.81	1.61	80.6	91.0	82.1	92.1	95.0	79.9	82.0	90.3	85.0
C083	33.75	29.36	17.20	6.15	3.88	3.28	4.77	1.62	80.6	91.0	82.1	92.1	95.0	79.9	82.0	90.3	85.0
C084	30.20	29.58	16.46	6.14	3.82	7.78	4.40	1.63	80.6	91.0	82.1	92.1	95.0	79.9	82.0	90.3	85.0
C085	32.13	32.16	14.44	8.46	2.76	4.28	4.07	1.71	80.6	91.0	82.1	92.1	95.0	79.9	82.0	90.3	85.0
C086	27.93	35.14	6.64	5.15	9.88	7.25	6.12	1.89	80.6	91.0	82.1	92.1	95.0	79.9	82.0	90.3	86.0
C087	33.19	25.25	21.14	6.56	3.44	4.94	4.15	1.34	80.6	91.0	82.1	92.1	95.0	79.9	82.0	90.3	85.0
C088	32.13	32.16	14.44	8.46	2.76	4.28	4.07	1.71	80.6	91.0	82.1	92.1	95.0	79.9	82.0	90.3	85.0
C089	32.42	25.56	19.59	7.89	4.07	5.00	4.21	1.26	80.6	91.0	82.1	92.1	95.0	79.9	82.0	90.3	85.0
C090	31.17	32.44	14.79	7.53	2.65	4.63	5.21	1.58	80.6	91.0	82.1	92.1	95.0	79.9	82.0	90.3	85.0
C091	32.13	24.61	21.16	8.44	2.59	5.58	3.92	1.58	80.6	91.0	82.1	92.1	95.0	79.9	82.0	90.3	85.0
C092	32.00	26.02	20.24	7.82	3.43	5.44	3.68	1.37	80.7	89.7	82.5	92.1	95.0	79.8	82.0	90.3	85.0
C093	22.16	26.23	24.64	10.38	3.23	6.82	4.99	1.55	80.7	89.7	82.5	92.1	95.0	79.8	82.0	90.3	84.9
C094	6.60	43.37	9.56	4.34	12.87	14.15	5.72	2.37	80.7	89.7	82.5	92.1	95.0	79.8	82.0	90.3	86.5
C095	34.14	29.91	15.00	7.49	2.75	3.12	4.63	1.55	80.7	89.7	82.5	92.1	95.0	79.8	82.0	90.3	85.0
C096	27.77	31.10	17.61	7.78	2.89	6.06	4.42	1.64	80.7	89.7	82.5	92.1	95.0	79.8	82.0	90.3	85.0
C097	31.83	39.22	0.00	11.40	2.17	6.14	7.10	2.13	80.7	89.7	82.5	92.1	95.0	79.8	82.0	90.3	86.0
C098	29.90	30.88	16.78	7.70	2.85	5.11	4.24	1.61	80.7	89.7	82.5	92.1	95.0	79.8	82.0	90.3	85.0
C099	31.33	31.08	17.18	7.21	2.58	3.94	4.06	1.69	81.1	92.1	82.4	91.7	95.0	79.9	82.0	90.3	85.0
C100	31.40	27.71	15.28	10.91	5.91	3.22	4.48	1.08	81.1	92.1	82.4	91.7	95.0	79.9	82.0	90.3	85.0
C101	33.83	27.69	21.75	4.84	2.33	3.50	4.40	1.66	81.1	92.1	82.4	91.7	95.0	79.9	82.0	90.3	85.0
C102	34.73	28.12	18.97	5.88	2.47	4.06	4.19	1.59	81.1	92.1	82.4	91.7	95.0	79.9	82.0	90.3	85.0
C103	33.68	27.15	19.37	6.01	3.46	4.04	4.72	1.58	81.1	92.1	82.4	91.7	95.0	79.9	82.0	90.3	85.0
C104	33.19	30.09	4.27	11.47	6.93	5.19	7.01	1.86	81.1	92.1	82.4	91.7	95.0	79.9	82.0	90.3	86.0
C105	37.01	23.25	18.96	6.57	3.72	4.46	4.62	1.42	81.1	92.1	82.4	91.7	95.0	79.9	82.0	90.3	85.0
C106	36.72	27.22	18.40	5.71	2.41	3.94	4.07	1.54	81.1	90.4	82.4	91.8	95.0	79.9	82.0	90.3	85.0
C107	36.63	24.27	20.18	10.21	3.68	3.53	0.00	1.49	81.1	90.4	82.4	91.8	95.0	79.9	82.0	90.3	85.0
C108	35.04	23.25	19.12	6.10	5.01	5.31	4.84	1.34	81.1	90.4	82.4	91.8	95.0	79.9	82.0	90.3	85.0
C109	34.18	31.70	18.82	2.22	2.90	3.40	4.96	1.83	81.1	90.4	82.4	91.8	95.0	79.9	82.0	90.3	85.0
C110	35.87	27.99	16.37	5.82	2.52	4.73	5.07	1.62	81.0	90.4	82.4	91.8	95.0	79.9	82.0	90.3	85.0
C111	35.10	28.45	18.52	3.72	3.76	3.44	5.39	1.63	81.0	90.4	82.4	91.8	95.0	79.9	82.0	90.3	85.0
C112	33.66	30.67	17.41	4.52	3.24	3.23	5.56	1.71	81.0	90.4	82.4	91.8	95.0	79.9	82.0	90.3	85.0
C113	34.08	28.23	17.11	6.31	3.08	4.84	4.86	1.49	80.9	90.4	82.4	91.8	95.0	79.9	82.0	90.3	85.0
C114	35.60	28.94	18.35	3.61	3.70	2.92	5.27	1.61	80.9	90.4	82.4	91.8	95.0	79.9	82.0	90.3	85.0
C115	34.68	28.53	17.58	5.62	3.29	4.27	4.49	1.53	80.9	90.5	82.5	91.7	95.0	80.0	82.0	90.3	85.0
C116	31.44	31.86	0.50	10.39	9.73	7.48	6.88	1.73	80.9	90.5	82.5	91.7	95.0	80.0	82.0	90.3	86.0
C117	34.84	25.16	18.78	6.61	4.50	3.40	5.33	1.37	80.9	90.5	82.5	91.7	95.0	80.0	82.0	90.3	85.0
C118	35.84	22.19	23.91	6.03	3.84	2.28	4.73	1.19	80.9	90.5	82.5	91.7	95.0	80.0	82.0	90.3	85.0
C119	39.54	27.37	17.79	4.14	4.64	0.00	5.06	1.47	80.9	90.5	82.5	91.7	95.0	80.0	82.0	90.3	85.0
C120	38.96	27.80	16.89	5.07	4.72	0.00	5.04	1.53	80.9	90.5	82.5	91.7	95.0	80.0	82.0	90.3	85.0
C121	32.00	32.30	17.84	5.27	1.59	3.83	5.43	1.74	80.9	90.5	82.5	91.7	95.0	80.0	82.0	90.3	85.0
C122	41.13	26.94	15.56	5.20	4.86	0.00	4.82	1.48	80.9	90.5	82.5	91.7	95.0	80.0	82.0	90.3	85.0
C123	36.91	31.49	19.24	1.80	3.87	0.00	5.06	1.64	80.9	90.5	82.5	91.7	95.0	80.0	82.0	90.3	85.0
C124	39.60	21.40	17.80	15.99	4.26	0.00	0.00	0.96	80.9	90.7	79.9	92.0	95.0	79.7	82.0	90.3	84.7
C125	35.94	37.60	11.97	0.51	12.25	0.00	0.00	1.73	80.9	90.7	79.9	92.0	95.0	79.7	82.0	90.3	86.0
C126	39.19	33.35	15.54	8.04	2.44	0.00	0.00	1.44	80.9	90.7	79.9	92.0	95.0	79.7	82.0	90.3	85.0
C127	36.64	33.60	16.71	6.41	3.01	2.11	0.00	1.51	80.9	90.7	79.9	92.0	95.0	79.7	82.0	90.3	85.0
C128	34.63	33.39	17.83	7.10	3.12	2.39	0.00	1.54	80.9	90.1	79.9	91.5	95.0	79.9	82.0	90.3	85.0
C129	34.87	34.52	16.19	8.15	2.65	2.15	0.00	1.46	80.9	90.1	79.9	91.5	95.0	79.9	82.0	90.3	85.0
C130	35.83	30.95	14.93	8.92	3.33	4.94	0.00	1.11	80.9	90.1	79.9	91.5	95.0	79.9	82.0	90.3	85.1
C131	33.74	33.31	16.11	7.85	2.87	5.02	0.00	1.10	80.9	90.1	79.9	91.5	95.0	79.9	82.0	90.3	85.0
C132	33.58	37.32	7.36	7.28	5.12	0.00	5.12	1.03	80.9	90.6	79.7	91.5	95.0	80.0	82.0	90.3	85.0
C133	38.05	33.41	10.59	6.70	2.79	2.80	0.00	1.12	80.9	90.6	79.7	91.5	95.0	80.0	82.0	90.3	85.0
C134	35.86	34.49	14.25	4.56	3.76	2.76	0.00	1.34	80.9	90.6	79.7	91.5	95.0	80.0	82.0	90.3	85.0
C135	37.17	29.78	14.19	7.00	3.93	4.56	0.00	0.00	80.9	90.9	82.1	92.0	95.0	79.7	82.0	90.3	85.0
C136	37.06	30.04	14.08	7.84	3.83	4.54	0.00	1.40	80.9	90.9	82.1	92.0	95.0	79.7	82.0	90.3	85.0
C137	37.25	29.91	14.66	7.60	3.75	4.48	0.00	1.42	80.9	91.1	81.7	91.6	95.0	80.0	82.0	90.3	85.0
C138	35.48	31.35	16.43	8.09	3.57	3.64	0.00	1.44	81.1	89.4	81.7	91.8	95.0	80.1	82.0	90.3	85.0
C139	36.16	33.39	15.28	7.53	3.08	2.84	0.00	1.73	81.1	89.4	81.8	91.6	95.0	80.3	82.0	90.3	85.0

Πίνακας 3 Δεδομένα ογκομετρικών συστάσεων και Α.Ο. MON των συστατικών των εργαστηριακά παρασκευασμένων μιγμάτων βενζινών

α/α	FCC	REF	ISO	ALK	MTBE	DIM	RON _{FCC}	RON _{REF}	RON _{ISO}	RON _{ALK}	RON _{MTBE}	RON _{DIM}	MON _{FCC}	MON _{REF}	MON _{ISO}	MON _{ALK}	MON _{MTBE}	MON _{DIM}	RON	MON
m01	80.00	20.00	0.00	0.00	0.00	0.00	93.7	102.0	84.0	94.1	115.0	93.7	81.7	90.4	82.1	92.1	95.0	80.0	95.9	83.2
m02	65.00	35.00	0.00	0.00	0.00	0.00	93.7	102.0	84.0	94.1	115.0	93.7	81.7	90.4	82.1	92.1	95.0	80.0	97.6	84.0
m03	40.00	35.00	25.00	0.00	0.00	0.00	93.7	102.0	84.0	94.1	115.0	93.7	81.7	90.4	82.1	92.1	95.0	80.0	93.8	84.9
m04	55.00	20.00	25.00	0.00	0.00	0.00	93.7	102.0	84.0	94.1	115.0	93.7	81.7	90.4	82.1	92.1	95.0	80.0	93.6	83.2
m05	59.00	41.00	0.00	0.00	0.00	0.00	93.7	102.0	84.0	94.1	115.0	93.7	81.7	90.4	82.1	92.1	95.0	80.0	97.0	84.9
m06	74.00	26.00	0.00	0.00	0.00	0.00	93.7	102.0	84.0	94.1	115.0	93.7	81.7	90.4	82.1	92.1	95.0	80.0	96.0	84.1
m07	34.00	41.00	25.00	0.00	0.00	0.00	93.7	102.0	84.0	94.1	115.0	93.7	81.7	90.4	82.1	92.1	95.0	80.0	95.4	84.9
m08	49.00	26.00	25.00	0.00	0.00	0.00	93.7	102.0	84.0	94.1	115.0	93.7	81.7	90.4	82.1	92.1	95.0	80.0	94.2	83.8
m09	47.00	35.00	0.00	18.00	0.00	0.00	93.7	102.0	84.0	94.1	115.0	93.7	81.7	90.4	82.1	92.1	95.0	80.0	97.1	85.9
m10	62.00	20.00	0.00	18.00	0.00	0.00	93.7	102.0	84.0	94.1	115.0	93.7	81.7	90.4	82.1	92.1	95.0	80.0	95.7	84.3
m11	22.00	35.00	25.00	18.00	0.00	0.00	93.7	102.0	84.0	94.1	115.0	93.7	81.7	90.4	82.1	92.1	95.0	80.0	95.6	86.1
m12	37.00	20.00	25.00	18.00	0.00	0.00	93.7	102.0	84.0	94.1	115.0	93.7	81.7	90.4	82.1	92.1	95.0	80.0	94.2	84.9
m13	41.00	41.00	0.00	18.00	0.00	0.00	93.7	102.0	84.0	94.1	115.0	93.7	81.7	90.4	82.1	92.1	95.0	80.0	98.2	86.6
m14	56.00	26.00	0.00	18.00	0.00	0.00	93.7	102.0	84.0	94.1	115.0	93.7	81.7	90.4	82.1	92.1	95.0	80.0	96.8	85.1
m15	16.00	41.00	25.00	18.00	0.00	0.00	93.7	102.0	84.0	94.1	115.0	93.7	81.7	90.4	82.1	92.1	95.0	80.0	97.0	86.9
m16	31.00	26.00	25.00	18.00	0.00	0.00	93.7	102.0	84.0	94.1	115.0	93.7	81.7	90.4	82.1	92.1	95.0	80.0	94.8	85.1
m17	55.00	35.00	0.00	0.00	0.00	10.00	93.7	102.0	84.0	94.1	115.0	93.7	81.7	90.4	82.1	92.1	95.0	80.0	97.7	83.5
m18	70.00	20.00	0.00	0.00	0.00	10.00	93.7	102.0	84.0	94.1	115.0	93.7	81.7	90.4	82.1	92.1	95.0	80.0	96.2	82.4
m19	30.00	35.00	25.00	0.00	0.00	10.00	93.7	102.0	84.0	94.1	115.0	93.7	81.7	90.4	82.1	92.1	95.0	80.0	96.1	84.1
m20	45.00	20.00	25.00	0.00	0.00	10.00	93.7	102.0	84.0	94.1	115.0	93.7	81.7	90.4	82.1	92.1	95.0	80.0	94.2	83.4
m21	49.00	41.00	0.00	0.00	0.00	10.00	93.7	102.0	84.0	94.1	115.0	93.7	81.7	90.4	82.1	92.1	95.0	80.0	98.0	84.9
m22	64.00	26.00	0.00	0.00	0.00	10.00	93.7	102.0	84.0	94.1	115.0	93.7	81.7	90.4	82.1	92.1	95.0	80.0	97.1	83.1
n01	35.00	30.00	14.00	10.00	6.00	5.00	93.6	102.5	84.0	94.0	115.0	94.5	81.0	90.9	82.1	92.0	95.0	79.7	97.5	88.5
n02	47.00	23.00	25.00	5.00	0.00	0.00	93.6	102.5	84.0	94.0	115.0	94.5	81.0	90.9	82.1	92.0	95.0	79.7	94.5	85.5
n03	34.00	31.00	14.00	10.00	6.00	5.00	93.6	102.5	84.0	94.0	115.0	94.5	81.0	90.9	82.1	92.0	95.0	79.7	97.1	86.6
n04	45.00	35.00	20.00	0.00	0.00	0.00	93.6	102.5	84.0	94.0	115.0	94.5	81.0	90.9	82.1	92.0	95.0	79.7	95.1	84.0
n05	48.00	38.00	0.00	14.00	0.00	0.00	93.6	102.5	84.0	94.0	115.0	94.5	81.0	90.9	82.1	92.0	95.0	79.7	97.5	85.7
n06	48.00	27.00	25.00	0.00	0.00	0.00	93.6	102.5	84.0	94.0	115.0	94.5	81.0	90.9	82.1	92.0	95.0	79.7	93.7	83.4
n07	45.00	30.00	25.00	0.00	0.00	0.00	93.6	102.5	84.0	94.0	115.0	94.5	81.0	90.9	82.1	92.0	95.0	79.7	94.0	83.9
n08	33.00	29.00	17.00	10.00	6.00	5.00	93.6	102.5	84.0	94.0	115.0	94.5	81.0	90.9	82.1	92.0	95.0	79.7	96.0	85.4
n09	45.00	20.00	25.00	0.00	0.00	10.00	93.6	102.5	84.0	94.0	115.0	94.5	81.0	90.9	82.1	92.0	95.0	79.7	93.5	83.0
n10	40.00	35.00	25.00	0.00	0.00	0.00	93.6	102.5	84.0	94.0	115.0	94.5	81.0	90.9	82.1	92.0	95.0	79.7	94.4	84.2
n11	37.00	38.00	25.00	0.00	0.00	0.00	93.6	102.5	84.0	94.0	115.0	94.5	81.0	90.9	82.1	92.0	95.0	79.7	95.0	83.7
n12	45.00	20.00	25.00	0.00	10.00	0.00	93.6	102.5	84.0	94.0	115.0	94.5	81.0	90.9	82.1	92.0	95.0	79.7	94.9	84.4
n13	48.00	38.00	14.00	0.00	0.00	0.00	93.6	102.5	84.0	94.0	115.0	94.5	81.0	90.9	82.1	92.0	95.0	79.7	95.1	84.7
n14	48.00	23.00	25.00	0.00	0.00	4.00	93.6	102.5	84.0	94.0	115.0	94.5	81.0	90.9	82.1	92.0	95.0	79.7	93.4	83.7
n15	45.00	20.00	25.00	10.00	0.00	0.00	93.6	102.5	84.0	94.0	115.0	94.5	81.0	90.9	82.1	92.0	95.0	79.7	93.2	84.0
p01	32.30	34.60	22.90	6.20	1.20	2.90	93.4	101.7	84.2	93.9	115.0	95.0	80.6	90.3	82.0	91.8	95.0	79.8	95.7	85.0
p02	27.90	30.50	27.70	6.30	2.80	4.90	93.4	101.7	84.2	93.9	115.0	95.0	80.6	90.3	82.0	91.8	95.0	79.8	95.4	84.7
q01	76.70	7.40	9.40	2.00	0.90	3.60	93.4	100.8	84.1	93.7	115.0	94.3	80.3	90.1	82.0	91.7	95.0	80.1	94.1	80.6
q02	31.80	27.50	23.70	7.70	4.50	4.80	93.4	100.8	84.1	93.7	115.0	94.3	80.3	90.1	82.0	91.7	95.0	80.1	95.2	84.6
r01	33.00	35.10	16.50	7.40	3.50	4.50	93.2	100.8	84.7	94.0	115.0	95.1	80.2	89.3	82.3	91.1	95.0	79.7	96.0	85.0
r02	37.50	35.10	16.50	7.40	3.50	0.00	93.2	100.8	84.7	94.0	115.0	95.1	80.2	89.3	82.3	91.1	95.0	79.7	95.8	85.0