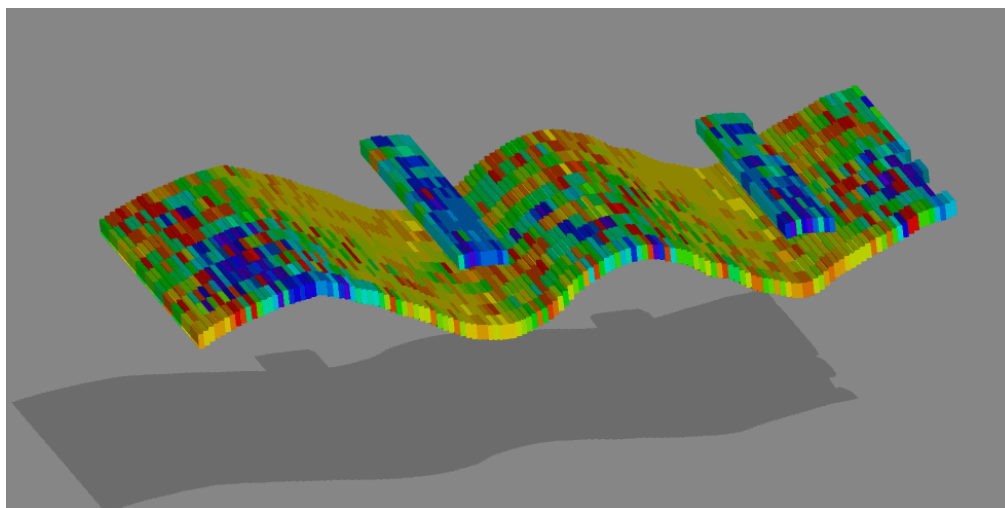


**ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΟΡΥΚΤΩΝ ΠΟΡΩΝ
ΕΡΓΑΣΤΗΡΙΟ ΕΦΑΡΜΟΣΜΕΝΗΣ ΓΕΩΦΥΣΙΚΗΣ**

**ΣΥΜΒΟΛΗ ΤΩΝ ΜΕΘΟΔΩΝ ΤΑΞΙΝΟΜΗΣΗΣ
ΔΕΔΟΜΕΝΩΝ ΓΕΩΡΑΝΤΑΡ ΣΤΗΝ ΕΠΙΛΥΣΗ
ΠΕΡΙΒΑΛΛΟΝΤΙΚΩΝ ΠΡΟΒΛΗΜΑΤΩΝ**

**ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ
ΣΤΟ ΠΛΑΙΣΙΟ ΤΟΥ ΠΡΟΓΡΑΜΜΑΤΟΣ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
«ΠΕΡΙΒΑΛΛΟΝΤΙΚΗ ΓΕΩΤΕΧΝΟΛΟΓΙΑ»**



ΣΠΑΝΟΥΔΑΚΗΣ ΝΙΚΟΛΑΟΣ

ΜΗΧΑΝΙΚΟΣ ΟΡΥΚΤΩΝ ΠΟΡΩΝ, ΜΔΕ

ΧΑΝΙΑ 2007

ΤΡΙΜΕΛΗΣ ΣΥΜΒΟΥΛΕΥΤΙΚΗ ΕΠΙΤΡΟΠΗ

Βαφείδης Αντώνιος (Καθηγητής, Τμήμα Μηχανικών Ορυκτών Πόρων, Πολυτεχνείο Κρήτης, Επιβλέπων)

Τσόκας Γρηγόριος (Καθηγητής, Τμήμα Γεωλογίας, ΑΠΘ, Συμβουλευτική Επιτροπή)

Τσούρλος Παναγιώτης (Επίκουρος Καθηγητής, Τμήμα Γεωλογίας, ΑΠΘ, Συμβουλευτική Επιτροπή)

ΕΠΤΑΜΕΛΗΣ ΕΞΕΤΑΣΤΙΚΗ ΕΠΙΤΡΟΠΗ

Βαφείδης Αντώνιος (Καθηγητής, Τμήμα Μηχανικών Ορυκτών Πόρων, Πολυτεχνείο Κρήτης, Επιβλέπων)

Τσόκας Γρηγόριος (Καθηγητής, Τμήμα Γεωλογίας, ΑΠΘ, Συμβουλευτική Επιτροπή)

Τσούρλος Παναγιώτης (Επίκουρος Καθηγητής, Τμήμα Γεωλογίας, ΑΠΘ, Συμβουλευτική Επιτροπή)

Γαλετάκης Μιχαήλ (Επίκουρος Καθηγητής, Τμήμα Μηχανικών Ορυκτών Πόρων, Πολυτεχνείο Κρήτης)

Λούης Ιωάννης (Καθηγητής, Τμήμα Γεωλογίας, Εθνικό Καποδίστριο Πανεπιστήμιο)

Πασαδάκης Νικόλαος (Επίκουρος Καθηγητής, Τμήμα Μηχανικών Ορυκτών Πόρων, Πολυτεχνείο Κρήτης)

Χριστόπουλος Διονύσιος (Καθηγητής, Τμήμα Μηχανικών Ορυκτών Πόρων, Πολυτεχνείο Κρήτης)

ΕΥΧΑΡΙΣΤΙΕΣ

Στο σημείο αυτό θα ήθελα να ευχαριστήσω όλους εκείνους που με τον τρόπο τους συνέβαλλαν στο να ολοκληρώσω τη διδακτορική μου διατριβή.

Τον Καθηγητή κύριο Αντώνη Βαφείδη που με εμπιστεύτηκε τα επτά χρόνια της συνεργασίας μας αποδεχόμενος να συμμετέχει ως επιβλέπων στη διπλωματική, την μεταπτυχιακή και την διδακτορική μου διατριβή. Επίσης που με τίμησε αναθέτοντας μου δραστηριότητες σε ερευνητικά προγράμματα, συμβάλλοντας με αυτό τον τρόπο ηθικά, επιστημονικά αλλά και οικονομικά ώστε να ολοκληρώσω βήμα βήμα όλες μου τις προσπάθειες.

Θα ήθελα επίσης να ευχαριστήσω θερμά τα μέλη της συμβουλευτικής επιτροπής.

Τον Καθηγητή κύριο Γρηγόρη Τσόκα ο οποίος παρακολουθεί την πορεία μου εδώ και χρόνια συμμετέχοντας στην συμβουλευτική επιτροπή της διπλωματικής αλλά και της μεταπτυχιακής μου εργασίας.

Τον Επίκουρο Καθηγητή κύριο Παναγιώτη Τσούρλο για το ενδιαφέρον που έδειξε για την εργασία μου.

Τον Καθηγητή κύριο Ιωάννη Λούη ο οποίος επίσης συμμετείχε ως μέλος της συμβουλευτικής επιτροπής και στη μεταπτυχιακή μου εργασία.

Τον Καθηγητή κύριο Διονύση Χριστόπουλο για ενδιαφέρον που έχει δείξει καθ' όλη τη διάρκεια της εκπόνησης της διατριβής μου.

Τον Επίκουρο Καθηγητή κύριο Νικόλαο Πασαδάκη για τις εποικοδομητικές συμβουλές του καθώς και για τη συνεργασία του.

Τον Επίκουρο Καθηγητή κύριο Μιχαήλ Γαλετάκη για τη συνεργασία του.

Τον Καθηγητή κύριο Machsım Bano ο οποίος διέθεσε τα δεδομένα του ελεγχόμενου πειράματος.

Τους φίλους και συνεργάτες Hamdan Hamdan, Νίκο Ανδρονικίδη, Γιώργο Κρητικάκη, Νίκο Οικονόμου, Αντώνη Αντωνίου, Μαρία Κουκαδάκη και Αγγελική Παπαβασιλείου.

Τον Θανάση Παπαδόπουλο και τον Ξενοφόντα Φραντζή για την υποστήριξη και τις επιστημονικές συμβουλές.

Τους γονείς μου και τον αδερφό μου οι οποίοι παρακολουθούν διακριτικά και στηρίζουν τις προσπάθειες μου.

Ιδιαίτερα ευχαριστώ τη φίλη μου Βανέσσα Μπάρμπα για την ηθική υποστήριξη την υπομονή που έκανε όλα αυτά τα χρόνια καθώς και για τις επιστημονικές και άλλες συμβουλές της.

Η εργασία αυτή είναι αφιερωμένη στη μνήμη του φίλου μου Πάνου Δήμιζα.

ΠΕΡΙΛΗΨΗ

Στην παρούσα διατριβή αναπτύσσεται μεθοδολογία για την επεξεργασία και ερμηνεία δεδομένων γεωραντάρ. Αυτή η μεθοδολογία είναι εμπνευσμένη από τη μέθοδο σεισμικής ανάκλασης και περιλαμβάνει τρία διακριτά στάδια, καθένα από οποία μπορεί να αποτελέσει αυτόνομο στάδιο χρήσιμο για την ερμηνεία των δεδομένων. Το πρώτο στάδιο αφορά στην εφαρμογή φίλτρων απαλοιφής του θορύβου, χωροθέτησης και ενίσχυσης του σήματος. Το δεύτερο στάδιο αφορά στον υπολογισμό στιγμιαίων και γεωμετρικών χαρακτηριστικών των δεδομένων. Η τεχνική αυτή άρχισε να αναπτύσσεται τις τελευταίες δεκαετίες και εξελίσσεται συνεχώς. Το τρίτο στάδιο αφορά στην ταξινόμηση των χαρακτηριστικών του σήματος με τη χρήση μεθόδων από τα πεδία της πολυδιάστατης στατιστικής, της ασαφούς λογικής, των νευρωνικών δικτύων και των γενετικών αλγόριθμων. Ο στόχος είναι να εντοπιστούν καταγραφές με παρόμοια χαρακτηριστικά. Οι παραπάνω τεχνικές εφαρμόζονται σε δεδομένα από ελεγχόμενο πείραμα που συλλέχθηκαν στο πλαίσιο του ερευνητικού προγράμματος HYGEIA (HYbrid Geophysical technology for the Evaluation of Insidious contaminated Areas) καθώς και δεδομένα από μετρήσεις στην αρχαία πόλη Άπτερα στο νομό Χανίων. Προτείνονται μεθοδολογίες, οι οποίες αφορούν στην ταξινόμηση δεδομένων που κατανέμονται σε κατακόρυφη τομή γεωραντάρ, σε οριζόντια τομή ή προέρχονται από ανακλαστήρα. Η μεθοδολογία αποδείχτηκε αποτελεσματική όσον αφορά στον εντοπισμό ρηχών στόχων αλλά και στην ερμηνεία δεδομένων γεωραντάρ.

ABSTRACT

In this PhD thesis, a Ground Penetrating Radar (GPR) data classification workflow is proposed. This methodology is inspired by seismic exploration and includes three discrete stages. Each stage can be considered as a distinct step for GPR data interpretation. The first stage includes filtering, gain and migration. The second includes the calculation of instantaneous and geometrical attributes, a technique employed the past few decades, for the qualitative and quantitative interpretation of seismic data. The third stage includes attributes classification using multivariate statistics, neural networks and fuzzy logic. The main purpose of attributes classification is to organize data with similar multi-attribute response. The above mentioned methodology is applied on a dataset acquired over a controlled experiment, under the EU program HYGEIA (HYbrid Geophysical technology for the Evaluation of Insidious contaminated Areas) for buried pipe detection and a dataset acquired in the ancient city of Aptera in Chania, Crete, Greece. The proposed methodology classifies GPR sections, slices and facies and is useful for GPR data interpretation and the detection of near surface buried targets.

ΠΕΡΙΕΧΟΜΕΝΑ

ΕΥΧΑΡΙΣΤΙΕΣ	iii
ΠΕΡΙΛΗΨΗ	v
ABSTRACT	vi
ΠΕΡΙΕΧΟΜΕΝΑ	vii
ΚΕΦΑΛΑΙΟ 1.	1
ΕΙΣΑΓΩΓΗ	1
1.1 Εισαγωγή.	1
1.2 Δομή της εργασίας.	2
1.3. Η συνεισφορά του γεωραντάρ στην επίλυση περιβαλλοντικών προβλημάτων.	3
1.4 Χαρακτηριστικά των καταγραφών του γεωραντάρ.	4
1.5 Ιστορική εξέλιξη και εφαρμογές των στιγμιαίων χαρακτηριστικών.	5
1.6 Νευρωνικά δίκτυα και ασαφής λογική: βιβλιογραφική έρευνα.	7
ΚΕΦΑΛΑΙΟ 2.	18
ΑΡΧΗ ΛΕΙΤΟΥΡΓΙΑΣ ΚΑΙ ΕΠΕΞΕΡΓΑΣΙΑ ΔΕΔΟΜΕΝΩΝ ΓΕΩΡΑΝΤΑΡ	18
2.1 Το γεωραντάρ.	18
2.2 Εφαρμογές του γεωραντάρ.	18
2.3 Βασικές αρχές.	19
2.4 Διατάξεις πομπού - δέκτη.	20
2.5 Επεξεργασία των δεδομένων του γεωραντάρ.	21
2.6 Απεικόνιση των δεδομένων.	23
2.7 Ταξινόμηση των δεδομένων του γεωραντάρ: Διάγραμμα ροής.	24
2.8 Λογισμικό επεξεργασίας δεδομένων γεωραντάρ.....	26
ΚΕΦΑΛΑΙΟ 3	28
ΜΕΘΟΔΟΙ ΤΑΞΙΝΟΜΗΣΗΣ ΔΕΔΟΜΕΝΩΝ ΓΕΩΡΑΝΤΑΡ	28
3.1 Ταξινόμηση των δεδομένων.	28
3.2 Νευρωνικά δίκτυα: βασικές αρχές.	30
3.3 Μέθοδοι αυτόματης ταξινόμησης.	30
3.3.1 Αλγόριθμος κ-μέσων τιμών.	31
3.3.2 Ανταγωνιστική μάθηση και κανόνας της συνείδησης.	32
3.3.3 Δίκτυα αυτο-οργανωμένης απεικόνισης.	33
3.3.4 Ασαφής αφαιρετική ταξινόμηση.	34
3.3.5 Ασαφής ταξινόμηση c-μέσων τιμών.	35
3.4 Μέθοδοι καθοδηγούμενης ταξινόμησης.	35
3.4.1 Δίκτυα διανυσματικής μάθησης.	35

3.4.2 Νευρωνικά δίκτυα πιθανότητας.....	36
3.5 Σύνοψη των μεθόδων ταξινόμησης.....	36
ΚΕΦΑΛΑΙΟ 4	39
ΕΡΜΗΝΕΙΑ ΕΡΓΑΣΤΗΡΙΑΚΩΝ ΔΕΔΟΜΕΝΩΝ ΓΕΩΡΑΝΤΑΡ.....	39
4.1 Εισαγωγή.....	39
4.2 Περιγραφή του ελεγχόμενου πειράματος.....	40
4.3 Επεξεργασία των δεδομένων.....	43
4.4 Αυτόματη ταξινόμηση των δεδομένων.....	54
4.5 Καθοδηγούμενη ταξινόμηση των δεδομένων με το δίκτυο LVQ.....	56
4.6 Αυτόματη ταξινόμηση ανακλαστήρων.....	65
4.7 Συμπεράσματα.....	65
ΚΕΦΑΛΑΙΟ 5	70
ΤΑΞΙΝΟΜΗΣΗ ΨΕΥΔΟ-ΤΡΙΣΔΙΑΣΤΑΤΩΝ ΔΕΔΟΜΕΝΩΝ ΓΕΩΡΑΝΤΑΡ	70
5.1 Εισαγωγή.....	70
5.2 Επεξεργασία των δεδομένων.....	70
5.3 Ερμηνεία των οριζόντιων τομών.....	73
5.4 Ανάλυση κυρίων συνιστωσών.....	75
5.5 Ταξινόμηση των δεδομένων.....	80
5.5.1 Αυτόματη ταξινόμηση.....	80
5.5.2 Καθοδηγούμενη ταξινόμηση οριζόντιας τομής.....	82
5.6 Σύγκριση των μεθόδων ταξινόμησης κ-μέσων τιμών και δικτύων αυτο- οργανωμένης απεικόνισης.....	83
5.7 Συμπεράσματα.....	92
ΚΕΦΑΛΑΙΟ 6	93
ΤΑΞΙΝΟΜΗΣΗ ΔΕΔΟΜΕΝΩΝ ΓΕΩΡΑΝΤΑΡ ΑΠΟ ΤΗΝ ΑΡΧΑΙΑ ΑΠΤΕΡΑ.....	93
6.1 Εισαγωγή.....	93
6.2 Ιστορικά στοιχεία για την αρχαία πόλη Άπτερα.....	94
6.3 Συλλογή των δεδομένων.....	95
6.4 Επεξεργασία των δεδομένων.....	98
6.5 Αυτόματη ταξινόμηση ανακλαστήρα.....	101
6.6 Αυτόματη ταξινόμηση των ψευδο-τρισδιάστατων δεδομένων.....	111
6.7 Αυτόματη ταξινόμηση των ανακλαστήρων.....	119
6.8 Συμπεράσματα.....	119
ΚΕΦΑΛΑΙΟ 7	121
ΤΑΞΙΝΟΜΗΣΗ ΔΕΔΟΜΕΝΩΝ ΓΕΩΡΑΝΤΑΡ ΑΠΟ ΤΟ ΟΡΟΠΕΔΙΟ ΤΟΥ ΟΜΑΛΟΥ	121
7.1 Εισαγωγή.....	121
7.2 Περιοχή Ομαλού.....	122
7.3 Σχεδιασμός διασκόπησης στο οροπέδιο του Ομαλού.....	125
7.4 Επεξεργασία των δεδομένων.....	125

7.5 Συνδυασμένη ερμηνεία γεωφυσικών δεδομένων.	126
7.6 Ταξινόμηση των δεδομένων	126
 ΚΕΦΑΛΑΙΟ 8	 131
ΣΥΜΠΕΡΑΣΜΑΤΑ - ΠΡΟΤΑΣΕΙΣ	131
8.1 Εισαγωγή.	131
8.2 Πρωτοτυπία της εργασίας.	132
8.3 Συμπεράσματα.	132
8.4 Μελλοντικοί στόχοι.	134
 ΒΙΒΛΙΟΓΡΑΦΙΑ.	 136
ΠΑΡΑΡΤΗΜΑ.....	146
Π.1 Ανάλυση κυρίων συνιστωσών.	146
Π.2 Μέθοδος των κ-μέσων τιμών.	148
Π.3 Δίκτυα αυτο-οργανωμένης απεικόνισης.	149
Π.3.1 Ανταγωνιστική διαδικασία.....	150
Π.3.2 Διαδικασία συνεργασίας.	151
Π.3.3 Διαδικασία προσαρμογής.....	152
Π.3.4 Σύνοψη των δικτύων αυτο-οργανωμένης απεικόνισης.....	155

ΚΕΦΑΛΑΙΟ 1.**ΕΙΣΑΓΩΓΗ****1.1 Εισαγωγή.**

Τα σεισμικά χαρακτηριστικά (*seismic attributes*) εισήχθησαν στη γεωφυσική τη δεκαετία του 1970, σαν τυπικά εργαλεία απεικόνισης, συνεισφέροντας στην ποιοτική ερμηνεία των σεισμικών δεδομένων. Αυτή η προσέγγιση όμως απαιτούσε τη χρονοβόρα μελέτη δεκάδων σεισμικών χαρακτηριστικών, με στόχο τον εντοπισμό ενδείξεων όπως καταγραφές χαμηλής συχνότητας κάτω από τους ταμιευτήρες, αντιστροφή της φάσης και της πολικότητας του σεισμικού σήματος κτλ, που σχετίζονται με την ύπαρξη υδρογονανθράκων. Τέτοιες ενδείξεις πολλές φορές οδηγούσαν σε εσφαλμένα συμπεράσματα σχετικά με τη θέση των παραγωγικών γεωτρήσεων.

Οι μέθοδοι ταξινόμησης, κυρίως με νευρωνικά δίκτυα, εισήχθησαν στην έρευνα με τη μέθοδο της σεισμικής ανάκλασης την προηγούμενη δεκαετία παρέχοντας ισχυρά και σύγχρονα εργαλεία για τη μελέτη των ταμιευτήρων πετρελαίου. Η διαδικασία αξιολόγησης των σεισμικών χαρακτηριστικών αυτοματοποιήθηκε, χρησιμοποιώντας επιπλέον δεδομένα διαγραφιών για τη βαθμονόμηση της ταξινόμησης.

Οι ομοιότητες στις κινηματικές ιδιότητες των σεισμικών δεδομένων και των δεδομένων του γεωραντάρ συνεπάγεται ότι οι παραπάνω τεχνικές μπορούν να εφαρμοστούν και στα δεδομένα του γεωραντάρ συνεισφέροντας στην ποιοτική τους ερμηνεία.

Τα δεδομένα του γεωραντάρ διαφέρουν από τα σεισμικά ως προς τα μετρούμενα μεγέθη. Ανακλώμενα ηλεκτρομαγνητικά κύματα προέρχονται από διαχωριστικές επιφάνειες μεταξύ γεωλογικών σχηματισμών, ρωγμές ή οποιαδήποτε διεπιφάνεια, η οποία χωρίζει πετρώματα με διαφορετικές ηλεκτρικές ιδιότητες. Οι ηλεκτρικές ιδιότητες των πετρωμάτων εξαρτώνται κυρίως από τη λιθολογική τους σύσταση και το βαθμό κορεσμού τους σε νερό. Μετά από κατάλληλη επεξεργασία των δεδομένων του γεωραντάρ (όπως αφαίρεση των απευθείας κυμάτων αέρα και κυμάτων εδάφους), αυτό που μένει μπορεί να συνδεθεί με τις ηλεκτρικές ιδιότητες του υπεδάφους (Lemke και Mankowski, 2000). Αυτές οι ιδιότητες επηρεάζονται επίσης και από άλλες πετροφυσικές παραμέτρους, όπως το πορώδες.

1.2 Δομή της εργασίας.

Το πρώτο κεφάλαιο αυτής της εργασίας αποτελεί μια εισαγωγή στο αντικείμενο της ταξινόμησης γεωφυσικών δεδομένων. Περιλαμβάνει εκτενή βιβλιογραφική έρευνα η οποία προέρχεται από την έρευνα με σεισμικές μεθόδους και αφορά στα χαρακτηριστικά των δεδομένων, καθώς και ιστορική ανασκόπηση για την ανάπτυξη των νευρωνικών δικτύων και της ασαφούς λογικής. Στο δεύτερο κεφάλαιο αναφέρεται η αρχή λειτουργίας του γεωραντάρ, όπως επίσης βασικές και προηγμένες τεχνικές επεξεργασίας και απεικόνισης των δεδομένων που συμβάλλουν στην επιτυχή ερμηνεία τους. Στη συνέχεια γίνεται εισαγωγή στη μεθοδολογία με την παρουσίαση διαγραμμάτων ροής και λογισμικού προγράμματος το οποίο κατασκευάστηκε για τις ανάγκες της διατριβής. Στο τρίτο κεφάλαιο αναφέρονται οι μέθοδοι ταξινόμησης που περιλαμβάνονται στο λογισμικό. Οι μέθοδοι αφορούν σε τεχνικές καθοδηγούμενης και αυτόματης ταξινόμησης. Στην τελευταία παράγραφο του κεφαλαίου γίνεται συγκριτική αποτίμηση των μεθόδων όπου και προτείνεται μεθοδολογία που πρόκειται να εφαρμοστεί στα επόμενα κεφάλαια. Στο τέταρτο κεφάλαιο προτείνεται μεθοδολογία για την ταξινόμηση

τομής γεωραντάρ με τη χρήση των μεθόδων των κυρίων συνιστωσών, των κ-μέσων τιμών και των δικτύων αυτό-οργανωμένης απεικόνισης. Παράλληλα, γίνεται διεξοδική παρουσίαση του λογισμικού επεξεργασίας GPR-Pro. Στο πέμπτο κεφάλαιο παρουσιάζεται η μεθοδολογία ταξινόμησης ψευδο-τρισδιάστατων δεδομένων γεωραντάρ, αλλά και οριζόντιων τομών. Το κεφάλαιο αυτό κλείνει με τη σύγκριση των μεθόδων ταξινόμησης κ-μέσων τιμών και του δικτύου αυτο-οργανωμένης απεικόνισης. Στο έκτο κεφάλαιο παρουσιάζεται ένας συστηματικός τρόπος επεξεργασίας και ταξινόμησης δεδομένων γεωραντάρ που συλλέχθηκαν στον αρχαιολογικό χώρο Άπτερα στο νομό Χανίων, ενώ στο έκτο κεφάλαιο παρουσιάζονται τα αποτελέσματα της επεξεργασίας δεδομένων γεωραντάρ που συλλέχθηκαν στο οροπέδιο του Ομαλού στο νομό χανίων για την περιχαράκωση καρστικών εγκοίλων. Στο τελευταίο κεφάλαιο γίνεται αποτίμηση της συνολικής δουλειάς, όπου προκύπτουν συμπεράσματα σχετικά με τις μεθόδους ταξινόμησης, αναφέρεται η πρωτοτυπία της διατριβής και προτείνονται τα αντικείμενα που χρίζουν περαιτέρω διερεύνηση.

1.3. Η συνεισφορά του γεωραντάρ στην επίλυση περιβαλλοντικών προβλημάτων.

Αναφέρονται συνοπτικά περιπτώσεις χρήσης του γεωραντάρ για την μελέτη και την επίλυση περιβαλλοντικών προβλημάτων.

Ο Brewster και οι συνεργάτες του (1995) πραγματοποίησαν ελεγχόμενο πείραμα στην канаδική βάση ενόπλων δυνάμεων κοντά στο Τορόντο. Ελευθέρωσαν 770lt τετραχλωροαιθυλενίου (PCE) εντός απομονωμένης περιοχής και παρακολουθούσαν το το φαινόμενο της κίνησης των ρύπων στο υπέδαφος επί 984 ώρες με τη χρήση του γεωραντάρ αλλά και άλλων γεωφυσικών μεθόδων. Σκοπός του ελεγχόμενου αυτού πειράματος ήταν η ανίχνευση και παρατήρηση ρύπανσης DNAPL κάτω από τον υδροφόρο ορίζοντα, καθώς και η περιγραφή της μετανάστευσης και τελικής κατανομής της ελεγχόμενης ρύπανσης. Τα αποτελέσματα αποδεικνύουν ότι φυσικοί ετερογενείς παράγοντες, ακόμα και σε σχετικά ομοιογενές περιβάλλον, μπορούν να προκαλέσουν μία πλευρική εξάπλωση ρύπανσης DNAPL σε πολύ μεγάλες περιοχές του υπεδάφους.

Ο Daniels και οι συνεργάτες του (1992) πραγματοποίησαν ελεγχόμενο πείραμα μέσα σε ένα ειδικά διαμορφωμένο χώρο από άμμο. Σκοπός του πειράματος ήταν η χρήση του

γεωραντάρ για την ανίχνευση ρύπανσης LNAPL. Επίσης πραγματοποίησαν μετρήσεις σε εργοστάσιο αποθήκευσης βενζίνης στη βόρειο Ιντιάνα (ΗΠΑ). Διαπίστωσαν ότι: α) η ρύπανση (βενζίνη) είναι πιθανόν να περιέχεται σε μικρές, διασκορπισμένες συγκεντρώσεις μέσα στην ακόρεστη ζώνη (*vadose zone*), β) τα προϊόντα των υδρογονανθράκων που δημιουργούν ρύπανση είναι δυνατόν να εμφανίζουν υψηλό βαθμό εξασθένισης. Αυτή η αιτιολόγηση θα μπορούσε να ερμηνεύσει τόσο την παρατηρούμενη μείωση στο πλάτος των ηλεκτρομαγνητικών κυμάτων που διαδίδονται μέσω των τριχοειδών αγγείων, όσο και το υψηλό ποσοστό ανάκλασης που εντοπίζεται στους κορεσμένους υδρογονάνθρακες.

Ο Grumman και οι συνεργάτες του (1995) πραγματοποίησαν πειραματικές μελέτες με το γεωραντάρ σε πρατήρια βενζίνης σε περιοχές κοντά στο Κλήβελαντ του Οχάιο και στη βόρεια Ιντιάνα των ΗΠΑ, τα οποία έκλεισαν εξαιτίας διαρροής υδρογονανθράκων. Στις περιοχές αυτές παρατηρήθηκε βενζίνη με τη μορφή ελεύθερης φάσης, καθώς και ρύπανση με τη μορφή ατμού. Οι έρευνες με χρήση υπεδάφειου ραντάρ είχαν σαν σκοπό την παρατήρηση της ικανότητας του γεωραντάρ να ανιχνεύσει την παρουσία ρύπανσης με τη μορφή ελεύθερης φάσης και φάσης ατμού. Παράλληλα, πραγματοποιήθηκαν ελεγχόμενα πειράματα προκειμένου να αξιολογηθεί η ευαισθησία του γεωραντάρ σε οργανικούς ρυπογόνους παράγοντες κάτω από ελεγχόμενες συνθήκες. Τα αποτελέσματα των πειραματικών και ελεγχόμενων μελετών δείχνουν ότι πραγματικά εμφανίζονται ευδιάκριτες αλλαγές στις μετρήσεις του γεωραντάρ, πάνω από περιοχές που χαρακτηρίζονται από οργανική ρύπανση. Τα πειραματικά δεδομένα έδειξαν ότι οι τρισδιάστατες τομές που προκύπτουν μπορούν να συνεισφέρουν στην ερμηνεία των εδαφολογικών συνθηκών, συμπεριλαμβανομένης και της οργανικής ρύπανσης. Σε αντίθεση, οι αρχικές ενδείξεις από την ελεγχόμενη μελέτη, δείχνουν ότι τα αποτελέσματα της οργανικής ρύπανσης μέσα στη ακόρεστη ζώνη είναι δύσκολο να αποκρυπτογραφηθούν με τη χρήση μόνο του γεωραντάρ.

1.4 Χαρακτηριστικά των καταγραφών του γεωραντάρ.

Υπάρχουν δυο κύριες κατηγορίες χαρακτηριστικών (Taner 1994, Robertson και Fischer 1988): τα στιγμιαία ή φυσικά (*instantaneous or physical attributes*) και τα

γεωμετρικά (*geometrical attributes*). Τα στιγμιαία χαρακτηριστικά αντιστοιχούν σε μεταβολές των φυσικών ιδιοτήτων του υπεδάφους και υπολογίζονται με τον μετασχηματισμό Hilbert. Μπορούν να υπολογιστούν δείγμα προς δείγμα (στιγμιαία), είτε στις κορυφές (*peaks*) του στιγμιαίου πλάτους (απόκρισης). Τα γεωμετρικά χαρακτηριστικά γεωλογικών σχηματισμών προκύπτουν από τη χωροχρονική μεταβολή των στιγμιαίων.

Στα δεδομένα γεωραντάρ, κάθε δείγμα περιγράφεται από μοναδικό συνδυασμό στιγμιαίων και γεωμετρικών χαρακτηριστικών και αντιμετωπίζεται σαν πολυδιάστατο διάνυσμα. Η ταξινόμηση ομαδοποιεί παρόμοια διανύσματα τα οποία απέχουν ορισμένη Ευκλείδεια απόσταση από το προκαθορισμένο κέντρο συσσώρευσης ομάδας δεδομένων (Vafidis και Spanoudakis 2003). Στην περίπτωση των νευρωνικών δικτύων το κέντρο συσσώρευσης ονομάζεται νευρώνιο (*neuron*) ή κόμβος (*neural node*) και αντιπροσωπεύει δεδομένα με παρόμοια χαρακτηριστικά.

1.5 Ιστορική εξέλιξη και εφαρμογές των στιγμιαίων χαρακτηριστικών.

Τα σεισμικά μιγαδικά χαρακτηριστικά εισήχθησαν περίπου το 1970 για την απεικόνιση και την ποιοτική ερμηνεία των σεισμικών δεδομένων. Ο Balch (1970) έκανε την πρώτη δημοσίευση για τον υπολογισμό του στιγμιαίου πλάτους και της στιγμιαίας συχνότητας, ενώ ο Anstey (1972) εισήγαγε τις χρωματικές απεικονίσεις των σεισμικών μιγαδικών χαρακτηριστικών. Οι Taner και συνεργάτες (1979) δημοσίευσαν τα αποτελέσματα των ερευνών τους και καθιέρωσαν τα μιγαδικά χαρακτηριστικά ως εργαλεία ερμηνείας των σεισμικών δεδομένων.

Ο Schmitt (1999) εφάρμοσε τα βασικά στιγμιαία μιγαδικά χαρακτηριστικά σε σεισμικά δεδομένα από περιοχή στα βορειοανατολικά της επαρχίας Αλμπέρτα του Καναδά. Επρόκειτο για κοίτασμα βαρέος πετρελαίου το οποίο φιλοξενούνταν σε αβαθείς ψαμμίτες του Κάτω Κρητιδικού. Από την απεικόνιση του στιγμιαίου πλάτους διαπίστωσε ότι τονίζονται οι ισχυροί ανακλαστήρες. Η αντίστοιχη απεικόνιση της στιγμιαίας συχνότητας είναι αρκετά πιο πολύπλοκη και αναδεικνύει την πλευρική συνέχεια των ανακλαστήρων. Παρατήρησε ότι στο κομμάτι των δεδομένων που αντιστοιχεί στον ταμιευτήρα αερίου, η στιγμιαία συχνότητα έχει υψηλές τιμές και μεταβάλλεται με αργό

ρυθμό, ενώ στην επαφή αερίου – πετρελαίου η στιγμιαία συχνότητα μεταβάλλεται απότομα.

Οι Robertson και Nogami (1984) παρουσίασαν το αποτέλεσμα της μελέτης τους που αφορούσε στην εφαρμογή του στιγμιαίου πλάτους και της στιγμιαίας συχνότητας σε συνθετικά σεισμικά δεδομένα, θέλοντας να αναδείξουν τη συμβολή των μιγαδικών χαρακτηριστικών στην μελέτη της στρωματογραφίας και ειδικότερα τη μελέτη λεπτών στρωμάτων. Διαπίστωσαν ότι, ένα λεπτό στρώμα ψαμμίτη, με χαμηλή ακουστική εμπέδηση και πάχος που είναι της τάξης του μισού μήκους κύματος του παλμού και περιβάλλεται από πετρώματα με υψηλές ταχύτητες, παρουσιάζει εξαιρετικά μεγάλες τιμές στιγμιαίου πλάτους. Ακόμα, διαπίστωσαν ότι όσο το πάχος του ψαμμιτικού στρώματος προσεγγίζει το ένα τέταρτο του μήκους κύματος του παλμού, η στιγμιαία συχνότητα μεταβάλλεται δίδοντας ακόμα υψηλότερες τιμές.

Ο Boadu (1997) εφάρμοσε τα μιγαδικά χαρακτηριστικά σε συνθετικά σεισμικά δεδομένα για πετρώματα με οριζόντιες ρωγμές, με στόχο να μελετήσει τη διαπερατότητα. Διαπίστωσε ότι οι τιμές των μιγαδικών χαρακτηριστικών τείνουν να αυξηθούν, καθώς αυξάνεται η διαπερατότητα που οφείλεται στις ρωγμές.

Η Orlando (2002) χρησιμοποίησε το στιγμιαίο πλάτος, τη στιγμιαία φάση και τη στιγμιαία συχνότητα δεδομένων γεωραντάρ για τον εντοπισμό ελαφρών διαλυμάτων μη υδατικής φάσης (LNAPL). Υποστηρίζει ότι τα μιγαδικά χαρακτηριστικά εξαρτώνται από τη συχνότητα του παλμού και από το πάχος του ρυπασμένου στρώματος και καταλήγει στο ότι είναι κατάλληλα εργαλεία για την οριοθέτηση ρυπασμένων περιοχών.

Πολλοί ερευνητές κατατάσσουν τα μιγαδικά χαρακτηριστικά σε κατηγορίες χρησιμοποιώντας διάφορα κριτήρια. Ο Taner (1992-2000) τα κατατάσσει αρχικά σε δυο κατηγορίες ανάλογα με το αν πρόκειται για δεδομένα πριν ή μετά την υπέρθεση. Περαιτέρω τα κατατάσσει σε στιγμιαία μιγαδικά χαρακτηριστικά, που αφορούν κάθε ίχνος της τομής ξεχωριστά, σε φυσικά ανάλογα με τη φυσική σημασία τους, σε γεωμετρικά που αφορούν τρισδιάστατα δεδομένα και σε διάφορες άλλες κατηγορίες (Κυματιδίων, ανακλαστικότητα, κ.α). Ο Brown (1996) τα κατατάσσει σε κατηγορίες που αφορούν στο χρόνο, στο πλάτος, στη συχνότητα και στην εξασθένιση. Οι Chen και Sidney (1997) τα ταξινομούν σε κατηγορίες που βασίζονται σε κινηματικά-δυναμικά χαρακτηριστικά και σε αυτά που αφορούν στις πετροφυσικές ιδιότητες του ταμιευτήρα

πετρελαίου. Τέλος, παραθέτουν πίνακες με τα μιγαδικά χαρακτηριστικά όπου διακρίνονται σε στιγμιαία ενός ή περισσότερων από ένα ιχνών. Αυτοί οι πίνακες περιλαμβάνουν ακόμα στοιχεία για τη χρησιμότητά τους.

1.6 Νευρωνικά δίκτυα και ασαφής λογική: βιβλιογραφική έρευνα.

Τα νευρωνικά δίκτυα, κλάδος της τεχνητής νοημοσύνης, προσπαθούν να μιμηθούν τον ανθρώπινο εγκέφαλο, υπό την έννοια ότι, μαθαίνουν από τα λάθη και την εμπειρία τους. Στην έρευνα με γεωφυσικές μεθόδους εισήχθησαν τα τελευταία μόλις χρόνια, αλλά προτάθηκαν από τους ψυχολόγους McCulloh και Pitts τη δεκαετία του 1940, στην προσπάθεια να μιμηθούν την ικανότητα του εγκεφάλου στη μάθηση. Αν και στην πορεία αμφισβητήθηκε η ικανότητά τους να λύνουν έστω απλά προβλήματα, τελικά καθιερώθηκαν σε πολλούς τομείς ως χρήσιμα εργαλεία επίλυσης, κυρίως μη γραμμικών προβλημάτων.

Η ασαφής λογική (fuzzy logic) αναπτύχθηκε τη δεκαετία του 1960 από τον Lofti Zadeh για την μελέτη δεδομένων αμφίβολης ποιότητας και ακρίβειας. Ο ίδιος υποστήριζε ότι “όσο αυξάνεται η πολυπλοκότητα, οι ακριβείς ισχυρισμοί χάνουν τη σημασία τους και οι σημαντικοί ισχυρισμοί χάνουν σε ακρίβεια”. Στη έρευνα με γεωφυσικές μεθόδους εισήχθησαν τα τελευταία χρόνια (Aminzadeh και Wilkinson 2004).

Από το 1998 μέχρι σήμερα έχουν δημοσιευτεί περισσότερα από 700 άρθρα που αφορούν στη χρήση-εφαρμογή των νευρωνικών δικτύων στη γεωφυσική. Όπως είναι φανερό καμία ανασκόπηση δε θα μπορούσε να τα συμπεριλάβει όλα. Οι εφαρμογές των νευρωνικών δικτύων στη γεωφυσική μπορούν να χωριστούν σε δυο περιόδους (Poulton 2002). Μέχρι το 1994 η προσοχή είχε επικεντρωθεί στο τι μπορούν να προσφέρουν τα νευρωνικά δίκτυα: οι επιστήμονες πειραματίζονταν κατά την προετοιμασία των γεωφυσικών δεδομένων, πριν τα εισαγάγουν σε αυτά, καθώς και με την ανάλυση των αποτελεσμάτων. Από το 1995 και μετά η προσοχή στράφηκε σε ποικίλες εφαρμογές κυρίως για τον χαρακτηρισμό ταμιευτήρων υδρογονανθράκων (reservoir characterization). Σε αυτή την ανασκόπηση θα γίνει μια σύντομη αναφορά σε σημαντικές δημοσιεύσεις, ενώ έμφαση θα δοθεί σε δημοσιεύσεις σχετικές με την παρούσα εργασία.

Οι Taner και συνεργάτες (1988) και Poulton και συνεργάτες (1989) ήταν οι πρώτοι που απέδειξαν ότι, τα νευρωνικά δίκτυα μπορούν να χρησιμοποιηθούν στην εκτίμηση των γεωφυσικών παραμέτρων.

Οι Carr και συνεργάτες (2001α) χρησιμοποίησαν δεδομένα από διαγραφίες και τρισδιάστατα σεισμικά δεδομένα για τη μελέτη ταμιευτήρα υδρογονανθράκων στο νότιο Τέξας. Η σεισμική διασκόπηση κάλυπτε έκταση πενήντα (50) τετραγωνικών μιλίων, ενώ είχαν πραγματοποιηθεί συνολικά τέσσερις (4) γεωτρήσεις. Με τη χρήση νευρωνικών δικτύων και συγκεκριμένα με τη βοήθεια των δικτύων αυτο-οργανωμένης απεικόνισης (Self Organizing Maps-SOM) ταξινομήθηκαν τα δεδομένα των διαγραφιών σε εκατό (100) τάξεις. Στη συνέχεια μείωσαν τον αριθμό των τάξεων σε δώδεκα αντιστοιχίζοντας τις σε γεωλογικούς σχηματισμούς. Το ίδιο έκαναν και με τα ταξινομημένα σεισμικά δεδομένα. Χρησιμοποιώντας το εκπαιδευμένο-βαθμονομημένο νευρωνικό δίκτυο πραγματοποίησαν την ταξινόμηση όλων των δεδομένων και έδειξαν ότι η μέθοδος είναι πολλά υποσχόμενη. Η ίδια μεθοδολογία εφαρμόστηκε (Carr και συνεργάτες 2001β) σε κοιτάσμα της Βόρειας Θάλασσας, όπου η σεισμική διασκόπηση κάλυπτε τριακόσια εικοσιπέντε (325) τετραγωνικά μίλια, ενώ ο αριθμός των γεωτρήσεων ήταν πέντε (5).

Οι Walls και συνεργάτες (2002) μελέτησαν ταμιευτήρα αέριων υδρογονανθράκων με σκοπό να χαρτογραφήσουν τη χωρική μεταβολή του πορώδους, χρησιμοποιώντας δεδομένα διαγραφιών, δείγματα πυρήνων από τις γεωτρήσεις καθώς και τρισδιάστατα σεισμικά δεδομένα. Η περιοχή μελέτης κάλυπτε συνολικά εικοσιπέντε (25) τετραγωνικά μίλια και υπήρχαν έξι (6) γεωτρήσεις. Η μελέτη πραγματοποιήθηκε σε δυο φάσεις που περιλάμβαναν: (α) ποιοτική ερμηνεία των σεισμικών μιγαδικών χαρακτηριστικών και (β) λεπτομερή μελέτη των στοιχείων από τις διαγραφίες. Στην πρώτη φάση υπολόγισαν ένα ικανό αριθμό μιγαδικών και γεωμετρικών χαρακτηριστικών των σεισμικών δεδομένων. Χρησιμοποιώντας το δίκτυο αυτο-οργανωμένης απεικόνισης ταξινόμησαν τα δεδομένα σε περιοχές με κοινά χαρακτηριστικά αναδεικνύοντας τις μεταβολές στη λιθολογία και το πορώδες. Στη δεύτερη φάση αναλύοντας τους πυρήνες και μελετώντας τις διαγραφίες έφτιαξαν συνθετικά ομοιώματα (μοντέλα) με στόχο να επιλέξουν σεισμικά μιγαδικά χαρακτηριστικά πιο ευαίσθητα στην παρουσία των υδρογονανθράκων. Κατόπιν εκπαίδευσαν ένα νευρωνικό δίκτυο, ώστε από τα σεισμικά μιγαδικά χαρακτηριστικά, τις τιμές του πορώδους και τη λιθολογία από τις διαγραφίες, να εκτιμήσουν το πορώδες σε

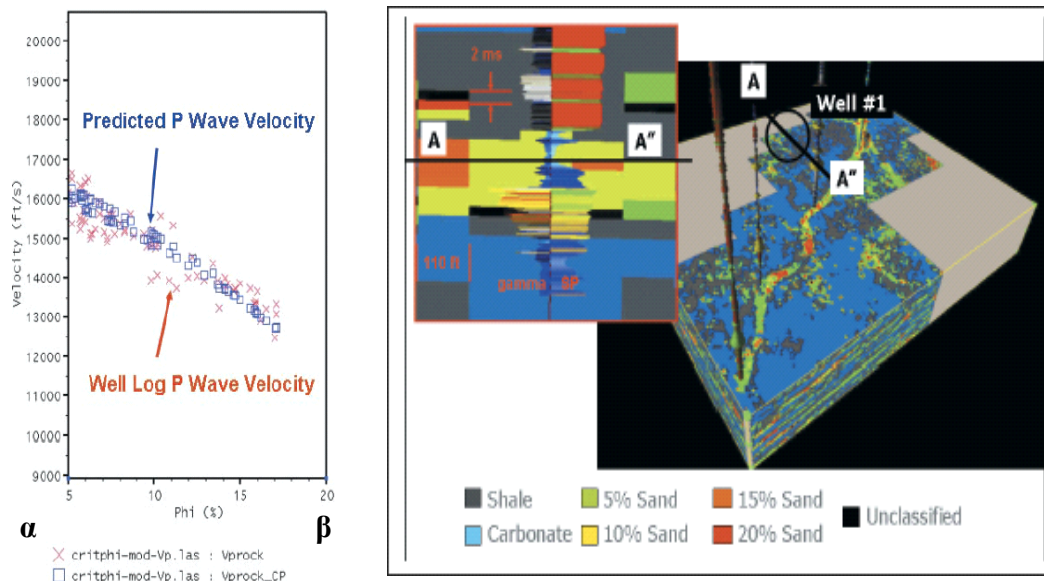
όλη την έκταση του ταμιευτήρα. Τα αποτελέσματα της έρευνας ήταν πολύ ενθαρρυντικά, αφού η ακρίβεια της εκτίμησης του πορώδους ξεπέρασε το 80%.

Οι Walls και συνεργάτες (1999, 2000) παρουσιάζουν δύο παραδείγματα όπου χρησιμοποιούν σεισμικά χαρακτηριστικά, πετροφυσικά δεδομένα και νευρωνικά δίκτυα με στόχο να προσδιορίσουν τη λιθολογία μεταξύ των γεωτρήσεων και να χαρτογραφήσουν τη μεταβολή του πορώδους. Στη πρώτη περίπτωση μελέτησαν την ευαισθησία των σεισμικών χαρακτηριστικών στη μεταβολή της λιθολογίας. Βασιζόμενοι σε διαγραφίες κατασκεύασαν συνθετικά γεωφυσικά δεδομένα και υπολόγισαν επτά (7) σεισμικά χαρακτηριστικά, με τα οποία τροφοδότησαν το νευρωνικό δίκτυο. Με αυτό το τρόπο απεικόνισαν σχιστόλιθους, ανθρακικούς σχηματισμούς και ψαμμίτες με πορώδες 5%, 10% και 15% αντίστοιχα. Στη δεύτερη περίπτωση προσδιόρισαν οκτώ (8) γεωλογικούς σχηματισμούς από δείγματα πυρήνων και διαγραφίες πέντε (5) συνολικά γεωτρήσεων. Τρεις από αυτές ήταν λίγο παραγωγικές, ενώ οι δύο δεν ήταν. Ο σκοπός της έρευνας ήταν να προσδιορίσουν παραγωγικές περιοχές ώστε να τις διατρήσουν. Το νευρωνικό δίκτυο εκπαιδεύτηκε με δεδομένα πυκνότητας, ολικού πορώδους, ταχύτητας διάδοσης των διαμήκων (V_p) και εγκάρσιων (V_s) κυμάτων και βαθμού κορεσμού σε νερό. Η πρόβλεψη που πέτυχαν με τη βοήθεια του δικτύου ήταν εξαιρετική σε γνωστή γεώτρηση (Σχήμα 1.1α). Στην επόμενη φάση της έρευνας συσχέτισαν τα σεισμικά χαρακτηριστικά με τα δεδομένα διαγραφιών, ώστε να προβλέψουν το πορώδες. Αφού εκπάιδευσαν το δίκτυο με συνθετικά δεδομένα παρήγαγαν τρισδιάστατη εικόνα για τους γεωλογικούς σχηματισμούς (Σχήμα 1.1β). Διαπίστωσαν ότι οι παραγωγικές γεωτρήσεις βρίσκονταν μέσα στην εκτιμώμενη παραγωγική ζώνη, ενώ οι μη παραγωγικές βρίσκονταν μακριά από αυτήν.

Οι Strecker και Uden (2002) χρησιμοποίησαν τα σεισμικά χαρακτηριστικά και το δίκτυο αυτο-οργανωμένης απεικόνισης για την ερμηνεία τρισδιάστατων σεισμικών δεδομένων. Η σεισμική διασκόπηση κάλυψε έκταση εβδομήντα δυο (72) τετραγωνικών μιλίων στη νότια Λουϊζιάνα των Η.Π.Α., ενώ στη μελέτη δε διέθεταν στοιχεία από γεωτρήσεις (πυρήνες ή διαγραφίες). Αρχικά υπολόγισαν σεισμικά χαρακτηριστικά (στιγμιαία, γεωμετρικά και απόκρισης). Έπειτα έκαναν τέσσερις διαφορετικούς συνδυασμούς από αυτά και τα ταξινόμησαν. Παρατήρησαν ότι όταν ο συνδυασμός περιέχει περισσότερα γεωμετρικά χαρακτηριστικά στην ταξινόμηση τονίζεται η δομή του

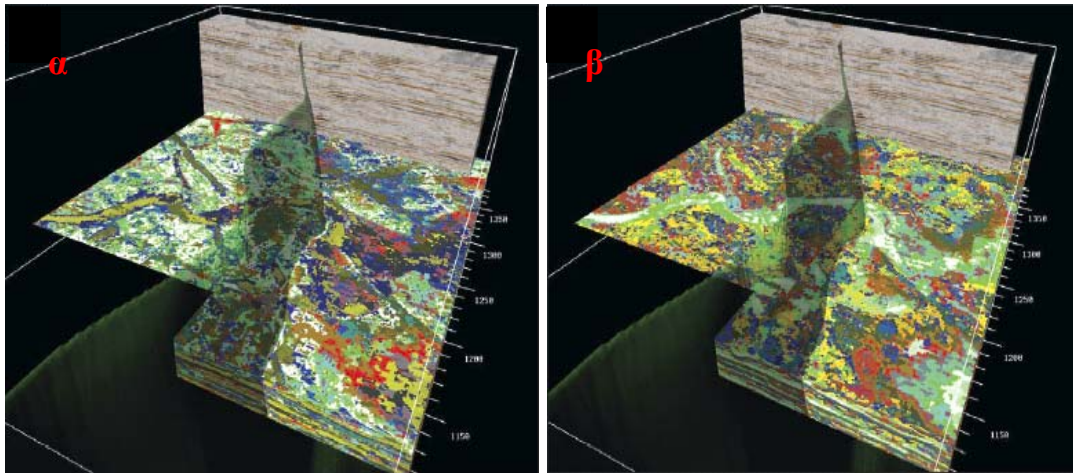
υπεδάφους, ενώ όταν ο συνδυασμός περιέχει περισσότερα στιγμιαία χαρακτηριστικά τονίζεται η διαφοροποίηση στη λιθολογία (σχήμα 1.2). Μελετώντας ταυτόχρονα τους παραπάνω συνδυασμούς προέκυψαν σημαντικά συμπεράσματα όσον αφορά στην ύπαρξη παλαιο-καναλιών, την ετερογένεια των γεωλογικών σχηματισμών καθώς και τις θέσεις για την πραγματοποίηση γεωτρήσεων.

Οι Banchs και Michelena (2002) διαθέτοντας τρισδιάστατα σεισμικά δεδομένα και στοιχεία από διαγραφίες φυσικού δυναμικού (SP) σε εικοσιπέντε (25) γεωτρήσεις, μελέτησαν κοίτασμα υδρογονανθράκων στην ανατολική Βενεζουέλα. Στόχος τους ήταν να δημιουργήσουν ένα τρισδιάστατο όγκο από ψευδό-διαγραφίες χρησιμοποιώντας τα σεισμικά χαρακτηριστικά, τα νευρωνικά δίκτυα καθώς και την προσομοίωση με τη μέθοδο Monte Carlo. Η διαδικασία που ακολούθησαν είναι εν συντομία η εξής (σχήμα 1.3): α) μετατροπή της κλίμακας βάθους των διαγραφιών σε κλίμακα χρόνου β) αύξηση του αριθμού των δειγμάτων στα σεισμικά δεδομένα (up-sampling) ώστε κάθε δείγμα να αντιστοιχεί σε μία μέτρηση της διαγραφίας. Κατόπιν, χρησιμοποιήθηκαν χαρακτηριστικά ιχνών που βρίσκονται κοντά στις γεωτρήσεις για την εκπαίδευση του δικτύου. Αναλυτικότερα, για περισσότερη ακρίβεια, χρησιμοποιήθηκαν όλα τα ίχνη γύρω από κάθε γεώτρηση, αντί του μέσου όρου τους. Επιπλέον, για κάθε μέτρηση της διαγραφίας χρησιμοποιήθηκαν δείγματα γειτονικών ιχνών που ανήκουν σε χρονικό παράθυρο γύρω από αυτό, αντί μιας τιμής. Οι συγγραφείς όμως εκφράζουν τους προβληματισμούς τους ως προς την επιλογή των σεισμικών χαρακτηριστικών που χρησιμοποιούνται στην εκπαίδευση του δικτύου. Η ανάλυση κυρίων συνιστωσών θεωρείται ότι δεν είναι το πλέον κατάλληλο εργαλείο, αφού η σχέση μεταξύ των σεισμικών χαρακτηριστικών δεν είναι γραμμική. Επίσης προβληματίζονται για την επιλογή των παραμέτρων του νευρωνικού δικτύου και συγκεκριμένα για το βέλτιστο αριθμό κρυφών επιπέδων και τον αριθμό των νευρώνων στο κάθε επίπεδο. Μάλιστα, υποστηρίζουν ότι για τη λύση των παραπάνω προβλημάτων απαιτείται μεγάλος αριθμός δοκιμών. Οι ίδιοι όμως κάνουν εβδομήντα (70) διαφορετικές δοκιμές και παίρνουν το βέλτιστο αποτέλεσμα από την τεχνική Monte Carlo. Η ακρίβεια πρόβλεψης των διαγραφιών είναι της τάξης του 80%.

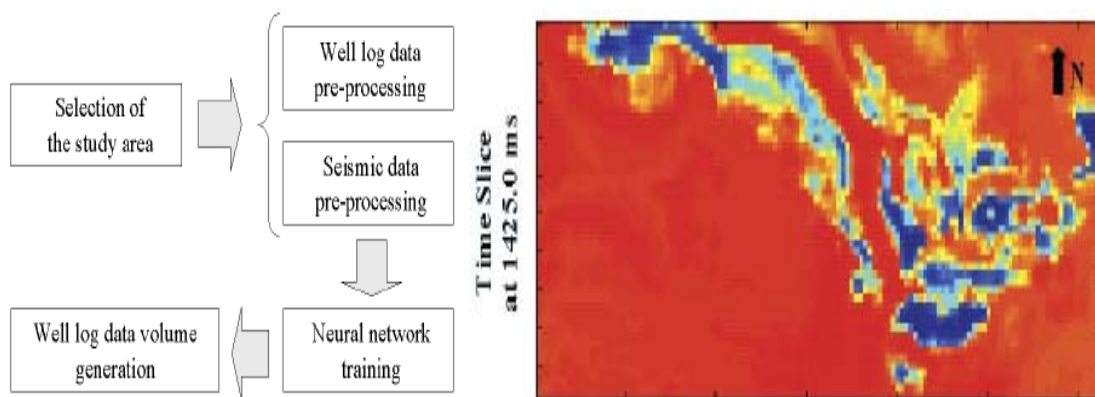


Σχήμα 1.1. (α) Διάγραμμα διασποράς όπου συγκρίνονται οι πραγματικές τιμές της ταχύτητας P των σεισμικών κυμάτων με τις εκτιμώμενες. (β) Από αυτές τα δεδομένα ταξινομήθηκαν σύμφωνα με τη λιθολογία (Walls και συνεργάτες 2002).

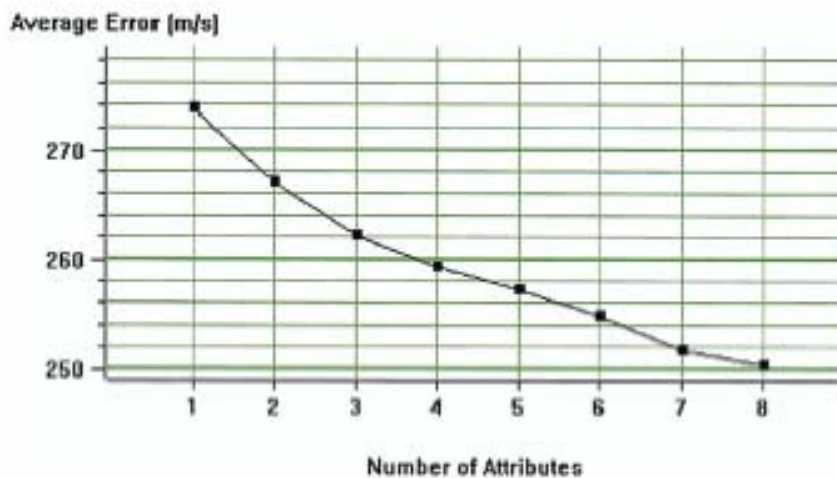
Οι Hampson και συνεργάτες (2000) εκτιμούν το πορώδες από σεισμικά χαρακτηριστικά χρησιμοποιώντας παλινδρόμηση (stepwise regression) και νευρωνικά δίκτυα. Η παλινδρόμηση χρησιμεύει στο να επιλεγεί ο βέλτιστος συνδυασμός σεισμικών χαρακτηριστικών. Διαπιστώνουν ότι, όσο περισσότερα σεισμικά χαρακτηριστικά προστίθενται στην παλινδρόμηση, τόσο μειώνεται το σφάλμα της εκτίμησης (σχήμα 1.4). Στην έρευνά τους χρησιμοποιούν έξι (6) σεισμικά χαρακτηριστικά και δυο διαφορετικά δίκτυα. Το δίκτυο της ανάστροφης διάδοσης λαμβάνει υπόψη όλες τις τιμές, ακόμα και αν αυτές είναι ακραίες. Αυτό είναι ένα σημαντικό προσόν του δικτύου, το οποίο όμως δεν είναι τόσο χρήσιμο στη συγκεκριμένη περίπτωση, αφού οι ακραίες τιμές δεν είναι χρήσιμες και πρέπει να αγνοούνται. Εναλλακτικά, χρησιμοποιούν δίκτυα πιθανότητας (εκτίμηση Bayes). Τελικά προτιμούν τα δίκτυα γενικευμένης παλινδρόμησης (generalized regression) αφού η εκτίμηση ακολουθεί τα γενικά χαρακτηριστικά των σεισμικών χαρακτηριστικών (όπως και της ανάστροφης διάδοσης). Το κύριο πρόβλημα αυτών των δικτύων είναι ότι απαιτούν μεγάλο χρόνο εφαρμογής.



Σχήμα 1.2. Ταξινομημένα τρισδιάστατα σεισμικά δεδομένα. Όταν χρησιμοποιούνται περισσότερα στιγμιαία χαρακτηριστικά τονίζεται η διαφοροποίηση στη λιθολογία (α) ενώ όταν χρησιμοποιούνται περισσότερα γεωμετρικά χαρακτηριστικά τονίζεται η δομή του υπεδάφους (β) (Strecker και Uden 2002).



Σχήμα 1.3. Το διάγραμμα ροής για την εκτίμηση ψευδο-διαγραφιών (δεξιά), από την οποία προέκυψε η οριζόντια τομή (δεξιά). Το κόκκινο χρώμα συμβολίζει την άμμο ενώ το μπλε, το πράσινο και το κίτρινο το σχιστόλιθο (Banchs και Michelena 2002).



Σχήμα 1.4. Όσο αυξάνεται ο αριθμός των χαρακτηριστικών τόσο μειώνεται το σφάλμα εκτίμησης (Hampson και συνεργάτες 2000).

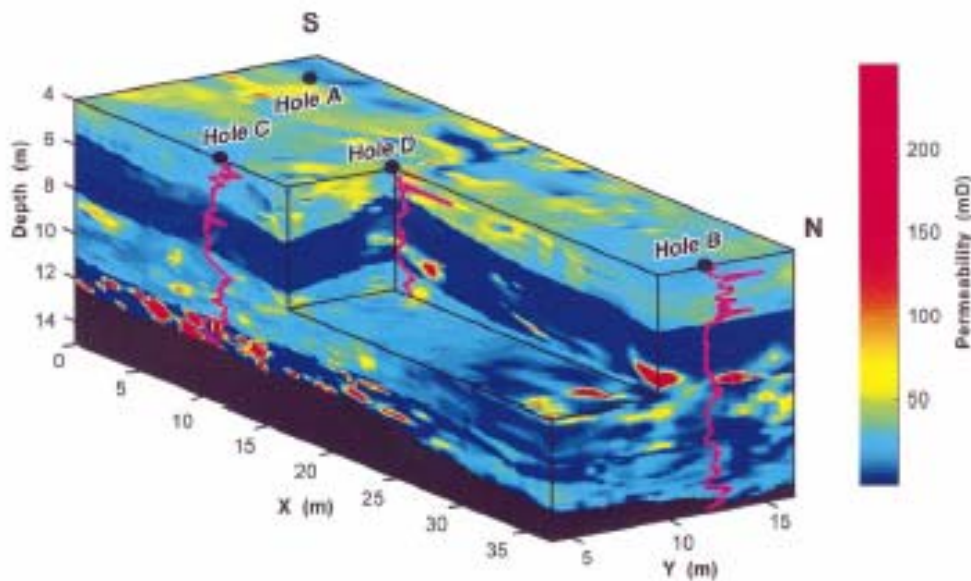
Οι Leiphart και Hart (2001) χρησιμοποίησαν τρισδιάστατα σεισμικά δεδομένα και δεδομένα διαγραφιών, για να προσδιορίσουν τη χωρική κατανομή του πορώδους σε ταμιευτήρα πετρελαίου. Η περιοχή μελέτης κάλυπτε δεκαέξι (16) τετραγωνικά μίλια και είχαν πραγματοποιηθεί εβδομήντα επτά (77) γεωτρήσεις. Σε αυτές είχαν πραγματοποιηθεί διαγραφίες ακτίνων γάμμα, πυκνότητας, ηλεκτρικής αντίστασης και ακουστικές. Ακολουθώντας το παράδειγμα του Hampson και των συνεργατών του (2000), για να επιλέξουν τα κατάλληλα σεισμικά χαρακτηριστικά χρησιμοποίησαν τη μέθοδο της παλινδρόμησης. Κατέληξαν να χρησιμοποιούν έξι (6) από αυτά, ενώ για την εκτίμηση του πορώδους χρησιμοποίησαν δυο τρόπους: γραμμική παλινδρόμηση και στοχαστικά νευρωνικά δίκτυα πιθανότητας. Αφού χαρτογράφησαν την υπό μελέτη περιοχή ως προς το πορώδες καταλήγουν ότι τα δίκτυα πιθανότητας είναι καλύτερα εργαλεία, αφού το μέσο σφάλμα πρόβλεψης είναι 1.4%, ενώ με τη γραμμική παλινδρόμηση το αντίστοιχο σφάλμα ήταν 2.2%.

Ο Corbeanu και οι συνεργάτες του (2002) χρησιμοποίησαν τρισδιάστατα δεδομένα γεωραντάρ και γεωτρήσεις, για να υπολογίσουν τη διαπερατότητα καθώς και να χαρτογραφήσουν τους ιλυόλιθους. Συγκεκριμένα, επέλεξαν τριάντα τέσσερις (34) γραμμές γεωραντάρ και μετρήσεις της διαπερατότητας των σχηματισμών σε πέντε (5)

γεωτρήσεις. Από τα δεδομένα του γεωραντάρ υπολόγισαν δυο μιγαδικά χαρακτηριστικά: το στιγμιαίο πλάτος και τη στιγμιαία συχνότητα. Αυτά τα δεδομένα προσπάθησαν να τα συσχετίσουν με τη μέθοδο της γραμμικής παλινδρόμησης. Θεώρησαν ότι η σχέση που συνδέει την διαπερατότητα με τα χαρακτηριστικά του γεωραντάρ έχει τη μορφή: $\ln(P(x, y, z)) = c_0 + c_1 A + c_2 F$, όπου \ln ο φυσικός λογάριθμος, P η διαπερατότητα, A το στιγμιαίο πλάτος και F η στιγμιαία συχνότητα, ενώ c_0, c_1, c_2 είναι σταθερές. Αφού προσδιόρισαν τους παραπάνω συντελεστές, τους εφάρμοσαν σε όλα τα δεδομένα του γεωραντάρ και διαπίστωσαν ότι η διαπερατότητα κυμαίνεται από 0.1md έως 200md. Αντίστοιχα, χρησιμοποιώντας μεθόδους ταξινόμησης, διαπίστωσαν ότι οι ιλύολιθοι έχουν μέγεθος το οποίο κυμαίνεται από 2m έως 20m (σχήμα 1.5).

Η Kalkomey (1997) αναλύει τα πιθανά λάθη που μπορεί να κάνει κάποιος, όταν χρησιμοποιεί τα σεισμικά χαρακτηριστικά, για να υπολογίσει τις πετροφυσικές παραμέτρους ενός ταμιευτήρα. Υποστηρίζει ότι η πιθανότητα εσφαλμένης εκτίμησης των παραμέτρων του ταμιευτήρα από τα σεισμικά χαρακτηριστικά είναι μεγάλη, όταν ο αριθμός των παρατηρήσεων που γίνονται μέσα στις διαθέσιμες γεωτρήσεις είναι μικρός ή ο αριθμός των ανεξάρτητων σεισμικών χαρακτηριστικών είναι πολύ μεγάλος. Επίσης όταν τα σεισμικά χαρακτηριστικά επιλέγονται με βάση μόνο εμπειρικά στοιχεία, η πιθανότητα λανθασμένης απόφασης και συνεπώς κακής διαχείρισης του ταμιευτήρα, είναι μεγάλη. Προτείνει την επιλογή σεισμικών χαρακτηριστικών, τα οποία συνδέονται με τις μετρούμενες στις γεωτρήσεις ποσότητες.

Ο Russel και οι συνεργάτες του (1997) επικεντρώνουν το ενδιαφέρον τους στην πρόβλεψη των ιδιοτήτων ενός ταμιευτήρα χρησιμοποιώντας τα σεισμικά χαρακτηριστικά και πολυδιάστατη στατιστική. Συγκεκριμένα, χρησιμοποιούν τη μέθοδο της γραμμικής παλινδρόμησης για την επιλογή των σεισμικών χαρακτηριστικών, που θα χρησιμοποιήσουν για την πρόβλεψη του πορώδους, ενώ επιλέγουν τα δείγματα της εκπαίδευσης από ευρύτερη περιοχή γύρω από τις γεωτρήσεις. Για τη εφαρμογή της μεθόδου σε όλα τα δεδομένα της σεισμικής γραμμής χρησιμοποιούν τους συντελεστές που προέκυψαν από την παλινδρόμηση. Ο Russel και οι συνεργάτες του (2001) χρησιμοποιούν την ίδια μεθοδολογία σε τρισδιάστατα σεισμικά δεδομένα για την εκτίμηση του πορώδους. Σε αυτή τη μελέτη χρησιμοποιούν την γραμμική παλινδρόμηση για την επιλογή των σεισμικών χαρακτηριστικών και τη μέθοδο εκτίμησης Kriging.



Σχήμα 1.5. Ψευδό-τρισδιάστατος όγκος δεδομένων γεωραντάρ, ο οποίος αναπαριστά στην κατανομή της διαπερατότητας, όπως αυτή εκτιμήθηκε από τα χαρακτηριστικά των δεδομένων και τις διαγραφίες (Corbeanu και συνεργάτες 2002).

Ο Schultz και οι συνεργάτες του (1994 α, β, γ) σε μια σειρά από άρθρα περιγράφουν την εμπειρία τους πάνω στην πρόβλεψη των πετροφυσικών παραμέτρων του ταμειυτήρα, χρησιμοποιώντας στοιχεία από διαγραφίες και τα σεισμικά χαρακτηριστικά. Μέσω κυρίως της γραμμικής παλινδρόμησης προσδιορίζουν την προσεγγιστική σχέση σεισμικών χαρακτηριστικών και πορώδους και προτείνουν τα δίκτυα πιθανότητας. Τέλος, δίνουν παραδείγματα από πραγματικά δεδομένα αποδεικνύοντας έτσι την ορθότητα της μεθοδολογίας τους. Τα άρθρα αυτά είναι πολύ κατατοπιστικά αφού προσδιορίζουν τον τρόπο εργασίας βήμα προς βήμα.

Ο Fogg (2000) χρησιμοποίησε τα σεισμικά χαρακτηριστικά για την πρόβλεψη των παραμέτρων ταμειυτήρα πετρελαίου, διαθέτοντας διαγραφίες ακτίνων γάμμα και πορώδους καθώς και τρισδιάστατα σεισμικά δεδομένα. Αρχικά, με τη μέθοδο της βηματικής γραμμικής παλινδρόμησης, επέλεξε τα βέλτιστα σεισμικά χαρακτηριστικά. Με τη χρήση νευρωνικών δικτύων πιθανότητας προέβλεψε το πορώδες όλης της περιοχής που καλύπτονταν από τη σεισμική διασκόπηση.

Ο Todorov και οι συνεργάτες του (1998 α, β) διέθεταν δεδομένα διαγραφιών πορώδους από δεκατρείς (13) γεωτρήσεις και τρισδιάστατα σεισμικά δεδομένα. Σκοπός τους ήταν να προβλέψουν το πορώδες των σχηματισμών της υπό έρευνα περιοχής. Για να επιτύχουν το στόχο τους, επέλεξαν τη μέθοδο της βηματικής γραμμικής παλινδρόμησης και νευρωνικά δίκτυα.

Τέλος, αναφέρεται η επιτυχής εφαρμογή των μεθόδων ταξινόμησης δεδομένων γεωραντάρ για τον εντοπισμό καρστικών εγκοίλων στο οροπέδιο του Ομαλού, στο νομό Χανίων (Spanoudakis και συνεργάτες 2004).

Οι αναφορές στη βιβλιογραφία για την πρόβλεψη πετροφυσικών παραμέτρων ταμιευτήρα με τη χρήση των σεισμικών χαρακτηριστικών και των νευρωνικών δικτύων είναι πολλές και στις περισσότερες αναφέρονται παραδείγματα εφαρμογών (Huang και συνεργάτες 1996, Gastaldi και συνεργάτες 1997, Todorov και συνεργάτες 1998β, Heggland 1999, Macias και συνεργάτες 2000, Moon και Kalatzis 2001, Tonn 2002, Asevedo και Pennigton 2003, Para και συνεργάτες 2003, Saggaf και συνεργάτες 2003α, Bachrach και συνεργάτες 2004, Bose και συνεργάτες 2004, Calvert και συνεργάτες 2004, Dorrington και Linkz 2004, Linari 2004, Dowd και Iguzquida 2005, Saggaf και συνεργάτες 2003α).

Τα τελευταία χρόνια τείνει να καθιερωθεί η ταξινόμηση ανακλαστήρων (Coleou και συνεργάτες 2002, Linari και συνεργάτες 2003, Saggaf και συνεργάτες 2003b, Basu και συνεργάτες 2004, Roupon και συνεργάτες 2004, Zeng 2004), η οποία εξηγείται και εφαρμόζεται σε δεδομένα, στο τέταρτο και έκτο κεφάλαιο.

Όσον αφορά στην ταξινόμηση δεδομένων γεωραντάρ σύμφωνα με τα χαρακτηριστικά τους, οι αναφορές και τα παραδείγματα είναι ελάχιστα (Matos και συνεργάτες 2004, Spanoudakis και συνεργάτες 2004).

Τα νευρωνικά δίκτυα δεν έχουν χρησιμοποιηθεί μόνο για την πρόβλεψη των παραμέτρων ταμιευτήρων και την ταξινόμηση σεισμικών δεδομένων. Οι Liu και συνεργάτες (1989) επικέντρωσαν την προσοχή τους στην ερμηνεία σεισμικών δεδομένων, οι Huang και συνεργάτες (1989) στον εντοπισμό ισχυρών ανακλάσεων (*bright spots*), ο McCormac (1990) στην επιλογή πρώτων αφίξεων, ενώ το 1991 επικεντρώθηκε στην ταξινόμηση μετρήσεων από διαγραφίες καθώς και στον εντοπισμό ελαττωματικών ιχνών (*trace editing*). Οι Schmidt και συνεργάτες (1992) επικεντρώθηκαν

στην ανάλυση ταχυτήτων, οι Wang και Mendel (1992) στην αποσυνέλιξη, οι Calderon και συνεργάτες (1997) με την εξασθένιση πολλαπλών καταγραφών και τη διόρθωση NMO.

Οι Chakravarthy και συνεργάτες (1999) χρησιμοποίησαν δίκτυα πιθανότητας (PNN) για τον προσδιορισμό των ορίων μεταξύ των γεωλογικών σχηματισμών χρησιμοποιώντας δεδομένα από διαγραφίες.

Αρκετοί έχουν χρησιμοποιήσει τα νευρωνικά δίκτυα για την αντιστροφή και την ανάλυση δυναμικών πεδίων (Fossati και συνεργάτες 1992, Guo και συνεργάτες 1992, Cartabia και συνεργάτες 1994, Vasco και Taylor 1991),

Τέλος, αναφέρονται δημοσιεύσεις που αφορούν καθαρά στην παρουσίαση μεθοδολογίας για τη χρήση τεχνικών όπως τα νευρωνικά δίκτυα, η ασαφής λογική και οι γενετικοί αλγόριθμοι (Le και Nyland 1990, McCormac 1991, McCormac και συνεργάτες 1999, Smith και συνεργάτες 1992).

ΚΕΦΑΛΑΙΟ 2.**ΑΡΧΗ ΛΕΙΤΟΥΡΓΙΑΣ ΚΑΙ ΕΠΕΞΕΡΓΑΣΙΑ ΔΕΔΟΜΕΝΩΝ ΓΕΩΡΑΝΤΑΡ****2.1 Το γεωραντάρ.**

Το γεωραντάρ είναι υψηλής ευκρίνειας ηλεκτρομαγνητική τεχνική, η οποία αναπτύχθηκε τα τελευταία 30 χρόνια κυρίως για τη μελέτη του ρηχού υπεδάφους, τον έλεγχο του οδοστρώματος, την εξέταση αστοχιών σε γέφυρες και κατασκευές (Daniels και συνεργάτες 1997, Daniels και συνεργάτες 1998). Η βασική αρχή λειτουργίας του γεωραντάρ είναι ίδια με αυτή της μεθόδου της σεισμικής ανάκλασης.

2.2 Εφαρμογές του γεωραντάρ.

Οι εφαρμογές του γεωραντάρ μπορούν να διακριθούν ουσιαστικά σε δύο κατηγορίες:

α) Για γεωλογικούς σκοπούς, όταν το βάθος διασκόπησης είναι σημαντικότερο από την διακριτική ικανότητα, γι' αυτό χρησιμοποιούνται κεραίες συχνότητας 2 MHz έως 50 MHz. β) Για εφαρμογές στις επιστήμες των μηχανικών και στις μη καταστροφικές δοκιμές (Non Destructive Testing, NDT) όπου η διακριτική ικανότητα είναι το

ζητούμενο, γι' αυτό χρησιμοποιούνται κεραίες συχνότητας 500 MHz - 2000 MHz. γ) για εφαρμογές στην αρχαιολογία όπου χρησιμοποιούνται κεραίες συχνότητας 50 MHz - 500 MHz.

Η ικανότητα του γεωραντάρ να προσφέρει λεπτομερείς πληροφορίες για το ρηχό υπέδαφος έχει αναφερθεί σε πλήθος δημοσιεύσεων. Αυτές αφορούν στον εντοπισμό ρωγμών και αστοχιών σε πετρώματα (Toshioka και συνεργάτες 1995), τον προσδιορισμό την στρωματογραφίας (Davis και Annan 1989, Dominic και συνεργάτες 1995), αρχαιολογική έρευνα (Sternberg και McGill 1995), τον έλεγχο της ποιότητας ασφαλοτάπητα (Hugenschmidt και συνεργάτες 1998), τη διάβρωση σκυροδέματος (Lorenzo και συνεργάτες 2001), τον προσδιορισμό της διαπερατότητας (Corbeanu και συνεργάτες 2002), τον εντοπισμό υγρών ρύπων στο υπέδαφος (Greenhouse και Brewster 1993, Brewster και Annan, 1994, Benson και Stubben 1997, Orlando 2002), τον εντοπισμό εγκοίλων (Spanoudakis και συνεργάτες 2004) και άλλες. Ο Olhoeft (1988) παρέχει εκτεταμένη βιβλιογραφική έρευνα για τις χρήσεις του γεωραντάρ.

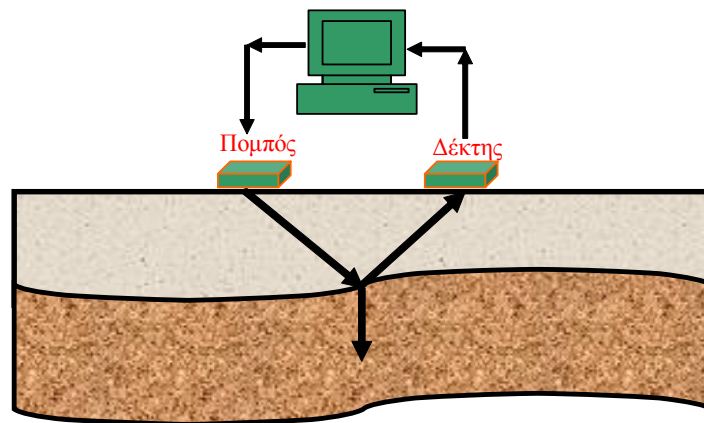
2.3 Βασικές αρχές.

Το γεωραντάρ εκπέμπει μικρής διάρκειας ηλεκτρομαγνητικό παλμό από κεραία-πομπό, ο οποίος ταξιδεύει στο υπέδαφος με ταχύτητα η οποία εξαρτάται κυρίως από την ηλεκτρική διαπερατότητα του υλικού (είναι αντιστρόφως ανάλογη με την τετραγωνική ρίζα της ηλεκτρικής διαπερατότητας). Το ηλεκτρομαγνητικό κύμα διαδίδεται στο υπέδαφος, ωστόσο συναντήσει διαχωριστική επιφάνεια διαφορετικών ηλεκτρικών ιδιοτήτων. Το κύμα ανακλάται και φτάνει στην κεραία δέκτη όπου και καταγράφεται (σχήμα 2.1).

Ο χρόνος που χρειάζεται το ηλεκτρομαγνητικό κύμα για να διαδοθεί από τον πομπό στο δέκτη ανεξαρτήτως της διαδρομής του, ονομάζεται χρόνος διαδρομής. Βασική μονάδα μέτρησης του χρόνου διαδρομής είναι το νανο-δευτερόλεπτο ($1ns = 10^{-9}s$). Αν το όργανο αρχίσει να καταγράφει τη στιγμή που εκπέμπει η κεραία-πομπός, δυο κύρια είδη κυμάτων πρόκειται να καταγραφούν. Στην πρώτη κατηγορία ανήκουν κύματα που ταξιδεύουν απ' ευθείας από τον πομπό στον δέκτη μέσω του αέρα. Αυτά τα κύματα, αφενός φτάνουν πρώτα, επειδή η ταχύτητα διάδοσης στον αέρα είναι μεγαλύτερη απ' ότι

σε οποιοδήποτε υλικό, αφετέρου έχουν μεγάλο πλάτος, αφού δεν υφίστανται σημαντική εξασθένιση. Στη δεύτερη κατηγορία ανήκουν κύματα που διαδίδονται στο υπέδαφος. Η καταγραφή του πλάτους των κυμάτων ονομάζεται ίχνος.

Όταν το κύμα συναντήσει υλικό με διαφορετική ηλεκτρική διαπερατότητα, αλλάζει κατεύθυνση δηλαδή διαθλάται ή ανακλάται. Περίθλαση συμβαίνει όταν το κύμα συναντήσει πολύ μικρού μεγέθους εμπόδια, τα οποία το αναγκάζουν να διαδίδεται προς όλες τις κατευθύνσεις. Αντήχηση συμβαίνει, όταν το κύμα “παγιδευτεί” σε κάποιο γεωλογικό στρώμα. Σε τέτοιες περιπτώσεις το κύμα ανακλάται πολλές φορές με συνέπεια τη δημιουργία πολλαπλών ανακλάσεων.

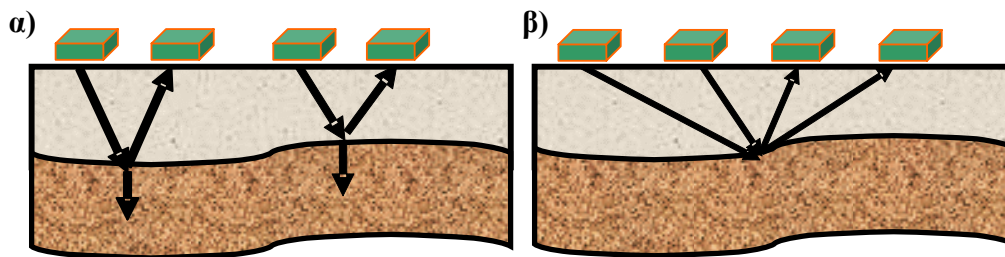


Σχήμα 2.1. Αρχή λειτουργίας του γεωραντάρ.

2.4 Διατάξεις πομπού - δέκτη.

Υπάρχουν διάφοροι τρόποι διάταξης πομπού - δέκτη ανάλογα με το σκοπό της έρευνας. Κύρια διάταξη για τη συλλογή δεδομένων γεωραντάρ είναι αυτή της σταθερής απόστασης πομπού δέκτη (*common offset*), όπου πομπός και δέκτης κινούνται κατά μήκος γραμμής μελέτης στην επιφάνεια του εδάφους, διατηρώντας την μεταξύ τους απόσταση σταθερή (σχήμα 2.2α). Η απόσταση μεταξύ διαδοχικών καταγραφών ονομάζεται ισο-απόσταση σταθμών παρατήρησης ή βήμα διασκόπησης. Όταν τα ίχνη απεικονίζονται το ένα δίπλα στο άλλο, σχηματίζεται η τομή γεωραντάρ (*GPR section*). Η μέθοδος αυτή χρησιμεύει κυρίως στην χαρτογράφηση του υπεδάφους. Αν είναι γνωστή η διηλεκτρική σταθερά του υπεδάφους, μπορεί να υπολογιστεί η ταχύτητα διάδοσης των

ηλεκτρομαγνητικών κυμάτων. Εναλλακτικά, για τον υπολογισμό της ταχύτητας χρησιμοποιείται η διάταξη κοινού σημείου βάθους (*common depth point-CDP*), όπου επιλέγεται στη επιφάνεια του εδάφους κάποιο σταθερό σημείο, εκατέρωθεν του οποίου τοποθετούνται πομπός και δέκτης (σχήμα 2.2β). Τέλος, αναφέρεται η μέθοδος της τομογραφίας, όπου πομπός και δέκτης τοποθετούνται ξεχωριστά σε κοντινές γεωτρήσεις, εκατέρωθεν τοίχων ή στοών μεταλλείων.



Σχήμα 2.2. Τρόποι συλλογής δεδομένων. **α)** Η διάταξη σταθερής απόστασης πομπού δέκτη και **β)** η διάταξη κοινού σημείου βάθους.

2.5 Επεξεργασία των δεδομένων του γεωραντάρ.

Η επεξεργασία των δεδομένων του γεωραντάρ είναι εμπνευσμένη από τη μέθοδο σεισμικής ανάκλασης. Επειδή το πεδίο είναι πολύ ευρύ, αναφέρονται τεχνικές αφαίρεσης του θορύβου και ενίσχυσης των δεδομένων (*gain*). Περισσότερη έμφαση δίδεται στην τεχνική εξαγωγής των χαρακτηριστικών του σήματος του γεωραντάρ (*GPR attributes*), η οποία κερδίζει ολοένα έδαφος στο στάδιο της ερμηνείας των δεδομένων.

Το αρχικό στάδιο επεξεργασίας περιλαμβάνει την εφαρμογή φίλτρων με στόχο να απομακρυνθεί το σήμα στις χαμηλές συχνότητες. Αυτό το βήμα είναι γνωστό ως “*dewowing*”. Το πολύ χαμηλής συχνότητας σήμα σχετίζεται με επαγωγικά φαινόμενα ή το περιορισμένο δυναμικό εύρος του εξοπλισμού.

Το επόμενο βήμα της επεξεργασίας περιλαμβάνει την εφαρμογή φίλτρου ενίσχυσης στα δεδομένα. Υπάρχουν διάφοροι τρόποι ενίσχυσης του σήματος. Μερικές φορές είναι σημαντικό να τονιστεί η οριζόντια εξάπλωση κάποιου γεωλογικού σχηματισμού εις βάρος της ακρίβειας. Σε αυτή την περίπτωση εφαρμόζεται ο αυτόματος έλεγχος

ενίσχυσης. Εναλλακτικά, εφαρμόζεται ενίσχυση που αντισταθμίζει τη γεωμετρική διασπορά του κύματος. Με αυτόν τον τρόπο συμπυκνώνεται το δυναμικό εύρος του παλμού δηλαδή μικραίνει ο λόγος της πιο ισχυρής προς την πιο ασθενή καταγραφή (Annan 1999).

Οι ειδικές τεχνικές περιλαμβάνουν την ανάλυση των χαρακτηριστικών των δεδομένων, αποσυνέλιξη, εξασθένιση των πολλαπλών ανακλάσεων, εφαρμογή φίλτρων στο χώρο συχνότητας – κυματαριθμού (*FK*) και απομάκρυνση του θορύβου υποβάθρου (Σπανουδάκης 2004). Για την απαλοιφή της αντήχησης της κεραίας του πομπού εφαρμόζεται φίλτρο διέλευσης υψηλών συχνοτήτων.

Η μιγαδική ανάλυση ιχνών αποτελεί τεχνική που εξασφαλίζει την τοπική λεπτομέρεια και έχει ως αποτέλεσμα τον φυσικό διαχωρισμό του πλάτους και της φάσης μιας κυματομορφής (Balch 1970, Taner και συνεργάτες 1979). Το χαρακτηριστικό πλάτος αναφέρεται ως ισχύς ανάκλασης ή στιγμιαίο πλάτος του σήματος, ενώ η φάση αναφέρεται ως στιγμιαία φάση. Η στιγμιαία φάση αποτελεί τη βάση για τον υπολογισμό της στιγμιαίας συχνότητας, ενώ συνδυασμοί των παραπάνω δίδουν χρήσιμα για την ερμηνεία χαρακτηριστικά (Barnes 1998). Σύμφωνα με την κατηγοριοποίηση που έγινε στο πλαίσιο της παρούσας εργασίας, τα χαρακτηριστικά των δεδομένων γεωραντάρ μπορεί να είναι στιγμιαία, απόκρισης ή γεωμετρικά.

Τα στιγμιαία χαρακτηριστικά αναφέρονται στις διακυμάνσεις των φυσικών ιδιοτήτων μιας χρονοσειράς και προκύπτουν από τον μετασχηματισμό Hilbert. Το πλήρες μαθηματικό υπόβαθρο για το μετασχηματισμό Hilbert, ο υπολογισμός πληθώρας στιγμιαίων χαρακτηριστικών καθώς και οι χρήσεις τους, αναπτύσσεται στη βιβλιογραφία (Vafidis και συνεργάτες 2002, Σπανουδάκης 2002). Τα παραπάνω, αν το επίπεδο του θορύβου είναι υψηλό, υπολογίζονται χρησιμοποιώντας διαδοχικές μετρήσεις του ίδιου ίχνους, χρονικής διάρκειας ίσης με αυτή του εκπεμπόμενου παλμού.

Τα χαρακτηριστικά απόκρισης ή κυματιδίων, υπολογίζονται στα τοπικά μέγιστα του στιγμιαίου πλάτους και σχετίζονται άμεσα με το μετασχηματισμό Fourier. Για παράδειγμα, η στιγμιαία συχνότητα που αντιστοιχεί στα τοπικά μέγιστα του πλάτους ισούται με τη μέση συχνότητα του φασματικού εύρους ενός παλμού.

2.6 Απεικόνιση των δεδομένων.

Η σωστή απεικόνιση των δεδομένων του γεωραντάρ διευκολύνει την ερμηνεία τους. Υπάρχουν διάφοροι τρόποι απεικόνισης. Η απεικόνιση ενός μοναδικού ίχνους καθιστά την ερμηνεία δύσκολη.

Μετακινώντας όμως τις κεραίες στην επιφάνεια του εδάφους με σταθερό βήμα και συνδυάζοντας περισσότερα ίχνη προκύπτει τομή γεωραντάρ, όπου ο οριζόντιος άξονας αντιστοιχεί στην θέση των κεραιών στην επιφάνεια, ενώ ο κατακόρυφος στο χρόνο διαδρομής (σχήμα 2.3). Αυτή η τομή καθιστά ευκολότερη την ερμηνεία. Το πλάτος των ηλεκτρομαγνητικών κυμάτων μπορεί επίσης να κωδικοποιηθεί με χρώματα.

Η τρισδιάστατη απεικόνιση των δεδομένων προϋποθέτει την πραγματοποίηση μετρήσεων σε ορθογώνιο πλέγμα δηλαδή σε παράλληλες και ισαπέχουσες γραμμές μελέτης. Πρέπει να τονιστεί ότι τα δεδομένα που προκύπτουν από αυτό τον τρόπο (σχήμα 2.4α) είναι ψευδο-τρισδιάστατα.

Ένας τρόπος απεικόνισης ψευδο-τρισδιάστατων δεδομένων είναι απλά η τοποθέτηση κάτω από τις παράλληλες γραμμές μελέτης. Η απεικόνιση βελτιώνεται όταν γίνει παρεμβολή μεταξύ των ιχνών. Χρήσιμο εργαλείο για την απεικόνιση τρισδιάστατων δεδομένων αποτελεί και η οριζόντια τομή (σχήμα 2.4β).

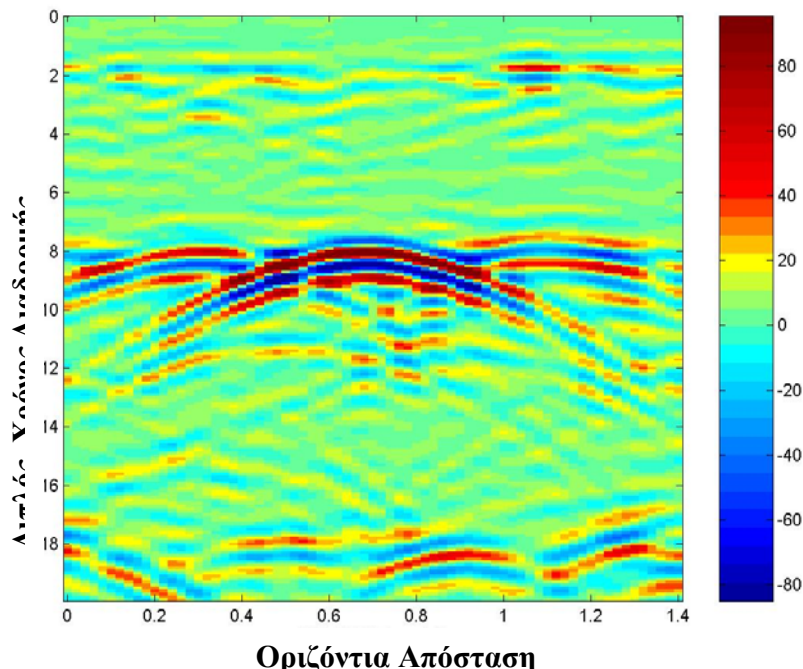
Η απλοποίηση της απεικόνισης μπορεί να βοηθήσει σημαντικά στη ερμηνεία των δεδομένων. Αυτό μπορεί να επιτευχθεί επιλέγοντας προσεκτικά τη χρωματική κλίμακα, την οπτική γωνία και τον αριθμό των χρωμάτων. Το ίδιο επιτυγχάνεται με την απεικόνιση τμήματος των δεδομένων, ή την αφαίρεση τμήματος των δεδομένων, κάνοντας διάφανο ή ημιδιάφανο κάποιο χρώμα.

Όσον αφορά στις οριζόντιες τομές, πολλές φορές βοηθά η απεικόνιση ισότιμων καμπυλών, επιφανειών σταθερού πλάτους, καθώς και χαρτών ανάγλυφου. Ο προσεκτικός συνδυασμός κάποιων από τους τρόπους απεικόνισης μπορεί να αποδειχθεί χρήσιμος για την ερμηνεία των δεδομένων.

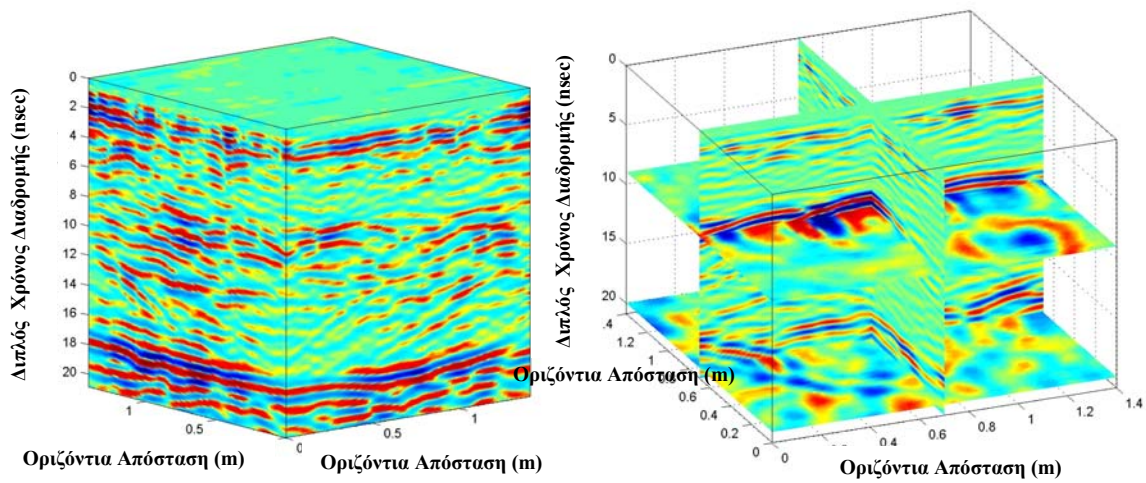
2.7 Ταξινόμηση των δεδομένων του γεωραντάρ: Διάγραμμα ροής.

Για την αυτόματη ταξινόμηση των καταγραφών γεωραντάρ ακολουθείται το παρακάτω γενικό διάγραμμα ροής (σχήμα 2.5). Αρχικά (βήμα 1), εφαρμόζονται φίλτρα στα δεδομένα με στόχο την απαλοιφή του θορύβου και των πολλαπλών ανακλάσεων. Ακολούθως (βήμα 2), υπολογίζονται κάποια συγκεκριμένα χαρακτηριστικά των καταγραφών. Η επιλογή των χαρακτηριστικών βασίζεται κυρίως στη εμπειρία, αφού δεν υπάρχει καθιερωμένος τρόπος. Για την ταξινόμηση των χαρακτηριστικών (βήμα 3) χρησιμοποιούνται είτε όλα τα δεδομένα είτε μέρος τους. Εναλλακτικά, πριν την ταξινόμηση, τα χαρακτηριστικά υφίστανται κάποια επεξεργασία, όπως την ανάλυση κυρίων συνιστωσών για την μείωση του αριθμού τους (βήμα 2β).

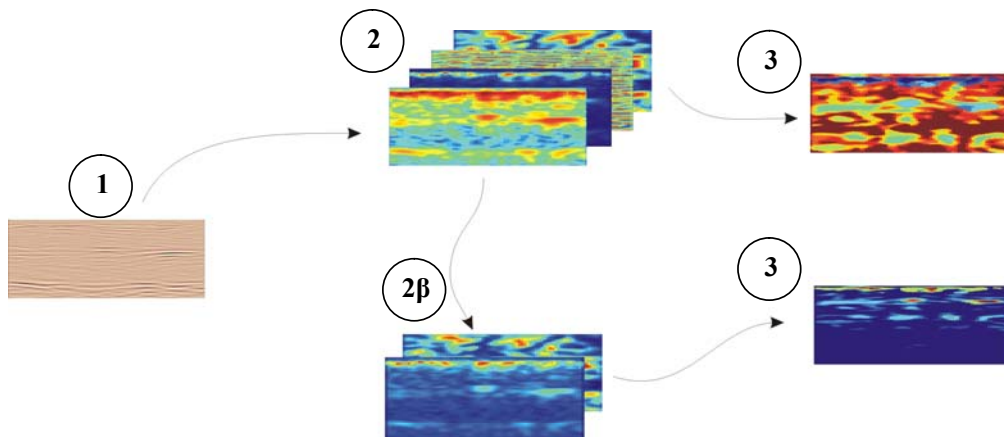
Για την καθοδηγούμενη ταξινόμηση η διαδικασία είναι η ίδια, με τη διαφορά ότι πριν την ταξινόμηση χρησιμοποιούνται επιπρόσθετες πληροφορίες, όπως άλλα γεωφυσικά δεδομένα, γεωτρήσεις κτλ (σχήμα 2.6). Εναλλακτικά, τα δεδομένα ταξινομούνται με αλγόριθμο αυτόματης ταξινόμησης και στη συνέχεια χρησιμοποιούνται πληροφορίες από γεωτρήσεις για την βαθμονόμηση των ταξινομημένων δεδομένων.



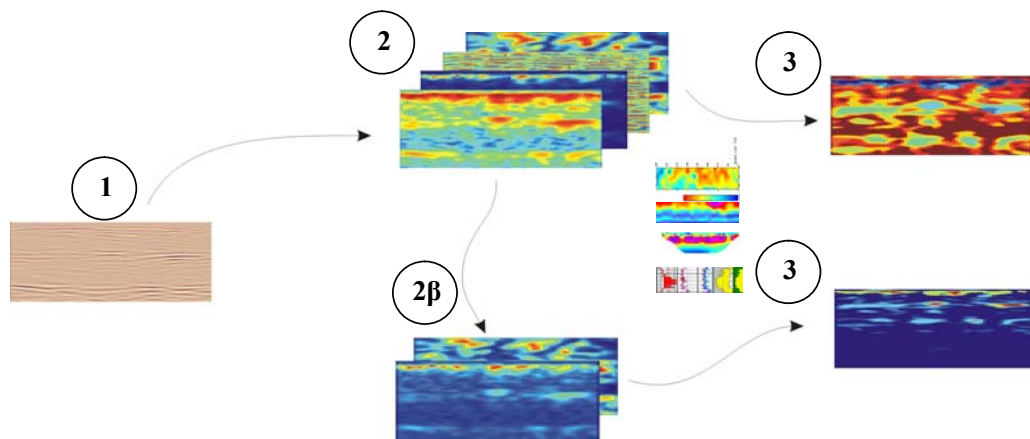
Σχήμα 2.3. Τομή γεωραντάρ. Ο οριζόντιος άξονας αντιστοιχεί στην θέση των κεραιών ενώ ο κάθετος στο χρόνο διαδρομής των κυμάτων. Η κλίμακα προσδιορίζει το πλάτος των H/M κυμάτων.



Σχήμα 2.4. Ψευδό-τρισδιάστατος όγκος δεδομένων γεωραντάρ (αριστερά). Η εξαγωγή οριζόντιων τομών (δεξιά) είναι χρήσιμη στην ερμηνεία των δεδομένων.



Σχήμα 2.5. Διάγραμμα ροής για την αυτόματη ταξινόμηση δεδομένων γεωραντάρ. 1) Εφαρμογή φίλτρων, ενίσχυσης κτλ. 2) Υπολογισμός χαρακτηριστικών. 3) Ταξινόμηση. Εναλλακτικά: 2β) Ανάλυση Κυρίων Συνιστωσών. 3) Ταξινόμηση.



Σχήμα 2.6. Διάγραμμα ροής για την καθοδηγούμενη ταξινόμηση δεδομένων γεωραντάρ. 1) Εφαρμογή φίλτρων, ενίσχυσης κτλ. 2) Υπολογισμός χαρακτηριστικών. 3) Ταξινόμηση (ή εκτίμηση παραμέτρων του υπεδάφους) με τη χρήση πρόσθετων πληροφοριών. Εναλλακτικά: 2β) Ανάλυση Κυρίων Συνιστωσών. 3) Ταξινόμηση (ή εκτίμηση παραμέτρων του υπεδάφους) με τη χρήση πρόσθετων πληροφοριών.

2.8 Λογισμικό επεξεργασίας δεδομένων γεωραντάρ.

Στο πλαίσιο εκπόνησης της μεταπτυχιακής αλλά και της διδακτορικής διατριβής μου αναπτύχθηκε ένας μεγάλος αριθμός αλγόριθμων, για την επεξεργασία των δεδομένων του γεωραντάρ σε περιβάλλον ανάπτυξης MATLAB. Η πληθώρα των αλγόριθμων κατέστησε άμεση ανάγκη την δημιουργία μιας επιφάνειας εργασίας, για την κατηγοριοποίησή τους και την εύκολη πρόσβαση και χρήση τους. Για το λόγο αυτό δημιουργήθηκε ένα λογισμικό που ονομάστηκε GPR-Pro. Παρόλο που δεν είναι ένα εμπορικό πακέτο, μπορεί να συγκριθεί με άλλα λογισμικά πακέτα επεξεργασίας δεδομένων γεωραντάρ, όπως το PulseEkko και το Reflexw. Κύριο μειονέκτημα των παραπάνω λογισμικών είναι ο περιορισμένος αριθμός αλγόριθμων που περιλαμβάνουν, αλλά και ο τρόπος διαχείρισης των δεδομένων που ακολουθεί μια διαδοχική σειρά. Αυτός ο τρόπος δεν επιτρέπει στο χρήστη να έχει πρόσβαση στα ενδιάμεσα αποτελέσματα.

Το λογισμικό υποστηρίζει τα συνηθέστερα πρότυπα αρχείων γεωραντάρ (PulseEkko, SEGY, Ramac, GSSI) αλλά και αρχεία Matlab. Περιλαμβάνει αλγόριθμους για την εφαρμογή φίλτρων σε ίχνη αλλά και τομές γεωραντάρ, τον υπολογισμό των μιγαδικών

χαρακτηριστικών αλλά και αλγόριθμους ταξινόμησης αυτών. Αναπτύχθηκε έτσι ώστε να είναι φιλικό στον χρήστη και παρέχει ένα πολύ μεγάλο αριθμό αλγόριθμων επεξεργασίας δεδομένων γεωραντάρ που προέρχονται από τα πεδία της εξερεύνησης με σεισμικές μεθόδους και την τηλεπισκόπηση. Εκτός των συμβατικών φίλτρων περιλαμβάνει τεχνικές όπως την ανάλυση κυρίων συνιστωσών, αλγεβρικές πράξεις πινάκων, υπολογισμό γεωμετρικών χαρακτηριστικών, ένα μεγάλο αριθμό στιγμιαίων χαρακτηριστικών καθώς και εργαλειοθήκη για την ταξινόμηση των δεδομένων.

Κύριο πλεονέκτημά του, έναντι των άλλων λογισμικών, είναι η διαχείριση των δεδομένων, η οποία είναι εμπνευσμένη από την εξερεύνηση των Windows (Windows Explorer) με την δένδροειδή μορφή. Αυτή καθιστά εύκολη τη παράλληλη εφαρμογή διαφορετικών αλγόριθμων στα δεδομένα έχοντας άμεση πρόσβαση στα ενδιάμεσα και τελικά αποτελέσματα. Επίσης, η απεικόνιση των δεδομένων είναι πιο αποτελεσματική παρέχοντας άμεσο έλεγχο στην επιλογή χρωματικών κλιμάκων, τον αριθμό των χρωμάτων, αλλά και το μέγεθος της εικόνας.

Αν και φιλικό προς το χρήστη, το λογισμικό κατασκευάστηκε για ακαδημαϊκή αλλά και βιομηχανική χρήση. Προϋποθέτει την εξοικείωση με εξειδικευμένες τεχνικές όπως την πολυδιάστατη στατιστική, την ταξινόμηση δεδομένων, και την προσέγγιση συναρτήσεων.

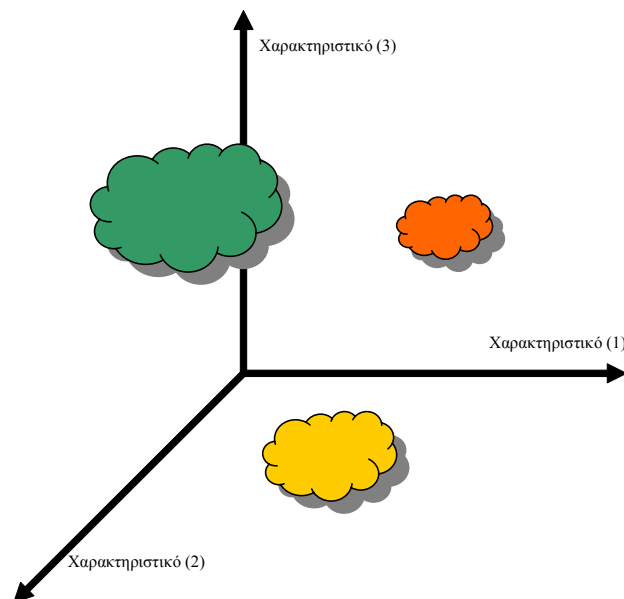
Το πακέτο είναι ανοιχτό, με την έννοια ότι είναι εύκολο να προστεθούν νέοι αλγόριθμοι. Η μελλοντική του ανάπτυξη θα περιλαμβάνει την επεξεργασία και απεικόνιση τρισδιάστατων δεδομένων. Κύριο μειονέκτημά του είναι η έλλειψη εγχειριδίου και η απαίτηση εγκατάστασης του λογισμικού MATLAB (έκδοση 6.5 ή νεότερη) στον υπολογιστή.

ΚΕΦΑΛΑΙΟ 3**ΜΕΘΟΔΟΙ ΤΑΞΙΝΟΜΗΣΗΣ ΔΕΔΟΜΕΝΩΝ ΓΕΩΡΑΝΤΑΡ****3.1 Ταξινόμηση των δεδομένων.**

Η χρήση των μεθόδων ταξινόμησης κερδίζει ολοένα έδαφος στην ερμηνεία των σεισμικών δεδομένων και των δεδομένων γεωραντάρ έναντι των συμβατικών μεθόδων επεξεργασίας. Δεν έχουν όμως καθιερωθεί ακόμα, επειδή δεν έχουν διερευνηθεί εκτενώς τα πλεονεκτήματα και οι περιορισμοί κάθε μεθόδου. Στις επόμενες παραγράφους περιγράφονται και συγκρίνονται οι μέθοδοι αυτόματης και καθοδηγούμενης ταξινόμησης που χρησιμοποιούνται συχνότερα στην ερμηνεία των γεωφυσικών δεδομένων, κυρίως της σεισμικής ανάκλασης και του γεωραντάρ. Στην αυτόματη ταξινόμηση χρησιμοποιούνται μόνο τα δεδομένα αυτά καθαυτά και όχι άλλες πληροφορίες, όπως για παράδειγμα στοιχεία από γεωτρήσεις κτλ. Αντίθετα, στην καθοδηγούμενη ταξινόμηση των δεδομένων χρησιμοποιούνται πληροφορίες και από γεωτρήσεις, ακόμα και η εμπειρία εκείνου που τα επεξεργάζεται. Συνήθως, τα δεδομένα του γεωραντάρ περιέχουν λεπτομερή πληροφορία σχετικά με το ρηχό υπέδαφος. Ο υπολογισμός των χαρακτηριστικών των δεδομένων βοηθά στην ποιοτική τους ερμηνεία, αλλά η μελέτη

των δεκάδων χαρακτηριστικών που μπορούν να υπολογιστούν είναι πολύ χρονοβόρα, παρόλο που καθένα από αυτά περιέχει διαφορετικού είδους χρήσιμες πληροφορίες. Η ανάλυση των δεδομένων με μεθόδους ταξινόμησης μπορεί να αναδείξει χρήσιμες πληροφορίες, αφού συνδυάζει το περιεχόμενο όλων των χαρακτηριστικών σε μια μόνο εικόνα.

Σε κάθε μέτρηση με το γεωραντάρ αντιστοιχούν στιγμιαία και γεωμετρικά χαρακτηριστικά, τα οποία ορίζουν ένα πολυδιάστατο διάνυσμα. Κάθε τέτοια μέτρηση μπορεί να απεικονιστεί σε διάγραμμα διασποράς από σημείο με συντεταγμένες τις τιμές των χαρακτηριστικών. Η εγγύτητα των σημείων στο διάγραμμα διασποράς συνεπάγεται παρόμοια χαρακτηριστικά. Οι τεχνικές ταξινόμησης ελέγχουν την οργάνωση των σημείων στο πολυδιάστατο διάγραμμα διασποράς αναζητώντας πληθυσμούς σημείων με παρόμοιους συνδυασμούς χαρακτηριστικών (σχήμα 3.1). Οι τεχνικές αυτές διαφέρουν μεταξύ τους ως προς τον τρόπο οργάνωσης των ομάδων.



Σχήμα 3.1. Τρισδιάστατο διάγραμμα διασποράς. Διακρίνονται τρεις πληθυσμοί δεδομένων με παρόμοιο συνδυασμό χαρακτηριστικών.

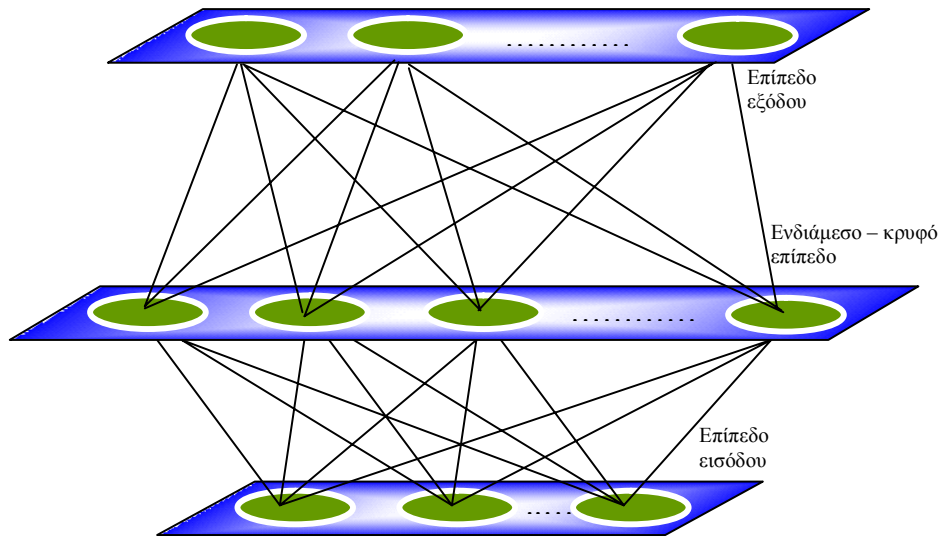
3.2 Νευρωνικά δίκτυα: βασικές αρχές.

Τα τεχνητά νευρωνικά δίκτυα αποτελούνται από διαφορετικά επίπεδα επεξεργασίας δεδομένων. Στο πρώτο επίπεδο εισάγονται τα δεδομένα στο δίκτυο. Στα ενδιάμεσα επίπεδα γίνεται η επεξεργασία των δεδομένων, ενώ το τελευταίο επίπεδο αποτελεί την έξοδο δηλαδή δίνει το τελικό αποτέλεσμα. Κάθε επίπεδο αποτελείται από τις βασικές μονάδες επεξεργασίας, οι οποίες ονομάζονται νευρώνια. Κάθε νευρωνικό δίκτυο διαφέρει από τα υπόλοιπα ως προς τον τρόπο λειτουργίας των ενδιάμεσων επιπέδων, δηλαδή τον τρόπο μάθησης. Για παράδειγμα, όλα τα δίκτυα έχουν τόσα νευρώνια στην είσοδο όσες και οι διαστάσεις του διανύσματος εισόδου. Αν, για παράδειγμα, εισαχθεί μια καταγραφή γεωραντάρ η οποία περιγράφεται από τρία χαρακτηριστικά, τα νευρώνια εισόδου θα είναι τρία. Η έξοδος του δικτύου αποτελείται από τόσα νευρώνια όσα και ο αριθμός των ομάδων που καλείται να ταξινομήσει (σχήμα 3.2).

Ο πιο διαδεδομένος τρόπος μάθησης για την ταξινόμηση των δεδομένων είναι η ανταγωνιστική. Κάθε νευρόνιο στο ανταγωνιστικό επίπεδο χαρακτηρίζεται από διάνυσμα με διαστάσεις ίσες με το διάνυσμα εισόδου το οποίο αποκαλείται διάνυσμα συναπτικού βάρους. Όλα τα νευρώνια ανταγωνίζονται μεταξύ τους για το ποιο από αυτά βρίσκεται πιο κοντά στο διάνυσμα εισόδου στο διάγραμμα διασποράς. Το νευρόνιο που νικάει στον ανταγωνισμό ενεργοποιείται, δηλαδή μεταβάλλει το διάνυσμα συναπτικού βάρους του, ώστε να έρθει πιο κοντά στο διάνυσμα εισόδου. Υπάρχουν διάφορες παραλλαγές της ανταγωνιστικής μάθησης, όπως ο αλγόριθμος της συνείδησης (*conscience rule*) και τα δίκτυα αυτο-οργανωμένης απεικόνισης (*Self Organizing Maps – SOM*).

3.3 Μέθοδοι αυτόματης ταξινόμησης.

Στις παρακάτω παραγράφους ακολουθεί σύντομη περιγραφή ορισμένων από τους πιο συνηθισμένους αλγόριθμους αυτόματης ταξινόμησης από τα πεδία της πολυδιάστατης στατιστικής (κ-μέσων τιμών (Richards 1998)), των νευρωνικών δικτύων (ανταγωνιστική μάθηση, δίκτυα αυτό-οργανωμένης απεικόνισης (Haykin 1999, Kohonen 2001)) και της ασαφούς λογικής (c- μέσων τιμών (Aminzadeh και Wilkinson 2004)). Το μαθηματικό απόβαθρο των σημαντικότερων αλγόριθμων παρατίθεται στο παράρτημα.



Σχήμα 3.2. Τυπικό νευρωνικό δίκτυο για ταξινόμηση δεδομένων. Αποτελείται από τα επίπεδα εισόδου, εξόδου και από ενδιάμεσο (κρυφό) επίπεδο επεξεργασίας.

3.3.1 Αλγόριθμος κ-μέσων τιμών.

Στον αλγόριθμο των K μέσων τιμών το κριτήριο για την ταξινόμηση των δεδομένων είναι η ελαχιστοποίηση της μέσης τετραγωνικής απόστασης όλων των δεδομένων από το αντίστοιχο κέντρο ομαδοποίησης και η μεγιστοποίηση της απόστασης μεταξύ των κέντρων ομαδοποίησης στο διάγραμμα διασποράς (Richards 1998).

Ένα ενδεχόμενο-κρίσιμο πρόβλημα του αλγόριθμου είναι η επιλογή του αριθμού K των ομάδων. Αλλάζοντας τον αριθμό K προκύπτουν τελείως διαφορετικές ομάδες. Κρίσιμη είναι επίσης και η επιλογή των αρχικών κέντρων συγκέντρωσης, αφού κάποιες ομάδες μπορεί να μείνουν άδειες. Ο αλγόριθμος των K μέσων τιμών είναι υπολογιστικά απλός και γρήγορος και χρησιμοποιείται συνήθως για την ταξινόμηση μεγάλου αριθμού δεδομένων.

3.3.2 Ανταγωνιστική μάθηση και κανόνας της συνείδησης.

Στην ανταγωνιστική μάθηση τα νευρώνια ανταγωνίζονται μεταξύ τους, ώστε να ενεργοποιηθούν (Haykin 1999). Αντίθετα με άλλες μεθόδους μάθησης όπου μερικά από τα νευρώνια ενεργοποιούνται ταυτόχρονα, στην ανταγωνιστική μάθηση ενεργοποιείται μόνο ένα κάθε φορά. Αυτό το χαρακτηριστικό καθιστά την ανταγωνιστική μάθηση κατάλληλη για την διερεύνηση στατιστικά σημαντικών χαρακτηριστικών στα δεδομένα. Τρεις είναι οι κανόνες της ανταγωνιστικής μάθησης:

- Τα νευρώνια διαφέρουν ως προς τα διανύσματα συναπτικών βαρών τα οποία αποτελούνται από τυχαίους αριθμούς.
- Τα νευρώνια ανταγωνίζονται μεταξύ τους, ώστε για κάθε νέα είσοδο μόνο ένα από αυτά ενεργοποιείται, δηλαδή νικάει στον ανταγωνισμό. Το νευρώνιο αυτό αποκαλείται νευρώνιο νικητής.
- Επιβάλλεται σε κάθε νευρώνιο όριο, για να μη νικάει ή χάνει συνεχώς στον ανταγωνισμό.

Στην απλούστερη μορφή της ανταγωνιστικής μάθησης, το νευρωνικό δίκτυο έχει μονάχα ένα επίπεδο νευρωνίων στην έξοδο, όπου κάθε νευρώνιο συνδέεται με όλες τις εισόδους του δικτύου. Τα νευρώνια εξόδου μπορεί να συνδέονται μεταξύ τους και κάθε νευρώνιο ανταγωνίζεται με αυτά που συνδέεται άμεσα.

Αν ορισμένα νευρώνια δεν νικάνε ποτέ στον ανταγωνισμό (νεκρά νευρώνια), όσο και αν διαρκεί η εκπαίδευση τότε χρησιμοποιείται μια διαφορετική εκδοχή της ανταγωνιστικής μάθησης που περιέχει τον κανόνα της στατιστικής απόκλισης και αποκαλείται μάθηση με συνείδηση. Για να υλοποιηθεί αυτό δίνεται στα νευρώνια που νικάνε σπάνια ή και ποτέ στον ανταγωνισμό, ένα πλεονέκτημα σε σχέση με τα άλλα.

Κατά τη διάρκεια της εκπαίδευσης τίθεται σε λειτουργία μηχανισμός που μετράει πόσες φορές κάθε νευρώνιο νικάει στον ανταγωνισμό. Έτσι η παράμετρος απόκλισης για τα νευρώνια που νικάνε συχνά έχει μικρές τιμές, ενώ για τα νευρώνια που νικάνε σπάνια ή ποτέ έχει μεγάλες τιμές. Αυτή η παρέμβαση έχει δυο βασικά πλεονεκτήματα:

- Αν κάποιο νευρώνιο δε νικάει στον ανταγωνισμό, οι παράμετροι απόκλισης το βοηθάνε. Όταν αυτό αρχίσει να νικάει η παράμετρος απόκλισης μηδενίζεται. Οπότε λύνεται το πρόβλημα με τα νεκρά νευρώνια.
- Κάθε νευρώνιο αναμένεται να ταξινομήσει περίπου τον ίδιο αριθμό δεδομένων. Αφού συμμετέχουν όλα τα νευρώνια στον ανταγωνισμό, η ταξινόμηση είναι πιο λεπτομερής, έστω και αν κάποιο από τα νευρώνια τείνει να ταξινομήσει περισσότερα δεδομένα σε σχέση με τα υπόλοιπα.

3.3.3 Δίκτυα αυτο-οργανωμένης απεικόνισης.

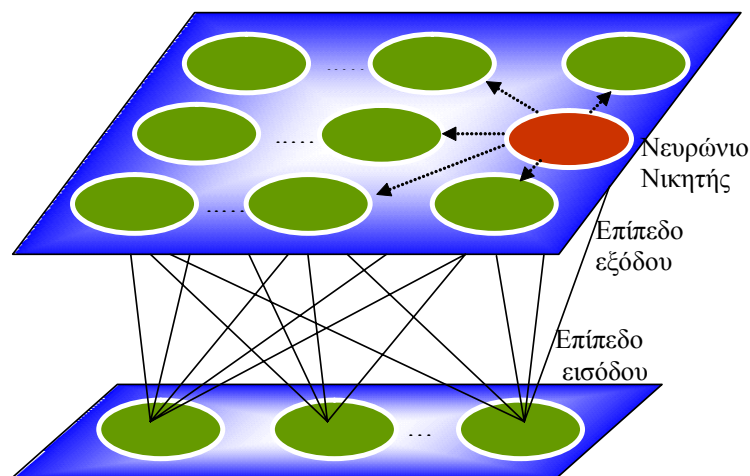
Το δίκτυο αυτο-οργανωμένης απεικόνισης είναι ειδική περίπτωση νευρωνικών δικτύων ανταγωνιστικής μάθησης (Haykin 1999, Kohonen 2001). Τα νευρώνια σε αυτή την περίπτωση είναι τοποθετημένα στους κόμβους πλέγματος (σχήμα 3.3). Τα νευρώνια εκπαιδεύονται βάσει των δειγμάτων εισόδου, σύμφωνα με τους κανόνες της ανταγωνιστικής μάθησης. Το πλεονέκτημα της μεθόδου είναι ότι τα νευρώνια διατάσσονται το ένα σε σχέση με το άλλο και σχηματίζεται τοπολογικός χάρτης, όπου η πλευρική διάταξή τους είναι ενδεικτική των φυσικών στατιστικών χαρακτηριστικών που περιέχονται στα δεδομένα.

Το δίκτυο αυτο-οργανωμένης απεικόνισης είναι εμπνευσμένο από τον ανθρώπινο εγκέφαλο ο οποίος είναι οργανωμένος ώστε γειτονικά νευρώνια να εκτελούν παρόμοιες λειτουργίες. Συγκεκριμένα, τα νευρώνια της αφής, της όρασης και της ακοής βρίσκονται σε διαφορετική θέση στον εγκέφαλο. Κάθε μία από τις αισθήσεις περιέχει πολλά νευρώνια τα οποία εκτελούν παρόμοιες λειτουργίες.

Ο αλγόριθμος της αυτο-οργανωμένης απεικόνισης ξεκινάει χωρίς οργανωμένη διάταξη. Τρεις είναι οι θεμελιώδεις λειτουργίες του δικτύου αυτο-οργανωμένης απεικόνισης:

1. *Ανταγωνισμός.* Τα νευρώνια ανταγωνίζονται μεταξύ τους και κάθε φορά που εισάγεται νέα μέτρηση, μόνο ένα νευρώνιο ανακηρύσσεται νικητής.
2. *Συνεργασία.* Το νευρώνιο-νικητής καθορίζει μια τοπολογική γειτονιά μέσα στην οποία τα νευρώνια συνεργάζονται.

3. *Προσαρμογή συναπτικών βαρών.* Αυτό το στάδιο είναι που διαφοροποιεί τα δίκτυα αυτό-οργανωμένης απεικόνισης από τα άλλα ανταγωνιστικά νευρωνικά δίκτυα. Σε αυτό το στάδιο διορθώνεται όχι μόνο το διάνυσμα συναπτικών βαρών του νευρωνίου νικητή, αλλά και των γειτονικών του νευρωνίων σε μικρότερο βαθμό, έτσι ώστε να έχουν τοπολογική συνέχεια.



Σχήμα 3.3. Το δίκτυο αυτο-οργανωμένης απεικόνισης (SOM). Τα νευρώνια στην έξοδο είναι τοποθετημένα στους κόμβους πλέγματος. Το νευρώριο νικητής (κόκκινο) καθορίζει την τοπολογική γειτονιά μέσα στην οποία τα νευρώνια συνεργάζονται.

3.3.4 Ασαφής αφαιρετική ταξινόμηση.

Οι περισσότεροι αλγόριθμοι αυτόματης ταξινόμησης είναι ουσιαστικά καθοδηγούμενοι υπό την έννοια ότι είναι απαραίτητο να προσδιοριστεί ο ακριβής αριθμός των ομάδων που υπάρχουν στα δεδομένα, πριν γίνει η ταξινόμηση. Αν ο αριθμός των ομάδων δεν είναι εκ των προτέρων γνωστός, είναι απαραίτητο να εφαρμοστεί αυτόματη διαδικασία υπολογισμού του αριθμού των ομάδων. Η αφαιρετική ταξινόμηση βασίζεται στην εύρεση περιοχών στο διάγραμμα διασποράς όπου υπάρχει υψηλή συγκέντρωση δεδομένων.

Επιλέγεται το κέντρο της ομάδας για την περιοχή υψηλής συγκέντρωσης δεδομένων και έπειτα όλα τα δεδομένα που βρίσκονται σε μία προκαθορισμένη ακτίνα γύρω από το κεντρικό σημείο αφαιρούνται και ο αλγόριθμος ψάχνει για νέο κέντρο. Η διαδικασία συνεχίζεται μέχρι να εξεταστούν όλα τα σημεία.

3.3.5 Ασαφής ταξινόμηση c-μέσων τιμών.

Σε αντίθεση με άλλους αλγόριθμους ταξινόμησης, η ασαφής ταξινόμηση c-μέσων τιμών, θεωρεί ότι κάθε διάνυσμα ανήκει σε περισσότερες από μια ομάδες με ένα βαθμό συμμετοχής. Αφού κάθε διάνυσμα μπορεί να ανήκει μόνο σε μια ομάδα, ορίζεται πίνακας με τις παρακάτω ιδιότητες:

- Το άθροισμα κάθε στήλης του ισούται με τη μονάδα.
- Το άθροισμα όλων των στοιχείων ισούται με τον αριθμό των διανυσμάτων εισόδου.

Μετά από πολλές επαναλήψεις-δόκιμες του αλγόριθμου, ο παραπάνω πίνακας σταθεροποιείται οπότε τα διανύσματα ταξινομούνται στη σωστή ομάδα.

3.4 Μέθοδοι καθοδηγούμενης ταξινόμησης.

Οι μέθοδοι καθοδηγούμενης ταξινόμησης προϋποθέτουν την ύπαρξη πρόσθετων πληροφοριών σχετικά με τα δεδομένα. Αυτές οι πληροφορίες προέρχονται από γεωτρήσεις ή και άλλα γεωφυσικά δεδομένα.

3.4.1 Δίκτυα διανυσματικής μάθησης.

Τα δίκτυα διανυσματικής μάθησης (LVQ) χρησιμοποιούνται στην καθοδηγούμενη ταξινόμηση των δεδομένων και αποτελούνται από δυο επίπεδα επεξεργασίας (σχήμα

3.4): Το πρώτο είναι ανταγωνιστικό, ενώ το δεύτερο γραμμικό. Το πρώτο μαθαίνει αυτόματα να κατηγοριοποιεί τα δεδομένα, όπως ακριβώς τα ανταγωνιστικά δίκτυα. Το γραμμικό στρώμα μετατρέπει τις ομάδες του ανταγωνιστικού επιπέδου σε ομάδες στόχους, όπως είχαν οριστεί αρχικά. Οι τάξεις που προκύπτουν από το ανταγωνιστικό επίπεδο αναφέρονται σαν υπο-τάξεις, ενώ οι τάξεις που προκύπτουν από το γραμμικό επίπεδο σαν τάξεις-στόχοι.

Και τα δύο επίπεδα διαθέτουν από ένα νευρώριο για κάθε τάξη. Όμως το ανταγωνιστικό επίπεδο μαθαίνει να κατηγοριοποιεί τα δεδομένα σε N_1 τάξεις, ενώ το γραμμικό επίπεδο σε N_2 τάξεις όπου $N_1 \gg N_2$.

3.4.2 Νευρωνικά δίκτυα πιθανότητας.

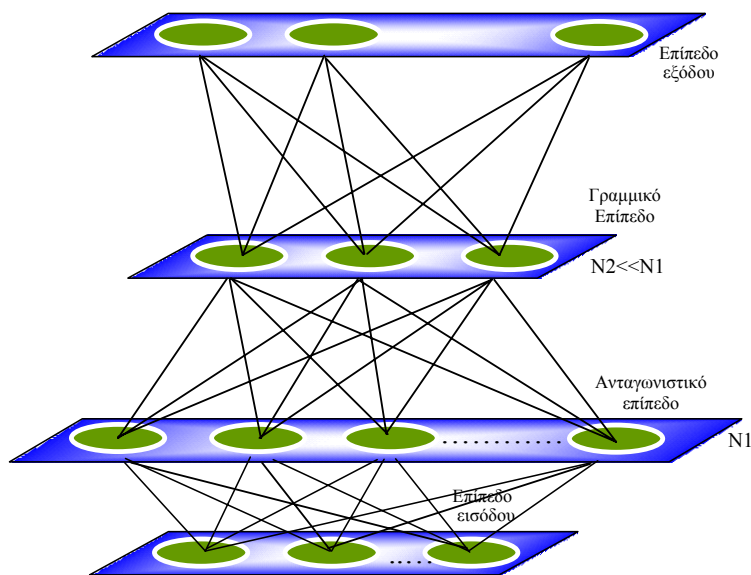
Στα δίκτυα πιθανότητας χρησιμοποιούνται το κρυφό επίπεδο και το ανταγωνιστικό επίπεδο εξόδου. Η εκπαίδευση είναι απλή και γρήγορη υπολογιστικά: Όταν εισάγεται ένα διάνυσμα στο δίκτυο, υπολογίζεται (στο κρυφό επίπεδο) η Ευκλείδεια απόσταση του διανύσματος αυτού από κάθε δείγμα εκπαίδευσης εισόδου και παράγεται ένα διάνυσμα του οποίου τα στοιχεία υποδεικνύουν την απόστασή τους. Το ανταγωνιστικό επίπεδο αθροίζει τις συνεισφορές για κάθε τάξη για τη δημιουργία διανύσματος πιθανοτήτων. Η ανταγωνιστική συνάρτηση μεταφοράς στην έξοδο του δεύτερου επιπέδου επιλέγει τη μέγιστη πιθανότητα και έχει σαν έξοδο τον αριθμό ένα (1) για την τάξη αυτή και μηδέν (0) για τις υπόλοιπες.

Τα δίκτυα αυτά είναι υπολογιστικά γρήγορα, αλλά για να δουλέψουν σωστά χρειάζεται μεγάλος αριθμός δειγμάτων εκπαίδευσης.

3.5 Σύνοψη των μεθόδων ταξινόμησης.

Από τη μελέτη των μεθόδων ταξινόμησης προκύπτει ότι οι περισσότερες είναι κατά μια έννοια καθοδηγούμενες, αφού χρειάζεται να προσδιοριστεί ο αριθμός των ομάδων που υπάρχει στα δεδομένα.

Ο αλγόριθμος κ-μέσων τιμών είναι ευαίσθητος στον αριθμό των ομάδων και τα αποτελέσματα διαφέρουν ριζικά, όταν αυτός μεταβάλλεται. Όμως είναι υπολογιστικά γρήγορος και συνιστάται όταν πρόκειται για μεγάλο αριθμό δεδομένων. Τα δίκτυα αυτο-οργανωμένης απεικόνισης (*SOM*), όχι μόνο απαιτούν τον αριθμό των ομάδων, αλλά είναι υπολογιστικά πολύ αργά και χρειάζονται εκπαίδευση από γνωστά δείγματα. Η τελευταία προϋπόθεση αναιρείται, όταν ο αριθμός των δεδομένων είναι μικρός (οπότε δείγματα εκπαίδευσης θεωρούνται όλα τα δεδομένα) ή υπάρχει διαθέσιμος ισχυρότατος υπολογιστής. Πλεονέκτημα του αλγόριθμου είναι ότι οι ομάδες είναι τοπολογικά οργανωμένες. Συνήθως ορίζεται μεγάλος αριθμός ομάδων και στη συνέχεια αυτές ενώνονται. Η ασαφής αφαιρετική ταξινόμηση είναι αυτή που μπορεί να χαρακτηριστεί σχεδόν αυτόματη, αφού προσδιορίζει τον αριθμό των ομάδων αυτόματα. Όμως πρέπει να προσδιοριστεί η ακτίνα επιρροής κάθε ομάδας πράγμα που προϋποθέτει γνώση της δομής των δεδομένων.



Σχήμα 3.4. Το δίκτυο LVQ αποτελείται από τέσσερα επίπεδα: εισόδου, εξόδου και δυο ενδιάμεσα. Ένα ανταγωνιστικό επίπεδο και ένα γραμμικό.

Ένας πρακτικός τρόπος μελέτης της δομής των δεδομένων είναι τα διαγράμματα διασποράς (*scatter plots*). Πρόκειται για διαγράμματα όπου κάθε χαρακτηριστικό των

δεδομένων απεικονίζεται σε άξονα καρτεσιανού συστήματος συντεταγμένων (Lemke και Mankowski 2000). Στα διαγράμματα αυτά απεικονίζονται ουσιαστικά, οι ομάδες στα δεδομένα. Σε περίπτωση που τα χαρακτηριστικά είναι περισσότερα από τρία (3) αυτά απεικονίζονται ανά τρία (3).

Όσον αφορά τις μεθόδους καθοδηγούμενης ταξινόμησης: Η μέθοδος LVQ, ανάλογα με το σχεδιασμό και την εκπαίδευση, μπορεί να ταξινομήσει τα δεδομένα σε προκαθορισμένο αριθμό ομάδων διατηρώντας την τοπολογική συνέχεια. Τα δίκτυα πιθανότητας (*PNN*) είναι υπολογιστικά γρήγορα, αλλά απαιτούν μεγάλο αριθμό δειγμάτων εκπαίδευσης, για να λειτουργήσουν σωστά.

ΚΕΦΑΛΑΙΟ 4**ΕΡΜΗΝΕΙΑ ΕΡΓΑΣΤΗΡΙΑΚΩΝ ΔΕΔΟΜΕΝΩΝ ΓΕΩΡΑΝΤΑΡ****4.1 Εισαγωγή.**

Στόχος της παρούσας εργασίας είναι η ανάπτυξη μεθοδολογίας για την ερμηνεία δεδομένων γεωραντάρ με τη χρήση μεθόδων ταξινόμησης. Για την επίτευξη του στόχου αυτού δημιουργήθηκαν σε περιβάλλον Matlab αλγόριθμοι επεξεργασίας δεδομένων γεωραντάρ (Σπανουδάκης 2002), όπως φίλτρα, αλγόριθμοι υπολογισμού των στιγμιαίων και γεωμετρικών χαρακτηριστικών των δεδομένων και εργαλεία της πολυδιάστατης στατιστικής, όπως ο μετασχηματισμός των κυρίων συνιστωσών. Τέλος χρησιμοποιήθηκαν αλγόριθμοι αυτόματης και καθοδηγούμενης ταξινόμησης από τα πεδία της πολυδιάστατης στατιστικής, της ασαφούς λογικής και των νευρωνικών δικτύων.

Για την παρουσίαση της μεθοδολογίας χρησιμοποιούνται δεδομένα από ελεγχόμενο πείραμα (Sandbox Experiment). Η ερμηνεία των δεδομένων αυτών θα ήταν δυνατή και

χωρίς τη χρήση των μεθόδων ταξινόμησης. Τα δεδομένα αυτά είναι κατάλληλα για την αξιολόγηση των μεθόδων ταξινόμησης ως προς τη δυνατότητά τους να ξεχωρίσουν τα θαμμένα αντικείμενα από το περιβάλλον τους αλλά και μεταξύ τους.

Παράλληλα, γίνεται παρουσίαση των δυνατοτήτων του λογισμικού GPR-Pro το οποίο κατασκευάστηκε για τις ανάγκες της μεταπτυχιακής εργασίας (Σπανουδάκης 2002) αλλά και της παρούσης διδακτορικής διατριβής.

4.2 Περιγραφή του ελεγχόμενου πειράματος.

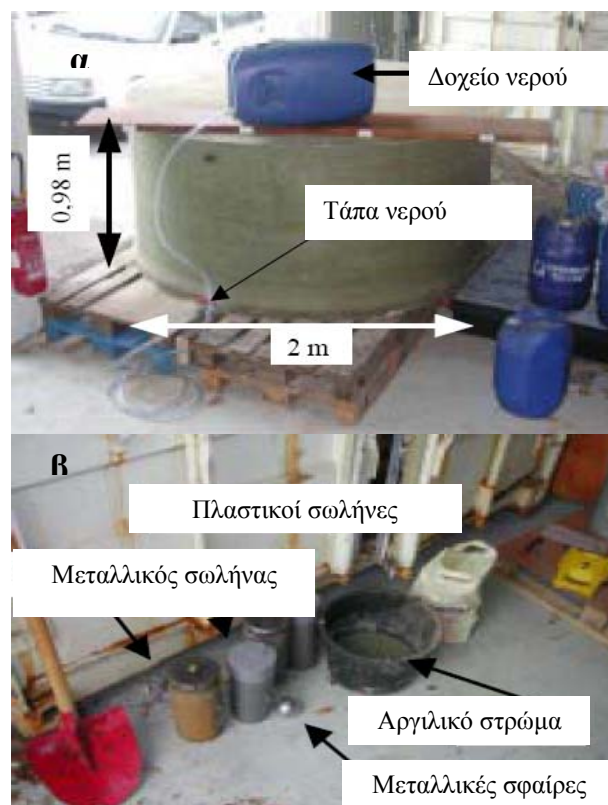
Το ελεγχόμενο πείραμα πραγματοποιήθηκε στο πανεπιστήμιο του Στρασβούργου στα πλαίσια του ερευνητικού προγράμματος HYGEIA. Χρησιμοποιήθηκε κυλινδρικό πλαστικό δοχείο διαμέτρου 2m και ύψους 0,98m το οποίο πληρώθηκε με κόκκους άμμου διαμέτρου 0.3-0.4mm. Στο κάτω μέρος του δοχείου υπήρχε στρόφιγγα για την εισαγωγή και εξαγωγή νερού στο δοχείο (σχήμα 4.1α) Μέσα στην άμμο τοποθετήθηκαν διάφορα αντικείμενα (σχήμα 4.1β) όπως ένας πλαστικός σωλήνας πληρωμένος με νερό (WPVC), ένας κενός πλαστικός σωλήνας (APVC), ένας σιδερένιος σωλήνας (Steel), τρεις μεταλλικές σφαίρες (P1, P2, P3) καθώς και ένα αργιλικό στρώμα (A). Οι διαστάσεις και το βάθος ταφής των αντικειμένων παρουσιάζονται στον πίνακα 4.1 (Loeffler και Bano 2004).

Οι μετρήσεις (σχήμα 4.1γ) πραγματοποιήθηκαν με το γεωραντάρ Pulse Ekko1000 (Sensors and Software). Μετά από πειράματα αποφασίστηκε η χρήση κεραιών συχνότητας εκπομπής 1200 MHz, διότι παρέχουν την καλύτερη ανάλυση καθώς και ικανό βάθος διείσδυσης, ώστε να καταγραφεί η ανάκλαση από τον πυθμένα του δοχείου, τουλάχιστον στην περίπτωση που η άμμος ήταν στεγνή. Διασκοπήθηκαν 71 παράλληλες γραμμές μελέτης με βήμα δειγματοληψίας 2 cm και ισοαπόσταση γραμμών ίση με 2 cm (σχήμα 4.2). Επίσης, πραγματοποιήθηκαν μετρήσεις κατά μήκος τριών κάθετων γραμμών μελέτης. Η πρώτη ομάδα δεδομένων συλλέχτηκε, όταν η άμμος ήταν στεγνή. Έπειτα εισήχθησαν 340lit νερού και σχηματίστηκε υδροφόρος ορίζοντας σε βάθος 72 cm και έπειτα άλλα 240lit, οπότε ο υδροφόρος ανέβηκε στα 48 cm. Για κάθε βάθος του υδροφόρου πραγματοποιήθηκαν οι ίδιες μετρήσεις που είχαν στόχο την εκτίμηση της περιεκτικότητας της άμμου σε νερό. Για την παρουσίαση της μεθοδολογίας επιλέχθηκε η

γραμμή μελέτης P36 (σχήμα 4.2) η οποία διασχίζει το δοχείο ακριβώς πάνω από το κέντρο των τριών σωλήνων την περίοδο πριν εκχυθεί οποιαδήποτε ποσότητα νερού ή πετρελαιοειδών. Στόχος είναι η αξιολόγηση των μεθόδων ταξινόμησης ως προς τη δυνατότητα να ξεχωρίσουν τους σωλήνες από το περιβάλλον τους, αλλά και μεταξύ τους.

Πίνακας 4.1. Διαστάσεις και βάθος ταφής των αντικειμένων.

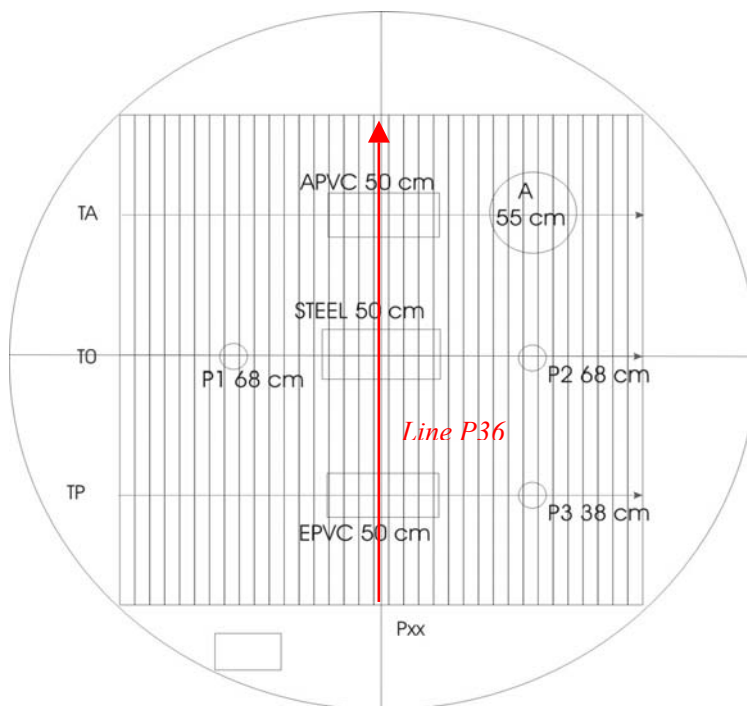
Αντικείμενα	Μήκος (mm)	Διάμετρος (mm)	Βάθος (cm)
Σιδερένιος σωλήνας (Steel)	316	140	50
Πλαστικοί σωλήνες (APVC, WPVC)	298	126	50
Μεταλλικές μπάλες (P1,P2, P3)	-	74	38, 68
Αργιλικό στρώμα	80 (πάχος)	230	55



Σχήμα 4.1. (α) Το δοχείο μέσα στο οποίο πραγματοποιήθηκε το ελεγχόμενο πείραμα και (β) τα αντικείμενα που τοποθετήθηκαν σε αυτό (Loeffler και Bano 2004).



Σχήμα 4.1 (συνέχεια). (γ) Συλλογή των δεδομένων του πειράματος.



Σχήμα 4.2. Κάτοψη του δοχείου. Διακρίνονται τα αντικείμενα, το βάθος ταφής τους καθώς και οι γραμμές μελέτης (Loeffler και Bano 2004).

4.3 Επεξεργασία των δεδομένων.

Η αρχική επεξεργασία των δεδομένων στόχευε στην απαλοιφή του θορύβου και την ενίσχυση των ασθενών καταγραφών. Αρχικά, εφαρμόστηκε το φίλτρο απαλοιφής της επαγωγικής συνιστώσας (*De-wow*) και έπειτα χωροθέτηση *Gazdag* (Gazdag 1978) για την ελαχιστοποίηση των περιθλώμενων κυμάτων από στόχους μικρών διαστάσεων. Στα πλαίσια της παρούσας εργασίας υλοποιήθηκε και εφαρμόστηκε ο αλγόριθμος χωροθέτησης, στο πεδίο συχνότητας – κυματαριθμού, του *Gazdag* κατά τον οποίο γίνονται οι παρακάτω παραδοχές: α) πομπός και δέκτης βρίσκονται στην ίδια θέση, β) η πρόσπτωση των H/M κυμάτων είναι κάθετη στα αντικείμενα, γ) το μέσον διάδοσης είναι οριζόντια στρωματομένο και δ) η προς τα κάτω συνέχεια του κυματικού πεδίου εφαρμόζεται διαδοχικά σε όλα τα βάθη χρησιμοποιώντας συνάρτηση της ταχύτητας με το βάθος και το κυματικό πεδίο στο υπερκείμενο στρώμα (Yilmaz 1987). Η ταχύτητα διάδοσης των H/M που χρησιμοποιήθηκε για τη χωροθέτηση ήταν ίση με 12.5 cm/nsec και υπολογίστηκε από τα περιθλώμενα κύματα.

Έπειτα εφαρμόστηκε το φίλτρο αντιστάθμισης της γεωμετρικής διασποράς του σήματος (*SEC*) και το φίλτρο κινητού μέσου όρου για την εξομάλυνση των δεδομένων. Στο σχήμα 4.3 φαίνονται τα δεδομένα μετά τη χωροθέτηση καθώς και τμήμα της καταγραφής από 6-11 nsec (σχήμα 4.3b), η οποία επιλέχθηκε για ταξινόμηση λόγω οικονομίας στον υπολογιστικό χρόνο.

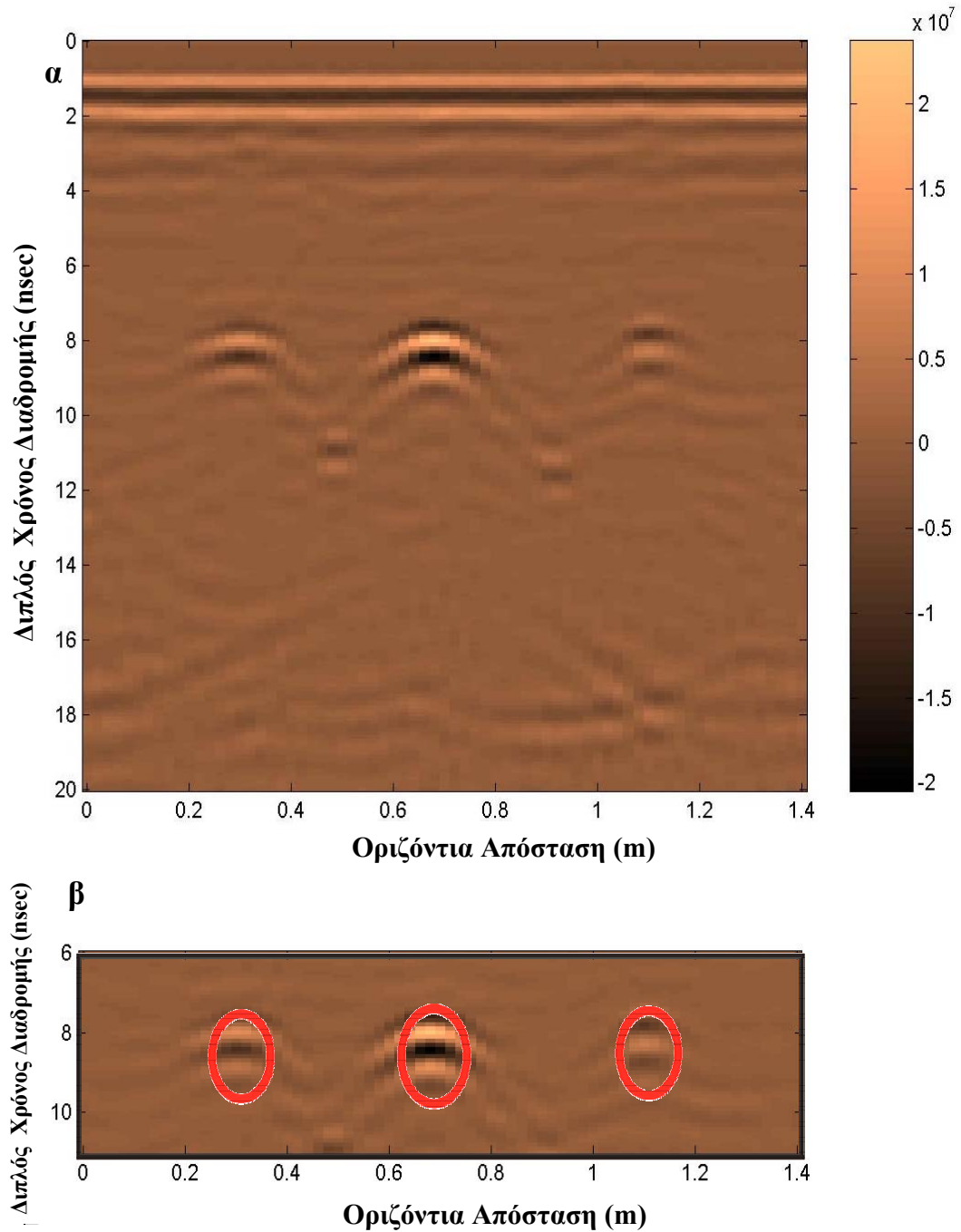
Για χρόνο καταγραφής 0-6 nsec περίπου παρατηρούνται κάποιες ασθενείς ανακλάσεις, οι οποίες μάλλον οφείλονται στις διαδοχικές στρώσεις της άμμου κατά την εισαγωγή της στο δοχείο. Αφού στρώθηκε η άμμος τοποθετήθηκαν τα αντικείμενα. Ανακλώμενα και περιθλώμενα κύματα από τους σωλήνες διακρίνονται στα 7 nsec περίπου. Ο σωλήνας EPVC είναι τοποθετημένος σε οριζόντια απόσταση 0.3 m από την αρχή της γραμμής μελέτης, ο μεταλλικός σωλήνας στα 0.7 m ενώ ο σωλήνας APVC, στα 1.1 m. Τέλος, ο οριζόντιος ανακλαστήρας στην βάση του δοχείου καταγράφεται στα 17 nsec και είναι παραμορφωμένος επειδή τα θαμμένα αντικείμενα επηρεάζουν την ταχύτητα διάδοσης του ανακλώμενου από αυτόν κύματος. Επίσης, παρατηρούνται ανακλώμενα κύματα από τα τοιχώματα του δοχείου, που δεν απαλείφθηκαν από την χωροθέτηση.

Στο επόμενο στάδιο υπολογίστηκε μια σειρά μιγαδικών χαρακτηριστικών. Υπολογίστηκαν τέσσερα (4) στιγμιαία χαρακτηριστικά και δύο (2) γεωμετρικά. Το στιγμιαίο πλάτος (σχήμα 4.4α) παρουσιάζει υψηλές τιμές για τους τρεις σωλήνες και παρόμοιες τιμές για την υπόλοιπη καταγραφή. Σε κλίμακα dB όμως (σχήμα 4.4β) τονίζονται τόσο οι υψηλές όσο και οι μικρές τιμές του. Η στιγμιαία φάση (σχήμα 4.4γ) και η στιγμιαία συχνότητα (σχήμα 4.4δ) παρουσιάζουν λεπτομέρειες σχετικά με τα ανακλώμενα κύματα από τους σωλήνες, αλλά και την άμμο, ενώ η συνάφεια των δεδομένων (σχήμα 4.4ε) και η συνάφεια της στιγμιαίας φάσης (σχήμα 4.4στ) δε δίνουν ιδιαίτερα στοιχεία για την παρουσία των στόχων.

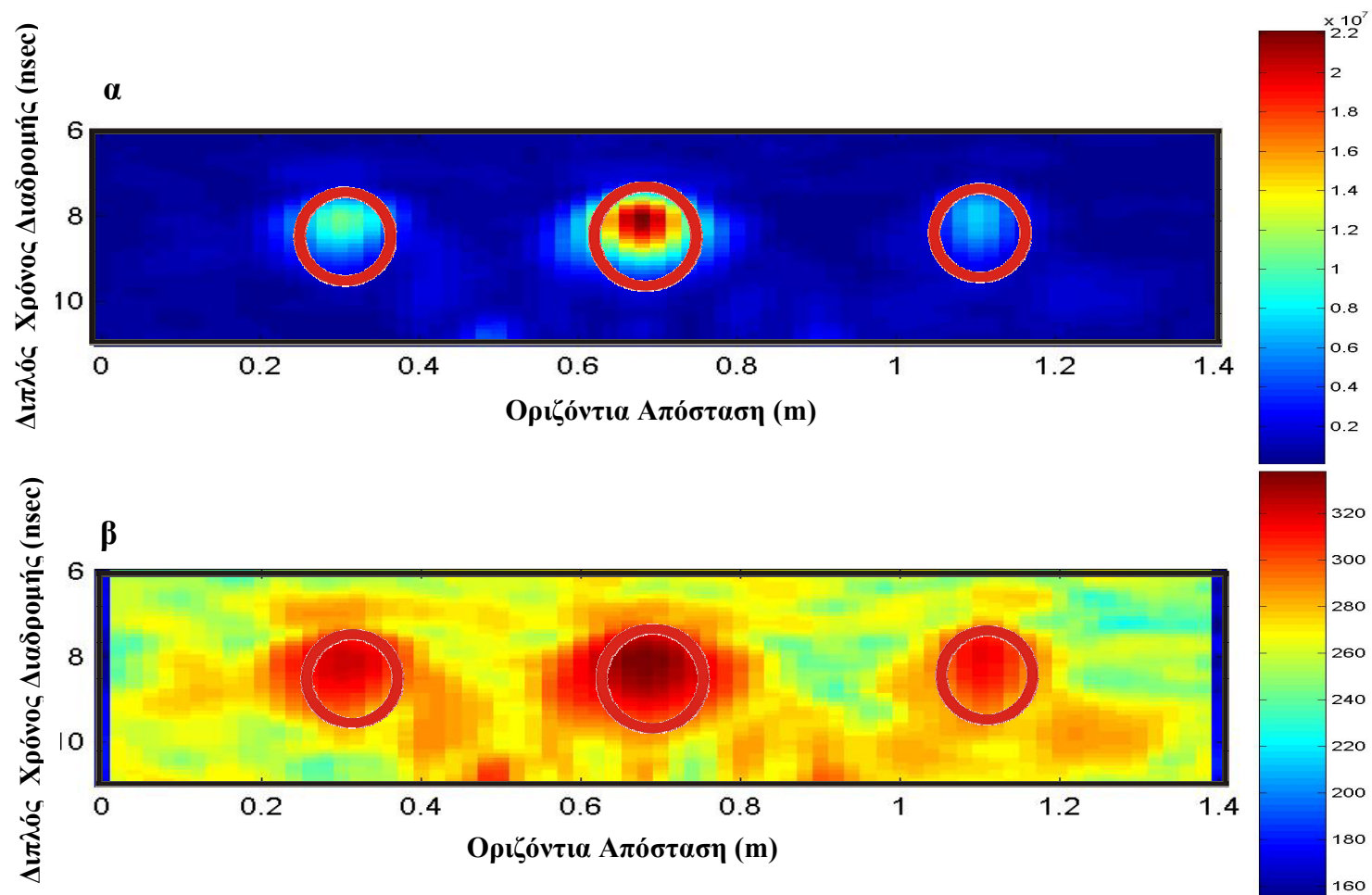
Στη συνέχεια εφαρμόστηκε η τεχνική της ανάλυσης κυρίων συνιστωσών με στόχο να περιοριστεί ο αριθμός των προς ταξινόμηση χαρακτηριστικών. Συγκεκριμένα, η ανάλυση πραγματοποιήθηκε σε ζευγάρια παρόμοιων χαρακτηριστικών. Για παράδειγμα, από το στιγμιαίο πλάτος και το στιγμιαίο πλάτος σε κλίμακα dB υπολογίστηκαν κύριες συνιστώσες εκ των οποίων η πρώτη συγκεντρώνει μεγάλο ποσοστό της πληροφορίας που περιείχαν από κοινού τα δύο χαρακτηριστικά (σχήμα 4.5α). Η ίδια τεχνική εφαρμόστηκε και στα ζευγάρια συνάφειας των αρχικών δεδομένων και της στιγμιαίας φάσης (σχήμα 4.5 β), καθώς και στιγμιαίας φάσης και στιγμιαίας συχνότητας (σχήμα 4.5γ).

Η συγκεκριμένη τεχνική ανάλυσης κυρίων συνιστωσών εφαρμόστηκε σε ζευγάρια χαρακτηριστικών διότι τα χαρακτηριστικά αυτά ανά δυο απεικονίζουν παρόμοιες ιδιότητες των δεδομένων. Σε επόμενα πειράματα (βλέπε, επόμενο κεφάλαιο) η τεχνική της ανάλυσης κυρίων συνιστωσών είτε δεν χρησιμοποιείται καθόλου, είτε εφαρμόζεται σε όλα τα χαρακτηριστικά των δεδομένων ταυτόχρονα.

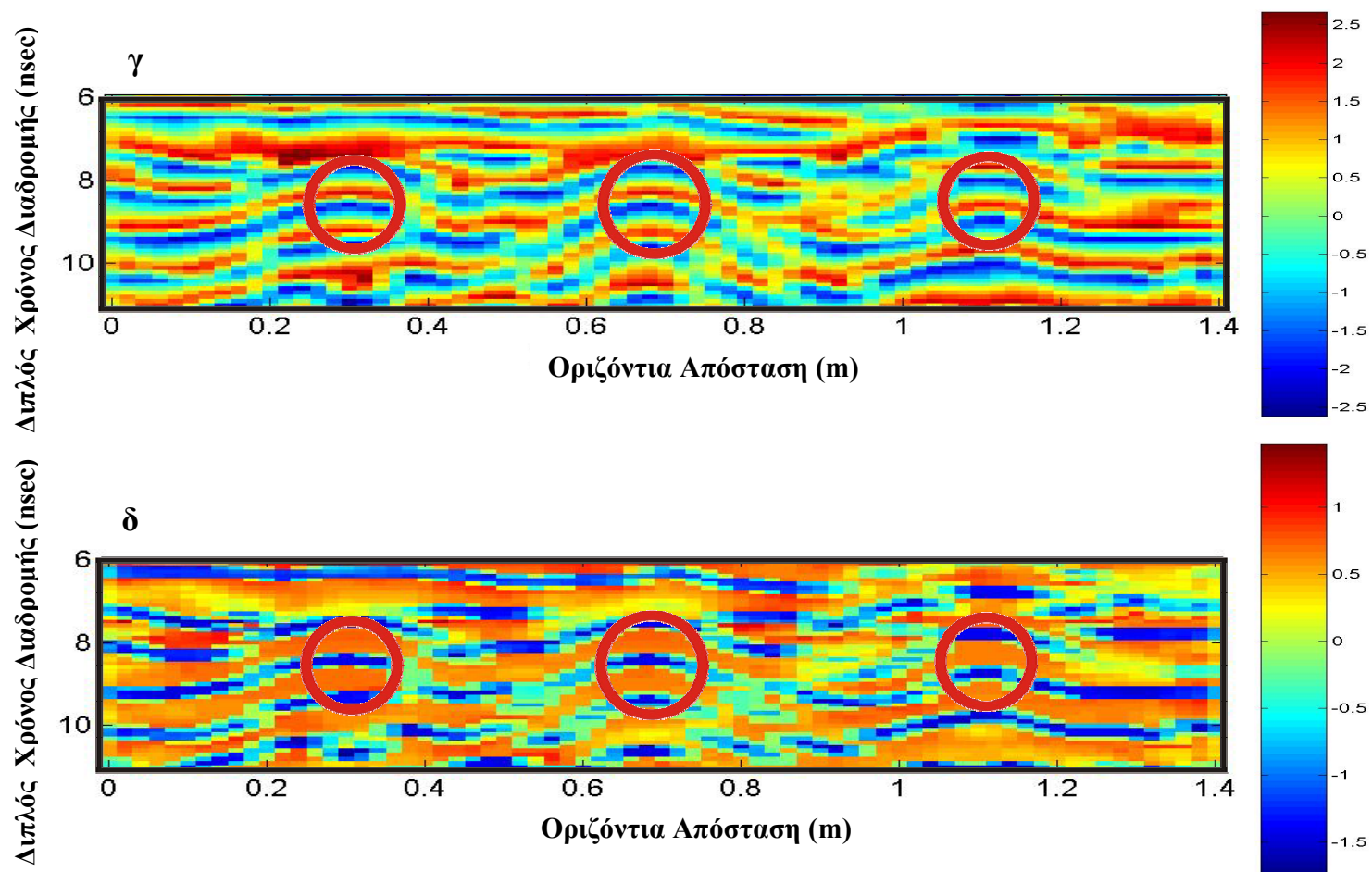
Στα σχήματα 4.6 - 4.7 φαίνονται λεπτομέρειες από το λογισμικό επεξεργασίας δεδομένων γεωραντάρ *GPR-Pro*, οι οποίες αφορούν το πρώτο στάδιο επεξεργασίας με τα φίλτρα, τη χωροθέτηση και την ενίσχυση, καθώς και τον υπολογισμό των χαρακτηριστικών των δεδομένων αντίστοιχα.



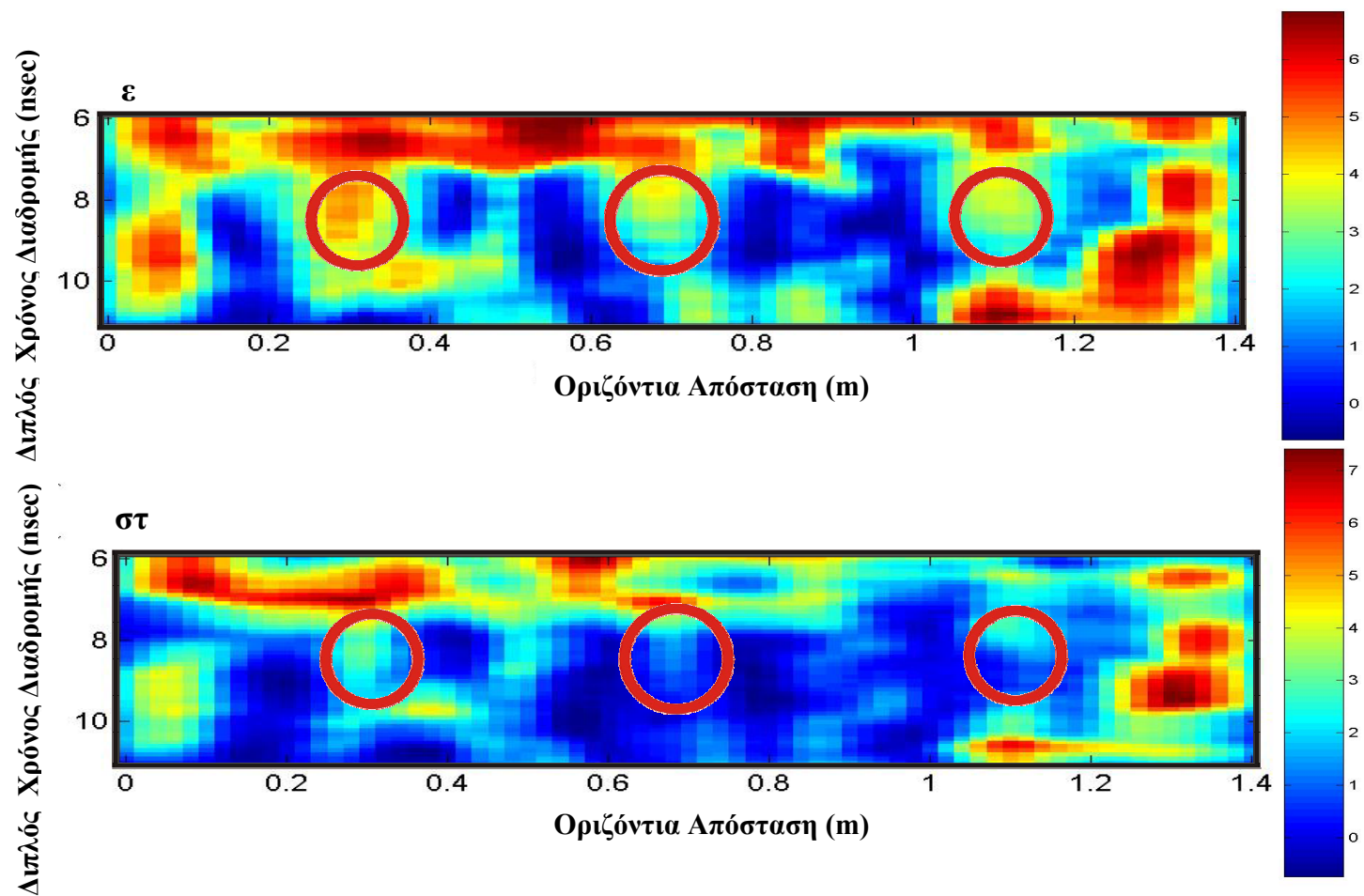
Σχήμα 4.3. (α) Τα δεδομένα της γραμμής P36 μετά από το πρώτο στάδιο επεξεργασίας, το οποίο περιλάμβανε την εφαρμογή ψηφιακών φίλτρων, χωροθέτηση και ενίσχυση. (β) Διακρίνεται λεπτομέρεια από την περιοχή ενδιαφέροντος (6-11 nsec).



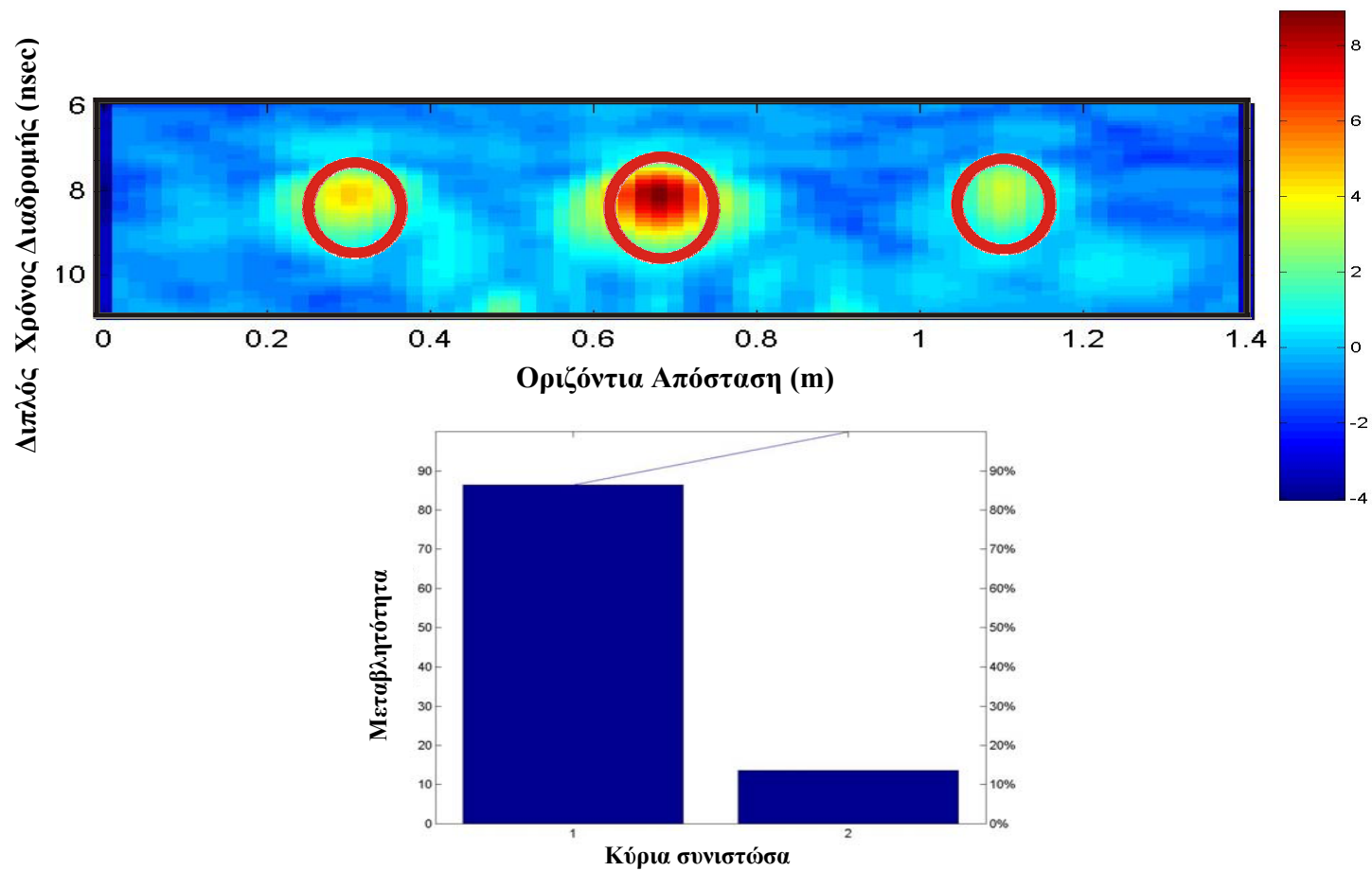
Σχήμα 4.4. Χαρακτηριστικά των δεδομένων του γεωραντάρ για την περιοχή 6-11 nsec. (α) Στιγμιαίο πλάτος, (β) Στιγμιαίο πλάτος σε κλίμακα dB.



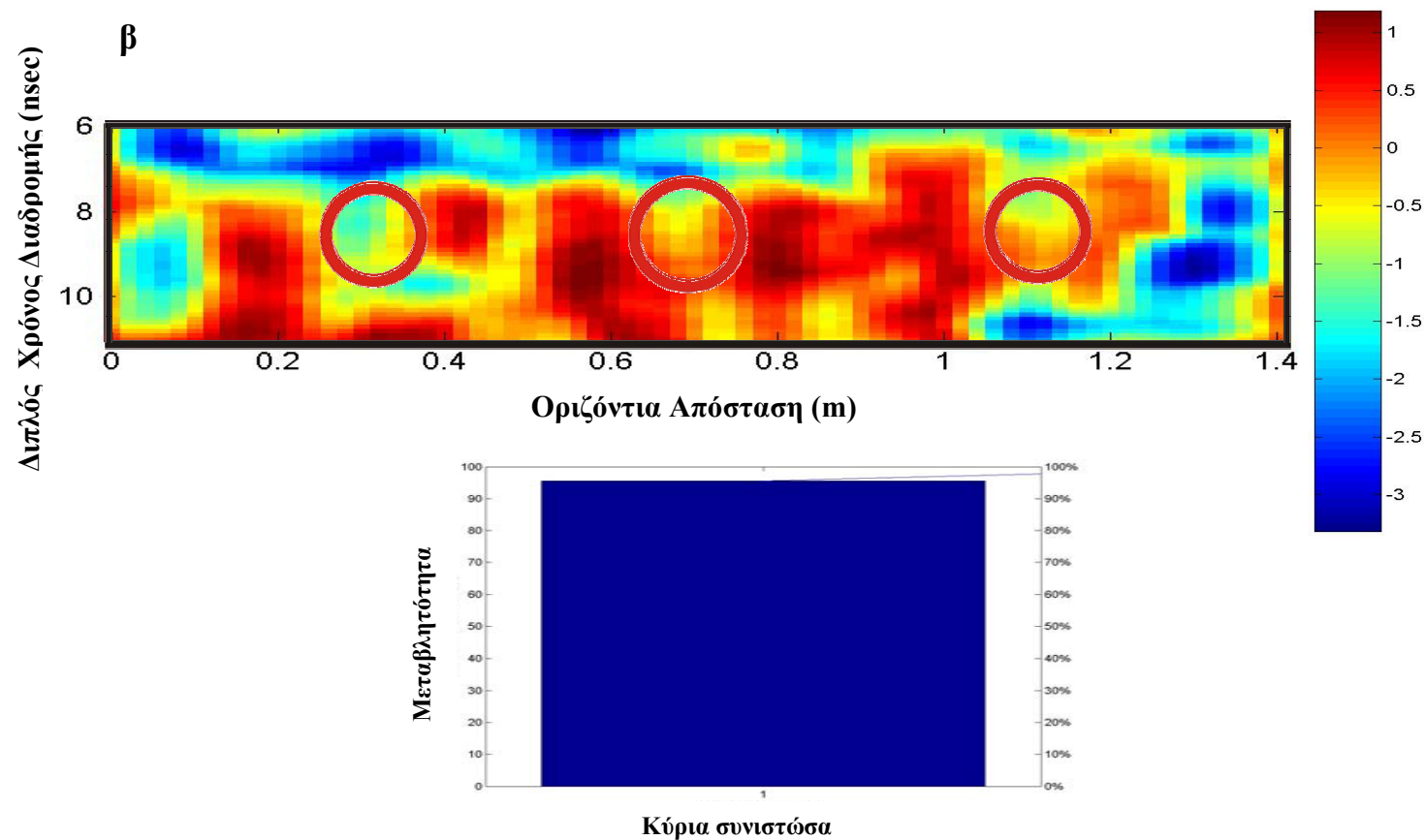
Σχήμα 4.4 (συνέχεια). Χαρακτηριστικά των δεδομένων του γεωραντάρ για την περιοχή 6-11 nsec. (γ) Στιγμιαία φάση, (δ) Στιγμιαία συχνότητα.



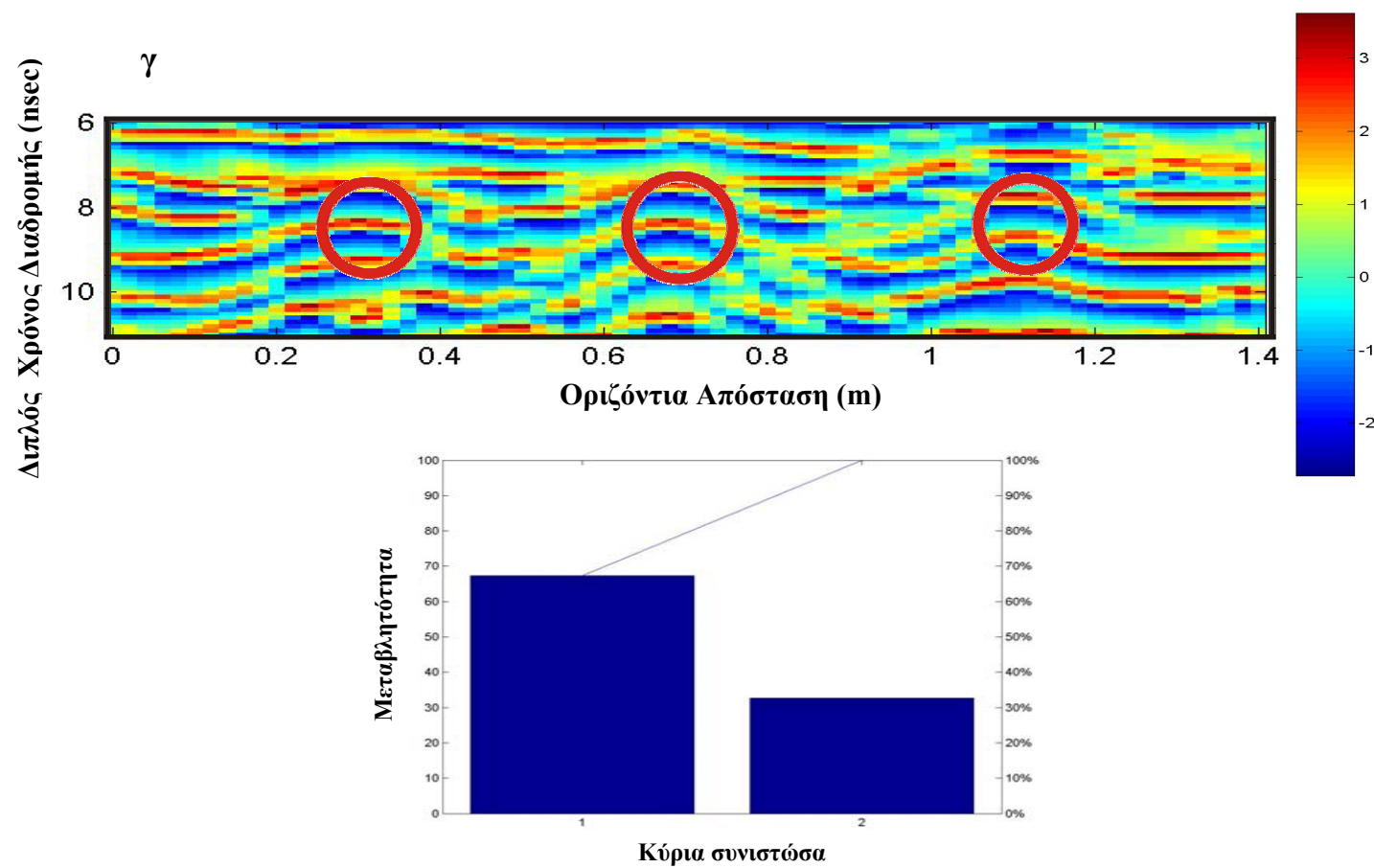
Σχήμα 4.4 (συνέχεια). Χαρακτηριστικά των δεδομένων του γεωραντάρ για την περιοχή 6-11 nsec. (ϵ) Συνάφεια, ($\sigma\tau$) Συνάφεια της στιγμιαίας φάσης.



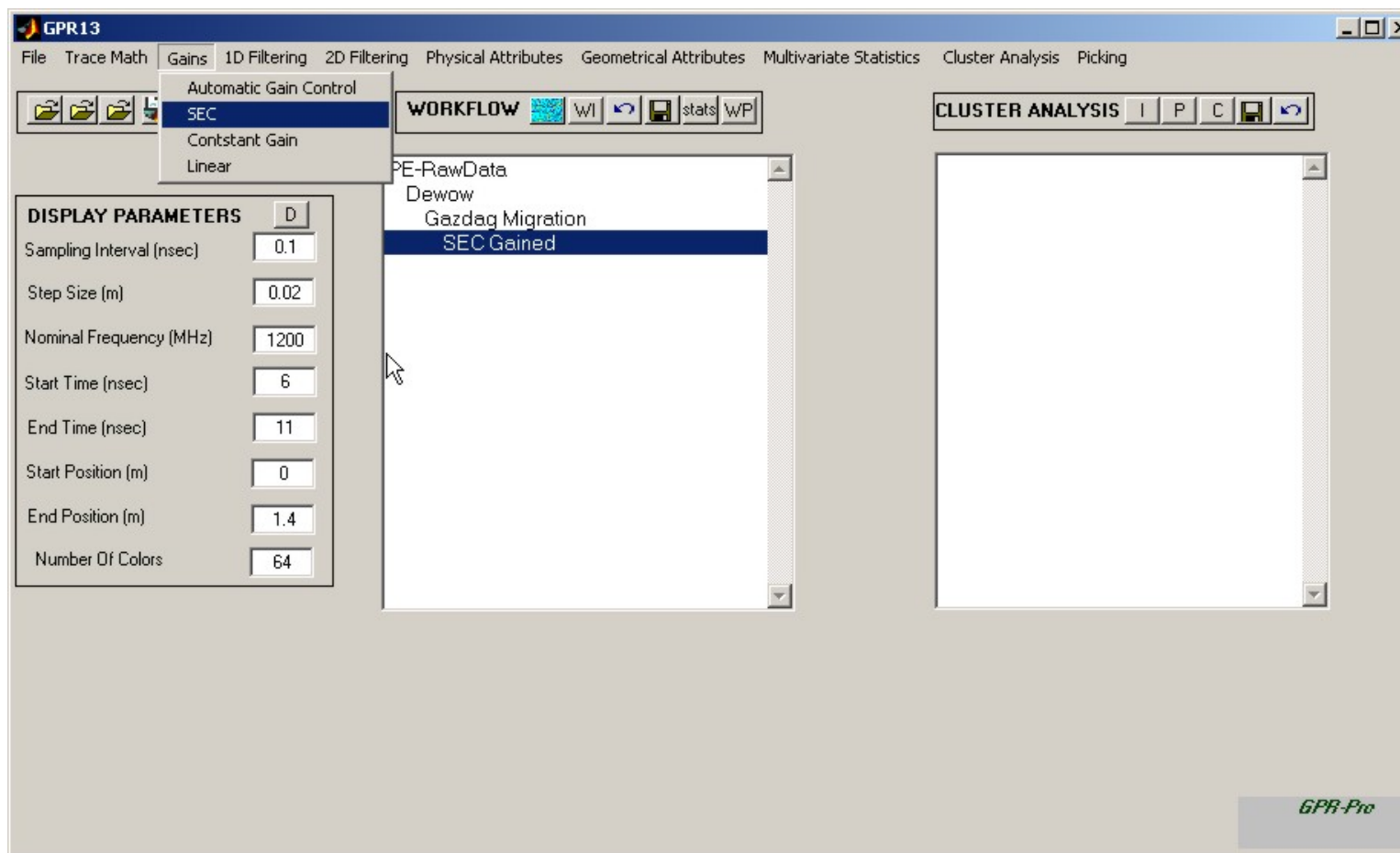
Σχήμα 4.5. Η πρώτη κύρια συνιστώσα, η οποία προέκυψε από την ανάλυση του στιγμιαίου πλάτους και του στιγμιαίου πλάτους σε κλίμακα dB περιέχει το 85% της πληροφορίας των δυο.



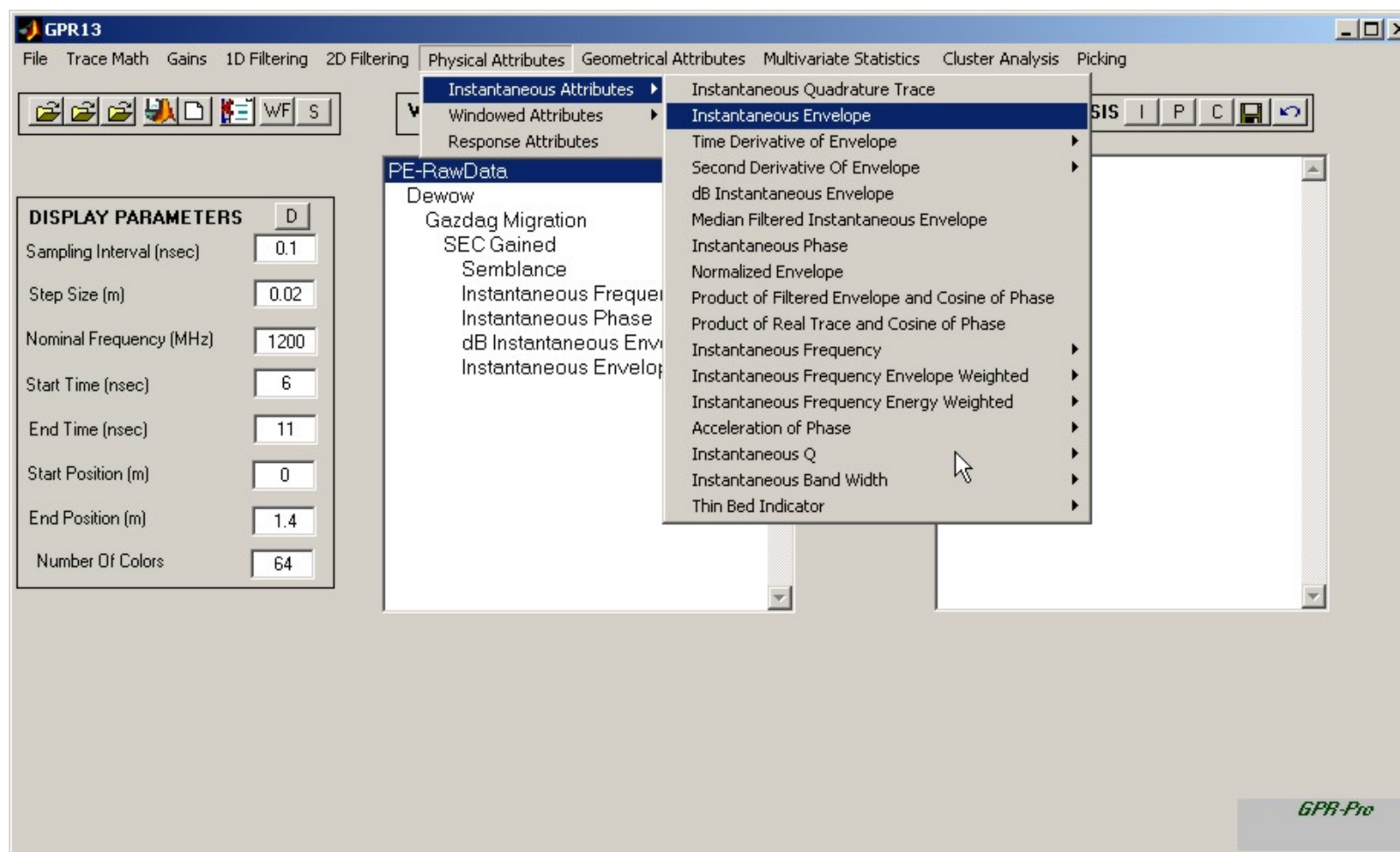
Σχήμα 4.5 (συνέχεια). Η πρώτη κύρια συνιστώσα η οποία προέκυψε από την ανάλυση της συνάφειας και της συνάφειας της στιγμιαίας φάσης περιέχει το 99% της πληροφορίας των δυο.



Σχήμα 4.5 (συνέχεια). Η πρώτη κύρια συνιστώσα η οποία προέκυψε από την ανάλυση της στιγμιαίας φάσης και της στιγμιαίας συχνότητας περιέχει το 70% της πληροφορίας των δυο.



Σχήμα 4.6. Τυπική απεικόνιση της διεπιφάνειας του λογισμικού *GPR-Pro*. Στο κέντρο διακρίνονται οι αλγόριθμοι που εφαρμόστηκαν κατά το πρώτο στάδιο επεξεργασίας των δεδομένων.



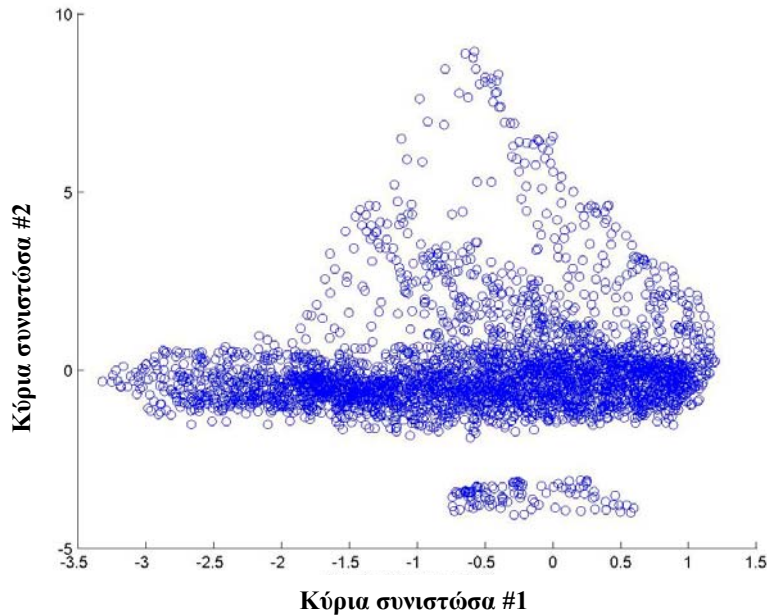
Σχήμα 4.7. Λογισμικό *GPR-Pro*. Υπολογισμός των χαρακτηριστικών των δεδομένων.

4.4 Αυτόματη ταξινόμηση των δεδομένων.

Τα νέα χαρακτηριστικά από το προηγούμενο στάδιο, δηλαδή οι κύριες συνιστώσες, πριν από την ταξινόμηση κανονικοποιούνται δηλαδή αφαιρείται η μέση τιμή και διαιρούνται με την τυπική απόκλιση. Παρακάτω παρουσιάζεται η ταξινόμηση των δεδομένων με διάφορες μεθοδολογίες. Στο σχήμα (4.8) παρουσιάζεται το διάγραμμα διασποράς δυο κυρίων συνιστωσών. Ο οριζόντιος άξονας αφορά στην κύρια συνιστώσα από το ζευγάρι στιγμιαίο πλάτος και στιγμιαίο πλάτος σε κλίμακα dB, ενώ ο κατακόρυφος την κύρια συνιστώσα από το ζευγάρι συνάφειας των αρχικών δεδομένων και συνάφειας της στιγμιαίας φάσης. Παρατηρείται μια μεμονωμένη ομάδα δεδομένων στο κάτω μέρος του διαγράμματος. Επίσης, παρατηρείται μια περιοχή με μεγάλη πυκνότητα δεδομένων για τιμή της τρίτης κύριας συνιστώσας περίπου ίση με 0.

Όπως έχει ήδη αναφερθεί, η μέθοδος του δικτύου αυτο-οργανωμένης απεικόνισης, αν και γρήγορη υπολογιστικά, χρειάζεται αρκετές επαναλήψεις για να συγκλίνει, σε αντίθεση με τη μέθοδο των κ-μέσων τιμών. Αρχικά λοιπόν για την εκπαίδευση του δικτύου αυτο-οργανωμένης απεικόνισης πραγματοποιείται μια αρχική ταξινόμηση στα δυο και έπειτα στα τρία (3) χαρακτηριστικά, με τη μέθοδο των κ-μέσων τιμών σε εκατό (100) ομάδες (λεπτομέρειες από το λογισμικό GPR-Pro στο σχήμα 4.9). Τα εκατό (100) κέντρα συσσώρευσης που προκύπτουν σε κάθε περίπτωση τροφοδοτούνται στο δίκτυο σαν δείγματα εκπαίδευσης και ταξινομούνται σε είκοσι (20) τάξεις αντίστοιχα (λεπτομέρειες από το λογισμικό GPR-Pro στο σχήμα 4.10). Κατόπιν τα δεδομένα ταξινομούνται σε είκοσι ομάδες.

Με τη χρήση δυο κυρίων συνιστωσών (σχήμα 4.11α), οι τρεις διαφορετικοί στόχοι διαχωρίζονται καλά και φαίνεται ότι ανήκουν στις ομάδες 8-12. Η διαφορά στα χρώματα είναι μικρή, που σημαίνει ότι βρίσκονται κοντά στο διάγραμμα διασποράς (4.11β). Στο διάγραμμα διασποράς φαίνεται η τοπολογική συνέχεια των ομάδων, όπου γειτονικές ομάδες έχουν και παρόμοιο χρώμα. Το υπόβαθρο ταξινομείται στις ομάδες 1-7, 11-20 (τόνοι του κόκκινου και του μπλε). Παρατηρείται ότι οι ομάδες που αφορούν στο υπόβαθρο βρίσκονται κοντά στο διάγραμμα διασποράς και στην περιοχή με μεγάλη πυκνότητα δεδομένων, ενώ τα δεδομένα που σχετίζονται με τους στόχους δεν έχουν μεγάλη πυκνότητα.



Σχήμα 4.8. Διάγραμμα διασποράς των δεδομένων. Οριζόντιος άξονας: κύρια συνιστώσα από τα στιγμιαία πλάτη. Κατακόρυφος άξονας: κύρια συνιστώσα από τις συνάφειες. Οι αρνητικές τιμές και στους δυο άξονες προκύπτουν λόγω της κανονικοποίησης των δεδομένων καθώς και λόγω της αλλαγής συστήματος συντεταγμένων.

Όταν χρησιμοποιούνται και οι τρεις κύριες συνιστώσες (σχήμα 4.12) οι στόχοι δε διαχωρίζονται και τόσο καλά μεταξύ τους, αφού φαίνεται να ανήκουν στις ομάδες 18-20 (κόκκινο χρώμα). Το υπόβαθρο ταξινομείται στις ομάδες 1-12 (μπλε-πράσινο), ενώ ανάμεσα από τους στόχους τα δεδομένα ταξινομούνται στις ομάδες 9-17 (τόνοι του κίτρινου-πορτοκαλί). Χρησιμοποιώντας όμως και τις τρεις συνιστώσες εξασφαλίζεται χρήσιμη λεπτομέρεια που αφορά στη φάση των καταγραφών.

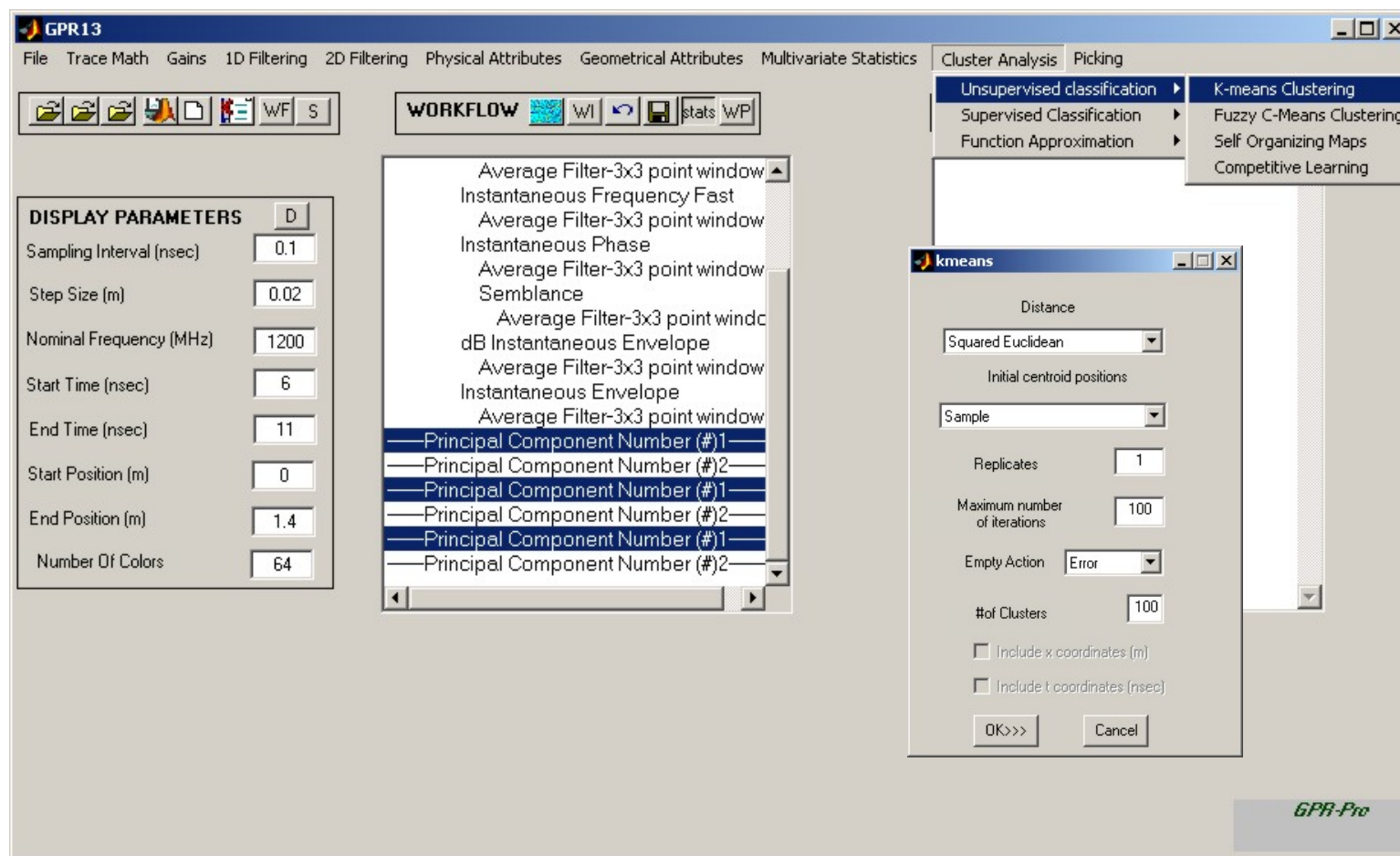
Μια άλλη προσέγγιση συνίσταται στην επιλογή, τυχαίων ή μη, δειγμάτων από τα δεδομένα με σκοπό την εκπαίδευση του αλγόριθμου. Σε αυτή την περίπτωση επιλέχθηκαν δείγματα που αντιστοιχούν σε ανακλώμενα κύματα από τους τρεις σωλήνες και το υπόβαθρο (σχήμα 4.13). Στο σχήμα 4.14 φαίνεται λεπτομέρεια από το λογισμικό *GPR-Pro* για την επιλογή δειγμάτων από τα δεδομένα κατά την εκπαίδευση του δικτύου αυτό-οργανωμένης απεικόνισης.

Σε αυτή την περίπτωση οι τρεις σωλήνες φαίνεται να ανήκουν σε διαφορετικές ομάδες (τόνοι του κόκκινου-πορτοκαλί), ενώ το υπόβαθρο ταξινομείται στις ομάδες που απεικονίζονται σε τόνους του μπλε-πράσινου (σχήμα 4.15 α). Το διάγραμμα διασποράς (σχήμα 4.15 β) είναι ομοιόμορφο όσον αφορά στην κατανομή των ομάδων. Το υπόβαθρο περιέχει καταγραφές στην περιοχή με τη μεγάλη πυκνότητα δεδομένων, ενώ οι ομάδες είναι τόσο κοντά (παραπλήσια χρώματα), που δε διακρίνεται η μια από την άλλη. Το αποτέλεσμα αυτό είναι το ιδανικότερο μέχρι στιγμής, αφού οι στόχοι διαχωρίζονται πλήρως μεταξύ τους και το υπόβαθρο φαίνεται σχεδόν ομοιογενές.

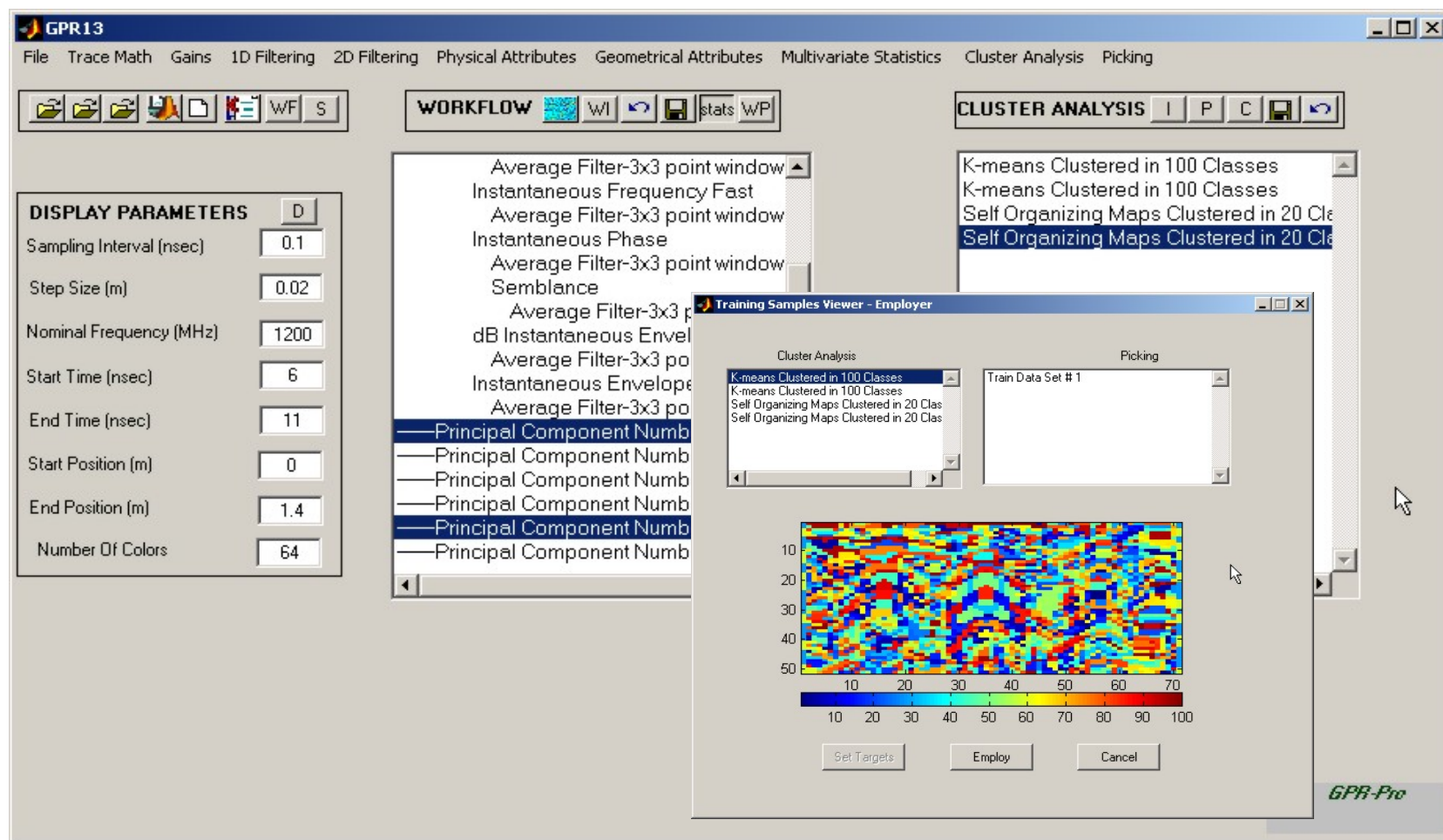
4.5 Καθοδηγούμενη ταξινόμηση των δεδομένων με το δίκτυο LVQ.

Σε περιπτώσεις που υπάρχουν δείγματα στα δεδομένα που αντιστοιχούν σε πρόσθετες πληροφορίες, εφαρμόζεται καθοδηγούμενη ταξινόμηση. Οι πρόσθετες πληροφορίες έρχονται από την εμπειρία, γεωλογικές ενδείξεις, γεωτρήσεις ή και άλλες γεωφυσικές μεθόδους. Το δίκτυο εκπαιδεύεται βάσει αυτών των πληροφοριών και ταξινομεί το σύνολο των δεδομένων. Τα δείγματα που επιλέχτηκαν ήταν τα ίδια με το προηγούμενο παράδειγμα (σχήμα 4.13).

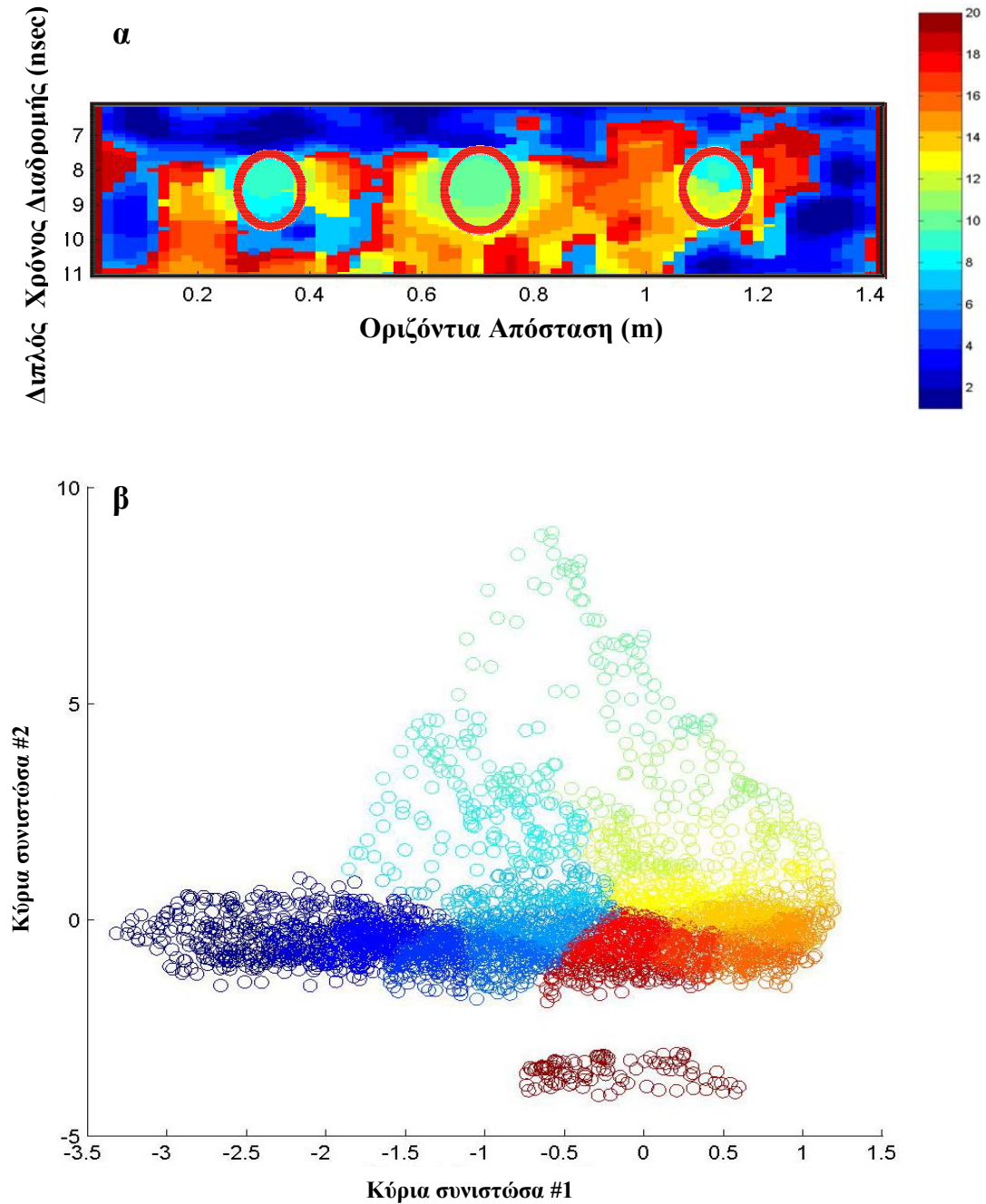
Όπως φαίνεται στο σχήμα 4.16α, οι τρεις σωλήνες ταξινομούνται σωστά σε διαφορετικές ομάδες και διαχωρίζονται από το περιβάλλον τους. Φυσικά υπάρχουν και καταγραφές οι οποίες ταξινομούνται λάθος. Στο τρισδιάστατο διάγραμμα διασποράς (σχήμα 4.16 β) οι καταγραφές που σχετίζονται με τους στόχους έχουν υψηλές τιμές για την κύρια συνιστώσα τρία (3), ενώ οι καταγραφές από την άμμο είναι γενικά τοποθετημένες στο κέντρο των αξόνων, δηλαδή κοντά στο μηδέν (0).



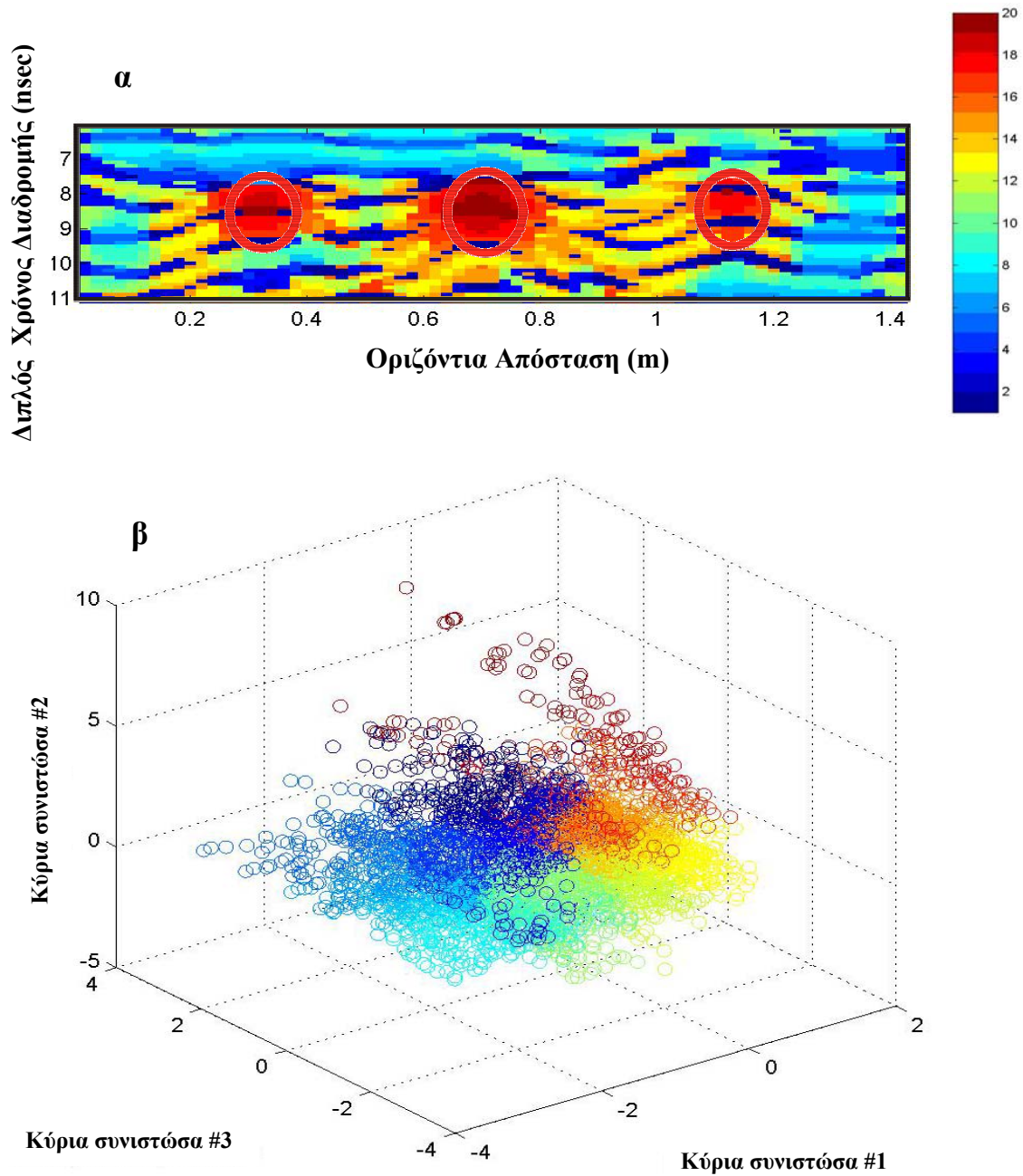
Σχήμα 4.9. Λεπτομέρεια του λογισμικού για την ταξινόμηση με τη μέθοδο κ-μέσων τιμών. Έχουν επιλεγεί οι τρεις κύριες συνιστώσες. Ζητείται από τον αλγόριθμο να ταξινομήσει τα δεδομένα σε 100 τάξεις.



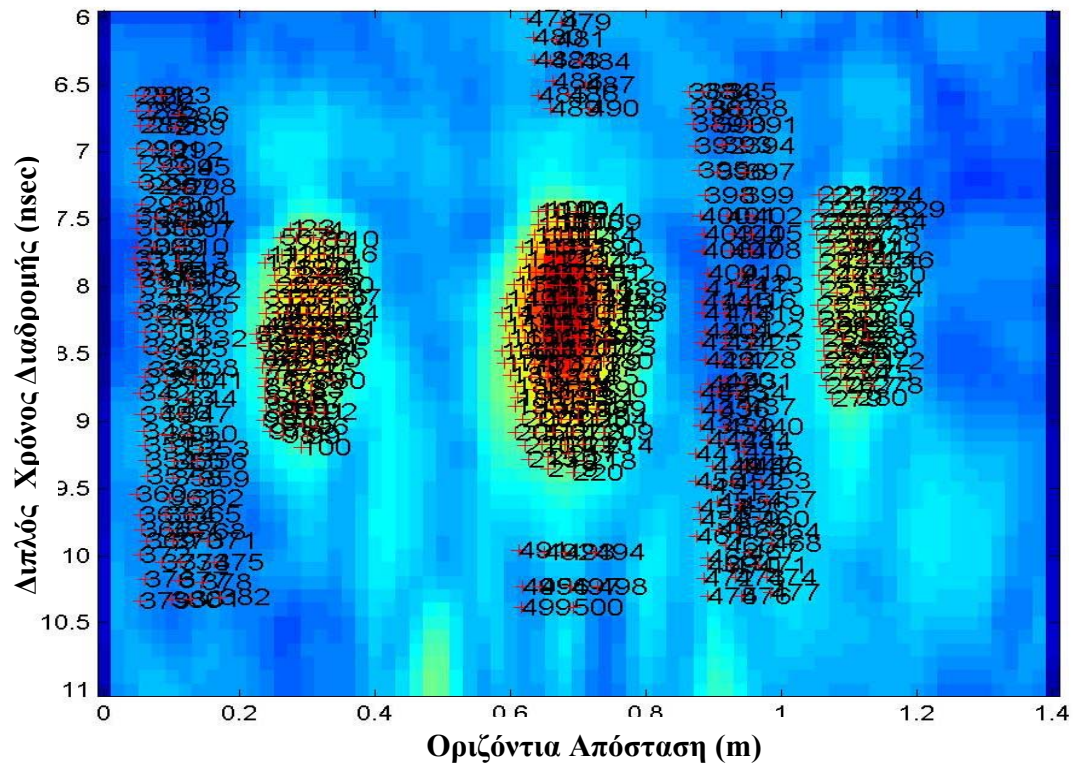
Σχήμα 4.10. Λεπτομέρεια του λογισμικού για την ταξινόμηση με τη μέθοδο δικτύου αυτο-οργανωμένης απεικόνισης. Έχουν επιλεγεί οι δυο κύριες συνιστώσες. Ζητείται από τον αλγόριθμο να πάρει ως δείγματα εκπαίδευσης τα 100 κέντρα που προέκυψαν από την ταξινόμηση με τη μέθοδο κ-μέσων τιμών.



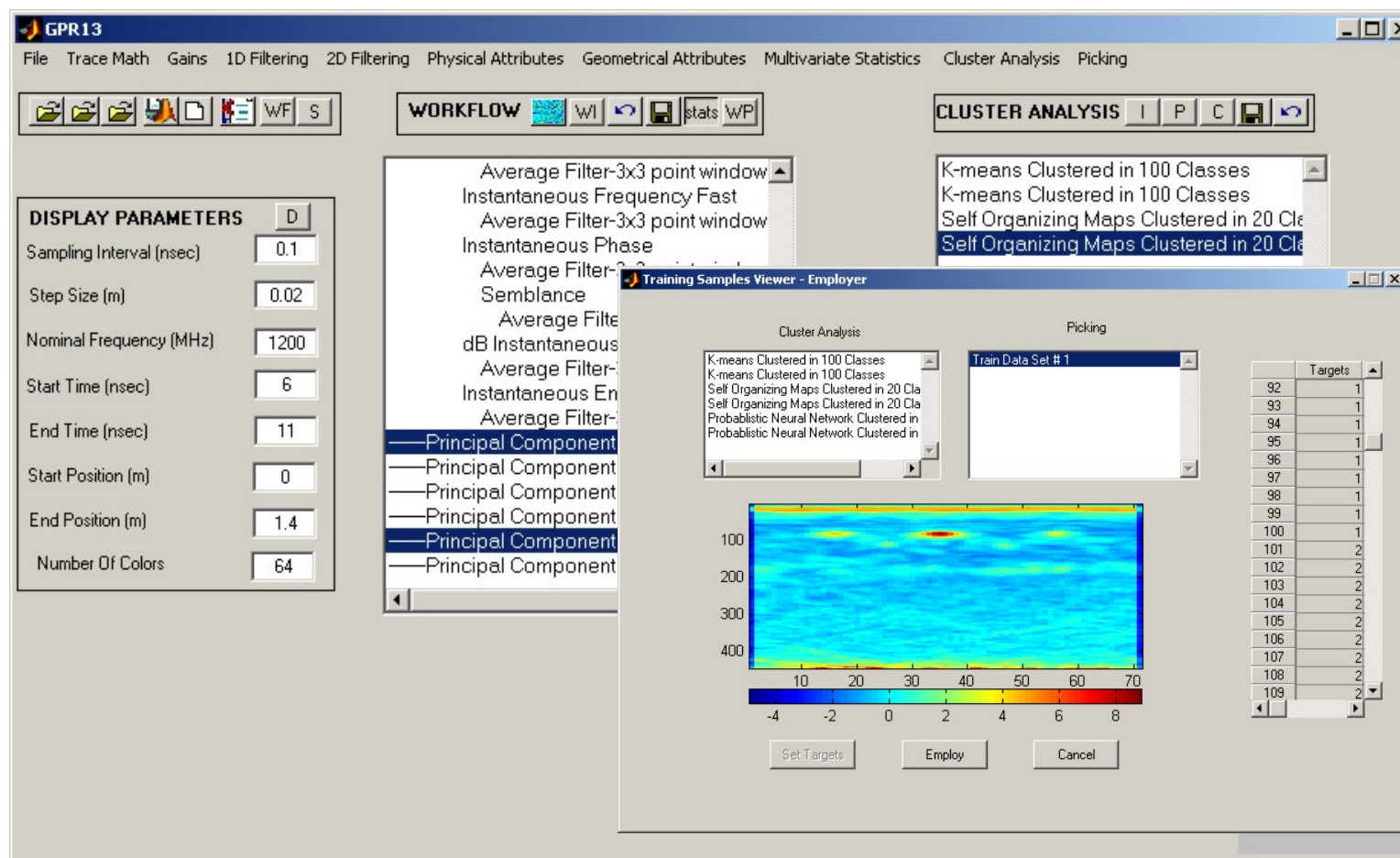
Σχήμα 4.11. (α) Το αποτέλεσμα της ταξινόμησης και το διάγραμμα διασποράς για το πείραμα με τη χρήση της μεθόδου των κ-μέσων τιμών (ταξινόμηση σε 100 τάξεις) και (β) του δικτύου SOM (ταξινόμηση σε 20 τάξεις) με τη χρήση δυο κυρίων συνιστωσών.



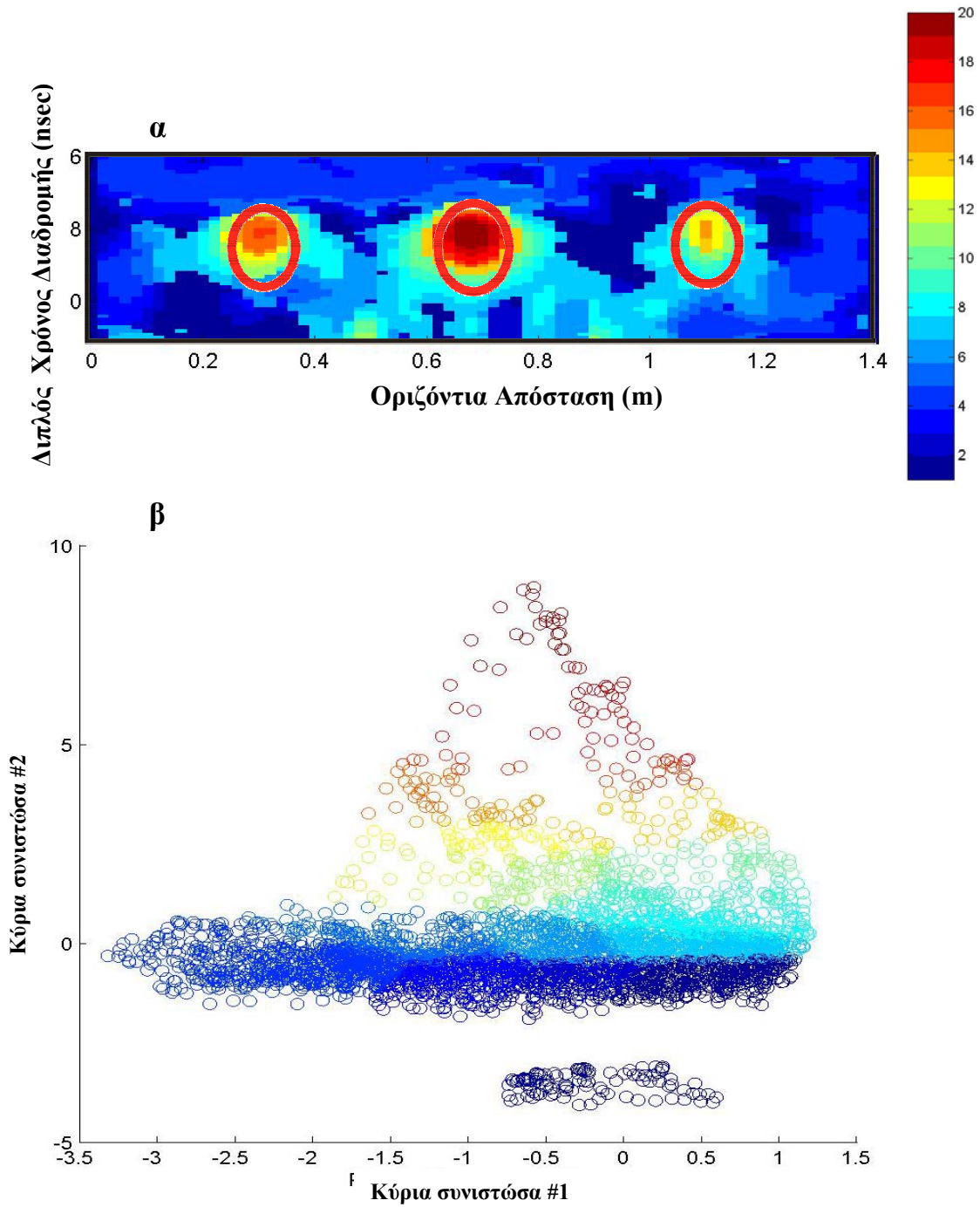
Σχήμα 4.12. (α) Το αποτέλεσμα της ταξινόμησης των τριών κυρίων συνιστωσών για το πείραμα με τη χρήση της μεθόδου των κ-μέσων τιμών (ταξινόμηση σε 100 τάξεις) και του δικτύου SOM (ταξινόμηση σε 20 τάξεις). (β) Τρισδιάστατο διάγραμμα διασποράς.



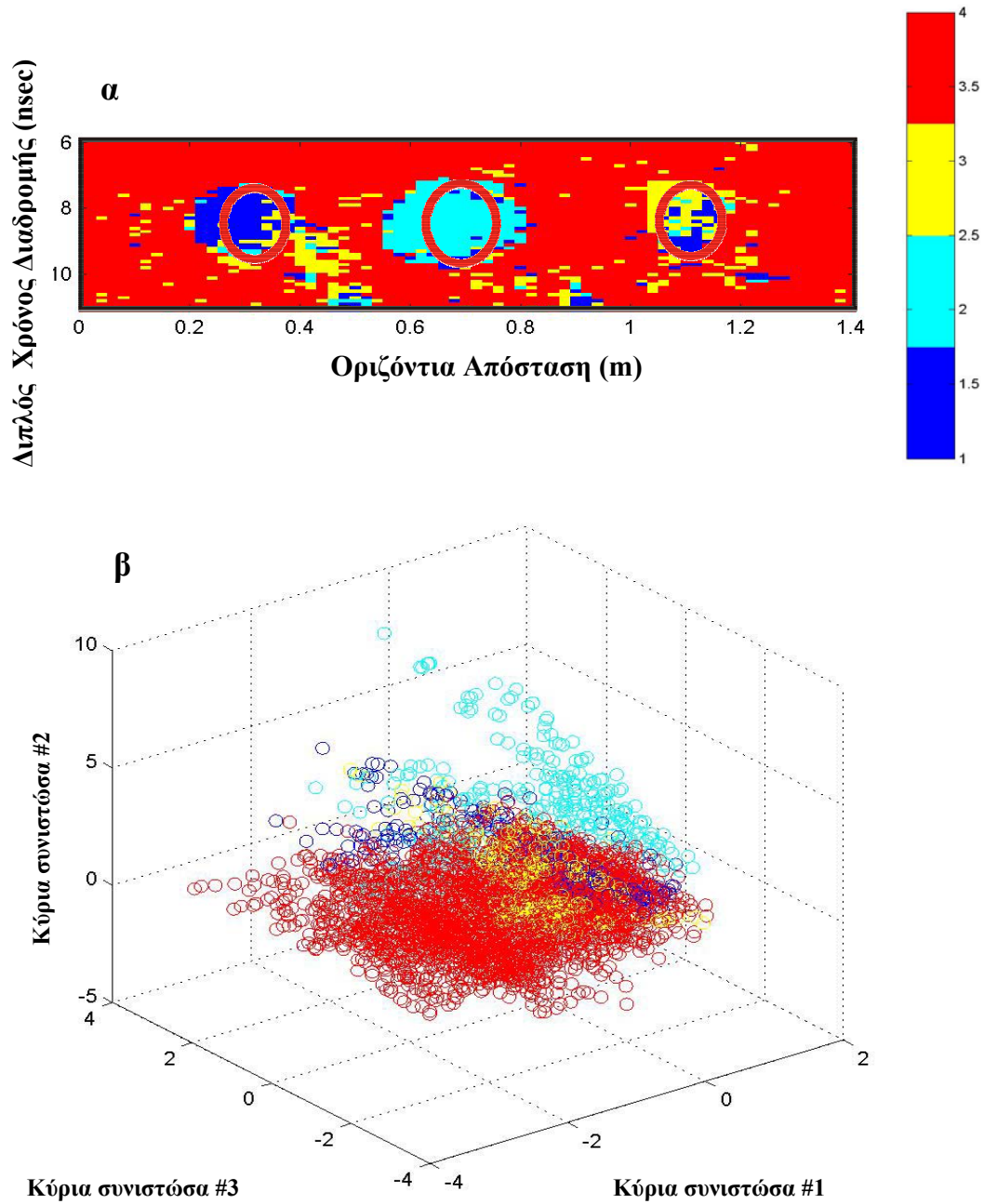
Σχήμα 4.13. Τα δείγματα που επιλέχθηκαν για την εκπαίδευση του δικτύου.



Σχήμα 4.14. Λεπτομέρεια του λογισμικού για την ταξινόμηση με τη μέθοδο δικτύου αυτο-οργανωμένης απεικόνισης. Ζητείται από τον αλγόριθμο να πάρει ως δείγματα εκπαίδευσης αυτά που επιλέχτηκαν χειροκίνητα.



Σχήμα 4.15. (α) Ταξινόμηση με το δίκτυο αυτο-οργανωμένης απεικόνισης σε 30 τάξεις με τη χρήση δειγμάτων που επιλέχτηκαν δια χειρός. (β) Το διάγραμμα διασποράς είναι ομοιόμορφο και καθιστά σαφέστατο το διαχωρισμό των ομάδων.



Σχήμα 4.16. (α) Καθοδηγούμενη ταξινόμηση. Οι στόχοι διαχωρίζονται πλήρως από το υπόβαθρο με μικρή συμμετοχή από λάθος ταξινομημένες καταγραφές. Στο διάγραμμα διασποράς (β) διακρίνεται ο διαχωρισμός των καταγραφών σε ομάδες.

4.6 Αυτόματη ταξινόμηση ανακλαστήρων

Για την παρουσίαση της μεθοδολογίας χρησιμοποιείται η γραμμή μελέτης ΤΑ η οποία διασχίζει κάθετα τις 71 γραμμές μελέτης (σχήμα 4.2). Η γραμμή μελέτης ΤΑ περνάει πάνω από τον πλαστικό σωλήνα (APVC) και το αργιλικό στρώμα. Η επεξεργασία των δεδομένων είναι ίδια με αυτή της γραμμής P36 (βλέπε παράγραφο 4.3). Τα δεδομένα αμέσως μετά την ενίσχυση τους παρουσιάζονται στο σχήμα 4.17.

Σε αυτή την περίπτωση επιλέγεται κάποιος ορίζοντας καθώς και τα δεδομένα γύρω από αυτόν που ανήκουν σε προκαθορισμένο παράθυρο. Το παράθυρο αυτό περιλαμβάνει τα ανακλώμενα κύματα από τον σωλήνα και το αργιλικό στρώμα. Η ταξινόμηση εφαρμόζεται σε καταγραφές που σχετίζονται με την ανάκλαση (σχήμα 4.18). Παρατηρείται ότι κατά την ανάκλαση στον πλαστικό σωλήνα το κυματίδιο παίρνει αρχικά θετικές τιμές, μετά αρνητικές και πάλι θετικές. Το αντίθετο συμβαίνει στο αργιλικό στρώμα όπου οι καταγραφές είναι επιπλέον πιο ισχυρές. Η επιλογή των δειγμάτων προς ταξινόμηση έγινε στα κεντρικά σημεία των καταγραφών που αφορούσαν κάθε ανάκλαση (σχήμα 4.19). Στο σχήμα (4.20) έχουν υπερτεθεί τα ταξινομημένα δεδομένα πάνω στα αρχικά, ενώ στο σχήμα (4.21) φαίνονται μόνο τα ταξινομημένα δεδομένα. Φαίνεται ότι οι ανακλάσεις λόγω της ανομοιογένειας της άμμου γύρω από τους στόχους ταξινομούνται στις τάξεις 7-8 (κίτρινο), οι ανακλάσεις από τον σωλήνα στις ομάδες 9-10 (κόκκινο), ενώ το αργιλικό στρώμα στις ομάδες 3-5 (γαλάζιο). Οι τρεις διαφορετικές ανακλάσεις διακρίνονται πλήρως και υπάρχει και πολύ καλή συμφωνία στις διαστάσεις των στόχων.

4.7 Συμπεράσματα.

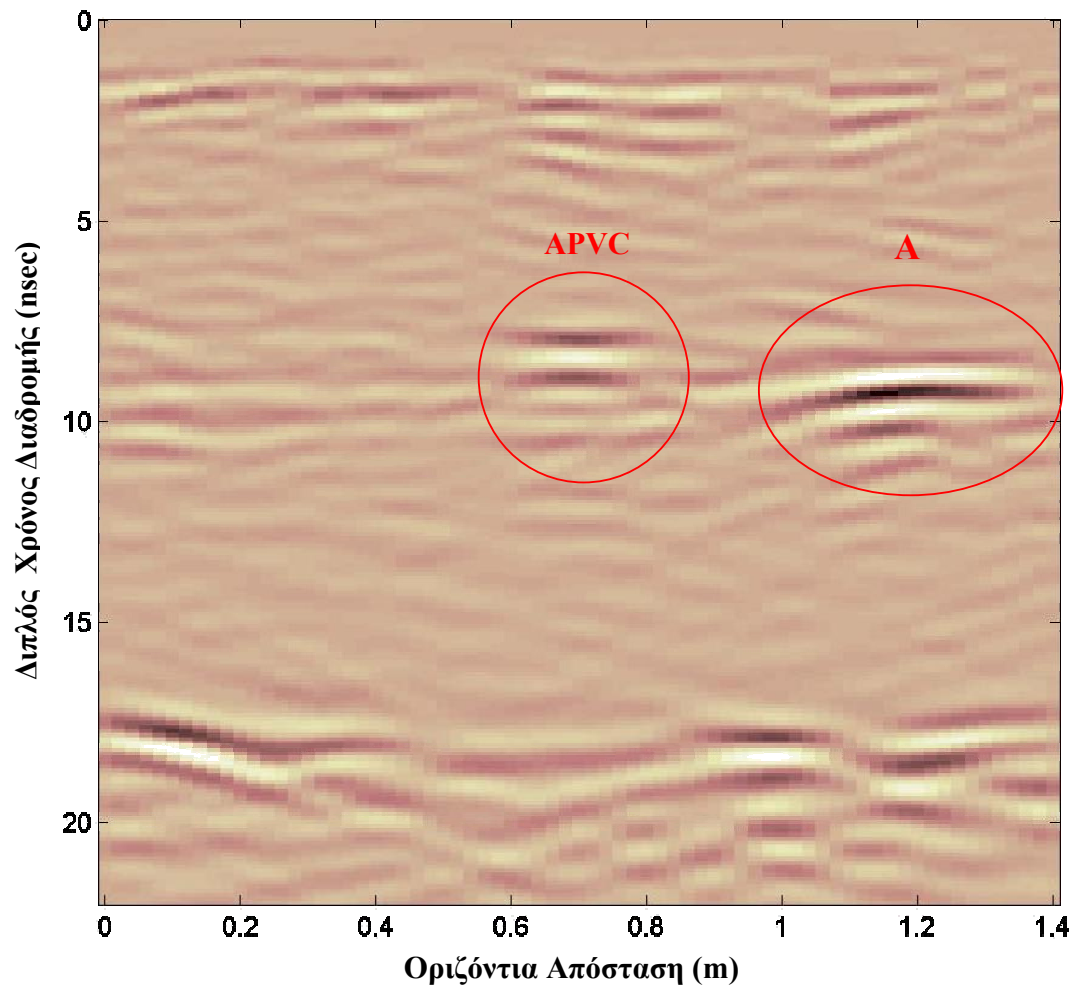
Στο κεφάλαιο αυτό χρησιμοποιήθηκαν μέθοδοι αυτόματης και καθοδηγούμενης ταξινόμησης για την ερμηνεία τομής γεωραντάρ από ελεγχόμενο πείραμα. Η ερμηνεία των συγκεκριμένων δεδομένων είναι δυνατή και χωρίς τη χρήση αυτών των μεθόδων. Σκοπός όμως ήταν να διερευνηθούν οι δυνατότητες των μεθόδων αυτών στο να διαχωρίσουν τα θαμμένα αντικείμενα τόσο μεταξύ τους όσο και από το περιβάλλον τους.

Όσον αφορά στην αυτόματη ταξινόμηση, χρησιμοποιήθηκε αρχικά η μέθοδος των κ-μέσων τιμών και έπειτα τα δίκτυα αυτο-οργανωμένης απεικόνισης με δείγματα εισόδου δύο και τρία χαρακτηριστικά. Τα θαμμένα αντικείμενα διαχωρίστηκαν από το περιβάλλον τους αλλά φαίνεται να ανήκουν στην ίδια ομάδα.. Στην περίπτωση όπου το δίκτυο αυτο-οργανωμένης απεικόνισης εκπαιδεύτηκε από επιλεγμένα δείγματα των καταγραφών, τα αποτελέσματα ήταν καλύτερα αφού τα θαμμένα αντικείμενα διαχωρίστηκαν τόσο μεταξύ τους όσο και από το περιβάλλον τους.

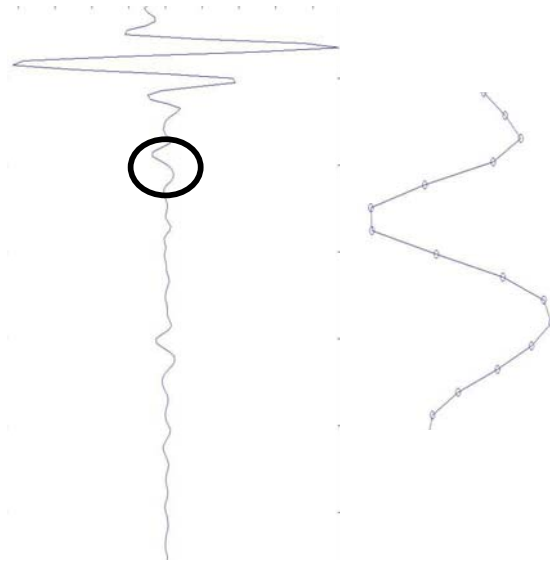
Τα ίδια επιλεγμένα δείγματα χρησιμοποιήθηκαν για την εκπαίδευση του δικτύου διανυσματικής μάθησης και τα θαμμένα αντικείμενα ταξινομήθηκαν σε διαφορετικές ομάδες.

Η αυτόματη ταξινόμηση ανακλαστήρων με τη χρήση του δικτύου αυτο-οργανωμένης απεικόνισης έδωσε καλά αποτελέσματα αφού διαχώρισε τα θαμμένα αντικείμενα από το περιβάλλον τους με σαφήνεια αναδεικνύοντας το διαφορετικό χαρακτήρα των καταγραφών.

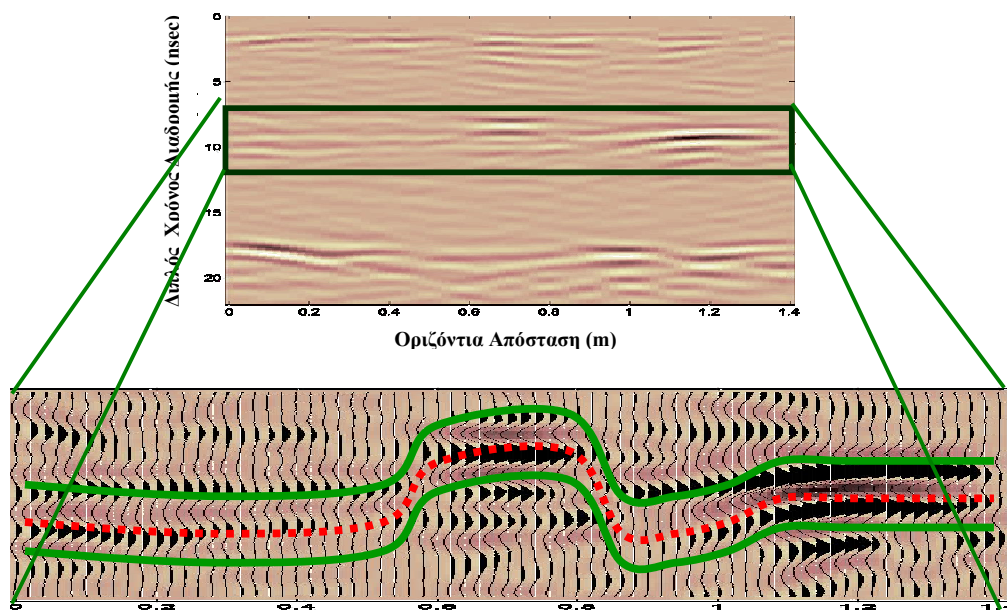
Από τα παραπάνω προκύπτει ότι όταν χρησιμοποιούνται χαρακτηριστικά δείγματα των καταγραφών για την εκπαίδευση του αλγόριθμου ταξινόμησης τα αποτελέσματα είναι καλύτερα είτε πρόκειται για αυτόματη είτε για καθοδηγούμενη ταξινόμηση, συμβάλλοντας στην ερμηνεία των αρχικών δεδομένων τα οποία δεν είναι άμεσα ερμηνεύσιμα.



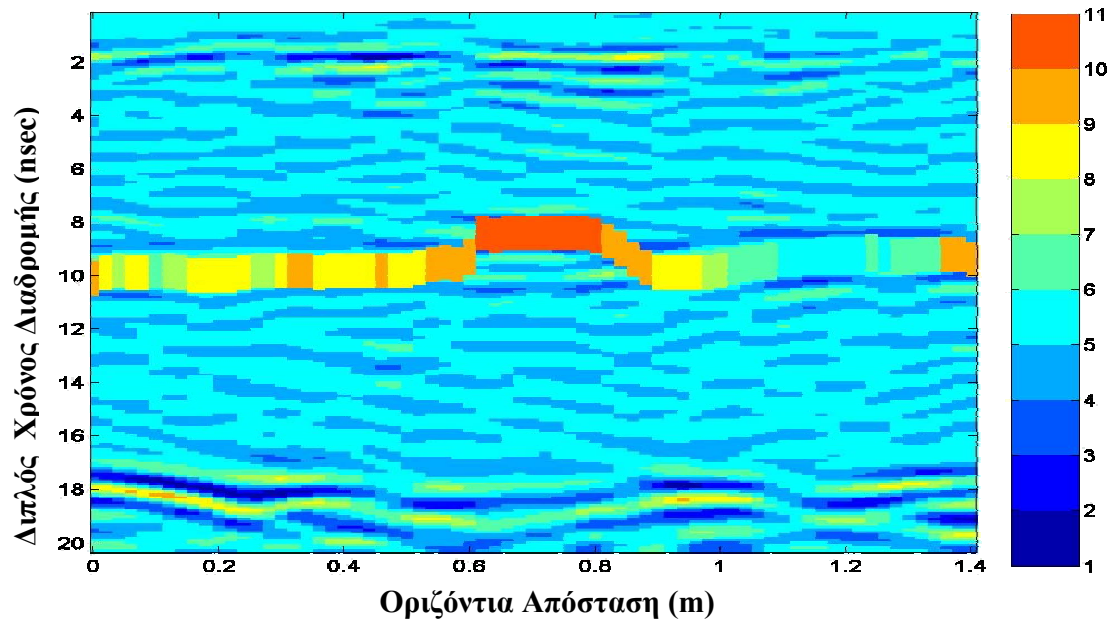
Σχήμα 4.17. Τα δεδομένα της γραμμής μελέτης ΤΑ. Διακρίνονται οι καταγραφές από τον πλαστικό σωλήνα (APVC) και την άργιλο (Α).



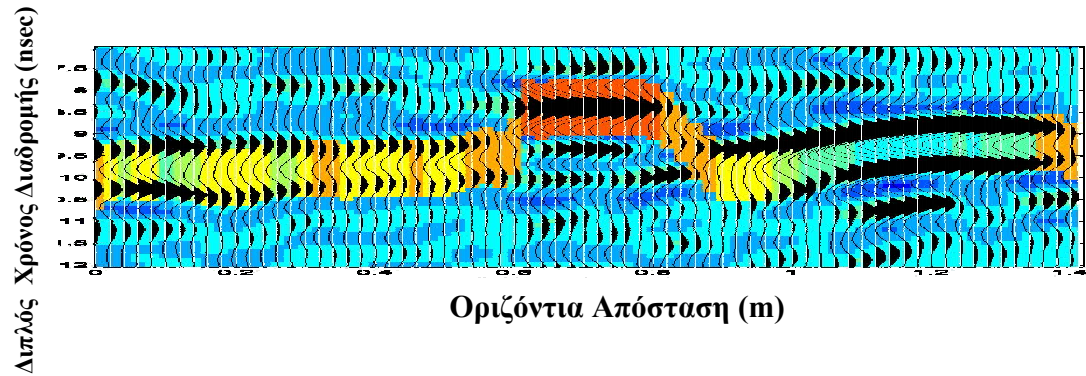
Σχήμα 4.18. Η καταγραφή του γεωραντάρ που συνδέεται με την ανάκλαση στον οριζόντιο ανακλαστήρα (κύκλος). Η κυματομορφή πολώνεται αρχικά με θετικές τιμές, έπειτα με αρνητικές και τελικά με θετικές τιμές.



Σχήμα 4.19. Η επιλογή των δειγμάτων (κόκκινη γραμμή) έγινε στην περιοχή 7-12 nsec. Ορίστηκε παράθυρο διάρκειας 1 nsec (πράσινες γραμμές) εκατέρωθεν των σημείων επιλογής (κόκκινη γραμμή).



Σχήμα 4.20. Τα ταξινομημένα δεδομένα έχουν υπερτεθεί πάνω στα αρχικά δεδομένα όπου διαχωρίζονται οι στόχοι από το υπόβαθρο.



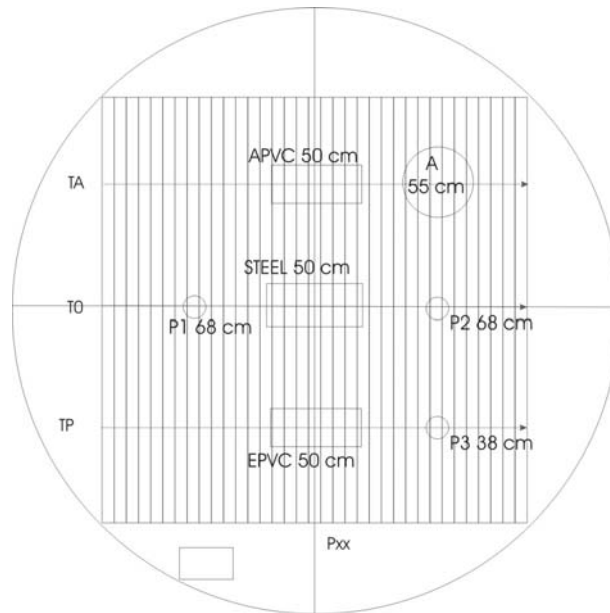
Σχήμα 4.21. Τα ταξινομημένα δεδομένα έχουν υπερτεθεί στα αρχικά ενώ φαίνονται και οι καταγραφές με τη μορφή ιχνών. Οι στόχοι διαχωρίζονται λόγω της διαφορετικής φάσης τους αλλά και λόγω των διαφορών στα πλάτη.

ΚΕΦΑΛΑΙΟ 5**ΤΑΞΙΝΟΜΗΣΗ ΨΕΥΔΟ-ΤΡΙΣΔΙΑΣΤΑΤΩΝ ΔΕΔΟΜΕΝΩΝ ΓΕΩΡΑΝΤΑΡ****5.1 Εισαγωγή.**

Σε αυτό το κεφάλαιο παρουσιάζεται μεθοδολογία ταξινόμησης η οποία εφαρμόζεται στα δεδομένα γεωραντάρ από το ελεγχόμενο πείραμα (Sandbox Experiment) που περιγράφηκε στο τέταρτο κεφάλαιο. Στην παρούσα περίπτωση ταξινομούνται οριζόντιες τομές από τις εβδομήντα μια (71) παράλληλες τομές γεωραντάρ του πειράματος (σχήμα 5.1).

5.2 Επεξεργασία των δεδομένων.

Αρχικά εφαρμόστηκε το φίλτρο απομάκρυνσης της επαγωγικής συνιστώσας (*Dewowing*) και έγινε χωροθέτηση των δεδομένων με τον αλγόριθμο του Gazdag. Σύμφωνα με τις μετρήσεις κοινού σημείου ανάκλασης η ταχύτητα είναι ίση με 0.14 m/nsec (Loeffler and Bano 2004). Εισάγοντας αυτή την ταχύτητα στον αλγόριθμο παραμορφώθηκαν οι ανακλαστήρες.

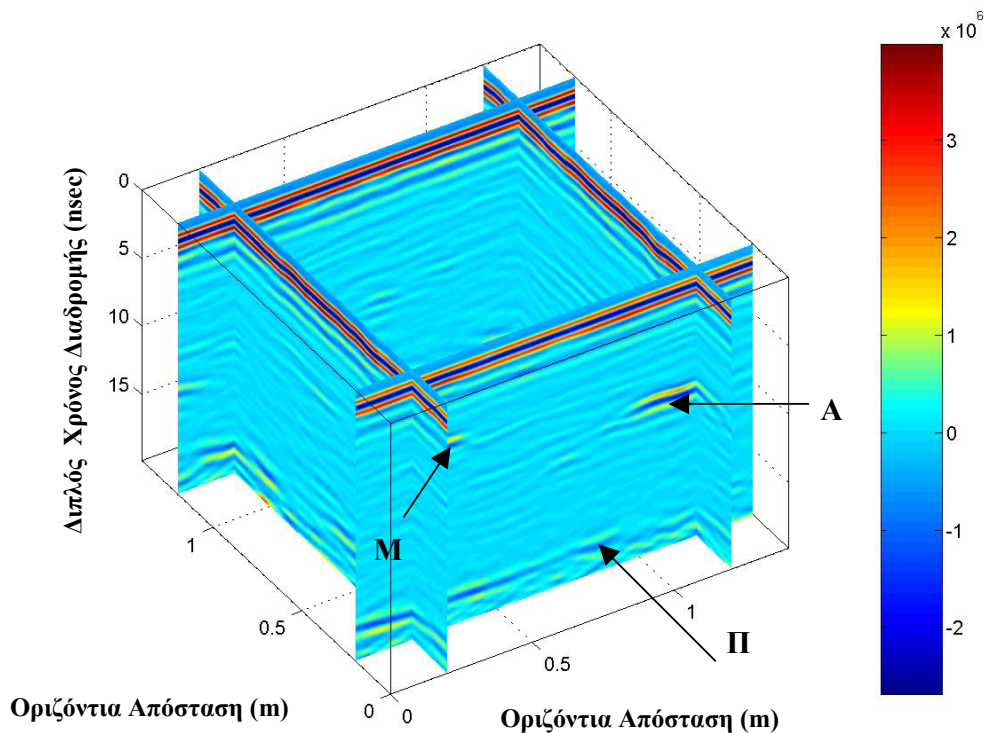


Σχήμα 5.1. Κάτοψη του καννάβου. Διακρίνονται οι στόχοι καθώς και το βάθος ταφής τους (P1-P3: μεταλλικές μπάλες θαμμένες σε βάθη 68 cm, 68 cm, 38 cm, A: αργιλικό στρώμα σε βάθος 55 cm και τρεις σωλήνες (APVC, EPVC και STEEL) θαμμένοι σε βάθος 50 cm. Διακρίνονται οι 71 γραμμές μελέτης και οι τρεις κάθετες σε αυτές (Loeffler and Bano 2004).

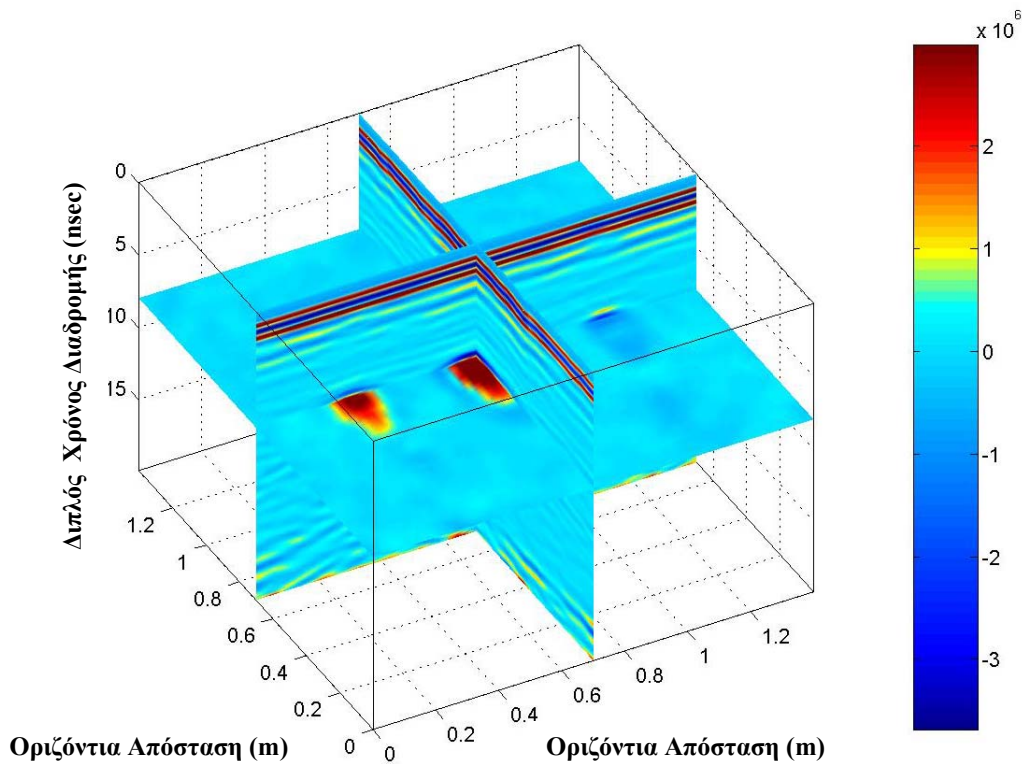
Μετά από διάφορες δοκιμές διαπιστώθηκε ότι ταχύτητα ίση με 0.125 m/nsec είναι κατάλληλη για τη χωροθέτηση των δεδομένων, αφού οι διαστάσεις των στόχων αποδίδονταν σωστά στις τομές του γεωραντάρ. Επίσης, εφαρμόστηκαν φίλτρο αντιστάθμισης για τη γεωμετρική διασπορά και φίλτρο εξομάλυνσης. Στη συνέχεια υπολογίστηκαν τέσσερα στιγμιαία μιγαδικά χαρακτηριστικά των δεδομένων (στιγμιαίο πλάτος, στιγμιαίο πλάτος σε κλίμακα dB, στιγμιαία φάση και στιγμιαία συχνότητα) και δυο γεωμετρικά (συνάφεια των δεδομένων και συνάφεια της στιγμιαίας φάσης). Στα σχήματα (5.2 -5.3) φαίνονται τομές από τα ψευδο-τρισδιάστατα δεδομένα μετά το πρώτο στάδιο επεξεργασίας. Οι οριζόντιες τομές προέκυψαν, λαμβάνοντας υπόψη δεδομένα που ανήκουν στον ίδιο χρόνο.

Στα δεδομένα μετά από την εξομάλυνση καθώς και στα χαρακτηριστικά τους, διαπιστώθηκε ότι τα ανακλώμενα κύματα από τους τρεις σωλήνες, το αργιλικό στρώμα και τη μεταλλική μπάλα που ήταν θαμμένη σε βάθος 38 cm παρουσιάζουν μεγάλα πλάτη σε σχέση με αυτά από τις δυο μεταλλικές μπάλες που ήταν θαμμένες σε βάθος 68 cm.

Ειδικά μετά από την χωροθέτηση αυτές οι μπάλες έδιναν τόσο ασθενείς καταγραφές που δεν διακρίνονταν. Αποφασίστηκε ότι τα επόμενα στάδια επεξεργασίας θα αφορούσαν τμήμα της περιοχής από 5.5 έως 10 nsec. Επίσης, για λόγους οικονομίας στο χρόνο επεξεργασίας, υπολογίστηκαν οριζόντιες τομές ανά 0.5 nsec. Διαπιστώθηκε ότι τα γεωμετρικά χαρακτηριστικά δεν έδιναν ενδείξεις για την ύπαρξη των στόχων οπότε δεν χρησιμοποιήθηκαν στα επόμενα στάδια επεξεργασίας.



Σχήμα 5.2. Τέσσερις τομές γεωραντάρ. Διακρίνονται τα ανακλώμενα κύματα από τον πυθμένα του δοχείου (Π) περίπου στα 18 nsec, καθώς και από τη μεταλλική μπάλα (M), η οποία είναι θαμμένη στα 38 cm, καθώς και το αργιλικό στρώμα (A).

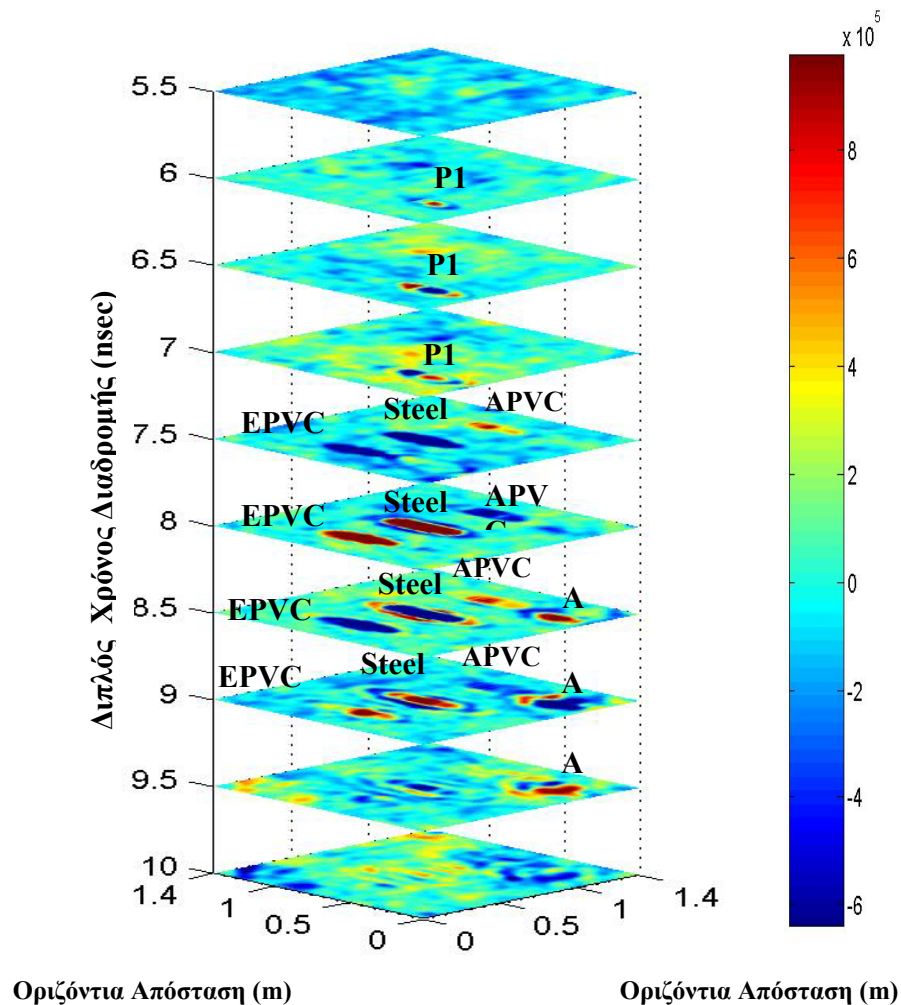


Σχήμα 5.3. Οριζόντια τομή στα 8 nsec, στην οποία διακρίνονται ανακλώμενα και περιθλώμενα κύματα από τους σωλήνες.

5.3 Ερμηνεία των οριζόντιων τομών.

Από τα χωροθετημένα δεδομένα προέκυψαν δέκα οριζόντιες τομές (5.4). Στις οριζόντιες τομές στα 6 nsec, 6.5 nsec, 7 nsec παρατηρούνται ανακλώμενα κύματα από μεταλλική μπάλα που είναι θαμμένη στα 38 cm. Στις διαδοχικές τομές αυτά τα κύματα εμφανίζουν θετικές, αρνητικές και θετικές τιμές αντίστοιχα. Από την τομή στα 7.5 nsec εμφανίζονται ανακλώμενα και περιθλώμενα κύματα από τους σωλήνες. Όσον αφορά στους σωλήνες Steel και EPVC, οι καταγραφές που σχετίζονται με αυτούς στα 7.5 nsec έχουν αρνητικές τιμές και μέχρι τα 9 nsec εναλλάσσονται με θετικές τιμές. Αντίθετα για τον σωλήνα APVC στα 7 nsec εμφανίζονται θετικές τιμές, οι οποίες μέχρι τα 9 nsec εναλλάσσονται με αρνητικές τιμές. Πρέπει επίσης να επισημανθεί ότι οι καταγραφές που σχετίζονται με τους σωλήνες EPVC και Steel έχουν πολύ μεγαλύτερο πλάτος από τις

καταγραφές του σωλήνα APVC. Ανακλώμενα κύματα από το αργιλικό στρώμα (A) καταγράφονται στα 8 nsec με αρνητικές τιμές και μικρότερο πλάτος. Στα 9.5 nsec φαίνονται μόνο αμυδρές καταγραφές από τους σωλήνες και την άργιλο, ενώ στα 10 nsec υπάρχουν καταγραφές μόνο από το αργιλικό στρώμα.



Σχήμα 5.4. Οριζόντιες τομές από τα δεδομένα μετά τη χωροθέτηση και την εξομάλυνση (5.5 nsec - 10 nsec). Διακρίνονται ανακλώμενα κύματα από τα θαμμένα αντικείμενα.

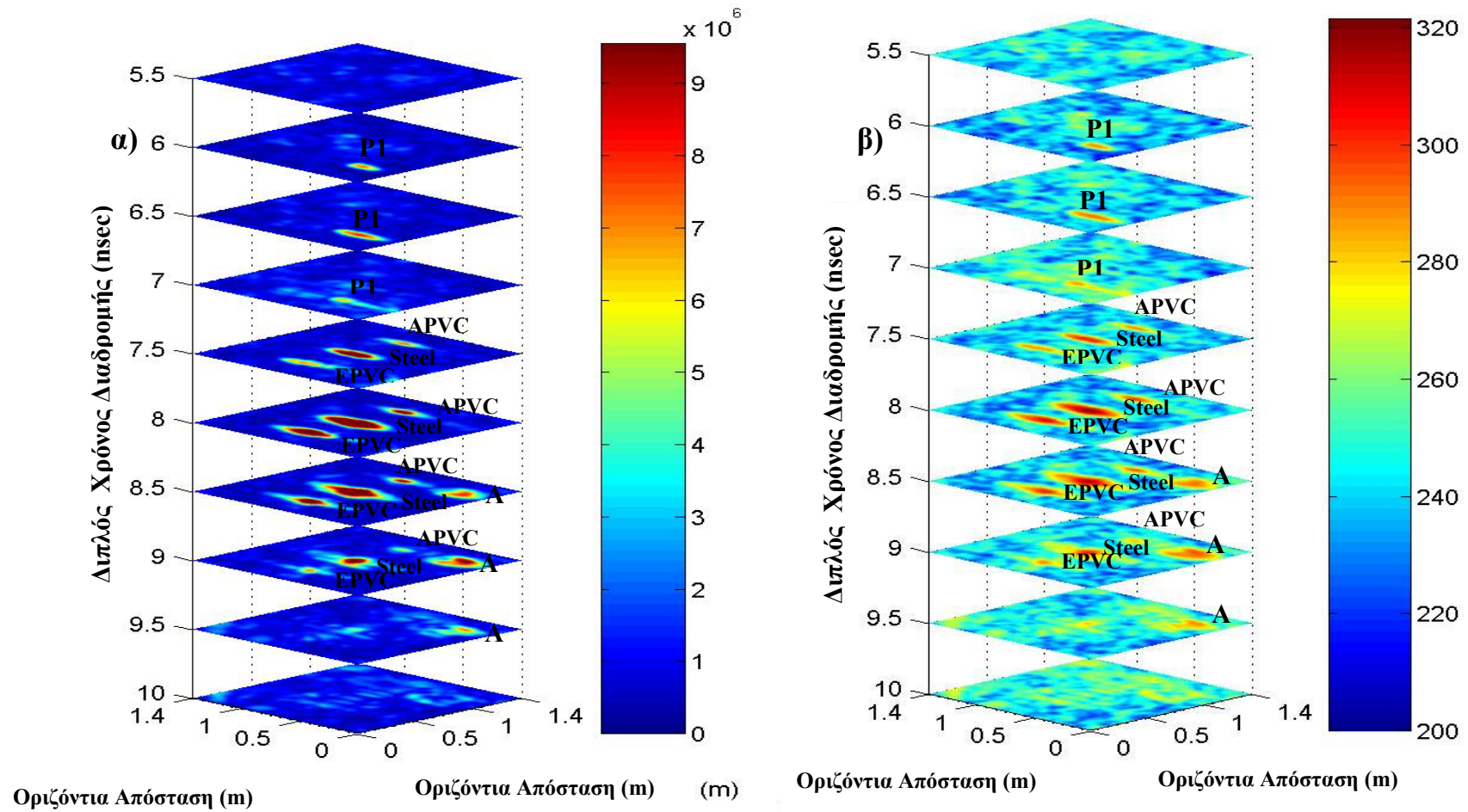
Στις διαδοχικές τομές του στιγμιαίου πλάτους (σχήμα 5.5α) οι καταγραφές με μεγάλο πλάτος (κόκκινο χρώμα) οφείλονται στα θαμμένα αντικείμενα. Το στιγμιαίο πλάτος σε κλίμακα dB (σχήμα 5.5β), τονίζει επίσης τα ανακλώμενα κύματα από τα θαμμένα αντικείμενα (αποχρώσεις του κόκκινου).

Η στιγμιαία φάση (σχήμα 5.5γ) και η στιγμιαία συχνότητα (σχήμα 5.5δ) τονίζουν το περίγραμμά τους, το οποίο έχει αντίθετες τιμές από τους ίδιους τους στόχους. Για παράδειγμα, στην οριζόντια τομή στα 8 nsec η στιγμιαία φάση παρουσιάζει τους σωλήνες να έχουν θετικές τιμές, ενώ γύρω από αυτούς η φάση είναι αρνητική. Η στιγμιαία συχνότητα αποδίδει τους σωλήνες EPVC και Steel με τον ίδιο τρόπο ενώ ο σωλήνας APVC φαίνεται να έχει αρνητικές τιμές ενώ περιβάλλεται από θετικές τιμές της συχνότητας.

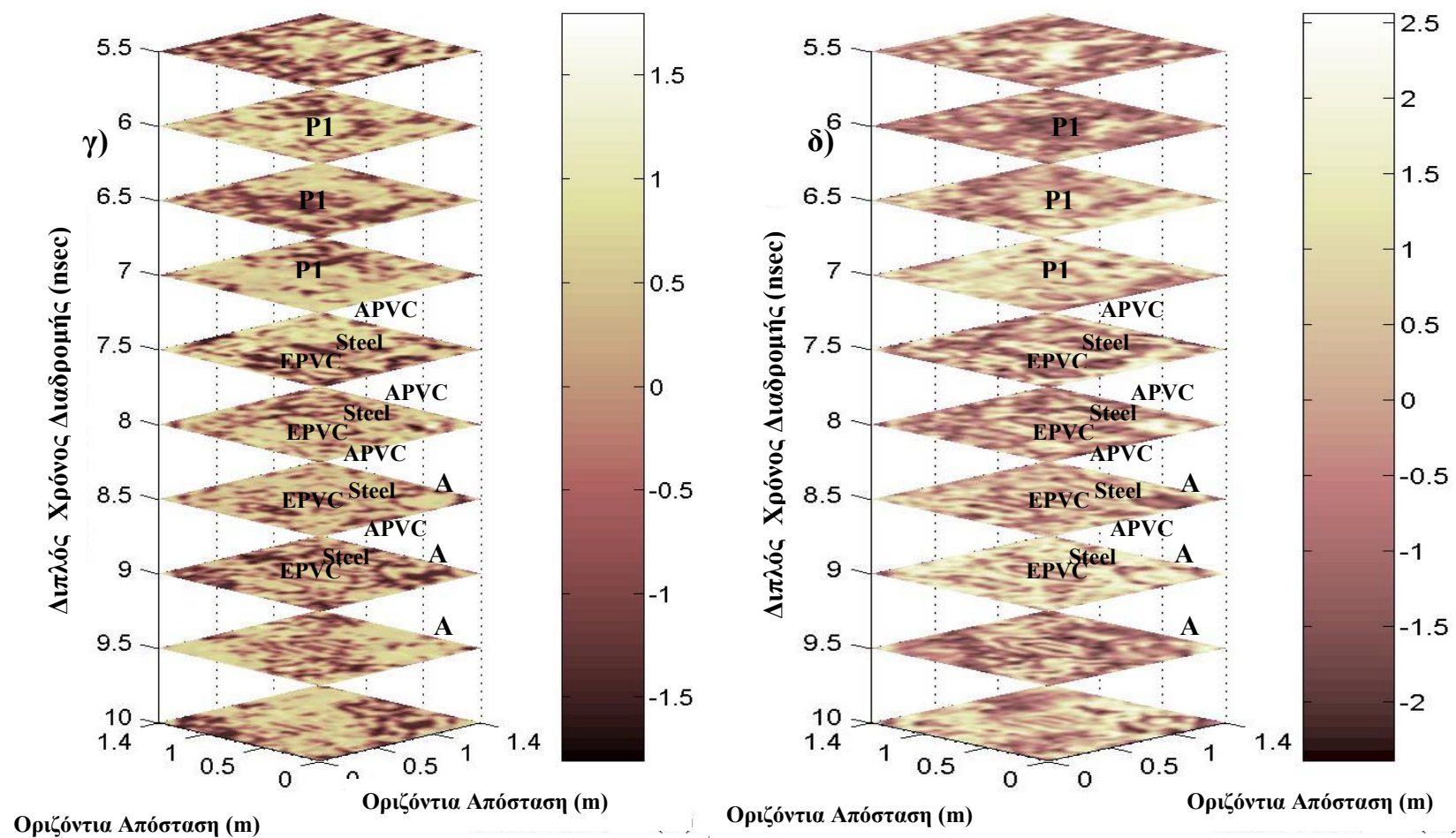
5.4 Ανάλυση κυρίων συνιστωσών.

Τα τέσσερα στιγμιαία χαρακτηριστικά καθώς και τα αρχικά δεδομένα μετά την εξομάλυνση χρησιμοποιήθηκαν για την ανάλυση κυρίων συνιστωσών. Τρεις συνιστώσες βρέθηκε ότι περιέχουν περίπου το 85% της συνολικής πληροφορίας των πέντε αρχικών (σχήμα 5.6α). Πρέπει να τονιστεί ότι τα δεδομένα, πριν εισαχθούν στον αλγόριθμο κανονικοποιήθηκαν.

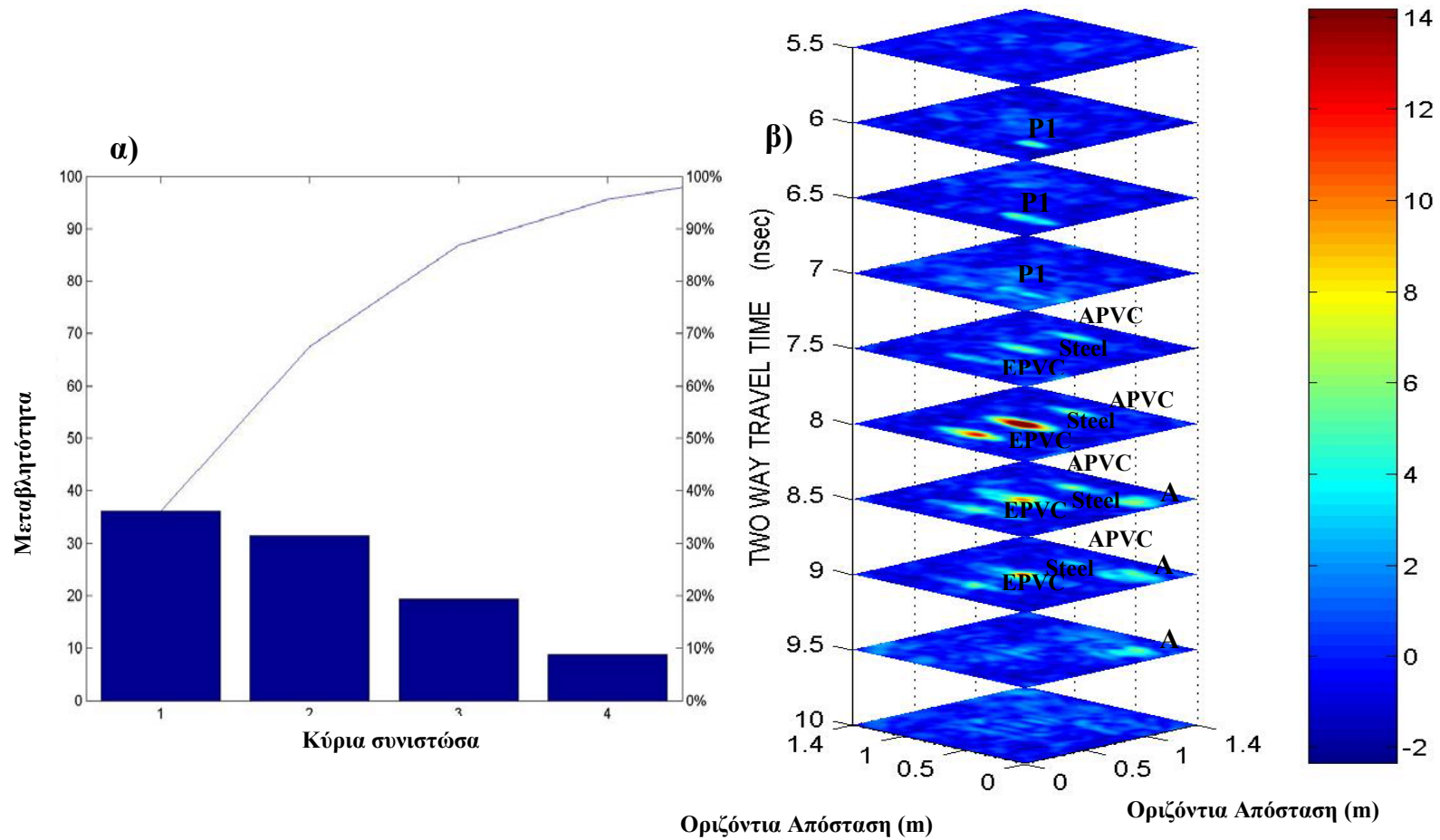
Η πρώτη κύρια συνιστώσα (σχήμα 5.6β) αντιστοιχεί στο στιγμιαίο πλάτος. Όλοι οι στόχοι παρουσιάζονται με μεγάλες τιμές (τόνοι του κόκκινου). Η δεύτερη κύρια συνιστώσα (σχήμα 5.6γ), αντιστοιχεί στη στιγμιαία φάση και παρουσιάζει τους στόχους με διαφορετικές τιμές στις διαδοχικές οριζόντιες τομές. Στα 7.5 nsec και στα 8.5 nsec οι σωλήνες EPVC και Steel παρουσιάζονται με αρνητικές τιμές (μπλε), σε αντίθεση με τον σωλήνα APVC ο οποίος παρουσιάζεται με θετικές τιμές (κόκκινο). Στα 8 nsec συμβαίνει το ακριβώς αντίθετο. Η τρίτη κύρια συνιστώσα (σχήμα 5.6δ) παρουσιάζει ομοιότητες με τα δεδομένα μετά την εξομάλυνση και το στιγμιαίο πλάτος σε κλίμακα dB.



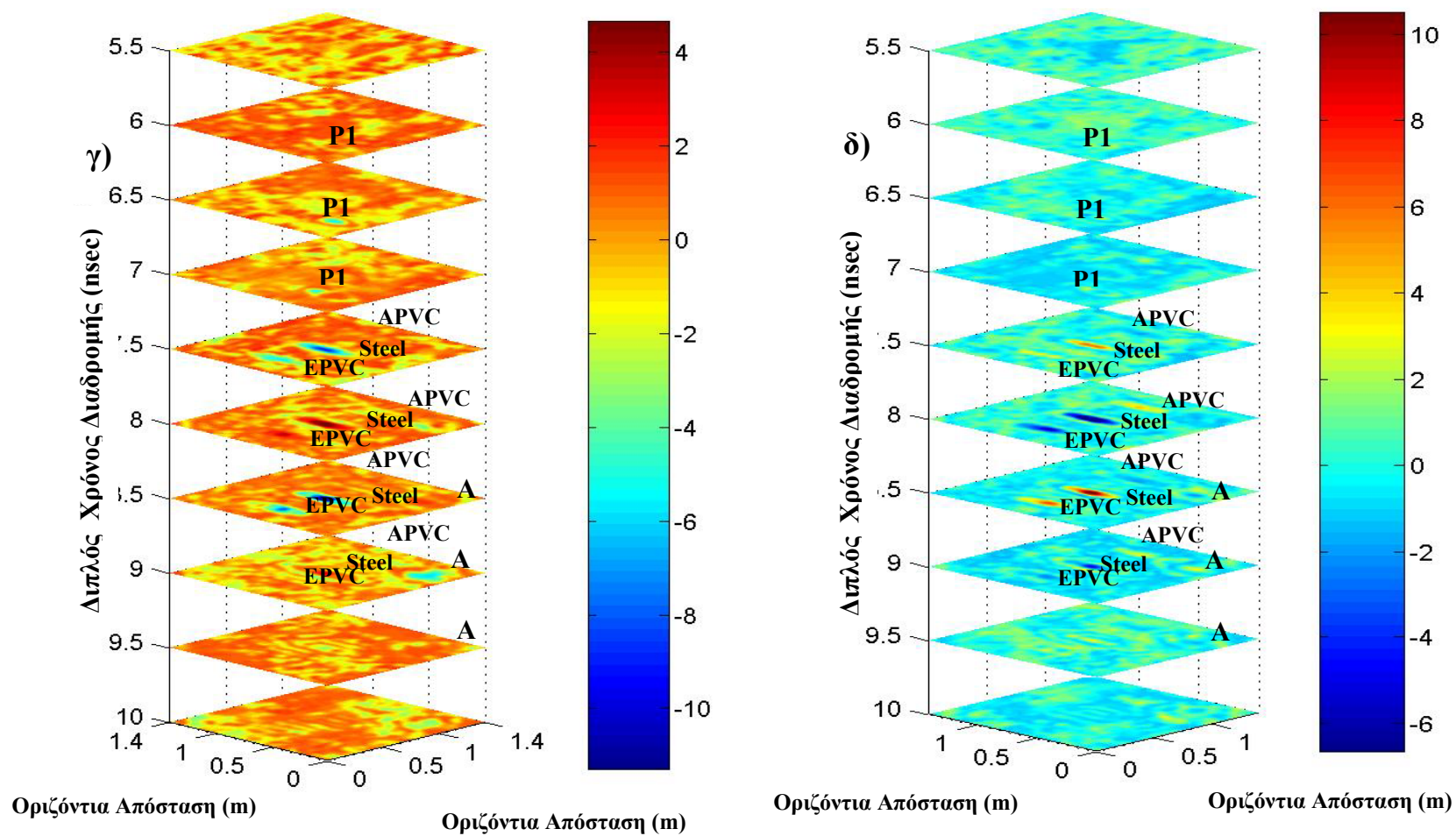
Σχήμα 5.5. (α) Το στιγμιαίο πλάτος και (β) το στιγμιαίο πλάτος σε κλίμακα dB. Τα ανακλώμενα κύματα από τους στόχους έχουν μεγάλα πλάτη.



Σχήμα 5.5 (συνέχεια). (γ) Η στιγμιαία φάση και (δ) η στιγμιαία συχνότητα.



Σχήμα 5.6. (α) Οι τρεις πρώτες κύριες συνιστώσες περιέχουν περισσότερο από 85% της συνολικής πληροφορίας. (β) Η πρώτη κύρια συνιστώσα επηρεάζεται από το στιγμιαίο πλάτος.



Σχήμα 5.6 (συνέχεια). (γ) Η δεύτερη και (δ) η τρίτη κύρια συνιστώσα.

5.5 Ταξινόμηση των δεδομένων.

Η ταξινόμηση των δεδομένων έγινε με τρεις τρόπους. Χρησιμοποιήθηκαν τόσο τα αρχικά δεδομένα μαζί με τα τέσσερα χαρακτηριστικά όσο και οι τρεις κύριες συνιστώσες οι οποίες περιέχουν περισσότερο από 80% της συνολικής πληροφορίας των πέντε. Συγκρίθηκαν οι μέθοδοι αυτόματης και καθοδηγούμενης ταξινόμησης ως προς τον υπολογιστικό τους χρόνο. Πρέπει να αναφερθεί ότι η επεξεργασία έγινε σε προσωπικό υπολογιστή (PC) με επεξεργαστή *Intel Pentium 3.2 MHz* και μνήμη *RAM 512 MB*. Για την αυτόματη ταξινόμηση χρησιμοποιήθηκαν οι μέθοδοι των κ-μέσων τιμών και των δικτύων αυτό-οργάνωσης (SOM), ενώ για την καθοδηγούμενη ταξινόμηση χρησιμοποιήθηκαν τα δίκτυα πιθανότητας (PNN).

5.5.1 Αυτόματη ταξινόμηση.

Η πρώτη δοκιμή αφορά στην αυτόματη ταξινόμηση των δεδομένων με τη χρήση της μεθόδου κ-μέσων τιμών και του δικτύου αυτο-οργανωμένης απεικόνισης. Αρχικά, τα δεδομένα δέκα (10) οριζόντιων τομών ταξινομούνται σε 200 τάξεις με τη μέθοδο των κ-μέσων τιμών. Τα αντίστοιχα 200 κέντρα εισάγονται στο δίκτυο αυτο-οργανωμένης απεικόνισης όπου ταξινομούνται σε 30 ομάδες βάσει της τοπολογικής τους συνέχειας. Η μέθοδος των κ-μέσων τιμών μειώνει το χρόνο επεξεργασίας στο δίκτυο, το οποίο είναι γρήγορο, αλλά χρειάζεται πολλές επαναλήψεις για να συγκλίνει (περίπου 500 φορές επί τον αριθμό των ομάδων).

Η ταξινόμηση των αρχικών δεδομένων (σχήμα 5.7α) επιτεύχθηκε μετά από 15000 επαναλήψεις, δηλαδή 10 ώρες. Αρχικά παρατηρείται ότι στις εικόνες υπάρχουν δυο κύριες ομάδες δεδομένων. Η πρώτη αποτελείται από τις ομάδες 1-12 (αποχρώσεις του μπλε) και περιλαμβάνει ανακλώμενα κύματα από τα θαμμένα αντικείμενα. Η δεύτερη αποτελείται από τις ομάδες 13-30 και αντιστοιχεί στην άμμο. Επιπρόσθετα, παρατηρείται ότι στις διαδοχικές οριζόντιες τομές από 7.5 nsec έως 8.5 nsec το ίδιο θαμμένο αντικείμενο ανήκει σε διαφορετική ομάδα. Η διαδικασία αυτή δεν είναι επιτυχής στο να ταξινομήσει σε διαφορετικές ομάδες ανακλώμενα κύματα από θαμμένα αντικείμενα με

αποτέλεσμα να μη διακριθούν με σαφήνεια από την ταξινόμηση οριζόντιας τομής γεωραντάρ.

Η δεύτερη δοκιμή αφορά στην ξεχωριστή ταξινόμηση κάθε οριζόντιας τομής με την μέθοδο των κ-μέσων τιμών σε 20 ομάδες. Από την ταξινόμηση των δέκα (10) οριζόντιων τομών προέκυψαν συνολικά 200 νέα κέντρα. Τα κέντρα αυτά τροφοδοτήθηκαν στο δίκτυο αυτο-οργανωμένης απεικόνισης και οργανώθηκαν σε 30 τάξεις. Αυτός ο τρόπος μείωσε τον υπολογιστικό χρόνο επεξεργασίας σε 1.5 ώρες. Η διαφορά στο χρόνο είναι πάρα πολύ μεγάλη αλλά το αποτέλεσμα είναι υποδεέστερο (σχήμα 5.7β). Δηλαδή είναι δύσκολη η συσχέτιση διαδοχικών οριζόντιων τομών μετά από την ταξινόμηση. Για παράδειγμα, στα 8.5 nsec οι στόχοι ταξινομούνται στις ομάδες 25-30 (EPVC, Steel) και 1-5 (APVC, A), ενώ η άμμος ταξινομείται στις ομάδες 6-24.

Οι παραπάνω δοκιμές εφαρμόστηκαν για την ταξινόμηση των τριών κυρίων συνιστωσών. Οι αντίστοιχοι χρόνοι επεξεργασίας ήταν 8.5 ώρες και 5.5 ώρες. Στην πρώτη περίπτωση (σχήμα 5.7γ) παρατηρείται κατ' αρχήν ότι τα ανακλώμενα κύματα από τα θαμμένα αντικείμενα ταξινομούνται στις ομάδες 1-10 (αποχρώσεις του μπλε), ενώ τα ανακλώμενα κύματα από την άμμο ταξινομούνται στις ομάδες 11-30 (αποχρώσεις πράσινου και κόκκινου). Για παράδειγμα οι σωλήνες EPVC και Steel ταξινομούνται στις ομάδες 1-4 ενώ ο σωλήνας APVC και η άργιλος A, ταξινομούνται στις ομάδες 5-10. Στη δεύτερη δοκιμή (σχήμα 5.7δ) οι στόχοι διακρίνονται καλύτερα από το περιβάλλον τους εκτός του σωλήνα APVC ο οποίος διακρίνεται μονάχα στις τομές 8 nsec – 8.5 nsec. Η διαδικασία αυτή επίσης δεν ήταν επιτυχής στο να ταξινομήσει σε διαφορετικές ομάδες ανακλώμενα κύματα από διαφορετικά αντικείμενα.

Οι επόμενες δοκιμές, συνδυάζουν την αυτόματη με την καθοδηγούμενη ταξινόμηση. Αρχικά θεωρήθηκε ότι η θέση των θαμμένων αντικειμένων είναι γνωστή, οπότε επιλέχθηκαν από την καταγραφή δείγματα που αντιστοιχούν σε ανακλώμενα κύματα, από κάθε αντικείμενο. Τα συνολικά 780 δείγματα εισήχθησαν στο δίκτυο αυτο-οργανωμένης απεικόνισης και οργανώθηκαν χωρίς φυσικά να καθοριστεί αν αυτά τα δείγματα ανήκουν σε ίδιες ή διαφορετικές ομάδες. Η πρώτη δοκιμή (σχήμα 5.8α) αφορούσε στα αρχικά τους δεδομένα, ενώ η δεύτερη στις τρεις κύριες συνιστώσες (σχήμα 5.8β). Στην πρώτη περίπτωση (διάρκεια επεξεργασίας 215 min) το αποτέλεσμα είναι ο πλήρης διαχωρισμός των θαμμένων αντικειμένων τόσο από το περιβάλλον τους

όσο και μεταξύ τους. Τα ανακλώμενα κύματα από την άμμο ταξινομούνται στις ομάδες 17-30 (τόνοι του κόκκινου), ενώ τα ανακλώμενα κύματα από τα θαμμένα αντικείμενα ταξινομούνται στις ομάδες 1-17 (τόνοι του μπλε). Στη δεύτερη δοκιμή (διάρκεια επεξεργασίας 213 min) τα ανακλώμενα κύματα από τα θαμμένα αντικείμενα ταξινομούνται στις ομάδες 1-11 (τόνοι του μπλε) ενώ τα ανακλώμενα κύματα από την άμμο ταξινομούνται στις ομάδες 12-30 (τόνοι του πράσινου και κόκκινου). Οι διαδικασίες αυτές κρίνονται επιτυχείς στο να ταξινομήσουν σε διαφορετικές ομάδες ανακλώμενα κύματα από θαμμένα αντικείμενα, με αποτέλεσμα να διακρίνονται αυτά τα αντικείμενα με σαφήνεια από την ταξινόμηση οριζόντιας τομής γεωραντάρ.

5.5.2 Καθοδηγούμενη ταξινόμηση οριζόντιας τομής.

Σε αυτή την παράγραφο παρουσιάζονται και συγκρίνονται ταξινομημένες εικόνες οριζόντιας τομής οι οποίες προέρχονται από διαφορετικές μεθόδους καθοδηγούμενης ταξινόμησης. Για το λόγο αυτό επιλέχθηκε οριζόντια τομή στα 8.5 nsec η οποία περιλαμβάνει καταγραφές τόσο από τους τρεις σωλήνες, όσο και από το αργίλικό στρώμα. Επιλέχθηκαν λοιπόν 230 καταγραφές οι οποίες προέρχονταν από τους τέσσερις στόχους καθώς και από την άμμο. Οι καταγραφές αυτές και τα αντίστοιχα μιγαδικά χαρακτηριστικά χρησιμοποιήθηκαν σαν δείγματα εκπαίδευσης των αλγόριθμων καθοδηγούμενης ταξινόμησης.

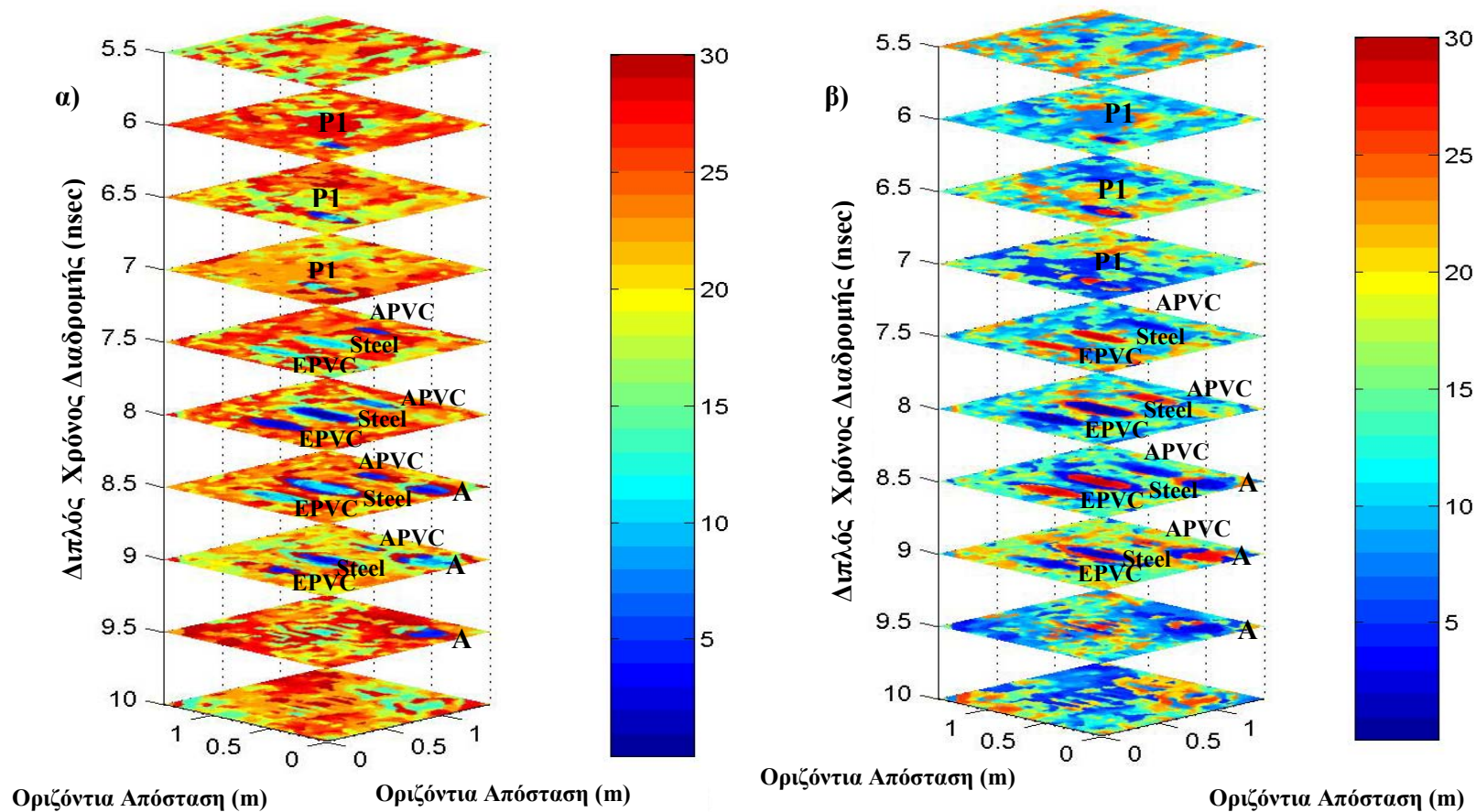
Η πρώτη ταξινόμηση πραγματοποιήθηκε με το δίκτυο πιθανότητας του οποίου η εκπαίδευση διήρκεσε μόλις 1 sec (σχήμα 5.9). Ο σωλήνας EPVC ταξινομήθηκε στην ομάδα 1, ο σωλήνας Steel στην ομάδα 2, ο σωλήνας APVC στην ομάδα 3, το αργίλικό στρώμα στην ομάδα 4, και η άμμος στην ομάδα 5, όπως ακριβώς είχε καθοριστεί κατά την εκπαίδευση. Οι λάθος ταξινομημένες καταγραφές, γύρω από τα θαμμένα αντικείμενα, είναι ελάχιστες. Η δεύτερη ταξινόμηση πραγματοποιήθηκε με το δίκτυο LVQ (σχήμα 5.10). Τα αποτελέσματα δε διαφέρουν σημαντικά από την ταξινόμηση με το δίκτυο πιθανότητας, αλλά ο χρόνος επεξεργασίας διήρκεσε 163 min για τις 15000 επαναλήψεις του αλγόριθμου.

5.6 Σύγκριση των μεθόδων ταξινόμησης κ-μέσων τιμών και δικτύων αυτο-οργανωμένης απεικόνισης.

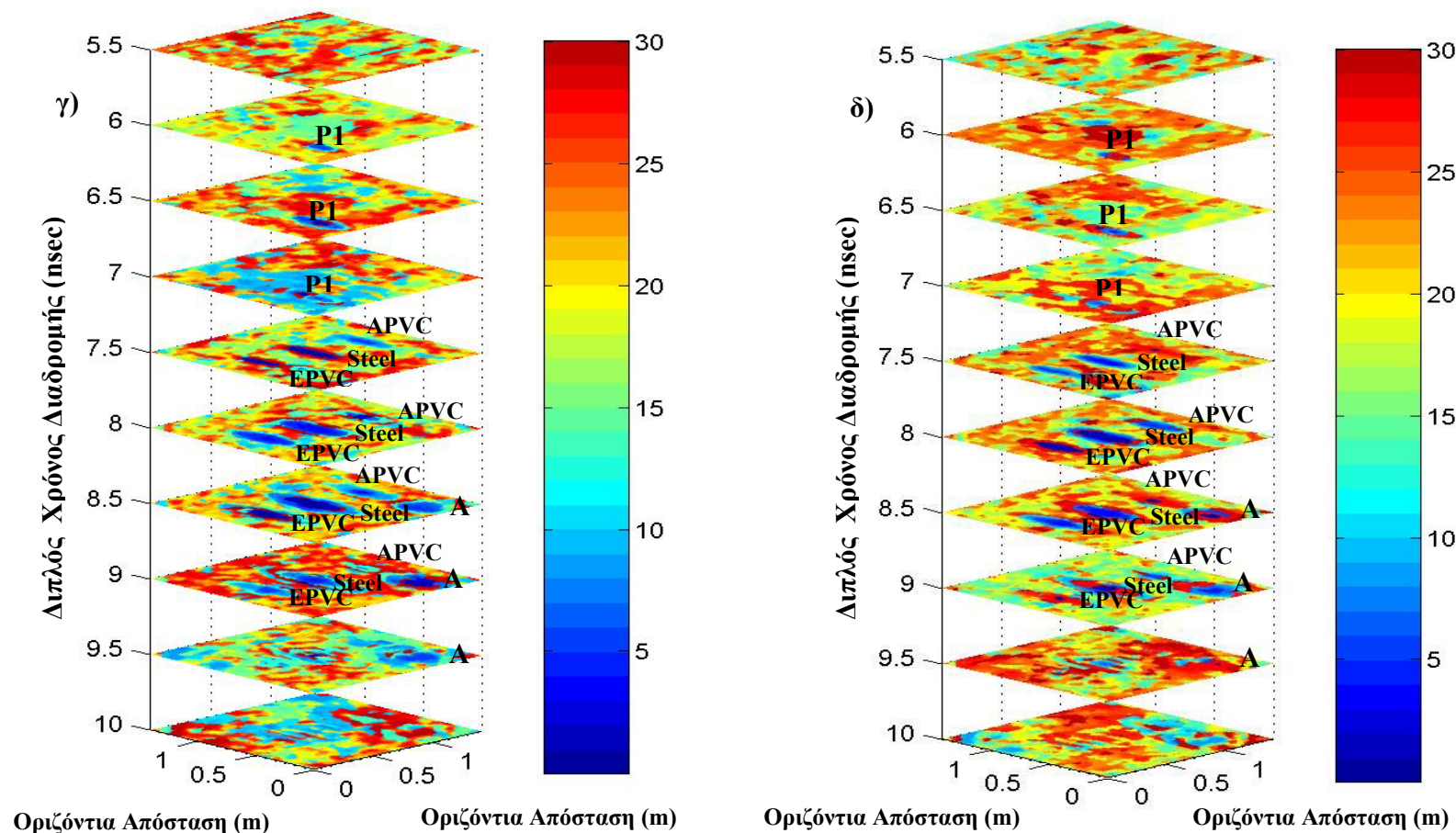
Για τη σύγκριση των μεθόδων ταξινόμησης κ-μέσων τιμών και δικτύων αυτο-οργανωμένης απεικόνισης χρησιμοποιήθηκαν τα δεδομένα της οριζόντιας τομής στα 8.5 nsec. Πέντε χαρακτηριστικά των ιχνών χρησιμοποιήθηκαν για την ανάλυση κυρίων συνιστωσών από τις οποίες οι πρώτες τρεις περιέχουν το 85% της συνολικής πληροφορίας. Οι κύριες συνιστώσες ταξινομήθηκαν σε 10 και 20 τάξεις.

Στις ταξινομημένες εικόνες με τη μέθοδο των κ-μέσων τιμών (5.11), οι στόχοι διαχωρίζονται από το περιβάλλον και μεταξύ τους αλλά τα ανακλώμενα κύματα από την άμμο ταξινομούνται σε πολλές ομάδες οι οποίες δεν έχουν τοπολογική συνέχεια προκαλώντας σύγχυση. Μάλιστα για τις 20 τάξεις το πρόβλημα είναι ακόμα πιο έντονο. Αντίθετα, το δίκτυο αυτο-οργανωμένης απεικόνισης οργάνωσε τις ομάδες με αποτέλεσμα οι ταξινομημένες εικόνες να μην διαφέρουν πολύ μεταξύ τους (σχήμα 5.12). Η διαφορά μεταξύ των ομάδων που αντιστοιχούν στους στόχους και των ομάδων που αντιστοιχούν στην άμμο είναι εμφανής, και παρόλο που η άμμος απεικονίζεται με πολλά χρώματα δείχνει αρκετά ομοιογενής σε σχέση με αυτό που παρουσιάζεται από την ταξινόμηση με τη μέθοδο των κ-μέσων τιμών.

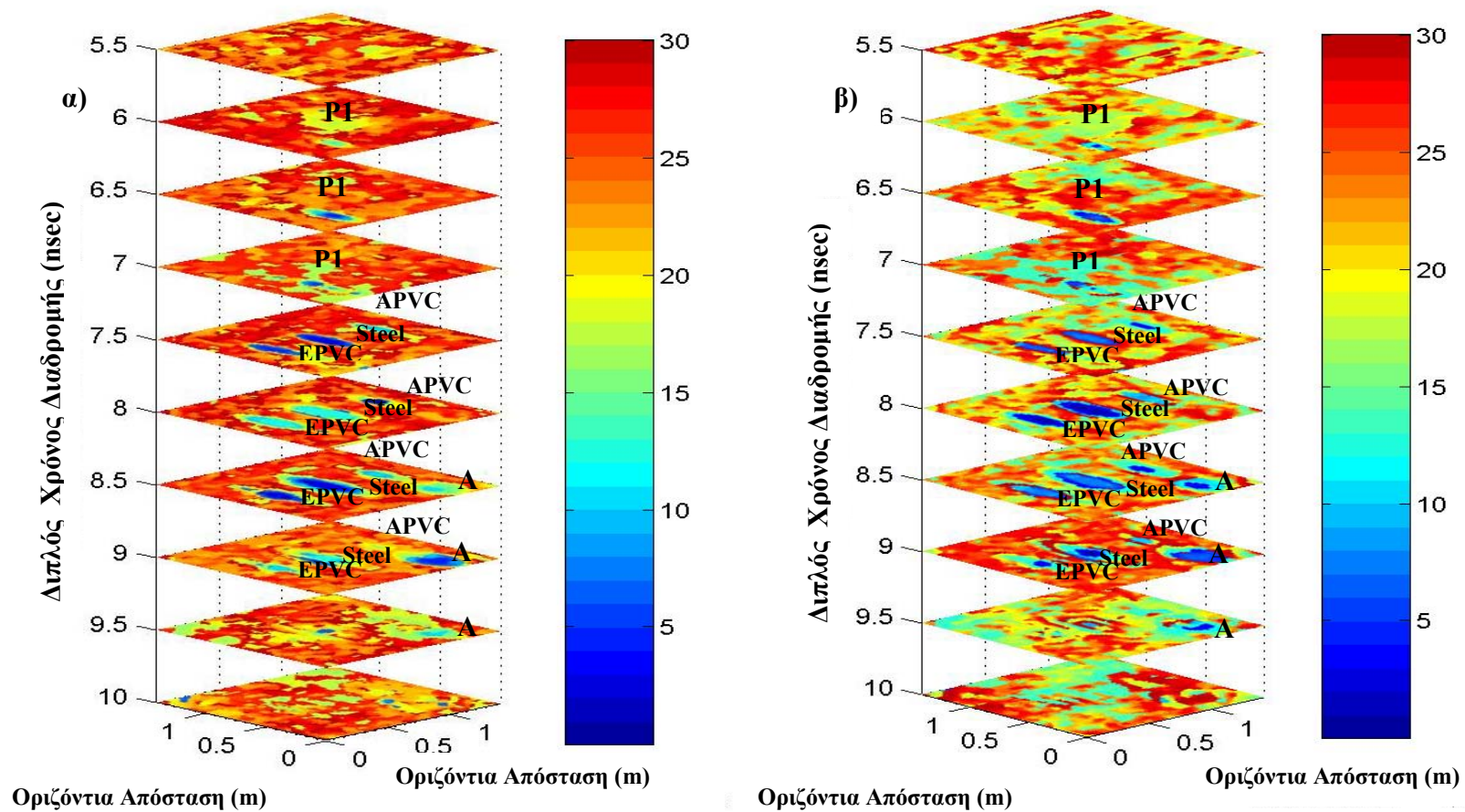
Η κύρια διαφορά μεταξύ των μεθόδων είναι στον τρόπο οργάνωσης των ομάδων. Με τον όρο οργάνωση εννοείται η συνάφεια μεταξύ των διαδοχικών ομάδων. Ο αλγόριθμος των κ-μέσων τιμών δημιουργεί ομάδες με κριτήριο την ελαχιστοποίηση της μέσης τετραγωνικής απόστασης των καταγραφών από τα αντίστοιχα κέντρα ομαδοποίησης. Ο αλγόριθμος μεγιστοποιεί τη μεταβλητότητα (variance) μεταξύ των ομάδων, ενώ την ελαχιστοποιεί μεταξύ των καταγραφών που ανήκουν στην ίδια ομάδα. Αντίθετα τα δίκτυα αυτό-οργανωμένης απεικόνισης δημιουργούν τοπολογικά οργανωμένες ομάδες συσχετίζοντας τις. Στα σχήματα (5.13 - 5.14) παρουσιάζονται διαγράμματα διασποράς των τριών κύριων συνιστωσών και οι ομάδες που δημιουργήθηκαν από τους αλγόριθμους των κ-μέσων τιμών και του δικτύου αυτό-οργανωμένης απεικόνισης όπου υπάρχει τοπολογική οργάνωση.



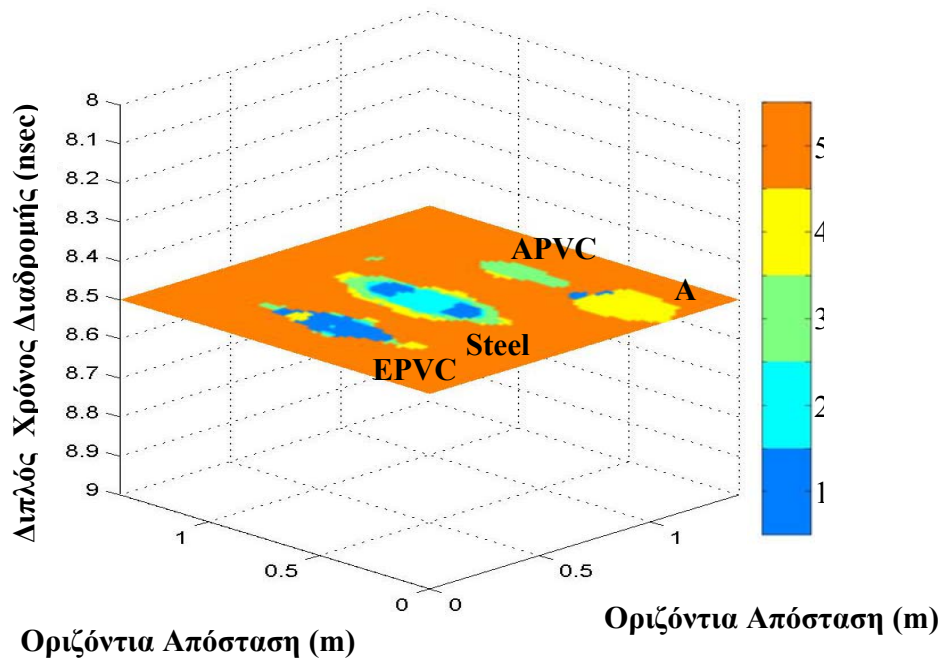
Σχήμα 5.7. Το αποτέλεσμα της ταξινόμησης των αρχικών δεδομένων με τις μεθόδους κ-μέσων τιμών και δικτύου αυτο-οργανωμένης απεικόνισης. Στην περίπτωση (α) τα δεδομένα ταξινομήθηκαν με τη μέθοδο κ-μέσων τιμών σε 200 τάξεις και μετά σε 30 με το δίκτυο αυτο-οργανωμένης απεικόνισης. Στην περίπτωση (β) ταξινομήθηκε κάθε τομή ξεχωριστά σε 20 τάξεις και μετά τα κέντρα τους οργανώθηκαν σε 30 τάξεις.



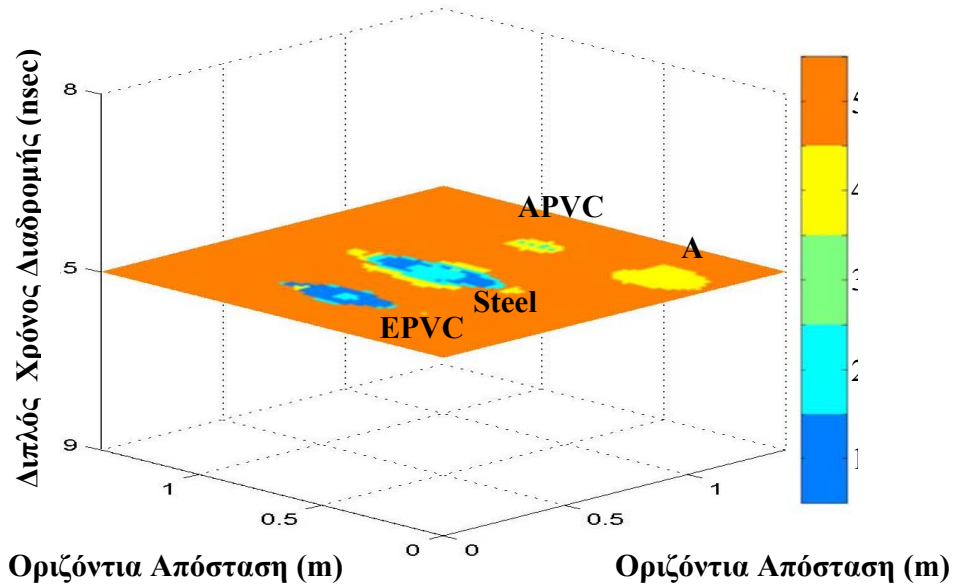
Σχήμα 5.7 (συνέχεια). Το αποτέλεσμα της ταξινόμησης των τριών κύριων συνιστωσών με τις μεθόδους κ-μέσων τιμών και δικτύου αυτο-οργανωμένης απεικόνισης. Στην περίπτωση (γ) οι κύριες συνιστώσες ταξινομήθηκαν με τη μέθοδο κ-μέσων τιμών σε 200 τάξεις και μετά σε 30 με το δίκτυο αυτο-οργανωμένης απεικόνισης. Στην περίπτωση (δ) ταξινομήθηκε κάθε τομή ξεχωριστά σε 20 τάξεις και μετά τα κέντρα οργανώθηκαν σε 30 τάξεις.



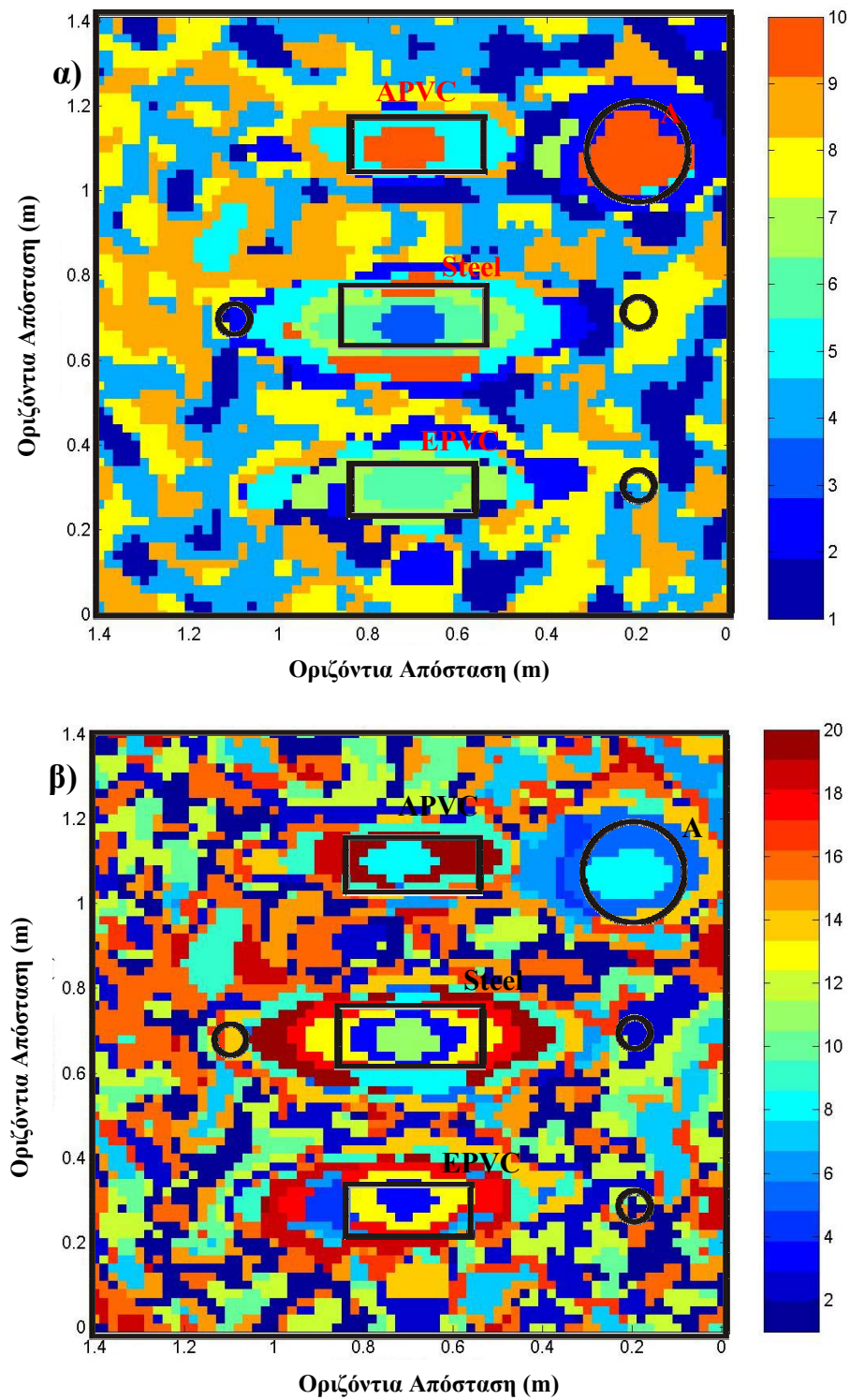
Σχήμα 5.8. Ταξινόμηση των χαρακτηριστικών (α) και των κύριων συνιστωσών (β) με το δίκτυο αυτό-οργανωμένης απεικόνισης, χρησιμοποιώντας σαν δείγματα εκπαίδευσης καταγραφές που προέρχονται από τους στόχους και την άμμο.



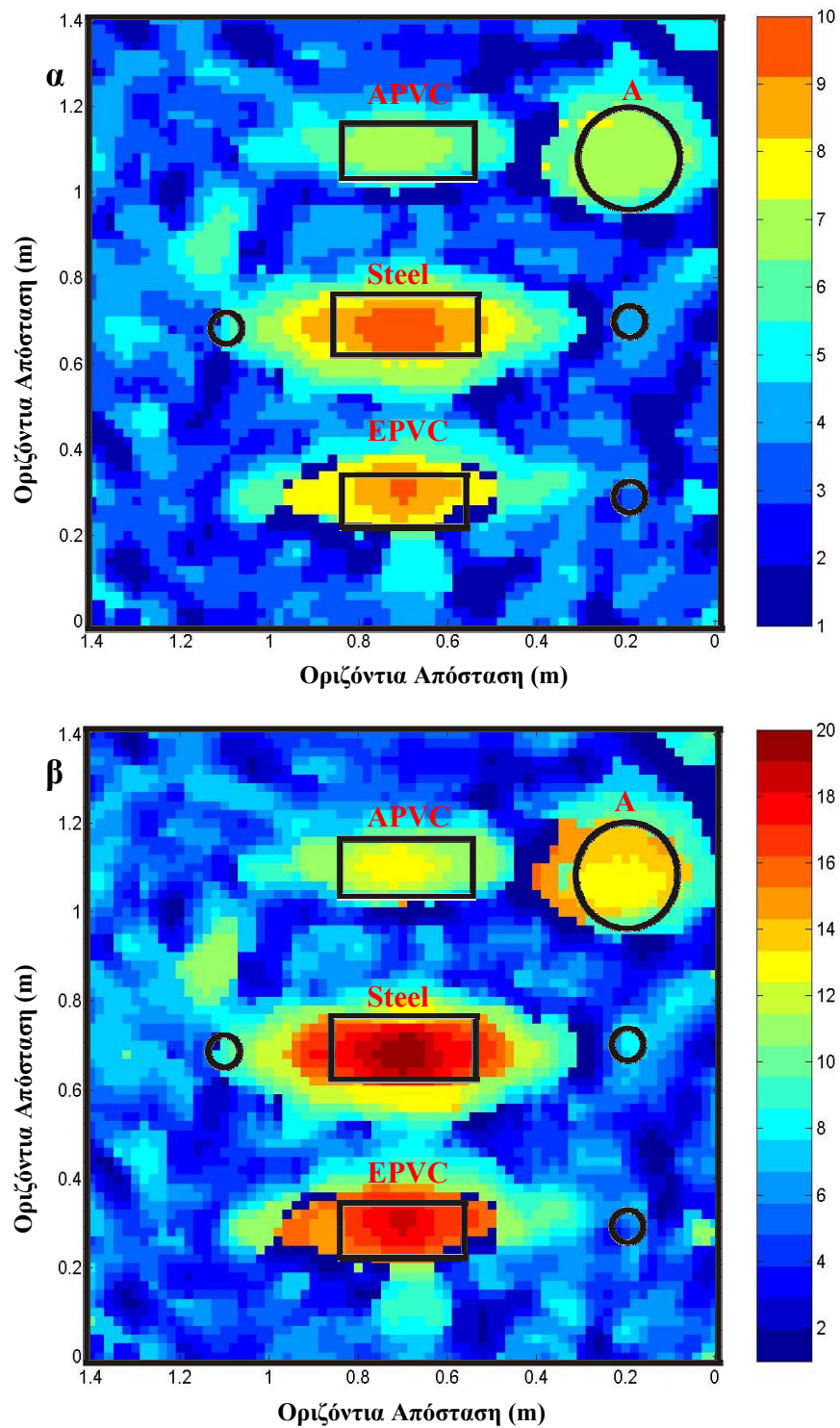
Σχήμα 5.9. Καθοδηγούμενη ταξινόμηση με το δίκτυο πιθανότητας σε πέντε ομάδες.



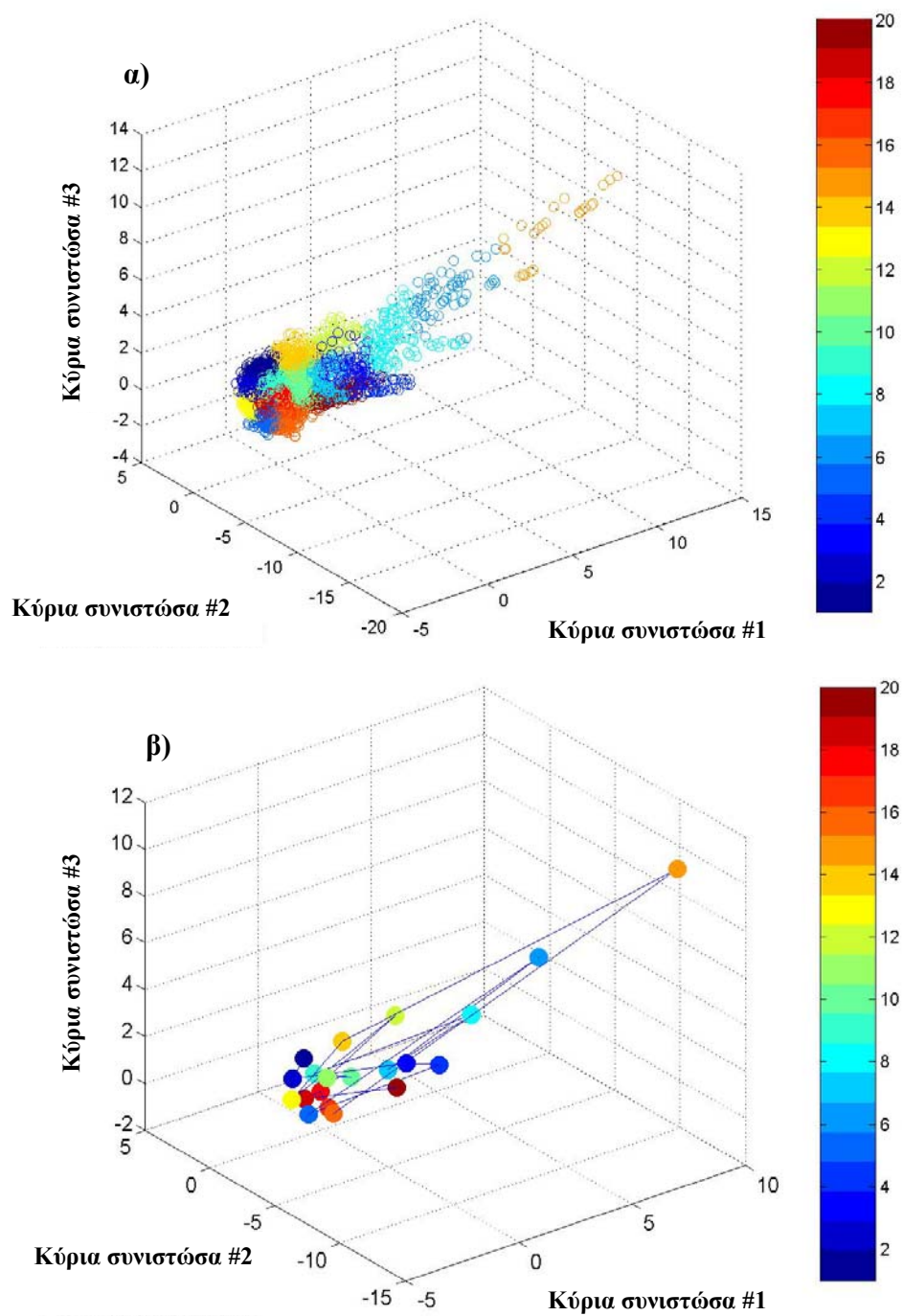
Σχήμα 5.10. Καθοδηγούμενη ταξινόμηση με το δίκτυο LVQ σε πέντε ομάδες.



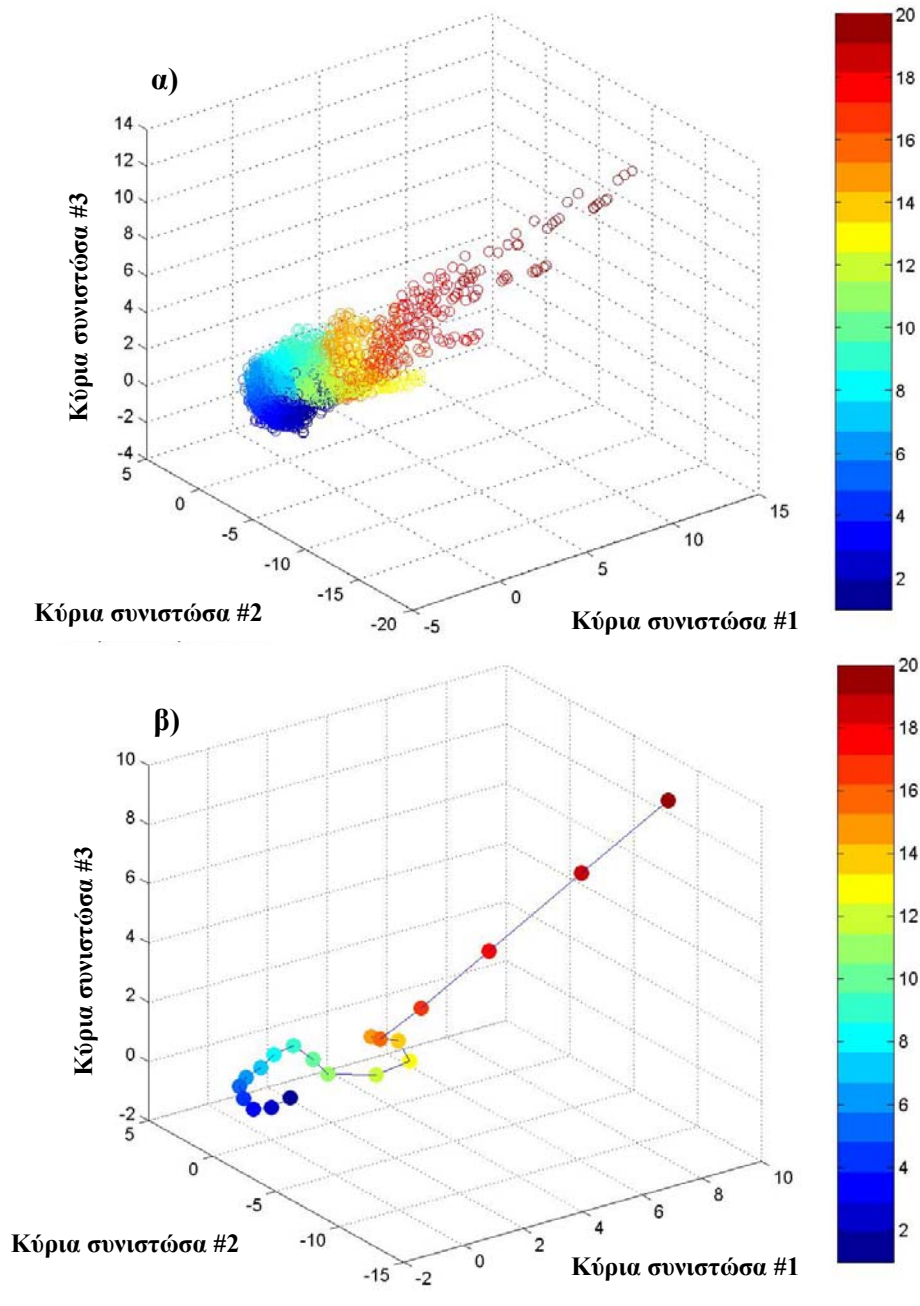
Σχήμα 5.11. Οριζόντια τομή ταξινομημένη με τη μέθοδο των κ-μέσων τιμών σε 10 τάξεις (α) και 20 τάξεις (β).



Σχήμα 5.12. Οριζόντια τομή ταξινομημένη με τη μέθοδο των δικτύων αυτό-οργανωμένης απεικόνισης σε 10 τάξεις (α) και 20 τάξεις (β).



Σχήμα 5.13. (α) Διάγραμμα διασποράς των δεδομένων με τη μέθοδο των κ-μέσων τιμών και (β) διάγραμμα διασποράς για τα κέντρα συγκέντρωσης.



Σχήμα 5.14. (α) Διάγραμμα διασποράς των δεδομένων με τη μέθοδο των δικτύων αυτο-οργανωμένης απεικόνισης και (β) διάγραμμα διασποράς για τα κέντρα συγκέντρωσης. Παρατηρείται ότι τα κέντρα είναι τοπολογικά οργανωμένα.

5.7 Συμπεράσματα.

Σε αυτό το κεφάλαιο προτείνονται μέθοδοι ταξινόμησης για οριζόντιες τομές γεωραντάρ. Όπως ήδη αναφέρθηκε οι τομές από το ελεγχόμενο πείραμα είναι δυνατό να ερμηνευθούν και χωρίς τη χρήση των μεθόδων αυτών, αλλά χρησιμοποιήθηκαν με στόχο να φανεί αν συμβάλλουν ουσιαστικά στην ερμηνεία οριζόντιων τομών γεωραντάρ.

Όσον αφορά στην αυτόματη ταξινόμηση προέκυψε ότι ο διαχωρισμός των θαμμένων αντικειμένων τόσο μεταξύ τους όσο και από το περιβάλλον τους είναι καλύτερος όταν για την εκπαίδευση του αλγόριθμου χρησιμοποιούνται χαρακτηριστικά δείγματα των καταγραφών.

Στις δοκιμές για την καθοδηγούμενη ταξινόμηση των δεδομένων, όπου υπάρχουν α-priori πληροφορίες σχετικά με τις καταγραφές, συγκρίθηκαν τα δίκτυα πιθανότητας και διανυσματικής μάθησης. Και οι δυο αλγόριθμοι έδωσαν εξίσου καλά αποτελέσματα. Ο υπολογιστικός χρόνος για την εκπαίδευση των δικτύων πιθανότητας ήταν κατά πολύ μικρότερος σε σχέση με τα δίκτυα διανυσματικής μάθησης.

Αποδείχτηκε ότι αυτές οι μέθοδοι μπορούν να ταξινομήσουν δεδομένα που αφορούν σε ανακλώμενα κύματα από θαμμένα αντικείμενα, σε διαφορετικές τάξεις, συμβάλλοντας έτσι στην ερμηνεία τους.

ΚΕΦΑΛΑΙΟ 6**ΤΑΞΙΝΟΜΗΣΗ ΔΕΔΟΜΕΝΩΝ ΓΕΩΡΑΝΤΑΡ ΑΠΟ ΤΗΝ ΑΡΧΑΙΑ ΑΠΤΕΡΑ****6.1 Εισαγωγή.**

Στο κεφάλαιο αυτό παρουσιάζεται η ταξινόμηση δεδομένων γεωραντάρ που προέκυψαν από μετρήσεις σε χώρο με αρχαιολογικό ενδιαφέρον. Συγκεκριμένα, οι μετρήσεις πραγματοποιήθηκαν στο πλαίσιο του μαθήματος ‘Μεγάλες Ασκήσεις Υπαίθρου IV’ του τμήματος Μηχανικών Ορυκτών Πόρων του Πολυτεχνείου Κρήτης τον Σεπτέμβριο του 2005, στην αρχαία πόλη Άπτερα, λίγα χιλιόμετρα ανατολικά από την πόλη των Χανίων.

Τα δεδομένα που παρουσιάζονται συλλέχθηκαν πάνω από τρίκλιτη δεξαμενή, της οποίας η οροφή επιχωματώθηκε από τους μοναχούς του μοναστηριού του Άγιου Ιωάννη του Θεολόγου την περίοδο του 12^{ου} αιώνα μ.Χ., για να χρησιμοποιηθεί ως αλώνι (σχήμα 6.1). Τα δεδομένα έχουν εκπαιδευτικό ενδιαφέρον, αφού από τα βόρεια η δεξαμενή είναι

επισκέψιμη. Επίσης, με αυτόν τον τρόπο είναι εύκολο να γίνει έλεγχος της ερμηνείας των δεδομένων.

6.2 Ιστορικά στοιχεία για την αρχαία πόλη Άπτερα.

Η αρχαία πόλη Άπτερα βρίσκεται κοντά στο χωριό Μεγάλα Χωράφια της επαρχίας Αποκορώνου του νομού Χανίων. Φημίζεται, εκτός από την αρχαιολογική της σημασία, για τη στρατηγική της θέση αφού προσφέρει πανοραμική θέα προς τον κόλπο της Σούδας καθώς και τα Λευκά Όρη.

Για την προέλευση του ονόματος της πόλης υπάρχουν διάφορες εκδοχές μεταξύ των οποίων η επικρατέστερη είναι ότι προέρχεται από το αντίστοιχο επίθετο της θεάς Άρτεμης (Άρτεμις Άπτερα). Άλλη εκδοχή αποδίδει το όνομα στο μουσικό αγώνα μεταξύ Μουσών και Σειρήνων κατά τον οποίο οι Σειρήνες ηττήθηκαν, απέβαλαν τα φτερά τους, έγιναν λευκές και έπεσαν στη θάλασσα.

Οι αρχαιότερες αναφορές για την πόλη χρονολογούνται την περίοδο του 14^{ου} – 13^{ου} αιώνα π.Χ. σε πινακίδες της Κνωσού σε Γραμμική Β. Η κατοίκηση της πόλης συνεχίστηκε κατά τα πρώιμα βυζαντινά χρόνια αλλά καταστράφηκε από ισχυρό σεισμό τον 7^ο αιώνα μ.Χ. Τον 12^ο αιώνα μ.Χ. είχε ήδη χτιστεί στον χώρο το μοναστήρι του Άγιου Ιωάννη του Θεολόγου, το οποίο διατηρείται μέχρι σήμερα. Το κάστρο που βρίσκεται στα ΒΑ του αρχαιολογικού χώρου χτίστηκε από τους Τούρκους με αφορμή την επανάσταση του 1866 χρησιμοποιώντας τα αρχαία ερείπια σαν οικοδομικά υλικά. Επίσης, τα αρχαία ερείπια χρησιμοποιήθηκαν ως οικοδομικά υλικά, τόσο κατά τη διάρκεια της Τουρκοκρατίας, όσο και της Ενετοκρατίας για την ανέγερση κάστρων αλλά και για την ανοικοδόμηση του οικισμού στα Μεγάλα Χωράφια. Από το τείχος της πόλης σώζεται μόνο ένα τμήμα μήκους τεσσάρων (4) χιλιομέτρων.

Αξιοθέατα του χώρου αποτελούν οι θολωτές δεξαμενές στα δυτικά (σχήματος Γ) και στα βορειοανατολικά (τρίκλιτη δεξαμενή (σχήμα 6.2)), το υδρευτικό σύστημα το οποίο αποτελείται από πολυάριθμα πηγάδια και κανάλια, ο ελληνιστικός τοίχος των επιγραφών, ο οποίος καταστράφηκε στις αρχές του περασμένου αιώνα από τους κατοίκους της περιοχής για να χρησιμοποιήσουν τα υλικά του, το διμερές ιερό, το μικρό θέατρο και ένας μικρός δωρικός ναός αφιερωμένος στον Απόλλωνα.

Τα περισσότερα ευρήματα των ανασκαφών εκτίθενται στο Αρχαιολογικό Μουσείο των Χανίων, ενώ υπεύθυνη για τις ανασκαφές οι οποίες συνεχίζονται μέχρι σήμερα είναι η ΚΕ Εφορεία Προϊστορικών και Κλασικών Αρχαιοτήτων.

6.3 Συλλογή των δεδομένων.

Για τη συλλογή των δεδομένων χρησιμοποιήθηκε το σύστημα *Terra SIRc* της εταιρίας *GSSI Inc.*, με κεντρική συχνότητα εκπομπής 400 MHz. Πραγματοποιήθηκαν έντεκα (11) παράλληλες γραμμές μελέτης μήκους δεκατεσσάρων (14) μέτρων στη διεύθυνση Δ-Α. Το διάστημα δειγματοληψίας ορίστηκε ίσο με 195 psec, η ισοαπόσταση των σταθμών παρατήρησης 10 cm και η ισοαπόσταση των γραμμών, 1 m.

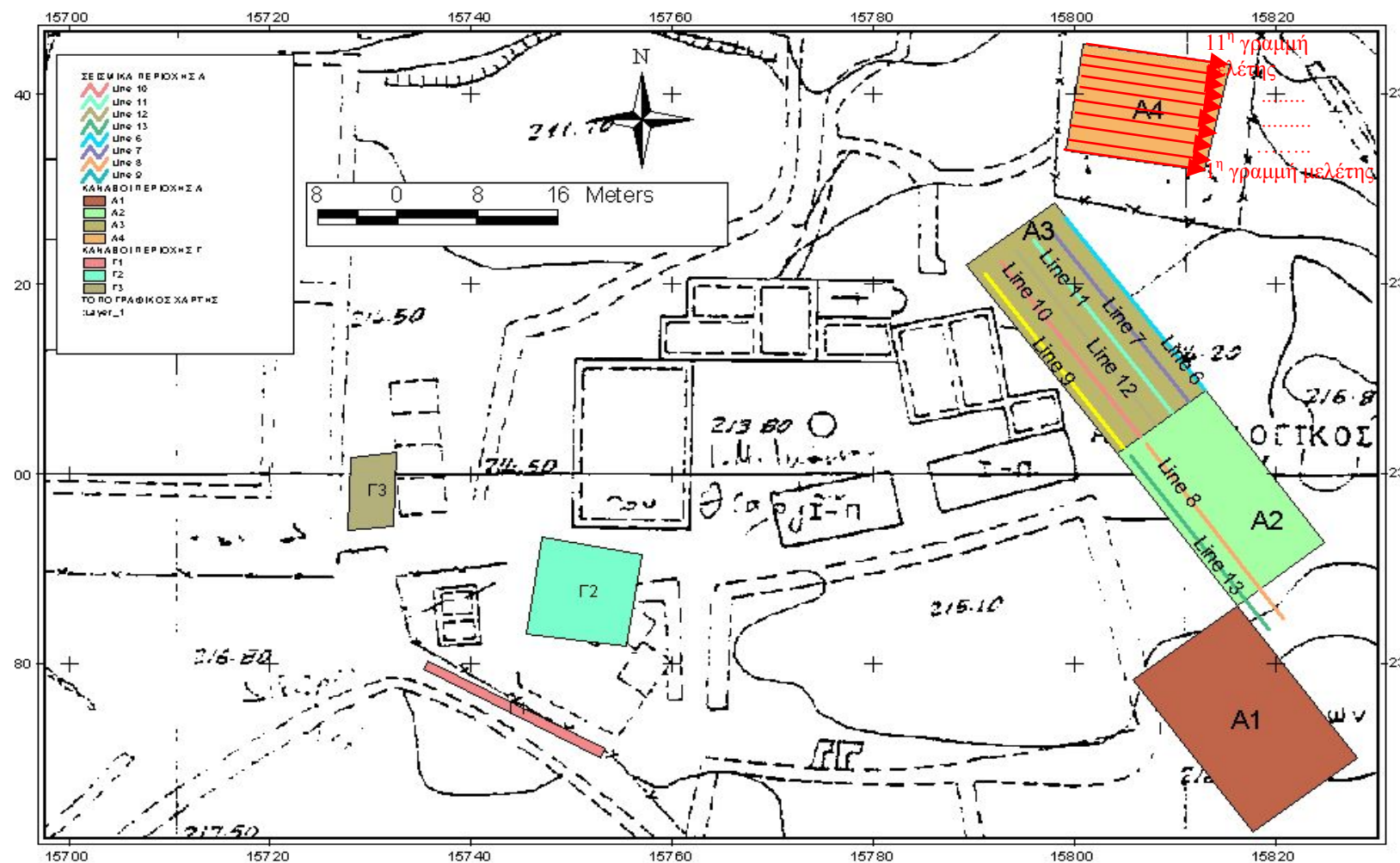
Ο κάνναβος σχεδιάστηκε μέσα στο αλώνι, ώστε οι γραμμές μελέτης να διασχίζουν κάθετα τη δεξαμενή, ενώ για λόγους ασφάλειας δεν πραγματοποιήθηκαν μετρήσεις στο βόρειο όριο της δεξαμενής. Στο σχήμα 6.3 φαίνεται ο τοπογραφικός χάρτης της περιοχής, όπου διακρίνεται ο κάνναβος A4 στον οποίο πραγματοποιήθηκαν οι μετρήσεις.



Σχήμα 6.1. Η περιοχή μελέτης. Παλαιότερα χρησιμοποιούνταν από τους μοναχούς σαν αλώνι.



Σχήμα 6.2. Το εσωτερικό της τρίκλιτης δεξαμενής.



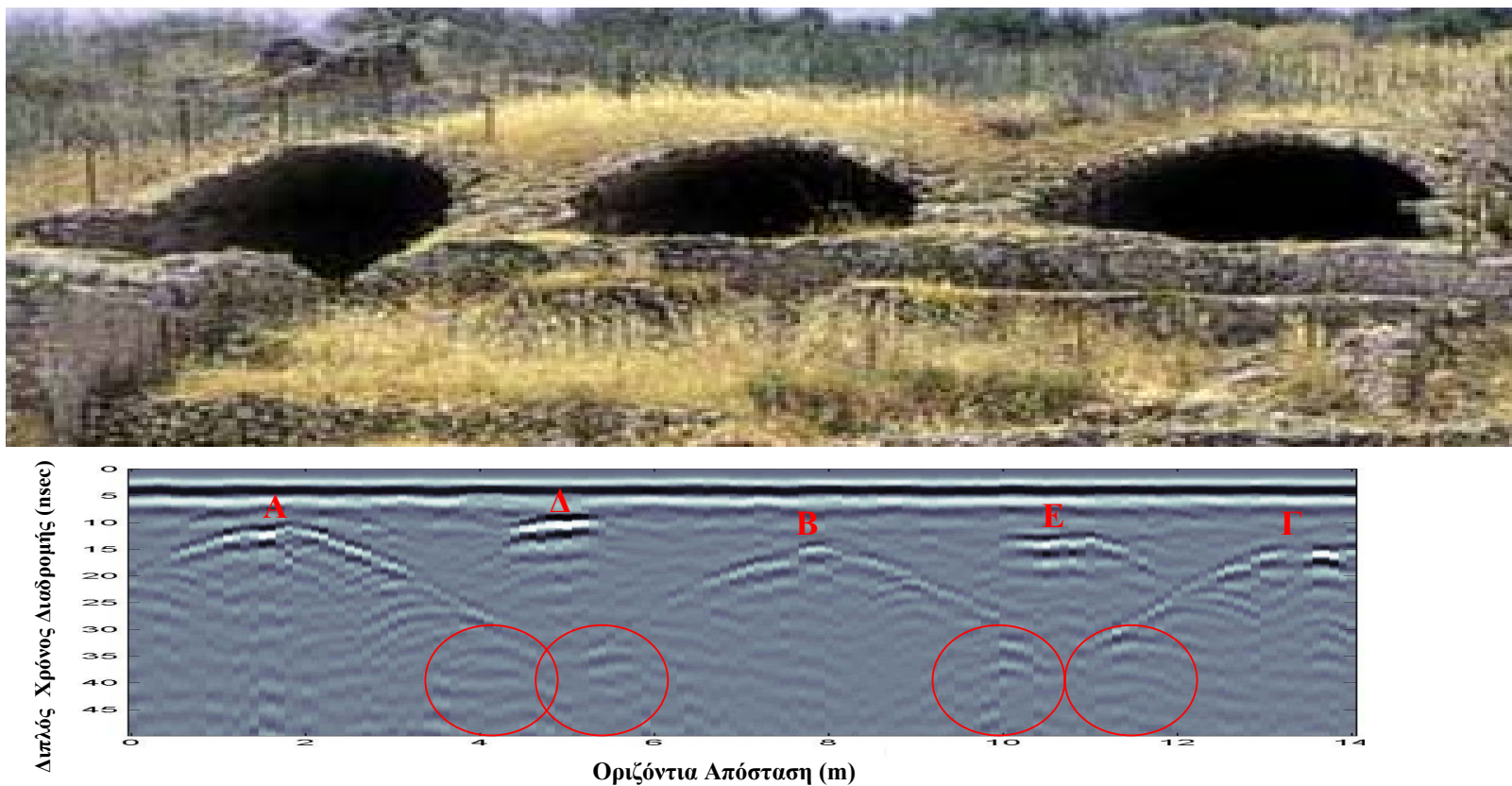
Σχήμα 6.3. Τοπογραφικός χάρτης της περιοχής. Διακρίνεται ο κήναβος Α4 στον οποίο πραγματοποιήθηκαν οι μετρήσεις.

6.4 Επεξεργασία των δεδομένων.

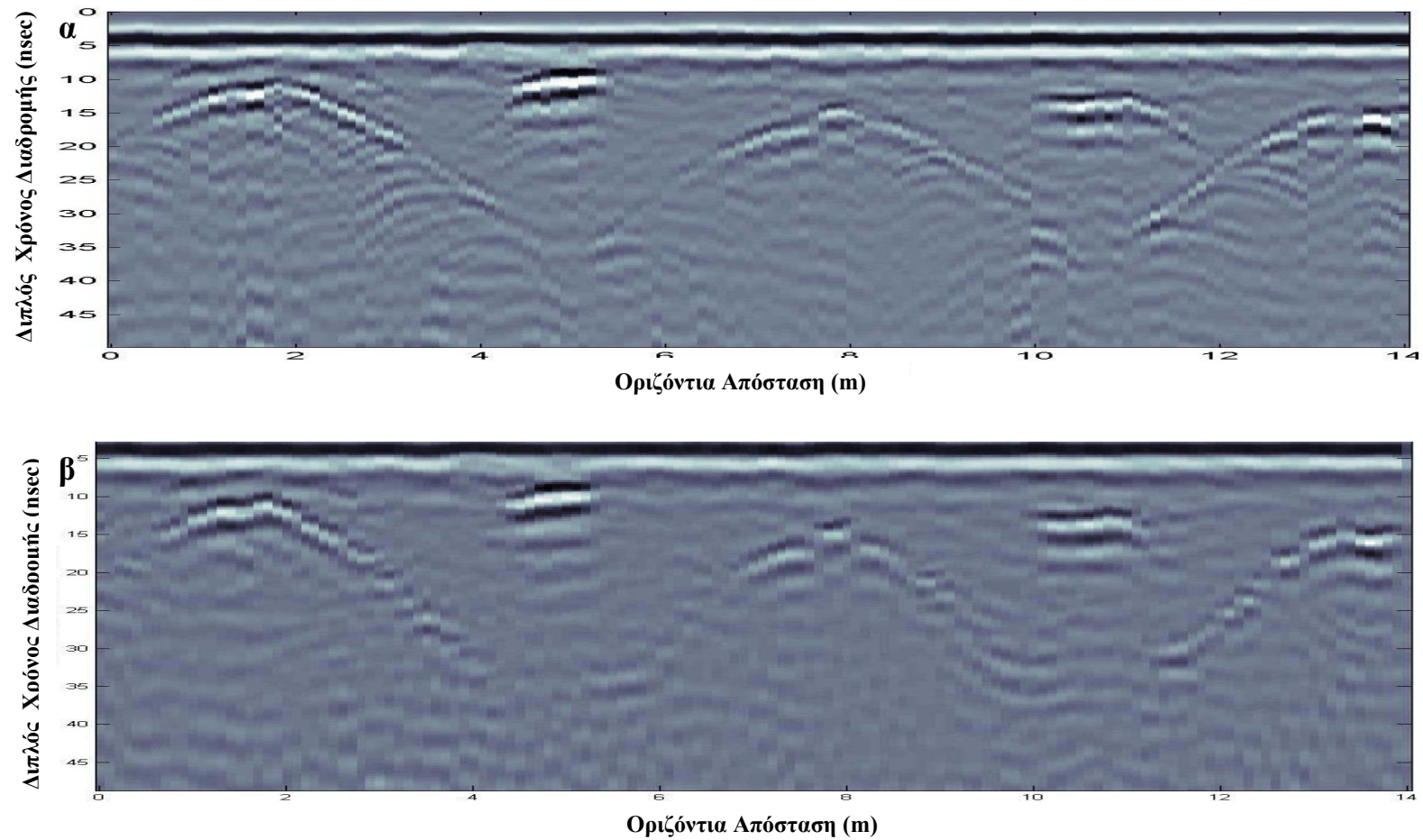
Τα πρώτα στάδια της επεξεργασίας στόχευαν στην απαλοιφή του υψίσυχνου θορύβου, την ενίσχυση των ασθενών καταγραφών και τη χωροθέτηση. Στην εικόνα του σχήματος 6.4 φαίνονται τα δεδομένα της 4^{ης} γραμμής μελέτης πριν τη χωροθέτηση καθώς και αντίστοιχη φωτογραφία από την δεξαμενή. Τα απ' ευθείας κύματα αέρα και εδάφους έχουν συνολική διάρκεια περίπου 5 nsec και φαίνεται να επηρεάζονται από την ανομοιογένεια στο ρηχό υπέδαφος. Με την πρώτη ματιά παρατηρεί κανείς τα ανακλώμενα κύματα από την οροφή της τρίκλιτης δεξαμενής (Α, Β, Γ). Ο χρόνος καταγραφής αυτών των κυμάτων κυμαίνεται από 10 nsec μέχρι 35 nsec. Οι καταγραφές για μικρούς χρόνους είναι ισχυρές και χωρίς πολύ θόρυβο, ενώ σε μεγαλύτερους είναι ασθενέστερες. Όταν εφαρμόστηκε ισχυρότερο φίλτρο ενίσχυσης, υποβαθμίστηκε η ποιότητα της εικόνας, αφού ενισχύθηκε κατά πολύ ο θόρυβος. Επίσης παρατηρούνται Η/Μ κύματα από δυο ισχυρούς ανακλαστήρες (αγωγοί, τετραγωνικής διατομής για τη συλλογή υδάτων στις δεξαμενές) σε ίδιους περίπου χρόνους και σε αποστάσεις 5m και 11m αντίστοιχα (Δ, Ε). Αξιοσημείωτες είναι οι καταγραφές των περιθλώμενων κυμάτων από τα όρια κάθε θόλου που έχουν σχήμα υπερβολής (κόκκινος κύκλος).

Πρέπει να τονιστεί ότι η χωροθέτηση των δεδομένων ήταν προβληματική, διότι υπήρχε μεταβολή της ταχύτητας προς όλες τις κατευθύνσεις και ήταν αδύνατο να κατασκευαστεί ένα σωστό μοντέλο ταχυτήτων. Το πρόβλημα υπεραπλουστεύτηκε και χρησιμοποιήθηκε ενιαία ταχύτητα διάδοσης των Η/Μ κυμάτων ίση με 10 cm/nsec. Στο σχήμα 6.5, η 4^η τομή γεωραντάρ πριν και μετά τη χωροθέτηση υποδεικνύει ότι, η χωροθέτηση είναι αποτελεσματική μέχρι εκεί που τα Η/Μ κύματα ανακλώνται στις καμάρες του τρίκλιτου, αφού η ταχύτητα είναι σταθερή.

Σε επόμενο στάδιο υπολογίστηκαν τα στιγμιαία και γεωμετρικά χαρακτηριστικά των καταγραφών, όπως το στιγμιαίο πλάτος και το στιγμιαίο πλάτος σε κλίμακα dB, η στιγμιαία φάση και η συνάφεια.



Σχήμα 6.4. Αντιπαράθεση της φωτογραφίας της τρίκλιτης δεξαμενής (επάνω) με τα δεδομένα της 4^{ης} γραμμής μελέτης (κάτω). Σε κόκκινο κύκλο διακρίνονται τα περιθλώμενα κύματα από τα όρια των θόλων



Σχήμα 6.5. Τα δεδομένα πριν (α) και μετά (β) τη χωροθέτηση. Το φαινόμενο της περίθλασης των κυμάτων στα όρια της κατασκευής έχουν απαλειφθεί.

Τέλος, τα δεδομένα ταξινομήθηκαν με τη μέθοδο των δικτύων αυτο-οργανωμένης απεικόνισης. Έγιναν δοκιμές για την αυτόματη ταξινόμηση των κυματιδίων των ανακλαστήρων σε 20 τάξεις και δοκιμές για την αυτόματη ταξινόμηση όλων των ψευδο-τρισιδιάστατων δεδομένων επίσης σε 20 τάξεις.

6.5 Αυτόματη ταξινόμηση ανακλαστήρα.

Η αυτόματη ταξινόμηση εφαρμόστηκε σε χωροθετημένα δεδομένα μετά την απομάκρυνση της επαγωγικής συνιστώσας και την αντιστάθμιση της γεωμετρικής διασποράς.

Επιλέχθηκαν δια χειρός δύο ανακλαστήρες με ενδιαφέρον (σχήμα 6.6). Ο πρώτος αντιστοιχεί στην οροφή της τρίκλιτης δεξαμενής (σχήμα 6.6α με την κόκκινη γραμμή), ενώ ο δεύτερος περιλαμβάνει τμήμα της οροφής (κοίλου μέρος των θόλων) καθώς και τους υδαταγωγούς (σχήμα 6.6β με κόκκινη γραμμή). Έπειτα, ορίστηκε συμμετρικά από αυτούς, χρονικό παράθυρο συνολικής διάρκειας 4 nsec (σχήματα 6.7α και 6.7β με τις κίτρινες γραμμές αντίστοιχα), όση δηλαδή η διάρκεια του παλμού. Τα δεδομένα αυτά τροφοδοτήθηκαν απευθείας σε δίκτυα αυτο-οργανωμένης απεικόνισης και ταξινομήθηκαν σε είκοσι (20) τάξεις αντίστοιχα. Στο σχήμα 6.8 φαίνεται το αποτέλεσμα της ταξινόμησης των δύο ανακλαστήρων. Τα ταξινομημένα δεδομένα (έγχρωμο) έχουν υπερτεθεί στα αρχικά (ασπρόμαυρο).

Όσον αφορά στον πρώτο ανακλαστήρα (σχήμα 6.8α), οι ισχυρές ανακλάσεις στα ρηγά ανήκουν στις ομάδες 14 έως 20, ενώ οι βαθύτερες ανακλάσεις στις ομάδες 1 έως 13. Εξαιτίας του ότι ο ανακλαστήρας αποτελείται από το ίδιο υλικό, όλα τα δεδομένα θα έπρεπε να ταξινομηθούν στην ίδια ομάδα. Όμως λόγω του ότι η οροφή της δεξαμενής, επιχρωματώθηκε με διάφορα υλικά για την κατασκευή του αλωνιού και η επιχρωμάτωση παρουσιάζει διαφορετικό πάχος, το σχήμα του κυματιδίου αλλάζει από θέση σε θέση. Στο σχήμα 6.9 απεικονίζονται τα ταξινομημένα κυματίδια σε χρώματα ανάλογα με την ομάδα στην οποία ανήκουν. Παρατηρείται ότι τα κυματίδια των ομάδων 1 έως 5 και 16 έως 20 είναι παρόμοια, το οποίο δικαιολογεί και τις μεταξύ τους εναλλαγές (σχήμα 6.8α). Επίσης, παρατηρούνται ομοιότητες στα κυματίδια των ομάδων 6 έως 10 καθώς και 11 έως 15.

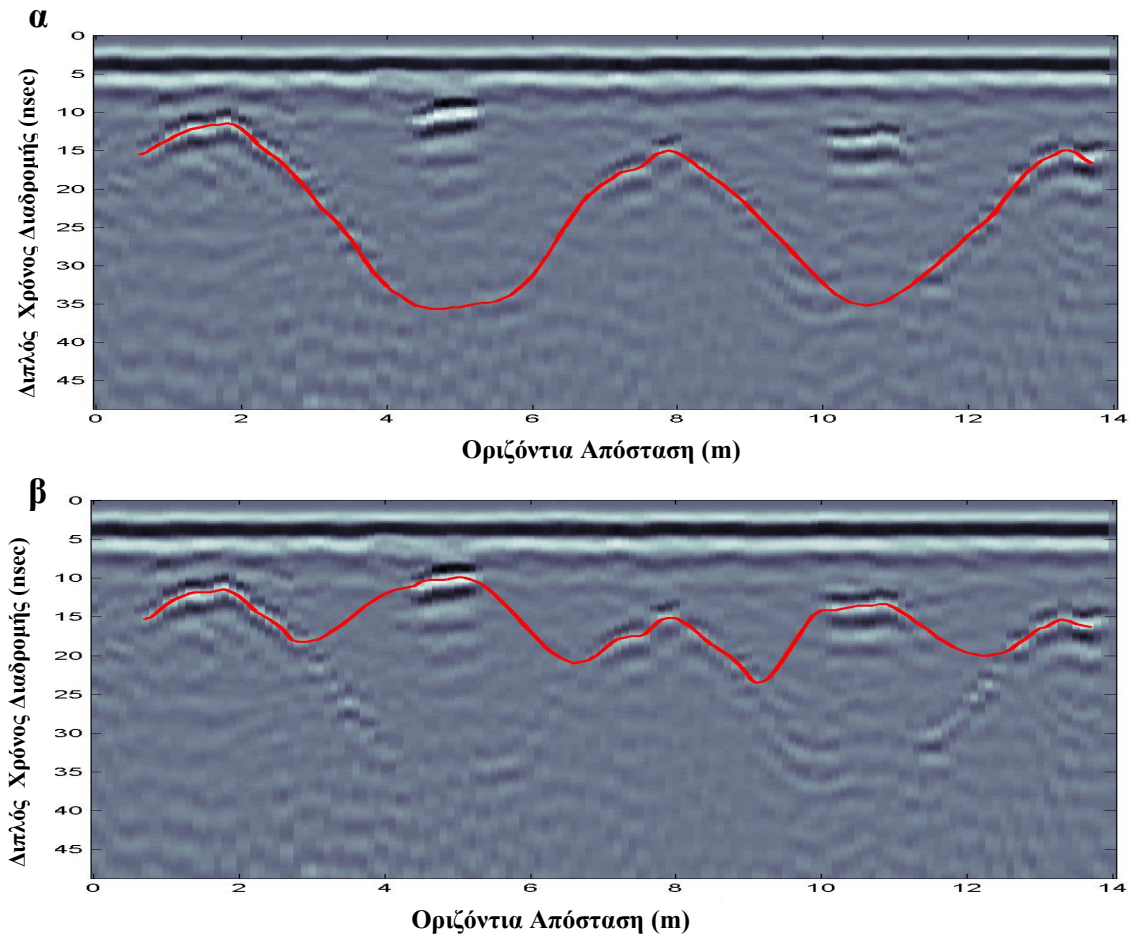
Όσον αφορά στο δεύτερο ανακλαστήρα (σχήμα 6.8β), παρατηρείται κατ' αρχήν σαφής διαχωρισμός μεταξύ των ανακλάσεων στους θόλους και στους υδαταγωγούς. Τα δεδομένα από τους υδαταγωγούς έχουν ταξινομηθεί στις ομάδες 3 έως 6. Οι ανακλάσεις ανάμεσα στους θόλους έχουν ταξινομηθεί στις ομάδες 19 και 20, ενώ οι θόλοι έχουν ταξινομηθεί στις ομάδες 7 έως 18 καθώς και 1 έως 2. Στο σχήμα 6.10 φαίνονται τα κυματίδια κάθε ομάδας.

Στο επόμενο στάδιο υπολογίστηκαν το στιγμιαίο πλάτος, το στιγμιαίο πλάτος σε κλίμακα dB, η στιγμιαία φάση και η συνάφεια των δεδομένων. Το στιγμιαίο πλάτος και το στιγμιαίο πλάτος σε κλίμακα dB (σχήματα 6.11α και 6.11β αντίστοιχα) τονίζουν τις ανακλάσεις από τους θόλους, αφού αυτές έχουν μεγάλες τιμές. Η στιγμιαία φάση (σχήμα 6.11γ) τονίζει μεν το περίγραμμα των ανακλαστήρων, αλλά έχει πολλές ασυνέχειες, οι οποίες οφείλονται είτε στο ανομοιογενές υπερκείμενο στρώμα είτε στη διάβρωση των θόλων. Η συνάφεια (σχήμα 6.11δ) των δεδομένων δίνει μικρές τιμές για τις ανακλάσεις από τους θόλους και μεγάλες τιμές για τους αγωγούς. Αυτό οφείλεται στο ότι οι αγωγοί καταγράφονται πιο ευδιάκριτα στις τομές του γεωραντάρ περιέχοντας λιγότερο θόρυβο από τα υπερκείμενα.

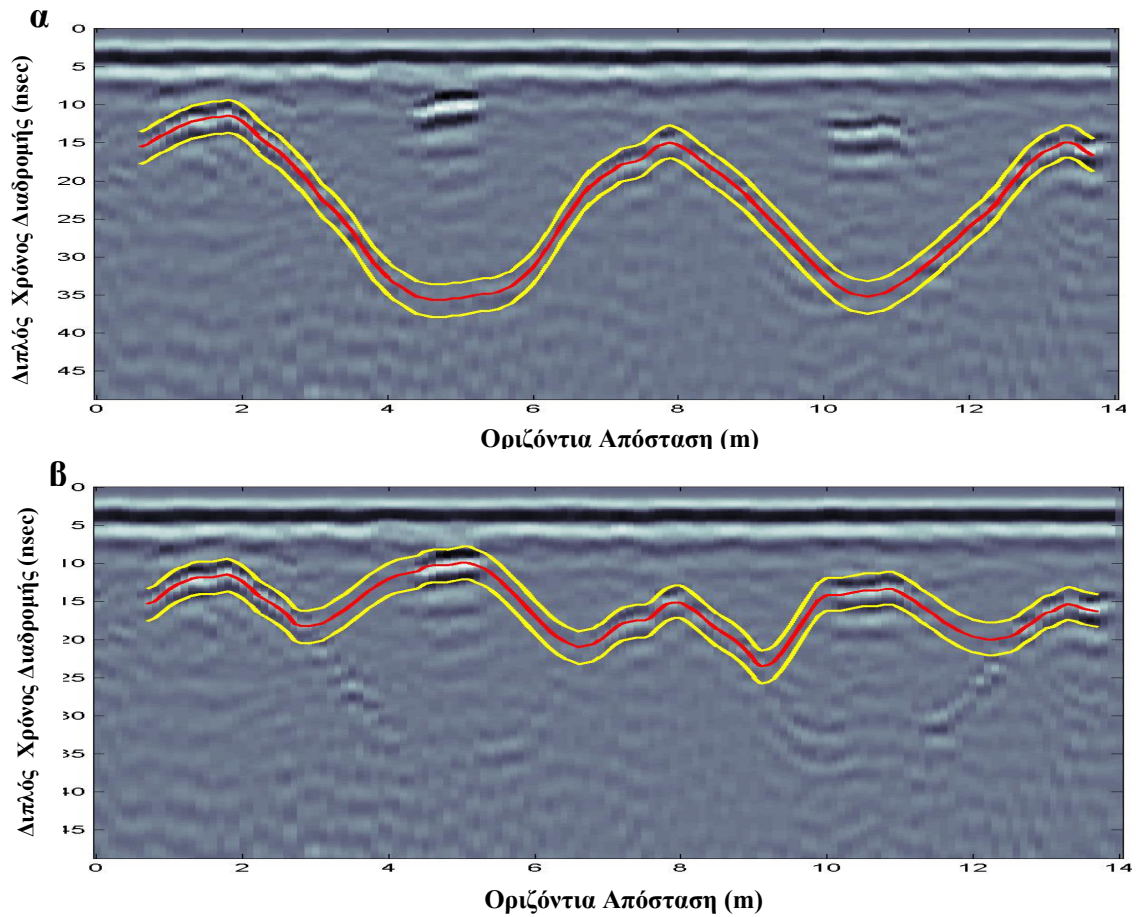
Από τα παραπάνω χαρακτηριστικά επιλέχθηκαν μόνο τα δεδομένα που περιέχονται σε χρονικό παράθυρο διάρκειας 4 nsec γύρω από τον υπό εξέταση ανακλαστήρα. Υπολογίστηκε η μέση τιμή του πλάτους και η μέση τιμή των χαρακτηριστικών του ίχνους και τροφοδοτήθηκαν σε δίκτυα αυτο-οργανωμένης απεικόνισης για ταξινόμηση σε είκοσι (20) τάξεις αντίστοιχα. Για κάθε ανακλαστήρα πραγματοποιήθηκαν τρεις (3) δοκιμές χρησιμοποιώντας διαφορετικούς συνδυασμούς χαρακτηριστικών. Ο πρώτος συνδυασμός περιελάμβανε όλα τα παραπάνω χαρακτηριστικά και τα χωροθετημένα δεδομένα, ο δεύτερος συνδυασμός μόνο τα χαρακτηριστικά, ενώ ο τρίτος όλα τα χαρακτηριστικά πλην της στιγμιαίας φάσης.

Στο σχήμα 6.12 παρουσιάζονται τα αποτελέσματα από τις τρεις δοκιμές για την ταξινόμηση του πρώτου ανακλαστήρα. Τα αποτελέσματα είναι παρόμοια, δηλαδή ο ανακλαστήρας στο πάνω μέρος των θόλων έχει ταξινομηθεί στις ομάδες 15 έως 20, ενώ όσο αυξάνεται το βάθος υπάρχει σταδιακή μείωση στον αύξοντα αριθμό της ομάδας όπου ανήκουν οι καταγραφές.

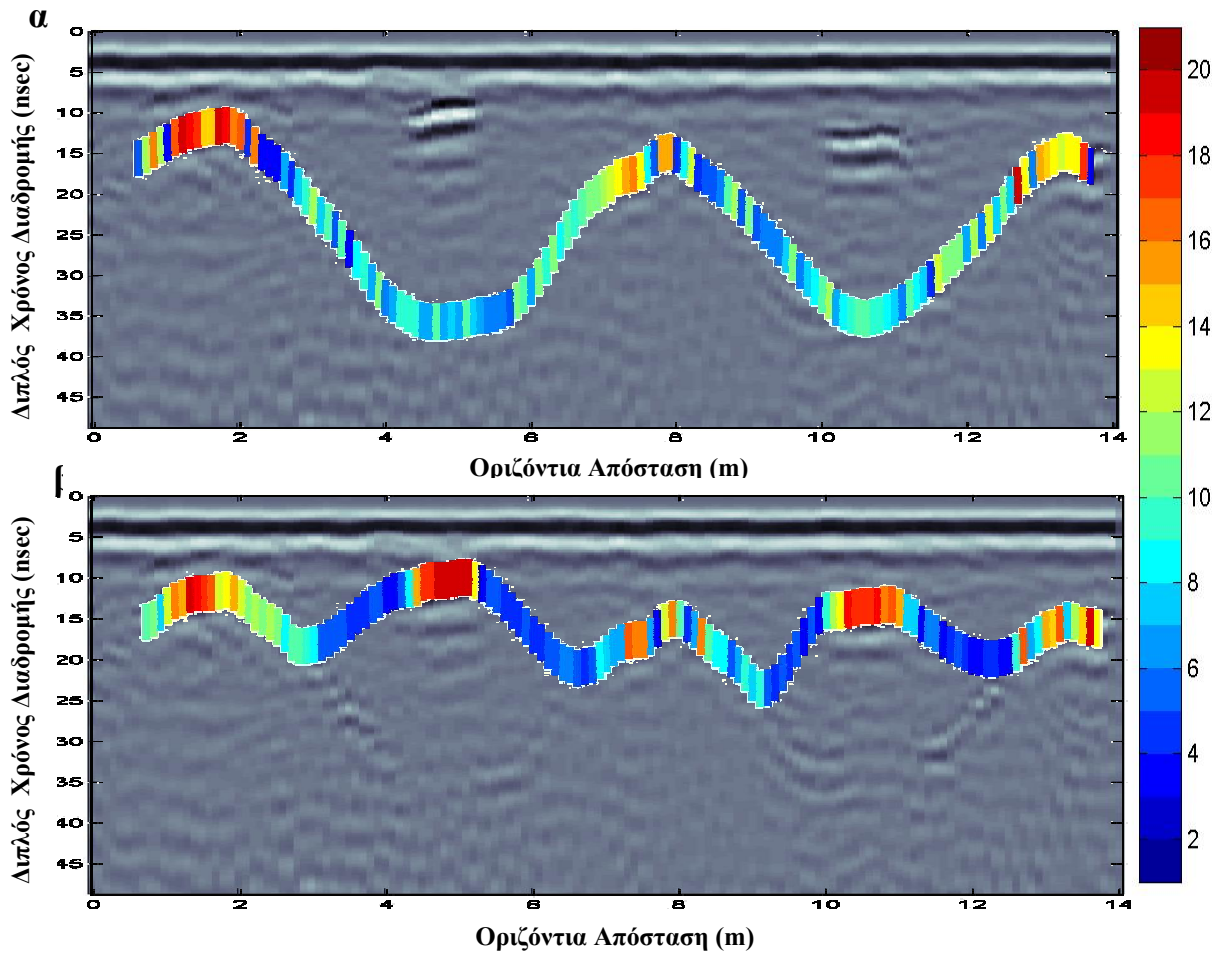
Στο σχήμα 6.13 παρουσιάζονται τα αποτελέσματα από τις δοκιμές για την ταξινόμηση του δεύτερου ανακλαστήρα. Όλος ο ανακλαστήρας ταξινομείται στις ομάδες 13 έως 20, ενώ ανάμεσα από αυτούς, τα δεδομένα ταξινομούνται στις ομάδες 1 έως 12. Εξαιρέση αποτελεί ο κεντρικός θόλος ο οποίος περιλαμβάνει και τις ομάδες 1 και 2.



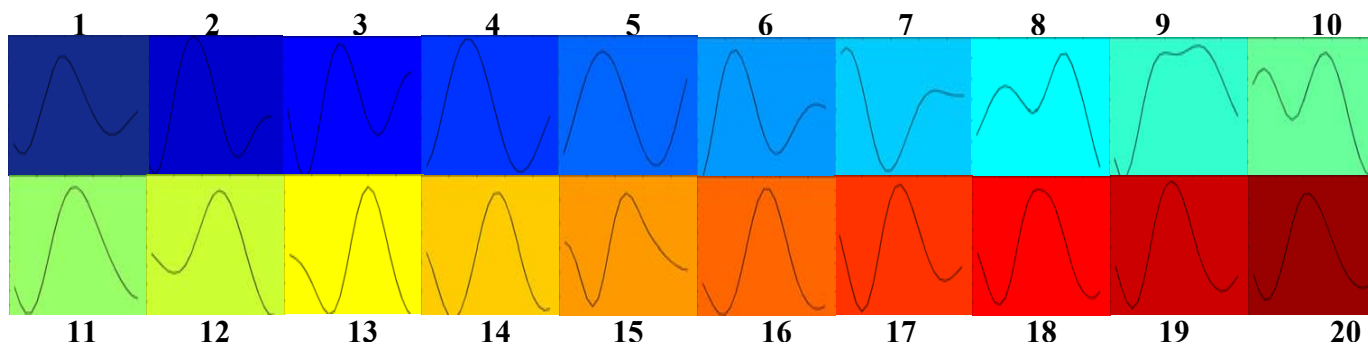
Σχήμα 6.6. Με κόκκινη γραμμή ορίστηκαν οι δυο ανακλαστήρες. Ο πρώτος (α) ακολουθεί το περίγραμμα της οροφής της τρίκλιτης δεξαμενής ενώ ο δεύτερος (β) περιλαμβάνει και τις ενδιάμεσες ανακλάσεις.



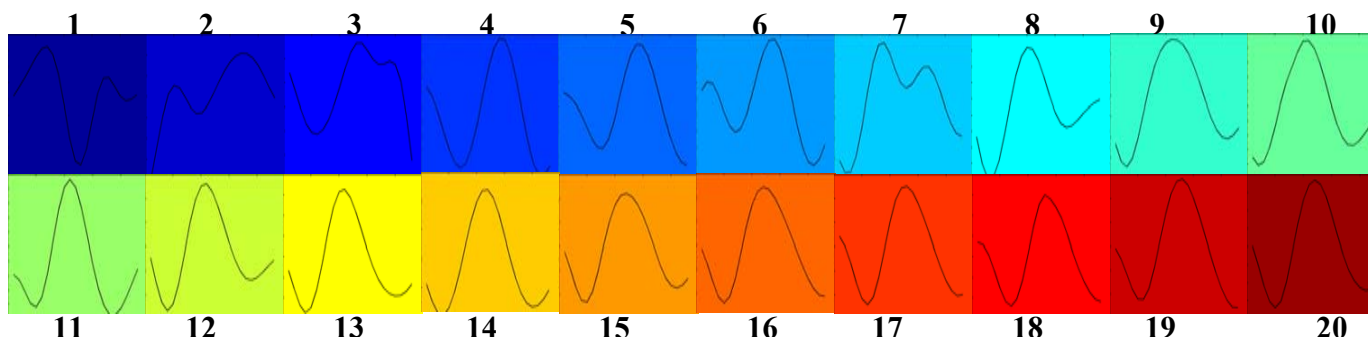
Σχήμα 6.7. Οι κίτρινες γραμμές οριοθετούν το παράθυρο μέσα στο οποίο ορίζονται τα κυματίδια που πρόκειται να ταξινομηθούν για τον πρώτο (α) και τον δεύτερο (β) ανακλαστήρα.



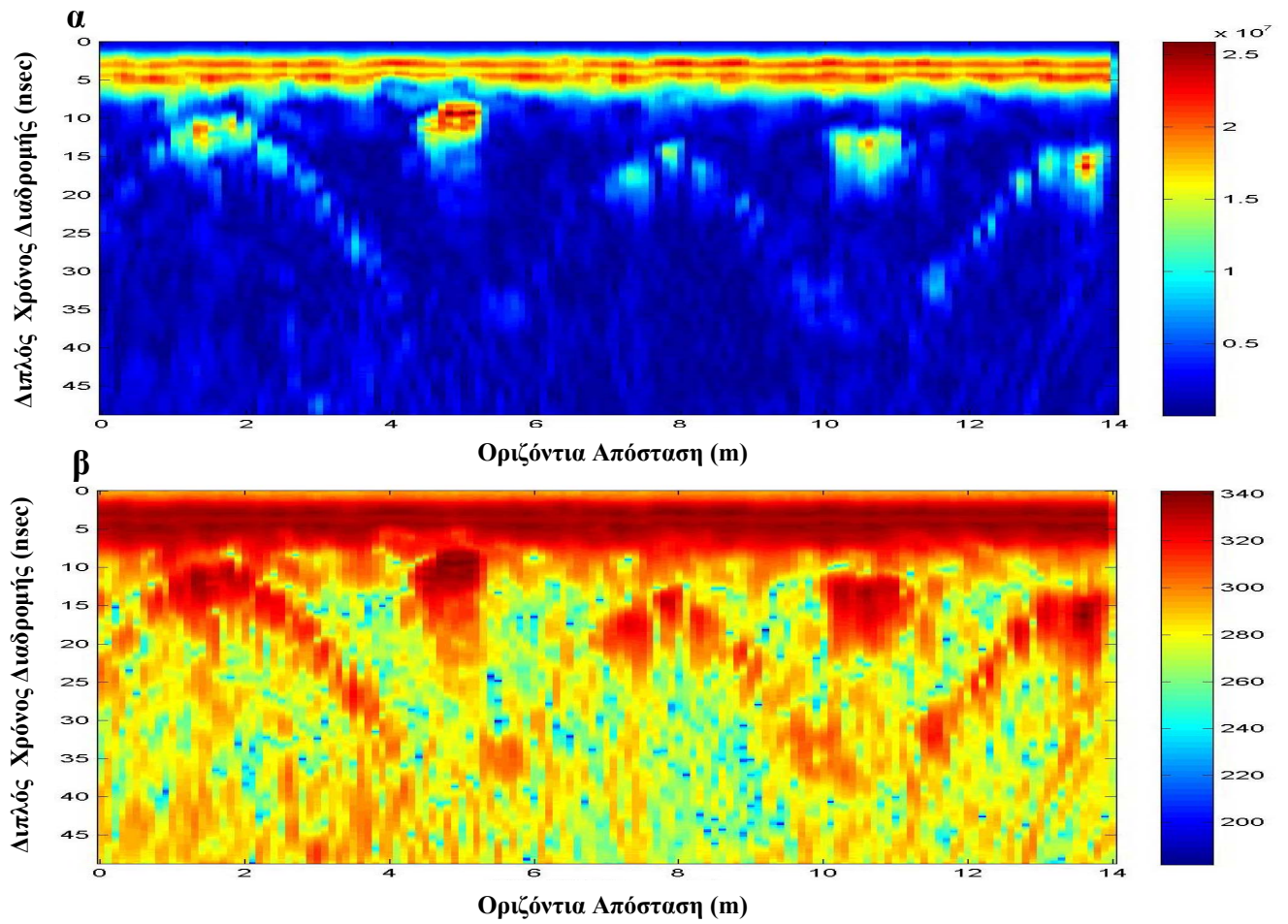
Σχήμα 6.8. Τα αποτελέσματα της ταξινόμησης του πρώτου (α) και του δεύτερου (β) ανακλαστήρα με τη χρήση των κυματιδίων έχουν υπερτεθεί των αρχικών δεδομένων, τα οποία απεικονίζονται σε κλίμακα του γκρι. Η χρωματική κλίμακα προσδιορίζει τον αύξοντα αριθμό των ομάδων.



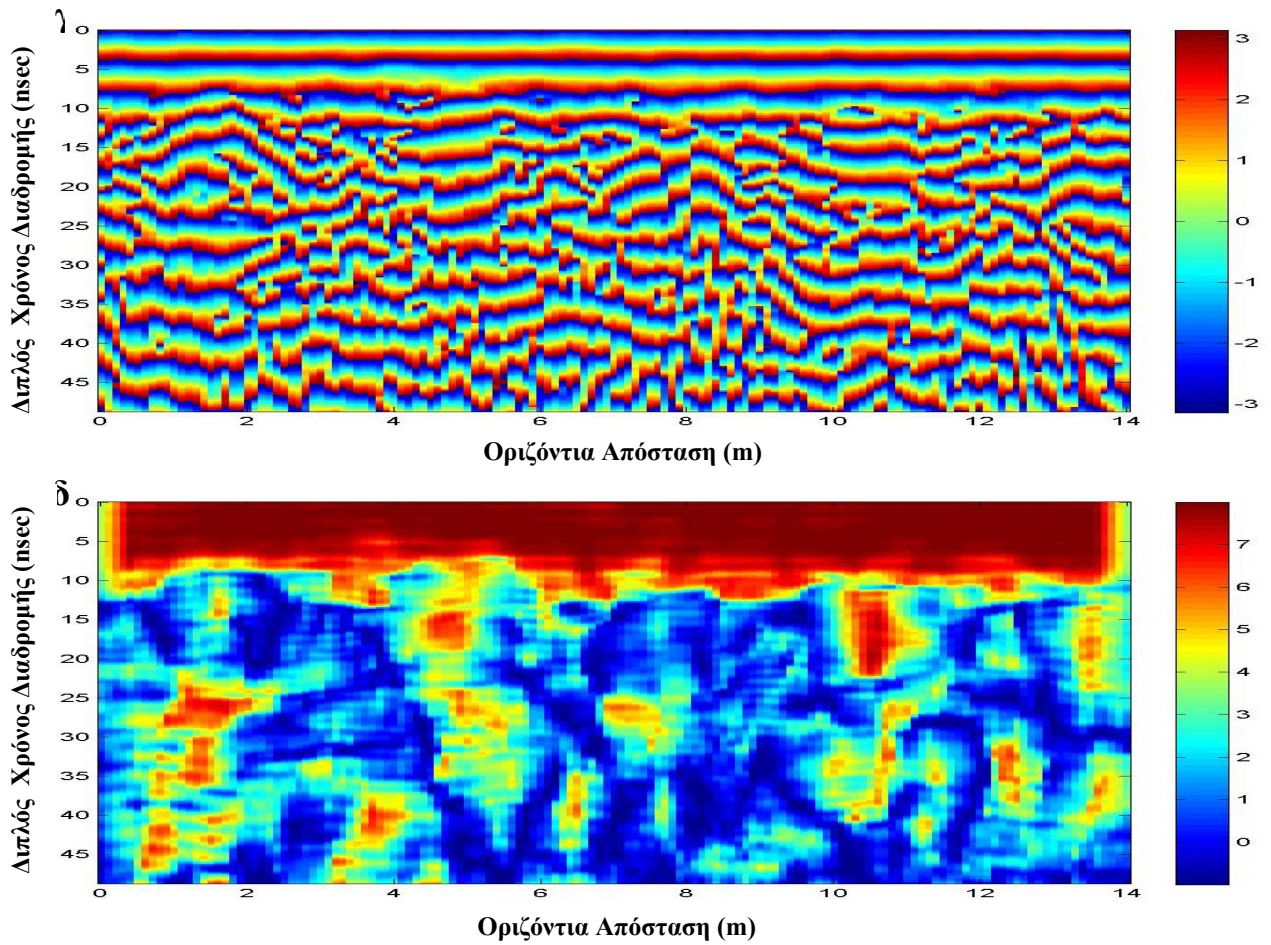
Σχήμα 6.9. Απεικόνιση των κυματιδίων σε χρώματα ανάλογα με την ομάδα που έχουν ταξινομηθεί στο πείραμα με τον πρώτο ανακλαστήρα.



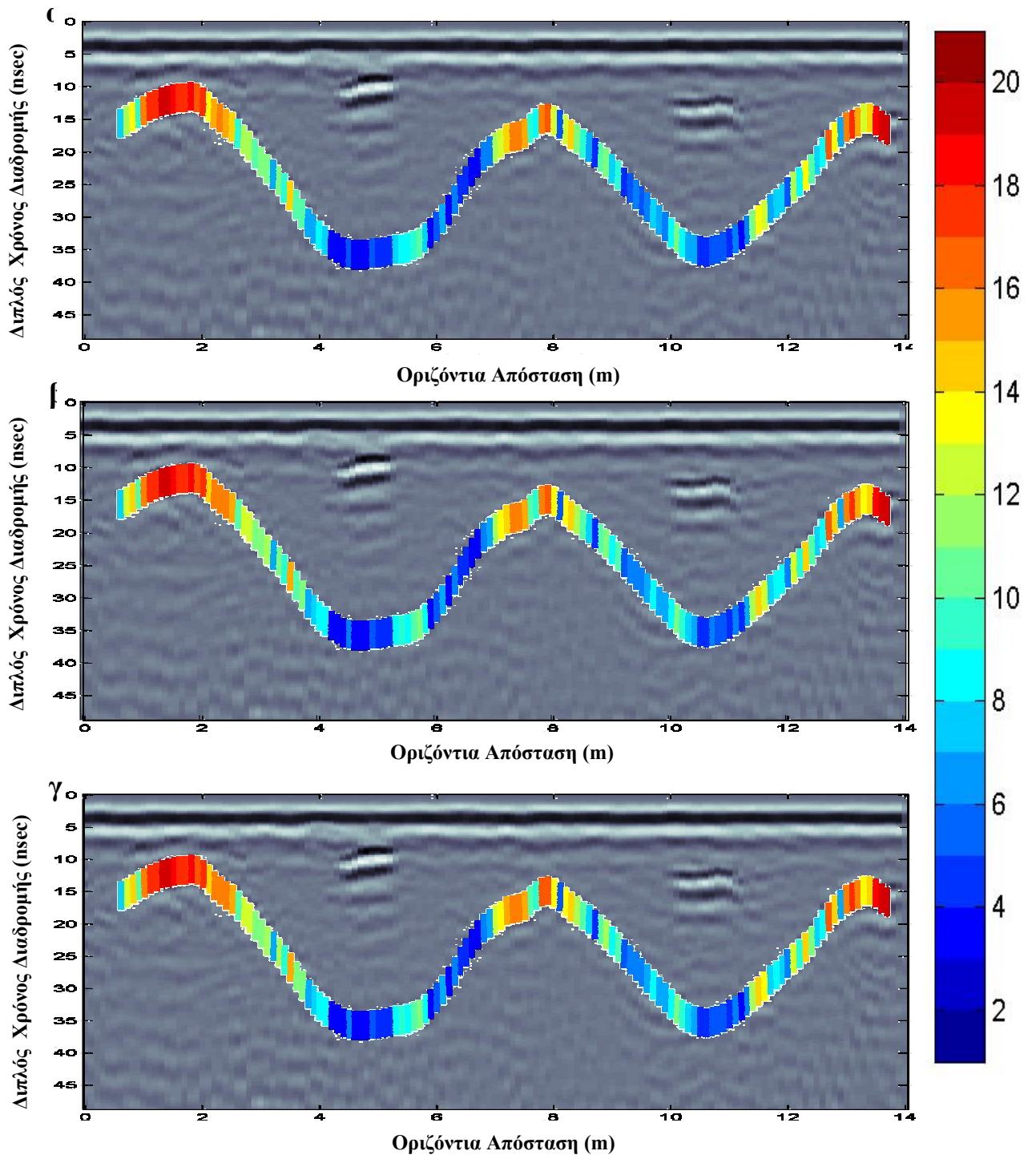
Σχήμα 6.10. Απεικόνιση των κυματιδίων σε χρώματα ανάλογα με την ομάδα που έχουν ταξινομηθεί στο πείραμα με τον δεύτερο ανακλαστήρα.



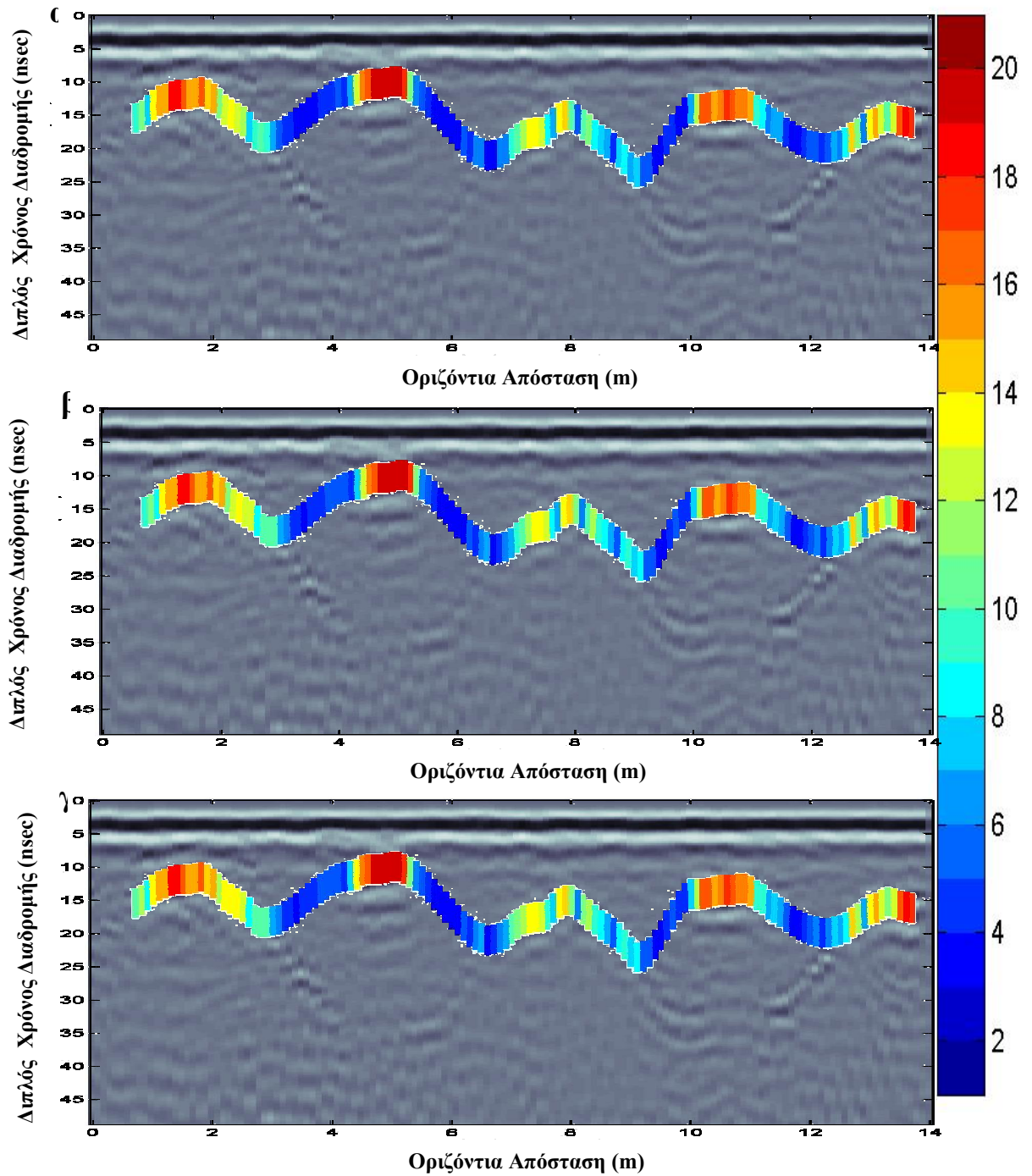
Σχήμα 6.11. Το στιγμιαίο πλάτος (α) και το στιγμιαίο πλάτος σε κλίμακα dB (β) δίνουν μεγάλες τιμές για τις ανακλάσεις στους τρεις θόλους αλλά και για τις ενδιάμεσες ανακλάσεις.



Σχήμα 6.11 (Συνέχεια). Η στιγμιαία φάση (γ) έχει ασυνέχειες οι οποίες υποδηλώνουν είτε την ανομοιογένεια στο υπερκείμενο στρώμα είτε φθορά των θόλων. Η συνάφεια (δ) έχει μεγάλες τιμές για τις ενδιάμεσες ανακλάσεις και μικρές τιμές για τις ανακλάσεις από τους θόλους.



Σχήμα 6.12. Τα αποτελέσματα της ταξινόμησης του πρώτου ανακλαστήρα με τη χρήση των χαρακτηριστικών. Στην πρώτη δοκιμή (α) χρησιμοποιήθηκαν και τα δεδομένα μετά την χωροθέτηση, ενώ στη δεύτερη δοκιμή (β) όχι. Στην τρίτη δοκιμή (γ) χρησιμοποιήθηκαν όλα τα χαρακτηριστικά πλην της στιγμιαίας φάσης. Τα αποτελέσματα όμως είναι παρόμοια.



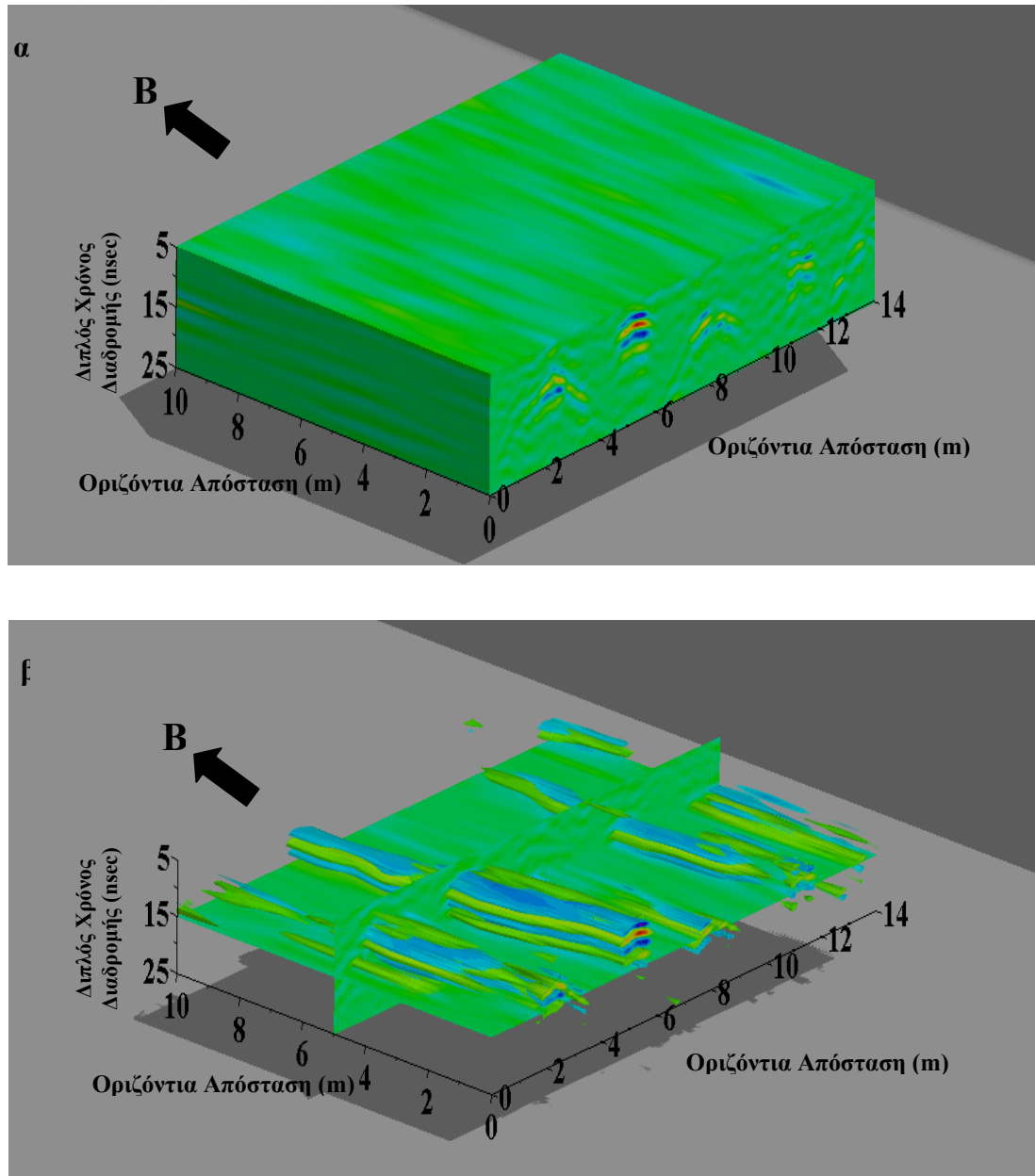
Σχήμα 6.13. Τα αποτελέσματα της ταξινόμησης του δεύτερου ανακλαστήρα με τη χρήση των χαρακτηριστικών. Στην πρώτη δοκιμή (α) χρησιμοποιήθηκαν και τα δεδομένα μετά την χωροθέτηση, ενώ στη δεύτερη δοκιμή (β) όχι. Στην τρίτη δοκιμή (γ) χρησιμοποιήθηκαν όλα τα χαρακτηριστικά πλην της στιγμιαίας φάσης. Τα αποτελέσματα και πάλι είναι παρόμοια.

6.6 Αυτόματη ταξινόμηση των ψευδο-τρισδιάστατων δεδομένων.

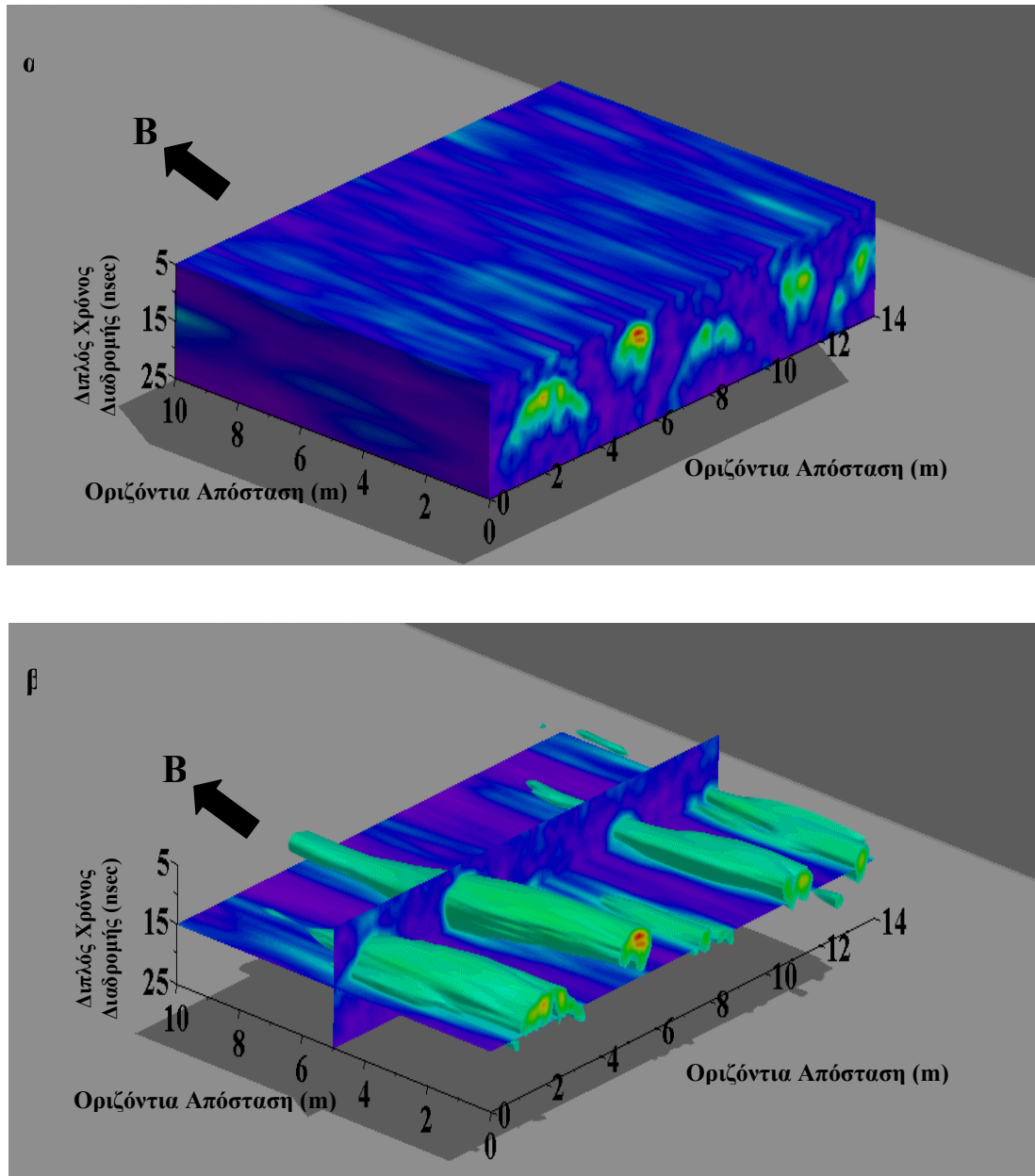
Και οι έντεκα (11) γραμμές μελέτης επεξεργάστηκαν με τον ίδιο τρόπο: αφαιρέθηκε η επαγωγική συνιστώσα, χωροθετήθηκαν και ενισχύθηκαν εφαρμόζοντας το φίλτρο αντιστάθμισης της γεωμετρικής διασποράς. Από τα ψευδο-τρισδιάστατα δεδομένα (από 5 nsec έως 55 nsec) υπολογίστηκαν τα χαρακτηριστικά των ιχνών. Για την καλύτερη απεικόνιση τόσο των χαρακτηριστικών, όσο και των ταξινομημένων δεδομένων, παρουσιάζονται εικόνες σε τρεις διαστάσεις όπως και τομές (οριζόντιες και κατακόρυφες). Οι τομές αφορούν τη διεύθυνση δύση – ανατολή στα 5.5 m καθώς και στα 15 nsec και έχουν στόχο να αναδείξουν τις διαστάσεις των ανακλαστήρων.

Στο σχήμα 6.14 φαίνονται τα δεδομένα μετά την πρώτη φάση επεξεργασίας. Απεικονίζονται μονάχα οι μικρές αρνητικές και μεγάλες θετικές τιμές του πλάτους ενώ οι υπόλοιπες τιμές είναι διάφανες. Με αυτό τον τρόπο διακρίνεται η τρισδιάστατη έκταση του κάθε ανακλαστήρα στην περιοχή μελέτης. Το στιγμιαίο πλάτος (σχήμα 6.15) και το στιγμιαίο πλάτος σε κλίμακα dB (σχήμα 6.16) αναδεικνύουν τον ανακλαστήρα ακόμα καλύτερα. Στις παραπάνω εικόνες απεικονίζονται μόνο οι πολύ μεγάλες τιμές του πλάτους. Στο βορειότερο τμήμα οι ανακλαστήρες είναι εξασθενημένοι. Η στιγμιαία φάση (σχήμα 6.17) δε δίνει τόσο προφανείς πληροφορίες όσο το στιγμιαίο πλάτος, αφού το εύρος τιμών της φάσης είναι μικρό και δεν είναι δυνατό να αναδειχτούν οι σημαντικοί ανακλαστήρες. Η συνάφεια, όπου διακρίνονται μόνο οι μεγάλες τιμές, δεν δίνει τόσο προφανή αποτελέσματα όσο το πλάτος (σχήμα 6.18). Στο επόμενο στάδιο τα χαρακτηριστικά του ίχνους τροφοδοτήθηκαν σε δίκτυο αυτο-οργανωμένης απεικόνισης και ταξινομήθηκαν σε είκοσι (20) τάξεις.

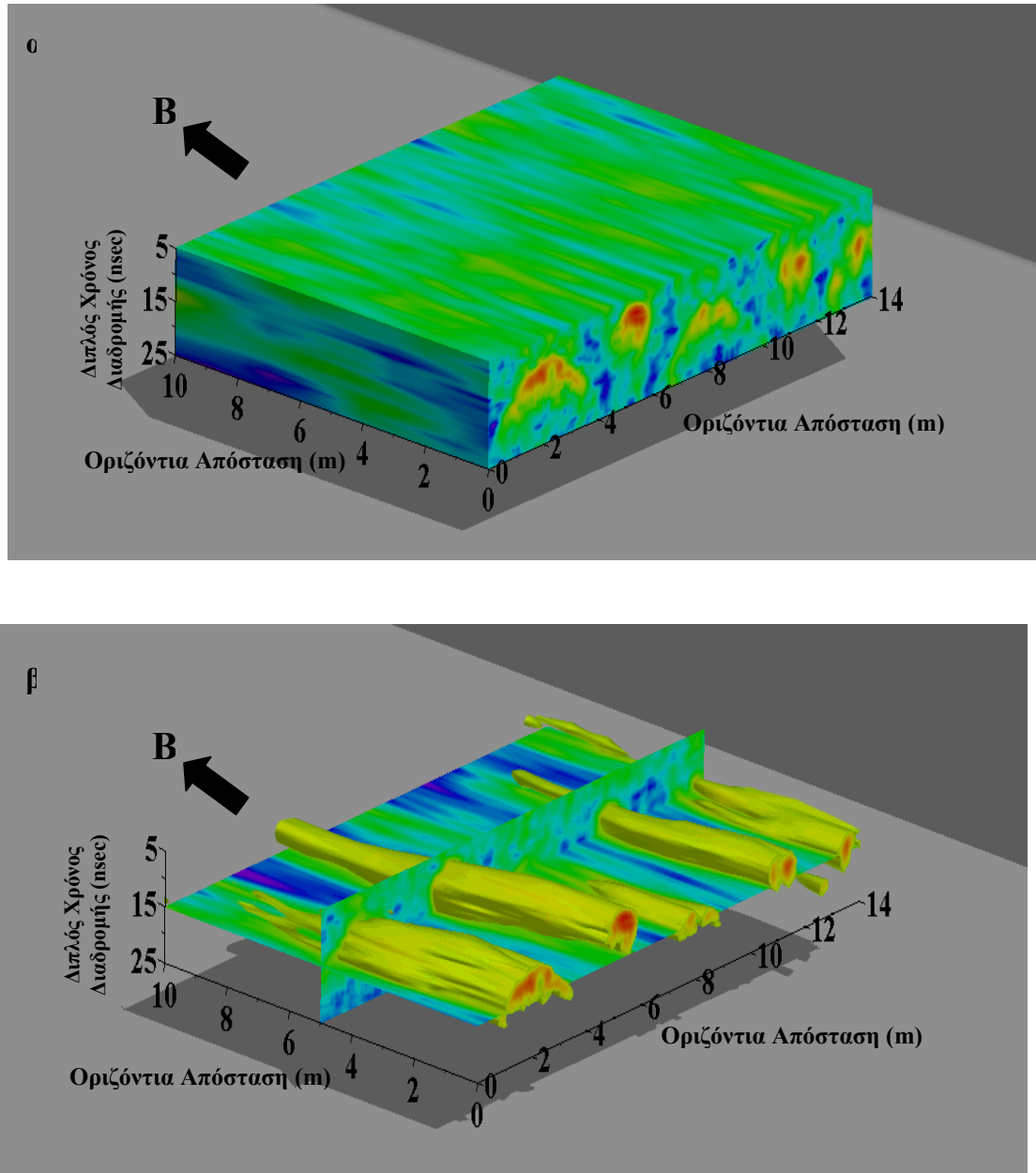
Η πρώτη δοκιμή περιλάμβανε τα τέσσερα χαρακτηριστικά και τα δεδομένα μετά το πρώτο στάδιο επεξεργασίας (σχήμα 6.19). Ο ανακλαστήρας από τους τρεις θόλους και οι ανακλάσεις από τους αγωγούς ταξινομήθηκαν στις ομάδες 19-20 (σκούρο πορτοκαλί και κόκκινο), ενώ το υπόβαθρο στις τάξεις 1-18. Όλοι οι ανακλαστήρες λεπταίνουν στα βόρεια και επίσης διακρίνονται λεπτομέρειες λόγω της χρήσης της στιγμιαίας φάσης στην ταξινόμηση.



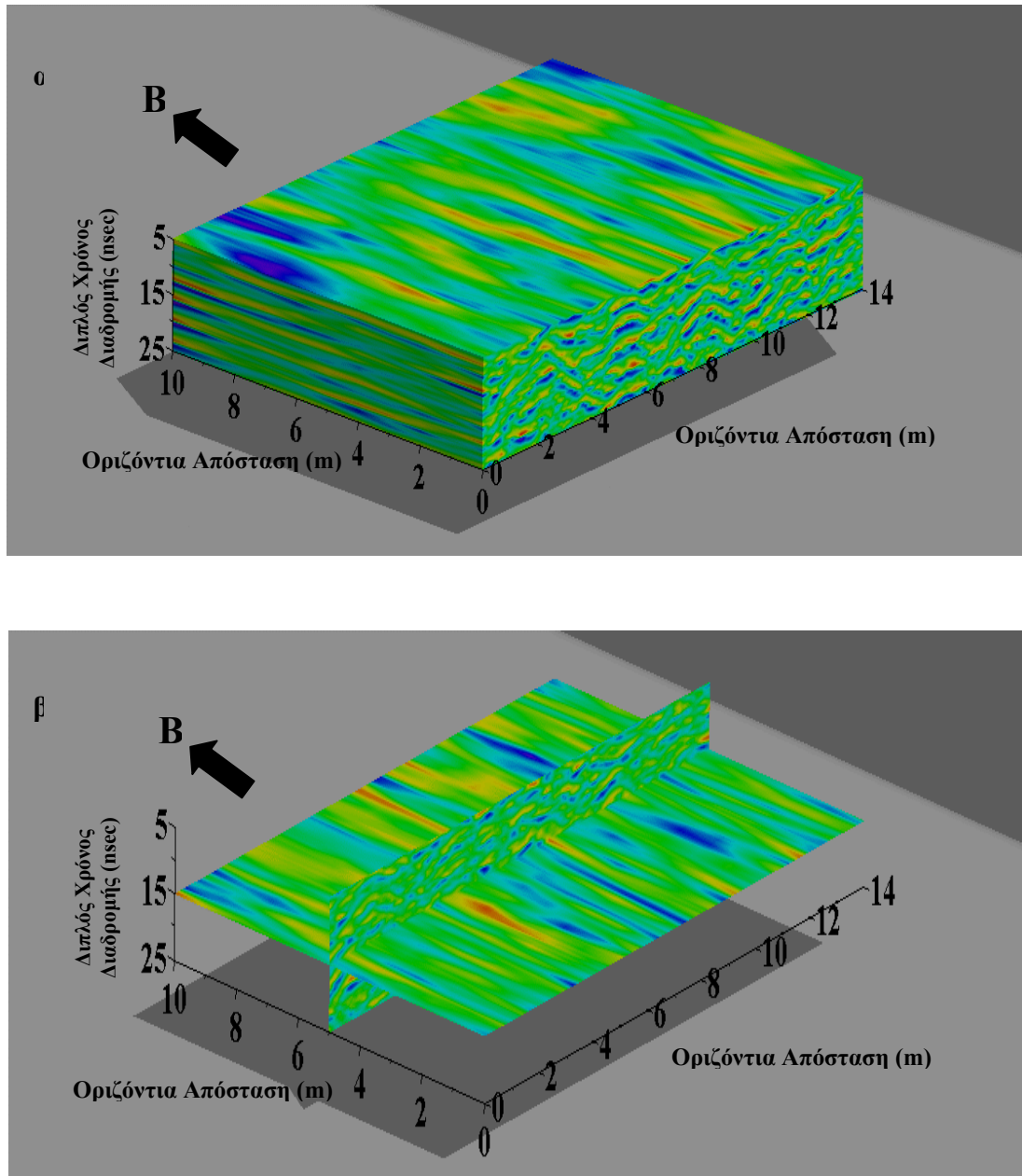
Σχήμα 6.14. (α) Ο ψευδο-τριδιάστατος όγκος των δεδομένων. (β) Διακρίνονται δυο τομές, μια οριζόντια τομή στα 15 nsec και μια γραμμή μελέτης στα 5.5 m, καθώς και οι μεγάλες θετικές τιμές του πλάτους (κίτρινο και κόκκινο χρώμα) και οι μικρές αρνητικές του πλάτους (μπλε). Οι ενδιάμεσες τιμές είναι διαφανείς.



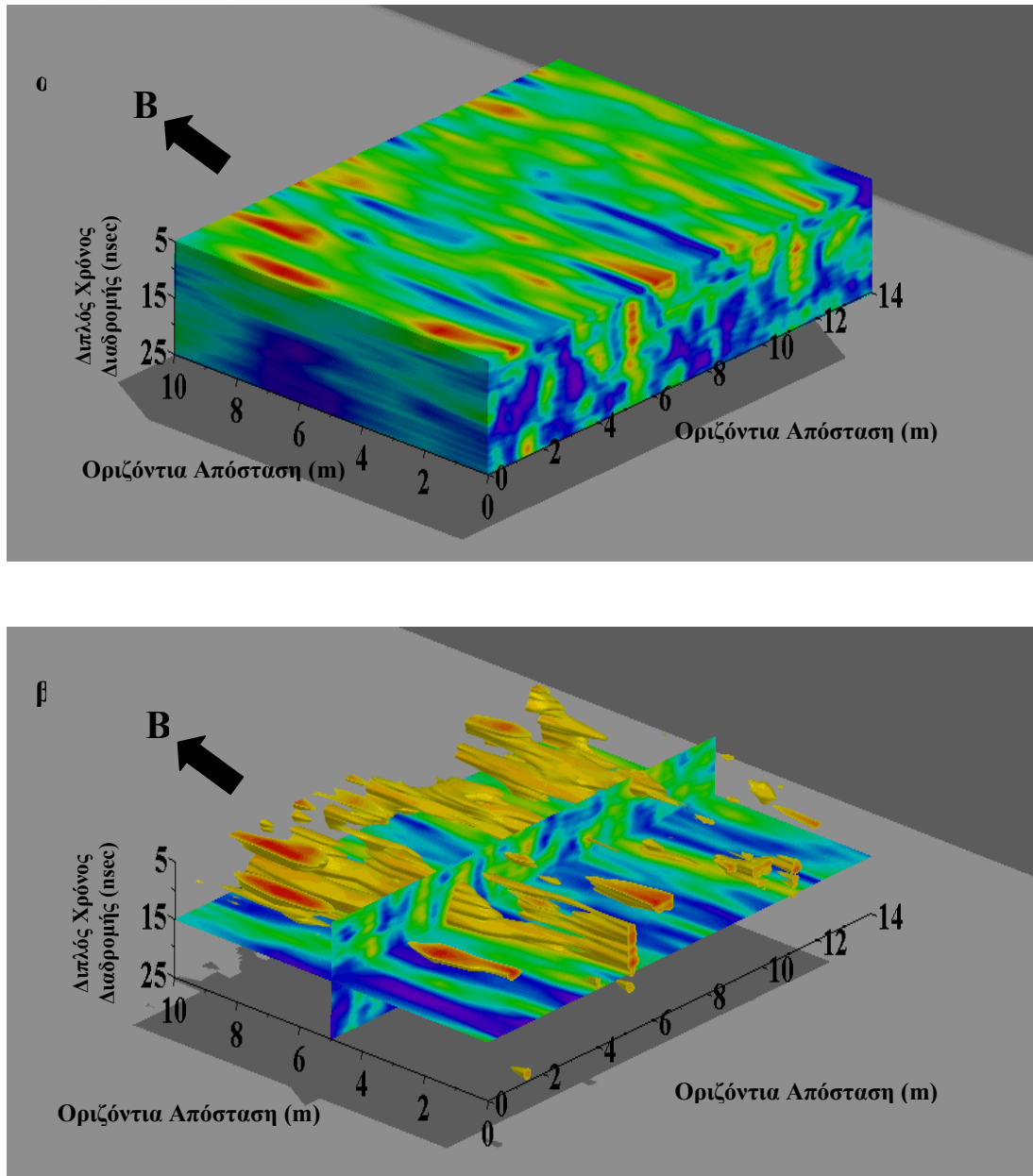
Σχήμα 6.15. (α) Το στιγμιαίο πλάτος του ψευδό-τρισεδιάστατου όγκου των δεδομένων. **(β)** Διακρίνονται δυο τομές, μια οριζόντια τομή στα 15 nsec και μια κάθετη τομή στα 5.5 m, καθώς και οι μεγάλες του πλάτους (κόκκινο έως γαλάζιο). Διακρίνονται πέντε (5) ανακλαστήρες με ενδιαφέρον.



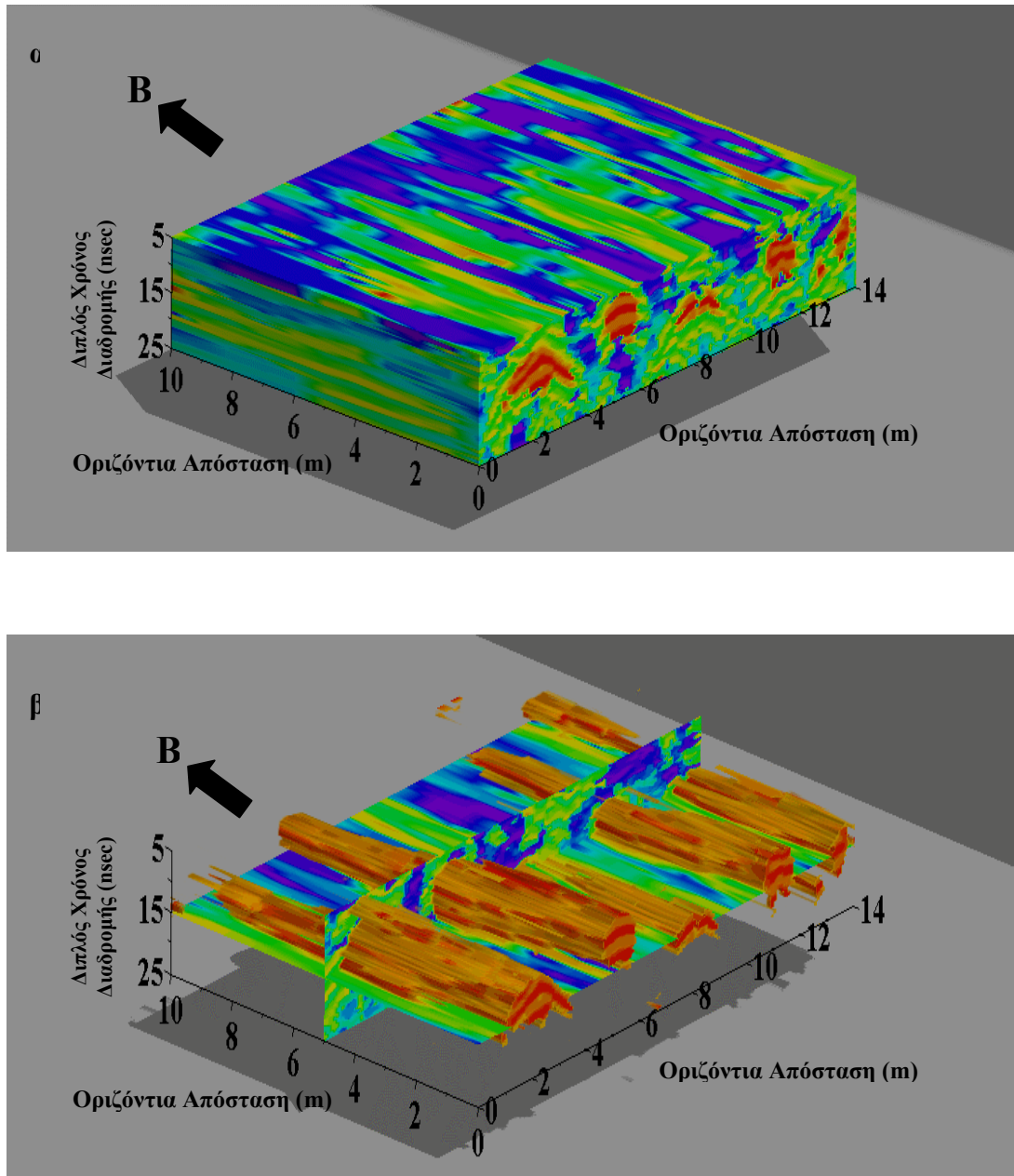
Σχήμα 6.16. (α) Το στιγμιαίο πλάτος σε κλίμακα dB του ψευδο-τρισδιάστατου όγκου των δεδομένων. **(β)** Διακρίνονται δυο τομές, μια οριζόντια τομή στα 15 nsec και μια κάθετη τομή στα 5.5 m, καθώς και οι μεγάλες του πλάτους (κόκκινο έως κίτρινο).



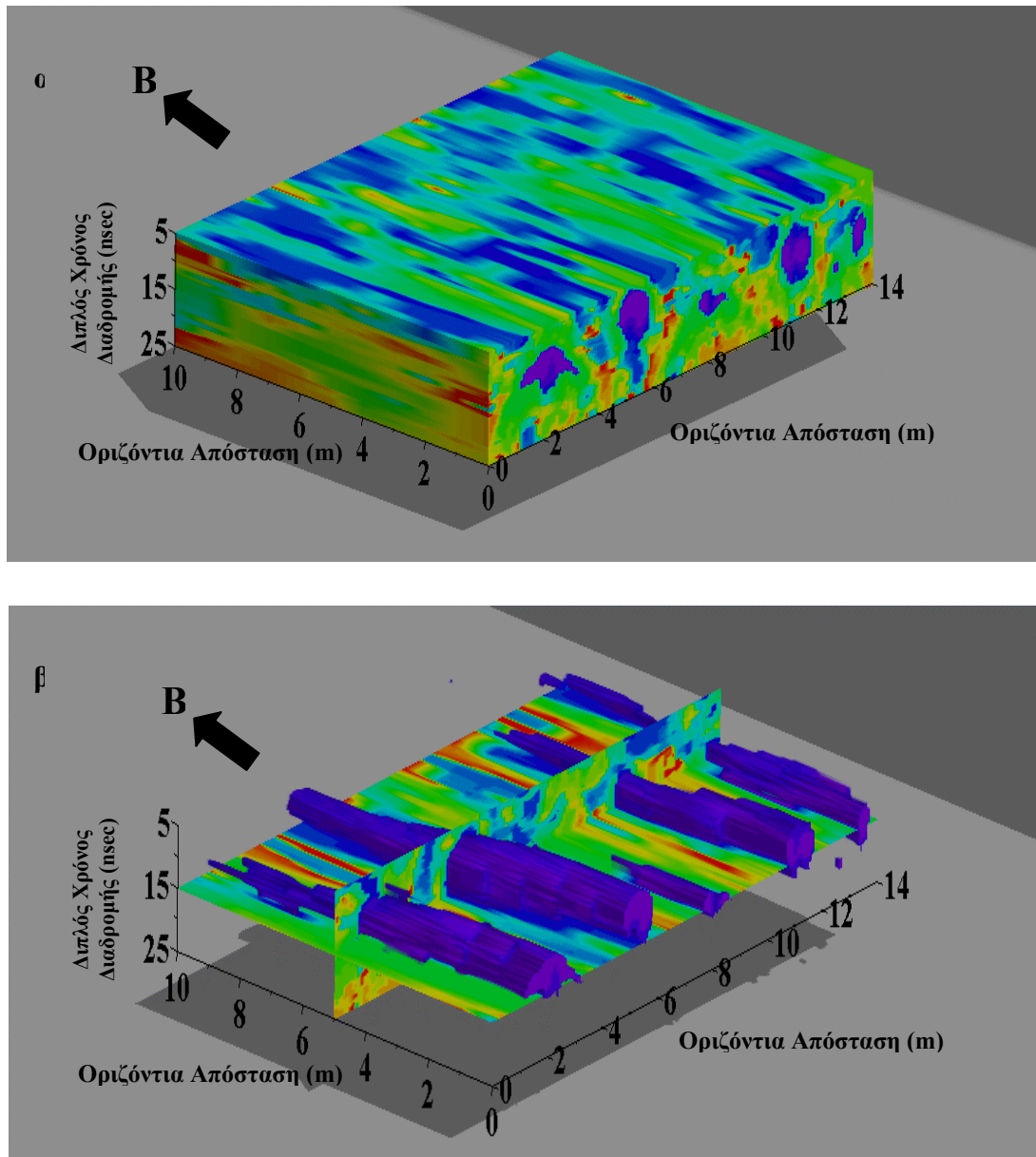
Σχήμα 6.17. (α) Η στιγμιαία φάση του ψευδο-τρισδιάστατου όγκου των δεδομένων. **(β)** Διακρίνονται δυο τομές, μια οριζόντια τομή στα 15 nsec και μια κάθετη τομή στα 5.5 m. Λόγω του μικρού εύρους τιμών της στιγμιαίας φάσης δεν είναι δυνατό να αποκαλυφθούν με σαφήνεια οι ανακλαστήρες.



Σχήμα 6.18. (α) Η συνάφεια του ψευδο-τριδιάστατου όγκου των δεδομένων. **(β)** Διακρίνονται δυο τομές, μια οριζόντια τομή στα 15 nsec και μια κάθετη τομή στα 5.5 m. Η συνάφεια των ενδιάμεσων ανακλαστήρων είναι μεγαλύτερη και παρουσιάζεται με κόκκινο και πορτοκαλί χρώμα, ενώ η συνάφεια των θόλων έχει μεγαλύτερο εύρος και δεν είναι δυνατόν να παρουσιαστεί με σαφήνεια.



Σχήμα 6.19. (α) Το αποτέλεσμα της ταξινόμησης των δεδομένων του ψευδο-τριδιάστατου όγκου των δεδομένων με τη χρήση όλων των χαρακτηριστικών και των δεδομένων μετά τη χωροθέτηση και την ενίσχυση. (β) Διακρίνονται δυο τομές, μια οριζόντια τομή στα 15 nsec και μια κάθετη τομή στα 5.5 m. Οι ανακλαστήρες έχουν ταξινομηθεί στις ομάδες 19-20 και διακρίνονται λεπτομέρειες λόγω της φάσης των δεδομένων. Το υπόβαθρο έχει ταξινομηθεί στις ομάδες 1-18 και είναι διάφανο στην εικόνα.



Σχήμα 6.20. (α) Το αποτέλεσμα της ταξινόμησης των δεδομένων του ψευδο-τρισδιάστατου όγκου των δεδομένων με τη χρήση του στιγμιαίου πλάτους, του στιγμιαίου πλάτους σε κλίμακα dB και της συνάφειας. (β) Διακρίνονται δυο τομές, μια οριζόντια τομή στα 15 nsec και μια κάθετη τομή στα 5.5 m. Οι ανακλαστήρες έχουν ταξινομηθεί στην ομάδα 20. Το υπόβαθρο έχει ταξινομηθεί στις ομάδες 1-19 και είναι διάφανο στην εικόνα.

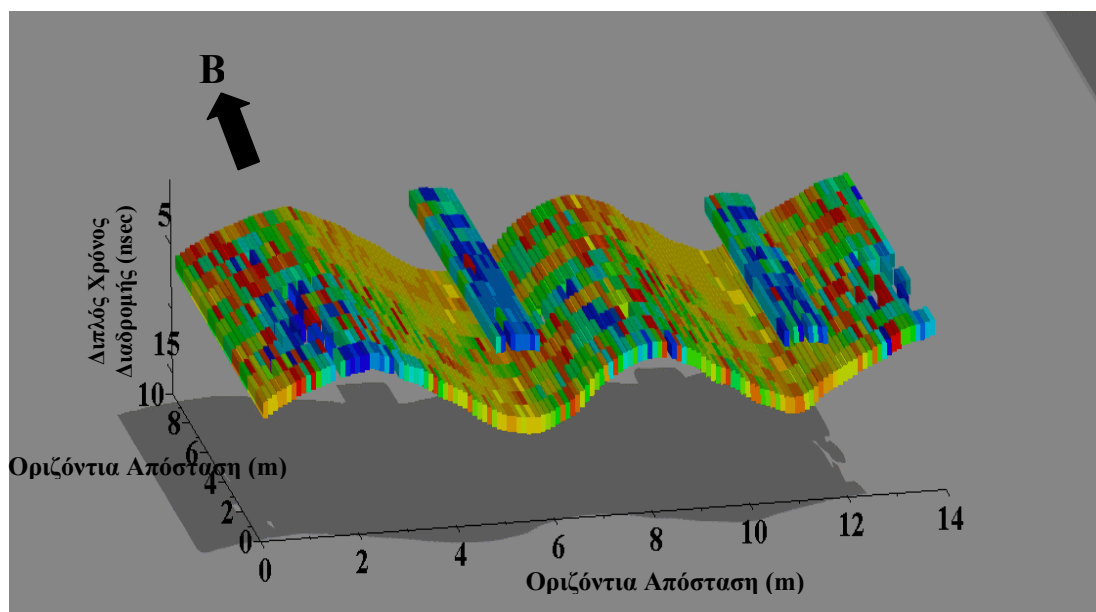
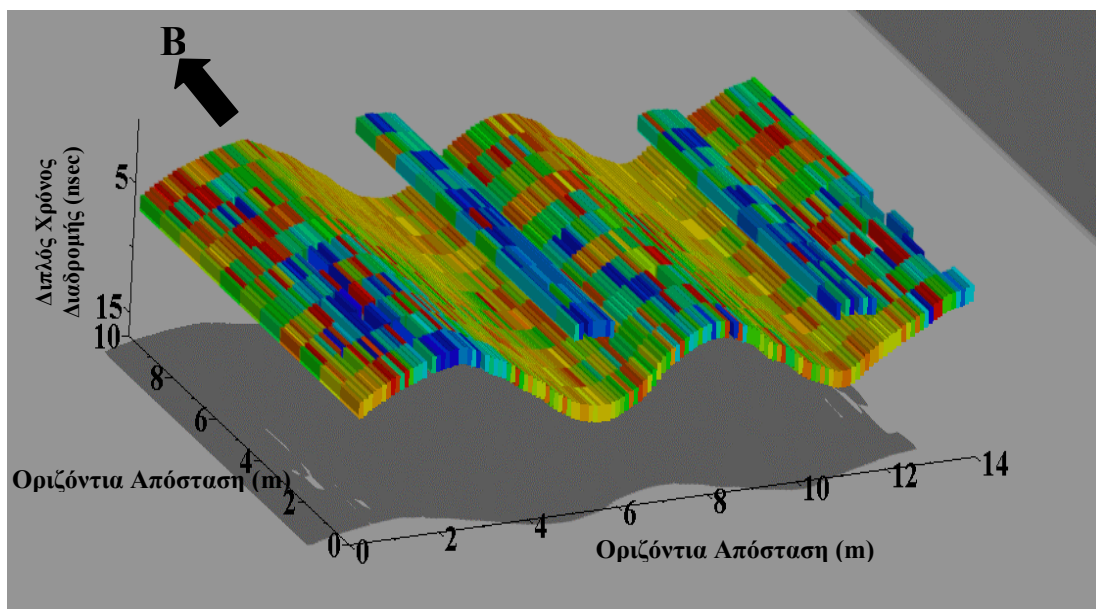
Η δεύτερη δοκιμή περιλάμβανε το στιγμιαίο πλάτος, στιγμιαίο πλάτος σε κλίμακα dB και τη συνάφεια. Οι ανακλαστήρες με ενδιαφέρον φαίνεται να ανήκουν στις τάξεις 19-20, ενώ το υπόβαθρο στις τάξεις 1-18 (σχήμα 6.20). Σε αυτή τη δοκιμή τονίζεται ακόμα καλύτερα το γεωμετρικό σχήμα των ανακλαστήρων και αναδεικνύονται οι διαστάσεις τους με περισσότερη σαφήνεια.

6.7 Αυτόματη ταξινόμηση των ανακλαστήρων.

Αυτή η δοκιμή περιλάμβανε τα παρακάτω στάδια επεξεργασίας τα οποία εφαρμόστηκαν και στις έντεκα (11) γραμμές μελέτης. Αρχικά, εφαρμόστηκε το φίλτρο απομάκρυνσης της επαγωγικής συνιστώσας. Έπειτα, χωροθετήθηκαν και ενισχύθηκαν με το φίλτρο αντιστάθμισης της γεωμετρικής διασποράς. Στη δεύτερη φάση έγινε χειροκίνητη επιλογή όλων των ανακλαστήρων με ενδιαφέρον, ορίστηκε παράθυρο διάρκειας 4 nsec και τα κυματίδια που περικλείονταν τροφοδοτήθηκαν σε δίκτυο αυτο-οργανωμένης απεικόνισης και ταξινομήθηκαν σε είκοσι (20) τάξεις. Στο σχήμα 6.21 φαίνεται το αποτέλεσμα της ταξινόμησης του ανακλαστήρα. Οι υδαταγωγοί ταξινομούνται στις ομάδες 1-5. Το κοίλο μέρος της οροφής ταξινομείται στις ομάδες 13-20, ενώ το κυρτό μέρος στις ομάδες 1-12.

6.8 Συμπεράσματα

Στο κεφάλαιο αυτό χρησιμοποιήθηκαν μέθοδοι αυτόματης ταξινόμησης για την ερμηνεία δεδομένων γεωραντάρ από διασκόπηση σε αρχαιολογικό χώρο. Ταξινομήθηκαν, η πλήρης κυματομορφή των ανακλώμενων κυμάτων από την οροφή της τρίκλιτης δεξαμενής, τα χαρακτηριστικά της και τα ψευδο-τρισδιάστατα δεδομένα. Φάνηκε ότι οι μέθοδοι ταξινόμησης συμβάλλουν στην ερμηνεία των αρχικών δεδομένων του γεωραντάρ.



Σχήμα 6.21. Ταξινόμηση των σημαντικών ανακλαστήρων σε 20 τάξεις. Οι αγωγοί έχουν ταξινομηθεί στις ομάδες 1-5, ενώ οι θόλοι στις ομάδες 6-20 με λίγες εξαιρέσεις καταγραφών που φαίνονται να ανήκουν στις ομάδες 1-5.

ΚΕΦΑΛΑΙΟ 7**ΤΑΞΙΝΟΜΗΣΗ ΔΕΔΟΜΕΝΩΝ ΓΕΩΡΑΝΤΑΡ ΑΠΟ ΤΟ ΟΡΟΠΕΔΙΟ ΤΟΥ
ΟΜΑΛΟΥ****7.1 Εισαγωγή**

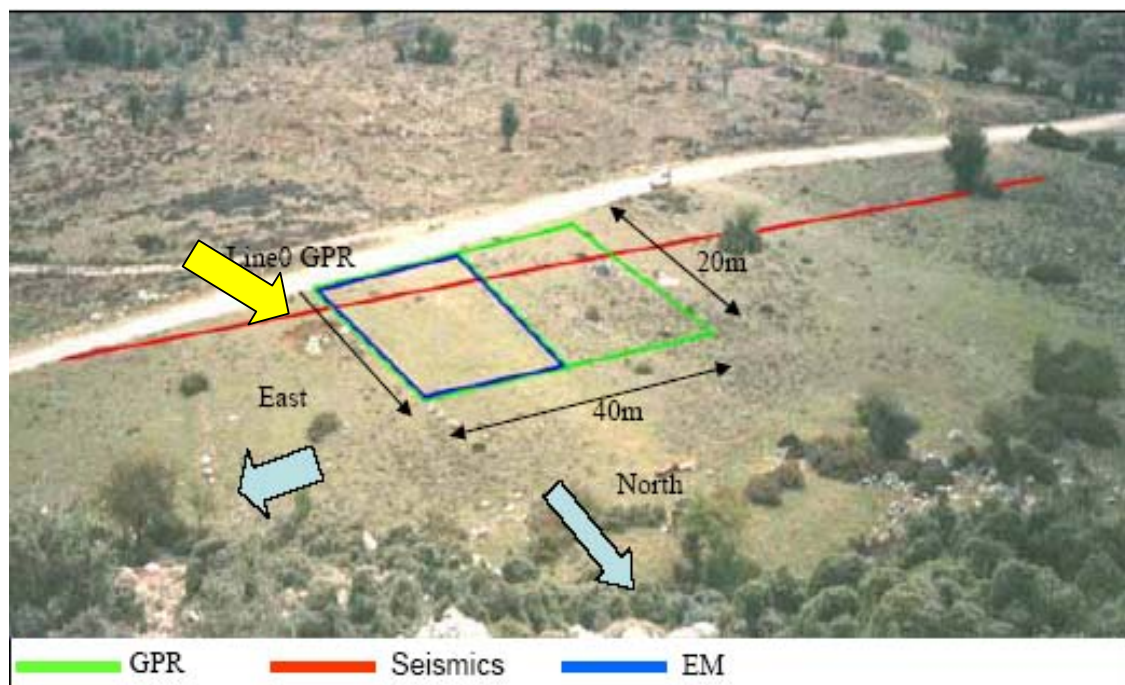
Στο κεφάλαιο αυτό παρουσιάζεται η ταξινόμηση ψευδοτριδιάστατων δεδομένων γεωραντάρ που συλλέχθηκαν στο οροπέδιο του Ομαλού, στο νομό Χανίων, με στόχο τον εντοπισμό και τη χαρτογράφηση καρστικών εγκοίλων. Οι μετρήσεις πραγματοποιήθηκαν στα πλαίσια του μαθήματος ‘Μεγάλες Ασκήσεις Υπαίθρου IV’ του τμήματος Μηχανικών Ορυκτών Πόρων του Πολυτεχνείου Κρήτης τον Σεπτέμβριο του 2003.

Τα δεδομένα του γεωραντάρ που παρουσιάζονται, συλλέχθηκαν σε περιοχή όπου υπάρχουν επιφανειακές εμφανίσεις δολινών (σχήμα 7.1). Τα αποτελέσματα συγκρίνονται με δεδομένα που προέκυψαν από την ηλεκτρομαγνητική χαρτογράφηση με το όργανο

EM- 32 καθώς και με σεισμικά δεδομένα που συλλέχθηκαν τη μέθοδο της σεισμικής διάθλασης.

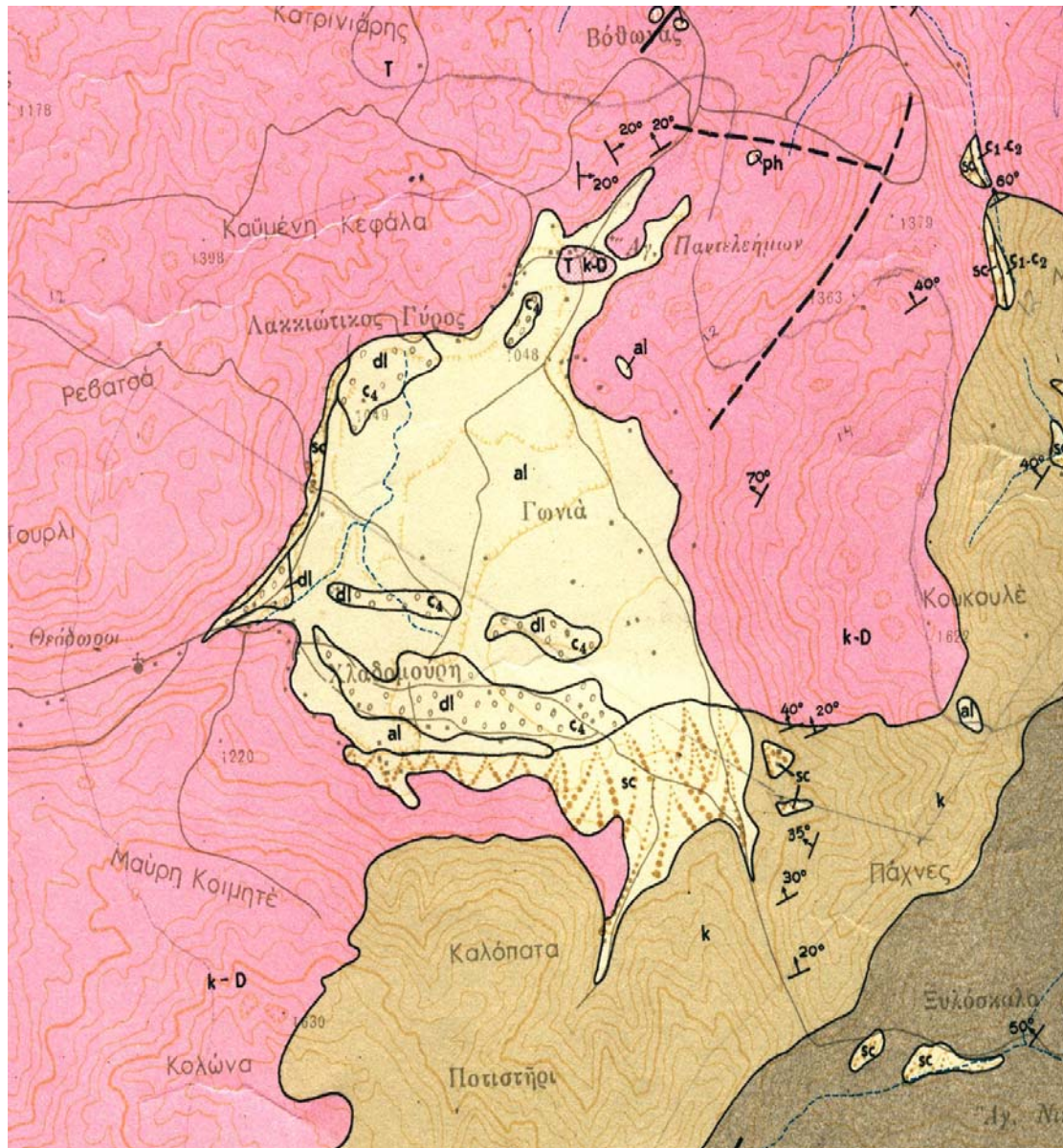
7.2 Περιοχή Ομαλού

Γεωγραφικά το οροπέδιο του Ομαλού βρίσκεται στα νότια του νομού Χανίων. Ο δρόμος για τον Ομαλό διέρχεται μέσα από πετρώματα του καλύμματος των φυλλιτών-χαλαζιτών και μετά το χωριό Λάκκοι από τα πετρώματα του Τρυπαλίου. Μεγάλες εμφανίσεις των φυλλιτικών πετρωμάτων υπάρχουν κατά μήκος όλης της διαδρομής. Γενικά πάντως, η γεωλογική τομή των Λευκών Ορέων στην περιοχή του Ομαλού είναι: πετρώματα φυλλιτών-χαλαζιτών, καλύμματος του Τρυπαλίου, τυπικοί πλακώδεις ασβεστόλιθοι του ομώνυμου καλύμματος και δολομίτης. Στο σχήμα 7. φαίνεται ο γεωλογικός χάρτης της περιοχής ενώ στον πίνακα. 7. φαίνεται το υπόμνημα του χάρτη.



Σχήμα 7.1: Η περιοχή μελέτης όπου πραγματοποιήθηκαν οι μετρήσεις. Διακρίνεται ο κάναβος του γεωραντάρ (πράσινη γραμμή), ο κάναβος της H/M χαρτογράφησης (μπλε

γραμμή) καθώς και η γραμμή όπου πραγματοποιήθηκαν μετρήσεις με τη μέθοδο της σεισμικής διάθλασης (κόκκινη γραμμή). Το κίτρινο βέλος δείχνει την θέση της επιφανειακής δολίνης.



Σχήμα 7.2. Γεωλογικός χάρτης του Ομαλού. Το υπόμνημα του πίνακα φαίνεται στον πίνακα 7.1

Πίνακας 7.1. Υπόμνημα του γεωλογικού χάρτη του σχήματος 7.2

1. Επιφανειακές σύγχρονες αποθέσεις	al, dl	Παράκτιοι άμμοι (al), προσχώσεις (al), κροκαλοπαγή (pl), ασβεστιτικοί ψαμμίτες (pl), κίτρινες ψαμμιτικές μάργες (pl), μαργαϊκοί ασβεστόλιθοι (pl), λεπτοστρωματώδεις χαλαζίτες με πιο λεπτές στρώσεις φυλλιτών και αργιλικών σχιστολίθων.
2. Σειρά φυλλιτών – χαλαζιτών	ph	
3. Ενότητα Τρυπαλίου	Ks-k	ραουβάκες, κρυσταλλικοί ασβεστόλιθοι και δολομιτικοί ασβεστόλιθοι, με μελανό ή τεφρομέλανο χρώμα και βιτουμενιώδη οσμή.
4. Μεταφλύσχης πλακωδών ασβεστόλιθων	T, k-D	Φυλλίτες και τους αργιλικούς σχιστόλιθους στους οποίους μεταβαίνουν σταδιακά οι πλακώδεις ασβεστόλιθοι και οι οποίοι αποτελούν τον μεταμορφωμένο φλύσχη. Στο κατώτερο τμήμα είναι τα μεταβατικά στρώματα που αποτελούνται κυρίως από τεφροπράσινους ή και ερυθρούς ασβεστοφυλλίτες.
5. Πλακώδεις ασβεστόλιθοι	(PC-k)	Πρόκειται για την γνωστή σειρά κρυσταλλικών ασβεστόλιθων με ενδιάμεσες στρώσεις ή βολβούς από κερατόλιθους. Η λιθολογική ομοιογένεια, η αδιατάρακτη ιζηματογένεση, το πάχος των στρώσεων και άλλα χαρακτηριστικά που παρατηρούνται σε όλες τις εμφανίσεις των πλακωδών ασβεστόλιθων οδηγούν στο συμπέρασμα ότι πρόκειται για πελαγικά ιζήματα.
6. Υποκείμενα στρώματα των πλακωδών ασβεστόλιθων	(ph-D-k)	Αποτελούνται από φυλλίτες, δολομίτες, μικρολατυποπαγείς ασβεστόλιθους, πυριτόλιθους και αργίλους. Συχνές εναλλαγές κλαστικών ιζημάτων (άργιλοι, ασβεστοφυλλίτες, ψαμμίτες, μικρολατυποπαγείς ασβεστόλιθοι).

7.3 Σχεδιασμός διασκόπησης στο οροπέδιο του Ομαλού

Για τη συλλογή των δεδομένων χρησιμοποιήθηκε το σύστημα γεωραντάρ PulseEKKO 1000 με κεντρική συχνότητα εκπομπής 110MHz. Πραγματοποιήθηκαν συνολικά σαράντα μία (41) παράλληλες γραμμές μελέτης μήκους είκοσι (20) μέτρων στη διαύθυνση N-B, με ισοαπόσταση μεταξύ των γραμμών ίση με ένα μέτρο (1m) και ισοαπόσταση μεταξύ των σταθμών παρατήρησης ίση με 0.25m. Ο κάνναβος σχεδιάστηκε ώστε η πρώτη γραμμή μελέτης να περνάει εφάπτομενικά από τα όρια της επιφανειακής δολίνης.

7.4 Επεξεργασία των δεδομένων

Το πρώτο στάδιο της επεξεργασίας περιλάμβανε την απομάκρυνση της επαγωγικής συνιστώσας, την αποκοπή του θορύβου υποβάθρου καθώς και την απομάκρυνση του υψίσυχνου θορύβου με τη χρήση τρισδιάστατου τελεστή εξομάλυνσης των δεδομένων.

Κατόπιν υπολογίστικαν τέσσερα στιγμιαία και δυο γεωμετρικά χαρακτηριστικά των δεδομένων του γεωραντάρ (πίνακας 7.1), ενώ το τελευταίο βήμα πριν από την ταξινόμηση αποτέλεσε η εξαγωγή των οριζόντιων τομών. Εξήχθησαν 30 οριζόντιες τομές από κάθε χαρακτηριστικό υπολογίζοντας τη μέση τιμή σε παράθυρο ίσο με 18 nsec (το οποίο αντιστοιχεί στο εύρος του εκπεμπόμενου παλμού). Στο σχήμα 7.3 φαίνονται οι τομές για τα χαρακτηριστικά των δεδομένων του γεωραντάρ που αντιστοιχούν σε χρόνο ίσο με 270 nsec.

Πίνακας 7.2. Τα χαρακτηριστικά των δεδομένων του γεωραντάρ που υπολογίστηκαν

Στιγμιαία χαρακτηριστικά		Γεωμετρικά χαρακτηριστικά	
➤	Στιγμιαίο πλάτος	➤	Συνάφεια των δεδομένων
➤	Στιγμιαίο πλάτος σε κλίμακα dB	➤	Συνάφεια της στιγμιαίας φάσης
➤	Στιγμιαία φάση		
➤	Στιγμιαία συχνότητα		

7.5 Συνδυασμένη ερμηνεία γεωφυσικών δεδομένων.

Στο σχήμα 7.4 φαίνονται τα αποτελέσματα των μετρήσεων με τη μέθοδο της σεισμικής διάθλασης όπου διακρίνονται δυο στρώματα. Το πρώτο εκτείνεται από την επιφάνεια μέχρι βάθος 8 m και αντιστοιχεί στις επιφανειακές προσχώσεις, ενώ το δεύτερο από τα 8 m έως τα 20 m αντιστοιχεί στο ασβεστολιθικό υπόβαθρο.

Στο χάρτη του EM (σχήμα 7.5) παρατηρείται μια ζώνη υψηλών αντιστάσεων (κόκκινο χρώμα) η οποία οφείλεται σε εξάρσεις του βράχου και μια ζώνη χαμηλών αντιστάσεων (μπλε χρώμα) στα βόρεια η οποία οφείλεται στο αργιλικό στρώμα.

Η οριζόντια τομή της στιγμιαίας φάσης του γεωραντάρ στα 270 nsec (σχήμα 7.6) αναδεικνύει μια ζώνη αρνητικών τιμών στα βόρεια η οποία οφείλεται στη ζώνη χαμηλής αντίστασης του EM, της οποίας το μέγιστο πλάτος είναι στα βόρεια. Πιθανόντατα οφείλεται σε τμήμα δολίνης η οποία φαίνεται επιφανειακά στα βόρεια του καννάβου και προεκτείνεται μέχρι τις παρυφές του βουνού.

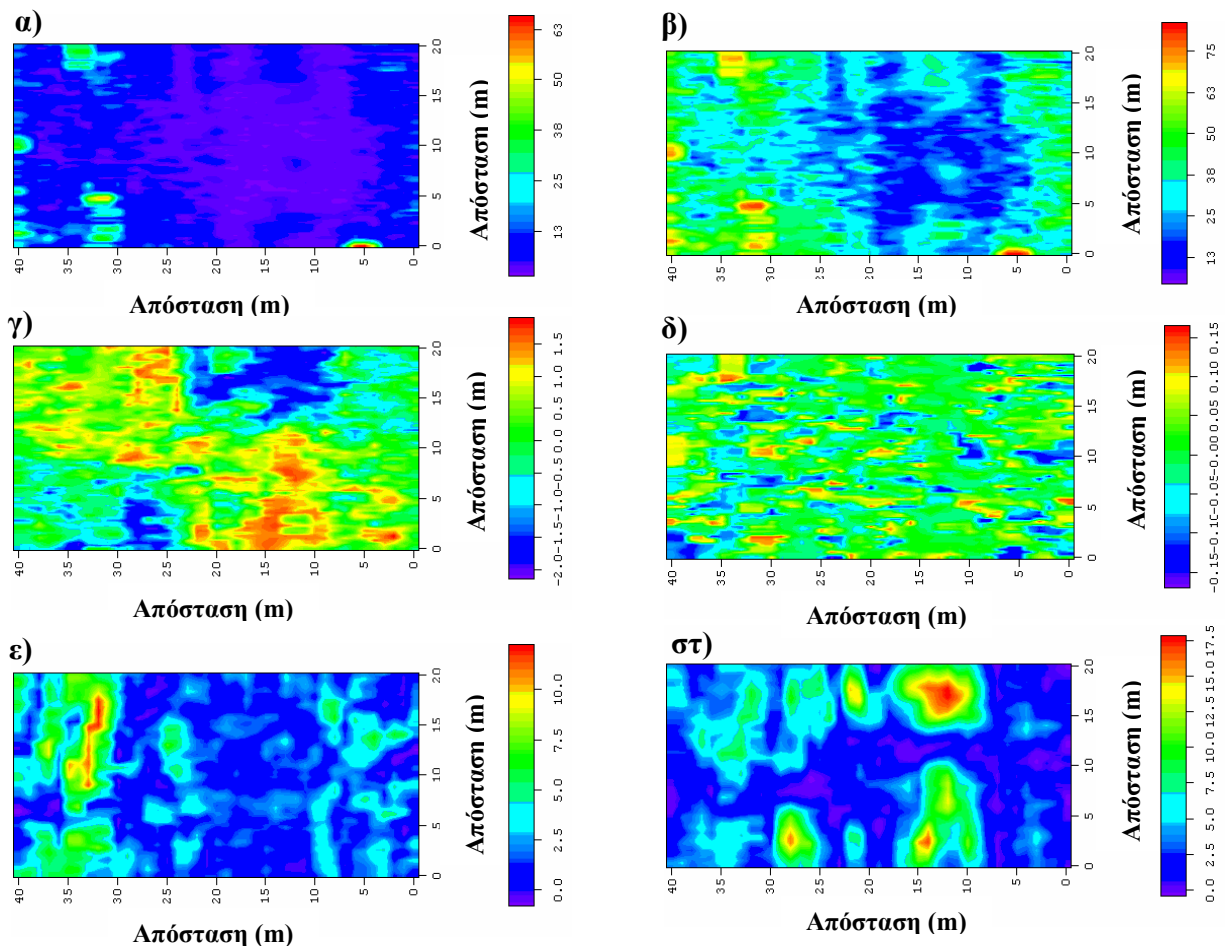
7.6 Ταξινόμηση των δεδομένων

Οι οριζόντιες τομές των χαρακτηριστικών από το προηγούμενο στάδιο κανονικοποιούνται, δηλαδή αφαιρείται η μέση τιμή και διαιρούνται με την τυπική απόκλιση. Πραγματοποιήθηκε σενάριο ταξινόμησης το οποίο περιλάμβανε την ταξινόμηση και των έξι χαρακτηριστικών.

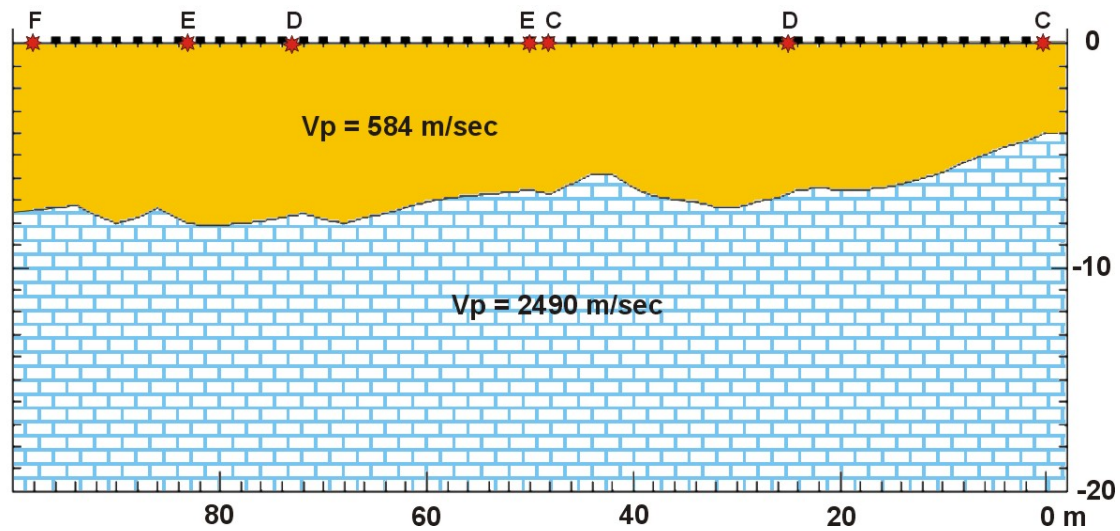
Η μέθοδος του δικτύου αυτο-οργανωμένης απεικόνισης, αν και γρήγορη υπολογιστικά, χρειάζεται αρκετές επαναλήψεις για να συγκλίνει, σε αντίθεση με τη μέθοδο των κ-μέσων τιμών. Αρχικά λοιπόν για την εκπαίδευση του δικτύου του δικτύου αυτο-οργανωμένης απεικόνισης πραγματοποιείται μια αρχική ταξινόμηση σε κάθε οριζόντια τομή με τη μέθοδο των κ-μέσων τιμών σε πενήντα (50) ομάδες. Τα χίλια πεντακόσια (1500) κέντρα συσσώρευσης που προκύπτουν σε κάθε περίπτωση τροφοδοτούνται στο δίκτυο αυτό-οργανωμένης απεικόνισης σαν δείγματα εκπαίδευσης και ταξινομούνται σε

είκοσι (20) τάξεις αντίστοιχα. Κατόπιν όλες οι τομές ταξινομούνται σε είκοσι ομάδες βάση των κέντρων συγκέντρωσης που προέκυψαν.

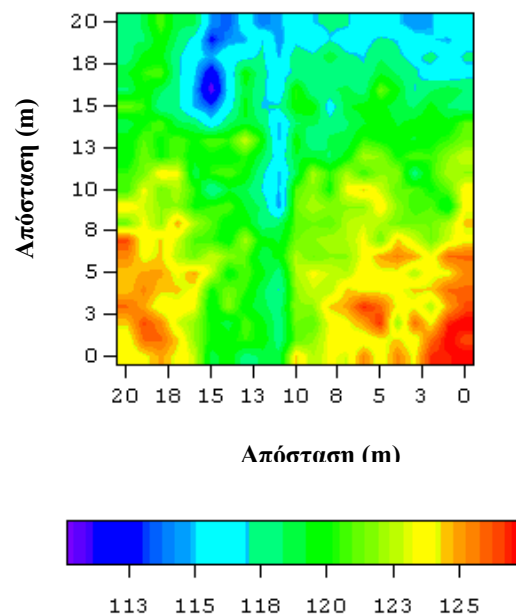
Στο σχήμα 7.7α φαίνεται το αποτέλεσμα της ταξινόμησης των 30 οριζόντιων τομών των χαρακτηριστικών των δεδομένων του γεωραντάρ. Στο σχήμα 7.7β διακρίνεται με κόκκινο χρώμα η εικοστή ομάδα των ταξινομημένων δεδομένων ενώ οι υπόλοιπες ομάδες είναι διάφανες, καθώς και μια οριζόντια τομή στα 270 nsec. Εκτός από την ανωμαλία που εντοπίστηκε από τη σύγκριση της στιγμιαίας φάσης με τις μετρήσεις του EM στα βόρεια, διακρίνεται άλλη μια ανωμαλία στα ΝΔ η οποία πιθανό να οφείλεται στην ύπαρξη δολίνης.



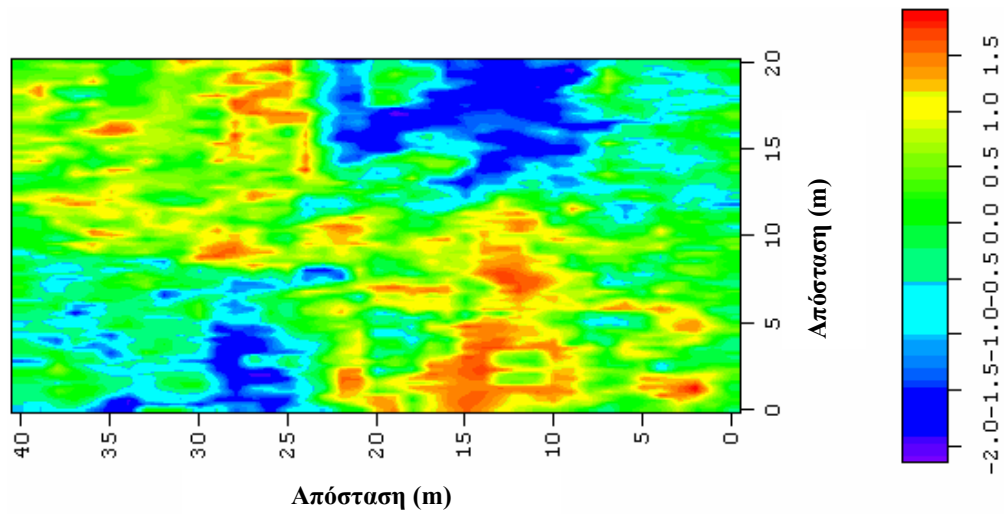
Σχήμα 7.3 Τα χαρακτηριστικά των δεδομένων του γεωραντάρ για την οριζόντια τομή στα 270 nsec. α) Στιγμιαίο πλάτος, β) Στιγμιαίο πλάτος σε κλίμακα dB, γ) Στιγμιαία φάση, δ) Στιγμιαία συχνότητα, ε) Συνάφεια των δεδομένων, στ) Συνάφεια της στιγμιαίας φάσης.



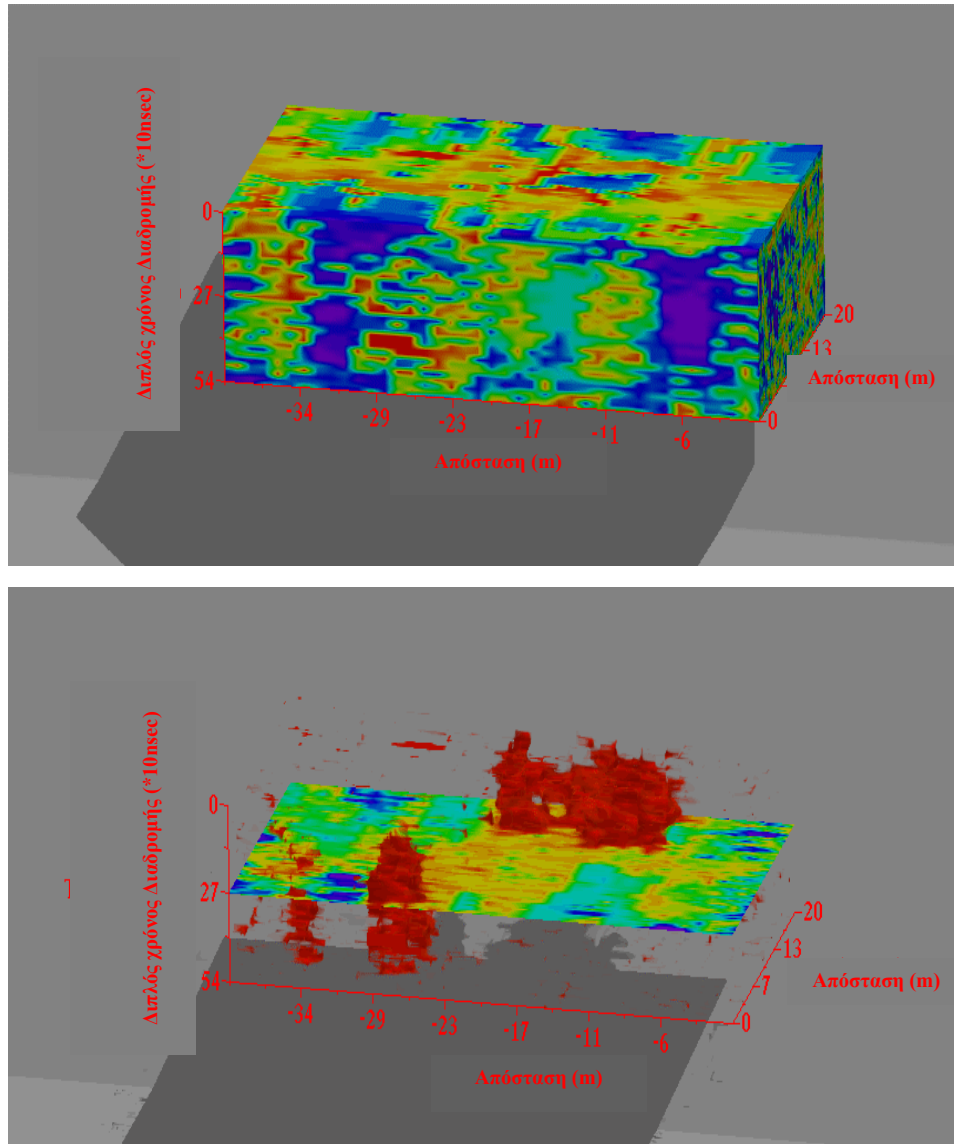
Σχήμα 7.4. Τα αποτελέσματα των δεδομένων σεισμικής διάθλασης. Διακρίνονται δυο στρώματα. Ένα επιφανειακό το οποίο αποδίδεται σε προσχώσεις το οποίο εκτείνεται μέχρι 8 m βάθος και ένα στρώμα το οποίο αποδίδεται στο ασβεστολιθικό υπόβαθρο.



Σχήμα 7.5. Ο χάρτης του EM. Οι χαμηλές τιμές (μπλε) στα βόρεια αποδίδονται σε αργιλικό στρώμα, ενώ οι υψηλές (κόκκινο) σε εξάρσεις του βράχου.



Σχήμα 7.6. Η στιγμιαία φάση των δεδομένων του γεωραντάρ για τη οριζόντια τομή στα 270 nsec. Οι αρνητικές τιμές στα βόρεια συμφωνούν με τις χαμηλές αντιστάσεις του χάρτη του EM και αποδίδονται σε δολίνη η οποία εκτείνεται προς τα βόρεια.



Σχήμα 7.7. α) Τρισδιάστατος όγκος της ταξινόμησης των οριζόντιων τομών των χαρακτηριστικών των δεδομένων του γεωραντάρ σε 20 ομάδες. β) Διακρίνονται τα δεδομένα που ανήκουν στην εικοστή ομάδα των δεδομένων (κόκκινο χρώμα) και μία οριζόντια τομή στα 270 nsec.

ΚΕΦΑΛΑΙΟ 8**ΣΥΜΠΕΡΑΣΜΑΤΑ - ΠΡΟΤΑΣΕΙΣ****8.1 Εισαγωγή.**

Στα πλαίσια της παρούσας διατριβής προτείνεται μια διαδικασία επεξεργασίας δεδομένων γεωραντάρ, η οποία μεταξύ άλλων περιλαμβάνει τη χρήση μεθόδων ταξινόμησης. Στόχος ήταν να ερευνηθεί το κατά πόσο οι μέθοδοι αυτές είναι χρήσιμες στην ερμηνεία των δεδομένων του γεωραντάρ, σε σχέση με τις κλασικές μεθόδους ερμηνείας.

Για την επίτευξη του στόχου δημιουργήθηκαν αλγόριθμοι επεξεργασίας δεδομένων στο Matlab, οι οποίοι εφαρμόστηκαν σε δυο διαφορετικά σύνολα δεδομένων. Το πρώτο προήλθε από ελεγχόμενο πείραμα στο εργαστήριο (sandbox experiment), όπου ήταν δυνατός ο έλεγχος των αποτελεσμάτων. Το δεύτερο σύνολο δεδομένων προέρχεται από γεωφυσική διασκόπηση στον αρχαιολογικό χώρο Άπτερα στο νομό Χανίων και συγκεκριμένα σε χώρο όπου υπάρχει τρίκλιτη ρωμαϊκή δεξαμενή.

8.2 Πρωτοτυπία της εργασίας.

Στα πλαίσια τη παρούσας διδακτορικής διατριβής αναπτύχθηκε μεθοδολογία επεξεργασίας και ερμηνείας δεδομένων γεωραντάρ η οποία περιλαμβάνει τον υπολογισμό χαρακτηριστικών των δεδομένων και τη συνδυασμένη ερμηνεία τους με τη χρήση μεθόδων ταξινόμησης. Η πρωτοτυπία της εργασίας έγκειται στο ότι τα χαρακτηριστικά των δεδομένων καθώς και η ταξινόμησή τους χρησιμοποιούνται σαν εργαλεία ερμηνείας και όχι για την επιβεβαίωση της ερμηνείας από δεδομένα που έχουν επεξεργαστεί με κλασικά εργαλεία.

Αναφορικά με την ιδέα της ταξινόμησης των δεδομένων του γεωραντάρ, συμπεριλήφθηκαν διάφορες τεχνικές από τα πεδία της πολυδιάστατης στατιστικής, των νευρωνικών δικτύων και της ασαφούς λογικής. Με τις μεθόδους αυτές ταξινομήθηκαν τόσο τα δεδομένα του γεωραντάρ και τα χαρακτηριστικά τους, οι κύριες συνιστώσες των χαρακτηριστικών και η πλήρης κυματομορφή ανακλώμενων κυμάτων. Αναφέρεται ότι η ανάλυση κυρίων συνιστωσών αντιμετωπίζεται σαν μέθοδος ταξινόμησης και όχι σαν ενδιάμεσο στάδιο για την ταξινόμηση (Coleou και συνεργάτες 2002).

8.3 Συμπεράσματα.

Για την ερμηνεία των δεδομένων γεωραντάρ της παρούσας διατριβής επιλέχθηκαν συγκεκριμένες μέθοδοι αυτόματης και καθοδηγούμενης ταξινόμησης. Για την αυτόματη ταξινόμηση χρησιμοποιήθηκαν οι μέθοδοι κ-μέσων τιμών και τα δίκτυα αυτο-οργανωμένης απεικόνισης, ενώ για την καθοδηγούμενη ταξινόμηση τα δίκτυα διανυσματικής μάθησης και πιθανότητας. Θα πρέπει να τονιστεί ότι τα ταξινομημένα δεδομένα δεν υποκαθιστούν ούτε τα αρχικά δεδομένα ούτε τα χαρακτηριστικά τους, αλλά ο ρόλος των μεθόδων ταξινόμησης είναι βοηθητικός για την ερμηνεία των δεδομένων γεωραντάρ.

Η μέθοδος των κ-μέσων τιμών είναι υπολογιστικά γρήγορη, αλλά είναι ευαίσθητη σε σχέση με τον αριθμό των τάξεων. Αυτό φάνηκε στο παράδειγμα της οριζόντιας τομής από το ελεγχόμενο πείραμα. Αντίθετα, η μέθοδος των δικτύων αυτο-οργανωμένης απεικόνισης είναι υπολογιστικά αργή, αλλά δεν είναι ευαίσθητη στον αριθμό των τάξεων. Λύση για ταξινόμηση μεγάλου όγκου δεδομένων γεωραντάρ αποτελεί η επιλογή

τυχαίων δειγμάτων εκπαίδευσης. Αυτή ενέχει όμως τον κίνδυνο να μη συμπεριληφθούν σημαντικές καταγραφές στην εκπαίδευση. Η επαναδειγματοληψία των δεδομένων του γεωραντάρ, ώστε να ελαττωθεί ο όγκος τους και να γίνει εκπαίδευση του δικτύου ενέχει τον ίδιο κίνδυνο. Η ταξινόμηση των δεδομένων με την μέθοδο των κ-μέσων τιμών σε πολλές τάξεις, και η τοπολογική οργάνωσή τους σε μικρότερο αριθμό τάξεων με τη χρήση των δικτύων αυτο-οργανωμένης απεικόνισης αποτελεί την καλύτερη λύση. Επίσης, προτείνεται η ταξινόμηση μέρους των δεδομένων με τα δίκτυα αυτο-οργανωμένης απεικόνισης, όπως άλλωστε έγινε στο παράδειγμα των δεδομένων από την αρχαία Άπτερα. Τέλος, αναφέρεται ότι τα παραπάνω προβλήματα έχουν μικρότερη βαρύτητα, όταν χρησιμοποιείται ισχυρός υπολογιστής ή πλατφόρμα Unix, όπου γίνεται καλύτερη διαχείριση των πόρων του συστήματος οπότε οι υπολογισμοί είναι πιο γρήγοροι.

Για την καθοδηγούμενη ταξινόμηση δεν τίθενται παρόμοιοι προβληματισμοί. Χρησιμοποιήθηκαν δυο διαφορετικές μέθοδοι καθοδηγούμενης ταξινόμησης (δίκτυα πιθανότητας και δίκτυα διανυσματικής μάθησης) με τα ίδια δεδομένα εκπαίδευσης και τα αποτελέσματα ήταν παρόμοια. Τα δίκτυα πιθανότητας απλώς ήταν πιο γρήγορα.

Όσον αφορά στην επιλογή των χαρακτηριστικών για την ταξινόμηση των δεδομένων του γεωραντάρ, δεν υπάρχει δυστυχώς κάποιος κανόνας που να υπαγορεύει ποια είναι τα κατάλληλα χαρακτηριστικά επομένως φαίνεται να είναι περισσότερο θέμα εμπειρίας. Αυτό φάνηκε στις δοκιμές για την ταξινόμηση των δεδομένων του ελεγχόμενου πειράματος, όπου σε κάποιες δοκιμές όλοι οι σωλήνες ταξινομήθηκαν στην ίδια ομάδα, ενώ σε άλλες δοκιμές ταξινομήθηκαν σε διαφορετικές ομάδες. Στο παράδειγμα πάλι από την αρχαία Άπτερα τα αποτελέσματα για διαφορετικούς συνδυασμούς χαρακτηριστικών ήταν σχεδόν ταυτόσημα. Προκύπτει ότι πρέπει να γίνονται περισσότερες από μια δοκιμές και τα αποτελέσματα να μελετώνται πάντα συγκριτικά με τα αρχικά δεδομένα. Απ' όλα τα σενάρια που εξετάστηκαν για την αυτόματη ταξινόμηση των δεδομένων προέκυψε ότι ο συνδυασμός χαρακτηριστικών πλάτους, φάσης και συνάφειας δίνει τα πιο ικανοποιητικά αποτελέσματα. Επίσης σε όλες τις περιπτώσεις που εξετάστηκαν επιλέχθηκε ο αριθμός των τάξεων ίσος με 20, ο οποίος είναι αρκετός για να περιγράψει τα δεδομένα του γεωραντάρ.

Στις μεθόδους ταξινόμησης ανακλαστήρων, η διαδικασία επιλογής των είναι επίπονη διαδικασία, ταυτόχρονα όμως σημαντική. Αυτή η ταξινόμηση τόσο των δεδομένων του ελεγχόμενου πειράματος όσο και των δεδομένων της αρχαίας Άπτερας αποδείχτηκε ιδιαίτερα χρήσιμη για την ερμηνεία τους.

Η ιδέα της ταξινόμησης των δεδομένων και των χαρακτηριστικών του γεωραντάρ προέκυψε ανατρέχοντας σε δημοσιεύσεις καθώς και ιστοσελίδες εταιριών που ασχολούνται με την ερμηνεία σεισμικών δεδομένων στην έρευνα για πετρέλαιο. Αυτή η διδακτορική διατριβή κατέδειξε ότι οι μέθοδοι ταξινόμησης αποτελούν χρήσιμο εργαλείο και για την ερμηνεία των δεδομένων του γεωραντάρ.

8.4 Μελλοντικοί στόχοι.

Στα πλαίσια της παρούσας διατριβής χρησιμοποιήθηκε πληθώρα αλγόριθμων επεξεργασίας δεδομένων γεωραντάρ με στόχο την ταξινόμησή τους. Σαν μελλοντικός στόχος έχει τεθεί η ανάπτυξη μεθόδων υπολογισμού νέων χαρακτηριστικών καθώς και συνδυασμένης χρήσης δεδομένων γεωραντάρ με διαγραφίες.

Όσον αφορά στα χαρακτηριστικά, ήδη από τη βιβλιογραφία φαίνεται η τάση της βιομηχανίας πετρελαίου για τη χρήση των χαρακτηριστικών υφής, όπως η τοπική εντροπία (*texture attributes*) (West και συνεργάτες 2002) καθώς και συνάφειας, όπως η ομοιότητα (*coherence attributes*) (Marfurt και συνεργάτες 1998) σε σεισμικές τομές και τρισδιάστατα σεισμικά δεδομένα. Αναφέρεται ότι ήδη γίνονται κάποιες πρώτες προσπάθειες για την δημιουργία αλγόριθμων υπολογισμού των χαρακτηριστικών υφής και δοκιμάζονται στα δεδομένα του ελεγχόμενου πειράματος.

Όσον αφορά στην ταξινόμηση των δεδομένων του γεωραντάρ κρίνεται χρήσιμη η δημιουργία ενός αλγόριθμου ο οποίος θα μπορεί να συνδυάζει μόνο τα προτερήματα των αλγόριθμων αυτόματης ταξινόμησης. Ένας τέτοιος αλγόριθμος θα πρέπει να είναι υπολογιστικά γρήγορος, να οργανώνει τοπολογικά τις ομάδες και να μην είναι ευαίσθητος στον αριθμό των ομάδων. Χρήσιμο θα είναι να προσδιορίζει μόνος του τον αριθμό των ομάδων, με βάση κάποια κριτήρια όπως για παράδειγμα την πυκνότητα των δεδομένων στο διάγραμμα διασποράς

Επίσης χρειάζεται περαιτέρω αξιολόγησης των μεθόδων ταξινόμησης με τη συνδυασμένη χρήση δεδομένων γεωραντάρ, δεδομένων από γεωτρήσεις (π.χ. διαγραφίες) καθώς και άλλων γεωφυσικών δεδομένων. Έρευνα σε αυτό το θέμα γίνεται σε μεγάλο βαθμό τα τελευταία δέκα χρόνια από τη βιομηχανία πετρελαίου, αλλά με τη χρήση μόνο των νευρωνικών δικτύων. Άλλες μέθοδοι, όπως για παράδειγμα οι γενετικοί αλγόριθμοι (*genetic algorithms*), οι μηχανές διανυσματικής υποστήριξης (*support vector machines*), ακόμα και οι κλασσικές γεωστατιστικές μέθοδοι, όπως η μέθοδος kriging, δεν εφαρμόζονται, παρά με ελάχιστες εξαιρέσεις. Αναφέρεται ότι σε έρευνες σχετικά με την παρακολούθηση σταθμών πυρηνικής ενέργειας (Dubois και Galmarini 2005) διαπιστώθηκε ότι οι μηχανές διανυσματικής υποστήριξης και η μέθοδος kriging, απέδωσαν καλύτερα αποτελέσματα από τα νευρωνικά δίκτυα.

Όσον αφορά στο λογισμικό επεξεργασίας δεδομένων γεωραντάρ, ήδη υπάρχουν κάποιες ιδέες ώστε να γίνει φιλικότερο στον χρήστη όπως, η δημιουργία πρότυπων διαγραμμάτων ροής. Τα διαγράμματα ροής θα είναι αποθηκευμένα και θα μπορούν να εφαρμόζονται αυτόματα σε κάθε σύνολο δεδομένων. Χρήσιμο, πιθανώς θα ήταν να γίνει η μεταγλώττιση του λογισμικού επεξεργασίας σε εκτελέσιμη μορφή (*stand alone application*). Όμως το λογισμικό βρίσκεται σε συνεχή εξέλιξη, προστίθενται νέοι αλγόριθμοι και δυνατότητες και επιπλέον δεν προορίζεται για εμπορική χρήση, ικανοποιώντας κυρίως τις ανάγκες του εργαστηρίου Εφαρμοσμένης Γεωφυσικής.

BIBΛΙΟΓΡΑΦΙΑ.

1. Aminzadeh F., de Groot P. (2004). Soft computing for qualitative and quantitative seismic object and reservoir property prediction. Part 1: Neural network applications. First Break Vol. 22, pp. 49-54.
2. Aminzadeh F., Wilkinson D. (2004). Soft computing for qualitative and quantitative seismic object and reservoir property prediction. Part 2: Fuzzy Logic Applications. First Break Vol. 22, pp. 69-78.
3. Annan A. P. (1999). Practical Processing of GPR Data. www.sensoft.on.ca
4. Anstey N. (1972). Seiscom '72. (Seiscom Limited internal report).
5. Asevedo H., Pennington D. (2003). Porosity and lithology prediction at Caballos Formation in the Puerto Colón Oil Field in Putumayo (Colombia). The Leading Edge, pp.1135-1141.
6. Bachrach R., Beller M., Chu Ching Liu, Perdomo J., Shelandier S., Dutta M., Benabentos M. (2004). Combining rock physics analysis, full waveform prestack inversion and high-resolution seismic interpretation to map lithology units in deep water: A Gulf of Mexico case study. The Leading Edge, pp. 378-383.
7. Balch, A.H. (1970), Color sonograms: A new dimension in seismic data interpretation: Geophysics, Vol. 36, pp. 1074-1098.
8. Banchs R. E., Michelena R. J. (2002). From 3D seismic attributes to pseudo-well-log volumes using neural networks: Practical considerations. The Leading Edge, pp. 996-1001.
9. Barnes A. E. (1992). The calculation of instantaneous frequency and instantaneous bandwidth. Geophysics, Vol. 57, No 11, pp 1520-1524.
10. Barnes A. E. (1998). The complex seismic trace made simple. The Leading Edge. Vol. 17, No 17, pp. 473-476.
11. Basu T., Claverie M., Nolan D., Kamarolzaman B., Yahya and Mustafa Suleiman (2004). Facies analysis: Integration of core and log data using a neural network as input for reservoir modeling in Betty Field, Malaysia. The Leading Edge, pp.794-797.

12. Benson A., Stubben, M. (1997). Mapping groundwater contamination using DC resistivity and VLF geophysical methods. A case study. *Geophysics*, Vol. 62, pp. 80-86
13. Boadu F. K. (1997). Relating the Hydraulic Properties of a fractured rock mass to seismic attributes. Theory and numerical experiments. *Int. J. Rock Mech. Min.* Vol 34, No 6, pp. 885-895.
14. Bose A., Singh V., Tandon A.K., B.S. Josyulu, and Mahesh Chandra, (2004). A case study of stratigraphic and lithologic interpretation of thin reservoirs through an integrated approach. *The Leading Edge*, pp. 966-972.
15. Brewster M., Annan P. (1994). Ground penetrating radar monitoring of a DNAPL release: 200MHz radar. *Geophysics*, Vol. 59, pp. 1211-1221.
16. Brown A. R. (1996). Seismic Attributes and their classification. *The Leading Edge*, Vol. 15, No 10, pp. 1090.
17. Calderon-Macias C., Sen M., Stoffa P. (1997). Hopfield neural networks and mean field annealing for seismic deconvolution and multiple attenuation. *Geophysics*, Vol. 92, pp. 992-1002.
18. Calvert C., Foreman L., Tingting Yao T., and Bishop G. (2004). Spectral component geologic modeling: A novel approach for integrating seismic data into geologic models. *The Leading Edge*, pp. 466-470.
19. Carr M., Cooper R., Smith M., Taner T. T., Taylor T., (2001 α). The generation of rock and fluid properties volume via the integration of multiple seismic attributes and log data. Internal paper at www.rocksolidimages.com
20. Carr M., Cooper R., Smith M., Taner T. T., Walls J., (2001 β). The integration of surface seismic and borehole data using artificial neural network clustering methods. ASEG 15th Geophysical conference and exhibition, August 2001, Brisbane, Extended abstracts.
21. Cartabia G., Zerilli A., Apolloni B. (1994). Lineament recognition for potential fields images using a learning algorithm for Boltzmann machines. SEG 64th Ann. Internat. Mtg, pp. 432-435.

22. Chakravarthy S., Chunduru R., Mezzatesta A., Fanini (1999). Detection of layer boundaries from array detection tool responses using neural networks. SEG 69th Ann. Internat. Mtg, pp. 140-143.
23. Chen Q., Sidney S. (1997). Seismic attribute technology for reservoir forecasting and monitoring. The Leading Edge. 1997 Yearbook. Vol 16, No 5, pp 445-456.
24. Coleou T., Poupon M, Azbel K, (2002) Unsupervised seismic facies classification: A review and comparison of techniques and implementation. The Leading Edge, pp. 942-953.
25. Corbeanu R. M., McMechan G. A., Szerbiak R. B., Soegard K. (2002). Prediction of 3D fluid permeability and mudstone distribution from ground penetrating radar (GPR) attributes: example from the Cretaceous Ferron Sandstone Member, east-central Utah. Geophysics, Vol. 67, N. 5, pp. 1495-1504.
26. Daniels J.J, Brower J., Baumgartner F. (1998). High resolution GPR at Brookhaven National Laboratory to delineate complex subsurface targets. Journal of Environmental and Engineering Geophysics, Vol. 3, pp. 1-5.
27. Daniels J., Roberts R., Vendl M. (1992). Site studies of ground penetrating radar for monitoring petroleum product contaminants. Proceedings of the Symposium on the application of Geophysics to engineering and environmental problems. Pp 597-609.
28. Daniels, J.J, Grumman D, Vendl M. (1997). Vertical Incident Three Dimensional GPR. Journal of Environmental and Engineering Geophysics, Vol. 2, pp. 1-9.
29. Davis, J. L., Annan, A. P. (1989), Ground penetrating radar for high resolution mapping of soil and rock stratigraphy: Geophysical Prospecting, 37, pp. 531-551.
30. Dominic D.F., Egan K., Carney C., Wolfe P.J., Boardman M.R. (1995). Delineation of shallow stratigraphy using ground penetrating radar. Journal of Applied Geophysics, Vol. 33, pp. 167-176.
31. Dorrington K. P. and Linkz C. A. (2004). Genetic-algorithm/neural-network approach to seismic attribute selection for well-log prediction. Geophysics, Vol. 59, No.1, pp. 12-21.

32. Dowd P. A. and Iguzquiza E. P. (2005) Estimating the boundary surface between geologic formations from 3D seismic data using neural networks and geostatistics. *Geophysics*, Vol. 70, No.1, pp. 1-12.
33. Dubois G., Galmarini S. (2005). Spatial Interpolation Comparison (SIC) 2004: introduction to the exercise and overview of the results. In G. Dubois (ed) *Automatic mapping algorithms for routine and emergency monitoring data*. Office for Official Publication of the European Communities, Luxemburg, EUR 21955 EN, EC. pp 7-18.
34. Fogg A. N. (2000). Petro-seismic classification using neural networks: UK onshore. *SEG 70th Ann. Internat. Mtg. Expanded abstracts*, pp.1426-1429.
35. Fossati, M., Zerilli A., Ronchini G., Apolloni B. (1992). Lineament analysis for potential fields data using neural networks. *SEG 62th Ann. Internat. Mtg*, pp. 6-9.
36. Gastaldi C., Biguenet J. P., Pazzis L. (1997) Reservoir characterization from seismic attributes: An example from the Peciko Field (Indonesia). *The Leading Edge*, pp. 263-266.
37. Gazdag J., (1978). Wave equation migration with the phase shift method. *Geophysics*, Vol.42, No 7, pp. 1342-1351.
38. Greenhouse J., Brewster M. (1993). Geophysics and solvents: The Borden experiment. *The Leading Edge*, pp. 261-267.
39. Grumman D. L., Daniels J. J. (1995) Experiments on the detection of organic contaminants in the Vadoze zone. *Journal of Environmental and Engineering Geophysics*. Vol. No. pp. 31-38.
40. Guo Y., Hansen R., Harthill N. (1992). Feature recognition from potential fields using neural networks. *SEG 62th Ann. Internat. Mtg*, pp. 1-5.
41. Hampson D., Schuelke J., Quieren J. (2000). Use of multi attribute transforms to predict log properties from seismic data. *Geophysics*, Vol. 56, pp. 220-236.
42. Haykin S. (1999). *Neural Networks. A comprehensive foundation*. Second edition, Prentice Hall.
43. Heggland R., Meldahl P., Bril B., Groot P. (1999). The chimney cube, an example of semi-automated detection of seismic objects by directive attributes and neural networks: Part I & II; interpretation. *SEG 1999, Expanded Abstracts*.

44. Huang K., Liu W., Chang I. (1989). Hopfield model of neural networks for detection of bright spots. 59th Ann. Internat. Mtg, pp. 444-446.
45. Huang Z, Shimeld J., Williamson M., Katsube J. (1996). Permeability prediction with artificial neural network modeling in the Venture gas field, offshore eastern Canada. *Geophysics*, Vol. 61, 2, pp. 422-436.
46. Hugenschmidt J., Partl M.N., DeWitte H. (1998). GPR inspection of a mountain motorway in Switzerland. *Journal of Applied Geophysics*, Vol. 40, pp. 95-104.
47. Kalkomey C. T. (1997). Potential risks when using seismic attributes as predictors of reservoir properties. *The Leading Edge*, 16, pp. 247-251.
48. Kohonen T. (2001) *Self Organizing Maps*. Third edition. Springer.
49. Le L. H. T., Nyland E. (1990). Pattern analysis of seismic records. *Geophysics*, Vol. 55, No.1, pp. 20-28.
50. Leiphart D. J., Hart B. S. (2001). Comparison of linear regression and a probabilistic neural network to predict porosity from 3D seismic attributes in Lower Brushy Canyon channeled sandstones, southeast New Mexico. *Geophysics*, Vol. 66, No. 5, pp. 1349-1358.
51. Lemke, S. R., and Mankowski, L. C. (2000), GPR attribute analysis: Symp. Appl. Geophys. Eng. Environ. Prob., Proceedings, pp. 263-272.
52. Linari V. (2004). A practical approach to well-seismic data calibration. *The Leading Edge*, pp. 774-775.
53. Linari V., Stantiago M., Pastore C., Azbel C., Poupon M. (2003). Seismic facies analysis based on 3D multiattribute volume classification, La Palma Field, Maracaibo, Venezuela. *The Leading Edge*, pp. 32-36.
54. Liu X., Xue P., Li Y. (1989). Neural network method for tracing seismic events. SEG 59th Ann. Internat. Mtg, Expanded Abstracts, pp. 716-718.
55. Loeffler O., Bano M. (2004). Ground Penetrating Radar Measurements in a Controlled Vadose Zone: Influence of the Water Content. *Vadose Vone Journal.*, Vol. 3. pp. 1082-1092.
56. Lorenzo H., Cuellar V., Hernadez M.C. (2001). Close range radar remote sensing of concrete degradation in a textile factory floor. *Journal of Applied Geophysics*, Vol. 47, pp. 327-336.

57. Macias C. C., Sen M. K., Stoffa P. L. (2000). Artificial neural networks for parameter estimation in geophysics. *Geophysical Prospecting*, Vol.48, pp. 21-47.
58. Marfurt K. J., Kirlin R. L., Farmer S. L., Bahorich M. S. (1998). 3D seismic attributes using a semblance-based coherency algorithm. *Geophysics*, Vol. 63, No 4, pp. 1150-1165.
59. Matos M. C., Schroeder P. R., Petrobras J. (2004). Using Matching Pursuit and Self Organizing Maps for Seismic Reservoir Characterization of a Deep-water Field, Campos Basin, Offshore Brazil. SEG 74th Annual meeting, Denver-Colorado.
60. McCormac M. (1990). Seismic trace editing and first break picking using neural networks. 60th Ann. Internat. Mtg, pp. 321-324.
61. McCormack M. D., (1991). Neural computing in geophysics. *The Leading Edge*, 10 (1), pp. 11-15.
62. McCormack M. D., Stoitsits R. F., MacAllister D. J., Crawford K. D., (1999). Applications of genetic algorithms in exploration and production. *The Leading Edge*, pp. 716-718.
63. McCulloch, W. S., and Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity: *Bull. Math. Biophys.*, 5, pp. 115–133.
64. Moon A. P., Kalantzis F. (2001). Reservoir characterization using seismic waveform and feedforward neural networks. *Geophysics*, Vol. 66, No.5, pp.1450-1456.
65. Olhoeft, G.R. (1988). Selected bibliography on ground penetrating radar. *Proceedings of symposium on Applications of Geophysics in Engineering and Environmental Problems*, Golden. CO, pp. 463-520.
66. Orlando L. (2002). Detection of LNAPL using the instantaneous amplitude and frequency of ground-penetrating radar. *Geophysical prospecting* 50, pp. 27-41.
67. Parra J.O., Hacker C., Bennett M., Collier H. A., (2003). Permeability and porosity images based on NMR, sonic, and seismic reflectivity: Application to a carbonate aquifer. *The Leading Edge*, pp. 1102-1108.
68. Poulton M., Glass C., Sternberg B., (1989). Recognizing EM ellipticity patterns with neural networks. SEG 59th Ann. Internat. Mtg, pp. 208-211.

69. Poulton, M. M., (2002). Neural networks as an intelligence amplification tool: A review of applications: *Geophysics*, Vol. 67, pp. 979–993.
70. Poupon M., Gil J., Vannaxay D, Cortiula B., (2004) Tracking Tertiary delta sands (Urdaneta West, Lake Maracaibo, Venezuela): An integrated seismic facies classification workflow. *The Leading Edge*, 21, pp. 909-912.
71. Richards A. J. (1998). *Remote sensing digital analysis*. Third edition. Springer
72. Robertson J. D., Fisher D.A. (1988). Complex seismic trace attributes. *The Leading Edge*, pp. 22-26.
73. Robertson J. D., Nogami H. H. (1984). Complex seismic trace analysis of thin beds. *Geophysics*, Vol. 49, pp. 344-352.
74. Russel B., Hampson D., Schuelke J., Quirein J. (1997). Multiattribute seismic analysis. *The Leading Edge*.
75. Russel B., Hampson D., Totorov T. (2001). Combining Geostatistics and multiattribute transforms-A channel sand case Study. SEG 71th Ann. Internat. Mtg.
76. Saggaf M. M, Toskoz M. N., Mustafa H. M., (2003a). Estimation of reservoir properties from seismic data by smooth neural networks. *Geophysics*, Vol. 68, No.6, pp.1969-1983.
77. Saggaf M. M., Toksozz M. N., Marhoon M. I. (2003b). Seismic facies classification and identification by competitive neural networks. *Geophysics*, Vol. 68, No. 6, pp. 1984-1999.
78. Schmidt J., Hadsell F. (1992). Neural network stacking velocity picker. 60th Ann. Internat. Mtg, pp. 18-21.
79. Schmitt D. R. (1999). Seismic attributes for monitoring of a heated heavy oil reservoir. A case study. *Geophysics*, Vol. 64, pp. 368-377.
80. Schultz P., Ronen S., Hattori M., Corbett C. (1994). Seismic Guided estimation of log properties. Part I: A data driven interpretation methodology. *The Leading Edge*, 13, pp. 305-311.
81. Schultz P., Ronen S., Hattori M., Corbett C. (1994). Seismic Guided estimation of log properties. Part II: Using Artificial Neural Networks for non linear attributes calibration. *The Leading Edge*, 13 pp. 674-678.

82. Schultz P., Ronen S., Hattori M., Corbett C. (1994). Seismic Guided estimation of log properties. Part III: A controlled study. *The Leading Edge*, 13, pp. 770-776.
83. Senechal P., Perroud H., Senechal G. (2000). Interpretation of reflection attributes in a 3D GPR survey at Valle d'Ossau Pyrenees, France. *Geophysics*, Vol. 65, No 5, pp. 1435-1445.
84. Smith M. L., Scales J. A., Fischer T. L., (1992) Global search and genetic algorithms. *The Leading Edge*, pp. 22-26.
85. Spanoudakis S. N., Vafidis A., Papavasiliou A. (2004). Delineating a doline system using 3D Ground Penetrating Radar (GPR) data, complex trace attributes and neural networks: a case study in Omalos Highlands, Chania, Crete. AMIREG (Advances in Mineral Resources Management and Environmental Geotechnology) 1st International Conference, 7-9 June, Hania, Crete, Greece.
86. Sternberg B.K., McGill J.W. (1995). Archaeology studies in southern Arizona using ground penetrating radar. *Journal of Applied Geophysics*, Vol. 33, pp. 209-215.
87. Strecker U., Uden R., (2002). Data mining of 3D poststack seismic attribute volume using Kohonen self-organizing maps. *The Leading Edge*, pp. 1032-1037.
88. Taner M. T. (1992-2000). Attributes Revisited. <http://www.rocksolidimages.com>
89. Taner M. T., Koehler F., Sheriff R. E. (1979). Complex seismic trace analysis. *Geophysics*, Vol. 44, No 6, pp. 1041-1063.
90. Taner T., Lu L., Baysal E. (1998). Unified method for 2D and 3D refraction statistics with first break picking by supervised learning. SEG 58th Ann. Internat. Mtg.
91. Taner, M. T., Schuelke J. S., O'Doherty R., and Baysal E. (1994), Seismic attributes revisited: 64th Annual International Meeting, SEG, Expanded Abstracts, pp. 1104-1106.
92. Taner M. T. (1997). Kohonen's Self Organizing Networks with "Conscience". <http://www.rocksolidimages.com>
93. Todorov T. I., Stewart R. R., Hampson D. P. (1998a) Porosity prediction using attributes from 3C-3D seismic data. CREWES Research Report Vol. 10.

94. Todorov T., I., Stewart R., R., Hampson D.P. (1998b). CREWES Research report, Volume 10.
95. Tonn R., (2002). Neural network seismic reservoir characterization in a heavy oil reservoir. *The Leading Edge*, 21, pp. 309-312.
96. Toshioka T., Tsuchida T., Sasahara K. (1995). Application of GPR to detecting and mapping cracks in rock slopes. *Journal of Applied Geophysics*, Vol. 33, pp. 119-124.
97. Vafidis A., Spanoudakis N. S. (2003). Cluster analysis and classification of GPR data for pollution indicators. Deliverable D5.3 realized in the framework of WP5 of the HYGEIA CEE project.
98. Vafidis A., Spanoudakis N.S., Kritikakis G. (2002). Attribute and surface wave analysis. Deliverable D5.1 realized in the framework of WP5 of the HYGEIA CEE project.
99. Vasco D., Taylor C. (1991). Inversion of airborne gravity gradient data, south-eastern Oklahoma. *Geophysics*, Vol. 56, pp. 21-101.
100. Walls J. D., Taner T. T., Taylor G., Smith M., Carr M., Derzi N., Drummond J., McGuire D., Morris S., Bregar J., Lakings J., (2002). Seismic reservoir characterization of a U.S. Midcontinent fluvial system using rock physics, post-stack seismic attributes and neural networks. *The Leading Edge* 21, pp 428-436.
101. Walls J. D., Taner T. T., Taylor G., Smith M., Derzi N., Carr M., Drummond J., McGuire D., Morris S., Bregar J. (2000). Seismic reservoir characterization of a mid-continent fluvial system using rock physics, post-stack seismic attributes, and neural networks: a case history. 70th Ann. Internat. Mtg. SEG, pp. 1437-1439.
102. Walls J., Derzi N., Dumas D., Guidish T., Taner T. T., Taylor G. (1999). North Sea reservoir characterization using rock physics, seismic attributes, and neural networks: a case history. 69th Ann. Internat. Mtg. SEG, pp. 1572-1575.
103. Wang L., Mendel J. (1992). Adaptive minimum prediction error deconvolution and source wavelet estimation using Hopfield neural networks. *Geophysics*, Vol. 57, pp. 670-679.

104. West B. P., May S. R., Eastwood J. E., Rossen, C. (2002). Interactive seismic facies classification using textural attributes and neural networks. *The Leading Edge* 21, pp. 1042-1049.
105. Yilmaz O. (1987). *Seismic data processing*. Society of Exploration Geophysics
106. Zadeh, L.A. (1965). Fuzzy sets. *Information and Control* 8, 333-53.
107. Zeng H. (2004). Seismic geomorphology-based facies classification. *The Leading Edge*, pp. 644-688.
108. Σπανουδάκης Νικόλαος (2002). *Ανάπτυξη μεθόδων επεξεργασίας δεδομένων γεωραντάρ και εφαρμογή σε περιβαλλοντικά προβλήματα*. Μεταπτυχιακή Εργασία, Τμήμα Μηχανικών Ορυκτών Πόρων, Πολυτεχνείο Κρήτης.

ΠΑΡΑΡΤΗΜΑ

Π.1 Ανάλυση κυρίων συνιστωσών.

Ένα συνηθισμένο πρόβλημα στην επεξεργασία δεδομένων είναι η επιλογή των χαρακτηριστικών της (feature selection). Η επιλογή των χαρακτηριστικών αναφέρεται στη διαδικασία όπου το πεδίο των δεδομένων μετασχηματίζεται στο πεδίο των χαρακτηριστικών το οποίο έχει τις ίδιες διαστάσεις με το αρχικό πεδίο. Ο μετασχηματισμός είναι σχεδιασμένος έτσι ώστε τα αρχικά δεδομένα να παρίστανται από ένα μειωμένο αριθμό ενεργών χαρακτηριστικών διατηρώντας όμως την ουσιαστική πληροφορία του περιέχεται. Πιο συγκεκριμένα αν ένα διάνυσμα \mathbf{x} έχει m διαστάσεις, αυτό μπορεί να περιγραφεί από l χαρακτηριστικά όπου $l < m$.

Από την ανάλυση κυρίων συνιστωσών, η οποία αποκαλείται και μετασχηματισμός *Karhunen-Loeve*, προσδιορίζεται ένα νέο σύστημα συντεταγμένων στο οποίο τα δεδομένα είναι στατιστικώς ασυσχέτιστα. Δηλαδή προσδιορίζεται ένα σύστημα συντεταγμένων \mathbf{y} όπου ο πίνακας μεταβλητότητας-συμμεταβλητότητας Σ των δεδομένων μετατρέπεται σε διαγώνιο πίνακα. Ο πίνακας μεταβλητότητας-συμμεταβλητότητας περιγράφεται από τις σχέσεις:

$$\sigma_{ik} = \frac{\sum_{j=1}^N (x_{i,j} - \mu_i)(x_{k,j} - \mu_k)}{N-1},$$

$$\sigma_{ii} = \sigma_i^2 = \frac{\sum_{j=1}^N (x_{i,j} - \mu_i)^2}{N-1}.$$

Ο πίνακας στροφής \mathbf{T} που μετασχηματίζει τα δεδομένα \mathbf{x} σε δεδομένα \mathbf{y} δίδεται από τη σχέση:

$$\mathbf{y} = \mathbf{T}\mathbf{x},$$

προσδιορίζεται από τις ιδιοτιμές και τα ιδιοδιανύσματα του πίνακα μεταβλητότητας-συμμεταβλητότητας Σ . Οι ιδιοτιμές του πίνακα προκύπτουν επιλύοντας τη σχέση:

$$[\Sigma - \lambda \mathbf{I}],$$

όπου \mathbf{I} ο μοναδιαίος πίνακας. Για κάθε ιδιοτιμή λ_i προσδιορίζεται το αντίστοιχο ιδιοδιάνυσμα \mathbf{q}_i , οπότε ο πίνακας στροφής δίδεται από τη σχέση:

$$\mathbf{T} = [\mathbf{q}_1 \ \mathbf{q}_2 \ \dots \ \mathbf{q}_m].$$

Γενικά για τις κύριες συνιστώσες ισχύουν οι παρακάτω ιδιότητες:

- Οι κύριες συνιστώσες είναι στατιστικώς ασυσχέτιστες και έχουν μεταβλητότητες ίσες με τις ιδιοτιμές του πίνακα μεταβλητότητας Σ :

$$Var(y_k) = \mathbf{e}_k^T \Sigma \mathbf{e}_k = \lambda_k, \quad k = 1, 2, \dots, m$$

- Ο πίνακας μεταβλητότητας-συμμεταβλητότητας των κυρίων συνιστωσών είναι διαγώνιος της μορφής:

$$\Sigma_y = Var(\mathbf{y}) = \Lambda = \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & 0 \\ 0 & \dots & \lambda_m \end{pmatrix}, \quad \lambda_1 \geq \lambda_2 \geq \dots \lambda_m$$

- Το ποσοστό της ολικής μεταβλητότητας από τη συνιστώσα k δίδεται από τη σχέση:

$$\frac{\lambda_k}{\lambda_1 + \lambda_2 + \dots + \lambda_m}, \quad k = 1, 2, \dots, m.$$

Αυτό σημαίνει ότι αν ένα μεγάλο ποσοστό της ολικής μεταβλητότητας (μεγαλύτερο του 80%) μπορεί να αποδοθεί στις πρώτες λίγες συνιστώσες, τότε αυτές μπορεί να αντικαταστήσουν τις m αρχικές μεταβλητές χωρίς να απολεσθούν σημαντικές πληροφορίες.

Π.2 Μέθοδος των κ-μέσων τιμών.

Στον αλγόριθμο των K μέσων τιμών το κριτήριο για την ταξινόμηση των δεδομένων σε ομάδες είναι η ελαχιστοποίηση της μέσης τετραγωνικής απόστασης όλων των σημείων από τα αντίστοιχα κέντρα ομαδοποίησης:

$$E_K = \sum_{j=1}^K \|m_j(k) - \mathbf{x}\|^2$$

Ο αλγόριθμος περιλαμβάνει τα παρακάτω τέσσερα (4) βήματα:

- *Βήμα 1.*-Επιλέγονται (είτε τυχαία είτε αυτά είναι γνωστά) K αρχικά κέντρα ομάδων $m_1(1), m_2(2), \dots, m_K(1)$.
- *Βήμα 2.* Κάθε άγνωστο δείγμα \mathbf{x} αντιστοιχείται στην ομάδα ω_j εάν:

$$\|m_j(k) - \mathbf{x}\| < \|m_i(k) - \mathbf{x}\|$$

για κάθε $i = 1, 2, \dots, K, i \neq j$

- *Βήμα 3.* Υπολογίζονται τα νέα κέντρα ομάδων από τη σχέση:

$$m_j(k+1) = \frac{1}{n} \sum_{\mathbf{x} \in \omega_j} \mathbf{x}$$

όπου n , ο αριθμός των δειγμάτων στην ομάδα j .

- *Βήμα 4* Αν $m_j(k+1) = m_j(k)$ για $j = 1, 2, \dots, K$, ο αλγόριθμος συγκλίνει, διαφορετικά επαναλαμβάνεται από το βήμα 2.

Ένα δυνητικό-κρίσιμο πρόβλημα του αλγόριθμου είναι η επιλογή του αριθμού K των κέντρων συγκέντρωσης. Αλλάζοντας τον αριθμό K των κέντρων προκύπτουν τελείως διαφορετικές ομάδες. Κρίσιμη είναι επίσης και η επιλογή των αρχικών κέντρων συγκέντρωσης $m_1(1), m_2(2), \dots, m_K(1)$, αφού κάποιες ομάδες μπορεί να μείνουν άδειες άν τα κέντρα απέχουν μακριά από την κατανομή των δεδομένων. Ο αλγόριθμος των K μέσων τιμών είναι υπολογιστικά απλός και γρήγορος και χρησιμοποιείται συνήθως για την ταξινόμηση μεγάλων συνόλων δεδομένων.

Π.3 Δίκτυα αυτο-οργανωμένης απεικόνισης.

Τα δίκτυα αυτο-οργανωμένης απεικόνισης είναι ειδικές περιπτώσεις νευρωνικών δικτύων που βασίζονται στην ανταγωνιστική μάθηση. Τα νευρώνια σε αυτή την περίπτωση είναι τοποθετημένα στους κόμβους ενός πλέγματος το οποίο είναι μίας ή δύο διαστάσεων. Υπάρχουν και πλέγματα τριών διαστάσεων τα οποία είναι ιδιαίτερα σπάνια. Τα νευρώνια εκπαιδεύονται βάση των δειγμάτων εισόδου σύμφωνα με τους κανόνες της ανταγωνιστικής μάθησης, δηλαδή μόνο ένα νευρόνιο νικάει κάθε φορά τον ανταγωνισμό. Η διαφορά έγκειται στο ότι τα νευρώνια διατάσσονται στο πλέγμα το ένα σε σχέση με το άλλο και σχηματίζεται ένας τοπογραφικός χάρτης όπου η πλευρική διάταξή τους είναι ενδεικτική των φυσικών στατιστικών χαρακτηριστικών που περιέχονται στα δεδομένα.

Τα δίκτυα αυτο-οργανωμένης απεικόνισης εμπνεύστηκαν από τον ανθρώπινο εγκέφαλο ο οποίος είναι οργανωμένος ώστε οι διαφορετικές αισθήσεις να βρίσκονται σε συγκεκριμένες θέσεις του και οι οποίες χαρακτηρίζονται από τοπολογική διάταξη. Με λίγα λόγια τα νευρώνια που βρίσκονται κοντά στον χάρτη εκτελούν παρόμοιες λειτουργίες. Συγκεκριμένα η αφή, η όραση και η ακοή βρίσκονται σε διαφορετική θέση

στον εγκέφαλο. Κάθε μία από τις αισθήσεις περιέχει πολλά νευρώνια τα οποία εκτελούν παρόμοιες λειτουργίες.

Ο αλγόριθμος ξεκινάει ορίζοντας τα διανύσματα συναπτικών βαρών από γεννήτρια τυχαίων αριθμών. Με αυτό τον τρόπο δεν υπάρχει οργανωμένη διάταξη από την αρχή. Τρεις είναι οι θεμελιώδεις λειτουργίες ενός χάρτη αυτοδιοργάνωσης:

4. *Ανταγωνισμός.* Τα νευρώνια ανταγωνίζονται μεταξύ τους για το πιο ανταποκρίνεται καλύτερα στα διανύσματα εισόδου. Κάθε φορά που εισάγεται μία νέα είσοδος μόνο ένα νευρώνιο ανακηρύσσεται νικητής.
5. *Συνεργασία.* Το νευρώνιο νικητής καθορίζει μια τοπολογική γειτονία μέσα στην οποία τα νευρώνια συνεργάζονται.
6. *Προσαρμογή συναπτικών βαρών.* Αυτό το στάδιο είναι που διαφοροποιεί τα δίκτυα αυτο-οργανωμένης απεικόνισης από τα ανταγωνιστικά νευρωνικά δίκτυα. Σε αυτό το στάδιο διορθώνεται όχι μόνο το διάνυσμα συναπτικών βαρών του νευρωνίου νικητή, ώστε να έρθει πιο κοντά στο διάνυσμα εισόδου, αλλά και τα διανύσματα συναπτικών βαρών της γειτονιάς του αλλά σε μικρότερο βαθμό. Αυτός είναι και ο λόγος όπου υπάρχει μια τοπογραφική συνέχεια στον τοπογραφικό χάρτη.

Π.3.1 Ανταγωνιστική διαδικασία.

Ας υποτεθεί ότι οι διαστάσεις του χώρου είναι m , ενώ ένα τυχαίο διάνυσμα του χώρου παρίσταται ως:

$$\mathbf{x} = [x_1, x_2, \dots, x_m]$$

Το διάνυσμα των συναπτικών βαρών κάθε νευρωνίου στο δίκτυο έχει τις ίδιες διαστάσεις με τον χώρο. Το διάνυσμα αυτό παρίσταται ως:

$$\mathbf{w}_j = [w_{j1}, w_{j2}, \dots, w_{jm}], \quad j = 1, 2, \dots, l$$

όπου l ο συνολικός αριθμός νευρώνων στο δίκτυο. Το νευρόνιο που νικάει στον ανταγωνισμό είναι αυτό που έχει το μεγαλύτερο εσωτερικό γινόμενο $\mathbf{w}_j^T \mathbf{x} = \sum_{i=1}^m x_i \cdot w_{ji}$ ή τη μικρότερη Ευκλείδια απόσταση μεταξύ των \mathbf{x} και \mathbf{w}_j η οποία ισούται με:

$$\sqrt{\sum_{i=1}^m [x_i - w_{ji}]^2}$$

Π.3.2 Διαδικασία συνεργασίας.

Το νευρόνιο νικητής καθορίζει το κέντρο μιας γειτονιάς στο πλέγμα των νευρώνων τα οποία συνεργάζονται. Η ερώτηση είναι πως θα καθοριστεί μια τοπολογική γειτονιά συνεργασίας η οποία είναι νευροβιολογικά σωστή. Στον ανθρώπινο εγκέφαλο κάθε νευρόνιο που διεγείρεται επηρεάζει τα άμεσα γειτονικά του περισσότερο από τα μακρινότερα. Οπότε γύρω από το νευρόνιο νικητή καθορίζεται μια γειτονιά η οποία μικραίνει ομαλά με την πλευρική απόσταση. Η γειτονιά αυτή παρίσταται με το σύμβολο h_{ij} η οποία έχει σα κέντρο το νευρόνιο νικητή i που περικλείεται από κάποια νευρόνια, έστω j ένα από αυτά. Η απόσταση μεταξύ των νευρώνων i και j συμβολίζεται με $d_{i,j}$ και ικανοποιεί τις παρακάτω συνθήκες:

- Η τοπολογική γειτονιά είναι συμμετρική γύρω από το μέγιστο σημείο το οποίο καθορίζεται από τη σχέση $d_{i,j} = 0$, δηλαδή παίρνει τη μέγιστη τιμή στο νευρόνιο νικητή όπου η απόσταση $d_{i,j}$ είναι ίση με μηδέν.
- Το πλάτος της γειτονιάς h_{ij} μειώνεται μονότονα με την αύξηση της απόστασης $d_{i,j}$ τείνοντας στο μηδέν για $d_{i,j} \rightarrow \infty$

Μια τυπική επιλογή της τοπολογικής γειτονιάς που ικανοποιεί τις παραπάνω συνθήκες είναι η γκαουσιανή συνάρτηση:

$$h_{j,i(x)} = \exp\left(-\frac{d_{j,i}^2}{2\sigma^2}\right)$$

η οποία είναι ανεξάρτητη από τη θέση του νευρωνίου νικητή. Η παράμετρος σ είναι το ενεργό πλάτος της γειτονιάς και καθορίζει το βαθμό συμμετοχής κάθε νευρωνίου στη διαδικασία της μάθησης. Βιολογικά η γκαουσιανή συνάρτηση είναι πιο κατάλληλη από μια τετραγωνική ενώ βοηθάει τον αλγόριθμο να συγκλίνει πιο γρήγορα. Η απόσταση $d_{i,j}$ μεταξύ του νευρωνίου νικητή και ενός διεγερμένου νευρωνίου δίδεται από τη σχέση:

$$d_{j,i}^2 = \|\mathbf{r}_j - \mathbf{r}_i\|^2$$

όπου το διάνυσμα \mathbf{r}_j συμβολίζει τη θέση του διαγεργμένου νευρωνίου ενώ το διάνυσμα \mathbf{r}_i συμβολίζει τη θέση του νευρωνίου νικητή.

Ένα χαρακτηριστικό του αλγόριθμου είναι ότι το μέγεθος της τοπολογικής γειτονιάς μειώνεται με τον διακριτό χρόνο n . Αυτή η απαίτηση ικανοποιείται κάνοντας το ενεργό πλάτος σ της συνάρτησης γειτονιάς h_{ij} να μειώνεται με το χρόνο. Η εξάρτηση αυτή περιγράφεται από τη σχέση:

$$\sigma(n) = \sigma_0 \exp\left(-\frac{n}{\tau_1}\right), \quad n = 0, 1, 2, \dots,$$

όπου σ_0 είναι η αρχική τιμή της σ κατά την έναρξη του αλγόριθμου και τ_1 είναι μια χρονική σταθερά. Οπότε όσο ο χρόνος n (ο αριθμός των επαναλήψεων) αυξάνει, το ενεργό πλάτος $\sigma(n)$ μειώνεται εκθετικά και η τοπολογική γειτονιά συρρικνώνεται με παρόμοιο τρόπο.

Π.3.3 Διαδικασία προσαρμογής.

Η τελευταία διαδικασία στον αλγόριθμο είναι η προσαρμογή. Για αυτό-οργανωθεί το δίκτυο το διάνυσμα συναπτικών βαρών \mathbf{w}_j για το νευρώριο j πρέπει να αλλάξει σε

σχέση με το διάνυσμα εισόδου \mathbf{x} . Αν συμβολιστεί με $\mathbf{w}_j(n)$ το διάνυσμα συναπτικών βαρών του νευρωνίου j στην επανάληψη n , τότε το εκσυγχρονισμένο διάνυσμα $\mathbf{w}_j(n+1)$ στην επανάληψη $n+1$ δίδεται από τη σχέση:

$$\mathbf{w}_j(n+1) = \mathbf{w}_j(n) + \eta(n)h_{j,i}(n)(\mathbf{x} - \mathbf{w}_j(n)),$$

όπου $\eta(n)$ είναι η παράμετρος ρυθμού εκμάθησης. Η παραπάνω σχέση εφαρμόζεται σε όλα τα νευρώνια τα οποία βρίσκονται στη τοπολογική γειτονιά του νευρωνίου νικητή.

Μετά από επαναλαμβανόμενη εισαγωγή των δεδομένων εκπαίδευσης στο δίκτυο, τα διανύσματα των συναπτικών βαρών τείνουν να ακολουθήσουν την κατανομή των δεδομένων επειδή εκσυγχρονίζεται όχι μόνο το νευρώνιο νικητής αλλά και τα γειτονικά νευρώνια. Έτσι ο αλγόριθμος επιτυγχάνει μια τοπολογική διαδοχή μεταξύ των βαρών των νευρώνων αφού γειτονικά νευρώνια στο πλέγμα έχουν παρόμοιες τιμές.

Όπως φαίνεται στην παραπάνω σχέση η παράμετρος ρυθμού εκμάθησης $\eta(n)$, εξαρτάται από τον διακριτό χρόνο. Ειδικότερα πρέπει να ξεκινάει από μια τιμή η_0 και να μειώνεται βαθμιαία με το χρόνο n . Αυτή η απαίτηση ικανοποιείται διαλέγοντας μια εκθετική μείωση όπως φαίνεται από τη σχέση:

$$\eta(n) = \eta_0 \exp\left(-\frac{n}{\tau_2}\right),$$

όπου τ_2 είναι μια άλλη χρονική σταθερά.

Η διαδικασία της προσαρμογής γίνεται σε δυο φάσεις. Μια φάση διαρύθμισης ή αυτοδιοργάνωσης η οποία ακολουθείται από τη φάση της σύγκλισης. Αυτές έχουν ως εξής:

1. *Φάση διαρύθμισης ή αυτοδιοργάνωσης.* Σε αυτή τη φάση γίνεται η τοπολογική διαρύθμιση των διανυσμάτων συναπτικών βαρών και απαιτούνται τουλάχιστον χίλιες (1000) επαναλήψεις ωσότου ολοκληρωθεί. Προσεκτική πρέπει να είναι η

επιλογή της παραμέτρου ρυθμού εκμάθησης $\eta(n)$ καθώς και της συνάρτησης γειτονιάς h_{ij} :

- Η παράμετρος ρυθμού εκμάθησης $\eta(n)$ πρέπει να ξεκινάει από μια τιμή κοντά στο 0.1, να μειώνεται βαθμιαία αλλά να μη γίνει μικρότερη από 0.01. Αυτές οι συνθήκες ικανοποιούνται επιλέγοντας:

$$\eta_0 = 0.1$$

$$\tau_2 = 1000$$

- Η συνάρτηση γειτονιάς h_{ij} , αρχικά πρέπει να περιλαμβάνει όλα τα νευρώνια του δικτύου γύρω από το νευρώριο νικητή και να συρρικνώνεται με το χρόνο. Έτσι η αρχική τιμή του ενεργού πλάτους τίθεται ίση με την ακτίνα του πλέγματος. Συνεπώς η σταθερά τ_1 τίθεται ως:

$$\tau_1 = \frac{1000}{\log \sigma_0}$$

2. *Φάση σύγκλισης.* Ο αριθμός των επαναλήψεων σε αυτή τη φάση πρέπει να είναι μεγαλύτερος από πεντακόσιες φορές τον αριθμό των νευρωνίων εισόδου του πλέγματος. Αυτή η φάση λοιπόν διαρκεί συνήθως για χιλιάδες επαναλήψεις:

- Για καλή στατιστική ακρίβεια η παράμετρος ρυθμού εκμάθησης $\eta(n)$ πρέπει να έχει μικρή τιμή της τάξης του 0.01. Σε καμία περίπτωση δε πρέπει να φτάσει στο μηδέν.
- Η συνάρτηση γειτονιάς h_{ij} πρέπει να περιέχει μόνο τα κοντινότερα νευρώνια του νικητή και μπορεί να περιέχει μόνο ένα ή και κανένα γειτονικό νευρώριο.

Π.3.4 Σύνοψη των δικτύων αυτο-οργανωμένης απεικόνισης.

Ο αλγόριθμος για τα δίκτυα αυτο-οργανωμένης απεικόνισης συνοψίζεται ως εξής:

1. Αρχικά επιλέγονται τυχαίες τιμές για τα διανύσματα συναπτικών βαρών $\mathbf{w}_j(0)$.

Ο μόνος περιορισμός είναι αυτές οι τιμές πρέπει να διαφέρουν μεταξύ τους για κάθε νευρώνιο $j = 1, 2, \dots, l$ όπου l , ο αριθμός των νευρώνων στο πλέγμα. Καλό είναι οι τιμές να είναι μικρές ώστε να μην κορεστεί το δίκτυο από τους υπολογισμούς. Ένας τρόπος για την τοποθέτηση αρχικών τιμών στα συναπτικά βάρη $\{\mathbf{w}_j(0)\}_{j=1}^l$ είναι να επιλεγθούν τυχαία δείγματα από το εύρος τιμών των διανυσμάτων εισόδου $\{\mathbf{x}_i\}_{i=1}^N$.

2. Υπολογίζεται το νευρώνιο νικητής $i(\mathbf{x})$ σε κάθε βήμα n με κριτήριο την ελάχιστη Ευκλείδεια απόσταση:

$$i(\mathbf{x}) = \arg \min_j \|\mathbf{x}(n) - \mathbf{w}_j\|, j = 1, 2, \dots, l$$

3. Τα συναπτικά βάρη όλων των νευρώνων ανανεώνονται χρησιμοποιώντας την εξίσωση:

$$\mathbf{w}_j(n+1) = \mathbf{w}_j(n) + \eta(n)h_{j,i(\mathbf{x})}(n)(\mathbf{x}(n) - \mathbf{w}_j(n)),$$

όπου $\eta(n)$ η παράμετρος ρυθμού εκμάθησης και $h_{j,i(\mathbf{x})}$ η τοπολογική γειτονιά γύρω από το νευρώνιο νικητή $i(\mathbf{x})$. Και οι δυο παραπάνω παράμετροι μεταβάλλονται κατά τη διαδικασία της μάθησης για καλύτερα αποτελέσματα.

4. Η διαδικασία επαναλαμβάνεται από το βήμα 2 ωσότου δεν παρατηρούνται σημαντικές διαφοροποιήσεις στον χάρτη.