

ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΟΡΥΚΤΩΝ ΠΟΡΩΝ



Πρόγραμμα Μεταπτυχιακών Σπουδών
«Γεωτεχνολογία & Περιβάλλον»

**«Εφαρμογή της μεθόδου των Σπαρτιάτικων Τυχαίων Πεδίων
στην γεωστατιστική ανάλυση της χωρικής κατανομής
περιβαλλοντικών ρύπων»**

Μεταπτυχιακή Διατριβή

Εμμανουήλ Α. Βαρουχάκης

Εξεταστική Επιτροπή

Δ. Χριστόπουλος, Αν. Καθηγητής (επιβλέπων)

Σ. Μερτίκας, Καθηγητής

Κ. Κομνίτσας, Αν. Καθηγητής

Χανιά
Φεβρουάριος 2005

Περίληψη

Η συγκεκριμένη μεταπτυχιακή εργασία ασχολείται με την χωρική εκτίμηση φυσικών μεγεθών (συγκεντρώσεις χρωμίου) σε σημεία μιας περιοχής ενδιαφέροντος όπου δεν υπάρχουν μετρήσεις. Από την μαθηματική πλευρά, ο γενικός στόχος είναι η αναπαραγωγή της χωρικής κατανομής της περιβαλλοντικής ρύπανσης στα σημεία ενός καννάβου, βάσει των μετρημένων τιμών στα ανομοιογενώς κατανεμημένα σημεία δειγματοληψίας. Από την σκοπιά των εφαρμογών, ο γενικός στόχος είναι η ανάλυση της περιβαλλοντικής επικινδυνότητας που προκύπτει από ένα περιορισμένο πλήθος μετρήσεων περιβαλλοντικών ρυπαντών. Η χωρική εκτίμηση είναι αποτέλεσμα των μεθόδων γεωστατιστικής ανάλυσης οι οποίες βασίζονται στην χωρική εξάρτηση των τιμών ενός δείγματος. Οι πιο διαδεδομένες είναι γνωστές με την ονομασία *Kriging*. Οι μέθοδοι αυτές χαρακτηρίζονται από ισχυρούς μαθηματικούς περιορισμούς ως προς την εφαρμογή τους. Όταν οι περιορισμοί αυτοί δεν ικανοποιούνται επηρεάζεται αρνητικά η ακρίβεια της χωρικής εκτίμησης. Επομένως, υπάρχει ανάγκη για την ανάπτυξη νέων, πιο ευέλικτων μεθόδων χωρικής εκτίμησης.

Μια καινοτόμος μέθοδος γεωστατιστικής ανάλυσης η οποία ονομάζεται μοντέλα Σπαρτιάτικων Τυχαίων Πεδίων παρέχει ένα διαφορετικό τρόπο ανάλυσης των χωρικών δεδομένων σε σχέση με τα κλασικά γεωστατιστικά πρότυπα. Ο όρος Σπαρτιάτικα Τυχαία Πεδία (ΣΤΠ) χαρακτηρίζει γεωστατιστικά μοντέλα που προσδιορίζονται από ένα σχετικά μικρό αριθμό παραμέτρων, οι οποίες δεν προϋποθέτουν τον υπολογισμό του ημιβαριογράμματος. Σε αντίθεση με το ημιβαριόγραμμα, ο προσδιορισμός των παραμέτρων των ΣΤΠ βασίζεται σε σχετικά απλούς στατιστικούς περιορισμούς, που υπολογίζονται αποτελεσματικά από το υπάρχον δείγμα.

Η μέθοδος αυτή εφαρμόζεται για πρώτη φορά σε πραγματικό δείγμα το οποίο αποτελείται από μετρήσεις συγκέντρωσης χρωμίου στο έδαφος σε 359 σημεία μέτρησης τα οποία βρίσκονται στην περιοχή των βουνών Jura (Ελβετία). Δίνεται έμφαση στον προσδιορισμό των παραμέτρων του μοντέλου βάσει των στατιστικών περιορισμών του δείγματος. Εξετάζονται δυο μέθοδοι για την εκτίμηση των

παραμέτρων: η πρώτη βασίζεται στον ορισμό ενός αυθαίρετου κανονικού καννάβου που καλύπτει την περιοχή της μελέτης, ενώ η δεύτερη χρησιμοποιεί τα τρίγωνα Delaunay. Συνάγεται από την ανάλυση ότι η μέθοδος των τριγώνων Delaunay χωρίς εξομάλυνση των δεδομένων δεν αποδίδει ικανοποιητικά.

Στην συνέχεια εξετάζεται η ακρίβεια της χωρικής εκτίμησης της μεθόδου ΣΤΠ, συγκρίνοντας τις εκτιμήσεις του μοντέλου με τις πραγματικές τιμές σε ένα σύνολο σημείων επιβεβαίωσης (validation points). Εξετάζεται η ακρίβεια της χωρικής εκτίμησης της μεθόδου ΣΤΠ και πραγματοποιείται συγκριτική ανάλυση με την μέθοδο *Kriging*.

Συγκρίνοντας τα αποτελέσματα των δύο μεθόδων διαπιστώνεται ότι η επίδοσή τους στην ανάλυση του συγκεκριμένου δείγματος είναι ανάλογη. Ως γενικό συμπέρασμα, τα αποτελέσματα της εφαρμογής της μεθόδου ΣΤΠ είναι ενθαρρυντικά, καθώς υπάρχουν πολλά περιθώρια βελτίωσης τόσο της εκτίμησης των παραμέτρων του μοντέλου όσο και της χωρικής εκτίμησης.

Abstract

Spartan Spatial Random Fields (SSRF's) are a recently proposed geostatistical model (Hristopulos, 2003) with applications in environmental risk assessment and natural resources estimation. The SSRF models are determined by a small number of parameters, the inference of which from the sample does not require the variogram calculation. The SSRF parameter inference is based on relatively simple statistical restrictions that are calculated effectively.

This study presents an application of a specific SSRF model in the estimation of environmental pollutants, at points where their concentration is unknown, using a real data set. The study focuses on practical methodological issues, such as the inference of the model parameters from points distributed on an irregular sampling grid, and the determination of an optimal correlation neighborhood for the SSRF estimator. The estimation error of the method is discussed and a detailed comparison of the SSRF model with a classical *kriging* method is applied. It is concluded that, at least within the scope of the present study, the SSRF framework provides a competitive alternative for spatial estimation.

Πρόλογος

Η εκπόνηση της συγκεκριμένης μεταπτυχιακής διατριβής δεν θα ήταν δυνατή χωρίς την καθοδήγηση και την επίβλεψη του Αναπληρωτή Καθηγητή κ. Διονύση Χριστόπουλου, τον οποίο και θα ήθελα να ευχαριστήσω ιδιαίτερα. Επίσης ευχαριστίες οφείλω στον Καθηγητή κ. Στυλιανό Μερτίκα και στον Αναπληρωτή Καθηγητή κ. Κωνσταντίνο Κομνίτσα για τις πολύτιμες συμβουλές και παρατηρήσεις τους. Μέρος της μεταπτυχιακής εργασίας χρηματοδοτήθηκε από τον ΕΛΚΕ Πολυτεχνείου Κρήτης στα πλαίσια του προγράμματος «Σπαρτιάτικα μοντέλα τυχαίων πεδίων για Γεωστατιστικές εφαρμογές» για την ενίσχυση της βασικής έρευνας.

Περιεχόμενα

Περίληψη	i
Abstract	iii
Πρόλογος	iv
Περιεχόμενα	v
Κατάλογος Σχημάτων	viii
Κατάλογος Πινάκων	xi
1. Εισαγωγή	1
1.1 Γεωστατιστική	1
1.2 Τυχαία πεδία	3
1.3 Βασικές έννοιες τυχαίων πεδίων	5
1.3.1 Μέση τιμή	6
1.3.2 Διασπορά	8
1.3.3 Συνάρτηση συνδιασποράς	9
1.3.4 Στατιστική ομοιογένεια	11
1.3.5 Στατιστική ισοτροπία	12
1.3.6 Χωρική εξάρτηση	12
1.3.7 Προσδιορισμός ημιβαριογράμματος	17
1.3.8 Χωρική εκτίμηση	21
1.4 Περιορισμοί κλασικών γεωστατιστικών μεθόδων	22
1.5 Γεωστατιστική ανάλυση τυχαίου πεδίου	25
2. Σπαρτιάτικα τυχαία πεδία	27
2.1 Εισαγωγή	27
2.2 Περιγραφή του μοντέλου των ΣΤΠ	28
2.3 Εκτίμηση παραμέτρων μοντέλου	31
2.3.1 Μέθοδος κυψελίδων σε κανονικό πλέγμα	32
2.3.2 Μέθοδος κυψελίδων σε ανομοιογενές πλέγμα	36
2.3.3 Μέθοδος τριγώνων Delaunay	40
2.3.4 Επιβεβαίωση της μεθόδου Delaunay	45
2.4 Εκτίμηση πεδίου σε σημεία χωρίς μετρήσεις	50

3. Εφαρμογή στην χαρτογράφηση περιεκτικότητας Cr στο έδαφος της περιοχής ενδιαφέροντος.....	55
3.1 Περιγραφή της γεωγραφικής περιοχής.....	55
3.2 Διερευνητική στατιστική ανάλυση.....	57
3.3 Ανάλυση χωρικής συνέχειας.....	61
3.3.1 Προσδιορισμός χωρικής τάσης.....	61
3.3.2 Προσδιορισμός ημιβαριογράμματος δείγματος χρωμίου.....	63
3.4 Μεθοδολογία γεωστατιστικής ανάλυσης.....	71
3.5 Γεωστατιστική ανάλυση με την μέθοδο Kriging.....	73
3.6 Εφαρμογή της μεθόδου κανονικού Kriging.....	80
3.7 Εφαρμογή μοντέλου Σπαρτιάτικων Τυχαίων Πεδίων.....	82
3.7.1 Εφαρμογή της μεθόδου Delaunay.....	88
3.7.2 Εφαρμογή ΣΤΠ με την μέθοδο της διασταυρωμένης επιβεβαίωσης.....	89
3.8 Σύγκριση μεθόδου ΣΤΠ με την μέθοδο Kriging.....	93
4. Συμπεράσματα.....	97
4.1 Προτάσεις για μελλοντική έρευνα.....	98
Ελληνική βιβλιογραφία.....	100
Διεθνής βιβλιογραφία.....	100
Δια-δικτυακές πηγές.....	101
Παραρτήματα.....	102
Παράρτημα Α «Προσδιορισμός αναλυτικών λύσεων στατιστικών περιορισμών για περιοδική συνάρτηση δείγματος».....	103
Παράρτημα Β «Υπολογισμός ισοτροπικής λύσης».....	109
Προγράμματα MATLAB.....	111
Παράρτημα Γ «Υπολογισμός στατιστικών περιορισμών με την μέθοδο Delaunay».....	112
Παράρτημα Δ «Δημιουργία τυχαίου πεδίου δύο διαστάσεων».....	113
Παράρτημα Ε «Υπολογισμός στατιστικών περιορισμών δείγματος (σύνολο εκπαίδευσης)».....	114
Παράρτημα Ζ «Προσδιορισμός συναρτησιακής απόστασης».....	115
Παράρτημα Η «Υπολογισμός παραμέτρων δείγματος (σύνολο εκπαίδευσης)».....	116

Παράρτημα Θ «Εκτίμηση στα σημεία του συνόλου επιβεβαίωσης»	118
Παράρτημα Ι «Εκτίμηση στα σημεία του συνόλου επιβεβαίωσης μετά την αφαίρεση των ακραίων τιμών»	121
Παράρτημα Κ «Υπολογισμός στατιστικών περιορισμών δείγματος».....	123
Παράρτημα Λ «Υπολογισμός παραμέτρων δείγματος»	124
Παράρτημα Μ «Εκτίμηση στα σημεία του δείγματος με την μέθοδο της διασταυρωμένης επιβεβαίωσης».....	125
Παράρτημα Ν «Εκτίμηση στα σημεία της περιοχής μελέτης όπου δεν υπάρχουν μετρήσεις»	128

Κατάλογος Σχημάτων

Σχήμα 1.1: Γράφημα τάσης τυχαίας μεταβλητής.	7
Σχήμα 1.2: Γράφημα συνημίτονου με εκθετικά μειούμενο πλάτος διακύμανσης (Χριστόπουλος 2004α).	9
Σχήμα 1.3: Σύστημα κύριων αξόνων (KA1, KA2) σε σχέση με το σύστημα συντεταγμένων x και y και ισοϋψής καμπύλη (έλλειψη) του ημιβαριογράμματος (Χριστόπουλος 2004β).	15
Σχήμα 1.4: Σχηματικό διάγραμμα μιας περιοχής $B(\mathbf{r})$ γύρω από το διάνυσμα απόστασης (Χριστόπουλος 2004β).....	18
Σχήμα 1.5: Παρουσίαση χαρακτηριστικών στοιχείων ημιβαριογράμματος (Surfer V.8.0.4, 2002).	20
Σχήμα 1.6: Ημιβαριόγραμμα της κατανομής περιεκτικότητας χρωμίου στα Ελβετικά βουνά Jura ως συνάρτηση της απόστασης μεταξύ δύο σημείων μέτρησης.....	24
Σχήμα 2.1: Απεικόνιση κανονικού πλέγματος (τετραγωνικό 10x10) το οποίο περιλαμβάνει τιμές ενός δείγματος κατανεμημένες πάνω στα σημεία του.	30
Σχήμα 2.2: Απεικόνιση τετραγωνικού πλέγματος 4x4 που καλύπτει μικρό δείγμα (Hristopulos 2003).	37
Σχήμα 2.3: Δειγματικά σημεία τυχαίου πεδίου (Mulchrone 2002).	41
Σχήμα 2.4: Διάγραμμα πολυγώνων Voronoi (Mulchrone 2002).....	42
Σχήμα 2.5: Παράδειγμα δημιουργίας τριγώνων Delaunay (Mulchrone 2002).....	43
Σχήμα 2.6: Εφαρμογή της μεθόδου τριγώνων Delaunay στα σημεία μέτρησης του δείγματος χρωμίου (Cr).	49
Σχήμα 3.1: Χωρική κατανομή συγκέντρωσης χρωμίου (Goovaerts 1997).....	56
Σχήμα 3.2: Διάγραμμα πιθανότητας κανονικής κατανομής δείγματος χρωμίου.	58
Σχήμα 3.3: Ιστόγραμμα κανονικής κατανομής Cr χρησιμοποιώντας την εντολή «hist» της Matlab®. Το ιστόγραμμα περιγράφει την συχνότητα εμφάνισης των μετρήσεων του δείγματος στα αντίστοιχα διαστήματα συγκέντρωσης χρωμίου.	60
Σχήμα 3.4: Ιστόγραμμα κανονικής κατανομής Cr χρησιμοποιώντας την εντολή «histfit» της Matlab®. Το ιστόγραμμα περιγράφει την συχνότητα εμφάνισης των μετρήσεων του δείγματος στα αντίστοιχα διαστήματα συγκέντρωσης χρωμίου.	61

Σχήμα 3.5: Η τάση του δείγματος παρουσιάζεται από την γραμμή (-) κόκκινου χρώματος ενώ οι τιμές του δείγματος από την γραμμή (-) μπλε χρώματος.	62
Σχήμα 3.6: Παράδειγμα απεικόνισης ανισοτροπίας (Surfer V.8.0.4, 2002).....	63
Σχήμα 3.7: Πειραματικό ημιβαριόγραμμα δείγματος Cr σαν συνάρτηση της απόστασης.....	65
Σχήμα 3.8: Απεικόνιση ανισοτροπίας στο δείγμα χρωμίου.....	67
Σχήμα 3.9: Προσαρμογή πειραματικού ημιβαριογράμματος Cr στο Εκθετικό με Φαινόμενο Πυρήνα μοντέλο ημιβαριογράμματος, κατά την κατεύθυνση του άξονα (y).....	68
Σχήμα 3.10: Προσαρμογή πειραματικού ημιβαριογράμματος Cr στο Εκθετικό με Φαινόμενο Πυρήνα μοντέλο ημιβαριογράμματος, κατά την κατεύθυνση του άξονα (x).....	68
Σχήμα 3.11: Προσαρμογή πειραματικού ημιβαριογράμματος Cr στο μοντέλο ημιβαριογράμματος, Φαινόμενου πυρήνα με Περιοχή Αντισυσχέτισης.	69
Σχήμα 3.12: Προσαρμογή πειραματικού ημιβαριογράμματος Cr στο Τετραγωνικό μοντέλο ημιβαριογράμματος	69
Σχήμα 3.13: Προσαρμογή πειραματικού ημιβαριογράμματος Cr στο Σφαιρικό με Φαινόμενο Πυρήνα μοντέλο ημιβαριογράμματος.	70
Σχήμα 3.14: Προσαρμογή πειραματικού ημιβαριογράμματος Cr στο Σφαιρικό μοντέλο ημιβαριογράμματος.	70
Σχήμα 3.15: Χάρτης σημείων όπου υπάρχουν μετρήσεις χρωμίου στα σημεία του συνόλου εκπαίδευσης και επιβεβαίωσης στο χώρο. Με κόκκινο κύκλο (ο) παρουσιάζονται τα σημεία του συνόλου επιβεβαίωσης ενώ με μαύρο αστεράκι (*) τα σημεία του συνόλου εκπαίδευσης.....	72
Σχήμα 3.16: Διάγραμμα διασποράς εκτιμώμενων και πραγματικών συγκεντρώσεων χρωμίου. Με κόκκινο (ο) δηλώνεται το κάθε σημείο του συνόλου επιβεβαίωσης το οποίο αντιστοιχίζει την πραγματική με την εκτιμώμενη συγκέντρωση χρωμίου. Ο άξονας (x) παρουσιάζει τις εκτιμώμενες συγκεντρώσεις χρωμίου ενώ ο άξονας (y) τις πραγματικές συγκεντρώσεις.	86
Σχήμα 3.17: Διάγραμμα διασποράς εκτιμώμενων και πραγματικών συγκεντρώσεων χρωμίου. Με κόκκινο (ο) δηλώνεται το κάθε σημείο του συνόλου επιβεβαίωσης το οποίο αντιστοιχίζει την πραγματική με την εκτιμώμενη συγκέντρωση χρωμίου. Ο άξονας (x) παρουσιάζει τις εκτιμώμενες συγκεντρώσεις χρωμίου ενώ ο άξονας (y) τις πραγματικές συγκεντρώσεις.	87
Σχήμα 3.18: Χάρτης περιεκτικότητας χρωμίου (σε ppm) χρησιμοποιώντας το Σφαιρικό με Φαινόμενο Πυρήνα μοντέλο ημιβαριογράμματος.	95

Σχήμα 3.19: Χάρτης περιεκτικότητας χρωμίου (σε ppm) χρησιμοποιώντας το μοντέλο Σπαρτιάτικων Τυχαίων Πεδίων.	96
--	----

Κατάλογος Πινάκων

Πίνακας 2.1: Αποτελέσματα επιβεβαίωσης της μεθόδου Delaunay για δείγμα τυχαίου πεδίου περιοδικής συνάρτησης.....	47
Πίνακας 2.2: Αποτελέσματα επιβεβαίωσης της μεθόδου Delaunay για το πρώτο ΣΤΧ.....	47
Πίνακας 2.3: Αποτελέσματα επιβεβαίωσης της μεθόδου Delaunay για το δεύτερο ΣΤΧ.....	48
Πίνακας 2.4: Αποτελέσματα επιβεβαίωσης της μεθόδου Delaunay για δείγμα τυχαίων πεδίων κατανεμημένων σε ανομοιογενές πλέγμα.....	49
Πίνακας 3.1: Επιτρεπτή συγκέντρωση χρωμίου σε έδαφος της Ολλανδίας (EC 2003).	56
Πίνακας 3.2: Στατιστικές ιδιότητες δείγματος χρωμίου.....	60
Πίνακας 3.3: Τάξεις απόστασης, πλήθος ζευγών ανά τάξη και τιμές πειραματικού ημιβαριογράμματος για την κατανομή Cr.	64
Πίνακας 3.4: Τιμές παραμέτρων των θεωρητικών μοντέλων ημιβαριογράμματος.	66
Πίνακας 3.5: Σφάλμα εκτίμησης μεθόδου κανονικού <i>Kriging</i> για διαφορετικά θεωρητικά μοντέλα ημιβαριογράμματος.	81
Πίνακας 3.6: Σφάλμα εκτίμησης μεθόδου κανονικού <i>Kriging</i> εξαιρουμένων την μέγιστης και της ελάχιστης συγκέντρωσης του δείγματος για διαφορετικά θεωρητικά μοντέλα ημιβαριογράμματος.	81
Πίνακας 3.7: Βέλτιστες τιμές παραμέτρων για διαφορετικούς συνδυασμούς αρχικών εκτιμήσεων.....	84
Πίνακας 3.8: Ανάλυση σφάλματος εκτίμησης ΣΤΠ σε 100 σημεία του συνόλου επιβεβαίωσης.	86
Πίνακας 3.9: Μέτρηση σφάλματος εκτίμησης σε 98 σημεία του συνόλου επιβεβαίωσης.	88
Πίνακας 3.10: Σφάλμα εκτίμησης μοντέλου Σπαρτιάτικων Τυχαίων Πεδίων.....	91
Πίνακας 3.11: Σφάλμα εκτίμησης μοντέλου Σπαρτιάτικων Τυχαίων Πεδίων.....	91
Πίνακας 3.12: Σφάλμα εκτίμησης μοντέλου Σπαρτιάτικων Τυχαίων Πεδίων εξαιρουμένων την μέγιστης και της ελάχιστης συγκέντρωσης του δείγματος.....	93

Πίνακας 3.13: Σφάλμα εκτίμησης μοντέλου Σπαρτιάτικων Τυχαίων Πεδίων εξαιρουμένων την μέγιστης και της ελάχιστης συγκέντρωσης του δείγματος.....	93
--	----

1. Εισαγωγή

1.1 Γεωστατιστική

Η επιστήμη της Γεωστατιστικής έχει γνωρίσει ιδιαίτερη ανάπτυξη τις τελευταίες δύο δεκαετίες. Το πεδίο εφαρμογών της έχει διευρυνθεί και περιλαμβάνει εφαρμογές που έχουν άμεση σχέση με τις ερευνητικές δραστηριότητες ανίχνευσης και εντοπισμού ορυκτών πόρων και περιβαλλοντικών ρύπων. Η γεωστατιστική ανάλυση ασχολείται με κατανομές στις οποίες τα χαρακτηριστικά της χωρικής εξάρτησης παίζουν πρωτεύοντα ρόλο. Οι μέθοδοι γεωστατιστικής ανάλυσης οδηγούν στον προσδιορισμό της χωρικής κατανομής μεταβλητών σε σημεία μιας περιοχής όπου δεν είναι γνωστές. Αυτές οι μεταβλητές χαρακτηρίζουν μεγέθη με οικονομική ή περιβαλλοντική σημασία. (Χριστόπουλος 2004β).

Οι κλασικές γεωστατιστικές μέθοδοι οι οποίες χρησιμοποιούνται για το σκοπό αυτό χαρακτηρίζονται από ισχυρούς μαθηματικούς περιορισμούς ως προς την εφαρμογή τους. Αυτό έχει ως αποτέλεσμα, όταν οι περιορισμοί αυτοί δεν ικανοποιούνται από το πειραματικό δείγμα, του υπό μελέτη φαινομένου, η χωρική εκτίμηση των μεταβλητών να περιλαμβάνει υψηλό ποσοστό σφάλματος (Hristopulos 2003).

Η Γεωστατιστική περιλαμβάνει ένα σύνολο στατιστικών τεχνικών που αφορούν τυχαίες μεταβλητές οι οποίες μεταβάλλονται στο χώρο (τυχαία πεδία). Οι τεχνικές αυτές βασίζονται στην υπόθεση ότι η χωρική διακύμανση της μεταβλητής εμπεριέχει τυχαίο χαρακτήρα, οπότε χρησιμοποιούν στατιστικές μεθοδολογίες (π.χ μέση τιμή, διασπορά κ.α.) για οποιαδήποτε εκτίμηση απορρέει από τις σημειακές μετρήσεις της μεταβλητής (Χριστόπουλος 2004β).

Πιο αναλυτικά, η Γεωστατιστική στηρίζεται στην μαθηματική έννοια του τυχαίου πεδίου. Η Γεωστατιστική έχει κοινά σημεία με τη Θεωρία Πιθανοτήτων και τη Στατιστική. Η Θεωρία Πιθανοτήτων ασχολείται με τους νόμους και τις ιδιότητες που διέπουν τις τυχαίες μεταβλητές. Η Στατιστική περιλαμβάνει το σύνολο των μεθόδων οι οποίες επιτρέπουν τον προσδιορισμό των παραμέτρων που χαρακτηρίζουν τις τυχαίες μεταβολές βάσει των δεδομένων. Η Θεωρία των Τυχαίων Πεδίων αποτελεί μια

γενίκευση της Θεωρίας Πιθανοτήτων που εφαρμόζεται σε τυχαίες μεταβλητές με χωρική εξάρτηση (Χριστόπουλος 2004β).

Η Γεωστατιστική αποσκοπεί στην εκτίμηση των στατιστικών παραμέτρων που προσδιορίζουν τη χωρική κατανομή βάσει του υπάρχοντος δείγματος (π.χ. τιμών συγκέντρωσης), καθώς και στη χρήση αυτών των παραμέτρων προκειμένου να εκτιμηθούν οι συγκεντρώσεις σε σημεία όπου δεν υπάρχουν μετρήσεις (Χριστόπουλος 2004b).

Η ανάγκη της εκτίμησης φυσικών μεγεθών σε σημεία όπου δεν υπάρχουν μετρήσεις, δεν είναι καινούρια. Στατιστικοί επιστήμονες, μηχανικοί μεταλλείων, μηχανικοί πετρελαίων, υδρολόγοι και γεωλόγοι οι οποίοι ασχολήθηκαν με το πρόβλημα ανέπτυξαν την επιστήμη της Γεωστατιστικής. Αρχικά η Γεωστατιστική εφαρμόστηκε κυρίως στην γεωλογία, στην μεταλλειολογία και στην υδρολογία. Στην συνέχεια βρήκε εφαρμογές και σε άλλα επιστημονικά και τεχνολογικά πεδία (Myers 2005). Μερικές από τις εφαρμογές στις οποίες χρησιμοποιείται σήμερα η γεωστατιστική περιλαμβάνουν (Χριστόπουλος 2004β):

- Την έρευνα κοιτασμάτων (π.χ. εκτίμηση έκτασης, βάθος και ποσοτικοποίηση συνολικής περιεκτικότητας κοιτάσματος).
- Την ωκεανογραφία (π.χ. χαρτογράφηση βυθού, ανάλυση κυματισμών).
- Την μορφολογική ανάλυση φυσικών και τεχνολογικών ανομοιογενών (π.χ. πορωδών) υλικών.
- Την χαρτογράφηση και την απεικόνιση συγκεντρώσεων ρυπαντών σε διάφορα περιβαλλοντικά μέσα (αέρας, υπέδαφος, επιφανειακοί-υπόγειοι υδατικοί πόροι).
- Το χαρακτηρισμό της ποιότητας βιομηχανικών προϊόντων (π.χ., προϊόντα χαρτιού, υλικά υψηλής τεχνολογίας όπως ημιαγωγοί).

- Την τοπογραφική ανάλυση και στα γεωγραφικά συστήματα πληροφορίας (GIS).
- Την ανάλυση βροχοπτώσεων σε περιοχές που υπάρχουν λίγοι βροχομετρικοί σταθμοί.
- Το προσδιορισμό γεωλογικών και υδρο-γεωλογικών δεδομένων (π.χ. τύπος υπεδάφους, υδραυλική αγωγιμότητα, πορώδες, αποθηκευτικότητα, εξατμισο-διαπνοή, υδραυλικό ύψος).
- Την εκτίμηση του περιβαλλοντικού κινδύνου και του κινδύνου για την ανθρώπινη υγεία (π.χ. εκτίμηση της περιεκτικότητας των ρυπαντών σε ελεγχόμενη περιοχή, προσδιορισμός των πιθανοτήτων υπέρβασης κρίσιμων ορίων)

Ο κοινός στόχος σε όλες τις εφαρμογές της Γεωστατιστικής είναι ο προσδιορισμός και ο έλεγχος της χωρικής κατανομής μεταβλητών που χαρακτηρίζουν μεγέθη με οικονομική ή περιβαλλοντική σημασία.

Οι πιο δημοφιλείς και διαδεδομένες μέθοδοι γεωστατιστικής ανάλυσης είναι γνωστές με την ονομασία *Kriging*. Ωστόσο οι μέθοδοι αυτές όπως προαναφέρθηκε χαρακτηρίζονται από μαθηματικούς περιορισμούς οι οποίοι επηρεάζουν την εφαρμογή τους. Αυτοί οι περιορισμοί αφορούν τον προσδιορισμό της χωρικής εξάρτησης των τυχαίων μεταβλητών (1.4). Για αυτό είναι επιθυμητή η ανάπτυξη νέων μεθόδων γεωστατιστικής ανάλυσης ώστε να υπάρχουν εναλλακτικές προτάσεις.

1.2 Τυχαία πεδία

Ως «τυχαίο πεδίο» μπορεί να θεωρηθεί ένα σύνολο τυχαίων μεταβλητών που περιγράφουν τη χωροχρονική μεταβολή του σχετικού φυσικού μεγέθους (π.χ τιμές συγκεντρώσεων ρύπων). Σε αντίθεση με τις συναρτήσεις οι οποίες έχουν μια συγκεκριμένη μαθηματική έκφραση, π.χ $f(x) = \cos(x)$, ένα τυχαίο πεδίο που δεν έχει σαφή μαθηματική έκφραση αντιπροσωπεύει ένα σύνολο δυνατών καταστάσεων.

Κάθε κατάσταση αποτελεί ένα **δείγμα** του πεδίου και χαρακτηρίζεται από μια πιθανότητα πραγματοποίησης που καθορίζεται από την πολυδιάστατη Συνάρτηση Πυκνότητας Πιθανότητας του πεδίου. Επομένως, ένα τυχαίο πεδίο μπορεί να θεωρηθεί ως μία πολυδιάστατη τυχαία μεταβλητή. Λόγω της αλληλεξάρτησης των φυσικών μεγεθών σε διαφορετικά σημεία του χώρου, τα τυχαία πεδία έχουν ιδιαίτερες μαθηματικές ιδιότητες που τα ξεχωρίζουν από ένα σύνολο ανεξάρτητων τυχαίων μεταβλητών (Χριστόπουλος 2004α).

Υπάρχουν διάφορες κατηγορίες τυχαίων πεδίων. Αν το πεδίο λαμβάνει τιμές μόνο από ένα μετρήσιμο σύνολο αριθμών ονομάζεται πεδίο διακριτών τιμών. Αν οι τιμές του πεδίου προέρχονται από ένα συνεχές διάστημα πραγματικών αριθμών, το πεδίο ονομάζεται πεδίο συνεχών τιμών. Όταν η μεταβολή ορίζεται σε ένα συνεχή χώρο, όπως π.χ. στα φυσικά πεδία, δημιουργείται ένα πεδίο συνεχούς χώρου. Αντίθετα όταν ορίζεται στις θέσεις ενός πλέγματος (καννάβου) ονομάζεται πλεγματικό πεδίο.

Τα πλεγματικά πεδία χρησιμοποιούνται σε υπολογιστικές (π.χ. προσομοίωση κατανομής ρυπαντών σε υδροφόρο ορίζοντα) αλλά και σε θεωρητικές μελέτες γιατί η συμμετρία του πλέγματος επιτρέπει την χρησιμοποίηση αριθμητικά αποτελεσματικών μεθόδων (π.χ. ταχείς μετασχηματισμούς Fourier). Επιπλέον τα πλεγματικά πεδία επιτρέπουν την σύγκριση των επιδόσεων διαφορετικών γεωστατιστικών μεθόδων (Χριστόπουλος 2004α).

Στην πράξη οι μετρήσεις αντιπροσωπεύουν ένα πεπερασμένο πλήθος σημείων, η κατανομή των οποίων στο χώρο δεν έχει κατ' ανάγκη τη συμμετρία ενός κανονικού πλέγματος. Σε αυτές τις περιπτώσεις το δίκτυο των σημείων δειγματοληψίας είναι ανομοιογενές. Επίσης χρησιμοποιείται ο όρος άτακτο πλέγμα (disordered lattice) και ο όρος εξωπλεγματική κατανομή (off lattice). Σε αυτές τις περιπτώσεις, χρειάζονται γεωστατιστικές μέθοδοι που να λειτουργούν ικανοποιητικά με τους περιορισμούς της εκάστοτε χωρικής κατανομής. Αν η κατανομή των δεδομένων είναι εξωπλεγματική, η εκτίμηση ή η προσομοίωση της διαδικασίας πραγματοποιείται πάνω σε ένα πλεγματικό υπόβαθρο που καλύπτει την περιοχή ενδιαφέροντος (Χριστόπουλος 2004α).

Η έννοια του τυχαίου πεδίου βασίζεται σε δύο άλλες βασικές έννοιες: την *τυχειότητα* και την *αλληλεξάρτηση* των τιμών φυσικών μεγεθών σε διαφορετικά σημεία του χώρου. Η τυχειότητα χαρακτηρίζει φαινόμενα στα οποία η γνώση μιας κατάστασης με απόλυτη ακρίβεια είναι αδύνατη λόγω διαφόρων περιορισμών. Τέτοιοι περιορισμοί προέρχονται από την μεταβλητότητα των διαφόρων φυσικών μεγεθών στον χώρο και την αβεβαιότητα που οφείλεται στον περιορισμένο αριθμό μετρήσεων. Σε αυτές τις περιπτώσεις το αποτέλεσμα (η τιμή του φαινομένου) καθορίζεται από μία συνάρτηση κατανομής πιθανότητας, η οποία προσδιορίζει τις πιθανότητες εμφάνισης κάθε κατάστασης.

Η χωρική εξάρτηση (αλληλεξάρτηση) είναι το ιδιαίτερο γνώρισμα των τυχαίων πεδίων και περιγράφει την εξάρτηση των τιμών του πεδίου σε διαφορετικά σημεία του χώρου μεταξύ τους. Η κατανομή πιθανότητας του πεδίου εμπερικλείει συσχετίσεις μεταξύ διαφορετικών σημείων, έτσι ώστε η πιθανότητα παρατήρησης μιας τιμής σε ένα σημείο να εξαρτάται από τις τιμές στα γειτονικά σημεία (Χριστόπουλος 2004α).

1.3 Βασικές έννοιες τυχαίων πεδίων

Ένα τυχαίο πεδίο συμβολίζεται ως $X(\mathbf{s})$, όπου \mathbf{s} ένα διάνυσμα θέσης, $\mathbf{s} = (x, y)$. Το $X(\mathbf{s})$ συμβολίζει το σύνολο των δυνατών καταστάσεων του πεδίου, ενώ ως $x(\mathbf{s})$ συμβολίζονται οι τιμές που αντιστοιχούν σε μια κατάσταση. Η Συνάρτηση Πυκνότητας Πιθανότητας (ΣΠΠ) του πεδίου παριστάνεται ως $f_X[x(\mathbf{s})]$. Ο δείκτης X δηλώνει το πεδίο, ενώ το όρισμα της συνάρτησης είναι οι τιμές της κατάστασης του πεδίου (π.χ. συγκεντρώσεις ρυπαντών).

Ένα παράδειγμα ΣΠΠ η οποία αντιστοιχεί σε τυχαίο πεδίο κανονικής κατανομής δίνεται

$$\text{από την εξίσωση: } f_X[x(\mathbf{s})] = \frac{1}{\sqrt{2\pi}\sigma_X} \exp\left[-\frac{(x(\mathbf{s}) - m_X(\mathbf{s}))^2}{2\sigma_X^2}\right].$$

Η Συνάρτηση Πυκνότητας Πιθανότητας τυχαίου πεδίου περιλαμβάνει τις τιμές σε ολόκληρο το χώρο όπου ορίζεται το πεδίο. Επομένως αποτελεί την κοινή ΣΠΠ για

ένα οποιοδήποτε πλήθος σημείων. Η μονοδιάστατη η αλλιώς σημειακή ΣΠΠ περιγράφει τις δυνατές καταστάσεις του πεδίου σε ένα συγκεκριμένο σημείο. Είναι δυνατόν η μονοδιάστατη ΣΠΠ να αλλάζει από σημείο σε σημείο και αυτό συμβαίνει όταν το πεδίο είναι ανομοιογενές. Ανάλογα, η δισδιάστατη ΣΠΠ του πεδίου εκφράζει την αλληλεξάρτηση των δυνατών καταστάσεων σε δύο σημεία και η πολυδιάστατη περιγράφει την αλληλεξάρτηση των δυνατών καταστάσεων για ένα σύνολο N σημείων (Χριστόπουλος 2004α).

Ένα άλλο είδος συναρτήσεων το οποίο δίνει πληροφορίες σχετικά με τις ιδιότητες ενός τυχαίου πεδίου είναι οι **στατιστικές ροπές**. Οι στατιστικές ροπές είναι αιτιοκρατικές συναρτήσεις, οι οποίες αντιπροσωπεύουν μέσες τιμές, ως προς όλες τις δυνατές καταστάσεις. Στην πράξη συνήθως είναι χρήσιμες οι ροπές χαμηλών τάξεων (μέχρι δεύτερης τάξης), όπως η μέση τιμή, η διασπορά, η συνάρτηση συνδιασποράς, και το ημιβαριόγραμμα (Χριστόπουλος 2004α).

1.3.1 Μέση τιμή

Η μέση τιμή ενός τυχαίου πεδίου δίνεται από την παρακάτω σχέση, $m_X(\mathbf{s}) = E[X(\mathbf{s})]$. Το σύμβολο $E[X(\mathbf{s})]$ δηλώνει τη μέση τιμή υπολογισμένη ως προς το σύνολο των καταστάσεων του πεδίου, δηλαδή

$$E[X(\mathbf{s})] = \int dx f_X(x; \mathbf{s}) x, \quad (1.1)$$

όπου x οι τιμές που αντιστοιχούν σε μια κατάσταση.

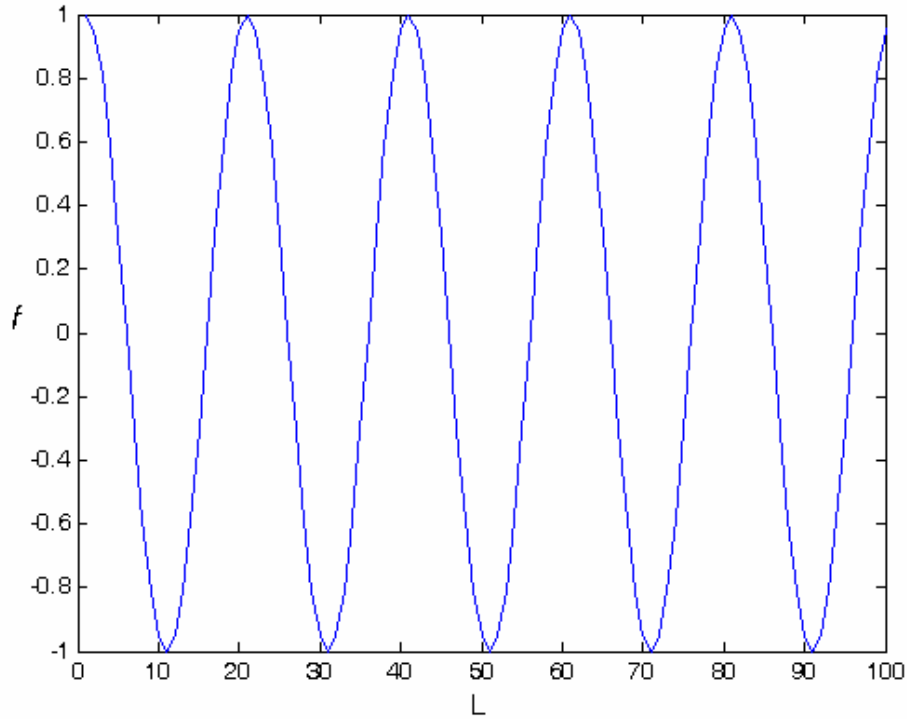
Τα όρια του ολοκληρώματος εξαρτώνται από τον χώρο στον οποίο ορίζεται το πεδίο X . Αν το πεδίο παίρνει όλες τις θετικές και αρνητικές τιμές το ολοκλήρωμα κυμαίνεται από $-\infty$ έως ∞ . Αν το πεδίο παίρνει μόνο θετικές τιμές τότε το ολοκλήρωμα κυμαίνεται από το 0 έως το ∞ . Αν είναι γνωστό ότι οι τιμές του πεδίου περιορίζονται σε ένα προκαθορισμένο διάστημα $[a, b]$, το ολοκλήρωμα υπολογίζεται σε αυτό το διάστημα.

Από την εξίσωση (1.1) παρατηρείται ότι η μέση τιμή μπορεί να έχει εξάρτηση από τη θέση \mathbf{s} , η οποία προέρχεται από πιθανή εξάρτηση της μονοδιάστατης

συνάρτησης πυκνότητας πιθανότητας από τη θέση. Επειδή η ΣΠΠ δεν είναι πάντα γνωστή εκ των προτέρων η μέση τιμή εκτιμάται από το δείγμα με στατιστικές μεθόδους. Αυτή δίνεται από το μέσο όρο των τιμών που περιλαμβάνονται στο δείγμα, (Χριστόπουλος 2004α),

$$\hat{m}_x(\mathbf{s}) = \frac{1}{M} \sum_{i=1}^M x_i(\mathbf{s}) . \quad (1.2)$$

Η μέση τιμή περιγράφει τις τάσεις μεγάλης εμβέλειας σε ένα τυχαίο πεδίο (π.χ. η τάση μιας τυχαίας μεταβλητής που ορίζεται σε διάστημα $[0, L]$, όπου $L=100$, και δίνεται από την συνάρτηση $f = \cos(2\pi L / 20)$ παρουσιάζεται στο σχήμα 1.1).



Σχήμα 1.1: Γράφημα τάσης τυχαίας μεταβλητής.

Η μέση τιμή $m_x(\mathbf{s})$ προσδιορίζεται από πρότυπες συναρτήσεις. Αυτές διαχωρίζονται σε πρότυπα γενικής και τοπικής εξάρτησης. Στην περίπτωση γενικής εξάρτησης μια μοναδική μαθηματική σχέση περιγράφει την μεταβολή σε όλη την περιοχή. Τέτοια πρότυπα εξάρτησης είναι:

- Η γραμμική εξάρτηση, π.χ. $m_X(\mathbf{s}) = m_0 + \mathbf{b} \mathbf{s}$, η οποία εκφράζει την ύπαρξη μιας σταθερής κλίσης
- Η πολωνυμική εξάρτηση, π.χ. $m_X(\mathbf{s}) = m_0 + b_1 s_1 + b_2 s_2$
- Η περιοδική εξάρτηση, π.χ. $m_X(\mathbf{s}) = m_0 + \sum_{n=1}^N A_n \cos(\mathbf{k}_n \mathbf{s} + \phi_n)$, όπου οι μεταβλητές \mathbf{k}_n αντιστοιχούν σε χωρικές συχνότητες και οι ϕ_n σε φάσεις
- Η υπέρθεση δύο η περισσότερων προτύπων, π.χ. ενός πολωνυμικού και ενός περιοδικού, $m_X(\mathbf{s}) = (m_0 + b_1 s_1 + b_2 s_2) + (m_0 + \sum_{n=1}^N A_n \cos(\mathbf{k}_n \mathbf{s} + \phi_n))$.

Σε περιπτώσεις όμως στις οποίες τα γενικά πρότυπα εξάρτησης δεν επαρκούν για τον ακριβή προσδιορισμό των τάσεων είναι προτιμότερη η χρήση τοπικών συναρτήσεων εξάρτησης (π.χ. τοπικά πολώνυμα). Τέτοιες μορφές εξάρτησης χρησιμοποιούνται στο μοντέλο της τοπικά ζυγισμένης παλινδρόμησης (Χριστόπουλος 2004a).

1.3.2 Διασπορά

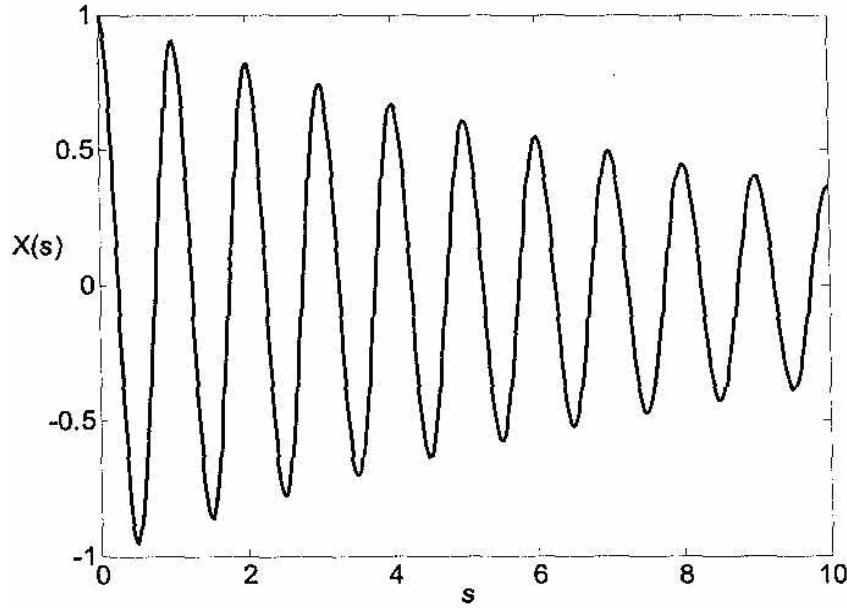
Η διασπορά ενός τυχαίου πεδίου δίνεται από τη μέση τιμή του τετραγώνου της διακύμανσης σύμφωνα με την εξίσωση (1.3).

$$\sigma_X^2(\mathbf{s}) \equiv E\left[\{X(\mathbf{s}) - m_X(\mathbf{s})\}^2\right] = E\left[\chi^2(\mathbf{s})\right] \quad (1.3)$$

Γενικά είναι δυνατόν η διασπορά να μεταβάλλεται από σημείο σε σημείο ενώ διατηρείται σταθερή μόνο όταν το πεδίο είναι στατιστικά ομοιογενές.

Η μεταβολή της διασποράς ενός τυχαίου πεδίου στον χώρο σημαίνει ότι οι διακυμάνσεις του πεδίου αλλάζουν μέγεθος από σημείο σε σημείο (Isaaks & Srivastava 1989). Στο επόμενο σχήμα (Σχήμα 1.2) φαίνεται το γράφημα μιας απλής συνάρτησης (όχι τυχαίου πεδίου) της οποίας το πλάτος διακύμανσης αλλάζει. Στην

περίπτωση ενός τυχαίου πεδίου η μεταβολή των διακυμάνσεων δεν είναι εξίσου κανονική και όχι πάντοτε ορατή με απλή επισκόπηση (Χριστόπουλος 2004α).



Σχήμα 1.2: Γράφημα συνημίτονου με εκθετικά μειούμενο πλάτος διακύμανσης (Χριστόπουλος 2004α).

1.3.3 Συνάρτηση συνδιασποράς

Μια άλλη ιδιότητα η οποία δίνει χρήσιμες πληροφορίες για ένα τυχαίο πεδίο είναι η κεντρική συνάρτηση συνδιασποράς (ΚΣΣ) η οποία ορίζεται με την βοήθεια της εξίσωσης (1.4), (Isaaks & Srivastava 1989),

$$c_X(\mathbf{s}_1, \mathbf{s}_2) \equiv E\left[\{X(\mathbf{s}_1) - m_X(\mathbf{s}_1)\}\{X(\mathbf{s}_2) - m_X(\mathbf{s}_2)\}\right]. \quad (1.4)$$

Το τυχαίο πεδίο $\chi(\mathbf{s}_1) \equiv X(\mathbf{s}_1) - m_X(\mathbf{s}_1)$ αντιστοιχεί στην **διακύμανση** του πεδίου $X(\mathbf{s}_1)$ γύρω από την μέση τιμή στο σημείο \mathbf{s}_1 . Η μέση τιμή του πεδίου διακύμανσης είναι ίση με το μηδέν, $E[\chi(\mathbf{s}_1)] = 0$. Βάσει των προηγούμενων η ΚΣΣ είναι ισοδύναμη με την **συνδιακύμανση**, δηλαδή $c_X(\mathbf{s}_1, \mathbf{s}_2) = E[\chi(\mathbf{s}_1)\chi(\mathbf{s}_2)]$.

Πιο συγκεκριμένα η ΚΣΣ περιγράφει ποσοτικά την εξάρτηση των διακυμάνσεων του πεδίου σε δύο διαφορετικά σημεία. Όταν τα σημεία της συνάρτησης συνδιασποράς

συμπίπτουν η τιμή της ισούται με την διασπορά του πεδίου στο συγκεκριμένο σημείο, $c_X(\mathbf{s}_1, \mathbf{s}_1) = \sigma_X^2(\mathbf{s}_1)$, όπως προκύπτει από την εξίσωση (1.3) και (1.4). Αντίθετα όταν η απόσταση μεταξύ δύο σημείων μεγαλώνει η εξάρτηση των διακυμάνσεων μειώνεται. Ένα παράδειγμα μεταβολής της συνάρτησης συνδιασποράς με την απόσταση δίνεται παρακάτω:

Έστω ότι η συνάρτηση συνδιασποράς μεταξύ δύο σημείων τυχαίου πεδίου δίνεται από το εκθετικό πρότυπο, $c_X(\mathbf{s}_1, \mathbf{s}_2) = \sigma_X^2 \exp\left(-\frac{\|\mathbf{r}\|}{\xi}\right)$, όπου $\|\mathbf{r}\|$ το Ευκλείδειο μέτρο του διανύσματος απόστασης μεταξύ δύο σημείων και ξ το μήκος συσχέτισης (1.3.5). Αν $\|\mathbf{r}\| = 0$ τότε $c_X(\mathbf{s}_1, \mathbf{s}_2) = \sigma_X^2$, ενώ αν $\|\mathbf{r}\| = \xi$ τότε $c_X(\mathbf{s}_1, \mathbf{s}_2) = 0.36\sigma_X^2$.

Σε γεωστατιστικές αναλύσεις η πειραματικά προσδιοριζόμενη χωρική εξάρτηση προσαρμόζεται σε ένα βέλτιστο πρότυπο το οποίο επιλέγεται από ένα σύνολο αποδεκτών θεωρητικών προτύπων (π.χ εκθετικό, γκαουσιανό). Επομένως είναι απαραίτητη η ύπαρξη συνθηκών αποδοχής για την συνάρτηση συνδιασποράς.

Οι **συνθήκες αποδοχής** καθορίζονται από το θεώρημα του Bochner (Bochner 1959). Αυτό εκφράζεται με την βοήθεια της φασματικής πυκνότητας ισχύος της συνδιασποράς η οποία δίνεται από τον μετασχηματισμό Fourier (Press et all 1992), της συνάρτησης συνδιασποράς. Η φασματική πυκνότητα ισχύος ορίζεται από το ολοκλήρωμα,

$$\tilde{c}_X(\mathbf{k}) = \int d\mathbf{r} \exp(-i\mathbf{k} \cdot \mathbf{r}) c_X(\mathbf{r}), \quad (1.5)$$

όπου \mathbf{r} το διάνυσμα απόστασης μεταξύ δύο σημείων, $\int d\mathbf{r} = \int dx \int dy$ και \mathbf{k} είναι το διάνυσμα της χωρικής συχνότητας (κυματάνυσμα).

Μια συνάρτηση $c_X(\mathbf{r})$ είναι αποδεκτή συνάρτηση συνδιασποράς αν ισχύουν οι παρακάτω τρεις συνθήκες:

- 1) Αν υπάρχει η φασματική πυκνότητα ισχύος $\tilde{c}_X(\mathbf{k})$ (δηλαδή αν υπάρχει ο μετασχηματισμός Fourier της συνάρτησης).
- 2) Αν η $\tilde{c}_X(\mathbf{k})$ είναι μη αρνητική σε όλο το πεδίο συχνοτήτων δηλαδή $\tilde{c}_X(\mathbf{k}) \geq 0$ για κάθε \mathbf{k} .
- 3) Αν το ολοκλήρωμα της $\tilde{c}_X(\mathbf{k})$ σε ολόκληρο το πεδίο συχνότητας είναι φραγμένο (δηλαδή αν υπάρχει η διασπορά).

Από πρακτική άποψη για να διαπιστωθεί αν μια συνάρτηση αποτελεί αποδεκτό πρότυπο συνάρτησης συνδιασποράς χρειάζεται να υπολογιστεί ο μετασχηματισμός Fourier της συνάρτησης (Χριστόπουλος 2004a).

1.3.4 Στατιστική ομοιογένεια

Ορισμένες παραδοχές που θέτουν περιορισμούς στις ιδιότητες ενός τυχαίου πεδίου μπορούν να οδηγήσουν σε μια πιο αποτελεσματική γεωστατιστική ανάλυση. Η πιο ευρέως χρησιμοποιημένη απλουστευτική παραδοχή είναι η **στατιστική ομοιογένεια**, η οποία αποτελεί επέκταση του κλασικού ορισμού της ομοιογένειας. Μια ιδιότητα είναι ομοιογενής αν η αντίστοιχη μεταβλητή έχει σταθερή τιμή στο χώρο. Αντίθετα ένα τυχαίο πεδίο $X(\mathbf{s})$ είναι στατιστικώς ομοιογενές αν η μέση τιμή είναι σταθερή, $m_X(\mathbf{s}) = m_X$, η συνάρτηση συνδιασποράς ορίζεται και εξαρτάται αποκλειστικά από το διάνυσμα απόστασης $\mathbf{r} = \mathbf{s}_1 - \mathbf{s}_2$ μεταξύ των δύο σημείων, $c_X(\mathbf{s}_1, \mathbf{s}_2) = c_X(\mathbf{r})$, και η διασπορά του πεδίου είναι επίσης σταθερή.

Οι παραπάνω προϋποθέσεις ορίζουν την στατιστική ομοιογένεια κατά την ασθενή έννοια. Ένα τυχαίο πεδίο είναι στατιστικώς ομοιογενές κατά την ισχυρή έννοια όταν η πολυδιάστατη ΣΠΠ για N σημεία (όπου N οποιοσδήποτε θετικός ακέραιος) παραμένει αμετάβλητη από μετασχηματισμούς οι οποίοι αλλάζουν την θέση των σημείων χωρίς να αλλάζουν τις μεταξύ τους αποστάσεις.

Επομένως η έννοια της στατιστικής ομοιογένειας είναι ότι οι στατιστικές ιδιότητες ενός τυχαίου πεδίου δεν εξαρτώνται από τις χωρικές συντεταγμένες των σημείων,

άρα και από το σύστημα αναφοράς. Πρακτικά η στατιστική ομοιογένεια προϋποθέτει ότι δεν υπάρχουν συστηματικές τάσεις έτσι ώστε η μεταβολή των τιμών του πεδίου μπορεί να αποδοθεί σε διακυμάνσεις γύρω από μια σταθερή στάθμη ίση με την μέση τιμή (Χριστόπουλος 2004α).

1.3.5 Στατιστική ισοτροπία

Μια άλλη ιδιότητα που μπορεί να φανεί χρήσιμη στην γεωστατιστική ανάλυση ενός τυχαίου πεδίου είναι η **στατιστική ισοτροπία**. Ένα πεδίο είναι στατιστικώς ισοτροπικό αν είναι στατιστικώς ομοιογενές και συγχρόνως η συνάρτηση συνδιασποράς εξαρτάται μόνο από το μέτρο (ευκλείδεια απόσταση), αλλά όχι από τη κατεύθυνση του διανύσματος απόστασης \mathbf{r} . Αυτό είναι σημαντικό από πρακτική άποψη επειδή διευκολύνει τον προσδιορισμό της χωρικής εξάρτησης. Αν μια συνάρτηση συνδιασποράς είναι στατιστικώς ισοτροπική είναι και εξ ορισμού στατιστικώς ομοιογενής, αλλά το αντίστροφο δεν ισχύει.

Στην περίπτωση στατιστικώς ισοτροπικών πεδίων οι δύο πιο σημαντικές παράμετροι οι οποίες προσδιορίζουν πολύ βασικά χαρακτηριστικά της συνάρτησης συνδιασποράς είναι η διασπορά $\sigma_X^2 = c_X(0)$ και το μήκος συσχέτισης ζ . Η διασπορά αποτελεί μέτρο του πλάτους των διακυμάνσεων του πεδίου. Το μήκος συσχέτισης ορίζει το διάστημα μέσα στο οποίο υπάρχει αλληλεξάρτηση, δηλαδή ορίζει την απόσταση μέσα στην οποία η τιμή του πεδίου σε ένα σημείο επηρεάζει την τιμή σε ένα άλλο σημείο (Χριστόπουλος 2004α).

1.3.6 Χωρική εξάρτηση

Υπάρχουν διάφοροι τρόποι μέτρησης της χωρικής εξάρτησης. Δύο από τους πλέον ευρέως χρησιμοποιούμενους είναι το **ημιβαριόγραμμα** και η **συνάρτηση συσχέτισης**. Και οι δύο συναρτήσεις περιγράφουν την εξάρτηση δύο σημείων στον χώρο κατά τη στατιστική έννοια καθώς και οι δύο συναρτήσεις αναφέρονται σε ζεύγη σημείων, επομένως η τιμή τους εξαρτάται από την απόσταση μεταξύ των σημείων. Ο όρος κατά τη στατιστική έννοια, προσδιορίζει ότι η περιγραφόμενη εξάρτηση προκύπτει ως μια μέση τιμή από ένα μεγάλο αριθμό

ζευγών και δεν χαρακτηρίζει ένα μεμονωμένο ζεύγος σημείων (Χριστόπουλος 2004β).

Η συνάρτηση συσχέτισης για ένα τυχαίο πεδίο ισούται με τον λόγο της συνάρτησης συνδιασποράς προς την διασπορά και δίνεται από την παρακάτω εξίσωση,

$$\rho_X = \frac{c_X(\mathbf{r})}{\sigma_X^2(\mathbf{s})} \quad (1.6)$$

ενώ το ημιβαριόγραμμα ενός τυχαίου πεδίου ορίζεται βάσει της εξίσωσης (1.7),

$$\gamma_X(\mathbf{s}, \mathbf{r}) = \frac{1}{2} E \left\{ [X(\mathbf{s} + \mathbf{r}) - X(\mathbf{s})]^2 \right\}. \quad (1.7)$$

Δηλαδή το ημιβαριόγραμμα ορίζεται σε σχέση με ένα ζεύγος σημείων με την βοήθεια της μέσης τιμής του τετραγώνου της διαφοράς $\delta X(\mathbf{s}; \mathbf{r}) \equiv X(\mathbf{s} + \mathbf{r}) - X(\mathbf{s})$. Το πεδίο της διαφοράς $\delta X(\mathbf{s}; \mathbf{r})$ αποκαλείται βήμα απόστασης \mathbf{r} . Αν το πεδίο $X(\mathbf{s})$ είναι στατιστικώς ομοιογενές το ημιβαριόγραμμα συνδέεται άμεσα με την συνάρτηση συνδιασποράς βάσει της εξίσωσης $\gamma_X(\mathbf{r}) = \sigma_X^2 - c_X(\mathbf{r})$.

Για στατιστικώς ομοιογενή πεδία το ημιβαριόγραμμα περιέχει την ίδια πληροφορία με την συνάρτηση συνδιασποράς. Αν η διαφορά $\delta X(\mathbf{s}; \mathbf{r})$ είναι στατιστικά ομοιογενής, το τυχαίο πεδίο $X(\mathbf{s})$ ονομάζεται πεδίο με στατιστικώς ομοιογενείς διαφορές. Σε αυτήν την περίπτωση το ημιβαριόγραμμα $\gamma_X(\mathbf{r})$ εξαρτάται αποκλειστικά από την απόσταση \mathbf{r} μεταξύ των σημείων και αυτό είναι απόρροια της στατιστικής ομοιογένειας του πεδίου διαφορών. Αν το πεδίο $X(\mathbf{s})$ είναι στατιστικά ομοιογενές το ίδιο ισχύει και για τις διαφορές $\delta X(\mathbf{s}; \mathbf{r})$, το αντίστροφο όμως δεν ισχύει απαραίτητα (Χριστόπουλος 2004β).

Οι παράμετροι του ημιβαριογράμματος καθορίζουν την χωρική εξάρτηση των τιμών του πεδίου σε δυο γειτονικά σημεία. Από τον ορισμό του ημιβαριογράμματος με την βοήθεια της μέσης τιμής του τετραγώνου των διαφορών προκύπτει ότι το ημιβαριόγραμμα είναι

ημιθετικά ορισμένο, $\gamma_X(\mathbf{r}) \geq 0$. Ωστόσο το αντίστροφο δεν ισχύει πάντοτε καθώς μια ημιθετικά ορισμένη συνάρτηση δεν είναι κατ' ανάγκη αποδεκτή ως ημιβαριόγραμμα.

Σε περίπτωση στατιστικά ομοιογενούς πεδίου, αν η χωρική εξάρτηση είναι ιστροπική, το ημιβαριόγραμμα προσδιορίζεται από δύο παραμέτρους: το *όριο* και το *μήκος συσχέτισης*. Η τιμή του ημιβαριογράμματος για πολύ μεγάλες αποστάσεις \mathbf{r} τείνει ασυμπτωτικά προς ένα όριο ίσο με την διασπορά σ_X^2 του τυχαίου πεδίου. Αυτή η ιδιότητα βασίζεται στην σχέση $\gamma_X(\mathbf{r}) = \sigma_X^2 - c_X(\mathbf{r})$ και το γεγονός ότι σε μεγάλες αποστάσεις η τιμή της συνάρτησης συνδιασποράς τείνει προς το μηδέν.

Η παρουσία σημαντικών τάσεων μεγάλης εμβέλειας σημαίνει ότι η προϋπόθεση της στατιστικής ομοιογένειας δεν ισχύει. Τότε το ημιβαριόγραμμα δεν προσεγγίζει κάποια τιμή ισορροπίας όταν η απόσταση τείνει προς το άπειρο (Χριστόπουλος 2004β).

Αν τα χαρακτηριστικά της συσχέτισης μεταβάλλονται σε διαφορετικές κατευθύνσεις στο χώρο τότε η εξάρτηση είναι ανιστροπική. Οι δύο κύριοι τύποι ανιστροπίας που συναντώνται στην πράξη είναι η γεωμετρική ανιστροπία και η ανιστροπία ζώνης. Η γεωμετρική ανιστροπία αναφέρεται σε περιπτώσεις στις οποίες το άνω φράγμα του ημιβαριογράμματος είναι ανεξάρτητο της κατεύθυνσης αλλά η ταχύτητα προσέγγισης του φράγματος εξαρτάται από την διεύθυνση (Hohn 1999). Σε αυτήν την περίπτωση το

ημιβαριόγραμμα εκφράζεται ως συνάρτηση $\gamma_X\left(\frac{r_1}{\xi_1}, \dots, \frac{r_d}{\xi_d}\right)$ των αδιάστατων

αποστάσεων $\frac{r_1}{\xi_1}, \dots, \frac{r_d}{\xi_d}$ όπου ξ_1, \dots, ξ_d είναι τα μήκη συσχέτισης στις αντίστοιχες

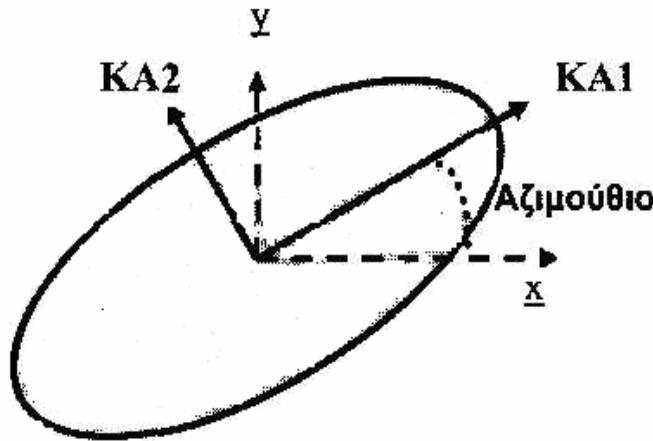
διευθύνσεις (Χριστόπουλος 2004β).

Η ανιστροπία ζώνης αναφέρεται στην περίπτωση στην οποία το άνω φράγμα εξαρτάται από την χωρική διεύθυνση. Το ημιβαριόγραμμα τότε μπορεί να εκφραστεί ως άθροισμα συνισταμένων (Hohn 1999), όπως φαίνεται στην εξίσωση (1.8),

$$\gamma_X(\mathbf{r}) = \gamma_{X,1}(r) + \gamma_{X,2}(\hat{\mathbf{r}}) \quad (1.8)$$

Στην παραπάνω σχέση η συνάρτηση $\gamma_{x,1}(r)$, όπου $r = \|\mathbf{r}\|$, περιγράφει μια ισοτροπική εξάρτηση ενώ η συνάρτηση $\gamma_{x,2}(\hat{\mathbf{r}})$ περιγράφει την ανισοτροπική εξάρτηση του άνω φράγματος από την κατεύθυνση του μοναδιαίου διανύσματος $\hat{\mathbf{r}}$ (Χριστόπουλος 2004β).

Στην περίπτωση της γεωμετρικής ανισοτροπίας χρειάζονται περισσότερα του ενός μήκη συσχέτισης ξ_1, \dots, ξ_d . Κάποια από αυτά, αλλά όχι όλα, μπορεί να είναι ίσα μεταξύ τους. Επομένως χρειάζονται πρόσθετες παράμετροι για τον προσδιορισμό της ανισοτροπίας του ημιβαριογράμματος. Σε ένα δισδιάστατο σύστημα αν, ξ_x και ξ_y , δηλώνουν τα μήκη συσχέτισης κατά μήκος των κυρίων αξόνων, οι ανισοτροπικές παράμετροι είναι: (1) Ο λόγος $\rho_{y/x} \equiv \xi_y / \xi_x$ που ονομάζεται ανισοτροπικός λόγος. (2) Η γωνία προσανατολισμού που καθορίζει τον προσανατολισμό των κυρίων αξόνων ανισοτροπίας σε σχέση με το καρτεσιανό σύστημα συντεταγμένων (Σχήμα 1.3).



Σχήμα 1.3: Σύστημα κύριων αξόνων (KA1, KA2) σε σχέση με το σύστημα συντεταγμένων x και y και ισοϋψής καμπύλη (έλλειψη) του ημιβαριογράμματος (Χριστόπουλος 2004β).

Για να κατανοηθεί η έννοια της γωνίας προσανατολισμού ορίζονται οι ισοϋψείς καμπύλες ως ο γεωμετρικός τόπος των σημείων (r_x, r_y) στα οποία η τιμή του ημιβαριογράμματος είναι σταθερή. Σε πολλές περιπτώσεις οι καμπύλες αυτές έχουν την μορφή ελλείψεως καθώς αυτό συμβαίνει για διάφορα πρότυπα ημιβαριογράμματος, όπως το εκθετικό και το γκαουσιανό ανισοτροπικό ημιβαριόγραμμα. Η γωνία προσανατολισμού είναι τότε η γωνία που σχηματίζει ο

μεγάλος άξονας (KA1) της έλλειψης με τον οριζόντιο άξονα του συστήματος συντεταγμένων (Σχήμα 1.3), (Χριστόπουλος 2004β).

Το ημιβαριόγραμμα εν γένει αυξάνεται ανάλογα, όχι κατ' ανάγκη γραμμικά, με την απόσταση μεταξύ των σημείων, ενώ αντίστροφα η συνάρτηση συσχέτισης μειώνεται. Αυτό συμβαίνει επειδή η συνάρτηση συσχέτισης περιγράφει την εξάρτηση των τιμών του πεδίου σε δύο διαφορετικά σημεία του χώρου, και η εξάρτηση μειώνεται όσο αυξάνει η απόσταση. Το ημιβαριόγραμμα αντίθετα μετρά πόσο διαφέρουν μεταξύ τους οι διακυμάνσεις του πεδίου ως συνάρτηση της απόστασης. Επομένως, οι τιμές του ημιβαριογράμματος αυξάνονται όσο μεγαλώνει η απόσταση.

Για στατιστικά ομοιογενή πεδία, οι δύο συναρτήσεις είναι ισοδύναμες, δηλαδή εμπεριέχουν την ίδια πληροφορία σε διαφορετική μορφή. Υπάρχουν ωστόσο περιπτώσεις τυχαίων πεδίων στις οποίες το ημιβαριόγραμμα είναι αποκλειστικά συνάρτηση της απόστασης μεταξύ δύο σημείων, ενώ η συνάρτηση συσχέτισης εξαρτάται τόσο από την απόσταση όσο και από τις ακριβείς θέσεις των σημείων στον χώρο (Χριστόπουλος 2004β).

Ευρέως χρησιμοποιημένα μοντέλα ημιβαριογραμμάτων τα οποία είναι επιτρεπτά και σε πρακτικές εφαρμογές είναι το εκθετικό, το γκαουσιανό, το σφαιρικό, το γενικευμένο, το δυναμονομικό ή αλγεβρικό και το φαινόμενο πυρήνα. Το εκθετικό μοντέλο χαρακτηρίζει κατανομές με απότομες χωρικές μεταβολές σε αντίθεση με το γκαουσιανό το οποίο χαρακτηρίζει ομαλότερες αυξομειώσεις. Το αλγεβρικό μοντέλο χαρακτηρίζει εξάρτηση με μακριά χωρική εμβέλεια ενώ το μοντέλο πυρήνα αντιστοιχεί σε μεταβολές που συντελούνται σε αποστάσεις μικρότερες από τη διακριτική ικανότητα που επιτρέπει το δείγμα (Χριστόπουλος 2004β).

Ένας άλλος τρόπος προσδιορισμού της χωρικής εξάρτησης τυχαίου πεδίου, ο οποίος θα εφαρμοστεί στην συγκεκριμένη εργασία είναι με την μέθοδο των Σπαρτιάτικών Μοντέλων Τυχαίων Πεδίων. Στην συγκεκριμένη μέθοδο για να προσδιοριστεί η χωρική εξάρτηση του τυχαίου πεδίου απαιτείται ο υπολογισμός ενός αριθμού παραμέτρων χωρίς να είναι απαραίτητος ο υπολογισμός του ημιβαριογράμματος. Για να υπολογιστούν αυτές

οι παράμετροι ορίζονται στατιστικοί περιορισμοί οι οποίοι μπορούν να υπολογιστούν από το δείγμα (Hristopulos 2003).

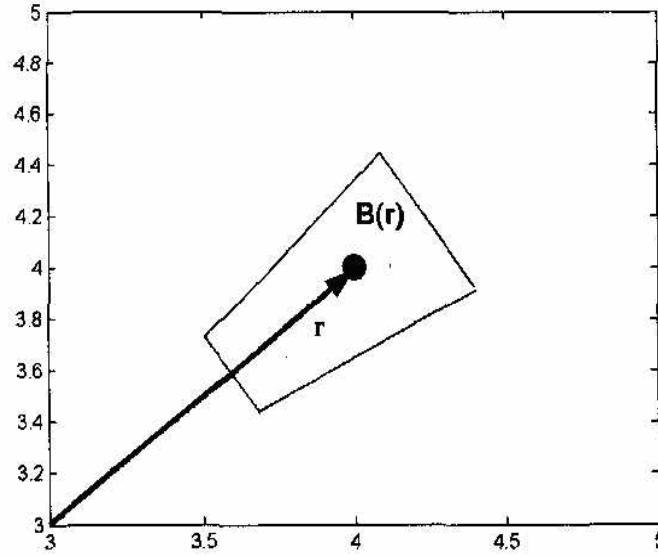
1.3.7 Προσδιορισμός ημιβαριογράμματος

Στην περίπτωση γεωγραφικών κατανομών, μεταλλευτικών κοιτασμάτων, και κατανομών περιβαλλοντικών ρυπαντών όπου τα δεδομένα περιορίζονται συνήθως σε ένα μοναδικό δείγμα επιχειρείται να προσδιοριστεί μια εκτίμηση του πραγματικού ημιβαριογράμματος από αυτό. Η εκτίμηση αυτή ονομάζεται *δειγματικό ημιβαριόγραμμα* και υπολογίζεται βάσει των τιμών του δείγματος ως εξής:

$$\hat{\gamma}_X(\mathbf{r}_k) = \frac{1}{2n(\mathbf{r}_k)} \sum_{i,j=1}^N \left\{ [X(\mathbf{s}_i) - X(\mathbf{s}_j)]^2 \right\} \vartheta_{ij}(\mathbf{r}_k), \quad (k=1, \dots, N_c) \quad (1.9)$$

$$\vartheta_{ij}(\mathbf{r}_k) = \begin{cases} 1, & \mathbf{s}_i - \mathbf{s}_j \in B(\mathbf{r}_k) \\ 0, & \text{otherwise} \end{cases} \quad (1.10)$$

- Η συνάρτηση τάξης $\vartheta_{ij}(\mathbf{r}_k)$ ορίζει διαφορετικές ομάδες (τάξεις) διανυσμάτων απόστασης, επιλέγοντας τα διανύσματα εκείνα που βρίσκονται σε μια κλειστή περιοχή $B(\mathbf{r}_k)$ γύρω από το διάνυσμα \mathbf{r}_k (Σχήμα 1.4).
- Η μεταβλητή $n(\mathbf{r}_k)$ είναι ίση με το πλήθος των ζευγών σημείων που περιέχονται μέσα στην τάξη $B(\mathbf{r}_k)$.
- Το δειγματικό ημιβαριόγραμμα ορίζεται για ένα διακριτό και πεπερασμένο σύνολο αποστάσεων \mathbf{r}_k , ($k=1, \dots, N_c$) το πλήθος των οποίων είναι ίσο με τον συνολικό αριθμό τάξεων N_c .



Σχήμα 1.4: Σχηματικό διάγραμμα μιας περιοχής $B(\mathbf{r})$ γύρω από το διάνυσμα απόστασης (Χριστόπουλος 2004β).

Επομένως, ο υπολογισμός αυτός προσδιορίζει μια τιμή του δειγματικού ημιβαριογράμματος για κάθε \mathbf{r}_k βάσει του μέσου όρου των διαφορών $[X(\mathbf{s}_i) - X(\mathbf{s}_j)]^2$ σε όλα τα ζεύγη σημείων, το διάνυσμα απόστασης των οποίων ανήκει στην περιοχή $B(\mathbf{r}_k)$. Το $\hat{\gamma}_X(\mathbf{r}_k)$ είναι ένας καλός εκτιμητής του $\gamma_X(\mathbf{r}_k)$ όταν ο μέσος όρος των διαφορών στην τάξη του \mathbf{r}_k προσεγγίζει με ακρίβεια την μέση τιμή $E[X(\mathbf{s}) - X(\mathbf{s} + \mathbf{r}_k)]^2$ (Χριστόπουλος 2004β).

Αυτό ισχύει όταν εκπληρώνεται η εργοδική υπόθεση (Christakos & Hristopulos 1998) η οποία επιτρέπει την εναλλαγή του στοχαστικού με τον δειγματικό μέσο. Στον υπολογισμό του ημιβαριογράμματος για να ισχύει η εργοδική ιδιότητα πρέπει να εκπληρώνονται κάποιες προϋποθέσεις, όπως: (1) το πεδίο διαφοράς $X(\mathbf{s}) - X(\mathbf{s} + \mathbf{r}_k)$ να είναι στατιστικά ομοιογενές, (2) το πλήθος ζευγών σε κάθε τάξη να είναι μεγάλο ώστε ο δειγματικός μέσος του τετραγώνου της διαφοράς να υπολογίζεται με καλή στατιστική ακρίβεια και (3) ο αριθμός των τάξεων πρέπει να είναι αρκετά μεγάλος ώστε να επιτρέπει μια πυκνή προσέγγιση της μεταβολής του ημιβαριογράμματος σαν συνάρτηση της απόστασης (Χριστόπουλος 2004β).

Αφού υπολογιστεί το πειραματικό ημιβαριόγραμμα προσαρμόζεται σε ένα θεωρητικό μοντέλο, π.χ.,

- Εκθετικό: $\gamma_X(\mathbf{r}) = \sigma_X^2 \left[1 - \exp\left(-\frac{\mathbf{r}}{\xi}\right) \right]$
- Γκαουσιανό: $\gamma_X(\mathbf{r}) = \sigma_X^2 \left[1 - \exp\left(-\frac{\mathbf{r}^2}{\xi^2}\right) \right]$

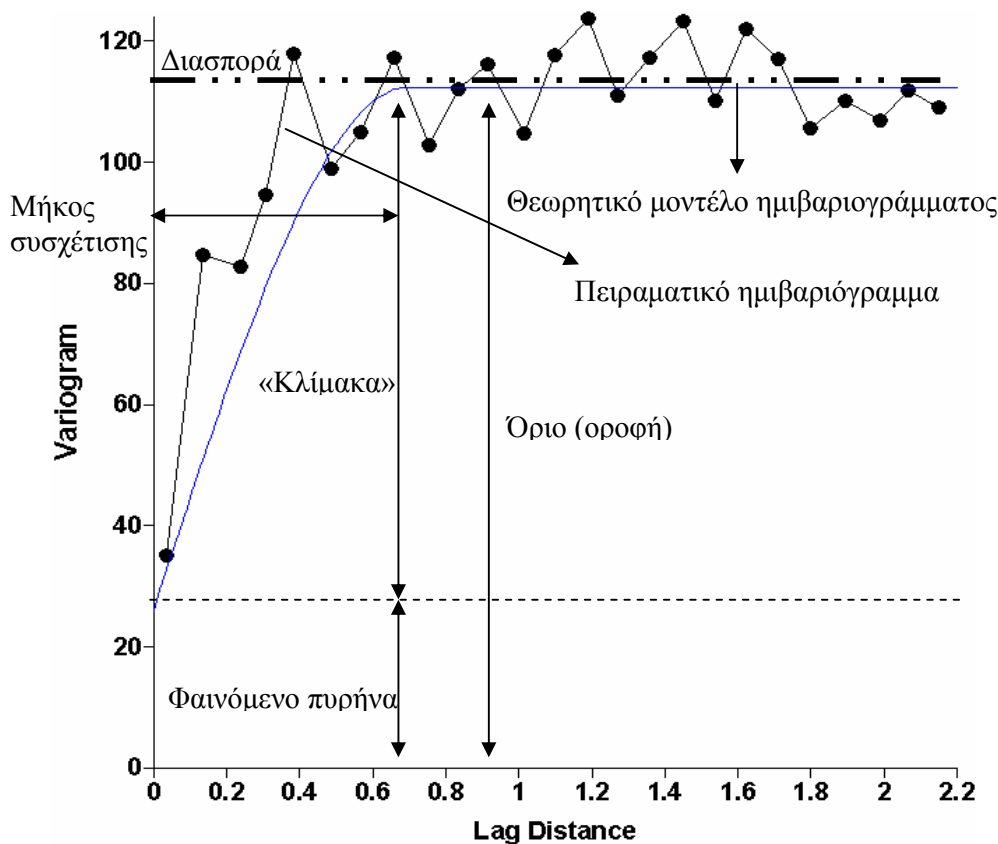
το οποίο επιτρέπει τον υπολογισμό του ημιβαριογράμματος για οποιαδήποτε απόσταση. Αυτό επιτυγχάνεται χρησιμοποιώντας την αρχή των ελαχίστων τετραγώνων, από την οποία υπολογίζονται και οι βέλτιστες παράμετροι ξ και σ_X^2 του θεωρητικού μοντέλου.

Το θεωρητικό πρότυπο χρειάζεται για την εκτίμηση των τιμών του πεδίου σε σημεία όπου δεν υπάρχουν μετρήσεις. Στην συνέχεια για να γίνει αποδεκτό το ημιβαριόγραμμα και να χρησιμοποιηθεί στην γεωστατική ανάλυση ελέγχεται σύμφωνα με τις συνθήκες αποδοχής ημιβαριογράμματος (Χριστόπουλος 2004b).

Αν υπάρχει ανισοτροπική χωρική εξάρτηση θα πρέπει να υπολογιστεί το ημιβαριόγραμμα σε διάφορες κατευθύνσεις στον χώρο, προκειμένου να προσδιοριστούν οι κύριες κατευθύνσεις της ανισοτροπίας. Αυτό προϋποθέτει τον ορισμό τάξεων σε σχέση όχι μόνο με το μέτρο αλλά και την κατεύθυνση του διανύσματος απόστασης. Κάθε τάξη έχει μία ανοχή ως προς το μέτρο ($2\delta r$) και μια γωνιακή ανοχή ($2\delta\phi$) ως προς την κατεύθυνση του διανύσματος, έτσι ώστε κάθε τάξη να περιλαμβάνει ένα ικανό αριθμό σημείων.

Το ημιβαριόγραμμα υπολογίζεται συνήθως στις διευθύνσεις Βορράς-Νότος και Ανατολή-Δύση ενώ για την γωνιακή ανοχή χρησιμοποιούνται οι τιμές 5° , 10° , 20° και 45° . Προς το παρόν όμως δεν υπάρχει απόλυτα ασφαλής ποσοτική μέθοδος ανίχνευσης και χαρακτηρισμού της ανισοτροπίας στις γεωστατιστικές μελέτες (Χριστόπουλος 2004α).

Το σχήμα 1.5 παρουσιάζει τα χαρακτηριστικά ενός ημιβαριογράμματος τα οποία εξετάζονται αναλυτικά παρακάτω.



Σχήμα 1. 5: Παρουσίαση χαρακτηριστικών στοιχείων ημιβαριογράμματος (Surfer V.8.0.4, 2002).

- Το φαινόμενο πυρήνα ποσοτικοποιεί την διασπορά του δειγματικού σφάλματος καθώς και την μικρής κλίμακας μεταβλητότητα, π.χ. την χωρική μεταβλητότητα που υπάρχει σε αποστάσεις μικρότερες από τις αποστάσεις μεταξύ των σημείων του δείγματος.
- Το «όριο» (οροφή) είναι η τιμή που πλησιάζει ασυμπτωτικά το δειγματικό ημιβαριόγραμμα.
- Η «κλίμακα» (scale) είναι η διαφορά της «οροφής» από το «φαινόμενο πυρήνα» και δηλώνει την μεταβλητότητα των συσχετισμένων διακυμάνσεων.
- Το μήκος συσχέτισης είναι η απόσταση στην οποία το ημιβαριόγραμμα προσεγγίζει πολύ κοντά, π.χ. κατά 95-97%, την τιμή οροφής.

- Η διασπορά είναι η μέση τετραγωνική απόκλιση κάθε τιμής του δείγματος από την μέση τιμή και δηλώνεται από την διακεκομμένη οριζόντια γραμμή στο σχήμα.
- Το πειραματικό ημιβαριόγραμμα παρουσιάζει τις ομάδες των ζευγών με τις αντίστοιχες δειγματικές τιμές του ημιβαριογράμματος.
- Το θεωρητικό μοντέλο ημιβαριογράμματος αποτελεί μια συνεχή θεωρητική καμπύλη που προσαρμόζεται στο πειραματικό.

1.3.8 Χωρική εκτίμηση

Ο προσδιορισμός της χωρικής εξάρτησης, καθώς τόσο της τάσης όσο και των διακυμάνσεων των τιμών του πεδίου οδηγεί σε δύο βασικές διαδικασίες της Γεωστατιστικής που είναι η **χωρική εκτίμηση** και η **προσομοίωση**. Και οι δύο διαδικασίες βοηθούν στην αναπαράσταση ενός τυχαίου πεδίου σε σημεία όπου δεν υπάρχουν ακριβείς τιμές με βάση την διαθέσιμη γνώση (π.χ. μετρήσεις σε γειτονικά σημεία, εκτιμήσεις ειδικών, γεωλογικά δεδομένα). Η διαθέσιμη γνώση χρησιμοποιείται για να επιβληθούν στατιστικοί περιορισμοί. Χρησιμοποιώντας στατιστικά πρότυπα χωρικής εξάρτησης (ημιβαριογράμματα) προσδιορίζονται οι άγνωστες τιμές βάσει της συσχέτισης. Η επανάληψη αυτής της διαδικασίας σε όλα τα σημεία ενός υπολογιστικού πλέγματος επιτρέπει τη χαρτογράφηση μιας ολόκληρης περιοχής (Χριστόπουλος 2004β).

Πιο συγκεκριμένα, ο όρος χωρική εκτίμηση περιλαμβάνει όλες τις μαθηματικές διαδικασίες που επιτρέπουν τον υπολογισμό των τιμών του πεδίου σε σημεία στα οποία δεν υπάρχουν μετρήσεις μιας ιδιότητας. Η χωρική εκτίμηση μπορεί να είναι είτε σημειακή, αν αναφέρεται στην τιμή του πεδίου σε ένα συγκεκριμένο σημείο, είτε γενική, οπότε αποσκοπεί στον υπολογισμό μιας χαρακτηριστικής τιμής που περιγράφει ολόκληρη την περιοχή.

Η χωρική εκτίμηση του πεδίου προϋποθέτει την ύπαρξη ενός προτύπου χωρικής εξάρτησης, έτσι ώστε η τιμή της ιδιότητας να επηρεάζεται από τις γειτονικές τιμές του πεδίου. Αυτή η αλληλεξάρτηση επιτρέπει εκτίμηση του πεδίου σε σημεία όπου

δεν υπάρχουν μετρήσεις βάσει της συμπεριφοράς σε γειτονικά μετρηθέντα σημεία. Σε πολλές περιπτώσεις, ο τελικός στόχος είναι η εκτίμηση σε ένα σύνολο σημείων και όχι σε ένα μεμονωμένο σημείο. Αυτό μπορεί να γίνει με επανάληψη της σημειακής εκτίμησης σε όλα τα σημεία ενδιαφέροντος.

Υπάρχουν ωστόσο διάφοροι μέθοδοι χωρικής εκτίμησης οι οποίες στηρίζονται σε παρόμοιες αρχές. Η βασική ιδέα είναι πως η τιμή στο σημείο εκτίμησης δίνεται από ένα συνδυασμό, γραμμικό ή μη γραμμικό των γειτονικών τιμών. Η εκτιμώμενη τιμή προκύπτει από την βελτιστοποίηση κάποιου στατιστικού μέτρου, π.χ., από την μεγιστοποίηση της πιθανότητας ή από την ελαχιστοποίηση του σφάλματος της εκτίμησης. Οι πλέον διαδεδομένες μέθοδοι βασίζονται στη γραμμική παρεμβολή σε συνδυασμό με την ελαχιστοποίηση του τετραγωνικού σφάλματος της εκτίμησης. Αυτό το σύνολο των μεθόδων είναι γνωστό ως «*Kriging*» (Χριστόπουλος 2004β).

Η διαδικασία της προσομοίωσης αντιθέτως, αποσκοπεί στο να παράγει πολλές από τις πιθανές καταστάσεις του πεδίου οι οποίες βρίσκονται σε καλή συμφωνία με τους υπάρχοντες στατιστικούς περιορισμούς που προκύπτουν από το πειραματικό δείγμα π.χ., οι προσομοιούμενες καταστάσεις να έχουν την ίδια μέση τιμή, τυπική απόκλιση και ημιβαριόγραμμα με αυτό που υπολογίζεται βάσει των μετρήσεων. Επομένως, η προσομοίωση αποσκοπεί στη παραγωγή πολλών εναλλακτικών σεναρίων, τα οποία είναι εφικτά με βάση τις υπάρχουσες μετρήσεις (Χριστόπουλος 2004β).

1.4 Περιορισμοί κλασικών γεωστατιστικών μεθόδων

Οι κλασικές γεωστατιστικές μέθοδοι που εφαρμόζονται ευρέως υστερούν σε δύο τουλάχιστον σημεία: Στον υπολογισμό του δειγματικού ημιβαριογράμματος και στην απουσία φυσικών μοντέλων (π.χ. διαφορικές εξισώσεις συνέχειας) από τον προσδιορισμό της χωρικής εξάρτησης. Όπως προαναφέρθηκε το ημιβαριόγραμμα είναι μία συνάρτηση δύο σημείων, ανάλογη της συνάρτησης συσχέτισης, που εκφράζει πως μεταβάλλεται η εξάρτηση δύο μεταβλητών ως συνάρτηση της μεταξύ τους απόστασης. Ένα χαρακτηριστικό παράδειγμα πειραματικού ημιβαριογράμματος παρουσιάζεται στο Σχήμα 1.6.

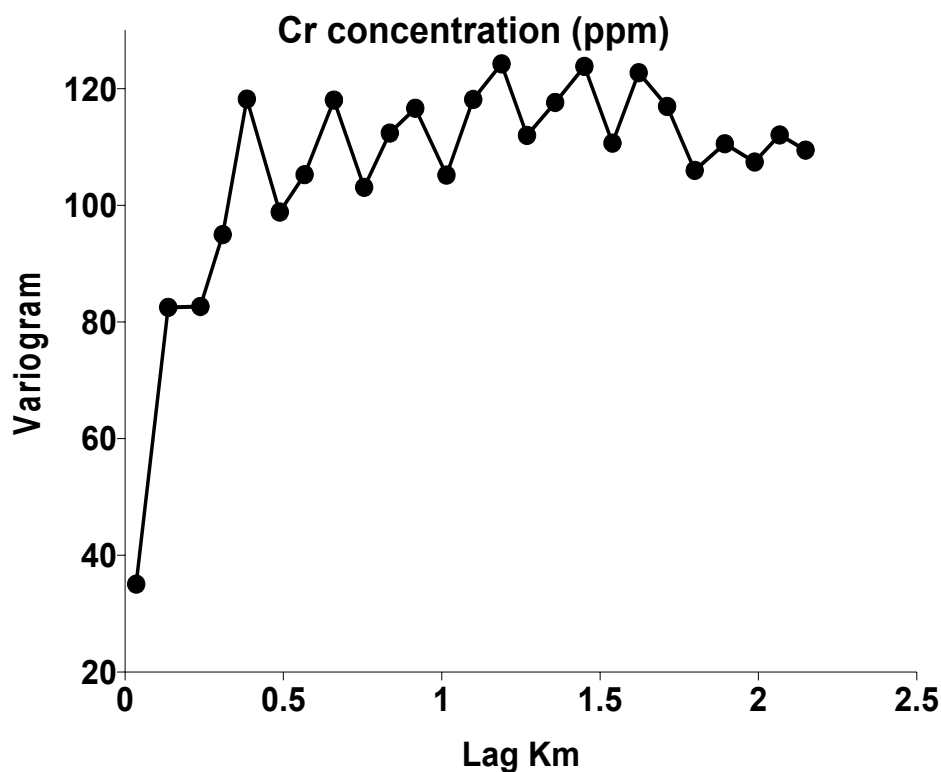
Η τυπική διαδικασία προσδιορισμού περιλαμβάνει αρχικά τον υπολογισμό του πειραματικού ημιβαριογράμματος και έπειτα την (μη γραμμική) προσαρμογή του ημιβαριογράμματος σε κάποιο από τα ευρέως χρησιμοποιούμενα θεωρητικά μοντέλα. Τα μοντέλα αυτά συνήθως προσδιορίζονται από $d + 1$ παραμέτρους, όπου d είναι ο αριθμός των διευθύνσεων ανισοτροπίας.

Οποιαδήποτε φυσική γνώση (π.χ. νόμοι διατήρησης μάζας, ενέργειας) σχετικά με το φαινόμενο δεν λαμβάνεται υπόψη, αλλά θεωρείται ότι οι σχετικές πληροφορίες περιέχονται ήδη στο πειραματικό ημιβαριόγραμμα. Ωστόσο, η διαδικασία προσαρμογής δεν εξασφαλίζει ότι τα θεωρητικά μοντέλα ημιβαριογράμματος διατηρούν τη φυσική πληροφορία. Το ημιβαριόγραμμα αποτελεί ένα χρήσιμο εργαλείο για την ανάλυση της χωρικής εξάρτησης, όμως υπάρχουν διάφορα προβλήματα σχετικά με τον προσδιορισμό και την χρήση του όπως:

- 1) Η υπολογιστική πολυπλοκότητα αυξάνεται με το τετράγωνο του αριθμού των σημείων. Αυτή η συμπεριφορά συνεπάγεται μεγάλο χρόνο υπολογισμού για δείγματα που περιέχουν μετρήσεις σε πολλά σημεία. Για να επιλυθεί αυτό το πρόβλημα χρησιμοποιείται συχνά ο υπολογισμός του ημιβαριογράμματος κατά μήκος συγκεκριμένων κατευθύνσεων. Ωστόσο, αυτή η προσέγγιση προϋποθέτει γνώση των κυρίων αξόνων στατιστικής ανισοτροπίας, η οποία δεν είναι συνήθως εκ των προτέρων γνωστή.
- 2) Η χωρική εξάρτηση φυσικών μεταβλητών είναι συχνά ανισοτροπική. Ωστόσο, οι υπάρχουσες μέθοδοι για τον προσδιορισμό της ανισοτροπίας του ημιβαριογράμματος βασίζονται συνήθως στη μέθοδο δοκιμής και σφάλματος.
- 3) Απαιτείται ο υπολογισμός ενός σημαντικού αριθμού «στατιστικών περιορισμών» για να γίνει δυνατή η προσαρμογή του ημιβαριογράμματος σε ένα συνεχές θεωρητικό πρότυπο. Για παράδειγμα, όπως φαίνεται στο Σχήμα 1.6, οι τιμές του ημιβαριογράμματος υπολογίζονται σε όλα τα σημεία που συμβολίζονται με κύκλους.

- 4) Το σφάλμα εκτίμησης των τιμών του ημιβαριογράμματος είναι σημαντικό.
- 5) Τα φυσικά πρότυπα δεν λαμβάνονται υπόψη στη διαδικασία προσδιορισμού της χωρικής εξάρτησης, ακόμα και όταν αυτές οι εξισώσεις είναι γνωστές (Hristopulos 2003).

Σύμφωνα με τις παραπάνω διαπιστώσεις γίνεται κατανοητό ότι οι κλασσικές γεωστατιστικές μέθοδοι χαρακτηρίζονται από κάποια μειονεκτήματα ως προς την εφαρμογή τους, κάτι που κάνει επιτακτική την ανάγκη για μια διαφορετική γεωστατιστική προσέγγιση.



Σχήμα 1.6: Ημιβαριόγραμμα της κατανομής περιεκτικότητας χρωμίου στα Ελβετικά βουνά Jura ως συνάρτηση της απόστασης μεταξύ δύο σημείων μέτρησης.

1.5 Γεωστατιστική ανάλυση τυχαίου πεδίου

Στην συγκεκριμένη εργασία θα πραγματοποιηθεί γεωστατιστική ανάλυση σε δείγμα μετρήσεων συγκέντρωσης χρωμίου στο έδαφος μιας περιοχής. Η γεωστατιστική ανάλυση θα εφαρμοστεί χρησιμοποιώντας την κλασική γεωστατιστική μέθοδο αλλά και μια καινοτόμο προσέγγιση.

Πιο συγκεκριμένα σκοπός της εργασίας είναι η ολοκληρωμένη παρουσίαση της λειτουργίας της νέας γεωστατιστικής μεθόδου, η εφαρμογή της σε πραγματικά δεδομένα πεδίου (κάτι που δεν έχει πραγματοποιηθεί μέχρι σήμερα), η παρουσίαση των διαφορών και των πλεονεκτημάτων της έναντι της κλασσικής γεωστατιστικής μεθόδου, η επιβεβαίωση της μεθόδου εξετάζοντας την ακρίβεια των αποτελεσμάτων της και την αξιοπιστία της λειτουργίας της, η χρησιμοποίησή της για την εκτίμηση σε σημεία του χώρου όπου δεν υπάρχουν μετρήσεις και η σύγκριση των αποτελεσμάτων της με αυτά της κλασσικής γεωστατιστικής μεθόδου.

Αρχικά η εφαρμογή της γεωστατιστικής ανάλυσης θα πραγματοποιηθεί με την κλασσική γεωστατιστική μέθοδο *Kriging* και θα ακολουθήσει η εφαρμογή της καινοτόμου γεωστατιστικής μεθόδου των Σπαρτιάτικων Τυχαίων Πεδίων. Η διαδικασία προσδιορισμού των παραμέτρων εκτίμησης και της χωρικής εκτίμησης που θα ακολουθηθεί είναι ανάλογη αυτής που περιγράφηκε παραπάνω.

Το διαθέσιμο δείγμα συγκεντρώσεων χρωμίου περιγράφεται από συγκεκριμένες στατιστικές ιδιότητες όπως η μέση τιμή, η διασπορά, η συνάρτηση πυκνότητας πιθανότητας και η συνδιασπορά. Ο υπολογισμός τους μπορεί να δώσει πληροφορίες σχετικά με την στατιστική ομοιογένεια και την στατιστική ισοτροπία των συγκεντρώσεων. Οι πληροφορίες αυτές είναι πολύ σημαντικές καθώς από αυτές εξαρτάται ο τρόπος προσδιορισμού της χωρικής εξάρτησης. Αφού προσδιοριστεί με τον κατάλληλο τρόπο, (ανάλογα με την γεωστατιστική μέθοδο και την περίπτωση), η χωρική εξάρτηση των συγκεντρώσεων του τυχαίου πεδίου καθίσταται δυνατή η χωρική εκτίμηση.

Στην συνέχεια επιβεβαιώνονται τα αποτελέσματα των δύο γεωστατιστικών μεθόδων χρησιμοποιώντας γνωστές συγκεντρώσεις σε σημεία μέτρησης. Τέλος, πραγματοποιείται χωρική εκτίμηση στα σημεία ενός πλέγματος που καλύπτει την περιοχή όπου η συγκέντρωση χρωμίου δεν είναι γνωστή, προκειμένου να παραχθούν γεωστατιστικοί χάρτες.

2. Σπαρτιάτικα τυχαία πεδία

2.1 Εισαγωγή

Όπως προαναφέρθηκε η γεωστατιστική ανάλυση μπορεί να έχει ευρεία εφαρμογή στην έρευνα και την τεχνολογία, όμως η κλασσική γεωστατιστική μέθοδος διέπεται από ορισμένα μειονεκτήματα. Με βάση λοιπόν ότι η γεωστατιστική στηρίζεται στην βασική μαθηματική έννοια του τυχαίου πεδίου, θα επιχειρηθεί μέσα από τη κατανόηση των τυχαίων πεδίων να εφαρμοστεί μια νέα γεωστατιστική μέθοδος η οποία σε κάποια σημεία πλεονεκτεί της κλασσικής γεωστατιστικής και ονομάζεται μοντέλα Σπαρτιάτικων Τυχαίων Πεδίων (ΣΤΠ).

Ο όρος Σπαρτιάτικα Τυχαία Πεδία (ΣΤΠ) χαρακτηρίζει γεωστατιστικά μοντέλα που προσδιορίζονται από ένα σχετικά μικρό αριθμό παραμέτρων, οι οποίες δεν προϋποθέτουν τον υπολογισμό του ημιβαριογράμματος. Σε αντίθεση με το ημιβαριόγραμμα, ο προσδιορισμός των παραμέτρων των ΣΤΠ βασίζεται σε σχετικά απλούς στατιστικούς περιορισμούς, που υπολογίζονται αποτελεσματικά. Οι εφαρμογές των ΣΤΠ περιλαμβάνουν ανάλυση, χαρακτηρισμό και προσομοίωση φυσικών μεγεθών που μεταβάλλονται στον χώρο, είτε αυτά αφορούν περιβαλλοντικούς ρύπους είτε κοιτάσματα. Υπολογιστικοί αλγόριθμοι μπορούν να αναπτυχθούν τόσο για δείγματα με κανονικές πλεγματικές κατανομές όσο και για δείγματα σε πλέγματα άτακτης δομής που χαρακτηρίζουν συνήθως την κατανομή των επίγειων δεδομένων (Hristopulos 2003).

Τα υπολογιστικά πλεονεκτήματα της μεθόδου ΣΤΠ σε σχέση με τον υπολογισμό του ημιβαριογράμματος γίνονται πιο σαφή όσο αυξάνεται το μέγεθος του δείγματος. Στον προσδιορισμό της χωρικής εξάρτησης με την βοήθεια του ημιβαριογράμματος δεν λαμβάνονται υπόψη φυσικά μοντέλα και περιορισμοί που διέπουν την μετρούμενη διαδικασία, αλλά χρησιμοποιείται η προσαρμογή σε γενικά μαθηματικά μοντέλα. Αντίθετα, στην περίπτωση των ΣΤΠ είναι δυνατή η κατασκευή συναρτήσεων εξάρτησης που λαμβάνουν υπόψη φυσικές διεργασίες. Οπότε η ανάπτυξη μεθόδων Σπαρτιάτικων Τυχαίων Πεδίων προσδίδει υπολογιστική ταχύτητα, συμπαγή χαρακτηρισμό της χωρικής εξάρτησης, συστηματικό

προσδιορισμό των ανισοτροπικών παραμέτρων, και επιτρέπει την εισαγωγή φυσικών περιορισμών στην συνάρτηση πυκνότητας πιθανότητας (Hristopulos 2003).

2.2 Περιγραφή του μοντέλου των ΣΤΠ

Τα μοντέλα Σπαρτιάτικων Τυχαίων Πεδίων ΣΤΠ καθορίζονται από μια συνάρτηση πυκνότητας πιθανότητας για την μετρούμενη ποσότητα $X_\lambda(\mathbf{s})$ η οποία παρίσταται ως ένα τυχαίο πεδίο με διακριτική ικανότητα λ . Η συνάρτηση πυκνότητας πιθανότητας περιλαμβάνει όλη την πληροφορία σχετικά με την χωρική εξάρτηση. Γενικά τα μοντέλα ΣΤΠ μπορούν να εκφραστούν με την εξής μαθηματική σχέση,

$$f_x[X_\lambda(\mathbf{s})] = Z^{-1} \exp\{-H[X_\lambda(\mathbf{s})]\} \quad (2.1)$$

Στην παραπάνω εξίσωση, $f_x[X_\lambda(\mathbf{s})]$ είναι η κοινή συνάρτηση πυκνότητας πιθανότητας (ΣΠΠ), $Z = \sum_{X_\lambda(\mathbf{s})} \exp\{-H[X_\lambda(\mathbf{s})]\}$ μια σταθερά κανονικοποίησης η οποία εξασφαλίζει το βασικό θεώρημα των πιθανοτήτων, δηλαδή ότι το άθροισμα των πιθανών καταστάσεων του τυχαίου πεδίου είναι ίσο με την μονάδα, και $H[X_\lambda(\mathbf{s})]$ ένα συναρτησιακό χωρικής αλληλεπίδρασης το οποίο εκφράζει τις αλληλεπιδράσεις των τιμών του τυχαίου πεδίου $X_\lambda(\mathbf{s})$ μεταξύ διαφορετικών σημείων του χώρου. Επομένως βάσει της παραπάνω εξίσωσης, τα ΣΤΠ αποτελούν υποσύνολα των τυχαίων πεδίων Gibbs (Hristopulos 2003).

Το νέο στοιχείο που εισάγουν τα ΣΤΠ είναι η εξάρτηση του $H[X_\lambda(\mathbf{s})]$ από φυσικά μεγέθη που σχετίζονται με την κατανομή στο χώρο, όπως τοπική κλίση και ακτίνα καμπυλότητας της τοπογραφίας, όπως επίσης και ο πλήρης προσδιορισμός της χωρικής αλληλεπίδρασης από ένα μικρό αριθμό παραμέτρων που υπολογίζονται από τα πειραματικά δεδομένα. Ένα παράδειγμα τοπικής αλληλεξάρτησης δίνεται από το εξής συναρτησιακό,

$$H[X_\lambda] = \frac{1}{2\eta_0\xi^d} \int d\mathbf{s} [\{X_\lambda(\mathbf{s}) - E[X_\lambda(\mathbf{s})]\}^2 + \eta_1\xi^2 \{\nabla X_\lambda(\mathbf{s})\}^2 + \xi^4 \{\nabla X_\lambda(\mathbf{s})\}^2] \quad (2.2)$$

Το ολοκλήρωμα δηλώνει πως λαμβάνονται υπόψη όλα τα σημεία του πεδίου. Η μεταβλητή d συμβολίζει τις διαστάσεις του χώρου στον οποίο ορίζεται το πεδίο. Η παράμετρος η_0 (συντελεστής κλίμακας) προσδιορίζει το μέγεθος των διακυμάνσεων στον χώρο, η παράμετρος ξ (μήκος συσχέτισης) την ακτίνα συσχέτισης και η παράμετρος η_1 (συντελεστής μορφής) την μορφή της χωρικής εξάρτησης. Οι τρεις όροι του ολοκληρώματος αντιστοιχούν στην διασπορά των τιμών του πεδίου γύρω από τη μέση τιμή, στην τοπική κλίση και στην ακτίνα καμπυλότητας (Hristopulos 2003).

Η χωρική εξάρτηση δύο σημείων εκφράζεται σε σχέση με τις παραμέτρους η_0 , η_1 , ξ και την συνάρτηση πυρήνα $\tilde{Q}_\lambda(\mathbf{k})$, όπου k η χωρική συχνότητα (κυματάνισμα), σύμφωνα με την εξίσωση της φασματικής πυκνότητας συνδιασποράς,

$$\tilde{G}_{\mathbf{x};\lambda}(\mathbf{k}) = \frac{|\tilde{Q}_\lambda(\mathbf{k})|^2 \eta_0 \xi^d}{1 + \eta_1 (k\xi)^2 + (k\xi)^4} \quad (2.3)$$

Το θεώρημα Bochner επιβάλλει την ισχύ ορισμένων συνθηκών αποδοχής έτσι ώστε το συναρτησιακό (2.2) να αντιστοιχεί σε επιτρεπτή χωρική εξάρτηση. Οι συνθήκες αυτές εκφράζονται με τις επόμενες ανισο-ισότητες.

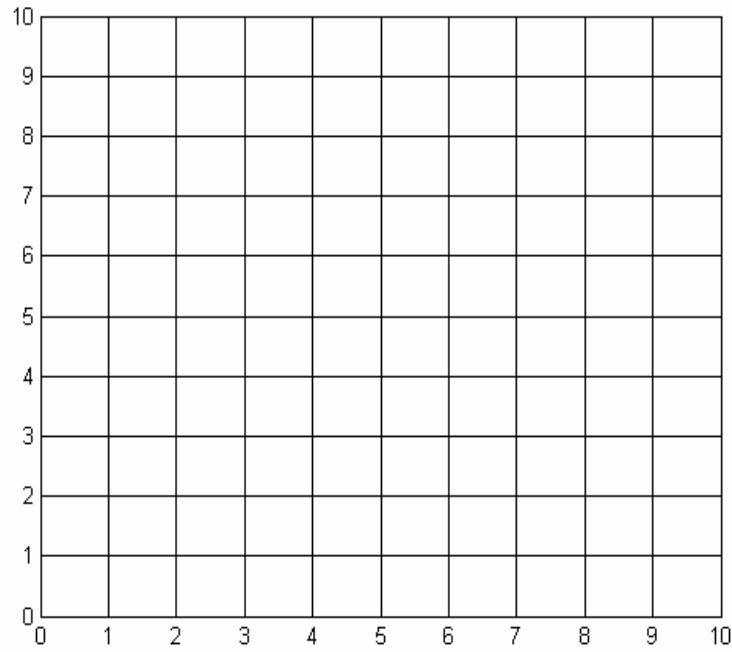
$$\begin{aligned} \eta_0 > 0, \xi > 0 \text{ και } \eta_1 &\geq 0 \\ \eta_0 > 0, \xi > 0 \text{ και } \eta_1 < 0, \eta_1^2 &< 4 \end{aligned}$$

Στον πραγματικό χώρο ωστόσο η συνάρτηση συνδιασποράς δίνεται από το παρακάτω ολοκλήρωμα

$$G_{\mathbf{x}}(\mathbf{r}) = \left(\frac{\eta_0 \xi^d \sqrt{\vartheta_d}}{r^{d/2-1}} \right) \int_0^\infty dk \frac{k^{d/2} J_{d/2-1}(kr)}{1 + \eta_1 (k\xi)^2 + (k\xi)^4}, \quad (2.4)$$

όπου $J_{d/2-1}(kr)$ είναι η συνάρτηση Bessel πρώτου είδους τάξης $\left(\frac{d}{2}-1\right)$ και ο συντελεστής $\vartheta_d = (2\pi)^{-d}$, (Hristopulos 2003).

Η εξίσωση (2.2) καθορίζει έναν Σπαρτιάτικο μοντέλο Τυχαίου Πεδίου στον συνεχή χώρο. Το αντίστοιχο τυχαίο πεδίο σε κανονικό πλέγμα περιλαμβάνει μόνο τις τιμές επί των σημείων του κανονικού πλέγματος (Σχήμα 2.1). Ένα ορθογώνιο κανονικό πλέγμα σε δύο διαστάσεις παράγεται από τα διανύσματα, $\hat{e}_1 = (1,0)$ οριζόντια διεύθυνση, $\hat{e}_2 = (0,1)$ κατακόρυφη διεύθυνση, με βήμα δικτυωτού πλέγματος a_1, a_2 .



Σχήμα 2.1: Απεικόνιση κανονικού πλέγματος (τετραγωνικό 10x10) το οποίο περιλαμβάνει τιμές ενός δείγματος καταναμημένες πάνω στα σημεία του.

Ως εκ τούτου, η συνάρτηση τροποποιείται ώστε να περιλαμβάνει πεπερασμένες διαφορές αντί των παραγώγων. Συνεπώς το συναρτησιακό αλληλεπίδρασης εκφράζεται ως εξής (Hristopulos 2003),

$$H[X_\lambda] = \frac{1}{2\eta_0\xi^2} \sum_{n=1}^N \left\{ [X_\lambda(\mathbf{s}_n)]^2 - [E[X_\lambda(\mathbf{s}_n)]]^2 + \eta_1\xi^2 \sum_{i=1}^2 \frac{[X_\lambda(\mathbf{s}_n + a_i\hat{e}_i) - X_\lambda(\mathbf{s}_n)]^2}{a_i^2} + \xi^4 \sum_{i=1}^2 \left[\frac{X_\lambda(\mathbf{s}_n + a_i\hat{e}_i) - 2X_\lambda(\mathbf{s}_n) + X_\lambda(\mathbf{s}_n - a_i\hat{e}_i)}{a_i^2} \right]^2 \right\} \quad (2.5)$$

2.3 Εκτίμηση παραμέτρων μοντέλου

Σε γεωστατιστικές εφαρμογές ο πρώτος στόχος είναι ο προσδιορισμός των παραμέτρων του ΣΤΠ βάσει των δεδομένων που υπάρχουν. Αυτό είναι ένα τυπικό αντίστροφο πρόβλημα. Ο καθορισμός των πρότυπων παραμέτρων απαιτεί πληροφορίες (πειραματικούς περιορισμούς) που μπορούν να υπολογιστούν αξιόπιστα από τα διαθέσιμα στοιχεία. Σε πολλές γεωστατιστικές εφαρμογές είναι δύσκολο να υπολογιστεί ακριβώς η μεγάλης ακτίνας χωρική εξάρτηση και συνεπώς το ημιβαριόγραμμα σε μεγάλες αποστάσεις. Επιπλέον, ο υπολογισμός του ημιβαριογράμματος περιλαμβάνει διάφορες υποθέσεις (όπως ο καθορισμός των τάξεων απόστασης, ο αριθμός σημείων ανά τάξη, κ.λ.π) και μπορεί να είναι υπολογιστικά πολύ εντατικός για μεγάλα δείγματα.

Σε τέτοιες περιπτώσεις έχει νόημα η εστίαση σε περιορισμένου φάσματος στοιχεία του τυχαίου πεδίου ώστε να διατυπωθούν οι κατάλληλοι περιορισμοί βασισμένοι στο πειραματικό δείγμα. Αν και αυτό δεν είναι απαραίτητο, είναι δυνατόν οι περιορισμοί να αντιστοιχούν στους όρους που περιέχονται στο συναρτησιακό της χωρικής αλληλεπίδρασης. Για παράδειγμα αν η χωρική αλληλεπίδραση δίνεται από την εξίσωση (2.2), είναι δυνατό οι στατιστικοί περιορισμοί να οριστούν ως η *διασπορά του δείγματος*, η *μέση τιμή του τετραγώνου της δειγματικής κλίσης* και η *μέση τιμή του τετραγώνου της καμπυλότητας*. Οι αντίστοιχες μέσες τιμές ως προς το σύνολο των καταστάσεων του πεδίου είναι δυνατόν να υπολογιστούν και θεωρητικά με την βοήθεια της συνάρτησης πυκνότητας πιθανότητας (2.1).

Η προσέγγιση αυτή οδηγεί στον καθορισμό των πρότυπων παραμέτρων «*ταιριάζοντας*» τους *πειραματικούς περιορισμούς* με τους αντίστοιχους πρότυπους *στοχαστικούς (θεωρητικούς) περιορισμούς*, οι οποίοι περιλαμβάνουν τις παραμέτρους του μοντέλου. Αυτό υποθέτει ότι το επιλεγμένο Σπαρτιάτικο Μοντέλο Τυχαίου Πεδίου αντιπροσωπεύει επαρκώς το πειραματικό δείγμα και είναι δυνατή η ανταλλαγή χωρικών και στοχαστικών μέσων όρων με την επίκληση της εργοδικής υπόθεσης. Για να ισχύει η εργοδική υπόθεση στον υπολογισμό των παραμέτρων του μοντέλου ΣΤΠ πρέπει να ισχύουν οι συνθήκες στατιστικής ομοιογένειας (μέση τιμή του πεδίου σταθερή και η συνάρτηση συνδιασποράς να εξαρτάται αποκλειστικά από

την απόσταση μεταξύ δύο σημείων) και το μήκος της περιοχής να υπερβαίνει την ακτίνα συσχέτισης κατά τουλάχιστον 30 φορές (Hristopulos 2003).

2.3.1 Μέθοδος κυψελίδων σε κανονικό πλέγμα

Ο υπολογισμός των περιορισμών εξαρτάται από τη χωρική κατανομή του πειραματικού δείγματος. Ειδικότερα εάν τα στοιχεία κατανέμονται σε ένα κανονικό πλέγμα η συνάρτηση πυκνότητας πιθανότητας (ΣΠΠ) ορίζεται με τη βοήθεια της εξίσωσης (2.5) και οι γείτονες κάθε σημείου (π.χ ο πλησιέστερος, δεύτερος πλησιέστερος, κ.λ.π.) προσδιορίζονται με ευκολία. Ως εκ τούτου, έχει νόημα να διατυπωθούν περιορισμοί βασισμένοι σε γραμμικούς συνδυασμούς των τιμών του πεδίου σε ομάδες γειτόνων.

Ο καθορισμός κοντινών γειτόνων είναι υπολογιστικά εντατικότερος για χωρικές κατανομές σε πλέγματα με άτακτη χωρική δομή. Στην απλούστερη προσέγγιση υπολογίζεται η απόσταση μεταξύ κάθε σημείου s και όλων των άλλων σημείων. Αυτός ο αλγόριθμος έχει υπολογιστική πολυπλοκότητα ίση με το τετράγωνο του συνολικού αριθμού των σημείων του δείγματος (N^2), και είναι του ίδιου βαθμού με την υπολογιστική πολυπλοκότητα που απαιτείται για τον υπολογισμό του ημιβαριογράμματος.

Υπάρχουν ωστόσο αποτελεσματικότερες προσεγγίσεις για τον υπολογισμό των ζευγών γειτόνων, όπως αυτές που είναι βασισμένες σε αποτελεσματικές υπολογιστικές δομές δέντρων (k-d trees), οι οποίες μειώνουν σημαντικά την πολυπλοκότητα υπολογισμού των κοντινότερων γειτόνων για δείγματα μεγάλων διαστάσεων (Hristopulos 2003).

Αρχικά θεωρείται ότι οι τιμές του δείγματος κατανέμονται πάνω σε ένα δισδιάστατο ορθογώνιο πλέγμα. Σε αυτήν την περίπτωση χρησιμοποιούνται οι κοντινότεροι γείτονες κάθε σημείου για να καθοριστούν οι πειραματικοί και στοχαστικοί περιορισμοί. Τόσο οι πειραματικοί όσο και οι στοχαστικοί περιορισμοί βασίζονται στα επόμενα μέτρα (συναρτησιακά) του πεδίου:

$$S_0 \equiv [X_\lambda(\mathbf{s})]^2 - \{E[X_\lambda(\mathbf{s})]\}^2 \quad (2.6)$$

$$S_1 \equiv \sum_{i=1}^2 \left[\frac{X_\lambda(\mathbf{s} + a_i \hat{\mathbf{e}}_i) - X_\lambda(\mathbf{s} - a_i \hat{\mathbf{e}}_i)}{2a_i} \right]^2 \quad (2.7)$$

$$S_2 \equiv \sum_{i,j=1}^2 \Delta_2^{(i)}[X_\lambda(\mathbf{s})] \Delta_2^{(j)}[X_\lambda(\mathbf{s})] \quad (2.8)$$

$$\text{Όπου } \Delta_2^{(i)}[X_\lambda(\mathbf{s})] = \frac{X_\lambda(\mathbf{s} + a_i \hat{\mathbf{e}}_i) + X_\lambda(\mathbf{s} - a_i \hat{\mathbf{e}}_i) - 2X_\lambda(\mathbf{s})}{a_i^2}$$

Ο πρώτος όρος μετρά το τετράγωνο της διακύμανσης (απόκλιση του τετραγώνου του πεδίου από το τετράγωνο της μέσης τιμής), ο δεύτερος όρος το τετράγωνο της τοπικής κλίσης, και ο τρίτος όρος ο οποίος υπολογίζεται από το άθροισμα, ως προς τις δύο ορθογώνιες διευθύνσεις, των γινομένων των διαφορών δεύτερης τάξης αντιπροσωπεύει το τετράγωνο της καμπυλότητας (Hristopulos 2003).

Οι στατιστικοί περιορισμοί του δείγματος δίνονται από το διάνυσμα $\bar{S} = (\bar{S}_0, \bar{S}_1, \bar{S}_2)$ τα στοιχεία του οποίου αντιστοιχούν στο μέσο δειγματικό όρο των S_0, S_1, S_2 και υπολογίζονται από τις παρακάτω εξισώσεις:

$$\bar{S}_0 = \frac{1}{N} \sum_{l=1}^N \left\{ X_\lambda(\mathbf{s}_l) - \left[\frac{1}{N} \sum_{l=1}^N X_\lambda(\mathbf{s}_l) \right] \right\}^2 \quad (2.9)$$

$$\bar{S}_1 = \frac{1}{N} \sum_{l=1}^N \left\{ \sum_{j=1}^2 \left[\frac{X_\lambda(\mathbf{s}_l + a_j \hat{\mathbf{e}}_j) - X_\lambda(\mathbf{s}_l - a_j \hat{\mathbf{e}}_j)}{2a_j} \right]^2 \right\} \quad (2.10)$$

$$\begin{aligned} \bar{S}_2 = \frac{1}{N} \sum_{l=1}^N \left\{ \sum_{i,j=1}^2 \left[\frac{X_\lambda(\mathbf{s}_l + a_i \hat{\mathbf{e}}_i) - 2X_\lambda(\mathbf{s}_l) + X_\lambda(\mathbf{s}_l - a_i \hat{\mathbf{e}}_i)}{a_i^2} \right] \right. \\ \left. \times \left[\frac{X_\lambda(\mathbf{s}_l + a_j \hat{\mathbf{e}}_j) - 2X_\lambda(\mathbf{s}_l) + X_\lambda(\mathbf{s}_l - a_j \hat{\mathbf{e}}_j)}{a_j^2} \right] \right\} \quad (2.11) \end{aligned}$$

Ενώ οι στατιστικοί περιορισμοί του μοντέλου (στοχαστικοί περιορισμοί) δίνονται από τις παρακάτω εξισώσεις:

$$E[S_0] = E\{[X_\lambda(\mathbf{s})]^2\} - E\{X_\lambda(\mathbf{s})\}^2 = \sigma_{x;\lambda}^2 \quad (2.12)$$

$$E[S_1] = \sum_{i=1}^2 \frac{2\sigma_{x;\lambda}^2 - G_{x;\lambda}(2a_i \hat{\mathbf{e}}_i) - G_{x;\lambda}(-2a_i \hat{\mathbf{e}}_i)}{4a_i^2} \quad (2.13)$$

$$E[S_2] = \sum_{i,j=1}^2 E\{\Delta_2^{(i)}[X_\lambda(\mathbf{s})] \Delta_2^{(j)}[X_\lambda(\mathbf{s})]\} \quad (2.14)$$

όπου $\sigma_{x;\lambda}^2$ είναι η διασπορά και $G_{x;\lambda}$ (14) η συνάρτηση συνδιασποράς του πεδίου (Hristopulos 2003).

Οι θεωρητικές μέσες τιμές περιλαμβάνουν τις τιμές των άγνωστων παραμέτρων η_0 , η_1 και ξ . Οι βέλτιστες παράμετροι υπολογίζονται επιλύοντας ένα πρόβλημα βελτιστοποίησης, δηλαδή βρίσκοντας τις τιμές $\hat{\eta}_0$, $\hat{\eta}_1$, $\hat{\xi}$ που ελαχιστοποιούν την διάφορα ανάμεσα στους δειγματικούς περιορισμούς και τους αντίστοιχους του μοντέλου. Οι τιμές των στατιστικών περιορισμών του δείγματος δεν αλλάζουν κατά την διαδικασία υπολογισμού των παραμέτρων ενώ οι θεωρητικοί περιορισμοί υπολογίζονται επαναληπτικά για κάθε σύνολο παραμέτρων το οποίο δοκιμάζεται από τον αλγόριθμο βελτιστοποίησης.

$$\Phi[X_\lambda(s)] \equiv \left| 1 - \sqrt{\frac{\bar{S}_1}{S_0} \frac{E[S_0]}{E[S_1]}} \right|^2 + \left| 1 - \sqrt{\frac{\bar{S}_2}{S_1} \frac{E[S_1]}{E[S_2]}} \right|^2 \quad (2.15)$$

Η μη αρνητική συνάρτηση (25) μετρά την απόσταση μεταξύ του μοντέλου και του δείγματος με βάση την απόκλιση των στοχαστικών περιορισμών $E[S_i]$ από τους αντίστοιχους \bar{S}_i ($i = 0, 1, 2$) περιορισμούς του δείγματος.

Οι βέλτιστες παράμετροι $\hat{\eta}_1$ και $\hat{\xi}$ καθορίζονται από την ελαχιστοποίηση της συναρτησιακής απόστασης $\Phi[X_\lambda(\mathbf{s})]$ ενώ η βέλτιστη παράμετρος κλίμακας δίνεται από τη σχέση:

$$\hat{\eta}_0 = \frac{\overline{S_0}}{E[S_0(\eta_0 = 1, \hat{\eta}_1, \hat{\xi})]} \quad (2.16)$$

Η συνάρτηση απόστασης $\Phi[X_\lambda(\mathbf{s})]$ εκφράζεται με βάση αδιάστατους συνδυασμούς των στοχαστικών περιορισμών οι οποίοι είναι ανεξάρτητοι του συντελεστή μεταβλητότητας η_0 . Αυτή η απλοποίηση, η οποία πραγματοποιείται επειδή ο συντελεστής η_0 πολλαπλασιάζει όλους τους όρους του συναρτησιακού αλληλεπίδρασης (2.2), μειώνει τον αριθμό των περιορισμών κατά ένα.

Επομένως, η βελτιστοποίηση της συνάρτησης $\Phi[X_\lambda(\mathbf{s})]$ υπολογίζει την παράμετρο η_1 που προσδιορίζει την μορφή της χωρικής εξάρτησης και την παράμετρο ξ που προσδιορίζει την ακτίνα συσχέτισης, αλλά δεν υπολογίζει την παράμετρο η_0 που προσδιορίζει την κλίμακα των διακυμάνσεων στο χώρο. Επειδή όμως η $E[S_0]$ μεταβάλλεται γραμμικά με την παράμετρο η_0 ισχύει η παρακάτω σχέση,

$$\eta_0 = \frac{E[S_0(\eta_0, \eta_1, \xi)]}{E[S_0(1, \eta_1, \xi)]} \quad (2.17)$$

η βέλτιστη τιμή της παραμέτρου $\hat{\eta}_0$ υπολογίζεται χρησιμοποιώντας τον περιορισμό $\overline{S_0} = E[S_0(\hat{\eta}_0, \hat{\eta}_1, \hat{\xi})]$ ο οποίος δίνει την εξίσωση (2.16) (Hristopulos 2003).

Η βελτιστοποίηση εφαρμόζεται αρχικά χρησιμοποιώντας τον αλγόριθμο simplex των Nelder και Mead. Ο αλγόριθμος δεν απαιτεί υπολογισμό της κλίσης των παραμέτρων της συνάρτησης, αλλά χρειάζεται αρχικές τιμές $\eta_{1,0}$ και ξ_0 για τις παραμέτρους του μοντέλου (Press et al. 1992). Η βελτιστοποίηση ολοκληρώνεται, όταν κατά την διαδικασία διαδοχικών βημάτων τόσο οι παράμετροι του μοντέλου όσο και η

απόσταση μεταξύ των δειγματικών και θεωρητικών στατιστικών περιορισμών έχουν διαφορά από τις προηγούμενες αντίστοιχες τιμές μικρότερη από ένα προκαθορισμένο όριο το οποίο για την συγκεκριμένη εφαρμογή ορίζεται 10^{-6} (Hristopulos 2003).

Εναλλακτικά χρησιμοποιείται ένας αλγόριθμος ο οποίος χρησιμοποιείται για επίλυση προβλημάτων μη γραμμικής βελτιστοποίησης υπό συνθήκες και υλοποιείται σε περιβάλλον Matlab® με την συνάρτηση «*fmincon*». Ο αλγόριθμος αυτός επίσης χρειάζεται αρχικές τιμές για τις παραμέτρους του μοντέλου ενώ επιτρέπει και τον σαφή ορισμό των συνθηκών αποδοχής του θεωρήματος Bochner. Οι συνθήκες τερματισμού της βελτιστοποίησης ορίζονται από ένα μέγιστο αριθμό επαναλήψεων του αλγορίθμου και μια βέλτιστη διαφορά μεταξύ των εκάστοτε αποτελεσμάτων τα οποία αντιπροσωπεύουν τις παραμέτρους του μοντέλου και την απόσταση μεταξύ των δύο κατηγοριών στατιστικών περιορισμών (Matlab 2002).

Στην συγκεκριμένη εφαρμογή ως μέγιστος αριθμός επαναλήψεων ορίστηκαν 1000, ενώ τόσο οι παράμετροι του μοντέλου όσο και η απόσταση μεταξύ των δειγματικών και θεωρητικών στατιστικών περιορισμών πρέπει να έχουν διαφορά, από τις προηγούμενες αντίστοιχες τιμές μικρότερη επίσης του 10^{-6} .

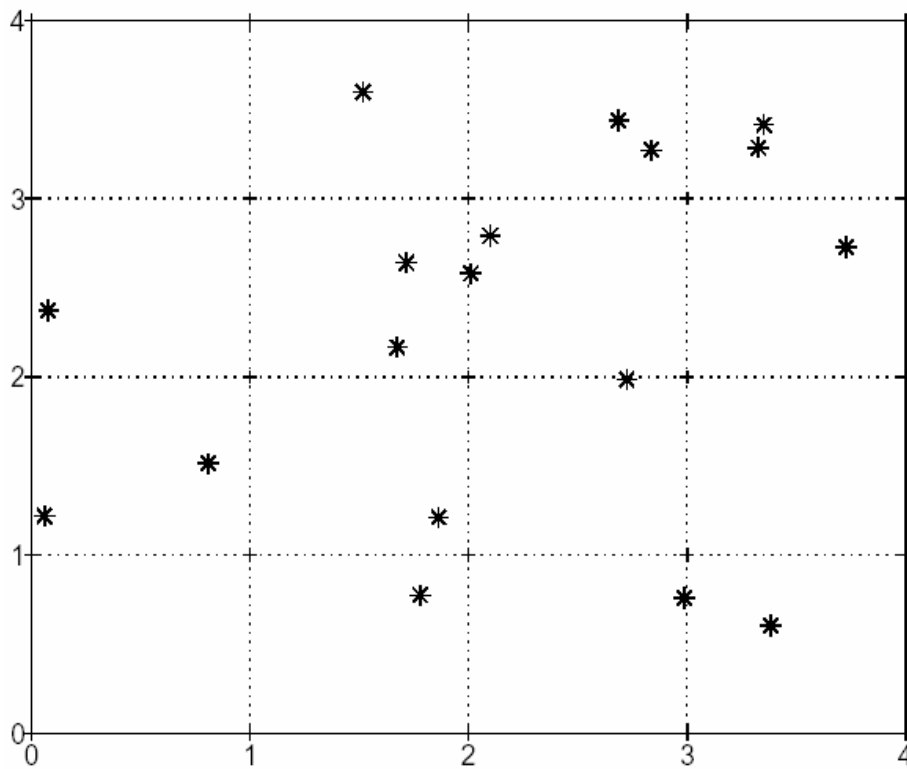
2.3.2 Μέθοδος κυψελίδων σε ανομοιογενές πλέγμα

Η θεωρητική ανάλυση του Σπαρτιάτικου μοντέλου Τυχαίου Πεδίου υπέθεσε ότι οι τιμές του πεδίου κατανέμονται πάνω σε ένα κανονικό ορθογώνιο πλέγμα. Ένα χαρακτηριστικό όμως των επίγειων παρατηρήσεων είναι ότι κατανέμονται ακανόνιστα. Σε αυτήν την περίπτωση, η συνάρτηση (2.2) και ο καθορισμός των στατιστικών περιορισμών του δείγματος απαιτούν περαιτέρω διερεύνηση. Οι πρακτικές επιλογές πρέπει να εξαρτηθούν από τη χωρική κατανομή των σημείων δειγματοληψίας. Η συνάρτηση (2.2) μπορεί να μετατραπεί ώστε να εκφράζεται χρησιμοποιώντας τους κοντινότερους γείτονες κάθε σημείου. (Hristopulos 2003).

Έστω μια τυχαία κατανομή των θέσεων δειγματοληψίας s_i , $i=1, \dots, N$. Ορίζεται ένα ορθογώνιο πλέγμα που καλύπτει την περιοχή του δείγματος. Λόγω απλότητας υποθέτονται τετραγωνικές κυψελίδες μοναδιαίου μήκους. Ο αριθμός των κόμβων ανά πλευρά του πλέγματος δίνεται από το $L = \sqrt{N}$. Οι θέσεις των κέντρων των

κυψελίδων ορίζονται από το z_k , όπου $k = 1, \dots, N_c$ είναι ένας δείκτης που αυξάνεται πρώτα κατά μήκος του y και στην συνέχεια κατά μήκος του x . Κάθε δείγμα θέσης s_i περιέχεται σε μια κυψελίδα $C(z_k)$ και κάθε κυψελίδα έχει έναν αριθμό $\eta(z_k)$, ο οποίος είναι ίσος με τον αριθμό των σημείων του δείγματος μέσα στην κυψελίδα. Ένα χαρακτηριστικό τετραγωνικό πλέγμα για μικρό αριθμό σημείων παρουσιάζεται στο σχήμα 2.2.

Οι πλησιέστεροι γείτονες της κυψελίδας $C(z_k)$ είναι οι κυψελίδες $C(z_k \pm \hat{e}_m)$, $m = x, y$. Η εξίσωση (2.5) μπορεί να χρησιμοποιηθεί τώρα για την μοντελοποίηση αλληλεπιδράσεων μεταξύ κυψελίδων. Οι στατιστικοί περιορισμοί του δείγματος περιλαμβάνουν ομάδες σημείων που ανήκουν σε γειτονικές κυψελίδες. Οι κυψελίδες μπορούν να περιλαμβάνουν σημεία περισσότερων του ενός ή μπορεί να είναι κενές. Ο αριθμός των σημείων στις γειτονικές κυψελίδες ποικίλλει ανάλογα με τη θέση της κυψελίδας και την κατεύθυνση των γειτόνων (Σχήμα 2.2).



Σχήμα 2.2: Απεικόνιση τετραγωνικού πλέγματος 4x4 που καλύπτει μικρό δείγμα (Hristopulos 2003).

Για να παρακολουθηθούν οι σχέσεις γειτονικών σημείων ορίζεται η παρακάτω συνάρτηση κυψελίδας:

Η δυαδική συνάρτηση $D_{k,m}(s_i, s_j)$ καθορίζει αν τα σημεία (s_i, s_j) είναι μέσα στις κοντινότερες γειτονικές κυψελίδες της κυψελίδας $C(\mathbf{z}_k)$ στην κατεύθυνση $\pm \hat{\mathbf{e}}_m$.

Η δυαδική μεταβλητή $Q_{k,mn} = (s_i, s_j, s_l)$ στην οποία $m \neq n$ μετρά τις ομάδες σημείων που αποτελούνται από μια κεντρική κυψελίδα $C(\mathbf{z}_k)$ και τους τέσσερις κοντινότερους γείτονες $C(\mathbf{z}_k \pm \hat{\mathbf{e}}_m), C(\mathbf{z}_k \pm \hat{\mathbf{e}}_n)$ έτσι ώστε κάθε κυψελίδα να περιέχει τουλάχιστον ένα σημείο δείγματος. Εάν $m = n$ προκύπτει μια συνάρτηση τριάδας η οποία μετρά τριάδες που αποτελούνται από μια κεντρική κυψελίδα και τους κοντινότερους γείτονες της στην κατεύθυνση $\pm \hat{\mathbf{e}}_m$.

Οι συναρτήσεις $D_{k,m}$ και $Q_{k,mn}$ ορίζονται παρακάτω:

$$D_{k,m}(s_i, s_j) = \begin{cases} 1, & \{s_i \in C(\mathbf{z}_k + \hat{\mathbf{e}}_m) \wedge s_j \in C(\mathbf{z}_k - \hat{\mathbf{e}}_m)\}_s \\ 0, & otherwise \end{cases} \quad (2.18)$$

Οι αγκύλες δηλώνουν $\{ \}_s$ ότι το περιεχόμενο τους πρέπει να λαμβάνει υπόψη όλους τους συνδυασμούς γειτονίας, δηλαδή τόσο την περίπτωση $s_i \in C(\mathbf{z}_k + \hat{\mathbf{e}}_m)$ και $s_j \in C(\mathbf{z}_k - \hat{\mathbf{e}}_m)$ όσο και την περίπτωση $s_j \in C(\mathbf{z}_k + \hat{\mathbf{e}}_m)$ και $s_i \in C(\mathbf{z}_k - \hat{\mathbf{e}}_m)$. Η δυαδική συνάρτηση παίρνει τιμή ίση με την μονάδα αν $s_j \in C(\mathbf{z}_k + \hat{\mathbf{e}}_m)$ και $s_i \in C(\mathbf{z}_k - \hat{\mathbf{e}}_m)$.

$$Q_{k,mn}(s_i, s_j, s_l) = \begin{cases} 1, & \{s_i \in C(\mathbf{z}_k + \hat{\mathbf{e}}_m) \wedge s_j \in C(\mathbf{z}_k - \hat{\mathbf{e}}_m) \wedge s_l \in C(\mathbf{z}_k)\}_s \\ 0, & otherwise \end{cases} \quad (2.19)$$

Οι αγκύλες δηλώνουν $\{ \}_s$ ότι το περιεχόμενο τους πρέπει να λαμβάνει υπόψη όλους τους συνδυασμούς γειτονίας, δηλαδή τις εξής περιπτώσεις:

- $\mathbf{s}_i \in C(\mathbf{z}_k + \hat{\mathbf{e}}_m)$, $\mathbf{s}_j \in C(\mathbf{z}_k - \hat{\mathbf{e}}_m)$ και $\mathbf{s}_l \in C(\mathbf{z}_k)$,
- $\mathbf{s}_i \in C(\mathbf{z}_k + \hat{\mathbf{e}}_m)$, $\mathbf{s}_l \in C(\mathbf{z}_k - \hat{\mathbf{e}}_m)$ και $\mathbf{s}_j \in C(\mathbf{z}_k)$,
- $\mathbf{s}_j \in C(\mathbf{z}_k + \hat{\mathbf{e}}_m)$, $\mathbf{s}_i \in C(\mathbf{z}_k - \hat{\mathbf{e}}_m)$ και $\mathbf{s}_l \in C(\mathbf{z}_k)$,
- $\mathbf{s}_j \in C(\mathbf{z}_k + \hat{\mathbf{e}}_m)$, $\mathbf{s}_l \in C(\mathbf{z}_k - \hat{\mathbf{e}}_m)$ και $\mathbf{s}_i \in C(\mathbf{z}_k)$,
- $\mathbf{s}_l \in C(\mathbf{z}_k + \hat{\mathbf{e}}_m)$, $\mathbf{s}_i \in C(\mathbf{z}_k - \hat{\mathbf{e}}_m)$ και $\mathbf{s}_j \in C(\mathbf{z}_k)$,
- $\mathbf{s}_l \in C(\mathbf{z}_k + \hat{\mathbf{e}}_m)$, $\mathbf{s}_j \in C(\mathbf{z}_k - \hat{\mathbf{e}}_m)$ και $\mathbf{s}_i \in C(\mathbf{z}_k)$,

Λαμβάνοντας υπόψη αυτούς τους ορισμούς οι στατιστικοί περιορισμοί του δείγματος αναδιαρθρώνονται ως εξής:

$$\overline{S_0} = \frac{1}{N} \sum_{i=1}^N \left\{ X_\lambda(\mathbf{s}_i) - \overline{X_\lambda(\mathbf{s}_i)} \right\}^2 \quad (2.20)$$

$$\overline{S_1} = \overline{S_{1,xx}} + \overline{S_{1,yy}} \quad (2.21)$$

$$\overline{S_{1,mm}} = \frac{1}{4N_{1,mm}} \sum_{k=1}^{N_C} \left\{ \sum_{\mathbf{s}_i \in C(\mathbf{z}_k + \hat{\mathbf{e}}_m)} \sum_{\mathbf{s}_j \in C(\mathbf{z}_k - \hat{\mathbf{e}}_m)} D_{k,m}(\mathbf{s}_i, \mathbf{s}_j) [X_\lambda(\mathbf{s}_i) - X_\lambda(\mathbf{s}_j)]^2 \right\} \quad (2.22)$$

$$N_{1,mm} = \sum_{\mathbf{z}_k} \eta(\mathbf{z}_k + \hat{\mathbf{e}}_m) \eta(\mathbf{z}_k - \hat{\mathbf{e}}_m), \quad m = x, y \quad (2.23)$$

$$\overline{S_2} = \overline{S_{2,xx}} + \overline{S_{2,yy}} + \overline{S_{2,xy}} \quad (2.24)$$

$$\begin{aligned} \overline{S_{2,mn}} = & \frac{1}{N_{2,mn}} \times \sum_{k=1}^{N_C} \left\{ \sum_{\mathbf{s}_i \in C(\mathbf{z}_k + \hat{\mathbf{e}}_m)} \sum_{\mathbf{s}_j \in C(\mathbf{z}_k - \hat{\mathbf{e}}_m)} \sum_{\mathbf{s}_l \in C(\mathbf{z}_k)} Q_{k,mn}(\mathbf{s}_i, \mathbf{s}_j, \mathbf{s}_l) \right. \\ & \times [X_\lambda(\mathbf{s}_i) + X_\lambda(\mathbf{s}_j) - 2X_\lambda(\mathbf{s}_l)] \\ & \times \left. \sum_{\mathbf{s}'_i \in C(\mathbf{z}_k + \hat{\mathbf{e}}_n)} \sum_{\mathbf{s}'_j \in C(\mathbf{z}_k - \hat{\mathbf{e}}_n)} \sum_{\mathbf{s}'_l \in C(\mathbf{z}_k)} Q_{k,mn}(\mathbf{s}'_i, \mathbf{s}'_j, \mathbf{s}'_l) \times [X_\lambda(\mathbf{s}'_i) + X_\lambda(\mathbf{s}'_j) - 2X_\lambda(\mathbf{s}'_l)] \right\} \end{aligned} \quad (2.25)$$

$$N_{2,mn} = \sum_{\mathbf{z}_k} \eta(\mathbf{z}_k) \eta(\mathbf{z}_k + \hat{\mathbf{e}}_m) \eta(\mathbf{z}_k - \hat{\mathbf{e}}_n) \times \sum_{\mathbf{z}_k} \eta(\mathbf{z}_k) \eta(\mathbf{z}_k + \hat{\mathbf{e}}_n) \eta(\mathbf{z}_k - \hat{\mathbf{e}}_m), \quad m = x, y \quad (2.26)$$

Οι εξισώσεις που χρησιμοποιήθηκαν για τον προσδιορισμό των στατιστικών περιορισμών του μοντέλου, μπορούν να χρησιμοποιηθούν και σε αυτή την περίπτωση όπως και οι μέθοδοι βελτιστοποίησης που περιγράφηκαν παραπάνω για τον υπολογισμό των παραμέτρων. (Hristopulos 2003).

Ο υπολογισμός των στατιστικών περιορισμών που αντιπροσωπεύουν την διασπορά, το τετράγωνο της κλίσης και της καμπυλότητας και η επίλυση του προβλήματος βελτιστοποίησης είναι σημαντικά ταχύτερος από τον υπολογισμό του ημιβαριογράμματος. Τα αίτια της διαφοράς αυτής οφείλονται στην διαφορετική υπολογιστική πολυπλοκότητα των δυο μεθόδων.

Αν το δείγμα περιέχει τιμές σε N σημεία ο υπολογιστικός χρόνος για το ημιβαριόγραμμα μεταβάλλεται ανάλογα με το N^2 , χωρίς να ληφθεί υπόψη ο απαιτούμενος χρόνος για τον υπολογισμό των παραμέτρων του μοντέλου, προσαρμόζοντας το πειραματικό ημιβαριόγραμμα σε θεωρητικά μοντέλα ημιβαριογράμματος. Αντίθετα ο χρόνος υπολογισμού των στατιστικών περιορισμών στην μέθοδο των Σπαρτιάτικων Τυχαίων Πεδίων μεταβάλλεται ανάλογα με το N .

Στην περίπτωση που τα σημεία βρίσκονται κατανεμημένα πάνω σε ένα ορθογώνιο κανονικό πλέγμα ο υπολογισμός των στατιστικών περιορισμών του δείγματος είναι ταχύτερος σε σχέση με την περίπτωση που τα σημεία είναι κατανεμημένα άτακτα σε ένα ορθογώνιο πλέγμα και σε πολλές περιπτώσεις η διαφορά είναι δυο τάξεις μεγέθους. Ωστόσο και σε αυτήν την περίπτωση ο υπολογισμός των στατιστικών περιορισμών είναι σημαντικά ταχύτερος από τον υπολογισμό του ημιβαριογράμματος. Επίσης, όσον αφορά στο πρόβλημα βελτιστοποίησης, η επίλυση του είναι ανεξάρτητη από το μέγεθος του δείγματος (Hristopulos 2003).

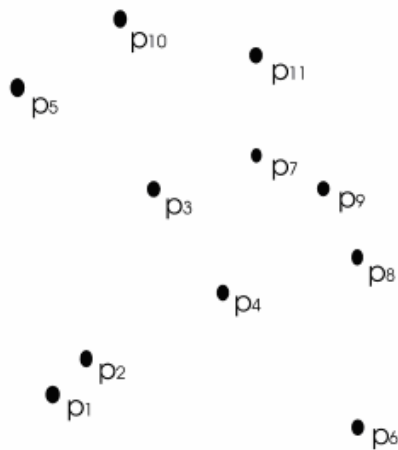
2.3.3 Μέθοδος τριγώνων Delaunay

Στις προηγούμενες δύο παραγράφους διατυπώθηκε ο τρόπος υπολογισμού των στατιστικών περιορισμών δείγματος και προσδιορισμού των παραμέτρων του

μοντέλου ΣΤΧ για κανονική και άτακτη κατανομή δειγματικών σημείων. Σε αυτή την παράγραφο θα περιγραφεί ο τρόπος υπολογισμού των στατιστικών περιορισμών του δείγματος σε μη κανονικό πλέγμα με την βοήθεια των τριγώνων Delaunay.

Όπως προαναφέρθηκε, τα πειραματικά δείγματα είναι συνήθως άτακτα κατανεμημένα. Για αυτό το λόγο η εκτίμηση των παραμέτρων με την βοήθεια ενός υποβάθρου ορθογώνιου πλέγματος μπορεί να οδηγήσει σε ανακριβή αποτελέσματα, καθώς υπάρχουν κυψελίδες με αρκετά σημεία δείγματος και άλλες οι οποίες είναι κενές. Αυτό έχει ως αποτέλεσμα την περιορισμένη ακρίβεια υπολογισμού των στατιστικών περιορισμών του δείγματος καθώς παρουσιάζονται σημεία με αρκετούς γείτονες στις διπλανές κυψελίδες ή με πολύ λίγους, ενώ μερικές φορές γείτονες υπάρχουν μόνο σε μια διεύθυνση.

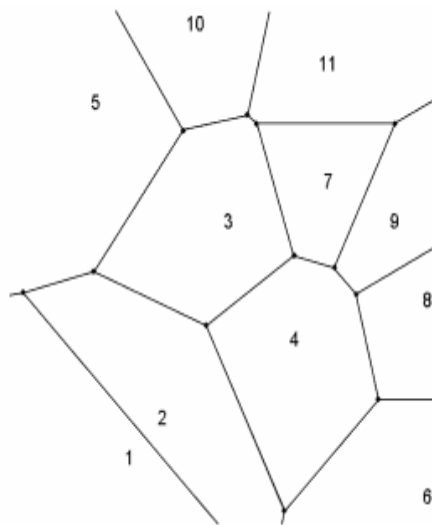
Για να προσδιοριστούν λοιπόν οι στατιστικοί περιορισμοί του δείγματος με όσο το δυνατόν μεγαλύτερη ακρίβεια εφαρμόστηκε η μέθοδος των τριγώνων Delaunay η οποία επικεντρώνεται στον προσδιορισμό των πλησιέστερων γειτόνων κάθε σημείου. Σε αυτήν περίπτωση για τον υπολογισμό των στατιστικών περιορισμών δεν χρειάζεται ο προσδιορισμός των πλησιέστερων γειτονικών σημείων σε γειτονικές κυψελίδες και η αλληλεπίδραση κυψελίδων, ώστε να είναι χρήσιμος ο αριθμός των σημείων αυτών στην καθεμία, αλλά μόνο σημεία που αντιπροσωπεύουν τους πλησιέστερους γείτονες.



Σχήμα 2.3: Δειγματικά σημεία τυχαίου πεδίου (Mulchrone 2002).

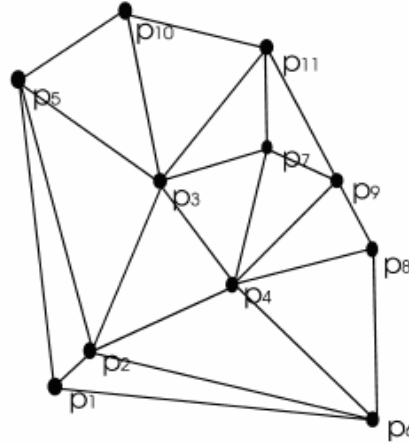
Τα τρίγωνα Delaunay αποτελούνται από ευθείες γραμμές που συνδέουν τους πλησιέστερους γείτονες των δειγματικών σημείων. Τα τρίγωνα Delaunay δημιουργούνται με τον εξής τρόπο: Αρχικά, θεωρείται ένα δείγμα τυχαίου πεδίου N δειγματικών σημείων (κόμβων), $P = \{p_1, p_2, p_3, \dots, p_N\}$ (Σχήμα 2.3).

Στην συνέχεια η περιοχή του δείγματος διαχωρίζεται σε πολύγωνα έτσι ώστε κάθε πολύγωνο να περιλαμβάνει ένα δειγματικό σημείο (κόμβο) p_i . Κάθε σημείο εντός ενός πολυγώνου βρίσκεται πιο κοντά στο δειγματικό σημείο (κόμβο) που αντιστοιχεί στο πολύγωνο, από οποιαδήποτε άλλο. Τα πολύγωνα αυτά που δημιουργούνται ονομάζονται πολύγωνα Voronoi. Ένα διάγραμμα Voronoi αποτελείται από όλα τα πολύγωνα που αντιστοιχούν στο σύνολο των δειγματικών σημείων P (Σχήμα 2.4).



Σχήμα 2.4: Διάγραμμα πολυγώνων Voronoi (Mulchrone 2002).

Όταν σχεδιαστεί το διάγραμμα Voronoi τα τρίγωνα Delaunay προκύπτουν συνδέοντας τον κάθε κόμβο p_i με αυτούς των οποίων τα πολύγωνα Voronoi συνορεύουν (Σχήμα 2.5). Αυτοί οι κόμβοι αποτελούν και τους πλησιέστερους γείτονες του p_i . (Mulchrone 2002).



Σχήμα 2.5: Παράδειγμα δημιουργίας τριγώνων Delaunay (Mulchrone 2002).

Οι στατιστικοί περιορισμοί του δείγματος υπολογίζονται μετά την εφαρμογή της μεθόδου Delaunay και τον προσδιορισμό των κοντινότερων γειτόνων. Η μέση τιμή του τετραγώνου της δειγματικής κλίσης S_1 και η μέση τιμή του τετραγώνου της καμπυλότητας S_2 του τυχαίου πεδίου υπολογίζονται χρησιμοποιώντας την μέθοδο των πεπερασμένων διαφορών (Finite Difference Method) για τον υπολογισμό τοπικών χωρικών παραγώγων. Η μέθοδος βασίζεται στο ανάπτυγμα σε σειρά Taylor.

Πιο συγκεκριμένα, θεωρείται το σημείο s_i όπου $i=1, \dots, N$ το οποίο έχει n πλησιέστερους γείτονες $s_j, j=1, \dots, n$. Οι πλησιέστεροι γείτονες τους είναι τα σημεία που συνδέονται με το s_i διαμέσου των πλευρών των τριγώνων Delaunay. Αν υποθεθεί ότι το πεδίο $X(s)$ είναι διαφορίσιμο, το ανάπτυγμα Taylor στα σημεία s_j γύρω από το σημείο s_i δίνεται από την παρακάτω εξίσωση,

$$\begin{aligned} X(s_j) = & X(s_i) + \frac{\partial X(s_i)}{\partial x} \delta x_j + \frac{\partial X(s_i)}{\partial y} \delta y_j + \frac{1}{2} \frac{\partial^2 X(s_i)}{\partial x^2} (\delta x_j)^2 \\ & + \frac{1}{2} \frac{\partial^2 X(s_i)}{\partial y^2} (\delta y_j)^2 + \frac{\partial^2 X(s_i)}{\partial x \partial y} (\delta x_j)(\delta y_j) + R_3 \end{aligned} \quad (2.27)$$

όπου R_3 συμβολίζει όρους παραγώγων υψηλότερης τάξης.

Σε αυτήν την εξίσωση $\delta x_j = x_j - x_i$ και $\delta y_j = y_j - y_i$, ενώ η διαφορά μεταξύ των τιμών του τυχαίου πεδίου στο σημείο s_i και των γειτόνων του, $\delta X_{j,i} = X(s_j) - X(s_i)$, μπορεί να εκφραστεί με την μορφή του παρακάτω γραμμικού συστήματος:

$$\begin{bmatrix} \delta X_{1,i} \\ \delta X_{2,i} \\ \dots \\ \delta X_{n,i} \end{bmatrix} = \begin{bmatrix} \delta x_1 & \delta y_1 & \delta x_1^2/2 & \delta y_1^2/2 & \delta x_1 \delta y_1 \\ \delta x_2 & \delta y_2 & \delta x_2^2/2 & \delta y_2^2/2 & \delta x_2 \delta y_2 \\ \dots & \dots & \dots & \dots & \dots \\ \delta x_n & \delta y_n & \delta x_n^2/2 & \delta y_n^2/2 & \delta x_n \delta y_n \end{bmatrix} \begin{bmatrix} \partial_x X(s_i) \\ \partial_y X(s_i) \\ \partial_{xx}^2 X(s_i) \\ \partial_{yy}^2 X(s_i) \\ \partial_{xy}^2 X(s_i) \end{bmatrix}. \quad (2.28)$$

Η παραπάνω εξίσωση μπορεί επίσης να εκφραστεί ως

$$\delta X = V D_X \quad (2.29)$$

όπου ο $n \times 1$ πίνακας δX περιλαμβάνει τις διαφορές των τιμών του πεδίου μεταξύ του κεντρικού σημείου και των γειτόνων του, V είναι ο $n \times 5$ πίνακας της διαφοράς των συντεταγμένων και D_X είναι ο 5×1 πίνακας των αγνώστων χωρικών παραγώγων. Το γραμμικό σύστημα (2.29) είναι απροσδιόριστο για $n < 5$ και υπέρπροσδιορισμένο για $n > 5$. Μπορεί να επιλυθεί σε περιβάλλον Matlab[®] χρησιμοποιώντας την εντολή της διαίρεσης πινάκων, $D_X = V \setminus \delta X$ (Hristopulos 2005a). Η εντολή της διαίρεσης πινάκων χρησιμοποιείται για την επίλυση γραμμικών συστημάτων που εκφράζονται με εξισώσεις της μορφής $AX=B$, όπως η εξίσωση (2.29). Αν ο πίνακας V είναι της μορφής $(n \times n)$ και δX ένας πίνακας $(n \times 1)$ τότε η λύση της εξίσωσης (2.29) $D_X = V \setminus \delta X$ δίνεται χρησιμοποιώντας την μέθοδο Gauss. Αν όμως ο πίνακας V είναι της μορφής $(m \times n)$ και δX ένας πίνακας $(m \times 1)$ τότε η λύση $D_X = V \setminus \delta X$ επιτυγχάνεται με την μέθοδο των ελαχίστων τετραγώνων (Matlab 2002).

Το γραμμικό σύστημα πρέπει να λυθεί σε κάθε σημείο του δείγματος για να υπολογιστούν οι τοπικές παράγωγοι. Σημεία τα οποία έχουν πολύ λίγους γείτονες, πολύ κοντινούς γείτονες, ή γείτονες σε μεγάλες αποστάσεις μπορεί να προκαλέσουν αριθμητικές αστάθειες. Σε τέτοιες περιπτώσεις οι παράγωγοι δεν θα υπολογιστούν σε

αυτά τα σημεία. Μια άλλη εναλλακτική λύση είναι να υπολογιστεί μόνο η τοπική κλίση στα σημεία τα οποία έχουν μόνο ένα μικρό αριθμό γειτόνων. Επίσης πρέπει να μελετηθεί αν γείτονες μεγάλης απόστασης πρέπει να εξαιρεθούν. Σε αυτή την περίπτωση ο υπολογισμός των χωρικών παραγώγων θα βασιστεί στους ενεργούς γείτονες, δηλαδή σε αυτούς των οποίων η απόσταση από το κεντρικό σημείο δεν ξεπερνά ένα συγκεκριμένο άνω όριο.

Στην συνέχεια για να υπολογιστούν οι στατιστικοί περιορισμοί του δείγματος που απαιτούνται από την μέθοδο των Σπαρτιάτικων Τυχαίων Πεδίων οι τοπικές παράγωγοι κάθε σημείου του δείγματος εισάγονται στις παρακάτω εξισώσεις:

$$S_1 = \frac{1}{N_G} \sum_{i=1}^{N_G} \left\{ [\partial_x X(s_i)]^2 + [\partial_y X(s_i)]^2 \right\} \quad (2.30)$$

$$S_2 = \frac{1}{N_L} \sum_{i=1}^{N_L} \left\{ [\partial_{xx}^2 X(s_i)]^2 + [\partial_{yy}^2 X(s_i)]^2 + [\partial_{xx}^2 X(s_i) \partial_{yy}^2 X(s_i)] \right\} \quad (2.31)$$

όπου N_G είναι ο ενεργός αριθμός των σημείων που χρησιμοποιήθηκαν για τον υπολογισμό της χωρικής παραγώγου πρώτου βαθμού και N_L ο αριθμός των σημείων που χρησιμοποιήθηκαν για τον υπολογισμό της χωρικής παραγώγου δευτέρου βαθμού. Ο στατιστικός περιορισμός S_0 υπολογίζεται και σε αυτήν την περίπτωση σύμφωνα με τη εξίσωση (2.6) καθώς περιγράφει την διασπορά του δείγματος ενώ δεν προϋποθέτει τον προσδιορισμό των πλησιέστερων γειτόνων (Hristopulos 2005a).

Όσον αφορά στους στατιστικούς περιορισμούς του μοντέλου προσδιορίζονται και σε αυτήν την περίπτωση από τις εξισώσεις (2.12, 2.13, 2.14). Η εκτίμηση των παραμέτρων πραγματοποιείται με την μέθοδο βελτιστοποίησης που περιγράφηκε στην παράγραφο 2.3.1.

2.3.4 Επιβεβαίωση της μεθόδου Delaunay

Για να είναι δυνατή η χρησιμοποίηση με ακρίβεια της μεθόδου *Delaunay* στον προσδιορισμό των στατιστικών περιορισμών ενός δείγματος και στην συνέχεια στον

υπολογισμό των παραμέτρων του μοντέλου ΣΤΠ είναι απαραίτητος ο έλεγχος της μεθόδου σε διάφορα δείγματα τυχαίων πεδίων και η επιβεβαίωση των αποτελεσμάτων της.

Η επιβεβαίωση των αποτελεσμάτων της μεθόδου *Delaunay* θα πραγματοποιηθεί συγκρίνοντας τα αποτελέσματα της μεθόδου με αριθμητικούς υπολογισμούς στην Matlab® και με αναλυτικές λύσεις που υπολογίζουν την μέση τιμή του τετραγώνου της κλίσης και της καμπυλότητας για δείγματα αιτιοκρατικών συναρτήσεων. Για τον σκοπό αυτόν θα χρησιμοποιηθούν συνθετικά δείγματα καταναμημένα σε κανονικό πλέγμα.

Πιο συγκεκριμένα για να πραγματοποιηθεί η επιβεβαίωση της μεθόδου, αρχικά προσδιορίζονται οι στατιστικοί περιορισμοί του δείγματος με την μέθοδο *Delaunay* χρησιμοποιώντας τον κώδικα «*delaunay_application*» (Παράρτημα Γ). Με την βοήθεια εντολών της Matlab® υπολογίζονται η μέση τιμή του τετραγώνου της κλίσης (εντολή «grad») και η μέση τιμή του τετραγώνου της καμπυλότητας (εντολή «del2») όταν το δείγμα είναι καταναμημένο σε κανονικό πλέγμα.

Και οι δύο εντολές βασίζονται στην μέθοδο των πεπερασμένων διαφορών (Matlab 2002). Στην συνέχεια πραγματοποιείται η σύγκριση και αξιολόγηση των αποτελεσμάτων. Έπεται μια δεύτερη σύγκριση μεταξύ των στατιστικών περιορισμών, οι οποίοι προέκυψαν με την μέθοδο *Delaunay* για δείγματα καταναμημένα σε άτακτο πλέγμα.

Το πρώτο συνθετικό δείγμα που θα εξεταστεί είναι καταναμημένο σε κανονικό πλέγμα διαστάσεων 100×100 και προέρχεται από μια περιοδική συνάρτηση της μορφής $\cos(k_x x + k_y y)$ όπου, $\mathbf{k} = k_x, k_y$ είναι το κυματόνυσμα το οποίο είναι ίσο με

$$k_x = \frac{2\pi}{J_x} \text{ και } k_y = \frac{2\pi}{J_y}, \text{ ενώ } J_x = \frac{L}{10} \text{ και } J_y = \frac{L}{5}. \text{ Η μεταβλητή } L \text{ δηλώνει το μήκος}$$

του πλέγματος στις δύο διαστάσεις.

Σε αυτή την περίπτωση είναι εφικτή και η απευθείας επιβεβαίωση των στατιστικών περιορισμών καθώς είναι δυνατό να υπολογιστούν αναλυτικά οι ποσότητες \bar{S}_1, \bar{S}_2 .

Αυτό επιτυγχάνεται υπολογίζοντας την μέση τιμή του τετραγώνου της πρώτης παραγώγου της περιοδικής συνάρτησης, $[f] = \frac{1}{L^2} \int_0^L dx \int_0^L dy (k_x^2 + k_y^2) \sin^2(k_x x + k_y y)$ αλλά και της δεύτερης $[f] = \frac{1}{L^2} \int_0^L dx \int_0^L dy (k_x^2 + k_y^2)^2 \cos^2(k_x x + k_y y)$, από την οποία προέρχεται το δείγμα (Παράρτημα Α).

Στον πίνακα 1 παρουσιάζονται τα αποτελέσματα της επιβεβαίωσης της μεθόδου για το πρώτο δείγμα.

Πίνακας 2.1: Αποτελέσματα επιβεβαίωσης της μεθόδου Delaunay για δείγμα τυχαίου πεδίου περιοδικής συνάρτησης.

Δείγμα Συνημίτονου	S ₁	S ₂
Μέθοδος Τριγώνων Delaunay	0.21	0.11
Μέθοδος Πεπερασμένων Διαφορών	0.22	0.11
Αναλυτικές Λύσεις	0.24	0.12

Σύμφωνα με τον πίνακα 2.1 συγκρίνοντας τα αποτελέσματα της μεθόδου *Delaunay* με τα αποτελέσματα τόσο των αριθμητικών υπολογισμών όσο και των αναλυτικών λύσεων παρατηρείται μια πολύ καλή συμφωνία μεταξύ τους. Συνεπώς ο υπολογισμός των στατιστικών περιορισμών του συγκεκριμένου δείγματος με την μέθοδο Τριγώνων *Delaunay* αποδεικνύεται ότι είναι ακριβής.

Το δεύτερο δείγμα αποτελείται από δύο καταστάσεις διαφορετικών τυχαίων πεδίων κατανεμημένων σε κανονικό πλέγμα διαστάσεων 100×100. Πιο συγκεκριμένα το δείγμα αποτελείται από δύο (ΣΤΠ) τα οποία προσδιορίζονται από συγκεκριμένες παραμέτρους. Το πρώτο ΣΤΠ προσδιορίζεται από τις παραμέτρους $\eta_0 = 1$, $\eta_1 = 0.2$ και $\xi = 5$ ενώ το δεύτερο από τις παραμέτρους $\eta_0 = 1$, $\eta_1 = -0.5$ και $\xi = 5$. Τα αποτελέσματα της επιβεβαίωσης παρουσιάζονται στους πίνακες 2 και 3.

Πίνακας 2.2: Αποτελέσματα επιβεβαίωσης της μεθόδου Delaunay για το πρώτο ΣΤΧ.

Κατάσταση 1 ^η	S ₁	S ₂
Μέθοδος Τριγώνων Delaunay	0.011	0.017
Μέθοδος Πεπερασμένων Διαφορών	0.012	0.021

Πίνακας 2.3: Αποτελέσματα επιβεβαίωσης της μεθόδου Delaunay για το δεύτερο ΣΤΧ.

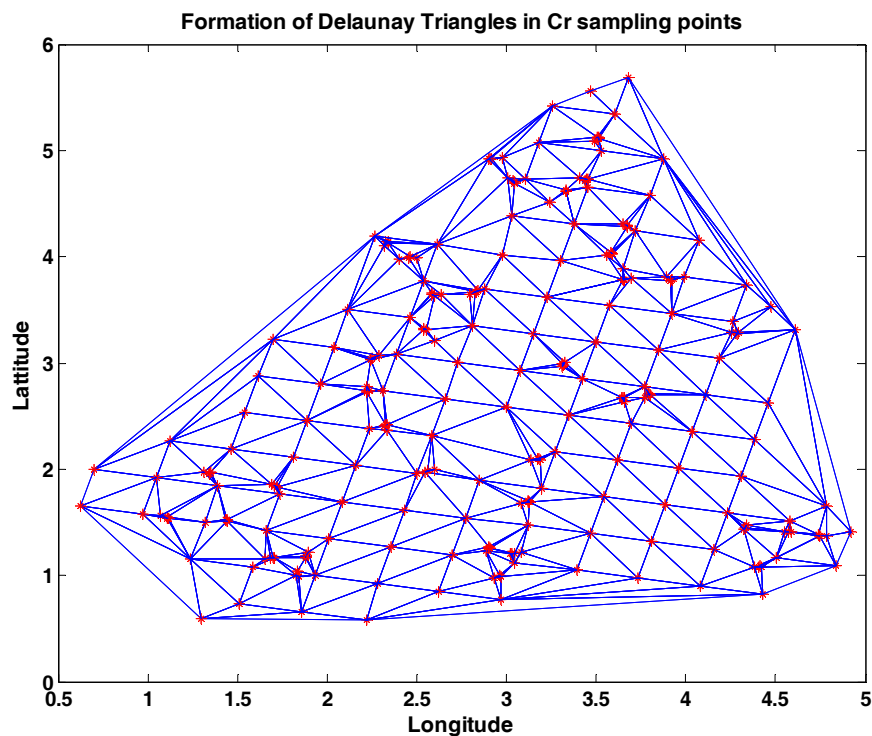
Κατάσταση 2 ^η	S ₁	S ₂
Μέθοδος Τριγώνων Delaunay	0.011	0.017
Μέθοδος Πεπερασμένων Διαφορών	0.015	0.021

Όπως παρατηρείται από τους πίνακες 2.2 και 2.3 και για αυτό το δείγμα η μέθοδος *Delaunay* δείχνει να υπολογίζει του στατιστικούς περιορισμούς με μεγάλη ακρίβεια καθώς τα αποτελέσματα της συγκρινόμενα με τα αποτελέσματα των αριθμητικών υπολογισμών με πεπερασμένες διαφορές έχουν πολύ μικρή διαφορά μεταξύ τους.

Τέλος, πραγματοποιείται μια νέα σύγκριση μεταξύ των αποτελεσμάτων της μεθόδου *Delaunay*. Η σύγκριση πραγματοποιείται μεταξύ των στατιστικών περιορισμών που υπολογίζονται με την μέθοδο *Delaunay* για το δείγμα συγκεντρώσεων χρωμίου και αυτών που υπολογίζονται με την ίδια μέθοδο για ένα δείγμα κατανεμημένο σε ανομοιογενές πλέγμα. Το δείγμα αυτό περιγράφει δύο καταστάσεις διαφορετικών τυχαίων πεδίων.

Πιο συγκεκριμένα είναι κατανεμημένο στις ίδιες θέσεις με το δείγμα (*Cr*) που εξετάζεται σε αυτήν την εργασία (Σχήμα 2.6). Τα δύο τυχαία πεδία έχουν τις ίδιες στατιστικές ιδιότητες, δηλαδή ακολουθούν την γκαουσιανή κατανομή, έχουν μέση τιμή $E[X(s)] = 0$, τυπική απόκλιση $\sigma = 1$, ακτίνα συσχέτισης $\xi = 0.1$ και λόγο ανισοτροπίας είναι $\rho_{y/x} = 1$. Ωστόσο το πρώτο τυχαίο πεδίο έχει γκαουσιανή συνάρτηση συσχέτισης ενώ το δεύτερο εκθετική. Τα δύο τυχαία πεδία δημιουργούνται από τον κώδικα «rf2dirreg» (Παράρτημα Δ) σε περιβάλλον Matlab[®].

Η σύγκριση των αποτελεσμάτων της μεθόδου *Delaunay* για τα δύο διαφορετικά τυχαία πεδία με αυτά για το δείγμα συγκεντρώσεων χρωμίου δεν αποτελεί επιβεβαίωση του προσδιορισμού των στατιστικών περιορισμών του δείγματος χρωμίου αλλά μια ποιοτική εκτίμηση της μεθόδου. Η επιβεβαίωση των αποτελεσμάτων της μεθόδου *Delaunay* σε αυτήν την περίπτωση δεν είναι δυνατή καθώς προς το παρόν δεν υπάρχει άλλη μέθοδος υπολογισμού των στατιστικών περιορισμών όταν το δείγμα είναι κατανεμημένο σε ανομοιογενές πλέγμα.



Σχήμα 2.6: Εφαρμογή της μεθόδου τριγώνων Delaunay στα σημεία μέτρησης του δείγματος χρωμίου (Cr).

Οι στατιστικοί περιορισμοί των δειγμάτων δεν αναμένεται να συμφωνούν καθώς τα τρία δείγματα είναι διαφορετικά, ωστόσο επειδή είναι κατανομημένα στο ίδιο ανομοιογενές πλέγμα είναι δυνατή μια πρώτη σύγκριση της τάξης μεγέθους των αποτελεσμάτων (Πίνακας 2.4). Τα αποτελέσματα της μεθόδου για το δείγμα χρωμίου τα οποία δίνονται στον πίνακα 2.4 έχουν διαιρεθεί με την διασπορά του δείγματος καθώς αυτή ισούται με $\sigma^2 = 120$, ενώ για τα δύο τυχαία πεδία ισούται με $\sigma^2 = 1$.

Πίνακας 2.4: Αποτελέσματα επιβεβαίωσης της μεθόδου Delaunay για δείγμα τυχαίων πεδίων κατανομημένων σε ανομοιογενές πλέγμα.

Μέθοδος Delaunay	S_1	S_2
1 ^ο Δείγμα Τυχαίου Πεδίου	91	1600
2 ^ο Δείγμα Τυχαίου Πεδίου	135	4600
Δείγμα Cr	59	3898

Από τον παραπάνω πίνακα διαπιστώνεται ότι οι τιμές των στατιστικών περιορισμών του δείγματος συγκεντρώσεων χρωμίου είναι της ίδιας τάξης μεγέθους με αυτές των δύο τυχαίων πεδίων. Στο τρίτο κεφάλαιο όπου θα παρουσιαστούν οι τιμές των

στατιστικών περιορισμών του δείγματος χρωμίου υπολογισμένες σύμφωνα με την μέθοδο κυψελίδων θα πραγματοποιηθεί σύγκριση των αποτελεσμάτων των δύο μεθόδων (3.6.1).

2.4 Εκτίμηση πεδίου σε σημεία χωρίς μετρήσεις

Όταν έχουν υπολογιστεί οι παράμετροι του χωρικού μοντέλου ακολουθεί η εκτίμηση του πεδίου σε σημεία του χώρου στα οποία δεν υπάρχουν δείγματα. Αν υποθεθεί ένα σημείο εκτίμησης \mathbf{s}_0 , απαιτείται η εύρεση της βέλτιστης τιμής $\hat{X}_\lambda(\mathbf{s}_0)$ η οποία επιτυγχάνεται μεγιστοποιώντας την κοινή συνάρτηση πυκνότητας πιθανότητας $f_x[X_\lambda(\mathbf{s}_0) | \mathbf{X}_\lambda(S)]$ όπου $\mathbf{X}_\lambda(S)$ είναι ένα διάνυσμα που αναπαριστά τις τιμές του πεδίου στα σημεία μέτρησης (Hristopulos 2005b).

Αν τα σημεία του πεδίου δεν είναι κατανεμημένα σε κανονικό πλέγμα η εκτίμηση μπορεί να βασιστεί στην συνεχή συνάρτηση πυκνότητας πιθανότητας (11) ακολουθώντας τα επόμενα βήματα (Hristopulos 2005b).

1. Θεωρείται η παράγωγος μεταβολής της συνάρτησης (2.2) ίση με το μηδέν (Mathematica 2005). Έτσι προκύπτει μια εξίσωση (2.32), για την διακύμανση στην γειτονιά του σημείου εκτίμησης,

$$\chi_\lambda(\mathbf{s}) - \eta_1 \xi^2 \nabla^2 \chi_\lambda(\mathbf{s}) + \xi^4 \nabla^4 \chi_\lambda(\mathbf{s}) = 0 \quad (2.32)$$

Υπολογίζοντας την διακύμανση είναι δυνατή η εκτίμηση στο σημείο αυτό βάσει της εξίσωσης,

$$\hat{X}_\lambda(\mathbf{s}) = m_X + \hat{\chi}(\mathbf{s}). \quad (2.33)$$

2. Στη συνέχεια επιλύεται η μερική διαφορική εξίσωση (2.32) τοπικά ως προς $\chi_\lambda(\mathbf{s})$.

3. Επικεντρώνοντας στο σημείο εκτίμησης \mathbf{s}_0 , ορίζεται η γειτονιά του εκτιμώμενου σημείου ως όλα τα σημεία $\mathbf{s}_j \in B(\mathbf{s}_0; r_s)$ σε μια σφαίρα ακτίνας r_s με κέντρο το εκτιμώμενο σημείο. Η ακτίνα r_s (ακτίνα επίδρασης) δύναται να είναι ίση με την ακτίνα συσχέτισης ξ , η οποία έχει υπολογιστεί ως παράμετρος του μοντέλου ΣΤΠ, ή να λαμβάνει μικρότερες αλλά και μεγαλύτερες τιμές. Η λύση εκφράζει τις διακυμάνσεις του τυχαίου πεδίου $\chi_\lambda(\mathbf{s})$ μέσα στην γειτονιά που έχει οριστεί.
4. Καθορίζονται οι τοπικές παράμετροι της λύσης, C_1, C_2, C_3, C_4 (2.34), οι οποίες ελαχιστοποιούν το τετραγωνικό σφάλμα στα σημεία της γειτονιάς.
5. Τέλος, καθορίζεται η εκτίμηση στο \mathbf{s}_0 χρησιμοποιώντας τις τιμές των παραμέτρων που υπολογίστηκαν στο τέταρτο βήμα.

Η διακύμανση $\hat{\chi}(\mathbf{s})$ στην γειτονιά του σημείου εκτίμησης δίνεται από την επόμενη εξίσωση,

$$\hat{\chi}(\mathbf{s}) = C_1 e^{i\mathbf{k}_1 \cdot \mathbf{s}} + C_2 e^{i\mathbf{k}_2 \cdot \mathbf{s}} + C_3 e^{i\mathbf{k}_3 \cdot \mathbf{s}} + C_4 e^{i\mathbf{k}_4 \cdot \mathbf{s}}, \quad \mathbf{k}_j = k_j \hat{\mathbf{e}}_j \quad (2.34)$$

όπου $\hat{\mathbf{e}}_j$ είναι το μοναδιαίο διάνυσμα κατεύθυνσης του κυματανύσματος \mathbf{k}_j .

Τα μέτρα k_j των κυματανυσμάτων της εξίσωσης (2.34) δίνονται από τις ρίζες του χαρακτηριστικού πολωνύμου της εξίσωσης (2.32),

$$1 + \eta_1 \xi^2 k^2 + \xi^4 k^4 = 0. \quad (2.35)$$

Οι ρίζες του πολωνύμου δίνονται από την παρακάτω εξίσωση,

$$k_j = \pm \frac{1}{\sqrt{2}\xi} \sqrt{-\eta_1 \pm \sqrt{\eta_1^2 - 4}} \quad (2.36)$$

Οι ρίζες μπορεί να είναι μιγαδικοί ή φανταστικοί αριθμοί, ανάλογα με τις τιμές που λαμβάνει η παράμετρος η_1 . Για $\eta_1 \geq 2$ οι ρίζες είναι φανταστικοί αριθμοί, ενώ για $|\eta_1| < 2$ οι ρίζες είναι μιγαδικοί αριθμοί. Αν η λύση η οποία αντιστοιχεί σε μια συγκεκριμένη κατεύθυνση \mathbf{e}_k του κυματανύσματος \mathbf{k} δίνεται από την εξίσωση (2.37)

$$\hat{\chi}(\mathbf{s}; \mathbf{e}_k) = C_1 e^{i\mathbf{k}_1 \cdot \mathbf{s}} + C_2 e^{i\mathbf{k}_2 \cdot \mathbf{s}} + C_3 e^{i\mathbf{k}_3 \cdot \mathbf{s}} + C_4 e^{i\mathbf{k}_4 \cdot \mathbf{s}}, \quad (2.37)$$

τότε η γενική ιστροπική λύση είναι το άθροισμα όλων των πιθανών κατευθύνσεων με τον ίδιο συντελεστή βαρύτητας για κάθε κατεύθυνση και δίνεται από το παρακάτω ολοκλήρωμα:

$$\hat{\chi}(\mathbf{s}) = \int d\mathbf{e}_k \hat{\chi}(\mathbf{s}; \mathbf{e}_k). \quad (2.38)$$

Στη συνέχεια ορίζεται ως ϕ η γωνία μεταξύ του διανύσματος θέσης \mathbf{s} και του κυματανύσματος \mathbf{k} . Σε δύο διαστάσεις χρησιμοποιώντας τον ορισμό της τροποποιημένης συνάρτησης Bessel πρώτου είδους και μηδενικής τάξης, δηλαδή

$$\int_0^\pi d\phi e^{\pm z \cos \phi} = I_0(z) \quad \text{προκύπτει η παρακάτω εξίσωση για τον υπολογισμό της}$$

ιστροπικής λύσης (2.38),

$$\hat{\chi}(\mathbf{s}) = \sum_{j=1}^4 C_j \int_0^{2\pi} d\phi e^{i(k_j s) \cos \phi} = 2\pi \sum_{j=1}^4 C_j I_0(ik_j s) \quad (2.39)$$

Εξετάζονται δύο περιπτώσεις ξεχωριστά, (Παράρτημα Β), οι οποίες οδηγούν σε δύο διαφορετικές συναρτήσεις $\hat{\chi}(\mathbf{s})$,

- για $\eta_1 \geq 2$

$$\hat{\chi}(\mathbf{s}) = 4\pi \left[AI_0 \left(\frac{s}{\sqrt{2\xi}} \sqrt{\eta_1 + \sqrt{\eta_1^2 - 4}} \right) + BI_0 \left(\frac{s}{\sqrt{2\xi}} \sqrt{\eta_1 - \sqrt{\eta_1^2 - 4}} \right) \right] \quad (2.40)$$

- για $|\eta|_1 < 2$

$$\hat{\chi}(\mathbf{s}) = 4\pi \left\{ A \operatorname{Re}[J_0(p_1 s + i u_1 s)] + B \operatorname{Im}[J_0(p_1 s + i u_1 s)] \right\}, \quad (2.41)$$

όπου (Re) το πραγματικό και (Im) το φανταστικό μέρος της συνάρτησης ενώ J_0 η συνάρτηση Bessel πρώτου είδους και μηδενικής τάξης.

Και οι δύο λύσεις (2.40) και (2.41) μπορούν να εκφραστούν με την μορφή

$$\hat{\chi}(\mathbf{s}) = A\psi_1(\mathbf{s}) + B\psi_2(\mathbf{s}), \quad (2.42)$$

όπου τα βάρη A και B είναι πραγματικοί αριθμοί. Οι συναρτήσεις βάσης ψ_1, ψ_2 εκφράζουν μορφές εξάρτησης που τείνουν να δώσουν υψηλή πιθανότητα πραγματοποίησης και δίνονται από τις εξισώσεις:

$$\psi_1(\mathbf{s}) = \begin{cases} 4\pi I_0\left(\frac{s}{2\xi} \sqrt{\eta_1 + \sqrt{\eta_1^2 - 4}}\right), \eta_1 \geq 2 \\ 4\pi \operatorname{Re}[J_0(p_1 s + i u_1 s)], |\eta_1| < 2 \end{cases} \quad (2.43)$$

$$\psi_2(\mathbf{s}) = \begin{cases} 4\pi I_0\left(\frac{s}{2\xi} \sqrt{\eta_1 - \sqrt{\eta_1^2 - 4}}\right), \eta_1 \geq 2 \\ 4\pi \operatorname{Im}[J_0(p_1 s + i u_1 s)], |\eta_1| < 2 \end{cases} \quad (2.44)$$

Οι πραγματικές τιμές της διακύμανσης στα σημεία $\mathbf{s}_j \in B(\mathbf{s}_0; r_s)$, $j = 1, \dots, M$ τα οποία ανήκουν στην γειτονιά του εκτιμώμενου σημείου προκύπτουν μέσω της εξίσωσης που ακολουθεί, όπου $\psi_{1,j}$ και $\psi_{2,j}$, $j = 1 \dots M$ είναι οι τιμές των συναρτήσεων βάσης ψ_1, ψ_2 στα γειτονικά σημεία, και ε_j , $j = 1 \dots M$ τα αντίστοιχα «σφάλματα».

$$\begin{bmatrix} \chi_1 \\ \vdots \\ \chi_M \end{bmatrix} = \begin{bmatrix} \psi_{1,1} & \psi_{2,1} \\ \vdots & \vdots \\ \psi_{1,M} & \psi_{2,M} \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_M \end{bmatrix} \quad (2.45)$$

Οι συντελεστές A , B της λύσης καθορίζονται ελαχιστοποιώντας το τοπικό τετραγωνικό σφάλμα της εκτίμησης στα σημεία στην γειτονιά του σημείου \mathbf{s}_0 , δηλαδή την εξής ποσότητα:

$$\varepsilon^2(A, B) = \sum_{j=1}^M [\hat{\chi}(\mathbf{s}_j) - \chi(\mathbf{s}_j)]^2, \quad \mathbf{s}_j \in B(\mathbf{s}_0; r_s) \quad (2.46)$$

Η ελαχιστοποίηση επιτυγχάνεται επιλύοντας τις εξισώσεις

$$\frac{\partial \varepsilon^2(A, B)}{\partial A} = \frac{\partial \varepsilon^2(A, B)}{\partial B} = 0 \quad (2.47)$$

Η ελαχιστοποίηση οδηγεί σε ένα γραμμικό σύστημα για την εκτίμηση της διακύμανσης που εκφράζεται με την παρακάτω εξίσωση,

$$\hat{\chi}(\mathbf{s}) = \mathbf{W}_s \mathbf{X}_s = \mathbf{Y}_{0,s} (\mathbf{Y}_s^T \mathbf{Y}_s)^{-1} \mathbf{Y}_s^T \mathbf{X}_s \quad (2.48)$$

όπου $\mathbf{X}_s^T = [\chi_1, \dots, \chi_M]$, $\mathbf{Y}_{0,s} = [\psi_1(\mathbf{s}_0) \quad \psi_1(\mathbf{s}_0)]$

και $\mathbf{Y}_s^T = \begin{bmatrix} \psi_1(\mathbf{s}_1) & \dots & \psi_1(\mathbf{s}_M) \\ \psi_2(\mathbf{s}_1) & \dots & \psi_2(\mathbf{s}_M) \end{bmatrix}$, (Hristopulos 2005b).

3. Εφαρμογή στην χαρτογράφηση περιεκτικότητας Cr στο έδαφος της περιοχής ενδιαφέροντος

Στο πρώτο κεφάλαιο περιγράφηκε τι είναι η γεωστατιστική ανάλυση, πώς εφαρμόζεται, πού εφαρμόζεται και τι αποτελέσματα δίνει. Στην συνέχεια στο δεύτερο κεφάλαιο αναλύθηκε η εφαρμογή μια νέας γεωστατιστικής μεθόδου, των μοντέλων Σπαρτιάτικων Τυχαίων Πεδίων, η οποία εκτιμάται ότι υπερτερεί έναντι των κλασσικών γεωστατιστικών μεθόδων. Σε αυτό το κεφαλαίο πραγματοποιείται η γεωστατιστική ανάλυση ρύπου χρωμίου στο έδαφος. Εφαρμόζεται τόσο η μέθοδος των μοντέλων ΣΤΠ, ώστε να διαπιστωθεί η αποτελεσματικότητα της σε πραγματικά δεδομένα πεδίου, όσο και η κλασσική γεωστατιστική μέθοδος. Αφού αξιολογηθούν και συγκριθούν τα αποτελέσματα των δύο μεθόδων πραγματοποιείται χαρτογράφηση της περιεκτικότητας του ρύπου στην περιοχή μελέτης.

3.1 Περιγραφή της γεωγραφικής περιοχής

Η διασπορά χρωμίου στο έδαφος θα μελετηθεί στην περιοχή «Jura» των Ελβετικών βουνών η οποία καταλαμβάνει έκταση 14.5km². Η περιοχή αποτελείται κατά 63% από λιβάδια, 22% βοσκοτόπια, 13% δάση και 2% από καλλιεργήσιμη γη. Το διαθέσιμο δείγμα αποτελείται από 359 μετρήσεις συγκέντρωσης Cr (σε mg/kg, ppm) σε συγκεκριμένες θέσεις οι οποίες προσδιορίζονται με την βοήθεια των συντεταγμένων τους (Goonaerts 1997). Η χωρική κατανομή του δείγματος περιγράφεται στο σχήμα 3.1.

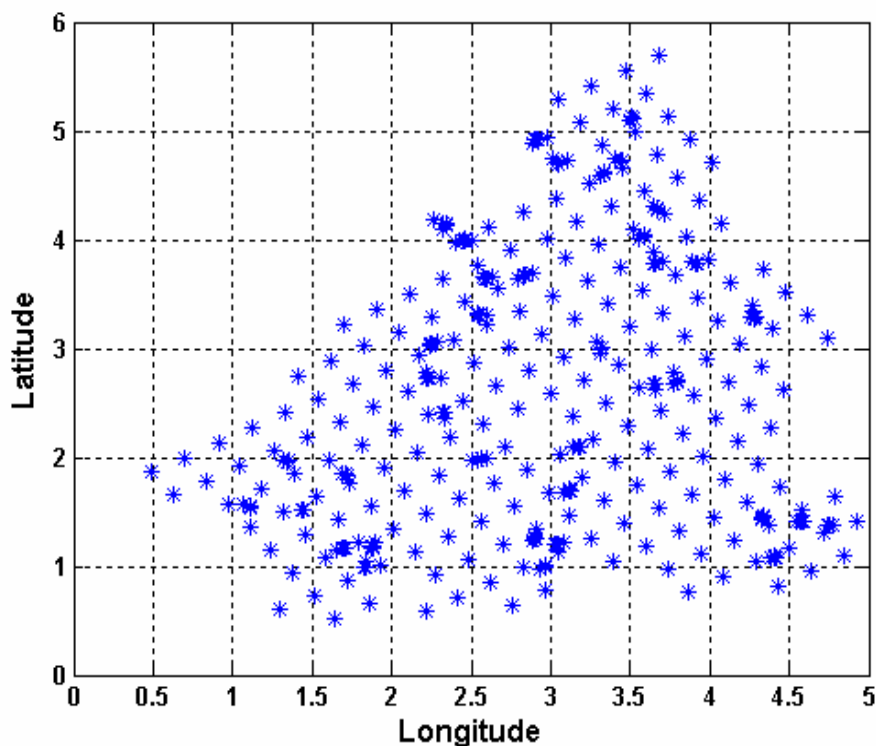
Για να διαπιστωθεί αν οι συγκεντρώσεις χρωμίου στο έδαφος της περιοχής είναι σημαντικές, αν δηλαδή αποτελούν πηγή περιβαλλοντικού κινδύνου, συγκρίνονται οι τιμές του δείγματος με τα επιτρεπτά όρια συγκέντρωσης χρωμίου στο έδαφος. Η Ελβετία έχει οριοθετήσει ως ανώτερη συγκέντρωση 75 ppm (Forest Management Institute 2005). Η Ευρωπαϊκή Ένωση αντίθετα δεν έχει θεσπίσει ακόμα συγκεκριμένα όρια συγκέντρωσης χρωμίου στο έδαφος τα οποία να μπορούν να χρησιμοποιηθούν σε περιπτώσεις ρύπανσης, αλλά στα πλαίσια της Οδηγίας για την χρήση λυματολάσπης ως βελτιωτικό χώματος ανώτατο όριο συγκέντρωσης χρωμίου έχουν οριστεί τα 100 ppm (EC 2003).

Άλλες Ευρωπαϊκές χώρες που έχουν θεσπίσει όρια συγκέντρωσης χρωμίου στο έδαφος, όπως η Μ. Βρετανία (BS 12457) και η Ολλανδία, έχουν ορίσει ως ανώτερη συγκέντρωση τα 100 ppm. Στην Ολλανδία έχουν τεθεί επίσης συγκεκριμένα όρια σημαντικότητας τα οποία καθορίζουν τις απαραίτητες ενέργειες ανά περίπτωση (Πίνακας 3.1). Τέλος η Δανία έχει θεσπίσει το αυστηρότερο όριο συγκέντρωσης χρωμίου στο έδαφος το οποίο είναι 30 ppm (EC 2003).

Πίνακας 3.1: Επιτρεπτή συγκέντρωση χρωμίου σε έδαφος της Ολλανδίας (EC 2003).

<i>Τιμή Αναφοράς (ppm)</i>	<i>Λεπτομερής Έλεγχος (ppm)</i>	<i>Άμεση Αποκατάστασ η (ppm)</i>
100	250	800

Λαμβάνοντας υπόψη τις παραπάνω πληροφορίες και το γεγονός ότι η μεγαλύτερη συγκέντρωση χρωμίου στο δείγμα είναι 70 ppm συμπεραίνεται ότι οι μετρημένες συγκεντρώσεις χρωμίου στην περιοχή βρίσκονται μέσα στα επιτρεπτά όρια που ισχύουν στην Ελβετία αλλά και στις άλλες Ευρωπαϊκές χώρες, εκτός της Δανίας.



Σχήμα 3.1: Χωρική κατανομή συγκέντρωσης χρωμίου (Goovaerts 1997).

3.2 Διερευνητική στατιστική ανάλυση

Πριν από την εφαρμογή του γεωστατιστικού μοντέλου πρέπει να πραγματοποιηθεί διερευνητικός στατιστικός έλεγχος του δείγματος. Πρέπει δηλαδή να διαπιστωθεί αν το δείγμα μπορεί να χρησιμοποιηθεί για αποτελεσματική γεωστατιστική ανάλυση ή αν πρέπει να υποστεί κάποιο μετασχηματισμό. Για να εφαρμοστούν οι κλασσικές μέθοδοι γεωστατιστικής ανάλυσης πρέπει να τηρούνται κάποιες προϋποθέσεις.

Η πιο βασική προϋπόθεση αφορά την κατανομή του δείγματος. Αν οι τιμές του δείγματος ακολουθούν την κανονική κατανομή απλοποιείται η εφαρμογή της γεωστατιστικής ανάλυσης, καθώς δεν είναι απαραίτητος ο μετασχηματισμός των τιμών του δείγματος. Επίσης, αν το δείγμα είναι στατιστικά ομοιογενές εξυπηρετεί σημαντικά στον υπολογισμό του ημιβαριογράμματος, καθώς σε αυτήν την περίπτωση το ημιβαριόγραμμα εξαρτάται αποκλειστικά από την απόσταση μεταξύ των σημείων του δείγματος

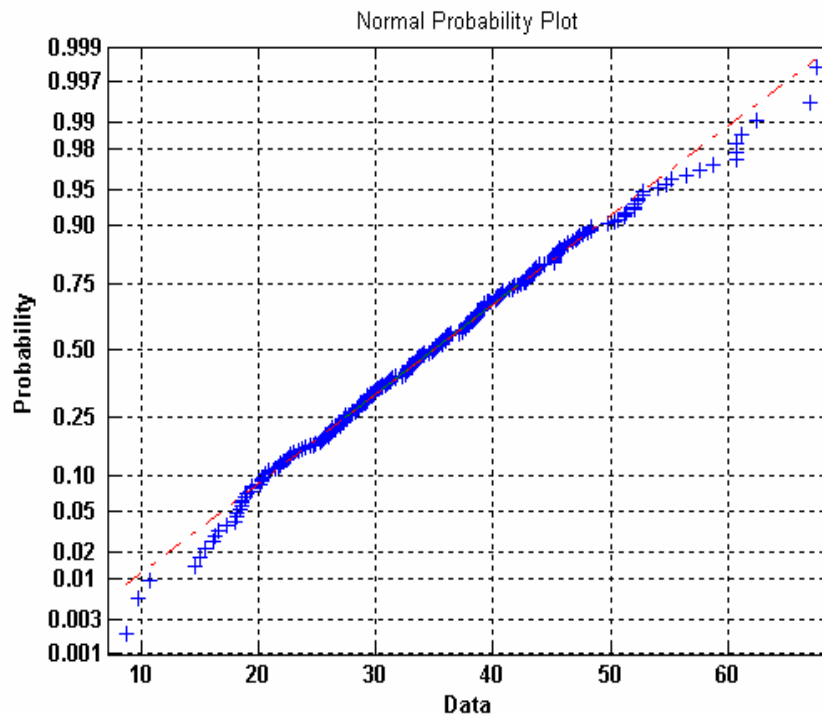
Ένας πρώτος τρόπος για να διαπιστωθεί αν το δείγμα ακολουθεί την κανονική κατανομή είναι η δημιουργία του διαγράμματος κανονικής κατανομής, το οποίο υλοποιείται από την εντολή «*normplot*» της Matlab®. Αν τα σημεία του δείγματος βρίσκονται πάνω στην ευθεία γραμμή του διαγράμματος τότε το δείγμα ακολουθεί την κανονική κατανομή. Όπως παρατηρείται από το Σχήμα 3.2, οι περισσότερες συγκεντρώσεις του δείγματος χρωμίου βρίσκονται κατά μήκος της ευθείας γραμμής του διαγράμματος και μόνο οι ακραίες αποκλίνουν.

Παρόμοιος έλεγχος, αλλά ποιοτικής μορφής, επιτυγχάνεται με την εντολή «*hist*» της Matlab® η οποία δείχνει με την χρήση ιστογράμματος (Σχήμα 3.3) αν οι συγκεντρώσεις του δείγματος ακολουθούν την κανονική κατανομή, αλλά και με την εντολή «*histfit*» επίσης της Matlab® η οποία προσαρμόζει την βέλτιστη κανονική συνάρτηση πυκνότητας πιθανότητας στο ιστόγραμμα των συγκεντρώσεων χρωμίου (Σχήμα 3.4). Από τα δύο σχήματα φαίνεται ότι οι συγκεντρώσεις χρωμίου ακολουθούν την κανονική κατανομή καθώς στο μεν πρώτο (Σχήμα 3.3) το ιστόγραμμα έχει την μορφή της καμπύλης «καμπάνας», στο δε δεύτερο (Σχήμα 3.4) προσαρμόζεται πολύ καλά στην καμπύλη της βέλτιστης κανονικής συνάρτησης πυκνότητας πιθανότητας.

Ένας άλλος τρόπος για να διαπιστωθεί αν η συγκέντρωση Cr ακολουθεί την κανονική κατανομή είναι η εφαρμογή στατιστικού ελέγχου. Ο πιο διαδεδομένος στατιστικός έλεγχος είναι ο έλεγχος Kolmogorov-Smirnov. Η μηδενική υπόθεση, θεωρεί ότι οι τιμές ενός δείγματος ακολουθούν την κανονική κατανομή. Ο έλεγχος Kolmogorov-Smirnov εφαρμόζεται συγκρίνοντας τις κανονικοποιημένες τιμές του δείγματος, οι οποίες δίνονται από την εξίσωση (3.1)

$$\frac{X(s) - \bar{X}(s)}{\sigma_X}, \text{ όπου } \bar{X}(s) \text{ ο δειγματικός μέσος όρος,} \quad (3.1)$$

με μια πρότυπη κανονική κατανομή η οποία έχει μέση τιμή $\bar{X} = 0$ και τυπική απόκλιση $\sigma = 1$. Ο έλεγχος υλοποιείται σε περιβάλλον Matlab[®] με την βοήθεια της εντολής «kstest». Στο συγκεκριμένο δείγμα προκύπτει ότι ισχύει η μηδενική υπόθεση σε στάθμη σημαντικότητας 5% και 10%, ενώ οι τιμές του Cr ακολουθούν την κανονική κατανομή με πιθανότητα 0.68 (Matlab 2002).



Σχήμα 3.2: Διάγραμμα πιθανότητας κανονικής κατανομής δείγματος χρωμίου.

Στην συνέχεια θα διερευνηθούν και άλλες βασικές στατιστικές ιδιότητες του δείγματος (Πίνακας 3.2). Ο συντελεστής ασυμμετρίας \hat{s}_X της κατανομής του δείγματος, ορίζεται με την βοήθεια της κεντρικής ροπής τρίτης τάξης (Χριστόπουλος 2003) ως εξής,

$$\hat{s}_X = \frac{\frac{1}{M} \sum_{i=1}^M (x_i - \hat{m}_X)^3}{\hat{\sigma}_X^3}, \text{ όπου } M \text{ το άθροισμα των σημείων του δείγματος,} \quad (3.2)$$

και υλοποιείται σε περιβάλλον Matlab[®] από την εντολή «skewnes». Για τις συγκεντρώσεις του δείγματος ο συντελεστής ασυμμετρίας υπολογίζεται σε $\hat{s}_X = 0.3$, δηλαδή είναι πολύ κοντά στην τιμή της κανονικής κατανομής που είναι ίση με το μηδέν.

Ο συντελεστής κύρτωσης ορίζεται με την βοήθεια της κεντρικής ροπής τέταρτης τάξης (Χριστόπουλος 2003) και δίνεται από την παρακάτω εξίσωση:

$$\hat{k}_X = \frac{\frac{1}{M} \sum_{i=1}^M (x_i - \hat{m}_X)^4}{\hat{\sigma}_X^4}. \quad (3.3)$$

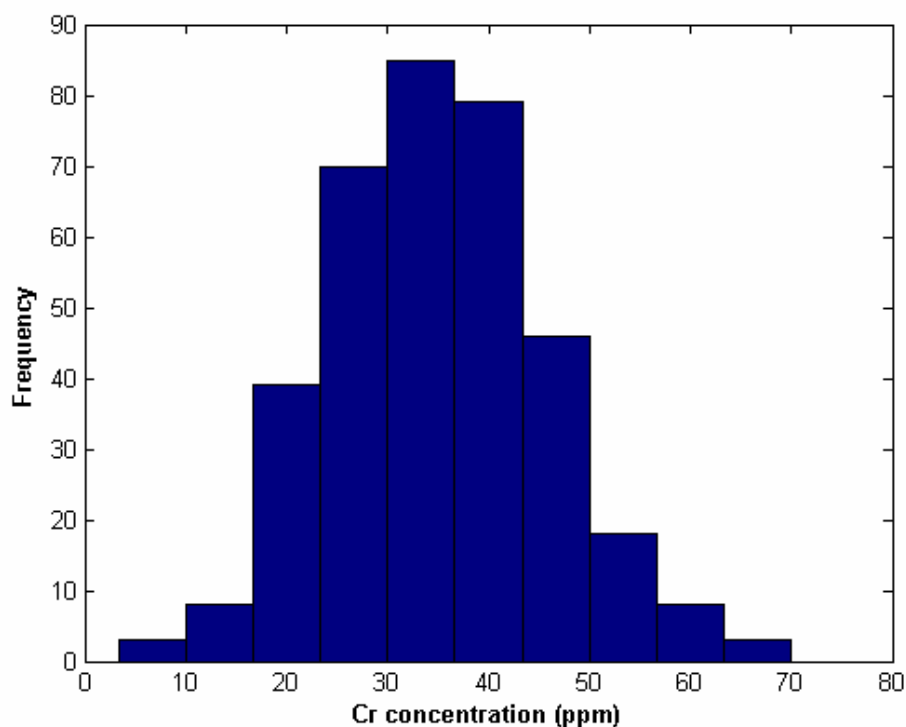
Ο συντελεστής κύρτωσης σε περιβάλλον Matlab[®] δίνεται από την εντολή «kurtosis» και για τις τιμές του δείγματος υπολογίζεται σε $\hat{k}_X = 3.32$. Η τιμή του συντελεστή κύρτωσης είναι πολύ κοντά στον συντελεστή κύρτωσης της κανονικής κατανομής ο οποίος ισούται με τρία.

Τέλος, η μικρότερη τιμή του δείγματος συγκεντρώσεων χρωμίου βρίσκεται ανάμεσα στο διάστημα εμπιστοσύνης 99.73% της κανονικής κατανομής το οποίο αντιστοιχεί σε διάστημα τριών τυπικών αποκλίσεων $[m - 3\sigma, m + 3\sigma]$ εκατέρωθεν της μέσης τιμής του δείγματος (Χριστόπουλος 2003). Αντίθετα η μέγιστη τιμή του δείγματος βρίσκεται εκτός αυτού του διαστήματος κατά περίπου μισή τυπική απόκλιση. Γνωρίζοντας ότι το 99.73% των τιμών του δείγματος βρίσκεται εντός του διαστήματος τριών τυπικών αποκλίσεων $[m - 3\sigma, m + 3\sigma]$ εκατέρωθεν της μέσης

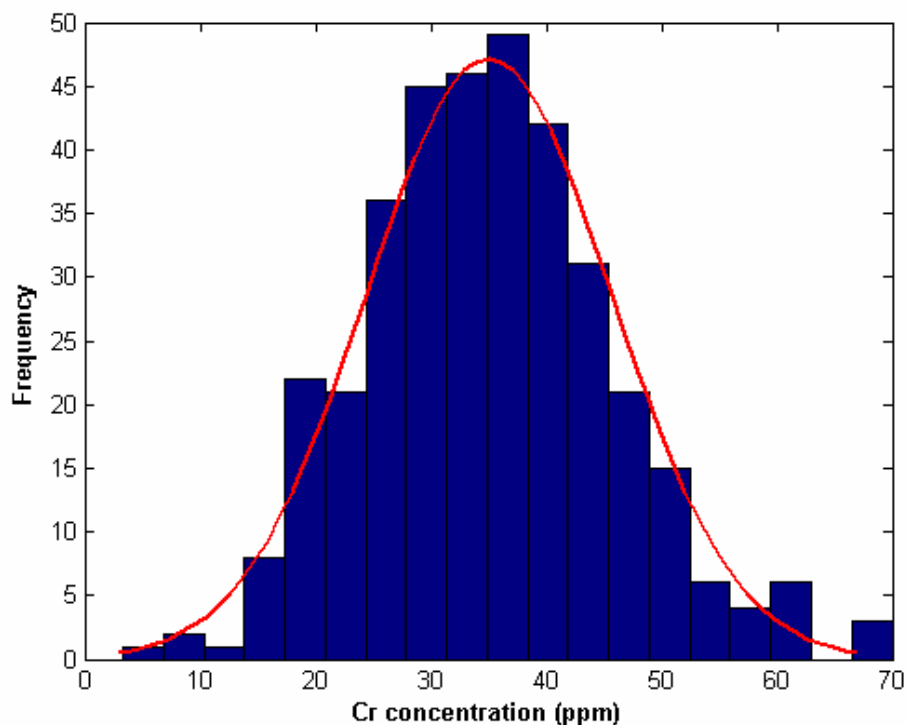
τιμής, μπορεί να υπολογιστεί ο “αναμενόμενος” αριθμός των τιμών που βρίσκονται εκτός: το 0.27% των τιμών του δείγματος σε σύνολο 359, αντιστοιχεί σε τουλάχιστον μία τιμή. Επομένως αυτό αποτελεί μια ένδειξη ότι το δείγμα συγκεντρώσεων χρωμίου ακολουθεί την κανονική κατανομή.

Πίνακας 3.2: Στατιστικές ιδιότητες δείγματος χρωμίου.

Μικρότερη Τιμή	Μεγαλύτερη Τιμή	Μέση Τιμή	Τυπική Απόκλιση	Ασυμμετρία Κατανομής	Συντελεστής Κύρτωσης
3.32ppm	70ppm	35.01ppm	10.66	0.30	3.32



Σχήμα 3.3: Ιστόγραμμα κανονικής κατανομής Cr χρησιμοποιώντας την εντολή «hist» της Matlab®. Το ιστόγραμμα περιγράφει την συχνότητα εμφάνισης των μετρήσεων του δείγματος στα αντίστοιχα διαστήματα συγκέντρωσης χρωμίου.



Σχήμα 3.4: Ιστόγραμμα κανονικής κατανομής Cr χρησιμοποιώντας την εντολή «histfit» της Matlab®. Το ιστόγραμμα περιγράφει την συχνότητα εμφάνισης των μετρήσεων του δείγματος στα αντίστοιχα διαστήματα συγκέντρωσης χρωμίου.

3.3 Ανάλυση χωρικής συνέχειας

3.3.1 Προσδιορισμός χωρικής τάσης

Ένας άλλος παράγοντας που πρέπει να εξεταστεί είναι η ύπαρξη χωρικής τάσης στο δείγμα συγκεντρώσεων χρωμίου. Η χωρική τάση ενός δείγματος μπορεί να προσδιοριστεί από πρότυπα γενικής και τοπικής εξάρτησης, όπως η πολλαπλή γραμμική παλινδρόμηση δοκιμάζοντας διάφορες συναρτήσεις χώρου (σταθερές, πολυωνυμικές, περιοδικές, σύνθετες), τα φίλτρα κινούμενου μέσου όρου, τα φίλτρα Savitzky-Golay και η μέθοδος της τοπικά ζυγισμένης παλινδρόμησης.

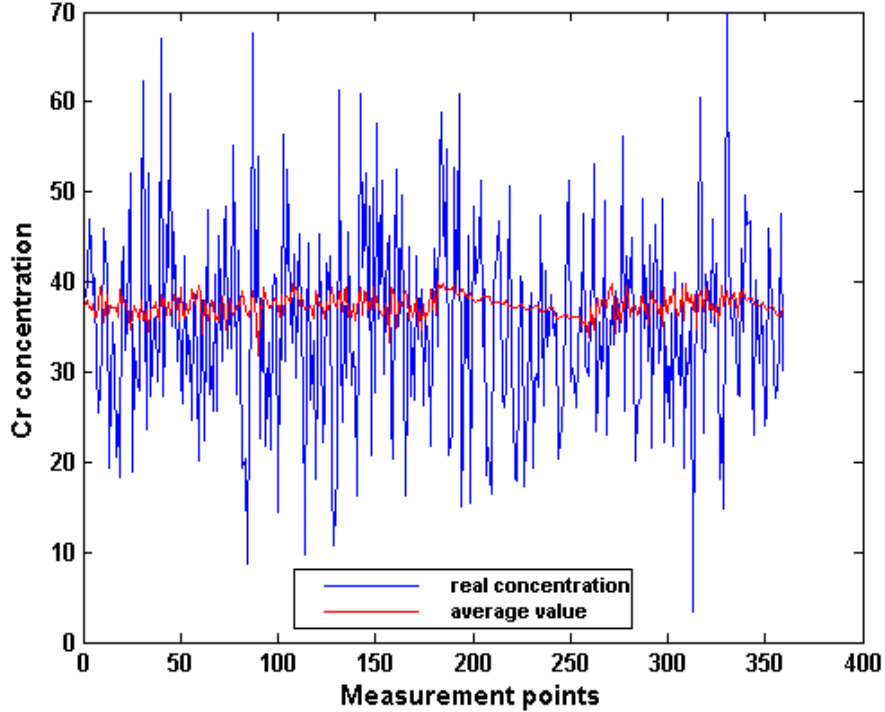
Η ανάλυση της χωρικής τάσης του δείγματος χρωμίου πραγματοποιήθηκε στα πλαίσια προγενέστερης εργασίας (Βαρουχάκης 2005), χρησιμοποιώντας τα παραπάνω πρότυπα. Το αποτέλεσμα της ανάλυσης απέδειξε ότι δεν υπάρχει σημαντική χωρική τάση στο δείγμα χρωμίου, συνεπώς το πεδίο τιμών μπορεί να χαρακτηριστεί ως στατιστικά ομοιογενές. Ο προσδιορισμός της χωρικής τάσης του

δείγματος συγκεντρώσεων χρωμίου από ένα πρότυπο τοπικής εξάρτησης, όπως η τοπικά ζυγισμένη παλινδρόμηση παρουσιάζεται στο σχήμα 3.5.

Κατά την μέθοδο της τοπικά ζυγισμένης παλινδρόμησης (Atkeson et all 1997), σταθμίζεται η απόσταση ανάμεσα στο σημείο εκτίμησης της τάσης και στα γειτονικά σημεία. Αυτό πραγματοποιείται με την βοήθεια μιας συνάρτησης αντιστάθμισης. Στην συγκεκριμένη περίπτωση χρησιμοποιείται η τρικυβική συνάρτηση αντιστάθμισης,

$$K(D) = \begin{cases} [1 - |D|^3]^3, & |D| \leq 1 \\ 0, & |D| \geq 1 \end{cases} \quad (3.4)$$

όπου $D = \frac{D_E}{h}$, και D_E είναι η ευκλείδεια απόσταση του αρχικού σημείου από τα υπόλοιπα ενώ το h είναι η ακτίνα αναζήτησης γειτονικών σημείων. Για τον προσδιορισμό της τάσης χρησιμοποιείται ένα γραμμικό πρότυπο τοπικής εξάρτησης.

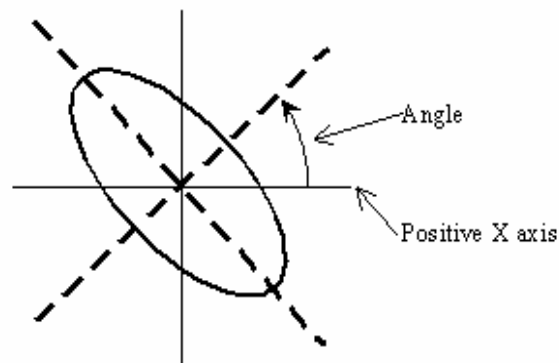


Σχήμα 3.5: Η τάση του δείγματος παρουσιάζεται από την γραμμή (-) κόκκινου χρώματος ενώ οι τιμές του δείγματος από την γραμμή (-) μπλε χρώματος.

3.3.2 Προσδιορισμός ημιβαριόγραμματος δείγματος χρωμίου

Στην συνέχεια υπολογίζεται το ημιβαριόγραμμα του δείγματος συγκεντρώσεων χρωμίου. Θεωρητικά το ημιβαριόγραμμα περιγράφει την εξάρτηση δύο σημείων στον χώρο, ενώ μαθηματικά καθορίζει την χωρική μεταβλητότητα των διακυμάνσεων του δείγματος. Η βασική αρχή είναι ότι δύο κοντινές παρατηρήσεις συμφωνούν σε μεγαλύτερο βαθμό σε σχέση με πιο μακρινές. Επειδή όμως οι τιμές των παρατηρήσεων μπορεί να αλλάζουν γρηγορότερα σε μια κατεύθυνση από ότι σε μια άλλη, δηλαδή να είναι ανισοτροπικές, το ημιβαριόγραμμα εν γένει μπορεί να είναι συνάρτηση της κατεύθυνσης.

Η ανισοτροπία παριστάνεται με την μορφή έλλειψης η οποία καθορίζεται από τα μήκη των δύο ορθογώνιων αξόνων της καθώς και από την γωνία κατεύθυνσης (Σχήμα 3.6). Ο έλεγχος της ανισοτροπίας περιλαμβάνει δύο παραμέτρους, τον λόγο ανισοτροπίας (ratio) και την γωνία προσανατολισμού (angle), οι οποίοι προσδιορίζουν το μέγεθος της ανισοτροπίας και την κατεύθυνση της.



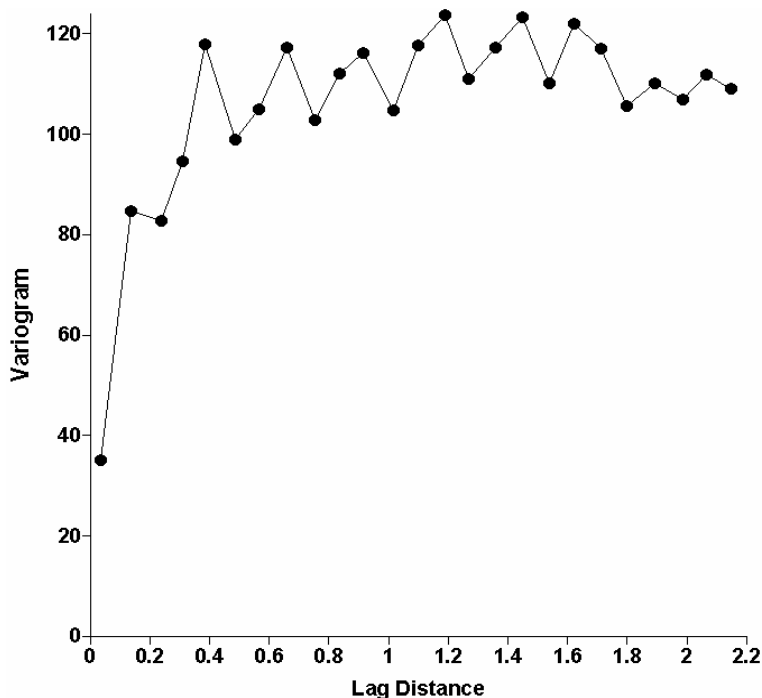
Σχήμα 3.6: Παράδειγμα απεικόνισης ανισοτροπίας (Surfer V.8.0.4, 2002).

Ο λόγος ανισοτροπίας δίνεται από τον λόγο του μεγαλύτερου μήκους άξονα προς τον μικρότερο ενώ η γωνία προσανατολισμού σχηματίζεται μεταξύ του θετικού άξονα (x) και του αντίστοιχου της έλλειψης και δείχνει την κατεύθυνση της ανισοτροπίας. Ανισοτροπία με λόγο μηκών συσχέτισης μικρότερη του δύο θεωρείται ήπια ενώ μεγαλύτερη του τέσσερα θεωρείται σημαντική (Surfer V.8.0.4, 2002).

Ο υπολογισμός του ημιβαριογράμματος του δείγματος συγκεντρώσεων χρωμίου πραγματοποιήθηκε χρησιμοποιώντας το λογισμικό πακέτο «Surfer V.8.0.4, 2002». Αρχικά υπολογίζεται ο αριθμός των ζευγών σημείων του δείγματος $n(r_k)$ ανά τάξη, για κάθε μία από τις εικοσιπέντε επιλεγθείσες τάξεις αποστάσεων r_k (Πίνακας 3.3), και στην συνέχεια υπολογίζεται μια τιμή του πειραματικού ημιβαριογράμματος $\hat{\gamma}(r_k)$ για κάθε r_k ώστε να προκύψει το πειραματικό ημιβαριόγραμμα του δείγματος (Σχήμα 3.7). Το επόμενο βήμα είναι η προσαρμογή του σε διάφορους τύπους θεωρητικών μοντέλων ημιβαριογραμμάτων, ώστε να επιλεγεί το βέλτιστο μοντέλο.

Πίνακας 3.3: Τάξεις απόστασης, πλήθος ζευγών ανά τάξη και τιμές πειραματικού ημιβαριογράμματος για την κατανομή Cr.

r_k	$n(r_k)$	$\hat{\gamma}(r_k)$
0.03	269	35.05
0.13	250	84.70
0.23	758	82.82
0.30	552	94.68
0.38	934	118.00
0.48	1138	98.92
0.56	1311	104.92
0.65	930	117.23
0.75	1869	102.72
0.83	1111	111.99
0.91	1469	116.16
1.01	1897	104.78
1.09	1975	117.80
1.18	1204	123.66
1.26	2362	111.02
1.35	1871	117.34
1.45	1864	123.20
1.53	2341	110.25
1.62	1707	122.09
1.71	1976	116.97
1.79	2424	105.59
1.89	1623	110.12
1.98	2263	106.85
2.06	1567	111.94
2.14	1823	109.09



Σχήμα 3.7: Πειραματικό ημιβαριόγραμμα δείγματος Cr σαν συνάρτηση της απόστασης.

Όπως προαναφέρθηκε το ημιβαριόγραμμα είναι συνάρτηση της κατεύθυνσης. Ένα παράδειγμα προσαρμογής πειραματικού ημιβαριόγραμματος σε ένα θεωρητικό μοντέλο σε δύο κατευθύνσεις, κατά την κατεύθυνση του άξονα (x) αλλά και κατά την κατεύθυνση του άξονα (y), δίνεται στα σχήματα 3.9 και 3.10. Οι τιμές των παραμέτρων (Πίνακας 3.4) των ημιβαριογραμμάτων που υπολογίζονται κατά τις δύο κατευθύνσεις είναι κατά προσέγγιση όμοιες.

Αυτό συμβαίνει γιατί ο λόγος ανισοτροπίας και στις δύο περιπτώσεις είναι μικρότερος ή ίσος με δύο και σύμφωνα με τον οδηγό λειτουργίας του λογισμικού πακέτου «Surfer V.8.0.4, 2002», για λόγο ανισοτροπίας μικρότερο του δύο η ανισοτροπία θεωρείται ήπια. Η γωνιακή ανοχή που χρησιμοποιείται για τον υπολογισμό του ανισοτροπικού λόγου είναι 90^0 , ωστόσο δοκιμάζοντας διαφορετικές τιμές γωνιακής ανοχής, 45^0 , 20^0 , 10^0 και 5^0 , διαπιστώνεται ότι η τιμή του ανισοτροπικού λόγου παραμένει ίδια. Αυτή η διαπίστωση οδηγεί στο συμπέρασμα ότι δεν υπάρχει ευαισθησία στην ανισοτροπία του δείγματος συγκεντρώσεων χρωμίου.

Λαμβάνοντας υπόψη τα παραπάνω συμπεράσματα θα παρουσιαστούν τα ημιβαριογράμματα μόνο κατά την κατεύθυνση του άξονα (x). Το πειραματικό

ημιβαριόγραμμα προσαρμόζεται σε διαφορετικούς τύπους θεωρητικών μοντέλων ημιβαριογράμματος καθώς και σε συνδυασμούς τους. Τα θεωρητικά μοντέλα που εξετάζονται είναι το μοντέλο με περιοχή αντισυσχέτισης, το εκθετικό, το γκαουσιανό, το σφαιρικό, το δυναμονομικό, το φαινόμενο πυρήνα, το λογαριθμικό, το τετραγωνικό, το κυβικό και το γραμμικό.

Για τον υπολογισμό του βέλτιστου πρότυπου ημιβαριογράμματος χρησιμοποιείται η μεθοδολογία της διασταυρωμένης επιβεβαίωσης (3.4). Από την προσαρμογή του πειραματικού στο θεωρητικό μοντέλο ημιβαριογράμματος υπολογίζονται οι παράμετροι του γεωστατιστικού μοντέλου, οι οποίοι στην συνέχεια χρησιμοποιούνται για την εκτίμηση με την μέθοδο *Kriging* χρησιμοποιώντας την μεθοδολογία της διασταυρωμένης επιβεβαίωσης.

Οι βέλτιστες τιμές των παραμέτρων υπολογίζονται ελαχιστοποιώντας το τετραγώνου της διαφοράς μεταξύ πειραματικού και θεωρητικού μοντέλου ημιβαριογράμματος. Το θεωρητικό μοντέλο ημιβαριογράμματος του οποίου οι παράμετροι δίνουν το μικρότερο σφάλμα εκτίμησης αποτελεί το βέλτιστο πρότυπο.

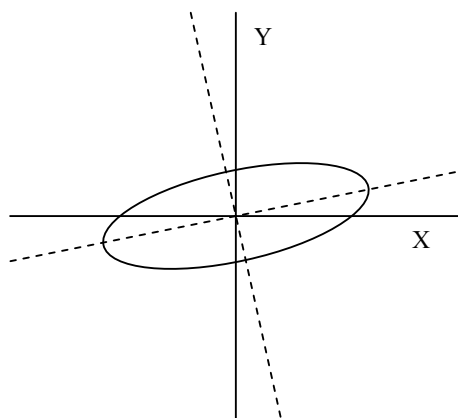
Πίνακας 3.4: Τιμές παραμέτρων των θεωρητικών μοντέλων ημιβαριογράμματος.

Θεωρητικό Μοντέλο Ημιβαριογράμματος	Κλίμακα	Μήκος συσχέτισης (km)	Λόγος Ανισοτροπίας	Γωνία Ανισοτροπίας (°)
Σφαιρικό & Φαινόμενο Πυρήνα	86.53	0.70	2.0	5.67
Σφαιρικό	112	0.55	1.99	2.00
Φαινόμενο πυρήνα & Μοντέλο με Περιοχή Αντισυσχέτισης	105	0.04	1.16	85.85
Εκθετικό & Φαινόμενο Πυρήνα	111	0.27	2.0	8.24
Τετραγωνικό	112	0.63	2.0	5.64

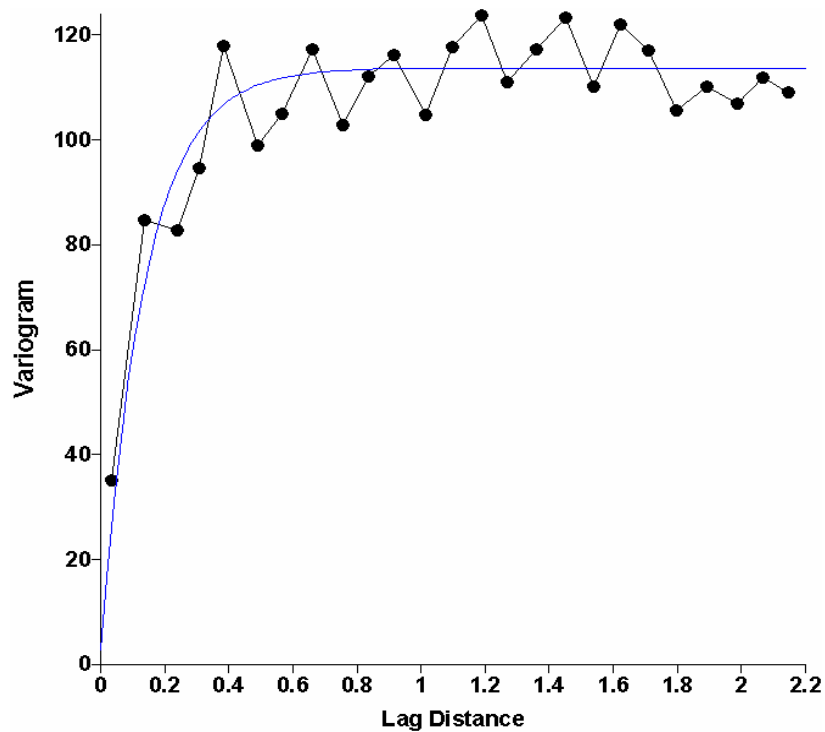
Το αποτέλεσμα της διασταυρωμένης επιβεβαίωσης δείχνει ότι το πειραματικό ημιβαριόγραμμα προσαρμόζεται ικανοποιητικά σε πέντε θεωρητικά μοντέλα ή συνδυασμούς τους (Σχήματα 3.10 έως 3.14) καθώς από αυτά προκύπτει το μικρότερο σφάλμα εκτίμησης (3.5.1). Οι βέλτιστες παράμετροι που υπολογίζονται σε κάθε περίπτωση παρουσιάζονται στον πίνακα 3.4.

Από τα πέντε θεωρητικά μοντέλα στα οποία προσαρμόζεται ικανοποιητικά το πειραματικό ημιβαριόγραμμα δείγματος χρωμίου δύο έχουν την καλύτερη προσαρμογή. Αυτά είναι ο συνδυασμός του Εκθετικού μοντέλου με το μοντέλο Φαινομένου Πυρήνα και ο συνδυασμός του Σφαιρικού μοντέλου με το μοντέλο Φαινομένου Πυρήνα (3.5.1).

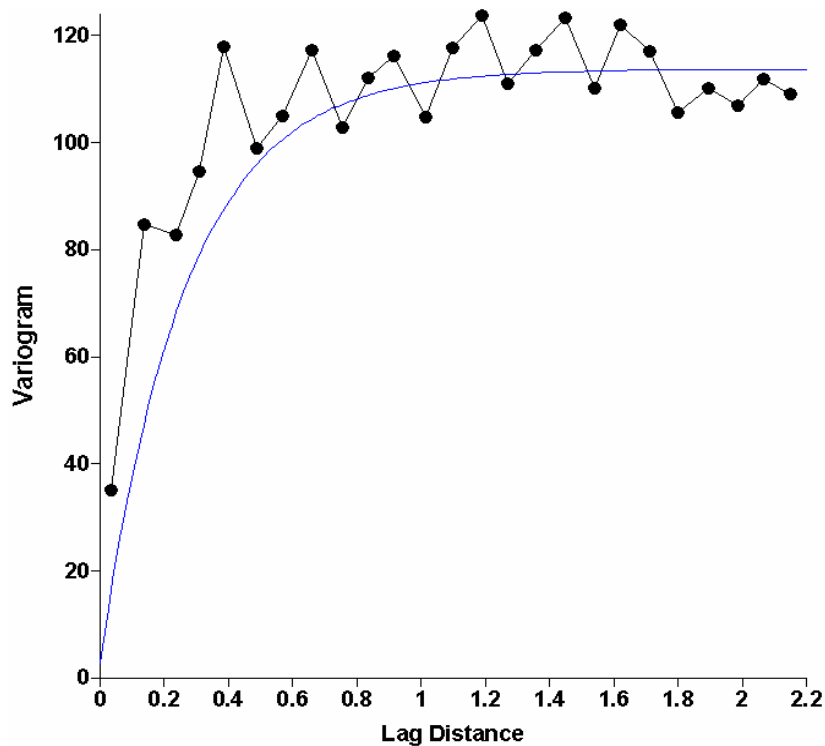
Στις δύο αυτές περιπτώσεις, όπως προαναφέρθηκε, πραγματοποιήθηκε έλεγχος ανισοτροπίας με την βοήθεια του λογισμικού πακέτου «Surfer V.8.0.4, 2002», κατά τον οποίον διαπιστώθηκε αν οι συσχετίσεις έχουν διαφορετική εξάρτηση σε διαφορετικές κατευθύνσεις. Στο δείγμα χρωμίου υπολογίζεται ανισοτροπία ίση με δύο και στις δύο περιπτώσεις των θεωρητικών μοντέλων ημιβαριογράμματος, η οποία σύμφωνα με την θεωρία που προαναφέρθηκε χαρακτηρίζεται οριακός ήπια, καθώς και η γωνία (angle) η οποία και για τις δύο περιπτώσεις υπολογίζεται περίπου σε οκτώ μοίρες. Σχηματικά η ανισοτροπία του δείγματος περιγράφεται από το Σχήμα 3.8, στο οποίο φαίνεται και η κατεύθυνση της καθώς παρατηρείται μια ελαφρά κλίση της έλλειψης από τα ανατολικά προς τα δυτικά.



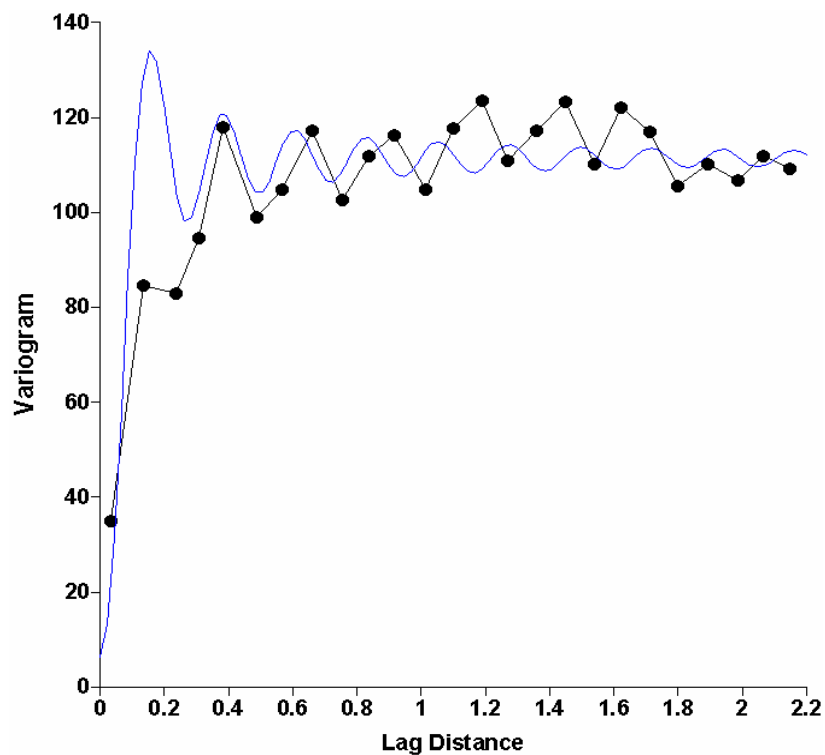
Σχήμα 3.8: Απεικόνιση ανισοτροπίας στο δείγμα χρωμίου.



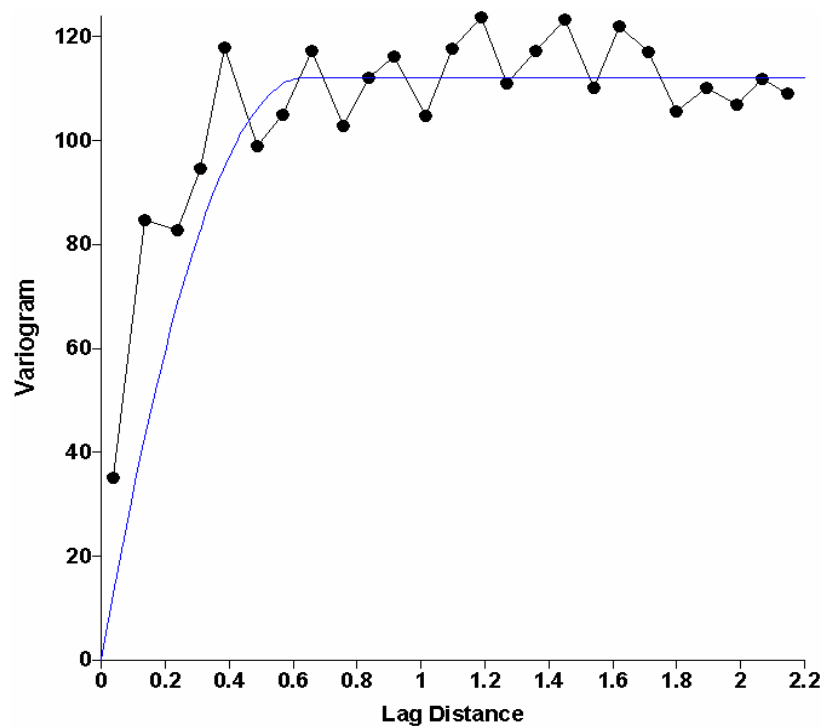
Σχήμα 3.9: Προσαρμογή πειραματικού ημιβαριογράμματος Cr στο Εκθετικό με Φαινόμενο Πυρήνα μοντέλο ημιβαριογράμματος, κατά την κατεύθυνση του άξονα (y).



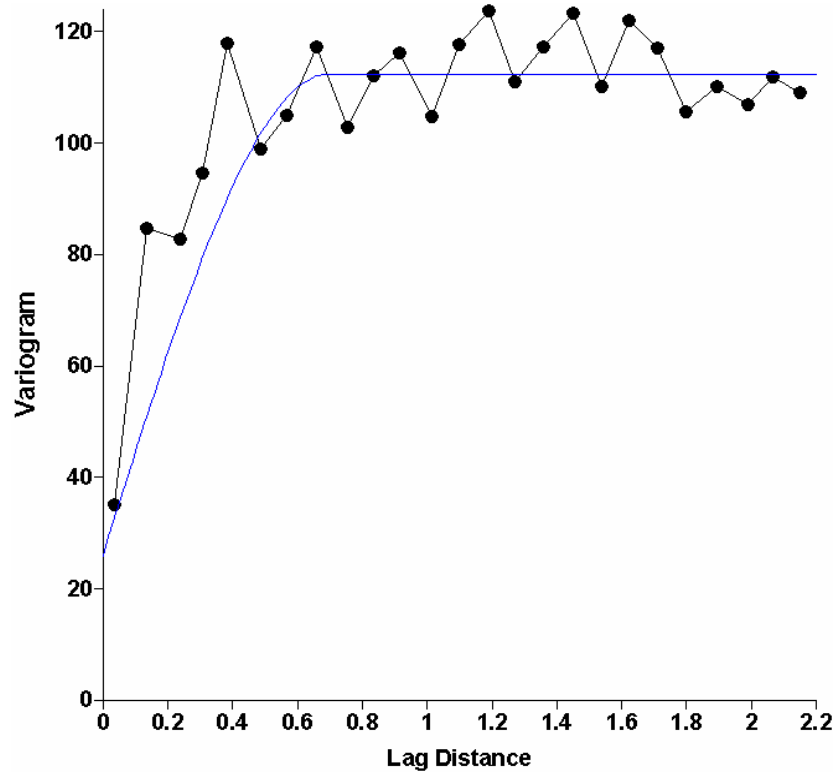
Σχήμα 3.10: Προσαρμογή πειραματικού ημιβαριογράμματος Cr στο Εκθετικό με Φαινόμενο Πυρήνα μοντέλο ημιβαριογράμματος, κατά την κατεύθυνση του άξονα (x).



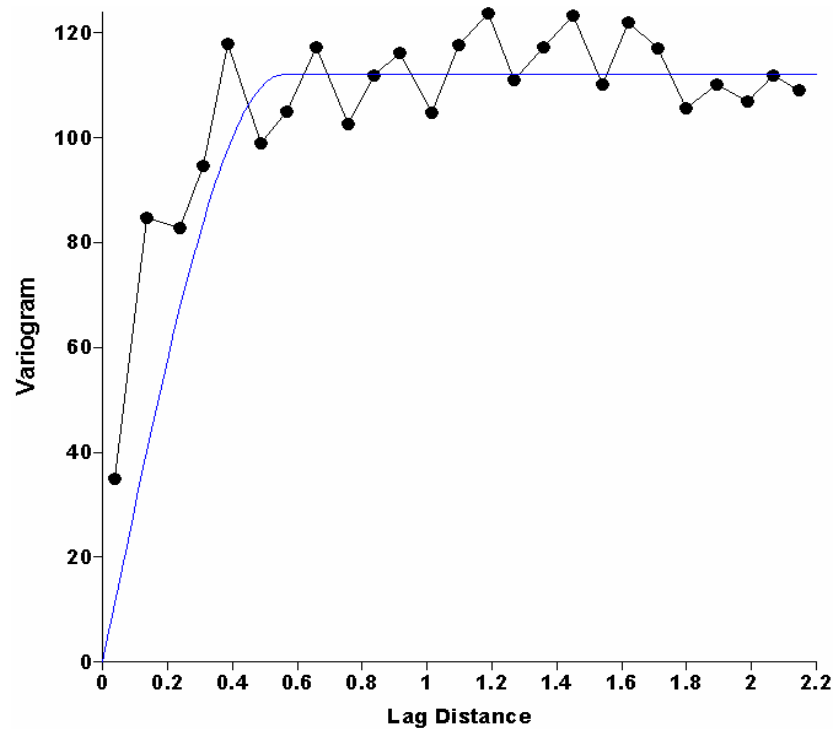
Σχήμα 3.11: Προσαρμογή πειραματικού ημιβαριογράμματος Cr στο μοντέλο ημιβαριογράμματος, Φαινόμενου πυρήνα με Περιοχή Αντισυσχέτισης.



Σχήμα 3.12: Προσαρμογή πειραματικού ημιβαριογράμματος Cr στο Τετραγωνικό μοντέλο ημιβαριογράμματος



Σχήμα 3.13: Προσαρμογή πειραματικού ημιβαριογράμματος Cr στο Σφαιρικό με Φαινόμενο Πυρήνα μοντέλο ημιβαριογράμματος.



Σχήμα 3.14: Προσαρμογή πειραματικού ημιβαριογράμματος Cr στο Σφαιρικό μοντέλο ημιβαριογράμματος.

3.4 Μεθοδολογία γεωστατιστικής ανάλυσης

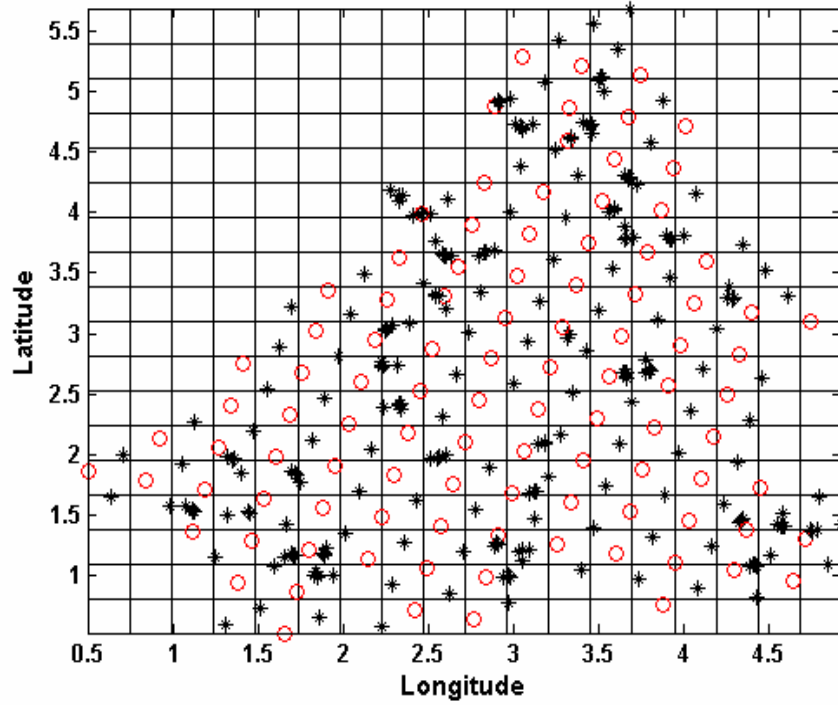
Η γεωστατιστική ανάλυση δείγματος τυχαίου πεδίου μπορεί να εφαρμοστεί με δύο διαφορετικές μεθόδους. Η πρώτη μέθοδος περιλαμβάνει τον διαχωρισμό του δείγματος σε σύνολα εκπαίδευσης και επιβεβαίωσης. Τα στοιχεία του συνόλου εκπαίδευσης αποτελούν τα εισαγωγικά δεδομένα του γεωστατιστικού μοντέλου ώστε να προσδιοριστούν οι παράμετροί του ενώ στην συνέχεια χρησιμοποιούνται για την εκτίμηση στα σημεία που ανήκουν στο σύνολο επιβεβαίωσης. Έτσι πραγματοποιείται επιβεβαίωση των εκτιμήσεων του γεωστατιστικού μοντέλου.

Η δεύτερη μεθοδολογία είναι η διασταυρωμένη επιβεβαίωση, η οποία χρησιμοποιείται σε μικρά σύνολα δεδομένων. Σε αυτήν την περίπτωση όλα τα σημεία του δείγματος αποτελούν το σύνολο εκπαίδευσης και χρησιμοποιούνται για τον υπολογισμό των παραμέτρων του γεωστατιστικού μοντέλου. Στην συνέχεια πραγματοποιείται επιβεβαίωση των εκτιμήσεων του γεωστατιστικού μοντέλου αφαιρώντας επαναληπτικά ένα σημείο από το δείγμα ενώ λαμβάνοντας υπόψη τα υπόλοιπα σημεία του δείγματος πραγματοποιείται εκτίμηση σε αυτό. Με αυτόν τον τρόπο πραγματοποιείται μια σύγκριση μεταξύ της εκτιμώμενης και της πραγματικής τιμής του δείγματος.

Στην συγκεκριμένη περιοχή μελέτης υπάρχουν 359 μετρήσεις συγκέντρωσης χρωμίου σε σημεία γνωστών συντεταγμένων οι οποίες ανάλογα την εφαρμοζόμενη μέθοδο μπορούν να διαχωριστούν σε σύνολο εκπαίδευσης και επιβεβαίωσης ή να αποτελέσουν το σύνολο εκπαίδευσης. Για την πρώτη μέθοδο οι πρώτες 259 από τις 359 μετρήσεις του δείγματος θα αποτελέσουν το σύνολο εκπαίδευσης ενώ οι υπόλοιπες 100 το σύνολο επιβεβαίωσης. Οι 259 μετρήσεις του συνόλου εκπαίδευσης θα χρησιμοποιηθούν για την εκτίμηση των παραμέτρων του γεωστατιστικού μοντέλου. Οι 100 μετρήσεις του συνόλου επιβεβαίωσης θα αποτελέσουν τα σημεία χωρικής εκτίμησης αλλά και επιβεβαίωσης καθώς η συγκέντρωση του ρύπου σε αυτά τα σημεία είναι γνωστή.

Στο σχήμα 3.15 αναπαρίσταται η κατανομή των δειγματικών σημείων στην περιοχή μελέτης. Με κόκκινο κύκλο (ο) παρουσιάζονται τα σημεία του συνόλου επιβεβαίωσης ενώ με μαύρο αστεράκι (*) τα σημεία του συνόλου εκπαίδευσης. Αν οι

εκτιμήσεις του γεωστατιστικού μοντέλου στα σημεία επιβεβαίωσης είναι ακριβείς, τότε η γεωστατιστική ανάλυση θεωρείται αξιόπιστη και είναι δυνατή η εκτίμηση της συγκέντρωσης του χρωμίου στα σημεία ενός κανονικού πλέγματος που καλύπτει την περιοχή.



Σχήμα 3.15: Χάρτης σημείων όπου υπάρχουν μετρήσεις χρωμίου στα σημεία του συνόλου εκπαίδευσης και επιβεβαίωσης στο χώρο. Με κόκκινο κύκλο (o) παρουσιάζονται τα σημεία του συνόλου επιβεβαίωσης ενώ με μαύρο αστεράκι (*) τα σημεία του συνόλου εκπαίδευσης.

Ο έλεγχος της ακρίβειας των εκτιμήσεων του γεωστατιστικού μοντέλου επιτυγχάνεται χρησιμοποιώντας τις παρακάτω εξισώσεις μέτρησης σφάλματος (Press et al 1992).

Μέσο Απόλυτο Σφάλμα (Mean absolute error):

$$\epsilon_{MA} = \frac{1}{N} \sum_{\nu=1}^N |X_{\lambda}^*(s_{\nu}) - X_{\lambda}(s_{\nu})| \quad (3.5)$$

όπου N είναι ο αριθμός των σημείων του συνόλου επιβεβαίωσης, $X_{\lambda}^*(s_{\nu})$ είναι οι εκτιμώμενες συγκεντρώσεις χρωμίου και $X_{\lambda}(s_{\nu})$ είναι οι πραγματικές συγκεντρώσεις χρωμίου.

Μέσο Σχετικό Απόλυτο Σφάλμα (Mean absolute relative error):

$$\varepsilon_{\text{MAR}} = \frac{1}{N} \sum_{\nu=1}^N \left| \frac{X_{\lambda}^*(\mathbf{s}_{\nu}) - X_{\lambda}(\mathbf{s}_{\nu})}{X_{\lambda}(\mathbf{s}_{\nu})} \right| \quad (3.6)$$

Τετραγωνική Ρίζα Μέσου Τετραγωνικού Σφάλματος (Root mean square error):

$$\varepsilon_{\text{RMS}} = \sqrt{\frac{1}{N} \sum_{\nu=1}^N [X_{\lambda}^*(\mathbf{s}_{\nu}) - X_{\lambda}(\mathbf{s}_{\nu})]^2} \quad (3.7)$$

Τετραγωνική Ρίζα Μέσου Σχετικού Τετραγωνικού Σφάλματος (Root mean square relative error):

$$\varepsilon_{\text{RMSR}} = \sqrt{\frac{1}{N} \sum_{\nu=1}^N \left[\frac{X_{\lambda}^*(\mathbf{s}_{\nu}) - X_{\lambda}(\mathbf{s}_{\nu})}{X_{\lambda}(\mathbf{s}_{\nu})} \right]^2} \quad (3.8)$$

Ένας άλλος τρόπος ελέγχου της ακρίβειας του γεωστατιστικού μοντέλου είναι ο έλεγχος της συσχέτισης μεταξύ εκτιμώμενων και πραγματικών συγκεντρώσεων που πραγματοποιείται με την χρησιμοποίηση του γραμμικού συντελεστή συσχέτισης R^2 .

Linear Correlation Coefficient:

$$R^2 = \frac{\sum_{\nu=1}^N [X_{\lambda}(\mathbf{s}_{\nu}) - m_X][X_{\lambda}^*(\mathbf{s}_{\nu}) - m_X]}{\sqrt{\sum_{\nu=1}^N [X_{\lambda}(\mathbf{s}_{\nu}) - m_X]^2} \sqrt{\sum_{\nu=1}^N [X_{\lambda}^*(\mathbf{s}_{\nu}) - m_X]^2}} \quad (3.9)$$

3.5 Γεωστατιστική ανάλυση με την μέθοδο Kriging

Η σημειακή εκτίμηση με μεθόδους *Kriging* εκφράζεται συνήθως βάσει ενός δείγματος $X(\mathbf{s}_i), i = 1, \dots, N$ εντός μιας περιοχής Ω . Η εκτιμώμενη τιμή του πεδίου στο σημείο εκτίμησης $\mathbf{u} \in \Omega$, το οποίο δεν συμπίπτει με κανένα από τα σημεία $\mathbf{s}_i, i = 1, \dots, N$, δίνεται από το $\hat{X}(\mathbf{u})$. Στην εκτίμηση με μεθόδους *Kriging* καθορίζεται μια γειτονιά

$\omega(\mathbf{u})$ του σημείου \mathbf{u} , η οποία περιλαμβάνει $n(\mathbf{u})$ σημεία από τα $\mathbf{s}_i (i=1, \dots, N)$. Σε ιστροπική περίπτωση η γειτονιά αποτελείται από τα σημεία εντός ενός κύκλου ακτίνας $r_s(\mathbf{u})$. Η ακτίνα της γειτονιάς (ακτίνα επίδρασης) $\omega(\mathbf{u})$ είναι της ίδιας τάξης μεγέθους με το μήκος συσχέτισης του πεδίου.

Το μήκος συσχέτισης, όπως έχει αναφερθεί προηγουμένως, προσδιορίζεται από την ανάλυση του ημιβαριογράμματος. Αν η ακτίνα της γειτονιάς υπερβαίνει σημαντικά ο μήκος συσχέτισης, η εκτίμηση χρησιμοποιεί μετρήσεις οι οποίες δεν έχουν σημαντική επίδραση στο σημείο εκτίμησης. Έτσι αυξάνεται άνευ αιτίας το αριθμητικό κόστος επίλυσης του συστήματος *Kriging*. Επιπλέον χρησιμοποιώντας μικρές γειτονιές είναι δυνατόν να διακρίνονται καλύτερα τοπικές μεταβολές της μέσης τιμής από γειτονιά σε γειτονιά (Goovaerts 1997; Journel 1989).

Η εκτιμώμενη τιμή δίνεται από τον εξής γραμμικό συνδυασμό:

$$\hat{X}(\mathbf{u}) - m_X(\mathbf{u}) = \sum_{\mathbf{s}_a \in \omega(\mathbf{u})} \lambda_a [X(\mathbf{s}_a) - m_X(\mathbf{s}_a)] = \sum_{a=1}^{n(\mathbf{u})} \lambda_a [X(\mathbf{s}_a) - m_X(\mathbf{s}_a)] \quad (3.10)$$

Τα λ_a αντιπροσωπεύουν τα γραμμικά βάρη της εκτίμησης. Επομένως, η παραπάνω εξίσωση εκφράζει τη διακύμανση στο σημείο εκτίμησης ως συνάρτηση της διακύμανσης στα υπόλοιπα σημεία της γειτονιάς εκτίμησης. Η εκτιμώμενη τιμή του πεδίου αντίστοιχα δίνεται από την εξίσωση:

$$\hat{X}(\mathbf{u}) = m_X(\mathbf{u}) + \sum_{a=1}^{n(\mathbf{u})} \lambda_a [X(\mathbf{s}_a) - m_X(\mathbf{s}_a)] \quad . \quad (3.11)$$

Τόσο η ακριβής τιμή του πεδίου $X(\mathbf{u})$ όσο και η εκτίμηση $\hat{X}(\mathbf{u})$ είναι τυχαίες μεταβλητές. Το σφάλμα εκτίμησης, $X(\mathbf{u}) - \hat{X}(\mathbf{u})$ είναι επίσης μια τυχαία μεταβλητή. Οι μέθοδοι *Kriging* υπολογίζουν την βέλτιστη τιμή $\hat{X}(\mathbf{u})$ χρησιμοποιώντας τα βάρη που ελαχιστοποιούν την διασπορά του σφάλματος εκτίμησης. Οι εκτιμήτριες που βασίζονται στη μέθοδο *Kriging* είναι γνωστές ως BLUE από

τα αρχικά των λέξεων Best Linear Unbiased Estimator, δηλαδή Βέλτιστη Γραμμική Αμερόληπτη Εκτιμήτρια.

Το σφάλμα εκτίμησης είναι μια τυχαία μεταβλητή που ορίζεται ως εξής:

$$\varepsilon(\mathbf{u}) \equiv X(\mathbf{u}) - \hat{X}(\mathbf{u}). \quad (3.12)$$

Οι μέθοδοι *Kriging* προσδιορίζουν την τιμή $\hat{X}(\mathbf{u})$ βάσει του κριτηρίου ελαχιστοποίησης της διασποράς του σφάλματος εκτίμησης η οποία ορίζεται ως εξής:

$$\sigma_E^2(\mathbf{u}) \equiv \text{Var} \{ X(\mathbf{u}) - \hat{X}(\mathbf{u}) \}. \quad (3.13)$$

Η διασπορά ελαχιστοποιείται υπό συνθήκη αμεροληψίας, η οποία εξασφαλίζει ότι η μέση τιμή του σφάλματος είναι μηδέν. Αν η συνθήκη δεν ικανοποιείται αυτόματα από την εξίσωση της γραμμικής παρεμβολής δημιουργεί ένα πρόσθετο περιορισμό για τα γραμμικά βάρη. Η συνθήκη αμεροληψίας εκφράζεται ως:

$$E[\hat{X}(\mathbf{u}) - X(\mathbf{u})] = 0. \quad (3.14)$$

Η ελαχιστοποίηση της διασποράς οδηγεί σε ένα γραμμικό σύστημα εξισώσεων ως προς τα γραμμικά βάρη. Τα γνωστά στοιχεία του συστήματος αποτελούν τα δεδομένα (τιμές του πεδίου στα σημεία μέτρησης) και οι τιμές της συνάρτησης συνδιασποράς (ή του ημιβαριογράμματος) στις αποστάσεις που αντιστοιχούν στα ζεύγη των μετρημένων σημείων, καθώς και στα ζεύγη που περιλαμβάνουν ένα μετρημένο σημείο και το σημείο εκτίμησης (Goovaerts 1997; Journel 1989).

Η πρακτική εφαρμογή των μεθόδων *Kriging* βασίζεται σε ορισμένες παραδοχές.

- 1) Υποτίθεται ότι το τυχαίο πεδίο είναι δυνατό να αναλυθεί σε δύο συνιστώσες, η πρώτη εκ των οποίων εκφράζει την τάση, δηλαδή μια αργά μεταβαλλόμενη εξάρτηση, ή μια περιοδική μεταβολή, και η δεύτερη την διακύμανση του πεδίου γύρω από την τάση.

- 2) Συνήθως θεωρείται ότι η διακύμανση περιγράφεται από ένα στατιστικά ομοιογενές πεδίο μηδενικής μέσης τιμής έτσι ώστε η συνάρτηση συνδιασποράς και το ημιβαριόγραμμα να εξαρτώνται αποκλειστικά από την απόσταση μεταξύ των σημείων.
- 3) Μια εκτίμηση που βασίζεται στην ελαχιστοποίηση της διασποράς του σφάλματος είναι ακριβής αν η συνάρτηση κατανομής πιθανότητας είναι συμμετρική και κατά προτίμηση Γκαουσιανή.

Οι παραδοχές (1) και (2) συνοψίζονται με την βοήθεια των παρακάτω εξισώσεων:

$$X(\mathbf{s}) = m_X(\mathbf{s}) + \chi(\mathbf{s}), \quad \mathbf{s} \in \Omega \quad (3.15)$$

$$E[X(\mathbf{s})] = m_X(\mathbf{s}) \quad (3.16)$$

$$E[\chi(\mathbf{s})] = 0 \quad (3.17)$$

$$E[\chi(\mathbf{s} + \mathbf{r})\chi(\mathbf{s})] = c_X(\mathbf{r}) \quad (3.18)$$

Υπάρχουν ωστόσο οι εξής πέντε κύριες παραλλαγές της μεθόδου *Kriging*:

1. Απλό *Kriging* (Simple *Kriging*): Εφαρμόζεται όταν η μέση τιμή του πεδίου είναι σταθερή και γνωστή σε όλη την περιοχή Ω , δηλαδή $m_X(\mathbf{s}) = m_X$, όπου m_X μία γνωστή σταθερά.
2. Κανονικό *Kriging* (Ordinary *Kriging*): Εφαρμόζεται όταν η μέση τιμή του πεδίου θεωρείται σταθερή στη γειτονιά $\omega(\mathbf{u})$ του σημείου εκτίμησης, αλλά η τιμή της $m_X(\mathbf{s})$ μπορεί να μεταβάλλεται από γειτονιά σε γειτονιά. Σε αυτή την περίπτωση θεωρείται ότι το $m_X(\mathbf{u})$ είναι άγνωστο.
3. Καθολικό *Kriging* με πρότυπο τοπικής τάσης (Universal *Kriging*):

Εφαρμόζεται όταν η μέση τιμή του πεδίου μεταβάλλεται εντός της γειτονιάς του σημείου εκτίμησης. Σε αυτή την περίπτωση χρησιμοποιείται ως πρότυπο τάσης ένας γραμμικός συνδυασμός γνωστών συναρτήσεων (π.χ., περιοδικών συναρτήσεων και πολυνύμων).

4. *Kriging* δεικτών (Indicator *Kriging*): Χρησιμοποιείται αν η συνάρτηση κατανομής είναι έντονα ασύμμετρη με σημαντική πυκνότητα πιθανότητας στις υψηλές τιμές. Σε αυτές τις περιπτώσεις δεν είναι δυνατό να χρησιμοποιηθεί το κανονικό (ή ακόμα και το λογαριθμοκανονικό) πρότυπο πιθανότητας. Το *Kriging* δεικτών υπολογίζει την πιθανότητα η τοπική τιμή να υπερβαίνει ένα καθορισμένο όριο. Επανάληψη της διαδικασίας για πολλαπλές τιμές του ορίου (multiple indicator *Kriging*) επιτρέπει την εκτίμηση της αθροιστικής κατανομής πιθανότητας.
5. Συνδυαστικό *Kriging* (Cokriging): Χρησιμοποιείται προκειμένου να συμπεριληφθούν στην διαδικασία εκτίμησης ιδιότητες συσχετισμένες με τις εκτιμώμενες μεταβλητές. Επομένως, οι εξισώσεις του συνδυαστικού *Kriging* περιλαμβάνουν και την διασταυρωμένη συνάρτηση συμμεταβολής των μεταβλητών (Goovaerts 1997; Journel 1989).

Στην συγκεκριμένη εργασία η μέθοδος *Kriging* θα υλοποιηθεί από το λογισμικό πακέτο «Surfer V.8.0.4, 2002» το οποίο χρησιμοποιεί την μέθοδο του κανονικού *Kriging*. Στο κανονικό *Kriging* όπως προαναφέρθηκε, η μέση τιμή μέσα σε κάθε γειτονιά παραμένει σταθερή αλλά μεταβάλλεται από γειτονιά σε γειτονιά. Χρησιμοποιώντας την βασική εξίσωση του *Kriging*, η εκτίμηση με την μέθοδο του κανονικού *Kriging* δίνεται από τις εξής εξισώσεις:

$$\hat{X}(\mathbf{u}) = \sum_{a=1}^{n(\mathbf{u})} \lambda_a X(\mathbf{s}_a) \quad (3.19)$$

$$\sum_{a=1}^{n(\mathbf{u})} \lambda_a = 1 \quad (3.20)$$

Η δεύτερη εξίσωση εκφράζει την συνθήκη αμεροληψίας. Η διακύμανση εκφράζεται με την βοήθεια της εξίσωσης $\chi(\mathbf{s}) = X(\mathbf{s}) - m_X(\mathbf{s})$, όπου το $m_X(\mathbf{s})$ είναι σταθερό μέσα σε κάθε γειτονιά. Στη μέθοδο του κανονικού *Kriging* δεν είναι απαραίτητη η γνώση της μέσης τιμής. Σε αυτή την περίπτωση πρέπει να υπολογιστεί το ελάχιστο σφάλμα με τον περιορισμό που επιβάλλει η συνθήκη αμεροληψίας στις τιμές των βαρών. Επομένως, χρειάζεται να χρησιμοποιηθεί η μέθοδο των συντελεστών Lagrange. Η διασπορά του σφάλματος υπολογίζεται σε αυτή την περίπτωση ως εξής:

$$\begin{aligned} \sigma_E^2(\mathbf{u}) = \text{Var}[X(\mathbf{u})] + \sum_{a=1}^{n(\mathbf{u})} \sum_{\beta=1}^{n(\mathbf{u})} \lambda_a \lambda_\beta E[\chi(\mathbf{s}_a) \chi(\mathbf{s}_\beta)] - 2 \sum_{a=1}^{n(\mathbf{u})} \lambda_a E[\chi(\mathbf{s}_a) \chi(\mathbf{u})] \\ + 2\mu \sum_{\beta=1}^{n(\mathbf{u})} (\lambda_\beta - 1). \end{aligned} \quad (3.21)$$

Η σταθερά 2μ είναι ο συντελεστής Lagrange για την συνθήκη αμεροληψίας.

Χρησιμοποιώντας την συνάρτηση συνδιασποράς η παραπάνω σχέση εκφράζεται ως εξής:

$$\begin{aligned} \sigma_E^2(\mathbf{u}) = \text{Var}[X(\mathbf{u})] + \sum_{a=1}^{n(\mathbf{u})} \sum_{\beta=1}^{n(\mathbf{u})} \lambda_a \lambda_\beta c_X(\mathbf{s}_a, \mathbf{s}_\beta) - 2 \sum_{a=1}^{n(\mathbf{u})} \lambda_a c_X(\mathbf{s}_a, \mathbf{u}) \\ + 2\mu \sum_{\beta=1}^{n(\mathbf{u})} (\lambda_\beta - 1). \end{aligned} \quad (3.22)$$

Η ελάχιστη τιμή του σφάλματος εκτίμησης προσδιορίζεται από την επίλυση του συστήματος των εξισώσεων που προκύπτει από τον μηδενισμό των παραγώγων του σφάλματος ως προς τα βάρη και την παράμετρο μ ,

$$\frac{\partial \sigma_E^2(\mathbf{u})}{\partial \lambda_a} = 0, \quad a = 1, \dots, n(\mathbf{u}) \quad (3.23)$$

$$\frac{\partial \sigma_E^2(\mathbf{u})}{\partial \mu} = 0 \quad (3.24)$$

Αυτές οι συνθήκες οδηγούν στις επόμενες εξισώσεις για τα γραμμικά βάρη,

$$\sum_{\beta=1}^{n(\mathbf{u})} \lambda_{\beta} c_X(\mathbf{s}_a - \mathbf{s}_{\beta}) + \mu = c_X(\mathbf{s}_a - \mathbf{u}), \quad a=1, \dots, n(\mathbf{u}) \quad (3.25)$$

$$\sum_{a=1}^{n(\mathbf{u})} \lambda_a = 1 \quad (3.26)$$

Το σύστημα των εξισώσεων για τα βάρη εκφράζεται από το επόμενο γραμμικό σύστημα,

$$\begin{bmatrix} c_X(\mathbf{s}_1, \mathbf{s}_1) & \dots & \dots & c_X(\mathbf{s}_1, \mathbf{s}_n) & 1 \\ c_X(\mathbf{s}_2, \mathbf{s}_1) & \dots & \dots & c_X(\mathbf{s}_1, \mathbf{s}_2) & 1 \\ \vdots & & & \vdots & \vdots \\ c_X(\mathbf{s}_n, \mathbf{s}_1) & & & c_X(\mathbf{s}_n, \mathbf{s}_n) & 1 \\ 1 & 1 & \dots & \dots & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ \mu \end{bmatrix} = \begin{bmatrix} c_X(\mathbf{s}_1, \mathbf{u}) \\ c_X(\mathbf{s}_2, \mathbf{u}) \\ \vdots \\ c_X(\mathbf{s}_n, \mathbf{u}) \\ 1 \end{bmatrix}. \quad (3.27)$$

Το μέσο τετραγωνικό σφάλμα εκτίμησης δίνεται από την ακόλουθη εξίσωση (Goovaerts 1997; Journel 1989)

$$\sigma_E^2(\mathbf{u}) = \sigma_X^2 - \sum_{a=1}^{n(\mathbf{u})} \lambda_a c_X(\mathbf{u}, \mathbf{s}_a) - \mu. \quad (3.28)$$

Στο κανονικό *Kriging* η εκτιμώμενη συγκέντρωση στο σημείο εκτίμησης εκφράζεται ως ένας γραμμικός συνδυασμός των τιμών των γειτονικών σημείων. Οι γραμμικοί συντελεστές (παράμετροι) του κανονικού *Kriging* προσδιορίζονται από το ημιβαριόγραμμα χρησιμοποιώντας ένα κριτήριο βελτιστοποίησης (ελαχιστοποίηση του μέσου τετραγωνικού σφάλματος). Η μέση τιμή των εκτιμήσεων είναι ίση με την μέση τιμή του τυχαίου πεδίου. Συνεπώς το κανονικό *Kriging* είναι ένας βέλτιστος γραμμικά αμερόληπτος τρόπος εκτίμησης.

3.6 Εφαρμογή της μεθόδου κανονικού Kriging

Η εφαρμογή της εκτίμησης με την μέθοδο κανονικού *Kriging* προϋποθέτει τον προσδιορισμό του βέλτιστου πρότυπου ημιβαριογράμματος. Όπως προαναφέρθηκε, αυτό επιτυγχάνεται χρησιμοποιώντας την μεθοδολογία της διασταυρωμένης επιβεβαίωσης. Πιο συγκεκριμένα η συγκεκριμένη μέθοδος χρησιμοποιείται για την εκτίμηση με την μέθοδο κανονικού *Kriging*, ωστόσο για να πραγματοποιηθεί η εκτίμηση είναι απαραίτητος ο υπολογισμός των βέλτιστων παραμέτρων της μεθόδου.

Οι παράμετροι προκύπτουν από την προσαρμογή του πειραματικού ημιβαριογράμματος του δείγματος σε ένα θεωρητικό μοντέλο. Για να διαπιστωθεί ποιες είναι οι βέλτιστες παράμετροι και ποιο το βέλτιστο θεωρητικό μοντέλο ημιβαριογράμματος από αυτά που εξετάστηκαν χρησιμοποιούνται στην εκτίμηση, η οποία αποτελεί και επιβεβαίωση των αποτελεσμάτων καθώς πραγματοποιείται σε γνωστά σημεία.

Από την επιβεβαίωση των αποτελεσμάτων της γεωστατιστικής ανάλυσης παρατηρείται ότι οι παράμετροι πέντε θεωρητικών μοντέλων ημιβαριογράμματος δίνουν το μικρότερο σφάλμα εκτίμησης. Τα αποτελέσματα της επιβεβαίωσης παρουσιάζονται στον πίνακα 3.5. Εξετάζοντας τα προσεκτικά προκύπτει ότι δύο θεωρητικά μοντέλα ημιβαριογράμματος δίνουν τα καλύτερα αποτελέσματα.

Ο συνδυασμός του *Εκθετικού* μοντέλου με το μοντέλο *Φαινομένου Πυρήνα*, όπως και ο συνδυασμός του *Σφαιρικού* μοντέλου ημιβαριογράμματος με το μοντέλο *Φαινομένου Πυρήνα*. Αυτά τα μοντέλα έχουν το μικρότερο σφάλμα εκτίμησης ενώ παρουσιάζουν τις υψηλότερες συσχετίσεις (R^2) μεταξύ των εκτιμώμενων συγκεντρώσεων και των πραγματικών (64% και 63% αντίστοιχα). Ο βασικός λόγος της ασυμφωνίας των εκτιμήσεων του γεωστατιστικού μοντέλου είναι η αδυναμία της μεθόδου να λάβει υπόψη με ακρίβεια τις ακραίες τιμές του κανονικού τυχαίου πεδίου καθώς και η ύπαρξη σημαντικής τυχαίας συνιστώσας (φαινόμενο πυρήνα).

Πίνακας 3.5: Σφάλμα εκτίμησης μεθόδου κανονικού *Kriging* για διαφορετικά θεωρητικά μοντέλα ημιβαριογράμματος.

Μοντέλο Ημιβαριογράμματος	MAE (3.5) (ppm)	MARE (3.6) (%)	RMSE (3.7) (ppm)	RMSRE (3.8) (%)	R ² (3.9)
Εκθετικό & Φαινόμενο Πυρήνα $r_s = 0.7km$	6.14	21.70	8.24	48.00	0.64
Φαινόμενο πυρήνα & Hole-Wave $r_s = 0.8km$	6.83	25.00	8.84	60.00	0.58
Σφαιρικό & Φαινόμενο Πυρήνα $r_s = 0.8km$	6.22	21.60	8.25	45.00	0.63
Σφαιρικό $r_s = 0.8km$	6.30	21.80	8.46	48.00	0.63
Τετραγωνικό $r_s = 0.7km$	6.28	22.00	8.45	48.50	0.63

Πίνακας 3.6: Σφάλμα εκτίμησης μεθόδου κανονικού *Kriging* εξαιρουμένων την μέγιστης και της ελάχιστης συγκέντρωσης του δείγματος για διαφορετικά θεωρητικά μοντέλα ημιβαριογράμματος.

Μοντέλο Ημιβαριογράμματος	MAE (3.5) (ppm)	MARE (3.6) (%)	RMSE (3.7) (ppm)	RMSRE (3.8) (%)	R ² (3.9)
Εκθετικό & Φαινόμενο Πυρήνα $r_s = 0.7km$	5.99	19.60	7.99	33.00	0.64
Φαινόμενο πυρήνα & Hole-Wave $r_s = 0.8km$	6.64	21.00	8.74	49.00	0.58
Σφαιρικό & Φαινόμενο Πυρήνα $r_s = 0.8km$	6.13	20.00	8.10	33.00	0.63
Σφαιρικό $r_s = 0.8km$	6.20	20.00	8.30	33.00	0.63
Τετραγωνικό $r_s = 0.7km$	6.17	20.00	8.30	33.00	0.63

Αφαιρώντας τις ακραίες τιμές, δηλαδή την μέγιστη και την ελάχιστη τιμή του δείγματος χρωμίου και εφαρμόζοντας ξανά την μέθοδο κανονικού *Kriging* στα αποτελέσματα της επιβεβαίωσης (Πίνακας 3.6), παρατηρείται ότι αυτές οι ακραίες τιμές επηρεάζουν την τετραγωνική ρίζα του μέσου σχετικού τετραγωνικού σφάλματος (RMSRE). Συγκρίνοντας το με το αντίστοιχο στον πίνακα 3.5, υπάρχει μια αξιοσημείωτη μείωση.

Τα θεωρητικά μοντέλα ημιβαριογράμματος (*Εκθετικό & Φαινόμενο Πυρήνα*) και (*Σφαιρικό & Φαινόμενο Πυρήνα*) δίνουν τα καλύτερα αποτελέσματα και σε αυτήν την περίπτωση, αφού έχουν το μικρότερο σφάλμα εκτίμησης και την υψηλότερη συσχέτιση (R^2). Σε αυτό το σημείο πρέπει να αναφερθεί ότι η αποτελεσματικότερη ακτίνα επίδρασης r_s που χρησιμοποιήθηκε σε όλα τα θεωρητικά μοντέλα ημιβαριογράμματος για τον προσδιορισμό των πλησιέστερων γειτόνων κάθε σημείου του δείγματος είναι μεταξύ 0.7 και 0.8km.

3.7 Εφαρμογή μοντέλου Σπαρτιάτικων Τυχαίων Πεδίων

Η εφαρμογή των μαθηματικών εξισώσεων οι οποίες περιγράφουν τη λειτουργία του γεωστατιστικού μοντέλου ΣΤΠ και παρουσιάζονται στο κεφάλαιο 2 πραγματοποιήθηκε σε περιβάλλον Matlab[®]. Στην συνέχεια το γεωστατιστικό μοντέλο εφαρμόστηκε στο δείγμα χρωμίου. Σε αυτήν την περίπτωση χρησιμοποιήθηκε η μεθοδολογία γεωστατιστικής ανάλυσης κατά την οποία διαχωρίζεται το δείγμα σε σύνολο εκπαίδευσης και σύνολο επιβεβαίωσης.

Η κατανομή του δείγματος χρωμίου μοντελοποιείται ως ένα κανονικό και στατιστικά ομοιογενές ισοτροπικό τυχαίο πεδίο. Η συνάρτηση πυκνότητας πιθανότητας του πεδίου προσδιορίζεται από τις εξισώσεις (2.1) και (2.2) λαμβάνοντας υπόψη ότι ο χώρος που ορίζεται το πεδίο είναι δύο διαστάσεων. Οι παράμετροι του μοντέλου Σπαρτιάτικων Τυχαίων Πεδίων η_0 , η_1 και ξ υπολογίζονται βάσει των εξισώσεων που περιγράφονται στην παράγραφο 2.3.

Αρχικά καθορίζεται το μέγεθος των κυψελίδων του τετραγωνικού πλέγματος που περιλαμβάνει τα σημεία του δείγματος καθώς το διαθέσιμο δείγμα κατανέμεται ακανόνιστα στον χώρο. Εξετάζοντας διαφορετικά μεγέθη κυψελίδων αποδεικνύεται

μετά από αρκετές δοκιμές ότι τα καλύτερα αποτελέσματα επιτυγχάνονται για μέγεθος κυψελίδων 0.34×0.34 km.

Στη συνέχεια για να προσδιοριστούν οι στατιστικοί περιορισμοί του δείγματος $\bar{S}_0, \bar{S}_1, \bar{S}_2$ οι 259 μετρήσεις του συνόλου εκπαίδευσης (συντεταγμένες και συγκεντρώσεις χρωμίου) εισάγονται στο γεωστατιστικό μοντέλο χρησιμοποιώντας τις εξισώσεις (2.20), (2.21), (2.24). Οι εξισώσεις αντιστοιχούν σε ακανόνιστη κατανομή δείγματος χρησιμοποιώντας υπόβαθρο τετραγωνικού πλέγματος. Οι τιμές των στατιστικών περιορισμών του δείγματος υπολογίζονται βάσει του κώδικα «*iregg_constraint3*» (Παράρτημα Ε) σε $S_0 = 119.6$, $S_1 = 55.9$, και $S_2 = 682.8$.

Οι στατιστικοί περιορισμοί του μοντέλου προσδιορίζονται από τις εξισώσεις (2.12), (2.13), (2.14) οι οποίες εφαρμόζονται από τον κώδικα «*distfunc2*» (Παράρτημα Ζ). Η συναρτησιακή απόσταση μεταξύ των στατιστικών περιορισμών του δείγματος και αυτών του μοντέλου υπολογίζεται από τον ίδιο κώδικα. Ο κώδικας αυτός καλείται από το πρόγραμμα «*dist_minimirreg*» (Παράρτημα Η) το οποίο υλοποιεί την βελτιστοποίηση, εκτιμώντας τελικά τις παραμέτρους του μοντέλου ΣΤΠ.

Πιο συγκεκριμένα οι παράμετροι του μοντέλου η_1 και ξ προσδιορίζονται από την ελαχιστοποίηση της «Φ» (2.15) η οποία επιτυγχάνεται από την έμφυτη συνάρτηση «*fmincon*». Ο αλγόριθμος χρησιμοποιείται για επίλυση προβλημάτων μη γραμμικής βελτιστοποίησης υπό συνθήκες σε περιβάλλον Matlab[®].

Ο αλγόριθμος «*fmincon*» απαιτεί αρχικές εκτιμήσεις για τις παραμέτρους $\eta_{1,0}$ και ξ_0 . Για αυτό πραγματοποιήθηκαν δοκιμές με αρκετούς συνδυασμούς (Πίνακας 3.7) ώστε να διαπιστωθεί η ευαισθησία των βέλτιστων παραμέτρων στις αρχικές εκτιμήσεις. Από τον παρακάτω πίνακα διαπιστώνεται ότι οι βέλτιστες τιμές των παραμέτρων δεν παρουσιάζουν μεγάλη ευαισθησία σε διαφορετικούς συνδυασμούς αρχικών εκτιμήσεων.

Πίνακας 3.7: Βέλτιστες τιμές παραμέτρων για διαφορετικούς συνδυασμούς αρχικών εκτιμήσεων.

Αρχική Τιμή $\eta_{l,0}$	Αρχική Τιμή ξ_0 (km)	Βέλτιστη Τιμή η_1	Βέλτιστη Τιμή ξ (km)
-1	1	378	1.32
0.5	0.5	360	1.35
1	1	386	1.31
1.5	0.5	355	1.33
1.5	1	392	1.30
2	1.5	377	1.31
3	1	380	1.31
4	2	383	1.31
5	1	371	1.32
1	2	371	1.32
1	3	374	1.32
2	4	386	1.31
3	4	387	1.31
1	5	390	1.30
2	5	394	1.30
3	5	396	1.31

Ως βέλτιστες παράμετροι του μοντέλου επιλέγονται οι τιμές $\eta_1=386$ και $\xi=1.31$ km.

Η βέλτιστη παράμετρος η_0 που εκφράζει το μέγεθος της μεταβλητότητας στον χώρο υπολογίζεται από την εξίσωση (2.17). Το αποτέλεσμα προκύπτει από τον κώδικα «*dist_minimirreg*» (Παράρτημα Η) και υπολογίζεται σε 5.78×10^4 .

Το επόμενο βήμα, μετά τον υπολογισμό των παραμέτρων του μοντέλου ΣΤΠ, είναι η εκτίμηση των συγκεντρώσεων του χρωμίου στα σημεία του συνόλου επιβεβαίωσης. Οι μαθηματικές εξισώσεις οι οποίες χρησιμοποιούνται περιγράφονται στην παράγραφο 2.4 και εφαρμόζονται από τον κώδικα «*prediction_SSRF*» (Παράρτημα Θ).

Για να επιτευχθεί αυτός ο στόχος χρησιμοποιείται η εξίσωση (2.40) η οποία αντιστοιχεί σε μορφή χωρικής εξάρτησης $\eta_1 \geq 2$ καθώς η τιμή της παραμέτρου η_1 του μοντέλου υπολογίστηκε ίση με 386. Βάσει των βέλτιστων παραμέτρων του μοντέλου πραγματοποιείται η επιλογή των πλησιέστερων γειτόνων του κάθε σημείου εκτίμησης

ώστε να καθοριστούν οι τιμές των τοπικών βαρών A , B και να πραγματοποιηθεί η εκτίμηση. Η επιλογή των πλησιέστερων γειτόνων πραγματοποιείται χρησιμοποιώντας την ακτίνα επίδρασης r_s , η οποία καθορίζεται με βάση την βέλτιστη ακτίνα συσχέτισης ξ .

Η αποτελεσματικότερη ακτίνα επίδρασης r_s που χρησιμοποιήθηκε για τον προσδιορισμό των πλησιέστερων γειτόνων κάθε εκτιμώμενου σημείου είναι ίση με 0.86km. Ο υπολογισμός των βαρών A και B σε κάθε σημείο εκτίμησης πραγματοποιείται σύμφωνα με την εξίσωση (2.45), η οποία επιλύεται με την μέθοδο της πολλαπλής γραμμικής παλινδρόμησης και υλοποιείται στην Matlab® με την εντολή «regress».

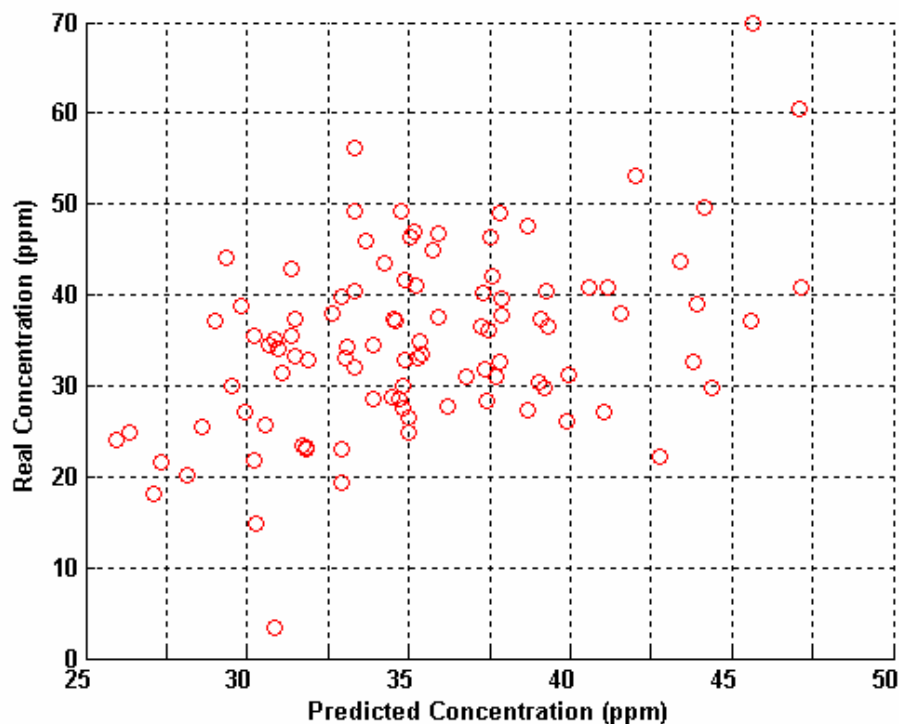
Σε αυτήν την εξίσωση για τον υπολογισμό των τοπικών βαρών χρησιμοποιούνται οι τιμές των συγκεντρώσεων χρωμίου και οι συναρτήσεις βάσης (2.43) και (2.44) στα γειτονικά σημεία. Οι συναρτήσεις βάσης θεωρούνται ισοτροπικές. Το τελευταίο βήμα είναι η εκτίμηση των συγκεντρώσεων χρωμίου, στα αντίστοιχα σημεία. Η παράμετρος η_0 δεν επηρεάζει τις εκτιμώμενες συγκεντρώσεις, αλλά επιδρά στην αβεβαιότητα των εκτιμήσεων.

Τα αποτελέσματα περιλαμβάνουν τις εκτιμώμενες συγκεντρώσεις χρωμίου σε 100 σημεία μέτρησης γνωστών συντεταγμένων, τα οποία αποτελούν το σύνολο επιβεβαίωσης. Η σύγκριση μεταξύ των εκτιμώμενων και των πραγματικών συγκεντρώσεων στα αντίστοιχα σημεία μέτρησης παρέχει μια εικόνα της ακρίβειας του μοντέλου. Το σχήμα 3.16 είναι ένα διάγραμμα διασποράς στο οποίο συγκρίνονται οι εκτιμώμενες συγκεντρώσεις χρωμίου με τις πραγματικές στα αντίστοιχα σημεία εκτίμησης.

Για να διαπιστωθεί η ακρίβεια της εκτίμησης του μοντέλου είναι απαραίτητος ο υπολογισμός του σφάλματος εκτίμησης το οποίο ποσοτικοποιείται από τις εξισώσεις μέτρησης σφάλματος (3.5 έως 3.9). Εισάγοντας τις εξισώσεις στην Matlab®, εκτός από τον γραμμικό συντελεστή συσχέτισης ο οποίος δίνεται από την εντολή «corrcoef», υπολογίζεται το σφάλμα εκτίμησης που παρουσιάζεται στον πίνακα 3.8.

Πίνακας 3.8: Ανάλυση σφάλματος εκτίμησης ΣΤΠ σε 100 σημεία του συνόλου επιβεβαίωσης.

Μέσο Απόλυτο Σφάλμα (MAE)	7.02 ppm
Μέσο Σχετικό Απόλυτο Σφάλμα (MARE):	29.00 %
Μέσο Τετραγωνικό Σφάλμα (RMSE):	8.88 ppm
Μέσο Σχετικό Τετραγωνικό Σφάλμα (RMSRE):	87.00 %
Συντελεστής Συσχέτισης (R^2)	0.44



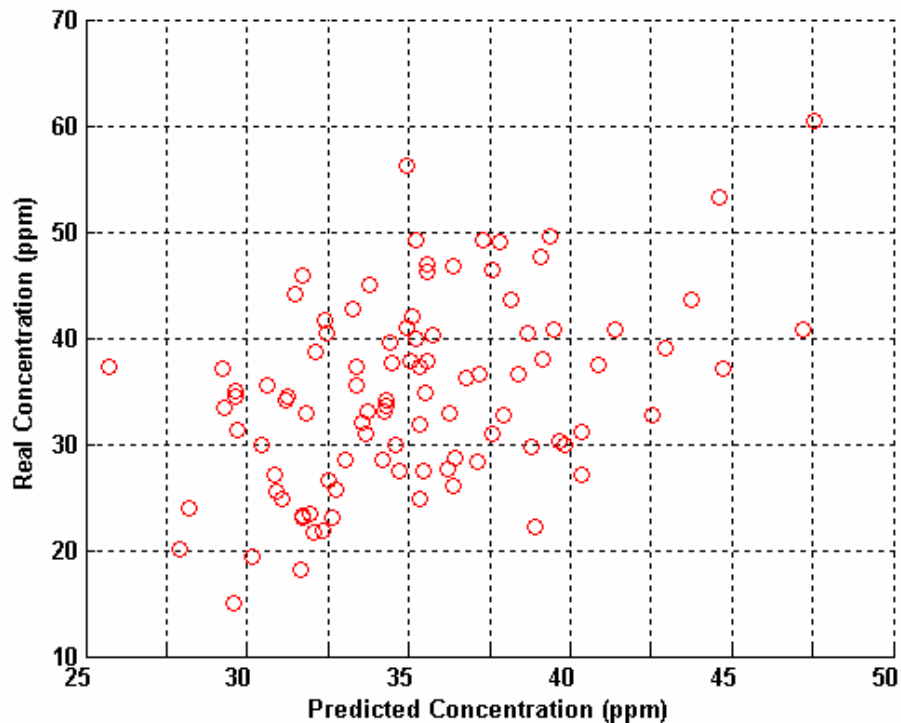
Σχήμα 3.16: Διάγραμμα διασποράς εκτιμώμενων και πραγματικών συγκεντρώσεων χρωμίου. Με κόκκινο (ο) δηλώνεται το κάθε σημείο του συνόλου επιβεβαίωσης το οποίο αντιστοιχίζει την πραγματική με την εκτιμώμενη συγκέντρωση χρωμίου. Ο άξονας (x) παρουσιάζει τις εκτιμώμενες συγκεντρώσεις χρωμίου ενώ ο άξονας (y) τις πραγματικές συγκεντρώσεις.

Όπως παρατηρείται από τον πίνακα 3.8 το σφάλμα εκτίμησης είναι αρκετά μεγάλο. Εξετάζοντας ξανά το δείγμα γίνεται αντιληπτό ότι στα σημεία που χρησιμοποιήθηκαν για την επιβεβαίωση των εκτιμήσεων του μοντέλου η μικρότερη πραγματική συγκέντρωση είναι 3.3ppm και η μεγαλύτερη 70ppm. Στα σημεία που χρησιμοποιήθηκαν για τον υπολογισμό των παραμέτρων του μοντέλου η μικρότερη συγκέντρωση είναι 18ppm και η μεγαλύτερη 65ppm.

Αυτό σημαίνει ότι δημιουργούνται δυο ακραίες τιμές που είναι αδύνατο να προβλεφθούν από το μοντέλο καθώς το αριθμητικό διάστημα των τιμών του συνόλου

επιβεβαίωσης είναι μεγαλύτερο (3.3ppm-70ppm), από το διάστημα τιμών του συνόλου εκπαίδευσης (18ppm-65ppm). Για το λόγο αυτό επαναλαμβάνεται η εφαρμογή της μεθόδου εκτίμησης αφαιρώντας τις ακραίες τιμές και τις αντίστοιχες συντεταγμένες τους από το σύνολο επιβεβαίωσης.

Η εκτίμηση τώρα πραγματοποιείται από τον κώδικα «*prediction_SSRF_98*» (Παράρτημα Ι) σε 98 από τα 100 σημεία. Η αποτελεσματικότερη ακτίνα επίδρασης r_s που χρησιμοποιήθηκε για τον προσδιορισμό των πλησιέστερων γειτόνων ισούται με 1km. Τα αποτελέσματα του γεωστατιστικού μοντέλου σε αυτή την περίπτωση παρουσιάζονται στο σχήμα 3.17.



Σχήμα 3.17: Διάγραμμα διασποράς εκτιμώμενων και πραγματικών συγκεντρώσεων χρωμίου. Με κόκκινο (o) δηλώνεται το κάθε σημείο του συνόλου επιβεβαίωσης το οποίο αντιστοιχίζει την πραγματική με την εκτιμώμενη συγκέντρωση χρωμίου. Ο άξονας (x) παρουσιάζει τις εκτιμώμενες συγκεντρώσεις χρωμίου ενώ ο άξονας (y) τις πραγματικές συγκεντρώσεις.

Το μοντέλο χωρίς τις ακραίες τιμές ανταποκρίνεται καλύτερα, αλλά αυτό πρέπει να ελεγχθεί υπολογίζοντας ξανά το σφάλμα εκτίμησης (Πίνακας 3.9).

Πίνακας 3.9: Μέτρηση σφάλματος εκτίμησης σε 98 σημεία του συνόλου επιβεβαίωσης.

Μέσο Απόλυτο Σφάλμα (MAE)	6.64 ppm
Μέσο Σχετικό Απόλυτο Σφάλμα (MARE):	20.00 %
Μέσο Τετραγωνικό Σφάλμα (RMSE):	7.91 ppm
Μέσο Σχετικό Τετραγωνικό Σφάλμα (RMSRE):	27.00 %
Συντελεστής Συσχέτισης (R^2)	0.44

Η εκτίμηση βελτιώνεται καθώς το μέσο σχετικό απόλυτο σφάλμα μειώνεται σε 20% από 29% ενώ το μέσο σχετικό τετραγωνικό σφάλμα σε 27% από 87%. Η ασυμφωνία αυτή των εκτιμήσεων του γεωστατιστικού μοντέλου οφείλεται σε τρεις παράγοντες: (1) Η παρούσα υλοποίηση της μεθόδου ΣΤΠ δεν παρέχει μεγάλη ακρίβεια στις ακραίες τιμές. (2) Η λύση των εξισώσεων εκτίμησης υποθέτει ισοτροπική στατιστική εξάρτηση ενώ το δείγμα συγκεντρώσεων χρωμίου είναι οριακά ισοτροπικό. (3) Οι συναρτήσεις βάσης είναι ισοτροπικές. Αυτός ο περιορισμός είναι ισχυρότερος από την στατιστική ισοτροπία, ωστόσο υπάρχει δυνατότητα να αρθεί σε μελλοντικό στάδιο.

3.7.1 Εφαρμογή της μεθόδου Delaunay

Ο δεύτερος τρόπος υπολογισμού των παραμέτρων του μοντέλου ΣΤΠ βασίζεται στα τρίγωνα *Delaunay*. Με αυτόν τον τρόπο επιχειρείται μια διαφορετική προσέγγιση για τον προσδιορισμό των στατιστικών περιορισμών του δείγματος. Μέχρι τώρα για τον προσδιορισμό των στατιστικών περιορισμών του δείγματος τυχαίου πεδίου θεωρείται ότι τα σημεία του κατανέμονται στις κυψελίδες ενός τετραγωνικού πλέγματος υποβάθρου. Εδώ επιχειρείται μια διαφορετική προσέγγιση βασισμένη στην μέθοδο *Delaunay* η οποία περιγράφεται στην παράγραφο 2.3.3. Η μέθοδος βασίζεται στους πλησιέστερους γείτονες του εκάστοτε σημείου του δείγματος, ενώ οι στατιστικοί περιορισμοί του υπολογίζονται χρησιμοποιώντας την μέθοδο πεπερασμένων διαφορών (*Finite Difference Method*).

Η εφαρμογή τόσο της μεθόδου των τριγώνων *Delaunay* όσο και των μαθηματικών εξισώσεων της μεθόδου πεπερασμένων διαφορών υλοποιείται από τον κώδικα «*delaunay_application*» (Παράρτημα Β). Λαμβάνοντας υπόψη τα συμπεράσματα που εξήχθησαν από την επιβεβαίωση των αποτελεσμάτων της μεθόδου *Delaunay* στην

παράγραφο 2.3.4 εφαρμόζεται η μέθοδος στα σημεία μέτρησης του συνόλου εκπαίδευσης για τον προσδιορισμό των πλησιέστερων γειτόνων του εκάστοτε σημείου (Σχήμα 2.6).

Εφαρμόζοντας το ανάπτυγμα Taylor σύμφωνα με την εξίσωση (2.27) στην γειτονία κάθε σημείου μέτρησης του συνόλου εκπαίδευσης και επιλύοντας το γραμμικό σύστημα (2.28) προσδιορίζονται οι τοπικές παράγωγοι. Έπειτα εισάγονται οι τοπικές παραγωγοί στις εξισώσεις (2.30), (2.31) και υπολογίζονται οι στατιστικοί περιορισμοί του δείγματος $S_1 = 7043$, $S_2 = 467200$. Ο στατιστικός περιορισμός $S_0 = 120$ προσδιορίζεται όπως και στην μέθοδο κυψελίδων από την εξίσωση (2.6) αφού δεν απαιτεί πληροφορίες για την μορφή του πλέγματος.

Οι τιμές των στατιστικών περιορισμών S_1 και S_2 του δείγματος με την μέθοδο *Delaunay* είναι πολύ μεγαλύτερες σε σχέση με τις αντίστοιχες, ($S_1 = 55.9$ και $S_2 = 682.8$) που υπολογίστηκαν με την μέθοδο των κυψελίδων. Αυτό έχει ως αποτέλεσμα η μέθοδος βελτιστοποίησης για τον υπολογισμό των παραμέτρων του μοντέλου να μην συγκλίνει μετά από 1000 επαναλήψεις που έχουν τεθεί ως όριο.

Το μέγεθος των τιμών των στατιστικών περιορισμών S_1 και S_2 οφείλεται στην ύπαρξη θορύβου, δηλαδή σε ασυσχέτιστες μεταβολές μεταξύ γειτονικών σημείων οι οποίες επηρεάζουν σημαντικά τους υπολογισμούς των παραγώγων. Άρα η μέθοδος *Delaunay* δεν χρησιμοποιείται περαιτέρω στην γεωστατιστική ανάλυση του δείγματος συγκεντρώσεων χρωμίου.

3.7.2 Εφαρμογή ΣΤΠ με την μέθοδο της διασταυρωμένης επιβεβαίωσης

Για να είναι εφικτή η σύγκριση των αποτελεσμάτων των δύο γεωστατιστικών μοντέλων (Κανονικό *Kriging* και ΣΤΠ) η μεθοδολογία της γεωστατιστικής ανάλυσης του μοντέλου ΣΤΠ τροποποιείται ώστε να επιτρέπει την εφαρμογή της μεθόδου της διασταυρωμένης επιβεβαίωσης που χρησιμοποιείται στο λογισμικό πακέτο «Surfer V.8.0.4, 2002»

Οι εξισώσεις που περιγραφούν το μοντέλο ΣΤΠ δεν τροποποιούνται. Αλλάζει μόνο το σύνολο εκπαίδευσης το οποίο σε αυτήν τη περίπτωση περιλαμβάνει όλα τα σημεία, (359), του δείγματος. Η διαδικασία υπολογισμού των στατιστικών περιορισμών και των παραμέτρων του δείγματος είναι ανάλογη με αυτήν που περιγράφηκε στην παράγραφο 3.6.

Σε αυτήν την περίπτωση οι στατιστικοί περιορισμοί του δείγματος υπολογίζονται από τον κώδικα «*bckg_constr*» (Παράρτημα Κ). Για τον υπολογισμό των βέλτιστων παραμέτρων η_1 και ξ του μοντέλου χρησιμοποιείται ο αλγόριθμος «*fmincon*» της Matlab[®], ο οποίος απαιτεί αρχικές τιμές για τις παραμέτρους. Η μέθοδος βελτιστοποίησης εφαρμόζεται από το πρόγραμμα «*constr_match*» (Παράρτημα Λ) σε συνδυασμό με το πρόγραμμα «*distfunc2*» (Παράρτημα Ζ) το οποίο προσδιορίζει τους στατιστικούς περιορισμούς του μοντέλου και την συναρτησιακή απόσταση μεταξύ των στατιστικών περιορισμών του δείγματος και του μοντέλου.

Η μέθοδος χωρικής εκτίμησης είναι ανάλογη αυτής που περιγράφηκε στην παράγραφο 3.6. Αντί να υπάρχει ένα σύνολο σημείων επιβεβαίωσης στα οποία η εκτίμηση εξαρτάται από τα σημεία του συνόλου εκπαίδευσης, χρησιμοποιούνται για την εκτίμηση όλα τα σημεία του δείγματος. Αφαιρώντας ένα σημείο από το δείγμα κάθε φορά και βασιζόμενοι στα υπόλοιπα πραγματοποιείται η εκτίμηση στο σημείο αυτό. Ο τροποποιημένος κώδικας σε περιβάλλον Matlab[®] που πραγματοποιεί την επιβεβαίωση των εκτιμήσεων σε όλα τα σημεία του δείγματος ονομάζεται «*infer_ssrrf2*» (Παράρτημα Μ).

Εφαρμόζοντας τον τροποποιημένο αλγόριθμο και εξετάζοντας διαφορετικά μεγέθη πλέγματος και διαφορετικούς συνδυασμούς αρχικών παραμέτρων, προσδιορίζονται δύο περιπτώσεις που δίνουν το μικρότερο σφάλμα εκτίμησης. Και στις δύο χρησιμοποιήθηκε πλέγμα μεγέθους 27x30. Οι στατιστικοί περιορισμοί του δείγματος υπολογίστηκαν $S_0 = 114$, $S_1 = 50$, και $S_2 = 574$.

Για τον προσδιορισμό των βέλτιστων παραμέτρων στην πρώτη δόθηκαν ως αρχικές παράμετροι $\eta_{1,0} = -1.5$ και $\xi_0 = 3$ ενώ στην δεύτερη $\eta_{1,0} = 100$ και $\xi_0 = 0.1$. Οι

παράμετροι του μοντέλου ΣΤΠ που προκύπτουν είναι $\eta_1 = 201.6$, $\xi = 0.35 \text{ km}$ και $\eta_1 = 100$, $\xi = 0.33 \text{ km}$ αντίστοιχα για τις δύο περιπτώσεις.

Η αποτελεσματικότερη ακτίνα επίδρασης r_s που χρησιμοποιήθηκε για τον προσδιορισμό των πλησιέστερων γειτόνων κάθε σημείου του δείγματος καθορίζεται με βάση την βέλτιστη ακτίνα συσχέτισης ξ και είναι $r_s = 0.37 \text{ km}$. Όπως φαίνεται και στις δυο περιπτώσεις η r_s είναι πολύ κοντά στην ακτίνα συσχέτισης. Η επιβεβαίωση του μοντέλου ΣΤΠ πραγματοποιείται χρησιμοποιώντας τις εξισώσεις (3.5) έως (3.9) και τα αποτελέσματα της επιβεβαίωσης παρουσιάζονται στους πίνακες 3.10 και 3.11.

Πίνακας 3.10: Σφάλμα εκτίμησης μοντέλου Σπαρτιάτικων Τυχαίων Πεδίων.

SSRF Model	MAE (ppm)	MARE (%)	RMSE (ppm)	RMSRE (%)	R ²
<i>Πλέγμα (G): 27×30</i>					
$\eta_{1,0} = -1.5$ $\xi_0 = 3$	6.65	23.00	8.69	50.00	0.58
$\eta_1 = 201.6$ $\xi = 0.35$					
$r_s = 0.37$					

Πίνακας 3.11: Σφάλμα εκτίμησης μοντέλου Σπαρτιάτικων Τυχαίων Πεδίων.

SSRF Model	MAE (ppm)	MARE (%)	RMSE (ppm)	RMSRE (%)	R ²
<i>Πλέγμα (G): 27×30</i>					
$\eta_{1,0} = 100$ $\xi_0 = 0.1$	6.59	22.74	8.66	49.40	0.58
$\eta_1 = 100$ $\xi = 0.33$					
$r_s = 0.37$					

Στην συνέχεια αφαιρείται η μέγιστη και η ελάχιστη τιμή συγκέντρωσης χρωμίου. Εφαρμόζεται ξανά η μέθοδος ΣΤΠ, κώδικας «*infer_ssrrf2*» (Παράρτημα Μ), εξετάζοντας διαφορετικά μεγέθη πλέγματος και διαφορετικούς συνδυασμούς αρχικών παραμέτρων. Οι νέοι στατιστικοί περιορισμοί του δείγματος υπολογίστηκαν

$S_0 = 108$, $S_1 = 51$, και $S_2 = 630$, ενώ δύο περιπτώσεις οι οποίες δίνουν το μικρότερο σφάλμα εκτίμησης και την υψηλότερη συσχέτιση παρουσιάζονται στους πίνακες 3.12 και 3.13.

Σε πλέγμα μεγέθους 27×30 η πρώτη περίπτωση περιλαμβάνει τις αρχικές παραμέτρους $\eta_{1,0} = -1.5$, $\xi_0 = 3$ και η δεύτερη $\eta_{1,0} = 1$, $\xi_0 = 1$. Οι παράμετροι του μοντέλου υπολογίζονται $\eta_1 = 539.2$, $\xi = 0.20$ km και $\eta_1 = 559.0$, $\xi = 0.20$ km αντίστοιχα για τις δύο περιπτώσεις. Η αποτελεσματικότερη ακτίνα επίδρασης r_s που χρησιμοποιήθηκε για τον προσδιορισμό των πλησιέστερων γειτόνων κάθε σημείου του δείγματος συγκεντρώσεων χρωμίου είναι $r_s = 0.25$ km. Όπως φαίνεται είναι πολύ κοντά στην υπολογισμένη ακτίνα συσχέτισης.

Στα αποτελέσματα της επιβεβαίωσης (Πίνακας 3.12, 3.13) παρατηρείται ότι η μέγιστη και η ελάχιστη τιμή του δείγματος επηρεάζουν σημαντικά το μέσο σχετικό τετραγωνικό σφάλμα (RMSRE). Συγκριτικά με το αντίστοιχο στους πίνακες 3.10 και 3.11 υπάρχει μια αξιοσημείωτη μείωση, ενώ αυξάνεται και η συσχέτιση (R^2) μεταξύ των εκτιμώμενων συγκεντρώσεων και των πραγματικών τιμών.

Η εκτίμηση με την μέθοδο ΣΤΠ δεν επιτρέπει τον υπολογισμό (εκτίμηση) σε όλα τα σημεία και αφήνει κάποια σημεία χωρίς εκτίμηση. Αυτό συμβαίνει γιατί η ακτίνα συσχέτισης ξ , είναι πολύ μικρή όπως και η αποτελεσματικότερη ακτίνα επίδρασης r_s όποτε και δεν υπάρχουν αρκετοί γείτονες σε αυτήν την απόσταση.

Στην εφαρμογή της μεθόδου έχει τεθεί ως όριο η ύπαρξη τουλάχιστον τριών γειτονικών σημείων για όσο το δυνατόν πιο αξιόπιστη εκτίμηση. Τα σημεία στα οποία δεν πραγματοποιείται εκτίμηση είναι 30 στον αριθμό και βρίσκονται κυρίως στα σύνορα της περιοχής μελέτης. Μεγαλύτερη r_s οδηγεί σε αποτελέσματα με πολύ χαμηλότερη τιμή συσχέτισης (0.38) και μεγαλύτερα σφάλματα. Επομένως προτιμήθηκε η μικρότερη ακτίνα συσχέτισης παρόλο που κάποια σημεία εκτίμησης εξαιρούνται.

Πίνακας 3.12: Σφάλμα εκτίμησης μοντέλου Σπαρτιάτικων Τυχαίων Πεδίων εξαιρουμένων την μέγιστης και της ελάχιστης συγκέντρωσης του δείγματος.

SSRF Model	MAE (ppm)	MARE (%)	RMSE (ppm)	RMSRE (%)	R ²
<i>Πλέγμα (G):</i> 27×30 $\eta_{h,0} = -1.5$ $\xi_0 = 3$ $\eta = 539.2$ $\xi = 0.20$ $r_s = 0.25$	6.33	20.00	8.92	33.00	0.61

Πίνακας 3.13: Σφάλμα εκτίμησης μοντέλου Σπαρτιάτικων Τυχαίων Πεδίων εξαιρουμένων την μέγιστης και της ελάχιστης συγκέντρωσης του δείγματος.

SSRF Model	MAE (ppm)	MARE (%)	RMSE (ppm)	RMSRE (%)	R ²
<i>Πλέγμα (G):</i> 27×30 $\eta_{h,0} = 1$ $\xi_0 = 1$ $\eta_h = 559.0$ $\xi = 0.20$ $r_s = 0.25$	6.30	20.00	8.90	33.00	0.61

3.8 Σύγκριση μεθόδου ΣΤΠ με την μέθοδο Kriging

Το νέο στοιχείο που εισάγει η γεωστατιστική μέθοδος των ΣΤΠ έναντι της κλασσικής γεωστατιστικής μεθόδου *Kriging* είναι ο προσδιορισμός της χωρικής εξάρτησης του δείγματος λαμβάνοντας υπόψη φυσικά μεγέθη που επηρεάζουν την κατανομή των τιμών στο χώρο. Αντίθετα, με την κλασσική γεωστατιστική μέθοδο χρησιμοποιείται, για τον προσδιορισμό της χωρικής εξάρτησης, η προσαρμογή σε γενικά μαθηματικά μοντέλα.

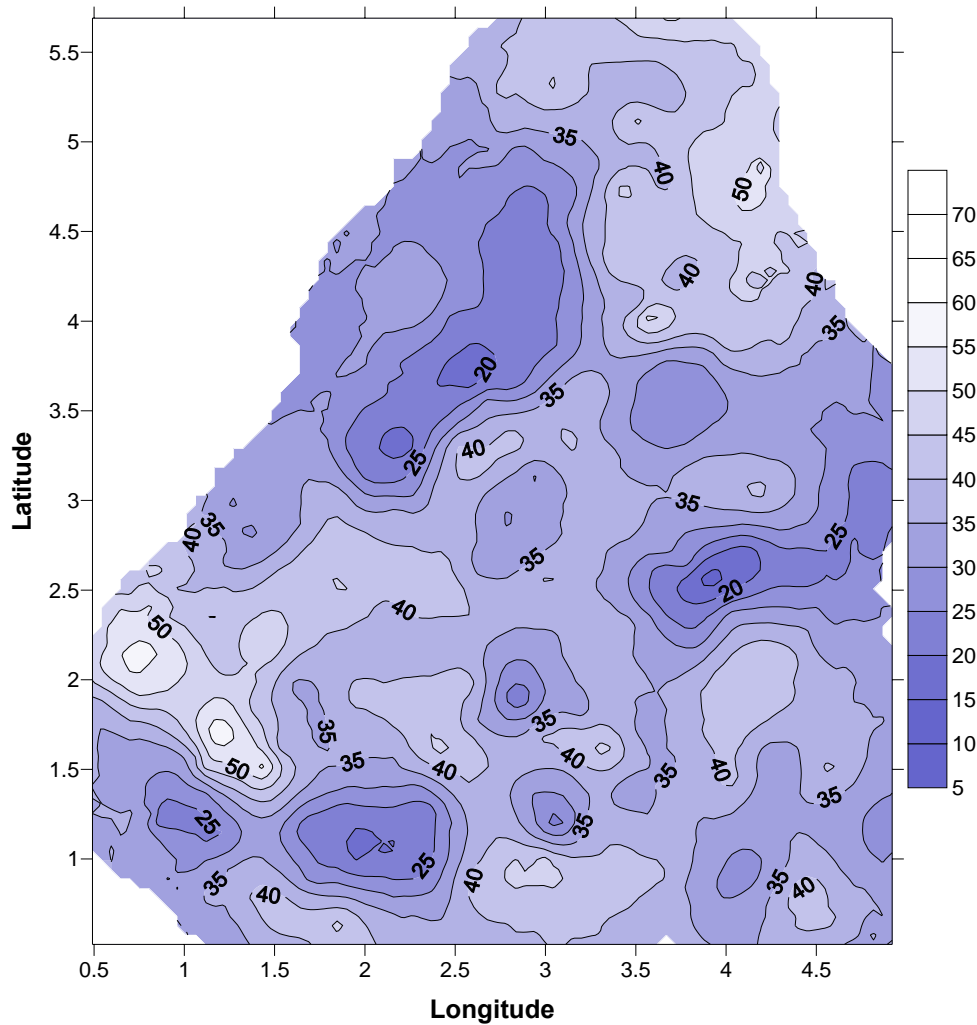
Η χωρική εξάρτηση δείγματος τυχαίου πεδίου χρησιμοποιώντας το γεωστατιστικό μοντέλο ΣΤΠ προσδιορίζεται από ένα μικρό αριθμό παραμέτρων. Για να υπολογιστούν αυτές οι παράμετροι ορίζονται στατιστικοί περιορισμοί που υπολογίζονται από το δείγμα. Αντίθετα στον προσδιορισμό της χωρικής εξάρτησης με την μέθοδο *Kriging* είναι απαραίτητος ο υπολογισμός του ημιβαριογράμματος. Όσον αφορά στη μέθοδο εκτίμησης ο υπολογισμός των χωρικών βαρών στο μοντέλο ΣΤΠ πραγματοποιείται βελτιστοποιώντας το τοπικό

τετραγωνικό σφάλμα, ενώ αντίθετα στη μέθοδος *Kriging* βελτιστοποιείται το μέσο τετραγωνικό σφάλμα.

Συγκρίνοντας τα αποτελέσματα των δύο μεθόδων για την συγκεκριμένη εργασία συμπεραίνεται ότι οι εκτιμήσεις με την μέθοδο ΣΤΠ είναι ανταγωνιστικές των εκτιμήσεων της μεθόδου κανονικού *Kriging* παρόλο που υστερούν ελαφρώς σε ακρίβεια σε σχέση με τις δεύτερες. Για παράδειγμα, το μικρότερο μέσο τετραγωνικό σφάλμα για τη μέθοδο κανονικού *Kriging* είναι 8.24ppm ενώ για το μοντέλο ΣΤΠ είναι 8.66ppm. Ο συντελεστής συσχέτισης, μεταξύ των εκτιμώμενων συγκεντρώσεων και των πραγματικών για το κανονικό *Kriging* είναι 0.64 ενώ ο αντίστοιχος για τα ΣΤΠ είναι 0.58.

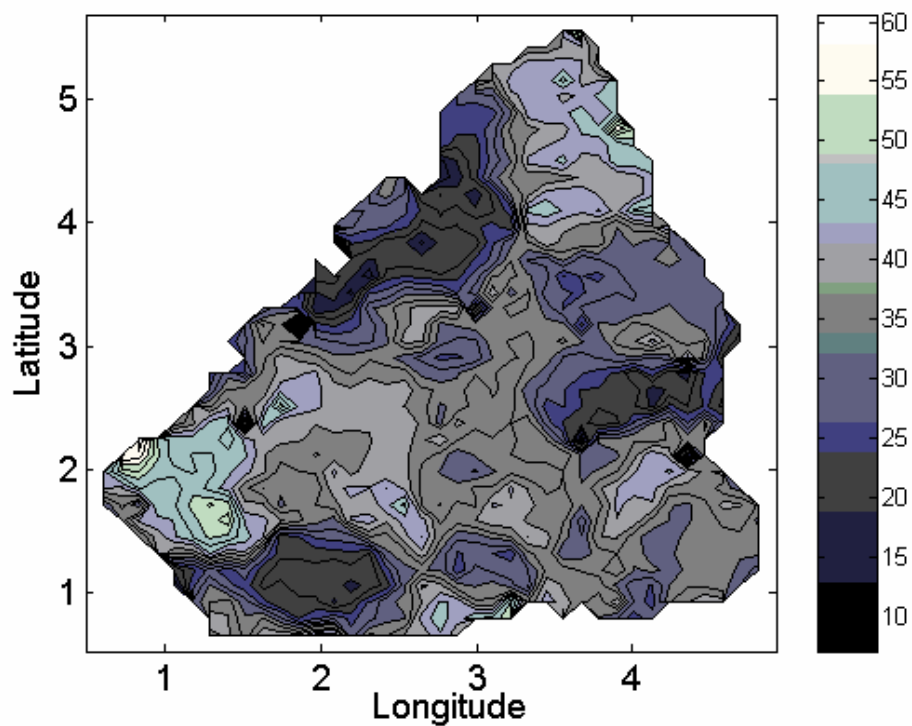
Όταν αφαιρεθούν οι ακραίες τιμές του δείγματος, δηλαδή η μέγιστη και η ελάχιστη τιμή, παρατηρείται μια αξιοσημείωτη μείωση στο μέσο σχετικό τετραγωνικό σφάλμα (RMSRE) το οποίο και με τις δύο μεθόδους μειώνεται σε 33%. Αυτό συμβαίνει γιατί οι ακραίες τιμές του δείγματος παρουσιάζουν το μεγαλύτερο σφάλμα εκτίμησης. Αντιστοίχως αυξάνεται η συσχέτιση μεταξύ των εκτιμώμενων και των πραγματικών συγκεντρώσεων για το μοντέλο ΣΤΠ σε 0.61 πολύ κοντά δηλαδή στη συσχέτιση που δίνει η μέθοδος *Kriging* η οποία είναι 0.64. Σε αυτή την περίπτωση παρατηρείται μια βελτίωση των αποτελεσμάτων του μοντέλου ΣΤΠ. Αντίθετα η συσχέτιση με την μέθοδο κανονικού *Kriging* δεν αλλάζει με την αφαίρεση των ακραίων τιμών.

Το επόμενο βήμα είναι η χαρτογράφηση της περιεκτικότητας χρωμίου στην περιοχή ενδιαφέροντος. Η χαρτογράφηση πραγματοποιείται και με τις δύο μεθόδους. Αρχικά με βάση τα αποτελέσματα της μεθόδου κανονικού *Kriging* στα οποία υπολογίστηκε ότι το μικρότερο σφάλμα εκτίμησης προκύπτει από τις παραμέτρους δύο συνδυαστικών θεωρητικών μοντέλων ημιβαριογράμματος (3.5.1). Χρησιμοποιώντας το λογισμικό πακέτο «Surfer V.8.0.4, 2002», δημιουργείται ένας χάρτης ισοβαρών καμπυλών συγκέντρωσης χρωμίου ο οποίος περιγράφει τις συγκεντρώσεις του σε ολόκληρη την περιοχή ενδιαφέροντος (Σχήμα 3.18).



Σχήμα 3.18: Χάρτης περιεκτικότητας χρωμίου (σε ppm) χρησιμοποιώντας το Σφαιρικό με Φαινόμενο Πυρήνα μοντέλο ημιβαριογράμματος.

Στην συνέχεια χρησιμοποιώντας το μοντέλο ΣΤΠ και λαμβάνοντας υπόψη τις παραμέτρους του οι οποίες δίνουν το μικρότερο σφάλμα εκτίμησης, πραγματοποιείται εκτίμηση της συγκέντρωσης χρωμίου σε όλα τα σημεία χωρίς μετρήσεις μέσα στα όρια της περιοχής ενδιαφέροντος. Οι εκτιμήσεις στα αντίστοιχα σημεία υπολογίζονται με την βοήθεια του κώδικα «*map_ssrf2*» (Παράρτημα Ν) χρησιμοποιώντας πλέγμα διαστάσεων 40x40. Το αποτέλεσμα της μεθόδου ΣΤΠ είναι η δημιουργία ενός χάρτη ισοβαρών καμπυλών συγκέντρωσης χρωμίου (Σχήμα 3.19).



Σχήμα 3.19: Χάρτης περιεκτικότητας χρωμίου (σε ppm) χρησιμοποιώντας το μοντέλο Σπαρτιάτικών Τυχαίων Πεδίων.

Συγκρίνοντας τους δύο χάρτες ισοβαρών καμπυλών παρατηρείται ότι παρουσιάζουν ποιοτικά παρόμοια πληροφορία καθώς στα περισσότερα τμήματα της περιοχής δίνουν σχεδόν τις ίδιες τιμές συγκέντρωσης χρωμίου. Ωστόσο στον χάρτη που παράχθηκε με την μέθοδο κανονικού *Kriging* (Σχήμα 3.18) οι ισοβαρές καμπύλες είναι ομαλές ενώ σε αυτόν που παράχθηκε με την μέθοδο των ΣΤΠ (Σχήμα 3.19) είναι τραχιές.

4. Συμπεράσματα

Δεδομένα περιβαλλοντική ρύπανσης (συγκέντρωσης χρωμίου στο έδαφος) αναλύθηκαν με την βοήθεια της μεθόδου ΣΤΠ και της κλασικής γεωστατιστικής μεθόδου του κανονικού *Kriging*. Χρησιμοποιώντας και τις δυο μεθόδους παρήχθησαν γεωστατιστικοί χάρτες ισοβαρών καμπυλών, οι οποίοι περιγράφουν την διασπορά της συγκέντρωσης χρωμίου στην περιοχή μελέτης. Στα πλαίσια αυτής της εργασίας, η μέθοδος ΣΤΠ εφαρμόστηκε σε πειραματικό στάδιο για πρώτη φορά σε πραγματικά δεδομένα. Η μέθοδος του κανονικού *Kriging* χρησιμοποιήθηκε ως μέτρο σύγκρισης.

Πιο συγκεκριμένα, όσον αφορά την εφαρμογή της μεθόδου ΣΤΠ, κατ' αρχήν υπολογίστηκαν οι παράμετροι του μοντέλου ΣΤΠ (δηλαδή συντελεστής κλίμακας, συντελεστής μορφής και μήκος συσχέτισης) βάσει των δεδομένων. Διερευνήθηκαν δυο μέθοδοι εκτίμησης, εκ των οποίων η μια βασίζεται στον υπολογισμό των στατιστικών περιορισμών με την βοήθεια ενός αυθαίρετου καννάβου υποβάθρου, ενώ η δεύτερη στη χρήση των τριγώνων Delaunay. Διαπιστώθηκε ότι η μέθοδος των τριγώνων Delaunay είναι αποτελεσματική σε συνθετικά δεδομένα κατανομημένα επί καννάβου. Ωστόσο, για την συγκεκριμένη εφαρμογή στην οποία τα σημεία του δείγματος έχουν τυχαία κατανομή στον χώρο, η πρώτη μέθοδος αποδίδει καλύτερα. Συμπεραίνεται ότι η μέθοδος των τριγώνων Delaunay χρήζει περαιτέρω βελτίωσης για την εφαρμογή της σε πραγματικά δεδομένα. Η επιθυμητή βελτίωση εστιάζεται κυρίως στην εξομάλυνση των ασυσχέτιστων διακυμάνσεων (φαινόμενο πυρήνα).

Τα αποτελέσματα της χωρικής εκτίμησης με το γεωστατιστικό μοντέλο ΣΤΠ χαρακτηρίζονται ενθαρρυντικά. Η τετραγωνική ρίζα του μέσου σχετικού τετραγωνικού σφάλματος της εκτίμησης υπολογίστηκε σε 33%, ενώ η συσχέτιση μεταξύ των εκτιμώμενων και των πραγματικών συγκεντρώσεων χρωμίου σε 0.61. Τα αντίστοιχα αποτελέσματα με την μέθοδο του κανονικού *Kriging* είναι 33% και 0.64. Από την μεταξύ τους σύγκριση συνεπάγεται ότι στα πλαίσια της συγκεκριμένης εργασίας, η επίδοση των δύο γεωστατιστικών μεθόδων είναι παραπλήσια.

Μια άλλη διαπίστωση είναι ότι το σφάλμα εκτίμησης και των δύο γεωστατιστικών μεθόδων, όπως προκύπτει από την μέθοδο της διασταυρωμένης επιβεβαίωσης,

επηρεάζεται από τις ακραίες τιμές του δείγματος καθώς αυτές δεν λαμβάνονται επαρκώς υπόψη στην εκτίμηση των παραμέτρων του μοντέλου. Αυτό οφείλεται στο γεγονός ότι και τα δυο γεωστατιστικά μοντέλα βασίζονται στην κανονική κατανομή πιθανότητας, στα πλαίσια της οποίας οι ακραίες τιμές έχουν εξαιρετικά μικρή πιθανότητα υλοποίησης.

Η τιμή του μήκους συσχέτισης του δείγματος διαφέρει για τις δύο γεωστατιστικές μεθόδους. Για την μέθοδο του κανονικού *Kriging* το μήκος συσχέτισης υπολογίζεται ότι κυμαίνεται, ανάλογα με το μοντέλο του ημιβαριογράμματος, από 0.27km - 0.7km. Στην μέθοδο των ΣΤΠ το μήκος συσχέτισης εκτιμάται ότι βρίσκεται στο διάστημα 0.33km - 0.35km αντίστοιχα. Επομένως, προκύπτει ότι και τα δυο γεωστατιστικά μοντέλα δίνουν παρόμοιες τιμές για την εμβέλεια της χωρικής εξάρτησης.

Επίσης διαπιστώθηκε με την βοήθεια της διασταυρωμένης επιβεβαίωσης ότι η αποτελεσματικότερη ακτίνα επίδρασης για την χωρική εκτίμηση με την μέθοδο ΣΤΠ διαφέρει σε μικρό βαθμό από το αντίστοιχο μήκος συσχέτισης. Αντίθετα, στην περίπτωση του κανονικού *Kriging* η αποτελεσματικότερη ακτίνα επίδρασης εξαρτάται από το μοντέλο του ημιβαριογράμματος: στην περίπτωση του σφαιρικού μοντέλου είναι παραπλήσια με το μήκος συσχέτισης, ενώ στην περίπτωση του εκθετικού είναι πολλαπλάσια (λόγω της ουράς του εκθετικού). Αν η αντιστοιχία του μήκους συσχέτισης με την ακτίνα επίδρασης στην μέθοδο ΣΤΠ τεκμηριωθεί και με άλλες μελέτες, αυτό θα διευκολύνει σημαντικά την (υπολογιστική) αναζήτηση της βέλτιστης ακτίνας επίδρασης. Πρέπει επίσης να αναφερθεί ότι οι βέλτιστες παράμετροι του μοντέλου ΣΤΠ δεν παρουσιάζουν ιδιαίτερη ευαισθησία στις αρχικές τιμές των παραμέτρων που χρησιμοποιούνται για την έναρξη του αλγορίθμου βελτιστοποίησης. Αντίθετα, πειραματισμοί με το συγκεκριμένο δείγμα έδειξαν ότι οι παράμετροι είναι ιδιαίτερα ευαίσθητες στις διαστάσεις και την γεωμετρία του πλέγματος (καννάβου) υποβάθρου.

4.1 Προτάσεις για μελλοντική έρευνα

Το γεωστατιστικό μοντέλο ΣΤΠ επιδέχεται διάφορες μεθοδολογικές βελτιώσεις, οι οποίες αναμένεται να αυξήσουν την αποτελεσματικότητα του μοντέλου. Η μελλοντική έρευνα θα επικεντρωθεί στην διερεύνηση της μεθόδου υπολογισμού των παραμέτρων, με κεντρικούς άξονες την εκτενέστερη μελέτη της μεθόδου Delaunay αλλά και την ανάπτυξη νέων μεθόδων εκτίμησης των παραμέτρων (π.χ., εκτίμηση βασισμένη σε συναρτήσεις πυρήνες),

οι οποίες κατά προτίμηση να παρακάμπτουν τον ορισμό ενός πλέγματος υποβάθρου. Αυτό είναι επιθυμητό επειδή ο ορισμός ενός τέτοιου πλέγματος είναι σε μεγάλο βαθμό αυθαίρετος, και διότι η χωρική εκτίμηση είναι ευαίσθητη στην επιλογή του πλέγματος, ώστε να απαιτείται βελτιστοποίηση ως προς το μέγεθος και τις διαστάσεις του.

Ένας άλλος παράγοντας που πρέπει να εισαχθεί στο μοντέλο ΣΤΠ είναι η ανισοτροπία, η οποία δεν λήφθηκε υπόψη κατά την παρούσα εφαρμογή. Πρέπει να σημειωθεί ότι εκτός από την υπόθεση της στατιστικής ισοτροπίας, η οποία προσεγγιστικά ικανοποιείται από το δείγμα που εξετάστηκε, η χωρική εκτίμηση πραγματοποιήθηκε χρησιμοποιώντας ισοτροπικές συναρτήσεις βάσης που εισάγουν πολύ πιο ισχυρούς περιορισμούς. Ωστόσο, η χρήση ισοτροπικών συναρτήσεων βάσης δεν είναι αναγκαίες και πρόκειται να αρθεί σε μελλοντικές προσπάθειες.

Επίσης είναι απαραίτητος ο υπολογισμός της αβεβαιότητας των εκτιμήσεων, ώστε να υπάρχει ολοκληρωμένη εικόνα της εκτίμησης με την βοήθεια των δύο γεωστατιστικών μεθόδων. Εξισώσεις για τον υπολογισμό της αβεβαιότητας των εκτιμήσεων με την μέθοδο ΣΤΠ έχουν αναπτυχθεί πρόσφατα (Hristopulos, 2005b), και επομένως είναι πλέον εφικτή μια ολοκληρωμένη σύγκριση με την μέθοδο *Kriging*.

Τέλος, η μέθοδος των ΣΤΠ θα πρέπει να εφαρμοστεί και σε άλλα δείγματα, έτσι ώστε να διαπιστωθεί η αποτελεσματικότητά της για διαφορετικές κατανομές σημείων δειγματοληψίας και φυσικών μεγεθών, αλλά και να καταγραφούν συγκριτικά πλεονεκτήματα ή αδυναμίες τα οποία θα ληφθούν υπόψη στην εξέλιξη της.

Ελληνική βιβλιογραφία

- 1) Βαρουχάκης, Ε., 2005, *Ανάλυση Χωρικής Κατανομής Συγκέντρωσης Χρωμίου (Cr)*, Μεταπτυχιακή εργασία στα πλαίσια του μαθήματος “Γεωστατιστική & Ρύπανση Περιβάλλοντος”, Τμήμα Μηχανικών Ορυκτών Πόρων, Πολυτεχνείο Κρήτης, Χανιά.
- 2) Χριστόπουλος, Δ., 2003, *Πανεπιστημιακές Σημειώσεις Εισαγωγή Στην Γεωστατιστική*, Τμήμα Μηχανικών Ορυκτών Πόρων, Πολυτεχνείο Κρήτης, Χανιά.
- 3) Χριστόπουλος, Δ., 2004α, *Μεταπτυχιακές Σημειώσεις Ανάλυσης Δεδομένων*, Τμήμα Μηχανικών Ορυκτών Πόρων, Πολυτεχνείο Κρήτης, Χανιά.
- 4) Χριστόπουλος, Δ., 2004β, *Πανεπιστημιακές Σημειώσεις Εισαγωγής Στην Έρευνα Κοιτασμάτων*, Τμήμα Μηχανικών Ορυκτών Πόρων, Πολυτεχνείο Κρήτης, Χανιά.

Διεθνής βιβλιογραφία

- 1) Atkeson, C. G, Moore, A. W. and Schaal, S., 1997, “Locally weighted learning”, *Artificial Intelligence Review*, Vol. 11, pp 11-73.
- 2) Bochner, S., 1959, *Lectures on Fourier Integrals*, Princeton University Press, Princeton, NJ.
- 3) Christakos, G., & Hristopulos, D. T., (1998), *Spatiotemporal Environmental Modelling*, Kluwer Academic Publishers, Boston.
- 4) Goovaerts, P., 1997, *Geostatistics for Natural Resources Evaluation*, Oxford university Press, New York.
- 5) Hohn, M.E., 1999, *Geostatistics & Petroleum Geology*, Kluwer Academic Publishers, The Netherlands.
- 6) Hristopulos, D. T., 2003, “Spartan Gibbs Random Field Models for Geostatistical Applications”, *SIAM Journal on Scientific Computing*, Vol. 24, no. 6, pp. 2125-2162.
- 7) Hristopulos, D. T., 2005a, *Estimation of Spatial Field Derivatives on Irregular Meshes*, Department of Mineral Resources Engineering, Technical University of Crete, Chania
- 8) Hristopulos, D. T., 2005b, “Spartan Spatial Random Field Models Inspired from Statistical Physics with Applications in the Geosciences”, *Next III Conference*, Kolymbari, Chania. <http://xxx.lanl.gov/abs/physics/0510035>

- 9) Issaks, E. & Srivastava R.M., 1989, *An Introduction to Applied Geostatistics*, Oxford University Press, New York.
- 10) Journel, A., 1989, *Fundamentals of Geostatistics in Five Lessons*, American Geophysical Union, Washington, D.C.
- 11) Matlab 2002, *The Language of Technical Computing*, Version 6, The Mathworks, Inc.
- 12) Mulchrone, K. F., 2002, "Application of Delaunay triangulation to the nearest neighbour method of strain analysis", *Journal of Structural Geology*, Vol. 25, pp 689-702.
- 13) Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P., 1992, *Numerical Recipes in Fortran*, Vol. 1, 2nd ed., Cambridge University Press, New York.
- 14) Surfer 2003, *Surface Mapping System*, Version 8.04, Golden Software, Inc.

Δια-δικτυακές πηγές

- 1) European Commission, Project ENV.E.3/ETU/2000/0058 2002, *Heavy Metals in Waste*, viewed December 2005.
<www.environmental-center.com/articles/article1208/heavymetalsreport.pdf>
- 2) European Commission, Council Regulation (EC) 2003, *Protection of the environment, and in particular of the soil, when sewage sludge is used in agriculture*, viewed December 2005.
<europa.eu.int/eur-lex/en/consleg/pdf/1991/en_1991L0308_do_001.pdf>
- 3) Forest Management Institute 2005, *Report on Chemical Analysis of Samples*, viewed December 2005.
<tofino.ex.ac.uk/euroturtle/medas/pdf/kazanlipollution.pdf>
- 4) Mathematica Applications Library 2005, *Document on Variational Methods*, Wolfram Research Documentation Center, viewed November 2005.
<<http://documents.wolfram.com/v5/AddonsLinks/StandardPackages/Calculus/VariationalMethods.html>>
- 5) Myers, D., 2005, *Spatial Statistics*, Department of Mathematics, University of Arizona, viewed September 2005.
< <http://www.u.arizona.edu/~donaldm/homepage/whatis.html>>

Παραρτήματα

Παράρτημα Α «Προσδιορισμός αναλυτικών λύσεων στατιστικών περιορισμών για περιοδική συνάρτηση δείγματος»

Οι στατιστικοί περιορισμοί του δείγματος \bar{S}_1, \bar{S}_2 μπορούν να υπολογιστούν αναλυτικά υπολογίζοντας την πρώτη και την δεύτερη παράγωγο της περιοδικής συνάρτησης από την οποία προέρχεται το δείγμα. Αυτή έχει την μορφή

$$f = \cos(k_x x + k_y y) \quad \text{A.1}$$

όπου, k_x, k_y είναι το κυματάνυσμα το οποίο είναι ίσο με $k_x = \frac{2\pi}{J_x}$ και $k_y = \frac{2\pi}{J_y}$, ενώ

$J_x = \frac{L}{10}$ και $J_y = \frac{L}{5}$. Η μεταβλητή L δηλώνει το μήκος του πλέγματος στις δύο διαστάσεις (100×100).

Αρχικά υπολογίζεται η μέση τιμή του τετραγώνου της πρώτης παράγωγου της περιοδικής συνάρτησης από την οποία προκύπτει ο στατιστικός περιορισμός \bar{S}_1 ο οποίος αντιστοιχεί στην μέση τιμή του τετραγώνου της κλίσης.

$$[f] = \frac{1}{L^2} \int dx \int dy (\nabla f(x, y))^2 \quad \text{A.2}$$

όπου $\nabla f(x, y) = (d_x f, d_y f)$, αντιστοιχεί στην βαθμίδα (κλίση). Για το τετράγωνο της μέσης τιμής της πρώτης παραγώγου ισχύει,

$$E[(\nabla f(x, y))^2] = E[(d_x f)^2 + (d_y f)^2] = [E(d_x f)^2 + E(d_y f)^2] \quad \text{A.3}$$

Η πρώτη παράγωγος της περιοδικής συνάρτησης κατά $(d_x f)$ είναι,

$$(d_x f) = -k_x \sin(k_x x + k_y y) \quad \text{A.4}$$

ενώ κατά $(d_y f)$ είναι,

$$(d_y f) = -k_y \sin(k_x x + k_y y) \quad \text{A.5}$$

Αντικαθιστώντας τις εξισώσεις (Α.4) και (Α.5) στην εξίσωση (Α.3) υπολογίζεται η πρώτη παράγωγος της περιοδικής συνάρτησης,

$$(\nabla f(x, y))^2 = (d_x f)^2 + (d_y f)^2 = (k_x^2 + k_y^2) \sin^2(k_x x + k_y y) \quad \text{A.6}$$

Στην συνέχεια αντικαθιστώντας την παραπάνω εξίσωση στην (Α.2) μπορεί να υπολογιστεί ο στατιστικός περιορισμός \bar{S}_1 επιλύοντας το διπλό ολοκλήρωμα.

$$[f] = \frac{1}{L^2} \int_0^L dx \int_0^L dy (k_x^2 + k_y^2) \sin^2(k_x x + k_y y)$$

Η λύση περιγράφεται από τις παρακάτω εξισώσεις.

$$\begin{aligned} [f] &= \frac{1}{L^2} (k_x^2 + k_y^2) \int_0^L dy \left[\frac{2k_x x + 2k_y y - \sin 2(k_x x + k_y y)}{4k_x} \right]_0^L \\ &= \frac{1}{L^2} (k_x^2 + k_y^2) \int_0^L dy \left[\frac{2k_x L + 2k_y y - \sin 2(k_x L + k_y y)}{4k_x} \right] - \left[\frac{2k_y y - \sin 2k_y y}{4k_x} \right] \\ &= \frac{1}{L^2} (k_x^2 + k_y^2) \frac{1}{4k_x} \left[\int_0^L (2k_x L + 2k_y y) dy - 2 \int_0^L \sin(k_x L + k_y y) dy \right. \\ &\quad \left. - 2 \int_0^L k_y y dy + 2 \int_0^L \sin k_y y dy \right] \\ &= \frac{1}{L^2} (k_x^2 + k_y^2) \frac{1}{4k_x} \left[\left[y(2k_x L + k_y y) \right]_0^L + \left[\frac{2 \cos(k_x L + k_y y)}{k_y} \right]_0^L \right. \\ &\quad \left. - \left[k_y y^2 \right]_0^L - \left[\frac{2 \cos k_y y}{k_y} \right]_0^L \right] \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{L^2} (k_x^2 + k_y^2) \frac{1}{4k_x} \left[L(2k_x L + k_y L) + \frac{2 \cos(k_x L + k_y L)}{k_y} - \frac{2 \cos(k_x L + 0)}{k_y} \right. \\
&\quad \left. - k_y L^2 - \frac{2 \cos k_y L}{k_y} + \frac{2 \cos 0}{k_y} \right] \\
&= \frac{1}{L^2} (k_x^2 + k_y^2) \frac{1}{4k_x} \left[2k_x L^2 + k_y L^2 + \frac{2 \cos k_x L}{k_y} + \frac{2 \cos k_y L}{k_y} - \frac{2 \cos k_x L}{k_y} - \frac{2}{k_y} \right. \\
&\quad \left. - k_y L^2 - \frac{2 \cos k_y L}{k_y} + \frac{2}{k_y} \right] \\
&= \frac{1}{L^2} (k_x^2 + k_y^2) \frac{1}{4k_x} \left[2k_x L^2 + k_y L^2 + \frac{2 \cos k_x L}{k_y} + 2 \cos L - \frac{2 \cos k_x L}{k_y} - k_y L^2 - 2 \cos L \right] \\
&= \frac{1}{L^2} (k_x^2 + k_y^2) \frac{1}{4k_x} 2k_x L^2 = 2 \left(\frac{k_x^2 + k_y^2}{4} \right).
\end{aligned}$$

Για πλέγμα 100×100 ισχύει $k_x = 0.6283$ και $k_y = 0.3142$, οπότε η μέση τιμή του τετραγώνου της κλίσης \bar{S}_1 υπολογίζεται,

$$[f] = 0.24$$

Ο στατιστικός περιορισμός \bar{S}_2 ο οποίος αντιστοιχεί στην μέση τιμή του τετραγώνου της καμπυλότητας δίνεται από την μέση τιμή του τετραγώνου της δεύτερης παραγώγου της περιοδικής συνάρτησης (Α.7),

$$[f] = \frac{1}{L^2} \int dx \int dy \left(\nabla^2 f(x, y) \right)^2 \quad \text{A.7}$$

όπου $\nabla^2 f(x, y)$, αντιστοιχεί στην Λαπλασιανή της περιοδικής συνάρτησης. Για το τετράγωνο της μέσης τιμής της δεύτερης παραγώγου ισχύει,

$$\begin{aligned} E \left[\left(\nabla^2 f(x, y) \right)^2 \right] &= E \left[\left(d_{xx}f \right) + \left(d_{yy}f \right) \right]^2 \\ &= E \left[\left(d_{xx}f \right)^2 \right] + E \left[\left(d_{yy}f \right)^2 \right] + E \left[2 \left(d_{xx}f \right) \left(d_{yy}f \right) \right] \end{aligned} \quad \text{A.8}$$

Η δεύτερη παράγωγος της περιοδικής συνάρτησης κατά $(d_{xx}f)$ είναι,

$$(d_{xx}f) = -k_x^2 \cos(k_x x + k_y y) \quad \text{A.9}$$

ενώ κατά $(d_{yy}f)$ είναι,

$$(d_{yy}f) = -k_y^2 \cos(k_x x + k_y y) \quad \text{A.10}$$

Αντικαθιστώντας τις εξισώσεις (Α.9) και (Α.10) στην εξίσωση (Α.8) υπολογίζεται η δεύτερη παράγωγος της περιοδικής συνάρτησης,

$$\begin{aligned} \left(\nabla^2 f(x, y) \right)^2 &= (d_{xx}f)^2 + (d_{yy}f)^2 + 2(d_{xx}f)(d_{yy}f) \\ &= \left(-k_x^2 \cos(k_x x + k_y y) \right)^2 + \left(-k_y^2 \cos(k_x x + k_y y) \right)^2 \\ &\quad + 2 \left(-k_x^2 \cos(k_x x + k_y y) \right) \left(-k_y^2 \cos(k_x x + k_y y) \right) \end{aligned} \quad \text{A.11}$$

Η εξίσωση (Α.11) μπορεί να γραφτεί και ως εξής,

$$\left(\nabla^2 f(x, y)\right)^2 = \left(k_x^2 + k_y^2\right)^2 \cos^2(k_x x + k_y y)$$

Στην συνέχεια αντικαθιστώντας την παραπάνω εξίσωση στην (Α.7) υπολογίζεται ο στατιστικός περιορισμός \bar{S}_2 .

$$[f] = \frac{1}{L^2} \int_0^L dx \int_0^L dy \left(k_x^2 + k_y^2\right)^2 \cos^2(k_x x + k_y y)$$

Η λύση περιγράφεται από τις παρακάτω εξισώσεις.

$$\begin{aligned} [f] &= \frac{1}{L^2} \left(k_x^2 + k_y^2\right)^2 \int_0^L dy \left[\frac{2k_x x + 2k_y y + \sin 2(k_x x + k_y y)}{4k_x} \right]_0^L \\ &= \frac{1}{L^2} \left(k_x^2 + k_y^2\right)^2 \int_0^L dy \left[\frac{2k_x L + 2k_y y + \sin 2(k_x L + k_y y)}{4k_x} \right] - \left[\frac{2k_y y + \sin 2k_y y}{4k_x} \right] \\ &= \frac{1}{L^2} \left(k_x^2 + k_y^2\right)^2 \frac{1}{4k_x} \left[\int_0^L (2k_x L + 2k_y y) dy + 2 \int_0^L \sin(k_x L + k_y y) dy \right. \\ &\quad \left. - 2 \int_0^L k_y y dy - 2 \int_0^L \sin k_y y dy \right] \\ &= \frac{1}{L^2} \left(k_x^2 + k_y^2\right)^2 \frac{1}{4k_x} \left[\left[y(2k_x L + k_y y) \right]_0^L - \left[\frac{2 \cos(k_x L + k_y y)}{k_y} \right]_0^L \right. \\ &\quad \left. - \left[k_y y^2 \right]_0^L + \left[\frac{2 \cos k_y y}{k_y} \right]_0^L \right] \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{L^2} (k_x^2 + k_y^2)^2 \frac{1}{4k_x} \left[L(2k_x L + k_y L) - \frac{2 \cos(k_x L + k_y L)}{k_y} + \frac{2 \cos(k_x L + 0)}{k_y} \right. \\
&\quad \left. - k_y L^2 + \frac{2 \cos k_y L}{k_y} - \frac{2 \cos 0}{k_y} \right] \\
&= \frac{1}{L^2} (k_x^2 + k_y^2)^2 \frac{1}{4k_x} \frac{1}{4k_x} \left[2k_x L^2 + k_y L^2 - \frac{2 \cos k_x L}{k_y} - \frac{2 \cos k_y L}{k_y} + \frac{2 \cos k_x L}{k_y} + \frac{2}{k_y} \right. \\
&\quad \left. - k_y L^2 + \frac{2 \cos k_y L}{k_y} - \frac{2}{k_y} \right] \\
&= \frac{1}{L^2} (k_x^2 + k_y^2)^2 \frac{1}{4k_x} \left[2k_x L^2 + k_y L^2 + \frac{2 \cos k_x L}{k_y} + 2 \cos L - \frac{2 \cos k_x L}{k_y} - k_y L^2 - 2 \cos L \right] \\
&= \frac{1}{L^2} (k_x^2 + k_y^2)^2 \frac{1}{4k_x} 2k_x L^2 = \left(\frac{(k_x^2 + k_y^2)^2}{2} \right).
\end{aligned}$$

Χρησιμοποιώντας τις τιμές που προαναφέρθηκαν για το κυματόνισμα k_x και k_y , η τιμή του τετραγώνου της καμπυλότητας \bar{S}_2 προκύπτει ίση με,

$$[f] = 0.12$$

Παράρτημα Β «Υπολογισμός ισοτροπικής λύσης»

Εξετάζονται οι δύο περιπτώσεις ξεχωριστά:

- Για $\eta_1 \geq 2$ οι ρίζες είναι φανταστικοί αριθμοί και δίνονται από την εξίσωση,

$$k_j = \mp \frac{i}{\sqrt{2}\xi} \sqrt{\eta_1 \mp \sqrt{\eta_1^2 - 4}}$$

Χρησιμοποιώντας την εξίσωση (2.39) και την ισότητα $I_0(z) = I_0(-z)$ λαμβάνεται η παρακάτω σχέση:

$$\hat{\chi}(s) = 4\pi \left[AI_0 \left(\frac{s}{\sqrt{2}\xi} \sqrt{\eta_1 + \sqrt{\eta_1^2 - 4}} \right) + BI_0 \left(\frac{s}{\sqrt{2}\xi} \sqrt{\eta_1 - \sqrt{\eta_1^2 - 4}} \right) \right]$$

όπου οι συντελεστές A και B είναι πραγματικοί αριθμοί.

- Για $|\eta_1| < 2$ οι ρίζες μπορούν να εκφραστούν ως $k_j = \pm p_1 \pm iu_1$, όπου $p_1 = (\sqrt{2}\xi)^{-1} \sqrt{1 - \eta_1/2}$, και $u_1 = (\sqrt{2}\xi)^{-1} \sqrt{1 + \eta_1/2}$

Η μετατροπή της εξίσωσης $I_0(ik_js) = I_0(\pm i p_1 s \mp u_1 s)$ είναι δυνατή χρησιμοποιώντας,

(α) την ισότητα $I_0(z) = J_0(iz)$ η οποία δίνει $I_0(\pm i p_1 s \pm u_1 s) = J_0(\pm p_1 s \mp iu_1 s)$ και

(β) την ισότητα $J_0(z) = J_0(-z)$ η οποία χρησιμοποιώντας την εξίσωση (2.39) δίνει την παρακάτω εξίσωση,

$$\hat{\chi}(s) = 2\pi [A_1 J_0(p_1 s + iu_1 s) + A_2 J_0(p_1 s - iu_1 s)]$$

Από τον ορισμό της συνάρτησης Bessel έπεται ότι $J_0(z^*) = J_0^*(z)$, έτσι προκύπτει

$$\begin{aligned}
 \hat{\chi}(\mathbf{s}) &= 2\pi \left\{ \frac{(A_1 + A_2)}{2} [J_0(p_1 s + i u_1 s) + J_0(p_1 s - i u_1 s)] \right. \\
 &\quad \left. + \frac{(A_1 - A_2)}{2} [J_0(p_1 s + i u_1 s) - J_0(p_1 s - i u_1 s)] \right\} \\
 &= 2\pi \left\{ (A_1 + A_2) \operatorname{Re}[J_0(p_1 s + i u_1 s)] + i(A_1 - A_2) \operatorname{Im}[J_0(p_1 s + i u_1 s)] \right\}
 \end{aligned}$$

Για να είναι πραγματική η λύση, $A_2 = A_1^*$. Αν $A_1 = A - iB$ όπου A και B είναι πραγματικοί αριθμοί εξάγεται:

$$\hat{\chi}(\mathbf{s}) = 4\pi \left\{ A \operatorname{Re}[J_0(p_1 s + i u_1 s)] + B \operatorname{Im}[J_0(p_1 s + i u_1 s)] \right\}$$

Προγράμματα MATLAB

Παράρτημα Γ «Υπολογισμός στατιστικών περιορισμών με την μέθοδο Delaunay»

% DELAUNAY_APPLICATION calculates the sample constraints by performing

% the delaunay method. The program uses the data points of the training set.

% INPUT

% x: X coordinates of data points

% y: Y coordinates of data points

% rf: Random sample

% rangemin: Minimum percentile distance cut-off for derivative calculation

% rangemax: Maximum percentile distance cut-off for derivative calculation

% OUTPUT

% s: [1 x 3] vector of sample constraints

% DELAUNAY_APPLICATION calls the functions:

% DELAUNAY, PRCTILE (MATLAB)

% NOTES

% The DELAUNAY function determines the nearest neighbours of each point in

% the training set. Then the sample constraints are calculated using the Finite

% Difference Method for the determination of the local derivatives.

% EXAMPLE

% [s]=delaunay_application(x,y,rf)

Παράρτημα Δ «Δημιουργία τυχαίου πεδίου δύο διαστάσεων»

% RF2DIRREG Normal, correlated 2D random fields
 % [RF]=RF2DIRREG(X,Y,FM,SF,B,AR,COR,NMOD)
 % returns the 2D random field RF at locations defined by the vectors X,Y.
 % The field is generated using the mode superposition method. This method does
 % not use the covariance matrix and is thus more appropriate for larger
 % systems than LU decomposition.

% INPUT VARIABLES:

% X: X coordinate vector (vector)
 % Y: Y coordinate vector (vector)
 % FM: Mean of RF
 % SF: Standard deviation of RF
 % B: Shortest correlation length (assumed along y)
 % AR: Anisotropic ratio $AR = b_x / b_y$ (real number > 1)
 % COR: Correlation model (string)
 % NMOD: Total number of modes

% At this point three models are supported:

% 'Gaus1: Gaussian: $\exp(-x_l^2 - y_l^2)$
 % 'Expo': Exponential: $\exp[-\sqrt{x_l^2 + y_l^2}]$
 % 'Sinc': Hole: $\sin(x_l) \sin(y_l) / (x_l y_l)$
 % where the reduced distances are: $x_l = x/b_x$ and $y_l = y/b_y$

% OUTPUT VARIABLES:

% RF: Random field values on the grid (sample)

Παράρτημα Ε «Υπολογισμός στατιστικών περιορισμών δείγματος (σύνολο εκπαίδευσης)»

```
% IRREG_CONSTRAINT3 calculates the sample constraints S0 S2 S4 of
% an irregularly spaced random field using a square background lattice
% of length L per side.
%
% [S,NC,GCELLS]=IRREG_CONSTRAINT3(X,Y,RF,L);
%
%
% INPUT
% x: X coordinates of data points
% y: Y coordinates of data points
% rf: Random sample (Chromium (Cr) in training set
% L: size of background lattice
%
% OUTPUT
% s: [1 x 3 ] vector of sample constraints
% Nc: [1 x 3 ] vector of number of cells used in the constraints calculation
% Gcells: Cell array containing 1D indices of data points inside each cell
%
% EXAMPLE
% [s,Nc,Gcells]=irreg_constraint3(x,y,rf,L);
```

Παράρτημα Z «Προσδιορισμός συναρτησιακής απόστασης»

```
% DISTFUNC2 Spartan model's distance functional for 2 parameters
% (given values of the Spartan Gibbs RF coefficients)
% and the appropriate sample constraints.
%
% [disf]=distfunc2(beta,d,samplec,kmax,Sd,normd)
%
%          INPUT
%   beta:      2-Coefficient vector [b \xi]
%   d:         Number of dimensions (d)
%   samplec:   Sample constraint vector [S1 S2]
%   kmax:      Wavevector cutoff (kmax)
%
%          The lowest permissible values are for b = -2 and for xi = 0
%
%          OPTIONAL PARAMETERS
%   Sd:        Surface area of unit sphere
%   normd:     Normalization constant (1/(2pi)^d)
%
%          OUTPUT
%   disf:      Value of the model distance functional
%
%          EXAMPLE
%   [disf]=distfunc2([2 1],d,[s1 s2],2*pi);
```


Παράρτημα Η «Υπολογισμός παραμέτρων δείγματος (σύνολο εκπαίδευσης)»

```
% DIST_MINIMIRREG statistics of Spartan model from given off-lattice sample.
% The function returns the values of the model coefficients, the COMPUTER TIME
% required for the minimization, and the number of iterations of the simplex search.
%
% The method used here is based on the minimization of the sum of squares of the
% difference between the experimental (sample) and theoretical (stochastic)
% moments.
%
% For the minimization is used the Matlab script FMINCON
% which employs the simplex search algorithm.
%
% The 2-coefficient parameterization b, \xi is used for the model coefficients.
% The third coefficient that is related to the variance has been eliminated.
%
% DIST_MINIMIRREG calls the function: DISTFUNC2
%
% INPUT
% s:      Sample constraints [variance, gradient, Laplacian]
% d:      Dimensionality
% kmax:    Wavevector cutoff
% par0:    Initial guess for [b(0) \xi(0)]
%
% OUTPUT
% par:     Solution of the constraint equation system [a, b, \xi]
%
% flag:    Type of convergence:
%          Flag > 0 convergence
%          Flag = 0 maximum no. of iterations reached
%          Flag < 0 failure to converge
% fval:    Value of the distance functional at minimum
```

```
%    time:    Total computational time
%    cput:    CPU Time used
%    output:  Structure containing the number of iterations taken
%             in OUTPUT iterations
%
%    EXAMPLE
%    [par,flag,fval,time,cput,output]=dist_minimirreg(s,d,kmax,par0)
```

Παράρτημα Θ «Εκτίμηση στα σημεία του συνόλου επιβεβαίωσης»

```
% PREDICTION_SSRF infers the SSRF model parameters from the data.
% The program uses the data points of the training set to infer SSRF parameters.
% The method is based on matching the sample and stochastic constraints.
%
% First, the sample constraints are calculated using the method of the
% background grid IRRREG_CONSTRAINT3.
%
% Then, the optimal parameters q2 (Shape coefficient) and q3 (Correlation length)
% are estimated by matching the sample and ensemble moments, using the function
% DIST_MINIMIRREG.
%
% The optimization is performed with the MATLAB constrained optimization
% function FMINCON.
%
% Next, an optimal search radius is determined (dist_nn) from the function
% LESS_THAN in terms of the correlation radius, and is used to determine
% the nearest neighbours of each point of the validation set.
%
% Then the first and the second basis function is calculated in the
% neighbourhood of each estimation (validation) point from the function BASIS1
%
% The coefficients of the estimate are determined using the function REGRESS.

% The SSRF error is defined as the difference between
% the actual value and the prediction. The procedure is repeated for all
% points and the root mean square (RMS) error is calculated.
% The procedure is repeated for various search neighbourhoods.
% Finally, the optimal (in the sense of minimum mean square error) case is selected.
%
% PREDICTION_SSRF calls the functions: FMINCON, REGRESS (MATLAB)
```

```
% PREDICTION_SSRF calls the functions: IRREG_CONSTRAINT3,
% DIST_MINIMIRREG, BASIS1, LESS_THAN,(USER DEFINED)
%
% INPUT PARAMETERS
%     x:   X coordinates of data points in training set
%     y:   Y coordinates of data points in training set
%  xlocva: X coordinates of data points in validation set
%  ylocva: Y coordinates of data points in validation set
%     cr:   Random sample in training set (Chromium, Cr)
%  crva:   Random sample in validation set (Chromium, Cr)

% OUTPUT PARAMETERS
%     xs:   SSRF predictions
%     mae:   Mean Absolute Error
%     rmse:  Root Mean Square Error
%     mare:  Mean Absolute Relative Error
%     rmsre: Root Mean Square Relative Error
%     R:     Linear Correlation Coefficient

% EXAMPLE
% [xs]=prediction_ssr(x,y,xlocva,ylocva,cr)

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% LESS_THAN. An optimal search radius is determined (dist_nn) in terms of the
% correlation radius, and is used to determine the nearest neighbours of each point of
% the validation set.
%
% INPUT
%     s1: Distances among the estimation (validation) point and the
%         sample points in the training set
%     q3: Correlation length
%     cr: Random sample in training set (Chromium, Cr)
```

```
% OUTPUT
% dist_nn: optimal search radius
%      s2: Set of distances from estimation point to nearest neighbours

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% BASIS1 calculates the values of the SSRF basis functions in 2D.
%
% INPUT
%  z:   Set of distances from estimation point to neighbours (vector)
%  q3:   SSRF Correlation length
%  q2:   SSRF Shape coefficient
%
% OUTPUT
%  y11:  First basis function (vector)
%  y22:  Second basis function (vector)

% EXAMPLE
%      [y11,y22]=basis1(q2,q3,z)
```

Παράρτημα Ι «Εκτίμηση στα σημεία του συνόλου επιβεβαίωσης μετά την αφαίρεση των ακραίων τιμών»

```
% PREDICTION_SSRF_98 infers the SSRF model parameters from the data.
% The program uses the data points of the training set to infer SSRF parameters.
% The method is based on matching the sample and stochastic constraints.
%
% First, the sample constraints are calculated using the method of the
% background grid IRREG_CONSTRAINT3.
%
% Then, the optimal parameters q2 (Shape coefficient) and q3 (Correlation length)
% are estimated by matching the sample and ensemble moments, using the function
% DIST_MINIMIRREG.
%
% The optimization is performed with the MATLAB constrained optimization
% function FMINCON.
%
% Next, an optimal search radius is determined (dist_nn) from the function
% LESS_THAN in terms of the correlation radius, and is used to determine
% the nearest neighbours of each point of the validation set.
%
% The validation set this time does not include the highest and the lowest
% concentration values and the correlated coordinates.
%
% Then the first and the second basis function is calculated in the
% neighbourhood of each estimation (validation) point from the function BASIS1
%
% The coefficients of the estimate are determined using the function REGRESS.
%
% The SSRF error is defined as the difference between
% the actual value and the prediction. The procedure is repeated for all
% points and the root mean square (RMS) error is calculated.
% The procedure is repeated for various search neighbourhoods.
```

```
% Finally, the optimal (in the sense of minimum mean square error) case is selected.

% PREDICTION_SSRF_98 calls the functions: FMINCON, REGRESS (MATLAB)
% PREDICTION_SSRF_98 calls the functions: IRREG_CONSTRAINT3,
% DIST_MINIMIRREG, BASIS1, LESS_THAN, (USER DEFINED)
%
% INPUT PARAMETERS
%      x:   X coordinates of data points in training set
%      y:   Y coordinates of data points in training set
% xlocva98: X coordinates of data points in validation set
% ylocva98: Y coordinates of data points in validation set
%      cr:   Random sample in training set (Chromium, Cr)
% crva98:   Random sample in validation set (Chromium, Cr)
%
% OUTPUT PARAMETERS
%      xs:   SSRF predictions
%      mae:   Mean Absolute Error
%      rmse:  Root Mean Square Error
%      mare:  Mean Absolute Relative Error
%      rmsre: Root Mean Square Relative Error
%      R:     Linear Correlation Coefficient

% EXAMPLE
% [xs]=prediction_ssr(x,y,xlocva,ylocva,cr)
```

Παράρτημα Κ «Υπολογισμός στατιστικών περιορισμών δείγματος»

```
% BCKG_CONSTR calculates the sample constraints S0 S2 S4 of
% an irregularly spaced random field using a square background lattice
% of length L per side.
%
%
% INPUT
% x: X coordinates of data points
% y: Y coordinates of data points
% rf: Random sample (Chromium (Cr))
% L: size of background lattice
% Ng: Integer determining the grid size
%
% OUTPUT
% s: [1 x 3 ] vector of sample constraints
% Nc: [1 x 3 ] vector of number of cells used in the constraints calculation
% Gcells: Cell array containing 1D indices of data points inside each cell
% x1: x coordinates of background grid
% y1: y coordinates of background grid
% dx: Grid step along x-axis
% dy: Grid step along y-axis
%
% NOTES
% Orthogonal background grid is used. The size of the grid in each direction
% is proportional to the extent of the modelled area in the same direction.
% Ng+2 is the number of nodes along the shortest direction.
% The step size is assumed to be the same in both directions;
% the number of nodes in the longer direction then follows automatically.
%
% EXAMPLE
% [s,Nc,Gcells,x1,y1,dx,dy]=bckg_constr(x,y,rf,L);
```


Παράρτημα Α «Υπολογισμός παραμέτρων δείγματος»

```
% CONSTR_MATCH calculates SRF parameters by matching sample & model
% constraints. The function returns the values of the model coefficients, the
% COMPUTER TIME required for the minimization, and the number of iterations of
% the simplex search. The method used here is based on the minimization of the sum
% of squares of the difference between the experimental (sample) and theoretical
% (stochastic) moments. For the minimization we use the Matlab constrained
% optimization method FMINCON.
%
% The 2-coefficient parameterization b, \xi is used for the model coefficients
% The third coefficient that is related to the variance has been eliminated.
%
% CONSTR_MATCH calls the function: DISTFUNC2
%
%
% INPUT
% s: Sample constraints [variance, gradient, Laplacian]
% d: Dimensionality
% kmax: Wavevector cutoff
% par0: Initial guess for [b(0) \xi(0)]
%
% OUTPUT
% par: Solution of the constraint equation system [a, b, \xi]
% flag: Type of convergence:
%       Flag > 0 convergence
%       Flag = 0 maximum no. of iterations reached
%       Flag < 0 failure to converge
% fval: Value of the distance functional at minimum
% time: Total computational time
% cput: CPU Time used
% output: Structure containing the number of iterations taken in
%
% EXAMPLE
% [par,flag,fval,time,cput,output]=constr_match(s,d,kmax,par0)
```

Παράρτημα Μ «Εκτίμηση στα σημεία του δείγματος με την μέθοδο της διασταυρωμένης επιβεβαίωσης»

```
% INFER_SSRF2 infers the SSRF model parameters from the data.
% The program uses all the data points to infer SSRF parameters.
% The method is based on matching the sample and stochastic constraints.
%
% First, the sample constraints are calculated using the method of the
% background grid (function BCKG_CONSTR).
%
% Then, the optimal parameters q2 (Shape coefficient) and q3 (Correlation length)
% are estimated by matching, the sample and ensemble moments, using the function
% CONSTR_MATCH.
%
% The optimization is performed with the MATLAB constrained optimization
% function FMINCON.
%
% SSRF prediction uses the function BASIS.
%
% Next, an optimal search radius is determined using a cross-validation procedure:
% One data point is removed and the value at this point is estimated using
% the SSRF predictor. The SSRF error is defined as the difference between
% the actual value and the prediction. The procedure is repeated for all
% points and the root mean square (RMS) error is calculated.
% The procedure is repeated for various search neighbourhoods.
% Finally, the optimal (in the sense of minimum mean square error) case is selected.
%
% INFER_SSRF2 calls the functions: FMINCON, REGRESS (MATLAB)
% INFER_SSRF2 calls the functions: BASIS, BCKG_CONSTR,
% CONSTR_MATCH (USER DEFINED)
```

```

%      INPUT PARAMETERS
%      d:   Number of dimensions
%      p1:   Percentile of the point-to-point distance distribution for lower bound on
%            correlation length
%      pinit: Percentile of the point-to-point distance distribution for initial
%            correlation length
%      p2:   Percentile of the point-to-point distance distribution for upper bound on
%            correlation length
%      eta0: Initial guess for the shape coefficient
%      xi0:  Initial guess for the correlation length
%      Ng:   Integer determining size of background grid
%      nbmin: Minimum number of neighbour data for a point in order to calculate
%            SSRF estimate
%
%      OUTPUT PARAMETERS
%      mae:  Mean Absolute Error
%      rmse: Root Mean Square Error
%      mare: Mean Absolute Relative Error
%      rmsre: Root Mean Square Relative Error
%      R:    Linear Correlation Coefficient
%      Cs:   Rank Correlation Coefficient
%      zssrf: SSRF predictions
%      zdat: Field values at data points
%
%  NOTES
%  The optimization algorithm gets stuck in local minima. Hence, in the
%  future a method like simulated annealing may be preferable.
%  To avoid over-smoothing smaller values of the correlation length are
%  preferable.
%
%  The SSRF predictor can not always determine reliable estimates.
%  Any resulting NAN values are removed. The errors are calculated
%  based on the points where there are numeric predictions.
%  This deserves further investigation.

```

%%%

```
% BASIS calculates the values of the SSRF basis functions in 2D.
%
%   INPUT
%   s:   Set of distances from estimation point to neighbours (vector)
%   xi:   SSRF Correlation length (number)
%   eta:   SSRF Shape coefficient (eta_{1})
%
%   OUTPUT
%   f1:   First basis function (vector)
%   f2:   Second basis function (vector)
%
%   EXAMPLE
%   [f1,f2]=basis(eta,xi,s)
```

Παράρτημα N «Εκτίμηση στα σημεία της περιοχής μελέτης όπου δεν υπάρχουν μετρήσεις»

```
% MAP_SSRF2 is used to provide contour maps of concentration
% using the SSRF method. The values of the concentration are
% estimated on a square grid inside the convex hull of the
% data points using the SSRF predictor with SSRF model parameters
% that have been estimated by means of the INFER_SSRF2 program.
%
% INPUT PARAMETERS
% d:   Number of dimensions
% q2:   Shape Coefficient
% q3:   Correlation Length
% rc:   Search radius
% L:   Size of prediction grid
% nbmin: Minimum number of neighbour data for a point in order to calculate
SSRF estimate
%
% OUTPUT PARAMETERS
% XI:   x coordinates of square grid
% YI:   y coordinates of square grid
% ZI:   SSRF estimates on square grid
% IND:  Indices of locations inside the Convex Hull where SSRF
%       does not return numeric values
%
% NOTES
% There are grid sizes that lead to unrealistic estimates.
% Also, for some points inside the convex hull there are no numerical estimates.
```