



ΠΟΛΥΤΕΧΝΕΙΟ ΚΡΗΤΗΣ  
ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ ΟΡΥΚΤΩΝ ΠΟΡΩΝ

ΓΕΩΣΤΑΤΙΣΤΙΚΗ ΑΝΑΛΥΣΗ ΔΕΔΟΜΕΝΩΝ  
ΥΠΟΓΕΙΩΝ ΥΔΑΤΩΝ ΣΕ ΠΕΡΙΟΧΗ  
ΜΕΤΑΛΛΕΥΤΙΚΗΣ ΔΡΑΣΤΗΡΙΟΤΗΤΑΣ

Ευγενία Διαμαντοπούλου

ΕΞΕΤΑΣΤΙΚΗ ΕΠΙΤΡΟΠΗ

Επίκουρος Καθ. Βαρουχάκης Εμμανουήλ (Επιβλέπων)

Καθ. Γαλετάκης Μιχαήλ

Αναπληρωτής Καθ. Στειακάκης Εμμανουήλ,

Χανιά

Φεβρουάριος, 2023

## Περίληψη

Οι γεωστατιστικές μέθοδοι χρησιμοποιούνται όλο και περισσότερο στις γεωεπιστήμες, τη μηχανική και την περιβαλλοντική μεταλλευτική για την ανάλυση δεδομένων στο χώρο και στο χρόνο. Σε περιοχές μεταλλευτικής δραστηριότητας είναι πολύ σημαντική η εκτίμηση της χωρικής μεταβολής ρυπαντικών φορτίων στο εδάφους και τα υπόγεια ύδατα τόσο για περιβαλλοντικούς λόγους, για τη λήψη μέτρων, όσο και για την καθοδήγηση της αποκατάστασης έπειτα από το πέρας των εργασιών εξόρυξης. Η παρούσα διπλωματική εργασία στοχεύει στη γεωστατιστική ανάλυση ποιοτικών δεδομένων υπογείων υδάτων από τρεις γειτονικές περιοχές μεταλλευτικής δραστηριότητας στη Βόρεια Ελλάδα. Η μελέτη αυτή είναι απαραίτητη καθώς η ενδεχόμενη ρύπανση των υπογείων υδάτων κατά τη μεταλλευτική εκμετάλλευση μπορεί να προκαλέσει σοβαρά περιβαλλοντικά προβλήματα.

Σκοπός της παρούσας εργασίας είναι η χωρική ανάλυση και η δημιουργία χαρτών εκτίμησης της χωρικής μεταβολής ρυπαντικών φορτίων Cd και Mn στα υπόγεια ύδατα της περιοχής μελέτης. Η γεωστατιστική ανάλυση δεδομένων περιλαμβάνει τα εξής στάδια: προσδιορισμό της κατανομής πιθανότητας, προσαρμογή τάσεων, υπολογισμό του χωρικής εξάρτησης των δεδομένων πεδίου και τέλος εκτίμηση τιμών σε μη παρατηρημένα σημεία στο χώρο. Για την εφαρμογή της γεωστατιστικής ανάλυσης, χρησιμοποιήθηκαν δεδομένα από 62 σημεία παρατήρησης υπογείων υδάτων και η μέθοδος Κανονικού Kriging. Η υλοποίηση της μεθοδολογίας βασίστηκε σε κώδικα που αναπτύχθηκε σε περιβάλλον Matlab.

Στο πρώτο κεφάλαιο αναλύονται οι στόχοι της Διπλωματικής εργασίας, ενώ στη συνέχεια γίνεται μια σύντομη αναφορά στη μεταλλευτική δραστηριότητα της εταιρείας που εκμεταλλεύεται τα μεταλλεία και στο περιβαλλοντικό πρόγραμμα παρακολούθησης που έχει εγκατασταθεί στην περιοχή. Στο δεύτερο κεφάλαιο αναφέρονται βασικές θεωρητικές έννοιες της γεωστατιστικής στις οποίες βασίστηκε η παρούσα εργασία και αναλύονται σε μαθηματικό

επίπεδο οι μέθοδοι που εφαρμόστηκαν για την ανάλυση των δεδομένων. Στο τρίτο κεφάλαιο παρουσιάζεται η εφαρμογή της γεωστατιστικής ανάλυσης και περιγράφονται τα αποτελέσματα. Τέλος, στο τέταρτο κεφάλαιο παρουσιάζονται τα συμπεράσματα της γεωστατιστικής ανάλυσης και προτάσεις για περαιτέρω έρευνα.

Τα διαθέσιμα δεδομένα ρυπαντικού φορτίου εξετάζονται για την καταλληλότητα τους ώστε να εφαρμοστεί αξιόπιστα η γεωστατιστική ανάλυση. Τα δεδομένα δεν ακολουθούν την κανονική κατανομή και για αυτό χρησιμοποιήθηκε ο μετασχηματισμός τους με εφαρμογή του λογαρίθμου. Στη συνέχεια προσδιορίζονται οι στατιστικές παράμετροι του δείγματος και το διάγραμμα Voronoi για να εξεταστεί η ύπαρξη στασιμότητας ή όχι. Έπειτα πραγματοποιήθηκε έλεγχος ύπαρξης τάσεων στο πεδίο και αφαίρεση δευτέρου βαθμού πολυωνυμικής τάσης. Η χωρική εξάρτηση των δεδομένων προσδιορίζεται με τον υπολογισμό του εμπειρικού βαριογράμματος και την προσαρμογή σε αυτό του κατάλληλου θεωρητικού προτύπου. Μετά από αρκετές δοκιμές, επιλέχθηκε και προσαρμόστηκε στο εμπειρικό βαριόγραμμα το σφαιρικό μοντέλο βαριογράμματος και για τα δύο δείγματα ρυπαντικού φορτίου. Η χωρική μεταβολή του ρυπαντικού φορτίου εκτιμήθηκε με τη μέθοδος Κανονικού Kriging (Ordinary Kriging) ενώ έγινε αντιστροφή του μετασχηματισμού των δεδομένων στην αρχική τους κλίμακα. Η ακρίβεια των αποτελεσμάτων είναι ικανοποιητική ενώ η αβεβαιότητα των εκτιμήσεων προσδιορίζει περιοχές όπου χρειάζεται ενίσχυση της δειγματοληψίας.

## Abstract

*Geostatistical prediction methods are increasingly used in earth sciences and engineering to improve our knowledge of space and time attributes. During mining activities, it is essential to estimate any contamination of the soil and groundwater in the area for environmental reasons and to guide the reclamation once mining operations are finished. This thesis presents the geostatistical analysis of the groundwater content in specific pollutants concentrations (Cd and Mn) in a group of mines in Northern Greece. The Correlation between Cd and Mn was investigated during the preliminary analysis of the data. Data normalization was performed using the logarithm, as well as trend approximation, to improve data stationarity. Then, the spatial dependence of the measurements was determined using variogram analysis and estimation. The spatial variability of the pollutants content in groundwater was estimated and presented employing maps using the Ordinary Kriging (OK) method and inversing the transformations.*

## Ευχαριστίες

Οφείλω να απευθύνω θερμές ευχαριστίες στον επιβλέπων αυτής της διπλωματικής κ. Εμμανουήλ Βαρουχάκη για την ανάθεση αυτού του θέματος αλλά και για την εξαιρετική συνεργασία που υπήρξε μεταξύ μας. Ιδιαίτερα θα ήθελα να ευχαριστήσω το μέλος του Εργαστηρίου Γεωστατιστικής κ. Ανδρέα Παυλίδη για την πολύτιμη βοήθεια του καθ' όλη την διάρκεια της εκπόνησης της διπλωματικής εργασίας.

Επίσης, θα ήθελα να ευχαριστήσω την εταιρία «Ελληνικός Χρυσός» καθώς χωρίς τα δεδομένα που έχουν μοιραστεί στην ιστοσελίδα τους δεν θα ήταν δυνατή η εκπλήρωση της εργασίας αυτής.

Τέλος, θα ήθελα να ευχαριστήσω την οικογένειά μου για τη στήριξη τους όλα αυτά τα χρόνια των σπουδών μου καθώς και τους φίλους που ήταν κοντά μου και με στήριζαν.

# Contents

|          |  |          |
|----------|--|----------|
| <b>1</b> | <b>Εισαγωγή</b>                            | <b>1</b> |
| 1.1      | Σημασία του καθαρού νερού . . . . .        | 1        |
| 1.2      | Σκοπός της Διπλωματικής Εργασίας . . . . . | 2        |
| 1.3      | Ελληνικός Χρυσός . . . . .                 | 5        |
| <b>2</b> | <b>Βασικές Θεωρητικές Έννοιες</b>          | <b>7</b> |
| 2.1      | Τυχαία μεταβλητή - Τυχαία Πεδία . . . . .  | 8        |
| 2.1.1    | Διάγραμμα Voronoi . . . . .                | 9        |
| 2.1.2    | Τάση . . . . .                             | 10       |
| 2.1.3    | Διακύμανση . . . . .                       | 11       |
| 2.2      | Παράμετροι κατανομής Πιθανότητας . . . . . | 11       |
| 2.2.1    | Παράμετροι Θέσης . . . . .                 | 12       |
| 2.2.2    | Παράμετροι Διασποράς . . . . .             | 13       |
| 2.2.3    | Παράμετροι Ασυμμετρίας . . . . .           | 13       |
| 2.2.4    | Παράμετροι Κύρτωσης . . . . .              | 14       |
| 2.2.5    | Ιστόγραμμα . . . . .                       | 14       |
| 2.3      | Συνάρτηση Συνδιασποράς . . . . .           | 14       |
| 2.3.1    | Στατιστική Ομοιογένεια . . . . .           | 15       |
| 2.3.2    | Στατιστική Ισοτροπία . . . . .             | 16       |
| 2.3.3    | Μοντέλα Συνδιασποράς . . . . .             | 16       |
| 2.4      | Χωρική Εξάρτηση . . . . .                  | 18       |
| 2.4.1    | Βαριόγραμμα . . . . .                      | 19       |

## Contents

|          |   |           |
|----------|---|-----------|
| 2.4.2    | Ιδιότητες Βαριογράμματος . . . . .                  | 20        |
| 2.4.3    | Μοντέλα Βαριογράμματος . . . . .                    | 23        |
| 2.5      | Εκτίμηση Τυχαίου Πεδίου (Μέθοδος Kriging) . . . . . | 24        |
| 2.5.1    | Απλό Kriging . . . . .                              | 26        |
| 2.5.2    | Κανονικό Kriging . . . . .                          | 28        |
| 2.5.3    | Αξιολόγηση των Εκτιμήσεων . . . . .                 | 29        |
| <b>3</b> | <b>Γεωστατιστική Ανάλυση Δεδομένων</b>              | <b>32</b> |
| 3.1      | Περιγραφή των δεδομένων . . . . .                   | 34        |
| 3.2      | Στατιστική Ανάλυση . . . . .                        | 35        |
| 3.2.1    | Παράμετροι Κατανομών Πιθανότητας . . . . .          | 36        |
| 3.2.2    | Voronoi . . . . .                                   | 39        |
| 3.2.3    | Τάση . . . . .                                      | 41        |
| 3.2.4    | Βαριογραφία . . . . .                               | 42        |
| 3.2.5    | Kriging . . . . .                                   | 44        |
| <b>4</b> | <b>Συμπεράσματα - Προτάσεις</b>                     | <b>51</b> |
| 4.1      | Συμπεράσματα . . . . .                              | 51        |
| 4.2      | Προτάσεις για μελλοντική έρευνα . . . . .           | 53        |
| <b>5</b> | <b>Παράρτημα Α'</b>                                 | <b>54</b> |
| 5.1      | Κώδικας MatLab . . . . .                            | 54        |

## List of Figures

|      |  |    |
|------|--|----|
| 2.1  | Παράδειγμα μορφής διαγράμματος Voronoi . . . . .   | 10 |
| 2.2  | Παράδειγμα Βαριογράμματος . . . . .  | 20 |
| 3.1  | Διάγραμμα ροής . . . . .   | 33 |
| 3.2  | Χάρτης περιοχής μελέτης με τις θέσεις σταθμών παρακολούθησης υπογείων<br>υδάτων . . . . .  | 34 |
| 3.3  | Θέσεις σταθμών παρακολούθησης υπογείων υδάτων . . . . .  | 35 |
| 3.4  | Συσχέτιση Cd-Mn με μετασχηματισμό λογαρίθμου . . . . .   | 36 |
| 3.5  | Ιστόγραμμα και Norm Plot Cd με μετασχηματισμό λογαρίθμου . . . . .   | 37 |
| 3.6  | Ιστόγραμμα και NormPlot Mn με μετασχηματισμό λογαρίθμου . . . . .  | 38 |
| 3.7  | Αθροιστική κατανομή λογαρίθμου . . . . .   | 39 |
| 3.8  | Voronoi διαγράμματα χωρίς μετασχηματισμό . . . . .   | 40 |
| 3.9  | Διαγράμματα Τάσης με μετασχηματισμό λογαρίθμου . . . . .   | 42 |
| 3.10 | Βαριόγραμμα Cd(ppb) και Mn(ppm). Οι κύκλοι (ο) περιγράφουν το Εμ-<br>πειρικό Βαριόγραμμα ενώ η γραμμή το σφαιρικό μοντέλο. . . . . | 43 |
| 3.11 | Χάρτης περιοχής μελέτης με γειτονιά εκτίμησης όσο η ακτίνα συσχέτισης για<br>το Cd. . . . .  | 44 |
| 3.12 | Χάρτες εκτίμησης λογαριθμού της συγκέντρωσης των στοιχείων. . . . .  | 45 |
| 3.13 | Χάρτες τυπικής απόκλισης σφάλματος Kriging με μετασχηματισμό λογαρίθμου  | 46 |
| 3.14 | CV errors με μετασχηματισμό λογαρίθμου . . . . .   | 47 |
| 3.15 | Χάρτες εκτίμησης συγκέντρωσης των στοιχείων (ppb,ppm) . . . . .  | 48 |
| 3.16 | CV Errors . . . . .  | 49 |



## List of Tables

|     |   |    |
|-----|---|----|
| 3.1 | Παράμετροι Κατανομών Πιθανότητας Cd-Mn με μετασχηματισμό λογαρίθμου   | 36 |
| 3.2 | Γραμμικοί Συντελεστές Τάσης με μετασχηματισμό λογαρίθμου . . . . .  | 41 |
| 3.3 | Παράμετροι Βαριογράμματος Cd και Mn: $\sigma$ είναι η τυπική απόκλιση (Standard deviation), $h$ είναι η ακτίνα συσχέτισης (Range) και $c_0$ το φαινόμενο κόκκου (Nuggett) . . . . .                       | 43 |
| 3.4 | Τιμές επαλήθευσης στη διακύμανση, όπου MAE είναι το Μέσο Απόλυτο Σφάλμα, RMSE η ρίζα του Μέσου Τετραγωνικού Σφάλματος, MaxAE η τιμή του Μέγιστου Απόλυτου Σφάλματος και $\rho$ ο συντελεστής συσχέτισης . | 47 |
| 3.5 | Τιμές επαλήθευσης στη διακύμανση και τάση . . . . .   | 49 |

# 1 Εισαγωγή

## 1.1 Σημασία του καθαρού νερού

Η ύπαρξη καθαρού νερού στο πλανήτη μας είναι πρωταρχική ανάγκη. Το μη καθαρό νερό είναι ένα από τα μεγαλύτερα προβλήματα υγείας και περιβάλλοντος παγκοσμίως. Βοηθά σε πολλές σημαντικές λειτουργίες που εκτελεί το σώμα μας καθημερινά και δίχως αυτό, ένα ζωντανό ον δεν μπορεί να επιβιώσει για πολλές μέρες. Αποβάλλει τις τοξίνες, στέλνει τις θρεπτικές ουσίες στο σώμα και βοηθά στην πέψη. Η επαρκής ποσότητα υγρών είναι απαραίτητη για τη διατήρηση της ισορροπίας των ηλεκτρολυτών, όπως το νάτριο, το κάλιο και το μαγνήσιο, που ρυθμίζουν τις κυτταρικές διεργασίες και ελέγχουν την επικοινωνία των νευρών, τη λειτουργία της καρδιάς και τη συστολή των μυών. Ακόμα, το καθαρό νερό μειώνει τις απειλές ασθενειών. Τα μικρόβια που μεταδίδονται από το νερό αποτελούν τεράστιο πρόβλημα υγείας που μπορεί να φτάσουν μέχρι και σε θάνατο.

Επιπλέον, η οικονομία μιας χώρας εξαρτάται από το καθαρό νερό, καθώς χρησιμοποιείται στην γεωργία, στην κτηνοτροφία και στην παραγωγή ενέργειας. Παρέχει αξιοπρεπή υγιεινή και θρεπτική τροφή με αποτέλεσμα να συμβάλει σημαντικά ακόμα και στον τουρισμό.

Για να επιτευχθεί η χρήση καθαρού νερού απαιτούνται σημαντικές εξελίξεις στην επεξεργασία του. Πρέπει να γίνεται ταξινόμηση σύμφωνα με την τοξικότητα του αλλά και να εξελιχθεί ο καθαρισμός του. Δυστυχώς, μια εξόρυξη επηρεάζει το νερό λόγω της μεγάλης χρήσης του στην επεξεργασία του μεταλλεύματος. Μέσω της διαδικασίας αυτής, τα ύδατα ρυπαίνονται από τα λύματα των ορυχείων που απορρίπτονται, τους σωρούς του εξορυγμένου μεταλλεύματος και τις αποθέσεις αποβλήτων. Η ρύπανση αυτή καταλήγει στα υπόγεια νερά, με αποτέλεσμα όλο και περισσότερο να απειλούνται οι πηγές νερού που εξαρτάται το πόσιμο νερό.

## 1.2 Σκοπός της Διπλωματικής Εργασίας

Το 1997 στην Ουάσιγκτον της Αμερικής, γράφτηκε ένα βιβλίο για το πως οι εξορύξεις ρυπαίνουν απερίσκεπτα τα νερά της χώρας[8]. Στο βιβλίο αναφέρεται το εξής: "Μόλις ένα ορυχείο τεθεί σε λειτουργία, η προστασία των υδάτων πρέπει να παραμείνει ο ύψιστος στόχος της εταιρείας, ακόμη και αν αυτό σημαίνει μειωμένη παραγωγικότητα σε μέταλλευμα. Η υιοθέτηση αυτής της ηθικής είναι ο μόνος τρόπος για να διασφαλίσουμε ότι τα χρυσά όνειρα της εξόρυξης δεν θα μετατραπούν στον εφιάλτη των δηλητηριασμένων ρεμάτων". Στόχος λοιπόν της παρούσας διπλωματικής εργασίας είναι η μελέτη και η γεωστατιστική ανάλυση με σκοπό να προσδιοριστεί η ποιότητα των υπόγειων νερών και οι δυνατότητες χρήσης τους μετά από την υποβάθμιση τους. Η ρύπανση των υδάτων δημιουργείται από την απελευθέρωση χημικών ουσιών της μεταλλευτικής δραστηριότητας, με αποτέλεσμα να επηρεάζουν τα φυσικά χαρακτηριστικά του νερού και να υποβαθμίζονται οι δυνατότητες χρήσης του. Σύμφωνα με την Ευρωπαϊκή Επιτροπή για την προστασία των υπόγειων υδάτων στην Ευρώπη τα υπόγεια νερά κινούνται αργά μέσα στο έδαφος, με αποτέλεσμα οι επιπτώσεις από τις ανθρώπινες δραστηριότητες να είναι σχετικά μακροπρόθεσμες. Αυτό σημαίνει ότι ρύπανση που εμφανίστηκε δεκαετίες πριν, ανεξαρτήτως προέλευσης μπορεί να εξακολουθεί να είναι επικίνδυνη για την ποιότητα των υπόγειων νερών.[26] Οι επιπτώσεις αυτές εξαρτώνται από διάφορους παράγοντες, όπως ο σχηματισμός πετρωμάτων που εξορύσσονται, το είδος του εδάφους, η τεχνολογία που χρησιμοποιείται κατά την εξόρυξη και τέλος, η περιβαλλοντική δέσμευση της κάθε εταιρίας με την γνώση και την ικανότητα του να επιβάλλουν τους περιβαλλοντικούς κανόνες. Ένα από τα μεγαλύτερα προβλήματα είναι πως η διαδικασία της εξόρυξης έχει αναπτυχθεί αρκετά με αποτέλεσμα την διαχείριση περισσότερων πετρωμάτων και μεταλλευτικών υλικών. Επομένως, τα απόβλητα των ορυχείων έχουν πολλαπλασιαστεί, ειδικά σε ανοιχτά ορυχεία όπου οι εκσκαφές περιλαμβάνουν μεγάλες ποσότητες άγονων υλικών. [21]

Υπάρχουν τέσσερις βασικοί τύποι επιπτώσεων της εξόρυξης στην ποιότητα των υδάτων. Η όξινη απορροή (Acid Rock Drainage - ARD) [19] γίνεται μια φυσική διαδικασία όπου παράγεται θειικό οξύ όταν σουλφίδια των πετρωμάτων έρθουν σε επαφή με τον αέρα και το

## Chapter 1. Εισαγωγή

νερό. Όταν γίνεται εξόρυξη σε ορυχείο μεγάλων ποσοτήτων πετρωμάτων που περιέχουν θειούχα ορυκτά γίνεται αντίδραση με το νερό ή το οξυγόνο και δημιουργείται θειικό οξύ. Το οξύ στη συνέχεια απομακρύνεται με το βρόχινο νερό ή την επιφανειακή αποστράγγιση και καταλήγει σε ποτάμια, λίμνες και υπόγεια ύδατα. Το ARD μπορεί να ρυπαίνει το νερό και να δημιουργήσει σοβαρά προβλήματα στην υδρόβια ζωή, καθώς και να υποβαθμίσει την ποιότητα του πόσιμου νερού. Η ρύπανση του νερού από βαρέα μέταλλα προκαλείται όταν μέταλλα όπως το κάδμιο, ο μόλυβδος, το αρσενικό έρχονται σε επαφή με το νερό. Καθώς το νερό χρησιμοποιείται στις διαδικασίες εμπλουτισμού και επεξεργασίας των μεταλλευμάτων, τα μέταλλα μεταφέρονται με αποτέλεσμα να καταλήγουν στα υπόγεια νερά και στον υδροφόρο ορίζοντα.[7] Ακόμα, χημικές ουσίες μπορεί να ρυπαίνουν το νερό όπως το κυάνιο ή το θειικό οξύ. Αυτές οι ουσίες χρησιμοποιούνται συχνά από τις εταιρείες στην διαδικασία διαχωρισμού του μεταλλεύματος και όταν διαρρέουν από το χώρο του ορυχείου και καταλήγουν σε κοντινά υδάτινα ρεύματα είναι τοξικές για τον άνθρωπο και τα ζώα. Τέλος, η μεταλλευτική δραστηριότητα αναδιατάσσει το έδαφος και τα πετρώματα σε ένα ορυχείο και σε ένα χώρο συγκέντρωσης αποβλήτων. Η διάβρωση του εδάφους και οι ανεπιθύμητες καθιζήσεις μπορεί να μεταφέρουν σημαντικές ποσότητες ιζημάτων σε ποτάμια και λίμνες. Σαν αποτέλεσμα, μπορούν να φράξουν τις κοίτες των ποταμών και να καλύψουν την βλάστηση μιας λεκάνης απορροής όπως και να ρυπαίνουν τους υδρόβιους οργανισμούς.[17]

Τα βαρέα μέταλλα είναι μέταλλα τα οποία έχουν ειδικό βάρος μεγαλύτερο από του σιδήρου και θεωρούνται χημικά στοιχεία τοξικά για τον άνθρωπο ακόμα και σε μικρές συγκεντρώσεις. Εισάγεται στον οργανισμό του ανθρώπου μέσω τροφών, του πόσιμου νερού και του αέρα. Το κάδμιο και το μαγγάνιο απελευθερώνονται από τα εδάφη, τα ιζήματα και τα πετρώματα μέσω της αποσάθρωσης. Κάποια από αυτά είναι απαραίτητα για τον ανθρώπινο μεταβολισμό και την σωστή λειτουργία του ανθρώπινου σώματος αλλά οι μεγάλες ποσότητες προκαλούν σοβαρά προβλήματα.[23] Η απόρριψη βιομηχανικών λυμάτων σε υδάτινους αποδέκτες αποτελεί μείζον περιβαλλοντικό πρόβλημα παγκοσμίως, επειδή περιέχει τοξικά βαρέα μέταλλα και άλλους οργανικούς και ανόργανους ρύπους. Τα βιομηχανικά απόβλητα συχνά περιέχουν ιόντα βαρέων μετάλλων τα οποία δεν είναι βιοδιασπώμενα και πολλά από αυτά είναι διαλυτά σε υδατικά μέσα και εύκολα διαθέσιμα στους ζωντανούς οργανισμούς.[3] Πολυάριθμες μέθοδοι και τεχνικές

## Chapter 1. Εισαγωγή

έχουν χρησιμοποιηθεί για την απομάκρυνση των βαρέων μετάλλων από το νερό, όπως η χημική καταβύθιση, η ιοντοανταλλαγή, η φυτοεσχύλιση, η υπερδιήθηση, η αντίστροφη ώσμωση και η προσρόφηση.[3] Στην παρούσα εργασία μελετάμε το κάδμιο (Cd) και το μαγγάνιο (Mn).

Το Κάδμιο (Cd) δεν υπάρχει συνήθως στο περιβάλλον σαν καθαρό μέταλλο, αλλά κυρίως ως μεταλλευμα με στοιχεία όπως το οξυγόνο (οξείδιο του καδμίου), το θείο (σουφλίδιο καδμίου) και το χλώριο (χλωριούχο κάδμιο). Ακόμα, βρίσκεται σε μεγάλες ποσότητες σε στερεή μορφή και χρησιμοποιείται ως υποπροϊόν σε άλλα μεταλλεύματα όπως ο χαλκός, ο μόλυβδος και ο ψευδάργυρος. Σε αυτές τις διαφορετικές μορφές το κάδμιο διαλύεται στο νερό σε διαφορετικά στάδια, με αποτέλεσμα το μεταλλικό κάδμιο να μην εξαφανίζεται από το περιβάλλον. Χρησιμοποιείται ευρέως στη βιομηχανία σε μπαταρίες, σε χρώματα, σε μεταλλικές επιστρώσεις αλλά και σε πλαστικά.[23]. Το επιτρεπτό όριο που μπορεί να βρίσκεται στο περιβάλλον σύμφωνα με την εταιρία «Ελληνικός Χρυσός» είναι 5 *ppb* [11, 1]. Ο Διεθνής Οργανισμός Έρευνας για τον καρκίνο όμως πιστεύει πως το κάδμιο είναι από τα πιο γνωστά καρκινογόνα βαρέα μέταλλα για τον άνθρωπο. Οι πιο συχνές μελέτες εργαζομένων όπου έχουν εκτεθεί σε υψηλά επίπεδα Cd παρουσιάζουν καρκίνο στους πνεύμονες. Επίσης, έχει παρατηρηθεί βλάβη στα νεφρά και διαταραχές στα οστά και στο αίμα λόγω ρύπανσης του πόσιμου νερού. Η ύπαρξη του καδμίου, η οποία μπορεί εύκολα να εκτεθεί στον άνθρωπο βρίσκεται στον καπνό, στο συκώτι μόσχου και στα όστρακα.[31] Όσον αφορά το περιβάλλον, το κάδμιο είναι πολύ τοξικό για τους υδρόβιους οργανισμούς όπως ακόμα μπορεί να προκαλέσει μακροχρόνιες αρνητικές επιδράσεις στο υδάτινο περιβάλλον. Υπολογίζεται, πως τόνοι καδμίου κάθε χρόνο απελευθερώνονται στο περιβάλλον όπου προέρχονται από την διάβρωση των πετρωμάτων αρχικά σε νερό των ποταμών και έπειτα στους ωκεανούς.

Η ύπαρξη μαγγανίου (Mn) στο νερό σε μεγάλες συγκεντρώσεις μπορεί επίσης να προκαλέσει ανεπιθύμητα αποτελέσματα. Υπάρχει στο νερό ως  $Mn^{2+}$  (δισθενές ιόν) και λόγω των οργανοληπτικών ιδιοτήτων του θεωρείται ρύπος. Το συνιστάμενο όριο Μαγγανίου στο πόσιμο νερό σύμφωνα με την Ελληνική Νομοθεσία είναι 0,05 ppm [24]. Έχει αποδειχθεί πως μικροοργανισμοί ευθύνονται για την ύπαρξη και την ανακύκλωση του μαγγανίου στο φυσικό περιβάλλον[12]. Η παρουσία του φαίνεται από το χρώμα του νερού που εμφανίζεται καφέ ή κοκκινωπό. Ακόμα, η γεύση του νερού αλλάζει αισθητά καθώς θυμίζει μεταλ-

λική. Δημιουργεί προβλήματα και στην διανομή νερού, καθώς ιζήματα μαγγανίου εναποθέτονται στους σωλήνες, με αποτέλεσμα να μειώνεται η διάμετρος του και να φράζει. Η συσσώρευση αυτή δημιουργεί ανάπτυξη βακτηρίων στο δίκτυο ύδρευσης και καταλήγει το νερό με δυσάρεστη οσμή και κακή γεύση. [13] Η χρόνια έκθεση στο μαγγάνιο προκαλεί ψυχολογικές διαταραχές, αναπνευστικά προβλήματα και λοιμώξεις.

### 1.3 Ελληνικός Χρυσός

Ο Ελληνικός Χρυσός είναι μια μεταλλευτική εταιρία που παράγει χρυσό (Au), ασήμι (Ag), χαλκό (Cu), μόλυβδο (Pb), και ψευδάργυρο (Zn) σε μορφή συμπυκνωμάτων τα οποία και εξάγει. Η αξία των εξαγωγών από τα συμπυκνώματα των μετάλλων αυτών ανέρχεται σε 180-190 εκατομμύρια ευρώ το χρόνο. Είναι ιδιοκτήτης των μεταλλείων Κασσάνδρας που βρίσκονται στην βορειο-ανατολική Χαλκιδική. Το έτος 2022 έχει δύο ενεργά μεταλλεία, το μεταλλείο της Ολυμπιάδας και το μεταλλείο Στρατωνίου ή Μαύρες Πέτρες, όπως επίσης και ένα ανενεργό, το μεταλλείο των Σκουριών. [29] Η εταιρία Ελληνικός Χρυσός δραστηριοποιείται στη Χαλκιδική και διαθέτει από τα καλύτερα προγράμματα περιβαλλοντικής παρακολούθησης και διαχείρισης στην Ελλάδα. Απο τα σημαντικότερα προτερήματα αυτού του προγράμματος είναι ότι ελέγχονται οι περιβαλλοντικές επιδόσεις όλο το εικοσιτετράωρο και ενημερώνεται η ηλεκτρονική πλατφόρμα του διαδικτύου ώστε να είναι διαθέσιμη σε όλους. Αυτή η παρακολούθηση κυρίως ελέγχει την ρύπανση της ατμόσφαιρας, του εδάφους, των υδάτων όπως ακόμα και την ύπαρξη και επικινδυνότητα της σκόνης, του θορύβου και των δονήσεων που προέρχονται από την εκμετάλλευση.

Σύμφωνα με το ήδη υπάρχον νομοθετικό πλαίσιο οι επενδυτές από το εξωτερικό επενδύουν στην εταιρία με μεγαλύτερη ασφάλεια, με αποτέλεσμα να χρησιμοποιούνται καινούριες τεχνολογίες όπως αυτή της Ξηράς Απόθεσης των αποβλήτων. Με αυτή την τεχνολογία μειώνεται το Περιβαλλοντικό αποτύπωμα στην περιοχή όπου αντί να αποθέτονται τα μεταλλευτικά κατάλοιπα της διεργασίας σε δύο χώρους, αποτίθενται σε ένα με σκοπό την μείωση του Περιβαλλοντικού αποτυπώματος στην περιοχή. Έχουν ήδη εφαρμοστεί έργα αποκατάστασης τέτοιας τεχνολογίας στην περιοχή της Ολυμπιάδας όπου τα υπολείμματα της προηγούμενης μεταλλευτικής δραστηριότητας έχουν μεταφερθεί σε έναν χώρο ονόματι Κοκκινόλακκας.[29]

## Chapter 1. Εισαγωγή

Η εταιρία σύμφωνα με Υπουργική Απόφαση ωφείλει να παρακολουθεί την ποιότητα των συστημάτων υπόγειων υδάτων στην ευρύτερη περιοχή του έργου. Έτσι, με υδρογεωτρήσεις, εκτός από την μηνιαία μέτρηση της στάθμης, συλλέγεται και δείγμα νερού όπου γίνονται πλήρεις αναλύσεις φυσικοχημικών παραμέτρων και βαρέων μετάλλων. Τα δείγματα αυτά συγκρίνονται σύμφωνα με τα επιτρεπτά όρια που ορίζει το Υπουργείο Περιβάλλοντος. Η αξιολόγηση τους γίνεται ξεχωριστά για τα τρία μεταλλεία Ολυμπιάδας, Μαύρων Πετρών και Σκουριών. Τα δείγματα καταλήγουν στο εργαστήριο ποιοτικού ελέγχου του Ελληνικού Χρυσού σε 24 ώρες από την δειγματοληψία τους.

Τα μεταλλευτικά κατάλοιπα μπορούν να αποτεθούν με διαφορετικούς τρόπους. Μπορούν να αποτεθούν απευθείας με την μορφή πολφού ή μετά από πύκνωση με την μορφή «πάστας», είτε μπορούν να διηθηθούν και να αποτεθούν σε στερεή μορφή όπως η τεχνολογία της Ξηράς Απόθεσης. Τα ξηρά κατάλοιπα δημιουργούνται από το υδαρές μίγμα του λεπτομερούς στείρου υλικού που προκύπτει από την διαδικασία ανάκτησης των μετάλλων. Το νερό απομακρύνεται έπειτα από φιλτράρισμα και το υλικό καταλήγει σε μορφή υγρής άμμου, όπου στο τέλος συμπιέζεται. Αυτή η τεχνολογία απομακρύνει και ανακυκλώνει το νερό μειώνοντας το αποτύπωμα στο περιβάλλον. Είναι ασφαλής ακόμα και σε έντονα καιρικά φαινόμενα όπως πλημμύρες και σεισμούς. [11]

Τα δεδομένα για αυτή την διπλωματική αναφορά πάρθηκαν από την ιστοσελίδα <https://environmental.hellas-gold.com/> της εταιρίας «Ελληνικός Χρυσός».

## 2 Βασικές Θεωρητικές Έννοιες

Η Γεωστατιστική είναι κλάδος της Χωρικής Στατιστικής που πραγματεύεται μελέτες και αναλύσεις στο χώρο. Οι μελέτες αυτές βασίζονται στη θεωρία των πιθανοτήτων και της στατιστικής. Η γεωστατιστική παρέχει μια μεθοδολογία για την ποσοτικοποίηση της χωρικής αβεβαιότητας. Η κατανομή πιθανοτήτων που παρουσιάζεται με την στατιστική είναι ένας τρόπος για την αναπαράσταση του εύρους πιθανών τιμών μιας παραμέτρου που μας ενδιαφέρει. [6] Χρησιμοποιεί τυχαία πεδία για την ανάλυση χωρικών δεδομένων με δεδομένα που προκύπτουν από συνεχείς υποκείμενες διαδικασίες. Συνεπώς σημειακά φαινόμενα αποκλείονται από την γεωστατιστική ανάλυση. Χρησιμοποιείται για την εκτίμηση αποθεμάτων σε εξορυκτικές δραστηριότητες και σε περιβαλλοντικές παρατηρήσεις όπως την εξάπλωση ρύπανσης σε εδαφικές ή υδάτινες εκτάσεις. Έχει εφαρμογές στη χωροχρονική χαρτογράφηση, στη μεταλλευτική, στην υδρογεωλογία ακόμα και στη γεωθερμία. Συγκεκριμένα, μπορεί να χρησιμοποιηθεί για να εκτιμήσει το βάθος και το πάχος ενός γεωλογικού στρώματος, το πορώδες και την διαπερατότητα των πετρωμάτων, τις ιδιότητες ενός εδάφους όπως επίσης και τις συγκεντρώσεις ρύπανσης σε περιοχές. Ο στόχος της Γεωστατιστικής είναι να προσεγγίσει και να αναγνωρίσει την μεταβλητότητα των φαινομένων της φύσης και να την ενσωματώσει σε μοντέλα στοχαστικής φύσης όπου θα εντοπίζονται οι σχέσεις μεταξύ των φαινομένων αυτών και θα χρησιμοποιούνται για την επίλυση συγκεκριμένων προβλημάτων.[15]

Σε αυτό το κεφάλαιο παρουσιάζονται με απλά και σύντομα λόγια οι βασικές ιδέες στις οποίες βασίστηκε η παρούσα Γεωστατιστική ανάλυση.



## 2.1 Τυχαία μεταβλητή - Τυχαία Πεδία

Η γνώση της μεταβλητής που μας ενδιαφέρει με απόλυτη ακρίβεια σε ένα σημείο του χώρου είναι πρακτικά αδύνατη. Αυτό οφείλεται στη μεταβλητότητα του πεδίου στην οποία δύναται να συνεισφέρουν διαφορετικά φαινόμενα όπως περιβαλλοντικοί παράγοντες, τοπογραφικό, μεταλλευτική δραστηριότητα κ.α. [20] Προκειμένου να περιγραφούν μαθηματικά οι μεταβολές αυτές χρησιμοποιείται η έννοια της *τυχαίας μεταβλητής*. Η Τυχαία Μεταβλητή είναι μια συνάρτηση όπου το πεδίο ορισμού της είναι ένας δειγματοληπτικός χώρος. Για παράδειγμα, μπορούμε να σκεφτούμε την ποσότητα του νερού σε οποιαδήποτε δεδομένη τοποθεσία ως τυχαία μεταβλητή που παίρνει διάφορες τιμές στο χρόνο. Δηλαδή, η έννοια της τυχαίας μεταβλητής συνδέεται με έναν αριθμό πιθανών καταστάσεων [15]. Η Τυχαία Μεταβλητή συμβολίζεται με το κεφαλαίο γράμμα  $X$  ενώ με το πεζό γράμμα  $x$  η τιμή της. Διακριτή μεταβλητή καλείται μια τυχαία μεταβλητή όπου οι πιθανές τιμές περιορίζονται,  $x_i$  με  $i = 1, \dots, N$ . Έχει δηλαδή έναν μετρήσιμο (ακέραιο) αριθμό ( $i$  έως  $N$ ) πιθανών τιμών. Οι αντίστοιχες κατανομές τους καλούνται διακριτές κατανομές. Αντίθετα, συνεχής μεταβλητή ονομάζεται η τυχαία μεταβλητή της οποίας η τιμή παίρνει τιμές από ένα συνεχές σύνολο, ενώ οι αντίστοιχες κατανομές τους καλούνται συνεχείς κατανομές.

Η αναμενόμενη τιμή μιας τυχαίας μεταβλητής  $X$  ή μέση τιμή για ένα σύνολο καταστάσεων συμβολίζεται με  $E[X]$ . Εάν η κατανομή πιθανοτήτων των  $X$  ακολουθεί μία συνάρτηση πυκνότητας πιθανότητας  $f(x)$ , τότε η αναμενόμενη τιμή είναι

$$E[X] = \int_{-\infty}^{+\infty} x f(x) dx \quad (2.1)$$

Ένα τυχαίο πεδίο μπορεί να θεωρηθεί ως μια συλλογή τυχαίων μεταβλητών που κατανέμονται στο χώρο. Μια βασική ιδιότητα των τυχαίων πεδίων είναι ότι οι τυχαίες μεταβλητές που τα αποτελούν συσχετίζονται. Από μαθηματική άποψη, τα τυχαία πεδία (χωρικά ή χωρο-χρονικά) αποτελούν μια περιοχή που μελετά τις τυχαίες συναρτήσεις.

Σε ένα πραγματικό πεδίο όπου συλλέγονται τα δείγματα με σκοπό την ανάλυση τους, είναι εύκολο να κατανοήσουμε πως είναι αδύνατη η δειγματοληψία να είναι συνεχής στο χώρο και στο χρόνο. Υπάρχουν δυσκολίες εφαρμογής που προκύπτουν από την γεωμετρία του

αμοιβαίου χώρου, δηλαδή είναι δύσκολη η συλλογή δεδομένων πάνω σε σταθερό πλέγμα, όπως κανονικά πολύγωνα ή τρίγωνα. Σχεδόν πάντα πραγματοποιούνται γεωτρήσεις που απέχουν διαφορετικές αποστάσεις μεταξύ τους και βρίσκονται σε μη κανονικό (τυχαίο) πλέγμα. Ωστόσο, αυτό το πλέγμα καλύπτει την περιοχή ενδιαφέροντος.[32] Η εκτίμηση δεν μπορεί να γίνει παντού στο χώρο. Συνήθως επιλέγεται ένα πυκνό, κανονικό πλέγμα για να γίνουν οι εκτιμήσεις. Στη χωρική εκτίμηση τα άγνωστα σημεία του πλέγματος παίρνουν τιμές που εξαρτώνται από τα γειτονικά τους σημεία, με μεγαλύτερη επιρροή από τα πιο κοντινά. Δηλαδή η Γεωστατιστική ανάλυση δημιουργεί ένα μοντέλο με συνεχείς τιμές σύμφωνα με τις αποστάσεις των γειτονικών σημείων και την χωρική συσχέτιση.

Σε αυτή την εργασία, ένα τυχαίο πεδίο θα συμβολίζεται με το κεφαλαίο γράμμα  $X(s)$ , και παρουσιάζει ένα σύνολο με πιθανές καταστάσεις. Το διάνυσμα  $s$  αντιπροσωπεύει τη θέση στη περιοχή μελέτης. Οι ποσότητες αυτών των πεδίων στο χώρο καθορίζονται από εξαρτημένες τυχαίες μεταβλητές. Αυτές συμβολίζονται με το μικρό γράμμα  $x$  και είναι μεταβλητές που μπορούν να πάρουν διαφορετικές αριθμητικές τιμές σύμφωνα με μία πιθανότητα.

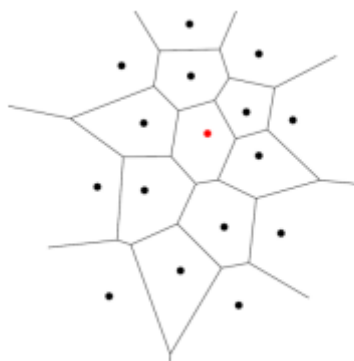
Η συλλογή τυχαίων μεταβλητών περιγράφει μια στοχαστική διαδικασία, η οποία αναπαριστά την βελτίωση ενός συστήματος τυχαίων τιμών με την πάροδο του χρόνου. Στη Γεωστατιστική θεωρούμε πως ένα τυχαίο πεδίο χωρίζεται σε ένα αιτιοκρατικό μέρος (που το λέμε Τάση  $m_x(s)$ ) και ένα στοχαστικό μέρος που ονομάζεται διακύμανση  $X'(s)$ .

$$X(s) = m_x(s) + X'(s) \quad (2.2)$$

### 2.1.1 Διάγραμμα Voronoi

Το διάγραμμα Voronoi θεωρείται από τις πιο χρήσιμες δομές δεδομένων για την επίλυση υπολογιστικών προβλημάτων. Είναι η απεικόνιση μιας περιοχής όπου διαχωρίζεται στο χώρο σύμφωνα με την απόσταση από ένα σύνολο διακριτών σημείων.

Αν, έστω  $P = \{p_1, p_2, \dots, p_n\}$  είναι ένα σύνολο διακριτών σημείων στο επίπεδο, όπου κάθε σημείο θεωρείται γεώτρηση, τότε το διάγραμμα Voronoi ορίζεται ως ένα επίπεδο που διαιρείται σε περιοχές με τέτοιο τρόπο, έτσι ώστε ένα σημείο  $u$  που βρίσκεται πιο κοντά στη γεώτρηση  $p_i$  με συντεταγμένες  $s_i$  να παίρνει την τιμή  $\hat{X}(u) = X(s_i)$ . Αυτές οι περιοχές



Εικόνα 2.1: Παράδειγμα μορφής διαγράμματος Voronoi

σχηματίζουν ένα κλειστό σχήμα και συγκεκριμένα πολυγωνικό.[4]

Ο κύριος λόγος εφαρμογής του είναι ότι με απλή δομή παρουσιάζει χρήσιμες πληροφορίες για μια μεγάλη περιοχή μελέτης.

### 2.1.2 Τάση

Οι τιμές των δεδομένων συνήθως σχετίζονται με την θέση στην οποία βρίσκονται (χωρικά εξαρτημένες). Δηλαδή σε ένα πλέγμα οι τιμές των σημείων σχετίζονται συνήθως με τις τιμές των διπλανών σημείων. Υπάρχει δηλαδή μια μαθηματική συνάρτηση που συσχετίζει τις τιμές των σημείων που μελετάμε με την θέση των σημείων αυτών στους άξονες x,y. Η επαλήθευση μιας τέτοιας μαθηματικής συνάρτησης περιγράφει όσον τον δυνατόν καλύτερα την τάση των δεδομένων στο χώρο. Αυτή η μαθηματική συνάρτηση μπορεί να είναι γραμμική ή πολυωνυμική εξίσωση. Υπάρχουν όμως περιπτώσεις όπου κοντινά σημεία έχουν αρκετά διαφορετικές τιμές μεταξύ τους. Τότε, η ανάλυση δεν μπορεί να συνεχιστεί και χρειάζεται να βρεθεί η μορφή της τάσης και να αφαιρεθεί για την συνέχεια της διαδικασίας. [2]

Η Γραμμική Χωρική Εξάρτηση υποδηλώνει πως η κλίση της μαθηματικής συνάρτησης παραμένει σταθερή σε μια εφαρμογή και παίρνει την παρακάτω μορφή για δύο διαστάσεις:

$$Mx = a_0 + a_1X + a_2Y \quad (2.3)$$

Η Πολυωνυμική Χωρική Εξάρτηση συνήθως είναι εξίσωση δευτέρου βαθμού που χρησι-

μπορείται για να περιγράψει την τάση των δεδομένων όταν δεν ταιριάζει η Γραμμική συνάρτηση. Εμφανίζεται με την παρακάτω εξίσωση:

$$Mx = a_0 + a_1X + a_2Y + a_3X^2 + a_4Y^2 + a_5XY \quad (2.4)$$

Αφού αφαιρεθεί η τάση τότε προκύπτει ένα άλλο πεδίο που θεωρείται στοχαστικό, η Διακύμανση.

### 2.1.3 Διακύμανση

Η διακύμανση είναι μια βασική παράμετρος για την περιγραφή της χωρικής μεταβλητότητας μιας περιοχής μέσω ενός τυχαίου πεδίου. [18] Οι διακυμάνσεις είναι αποτέλεσμα μιας στοχαστικής συνιστώσας ενός τυχαίου πεδίου. Μαθηματικά, προσδιορίζονται σύμφωνα με δομές κατανομών πιθανοτήτων. Δηλαδή, οι διακυμάνσεις αποδίδονται σε μετρήσεις που δεν μπορούν να χαρακτηριστούν πλήρως. Η διακύμανση του πεδίου  $X(s_1)$  συμβολίζεται με  $X'(s_1)$  και περιγράφεται από την παρακάτω σχέση γύρω από τη μέση τιμή στο σημείο  $s_1$ :

$$X'(s_1) = X(s_1) - m_X(s_1) \quad (2.5)$$

Η διακύμανση μπορεί να περιγραφεί ως ένα διαφορετικό τυχαίο πεδίο με μέση τιμή  $E[X'(s)] = 0$ . Άρα μπορεί να εκφράσει τις αποκλίσεις που υπάρχουν στις τιμές των μεταβλητών  $X$  από την μέση τιμή. [32]

## 2.2 Παράμετροι κατανομής Πιθανότητας

Ξεκινώντας μια γεωστατιστική ανάλυση πρέπει να υπολογιστούν οι Παράμετροι Κατανομής Πιθανότητας. Είναι στατιστικές παράμετροι, οι οποίοι χαρακτηρίζουν το τυχαίο πεδίο και τις τυχαίες κατανομές που το αποτελούν. Αυτές οι παράμετροι, καθορίζονται από την αθροιστική συνάρτηση κατανομής και τη συνάρτηση πυκνότητας πιθανότητας. Χωρίζονται σε μεγέθη θέσης (measures of location), μεγέθη διασποράς (measures of spread) και μεγέθη μορφής (measures of shape).[32]

### Ελάχιστη και Μέγιστη τιμή

Η Ελάχιστη  $x_{min}$  και η Μέγιστη  $x_{max}$  τιμή των δεδομένων προσδιορίζουν το εύρος των εμφανιζόμενων τιμών της δειγματοληψίας.

### 2.2.1 Παράμετροι Θέσης

#### Διάμεσος

Ο Διάμεσος  $x_{med}$  στην περίπτωση που μελετάμε δείγματα και οι τιμές ταξινομούνται σε αύξουσα σειρά ώστε να ισχύει  $x_i \leq x_{i+1}$ , αντιστοιχεί στο μέσο της κατανομής. Δηλαδή, η διάμεσος αντιστοιχεί στο σημείο όπου η καμπύλη πυκνότητας πιθανότητας χωρίζεται σε δύο ίσα μέρη. Ο υπολογισμός της γίνεται ως εξής: Αν το σύνολο του δείγματος είναι περιττός αριθμός, ο διάμεσος δίνεται από την κεντρική τιμή, ενώ αν είναι ζυγός αριθμός υπολογίζεται από τον μέσο όρο των δύο μεσαίων τιμών.

#### Μέση τιμή

Η έννοια της διαμέσου δεν πρέπει να μπερδεύεται με την μέση τιμή  $m_x = \bar{x}$ , η οποία δηλώνει τον μέσο όρο των τιμών της δειγματοληψίας.

Η Μέση Τιμή μιας τυχαίας μεταβλητής ορίζεται από τον τύπο:

$$m_x(s) = E[X(s)] \quad (2.6)$$

όπου το  $X(s)$  είναι ένα τυχαίο πεδίο και το  $E$  η μέση τιμή σε ένα σύνολο καταστάσεων. Ακόμα μπορεί να περιγραφεί ως η θέση στο κέντρο βάρους του σχήματος που ορίζει η συνάρτηση πυκνότητας πιθανότητας με τον οριζόντιο άξονα.

Η Μέση τιμή και η Διάμεσος ταυτίζονται μόνο αν η κατανομή είναι συμμετρική και έχει μία μόνο μέγιστη τιμή (Μονοκόρυφη). Αυτό όμως είναι πολύ σπάνιο.

### 2.2.2 Παράμετροι Διασποράς

Η Διασπορά (Variance) ενός τυχαίου πεδίου ορίζεται ως η μέση τιμή του τετραγώνου των διακυμάνσεων και δίνεται από τον τύπο:

$$\sigma_x^2(s) \equiv E[\{X(s) - m_x(s)\}^2] = E[X'^2(s)] \quad (2.7)$$

Ουσιαστικά δείχνει την κατανομή ενός τυχαίου πεδίου γύρω από την μέση τιμή του. Αν δηλαδή παρατηρηθεί μικρή διασπορά, σημαίνει πως υπάρχει συγκεντρωμένη κατανομή γύρω από την μέση τιμή. Η διασπορά είναι δυνατό να μεταβάλλεται από σημείο σε σημείο, ενώ αν το πεδίο είναι στατιστικά ομογενές τότε παραμένει σταθερή σε όλα τα σημεία.

#### Τυπική απόκλιση

Η Τυπική απόκλιση (standard deviation) συμβολίζεται με  $\sigma_x$  και ορίζεται ως η τετραγωνική ρίζα της διασποράς που προσδιορίζει ποσοτικά την απόκλιση της κατανομής από την μέση τιμή της. Αυτό το στατιστικό μέγεθος χρησιμοποιείται για να περιγράψει το πιθανό σφάλμα των πειραματικών μετρήσεων αλλά και το σφάλμα εκτίμησης των παραμέτρων μιας κατανομής πιθανότητας από ένα δείγμα.

### 2.2.3 Παράμετροι Ασυμμετρίας

Ο Δειγματικός Συντελεστής Ασυμμετρίας  $s_x$  (Coefficient of Skewness) είναι ένας συντελεστής ο οποίος παρουσιάζει το εύρος των υψηλών ή χαμηλών τιμών της δειγματοληψίας με αποτέλεσμα να δείχνει αν ένα ιστόγραμμα έχει θετική ή αρνητική ασυμμετρία. Για παράδειγμα, αν υπάρχουν πολλές χαμηλές τιμές το ιστόγραμμα παρουσιάζει αρνητική ασυμμετρία ενώ αν είναι υψηλές τιμές θετική ασυμμετρία. Ακόμα, ο συντελεστής ασυμμετρίας είναι μηδενικός όταν οι κατανομές είναι συμμετρικές.

## 2.2.4 Παράμετροι Κύρτωσης

Ο Συντελεστής Κύρτωσης  $kx$  περιγράφεται από τις ακραίες τιμές της κατανομής. Αυτός ο συντελεστής είναι ίσος με τρία. Κατανομές με συντελεστή κυρτότητας μεγαλύτερο από τρία ονομάζονται λεπτοκυρτικές, ενώ αυτές με συντελεστή κύρτωσης μικρότερο του τρία ονομάζονται πλατυκυρτικές. Στην πρώτη κατηγορία οι τιμές βρίσκονται στις ουρές παρά στην κανονική κατανομή, ενώ στη δεύτερη συγκεντρώνονται γύρω από το κέντρο.

## 2.2.5 Ιστόγραμμα

Το Ιστόγραμμα είναι ένα γράφημα παρουσίασης ομαδοποιημένων δεδομένων που βοηθάει στην κατανόηση της κατανομής των δεδομένων. Η ένδειξη της συμμετρίας, της μέσης τιμής και της τιμής του διαμέσου είναι η πρώτη εικόνα που εμφανίζει αν υπάρχει κανονική κατανομή ή όχι. Όταν ένα ιστόγραμμα είναι συμμετρικό και οι τιμές της μέσης τιμής και της διαμέσου είναι παρόμοιες τότε έχουμε κανονική κατανομή (normal distribution). Όταν αυτό δεν συμβαίνει τότε συνήθως χρειάζεται να μετατραπούν τα δεδομένα λογαριθμίζοντας τα ή χρησιμοποιώντας κάποιον άλλο μετασχηματισμό (πχ. μετασχηματισμό Box-Cox[30]).

## 2.3 Συνάρτηση Συνδιασποράς

Σε ένα τυχαίο πεδίο  $X_s$  με δύο τυχαία σημεία  $s_1, s_2$  χρησιμοποιείται η συνάρτηση συνδιασποράς για να περιγράψει την μεταξύ τους συσχέτιση. Η κεντρική συνάρτηση συνδιασποράς ορίζεται από τον παρακάτω τύπο:

$$c_X(s_1, s_2) = E[X(s_1)X(s_2)] - E[X(s_1)]E[X(s_2)] \quad (2.8)$$

Σύμφωνα με αυτή τη σχέση η μέση τιμή του πεδίου συνδιακύμανσης είναι ίση με το μηδέν, άρα προκύπτει και θεωρούμε στην συνέχεια ότι η Συνάρτηση Συνδιασποράς είναι ισοδύναμη με την Συνδιακύμανση.

$$c_X(s_1, s_2) = E[X'(s_1)]E[X'(s_2)] \quad (2.9)$$

Όταν δύο σημεία του τυχαίου πεδίου συμπίπτουν, τότε η τιμή της συνδιασποράς είναι

ίση με τη διασπορά στο σημείο αυτό.

$$c_X(\mathbf{s}_1, \mathbf{s}_2) = \sigma_X^2(\mathbf{s}_1) \quad (2.10)$$

### 2.3.1 Στατιστική Ομοιογένεια

Οι περιορισμοί των ιδιοτήτων ενός πεδίου μπορούν να καταλήξουν πολλές φορές σε αποτελεσματικότερες μεθόδους για το γεωστατιστικό χαρακτηρισμό και την απεικόνισή τους. Μία τέτοια υπόθεση που χρησιμοποιείται συχνά λέγεται στατιστική ομοιογένεια όπου αποτελεί επέκταση του κλασικού ορισμού της ομοιογένειας. Μια ιδιότητα είναι ομοιογενής αν η αντίστοιχη μεταβλητή έχει σταθερή τιμή στο χώρο. Συνεπώς, ένα τυχαίο πεδίο  $X(\mathbf{s})$  είναι στατιστικά ομοιογενές αν η μέση τιμή είναι σταθερή,  $m_x(\mathbf{s}) = m_x$ . Η συνάρτηση συνδιασποράς ορίζεται και εξαρτάται αποκλειστικά από το διάνυσμα της απόστασης  $\mathbf{r} = \mathbf{s}_1 - \mathbf{s}_2$  μεταξύ των δύο σημείων και η διασπορά του πεδίου είναι σταθερή. Ένα τυχαίο πεδίο είναι στατιστικά ομοιογενές χωρίς να αλλάζει η απόσταση μεταξύ των σημείων όταν η πολυδιάστατη Συνάρτηση Πυκνότητας Πιθανότητας για  $N$  σημεία (όπου  $N$  έναν θετικό ακέραιο αριθμό) παραμένει αμετάβλητη από μετασχηματισμούς. Επομένως, η έννοια της στατιστικής ομοιογένειας είναι ότι οι στατιστικές ιδιότητες του πεδίου δεν εξαρτώνται από τις χωρικές συντεταγμένες των σημείων. Από πρακτική άποψη, η στατιστική ομοιογένεια προϋποθέτει ότι δεν υπάρχουν χωρικές τάσεις, επομένως η διακύμανση στις τιμές πεδίου μπορεί να αποδοθεί σε διακυμάνσεις γύρω από ένα σταθερό επίπεδο ίσο με το μέσο όρο. [33, 25]

Η στασιμότητα είναι μια πολύ βολική υπόθεση που απλοποιεί σημαντικά τη δομή των χωρικών μοντέλων. Χρησιμοποιείται συχνά στην μοντελοποίηση των χωρικών μοντέλων όπου όταν η ακρίβεια και η ορθότητα του μοντέλου κρίνονται αποδεκτές, παρουσιάζεται μια προσέγγιση της πραγματικότητας. Όταν οι ιδιότητες ενός μη στάσιμου πεδίου πρέπει να ληφθούν από τα δεδομένα, υπάρχουν αβέβαια αποτελέσματα στο μοντέλο. Επιπλέον, τα μη στάσιμα μοντέλα θα πρέπει να χρησιμοποιούνται με προσοχή, καθώς η εισαγωγή της μη στασιμότητας σε λάθος στοιχείο του χωρικού μοντέλου μπορεί να οδηγήσει σε λάθη. Συνήθως η μη στασιμότητα προκαλείται από την παρουσία τάσεων.



### 2.3.2 Στατιστική Ισοτροπία

Μία ακόμη ιδιότητα στην γεωστατιστική ανάλυση ενός τυχαίου πεδίου είναι η Στατιστική Ισοτροπία. Λέμε ότι ένα πεδίο είναι στατιστικά ισοτροπικό όταν είναι στατιστικά ομοιογενές και η συνάρτηση συνδιασποράς εξαρτάται μόνο από το μέτρο  $r$  και όχι από την κατεύθυνση του διανύσματος απόστασης  $\mathbf{r}$ . Άρα όταν μια συνάρτηση συνδιασποράς είναι στατιστικά ισοτροπική σημαίνει πως είναι και στατιστικά ομοιογενείς, ενώ το αντίστροφο δεν ισχύει.

Αντίθετα, με τον όρο ανισοτροπία θεωρείται ότι η χωρική μεταβλητότητα εξαρτάται από την κατεύθυνση. Στην περίπτωση της συνδιασποράς ενός ανισοτροπικού τυχαίου πεδίου εξαρτάται και από την απόσταση  $r$  και από την κατεύθυνση του διανύσματος  $\mathbf{r}$ . [25]

Υπάρχουν ορισμένοι παράγοντες που καθορίζουν τα βασικά χαρακτηριστικά μιας συνάρτησης συνδιασποράς. Στην περίπτωση στατιστικών ισοτροπικών πεδίων οι πιο σημαντικές παράμετροι είναι η διασπορά  $\sigma_X^2 = c_X(0)$  και το μήκος συσχέτισης  $\xi$ . Όταν δηλαδή η τιμή του πεδίου σε ένα σημείο επηρεάζεται από την τιμή ενός άλλου σημείου, η διασπορά μετράει το πλάτος των μεταβολών των πεδίων, ενώ το μήκος συσχέτισης καθορίζει την απόσταση μεταξύ των σημείων αυτών. Αντίθετα, στην περίπτωση ανισοτροπικής εξάρτησης υπάρχουν διαφορετικά μήκη συσχέτισης για κάθε κύρια διεύθυνση ανισοτροπίας.

### 2.3.3 Μοντέλα Συνδιασποράς

Όταν έχουμε στασιμότητα και ισοτροπία, τα Μοντέλα Συνδιασποράς εξαρτώνται μόνο από την απόσταση και συγκεκριμένα από το μέτρο της Ευκλείδειας απόστασης  $\mathbf{r}$ . Ακολουθούν εξισώσεις για μερικά από τα πιο κοινά μοντέλα συνδιασποράς που χρησιμοποιούνται στις γεωστατιστικές αναλύσεις, με  $\sigma_X^2$  να αντιστοιχεί στην διασπορά,  $\xi$  το μήκος συσχέτισης και το  $\mathbf{r}$  η απόσταση. [33]

#### 1. Εκθετικό Μοντέλο:

$$c_x(\mathbf{r}) = \sigma_X^2 \exp\left(-\frac{\|\mathbf{r}\|}{\xi}\right) \quad (2.11)$$

2. Γκαουσσισιανό Μοντέλο:

$$c_x(\mathbf{r}) = \sigma_X^2 \exp\left(-\frac{\mathbf{r}^2}{\xi^2}\right) \quad (2.12)$$

3. Σφαιρικό Μοντέλο:

$$c_x(\mathbf{r}) = \sigma_X^2 \left[ 1 - 1.5 \left( \frac{\|\mathbf{r}\|}{\xi} \right) + 0.5 \left( \frac{\|\mathbf{r}\|}{\xi} \right)^3 \right] \theta(\xi - \|\mathbf{r}\|) \quad (2.13)$$

4. Γενικευμένο Εκθετικό Μοντέλο:

$$c_x(\mathbf{r}) = \sigma_X^2 \exp\left(-\frac{\|\mathbf{r}\|^\nu}{\xi^\nu}\right), 0 < \nu < 2 \quad (2.14)$$

5. Matérn:

$$c_x(\mathbf{r}) = \frac{\sigma_X^2}{2^{\nu-1}\Gamma(\nu)} \left( \frac{2\sqrt{\nu}}{R} \|\mathbf{r}\| \right)^\nu K_\nu \left( \frac{2\sqrt{\nu}}{R} \|\mathbf{r}\| \right) \quad (2.15)$$

6. Φαινόμενο πυρήνα ή κόκκου (nugget effect):

$$c_x(\mathbf{r}) = \begin{cases} 0, & \text{αν } \|\mathbf{r}\| \neq 0 \\ c_0, & \text{αν } \|\mathbf{r}\| = 0 \end{cases} \quad (2.16)$$

7. Συνάρτηση με οπή αντισυσχέτισης:

$$c_x(\mathbf{r}) = \sigma_X^2 \frac{\sin(\|\mathbf{r}\|/\xi)}{\|\mathbf{r}\|/\xi} \quad (2.17)$$

8. Δυναμονομικό - αλγεβρικό:

$$\lim_{\|\mathbf{r}\| \rightarrow \infty} c_x(\mathbf{r}) \approx \|\mathbf{r}\|^{2\alpha}, -1 < \alpha < 0 \quad (2.18)$$

### 9. Ρητής τετραγωνικής εξάρτησης:

$$c_x(\mathbf{r}) = \frac{\sigma_X^2}{\left[1 + \left(\frac{\|\mathbf{r}\|}{\xi}\right)^2\right]^\beta} \quad (2.19)$$

## 2.4 Χωρική Εξάρτηση

Ο ορισμός της Χωρικής Εξάρτησης ή Χωρικής Παρεμβολής είναι η διαδικασία υπολογισμού μιας τιμής σε ένα τυχαίο πεδίο  $X_s$  σε θέση που δεν έχει πραγματοποιηθεί μέτρηση, χρησιμοποιώντας μετρήσεις που έχουν γίνει σε κοντινά σημεία ή σε σημεία γενικά μέσα στο πεδίο. Η μέθοδος αυτή χρησιμοποιείται για να καλύψει μεγαλύτερη επιφάνεια στην περιοχή μελέτης και να μετατρέψει τα δεδομένα σε ένα σύνολο συνεχείς επιφάνειας. Η οπτικοποίηση αυτών των συνεχών επιφανειών παρουσιάζεται ως κανονικά ή ακανόνιστα πλαίσια (grids) και πολύγωνα. Οι τιμές και οι μεταβολές των τιμών αυτών διακρίνονται από ζώνες διαφορετικών χρωμάτων. Υπάρχουν δύο μέθοδοι για την διαδικασία της χωρικής παρεμβολής. Η Γενική (Global Interpolation) και η Τοπική Μέθοδος (Local Interpolation). Στην πρώτη η πρόβλεψη της τιμής μιας μεταβλητής ορίζεται από το σύνολο όλων των δεδομένων στο χώρο δειγματοληψίας. Αντίθετα, στην δεύτερη μέθοδο η τιμή ενός σημείου καθορίζεται από περιορισμένο αριθμό γειτονικών σημείων προς αυτό. Αρχικά, καθορίζεται μια περιοχή (ακτίνα) που θα δηλώνει το εύρος επιρροής του σημείου. Αυτή η ακτίνα ονομάζεται Ακτίνα Συσχέτισης, όπου κοντινά σημεία επηρεάζονται περισσότερο και υπάρχει μεγάλη συσχέτιση, ενώ όταν η απόσταση είναι μεγάλη, υπάρχει μικρή συσχέτιση ή και καμία συσχέτιση. Έπειτα αν ο αριθμός των περιβάλλον σημείων είναι ικανοποιητικός, επιλέγεται η κατάλληλη μαθηματική συνάρτηση που εκφράζει καλύτερα την μεταξύ τους μεταβολή. Οι πιο διαδεδομένες αιτιοκρατικές μέθοδοι που ανήκουν σε αυτή την κατηγορία είναι οι Nearest Neighbor, πολύγωνα Voronoi, Inverse Distance Weighting, Splines [28].

### 2.4.1 Βαριόγραμμα

Επειδή οι πληροφορίες που λαμβάνονται από την περιοχή μελέτης είναι διάσπαρτες στο χώρο, είναι απαραίτητο ένα μοντέλο να παρουσιάζει συμπεράσματα και στις άγνωστες περιοχές σε ένα πεδίο μελέτης. Για αυτό το *βαριόγραμμα* θεωρείται αρχική βάση σε μια γεωστατιστική μελέτη. Το βαριόγραμμα δείχνει πόσο είναι το εύρος των τιμών των δεδομένων σε σχέση με την απόστασή τους. Υπολογίζεται σύμφωνα με το τετράγωνο των διαφορών που βρίσκονται σε κάποιες αποστάσεις με τον τύπο  $\delta X(\mathbf{s}; \mathbf{r}) = X(\mathbf{s} + \mathbf{r}) - X(\mathbf{s})$ . Το ημιβαριόγραμμα ενός τυχαίου πεδίου ορίζεται σύμφωνα με την εξής συνάρτηση:

$$\gamma(\mathbf{s}, \mathbf{r}) = \frac{1}{2} E[\{X(\mathbf{s} + \mathbf{r}) - X(\mathbf{s})\}^2] \quad (2.20)$$

Στη πράξη όμως για τον υπολογισμό του *πειραματικού ημιβαριογράμματος* (Experimental Semivariogram) χρησιμοποιείται η παρακάτω εξίσωση:

$$\gamma(\mathbf{s}, \mathbf{r}) = \frac{1}{2N} \sum_{i=1}^N [\{X(\mathbf{s}_i + \mathbf{r}) - X(\mathbf{s}_i)\}^2] \quad (2.21)$$

όπου

- $\gamma(\mathbf{s}, \mathbf{r})$  είναι η τιμή του ημιβαριογράμματος
- $\mathbf{r}$  είναι η απόσταση μεταξύ δύο σημείων γεωτρήσεων
- $N$  είναι ο αριθμός των ζευγαριών σημείων
- $X(\mathbf{s}_i)$  είναι η τιμή του αρχικού σημείου από το ζευγάρι
- $X(\mathbf{s}_i + \mathbf{r})$  είναι η τιμή του τελικού σημείου από το ζευγάρι

Στόχος του ημιβαριογράμματος είναι να εξάγει την ημιδιασπορά όπου είναι απαραίτητη πληροφορία για στοχαστικές μεθόδους χωρικής ανάλυσης. Όταν το πεδίο  $X(\mathbf{s})$  είναι στατιστικά ομοιογενές, το βαριόγραμμα συνδέεται με την συνάρτηση συνδιασποράς βάσει της παρακάτω εξίσωσης.

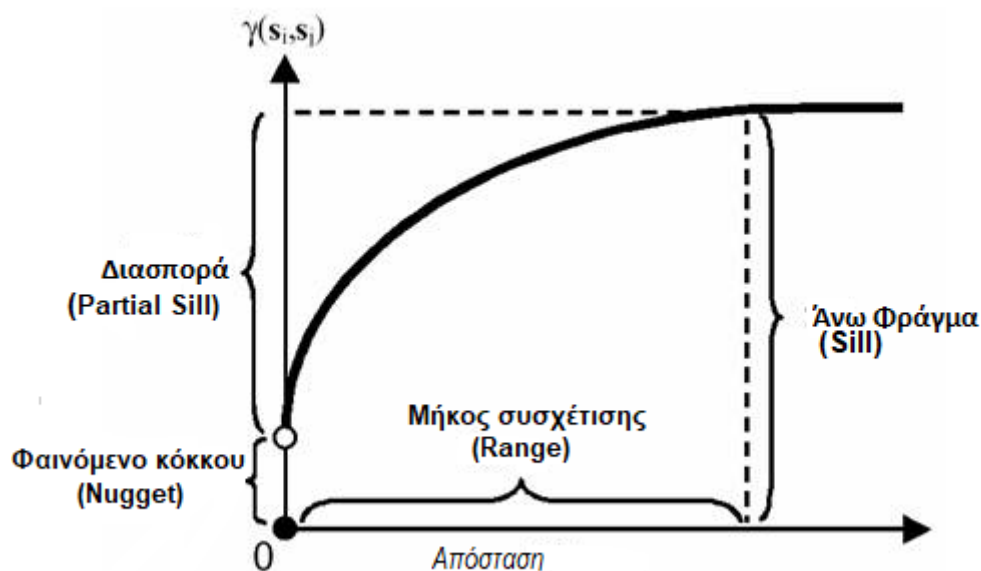
$$\gamma_X(\mathbf{r}) = \sigma_X^2 - \sigma_X(\mathbf{r}) \quad (2.22)$$

Συνεπώς, σύμφωνα με την παραπάνω σχέση το βαριόγραμμα τείνει προς την τιμή της διασποράς και περιέχει τις ίδιες πληροφορίες με την συνάρτηση της συνδιασποράς. [33]

### 2.4.2 Ιδιότητες Βαριογράμματος

Όταν ένα Πεδίο  $X(s)$  είναι στατιστικά ομοιογενές τότε η διαφορά  $\delta X(s; \mathbf{r})$  είναι επίσης στατιστικά ομοιογενής και ορίζεται πεδίο με στατιστικά ομοιογενείς διαφορές. Το βαριόγραμμα  $\gamma(s, \mathbf{r})$  σε μια τέτοια κατάσταση εξαρτάται μόνο από την απόσταση  $\mathbf{r}$  μεταξύ των σημείων. Ωστόσο το αντίστροφο δεν ισχύει. Δηλαδή, αν το πεδίο διαφοράς είναι στατιστικά ομοιογενές δεν είναι αναγκαστικά και το αρχικό πεδίο  $X(s)$  ομοιογενές.[33]

Το γράφημα του βαριογράμματος  $\gamma(s, \mathbf{r})$  είναι συνήθως μια θετική συνάρτηση, δηλαδή  $\gamma(s, \mathbf{r}) \geq 0$ . Οι παράμετροι ενός βαριογράμματος περιγράφουν την χωρική εξάρτηση των τιμών ενός πεδίου σε δύο γειτονικά σημεία. Μία μορφή ενός ημιβαριογράμματος παρουσιάζεται στη παρακάτω εικόνα, όπου διακρίνονται και οι παράμετροι του, άνω φράγμα (Sill), μήκος ή ακτίνα συσχέτισης (Range) και φαινόμενο κόκκου ή φαινόμενο πυρήνα (Nugget).



Εικόνα 2.2: Παράδειγμα Βαριογράμματος

Σύμφωνα με τη στασιμότητα του πεδίου, δηλαδή αν είναι στατιστικά ομοιογενές ή στατι-

στικά ανομοιογενές, το βαριόγραμμα υπολογίζεται με διαφορετικές παραμέτρους.

Το βαριόγραμμα σε ομοιογενές πεδίο με ισοτροπική χωρική εξάρτηση καθορίζεται από τις παραμέτρους άνω φράγμα (Sill) και το μήκος συσχέτισης (Range). Το μήκος συσχέτισης περιγράφει το πόσο γρήγορα το βαριόγραμμα πλησιάζει το άνω φράγμα. Πρακτικά, σημαίνει πως προσδιορίζει την ακτίνα όπου δύο σημεία θεωρούνται συσχετισμένα. Για πολύ μεγάλες αποστάσεις  $\mathbf{r}$  η τιμή του βαριογράμματος τείνει ασυμπτωματικά προς ένα άνω φράγμα που ισούται με την διασπορά του τυχαίου πεδίου  $\sigma_X^2$ . Δηλαδή, σε τέτοιες περιπτώσεις το φαινόμενο κόκκου τείνει στο 0 και η διασπορά ισούται με το άνω φράγμα. Αυτή η ιδιότητα βασίζεται στη σχέση  $\gamma_X(\mathbf{r}) = \sigma_X^2 - c_X(\mathbf{r})$  και στο ότι η τιμή της συνάρτησης συνδιασποράς σε μεγάλες αποστάσεις τείνει στο μηδέν.[33]

Όταν το βαριόγραμμα είναι ανισοτροπικό, το άνω φράγμα κάποιες φορές είναι ανεξάρτητο της κατεύθυνσης και το μήκος συσχέτισης εξαρτάται από τη διεύθυνση. Σε αυτή την περίπτωση, το βαριόγραμμα εκφράζεται σαν συνάρτηση  $\gamma_X\left(\frac{r_1}{\xi_1}, \dots, \frac{r_d}{\xi_d}\right)$  των αδιάστατων αποστάσεων  $\left(\frac{r_1}{\xi_1}, \dots, \frac{r_d}{\xi_d}\right)$  όπου τα μήκη  $\xi_1, \dots, \xi_d$  είναι τα μήκη συσχέτισης στις αντίστοιχες διευθύνσεις.

Στην περίπτωση της ανισοτροπίας ζώνης, το άνω φράγμα εξαρτάται από την χωρική κατεύθυνση, και το βαριόγραμμα εκφράζεται με την παρακάτω σχέση.

$$\gamma_X(\mathbf{r}) = \gamma_{X,1}(r) + \gamma_{X,2}(\hat{r}) \quad (2.23)$$

Η συνάρτηση  $\gamma_{X,1}(r)$  όπου  $r = \|\mathbf{r}\|$  είναι το μέτρο του διανύσματος απόστασης, αντιστοιχεί στην ισοτροπική εξάρτηση. Αντίθετα η συνάρτηση  $\gamma_{X,2}(\hat{r})$  αντιστοιχεί στην ανισοτροπική εξάρτηση του άνω φράγματος από την κατεύθυνση του μοναδιαίου διανύσματος  $\hat{r}$ .

Το Nugget συμβολίζεται συνήθως ως  $C_o$ , το partial Sill ως  $C_1$  ή  $\sigma^2$  και το Sill ως το άθροισμα αυτών  $C = C_o + C_1$ . Το nugget μπορεί να οφείλεται σε σφάλματα μέτρησης ή σε χωρικές πηγές διακύμανσης σε αποστάσεις μικρότερες από το διάστημα δειγματοληψίας ή και στα δύο. Το Range δείχνει την απόσταση όπου το ημιβαριόγραμμα φτάνει το ηψηλότερο σημείο (ταβάνι). Συμβολίζει τότε τα δεδομένα είναι συσχετισμένα και τότε ασυσχέτιστα[10]. Όταν το φαινόμενο Nugget είναι πολύ μεγάλο σε σχέση με το partial Sill σημαίνει πως υπάρχει πολύς θόρυβος στα δεδομένα και είναι συνετό να μην συνεχιστεί η ανάλυση, διότι

## Chapter 2. Βασικές Θεωρητικές Έννοιες

το αποτέλεσμα μιας τέτοιας ανάλυσης θα ήταν αναξιόπιστα και ίσως λανθασμένα.

Στην περίπτωση της στασιμότητας, το βαριόγραμμα έχει ορισμένες χαρακτηριστικές ιδιότητες [15]:

- Η τιμή του βαριογράμματος είναι μηδέν στο  $\|\mathbf{r}\| = 0$
- Η παρουσία Nugget είναι πάντα τιμή κοντινή στο μηδέν
- Το βαριόγραμμα τείνει να αυξάνεται καθώς αυξάνεται το  $\|\mathbf{r}\|$  αυξάνεται. Η αύξηση μπορεί να είναι μη μονοτονική για μοντέλα με ταλαντευόμενη εξάρτηση των συσχετίσεων
- Ο ρυθμός αύξησης του βαριογράμματος καθορίζεται από ένα χαρακτηριστικό μήκος, ενδεχομένως σε συνδυασμό και με άλλες παραμέτρους
- Το βαριόγραμμα τείνει σε ένα σταθερό Sill καθώς το  $\|\mathbf{r}\|$  τείνει το άπειρο.
- Ο ρυθμός με τον οποίον προσεγγίζεται το Sill αλλά και η τιμή του, εξαρτάται από την διεύθυνση του  $\mathbf{r}$  εάν υπάρχει ανισοτροπία

Τόσο το βαριόγραμμα όσο και η συνάρτηση συνδιακύμανσης περιγράφουν συσχετίσεις δύο σημείων και σχετίζονται μεταξύ τους με την παρακάτω σχέση.

$$\gamma_X(\mathbf{s}_1 - \mathbf{s}_2) = \sigma_X^2 - c_X(\mathbf{s}_1 - \mathbf{s}_2) \quad (2.24)$$

Και οι δύο συναρτήσεις περιέχουν την ίδια πληροφορία στην περίπτωση στάσιμων τυχαίων πεδίων. Ωστόσο, υπάρχουν σημαντικές διαφορές μεταξύ του βαριογράμματος και της συνδιασποράς ιδιαίτερα σε μη στάσιμα πεδία. Οι διαφορές αυτές είναι [15]:

- Εάν η συνδιασπορά είναι αμετάβλητη μετά από κάποιον μετασχηματισμό το βαριόγραμμα είναι επίσης αμετάβλητο σύμφωνα με την σχέση 2.24 αλλά το αντίστροφο δεν ισχύει πάντα.
- Σε μη στάσιμα τυχαία πεδία με στάσιμες προσαυξήσεις, το βαριόγραμμα εξαρτάται μόνο από την απόσταση των σημείων του ζεύγους, ενώ η συνάρτηση συνδιασποράς εξαρτάται από τις θέσεις και των δύο σημείων.

- Η συνάρτηση του βαριογράμματος μπορεί να είναι καλά ορισμένη για τυχαία πεδία με άπειρη διακύμανση, σε αντίθεση με τη συνδιασπορά.
- Η εκτίμηση του βαριογράμματος δεν απαιτεί τη γνώση της μέσης τιμής με αποτέλεσμα να πλεονεκτεί έναντι της συνδιασποράς σε σχέση με την εκτίμηση.

### 2.4.3 Μοντέλα Βαριογράμματος

Οι εξισώσεις που ακολουθούν χρησιμοποιούνται περισσότερο για να περιγράψουν τα μοντέλα βαριογράμματος. Το  $\xi$  είναι το μήκος συσχέτισης, το  $\sigma_X^2$  αντιστοιχεί στην διασπορά και το  $r$  είναι η απόσταση. [33]

#### 1. Εκθετικό Μοντέλο:

$$\gamma_X(r) = \sigma_X^2 [1 - \exp(-\|r\|/\xi)] \quad (2.25)$$

#### 2. Γκαουσσισιανό Μοντέλο:

$$\gamma_X(r) = \sigma_X^2 [1 - \exp(-\|r\|^2/\xi^2)] \quad (2.26)$$

#### 3. Σφαιρικό Μοντέλο:

$$\gamma_X(r) = \begin{cases} \sigma_X^2 \left[ 1.5 \left( \frac{\|r\|}{\xi} \right) - 0.5 \left( \frac{\|r\|}{\xi} \right)^3 \right], & \|r\| \leq \xi \\ \sigma_X^2 & \|r\| \geq \xi \end{cases} \quad (2.27)$$

#### 4. Γενικευμένο Εκθετικό Μοντέλο:

$$\gamma_X(\|r\|) = \sigma_X^2 [1 - \exp(-\|r\|^\nu/\xi^\nu)], 0 < \nu < 2 \quad (2.28)$$

#### 5. Δυναμονομικό (αλγεβρικό):

$$\gamma_X(\|r\|) = \alpha \|r\|^{2H}, 0 < H < 1, \alpha > 0 \quad (2.29)$$



6. Γραμμικό:

$$\gamma_X(\|\mathbf{r}\|) = \alpha\|\mathbf{r}\|, \alpha > 0 \quad (2.30)$$

7. Φαινόμενο Πυρήνα (nugget effect):

$$\gamma_X(r) = \begin{cases} 0, & r = 0 \\ c_0, & r \neq 0 \end{cases} \quad (2.31)$$

8. Κυματοειδές:

$$\gamma_X(\|\mathbf{r}\|) = \sigma_X^2[1 - \xi\|\mathbf{r}\|^{-1}\sin(\|\mathbf{r}\|/\xi)] \quad (2.32)$$

9. Βαριόγραμμα Matérn:

$$\gamma_X(\|\mathbf{r}\|) = \sigma_X^2 \left[ 1 - \frac{1}{2^{\nu-1}\Gamma(\nu)} \left( \frac{2\sqrt{\nu}}{R} \|\mathbf{r}\| \right)^{\nu} K_{\nu} \left( \frac{2\sqrt{\nu}}{R} \|\mathbf{r}\| \right) \right] \quad (2.33)$$

## 2.5 Εκτίμηση Τυχαίου Πεδίου (Μέθοδος Kriging)

Οι γραμμικές μέθοδοι εκτίμησης χρησιμοποιούν γραμμικούς συνδυασμούς για να προσδιορίσουν τα γειτονικά σημεία. Η εκτίμηση ενός τυχαίου πεδίου προϋποθέτει την ύπαρξη ενός μοντέλου χωρικής εξάρτησης καθώς πρέπει η τιμή να επηρεάζεται από τις γειτονικές τιμές του πεδίου. Έτσι θα υπάρχουν αποτελέσματα και σε σημεία που δεν υπήρχε μέτρηση. Οι πιο διαδεδομένες μέθοδοι βασίζονται στην γραμμική παρεμβολή και στην ελαχιστοποίηση του τετραγωνικού σφάλματος της εκτίμησης. Ο συνδυασμός αυτών των μεθόδων είναι γνωστός με την ονομασία Kriging. Όταν η τάση αφαιρεθεί από τις διακυμάνσεις αλλά εφαρμοστεί η μέθοδος kriging στις διακυμάνσεις τότε λέγεται πως χρησιμοποιείται Kriging με παλινδρόμηση (Regression kriging).[33, 25] Η παλινδρόμηση είναι μια στατιστική τεχνική μοντελοποίησης που ερευνά την συσχέτιση μίας εξαρτώμενης μεταβλητής και μιας ή περισσότερων ανεξάρτητων μεταβλητών.[14]

Στις μεθόδους kriging απαραίτητο είναι να προσδιορίζεται μια περιοχή όπου αποκαλείται γειτονιά συσχέτισης  $\omega(\mathbf{u})$  ενός σημείου  $\mathbf{u}$ . Περιλαμβάνει  $n(\mathbf{u}) \leq N$  σημεία από τα  $\mathbf{s}_i (i =$

$1, \dots, N$ ). Το μέγεθος της περιοχής αυτής ορίζεται από το μήκος (ή τα μήκη) συσχέτισης. Η διακύμανση του πεδίου στο σημείο εκτίμησης εκφράζεται με τον παρακάτω τύπο, όπου τα  $\lambda_\alpha$  αντιπροσωπεύουν τα γραμμικά βάρη. Δηλαδή, εκφράζεται η διακύμανση στο σημείο εκτίμησης σαν συνάρτηση της διακύμανσης στα υπόλοιπα σημεία εκτίμησης.

$$\hat{X}(\mathbf{u}) - m_X(\mathbf{u}) = \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_\alpha [X(\mathbf{s}_\alpha) - m_X(\mathbf{s}_\alpha)] \quad (2.34)$$

Ενώ η εκτιμώμενη τιμή του πεδίου δίνεται από την επόμενη εξίσωση. Η  $\hat{X}(\mathbf{u})$  και το σφάλμα εκτίμησης  $\varepsilon(\mathbf{u})$  είναι τυχαίες μεταβλητές διότι αποτελούνται από τιμές γραμμικού συνδυασμού ενός τυχαίου πεδίου.

$$\hat{X}(\mathbf{u}) = m_X(\mathbf{u}) + \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_\alpha [X(\mathbf{s}_\alpha) - m_X(\mathbf{s}_\alpha)] \quad (2.35)$$

Το σφάλμα εκτίμησης ορίζεται ως η διαφορά ανάμεσα στην πραγματική και την εκτιμώμενη τιμή του πεδίου.

$$\varepsilon(\mathbf{u}) = X(\mathbf{u}) - \hat{X}(\mathbf{u}). \quad (2.36)$$

Καθώς στη μέθοδο kriging χρησιμοποιούνται τα βάρη για να ελαχιστοποιηθεί η διασπορά του σφάλματος της εκτίμησης, υπολογίζεται η βέλτιστη τιμή  $\hat{X}(\mathbf{u})$ .

Σύμφωνα με την μέση τιμή και το πως επιλέγετε να χρησιμοποιηθεί η τάση  $m_X(\mathbf{s})$  και η διακύμανση στην ανάλυση του πεδίου, εφαρμόζονται διαφορετικά είδη kriging:

- Το Απλό Kriging (Simple kriging) χρησιμοποιείται όταν η μέση τιμή του πεδίου είναι γνωστή σε όλη την περιοχή έρευνας.
- Το Κανονικό Kriging (Ordinary Kriging) χρησιμοποιείται όταν η μέση τιμή του πεδίου είναι σταθερή μόνο στη γειτονιά συσχέτισης  $\omega(\mathbf{u})$  του σημείου εκτίμησης, ενώ η τιμή της  $m_X(\mathbf{s})$  μπορεί να αλλάζει από γειτονιά σε γειτονιά.
- Το Καθολικό Kriging (Universal Kriging) χρησιμοποιείται όταν η μέση τιμή αλλάζει μέσα στη γειτονιά του σημείου εκτίμησης.

- Το Kriging σε πεδίο δεικτών (Indicator Kriging) χρησιμοποιείται όταν η συνάρτηση κατανομής είναι ασύμμετρη με τις περισσότερες τιμές να είναι πολύ υψηλές. Αυτή η μέθοδος υπολογίζει πως η τοπική τιμή υπερβαίνει ένα καθορισμένο όριο, έτσι υπολογίζει την αθροιστική κατανομή πιθανότητας.
- Το Συνδυαστικό Kriging (Co-Kriging) εφαρμόζεται για να συμπεριλάβει ιδιότητες στην εκτίμηση που σχετίζονται με ιδιότητες που επιδιώκονται να εκτιμηθούν. [33]
- Kriging με παλινδρόμηση (Regression kriging) είναι μια τεχνική χωρικής παρεμβολής που συνδυάζει μια παλινδρόμηση της εξαρτημένης μεταβλητής σε βοηθητικές μεταβλητές (όπως παράμετροι της επιφάνειας του εδάφους, εικόνες τηλεπισκόπησης και θεματικούς χάρτες) με απλό kriging. Είναι μαθηματικά ισοδύναμη με τη μέθοδο παρεμβολής Universal Kriging. Το πλεονέκτημα της είναι η δυνατότητα επέκτασης της μεθόδου σε ένα ευρύτερο φάσμα τεχνικών παλινδρόμησης.[14]

### 2.5.1 Απλό Kriging

Το απλό kriging αναφέρεται μερικές φορές και ως «kriging με γνωστή μέση τιμή»[9]. Όταν στη μέθοδο kriging η μέση τιμή του πεδίου είναι γνωστή, όπως προϋποθέτει η χρήση του απλού kriging, αλλά και σταθερή  $E[X(\mathbf{u})] = m_X$ , η εκτίμηση προέρχεται από την παρακάτω εξίσωση. Με  $\lambda_\alpha$  συμβολίζονται οι γραμμικοί συντελεστές (βάρη) και υπολογίζονται χρησιμοποιώντας τη συνθήκη ελαχιστοποίησης του τετραγωνικού σφάλματος.

$$\hat{X}(\mathbf{u}) = \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_\alpha X(\mathbf{s}_\alpha) - m_X \left[ \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_\alpha - 1 \right] \quad (2.37)$$

Η τιμή του  $\left[ \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_\alpha - 1 \right]$  δεν είναι απαραίτητα 0. Αντιθέτως, όσο μεγαλύτερη είναι η τιμή τόσο η εκτίμηση πλησιάζει στην μέση τιμή. Γενικά, η τιμή του  $\left[ \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_\alpha - 1 \right]$  αυξάνεται σε περιοχές με σχετικά φτωχή δειγματοληψία.

Το απλό kriging υποθέτει ότι υπάρχει στασιμότητα δευτέρου βαθμού, δηλαδή σταθερή μέση τιμή, διακύμανση και συνδιακύμανση στον τομέα ή την περιοχή ενδιαφέροντος. Επειδή

μια τέτοια υπόθεση είναι συχνά πολύ περιοριστική, το κανονικό kriging χρησιμοποιείται συχνότερα.[9]

Η εξίσωση που περιγράφει τα γραμμικά βάρη μπορεί να εκφραστεί και ως εξίσωση πινάκων,

$$\mathbf{C}_{\alpha,\beta} \lambda_\beta = \mathbf{C}_{\alpha,u} \quad (2.38)$$

όπου ο πίνακας  $\mathbf{C}_{\alpha,\beta}$  αντιπροσωπεύει τον πίνακα συνδιασποράς  $\mathbf{C}_{\alpha,\beta} = c_X(\mathbf{s}_\alpha - \mathbf{s}_\beta)$ . Το διάνυσμα  $\mathbf{C}_{\alpha,u}$  αναπαριστά τις τιμές της συνάρτησης διασποράς των σημείων του δείγματος και της εκτίμησης  $\mathbf{C}_{\alpha,\beta} = c_X(\mathbf{s}_\alpha - \mathbf{u})$ . Επιπλέον, με την εξίσωση  $c_X(0) = \sigma_X^2$ , το γραμμικό σύστημα καταλήγει σε αναλυτική μορφή στον παρακάτω γραμμικό πίνακα,

$$\begin{bmatrix} \sigma_x^2 & \dots & \dots & c_X(\mathbf{s}_1 - \mathbf{s}_n) \\ c_X(\mathbf{s}_2 - \mathbf{s}_1) & \dots & \dots & c_X(\mathbf{s}_2 - \mathbf{s}_n) \\ \vdots & \vdots & \vdots & \vdots \\ c_X(\mathbf{s}_n - \mathbf{s}_1) & \dots & \dots & \sigma_X^2 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{bmatrix} = \begin{bmatrix} c_X(\mathbf{s}_1 - \mathbf{u}) \\ c_X(\mathbf{s}_2 - \mathbf{u}) \\ \vdots \\ c_X(\mathbf{s}_n - \mathbf{u}) \end{bmatrix} \quad (2.39)$$

όπου η λύση του εκφράζεται  $\lambda_\beta = C_{\beta,\alpha}^{-1} C_{\alpha,u}$  για κάθε  $\beta = 1, \dots, n(\mathbf{u})$ . [32]

Τα στοιχεία του πίνακα συνδιασποράς υπολογίζονται από την παρακάτω μαθηματική σχέση.

$$c_X(\mathbf{s}_\alpha, \mathbf{s}_\beta) = \sigma_X^2 - \gamma_X(\mathbf{s}_\alpha, \mathbf{s}_\beta) \quad (2.40)$$

Συμπεραίνεται από τις παραπάνω εξισώσεις πως :

- το γραμμικό σύστημα θα έχει λύσεις για τα  $\lambda_\beta$  μόνο όταν η συνάρτηση συνδιασποράς είναι ημιθετικά ορισμένη και δεν χρησιμοποιούνται άλλες μετρήσεις στο ίδιο σημείο.
- Οι τιμές που παίρνουν τα γραμμικά βάρη είναι ανεξάρτητα από το άνω φράγμα (Sill) του βαριογράμματος. Εξαρτώνται όμως από το πρότυπο συνδιασποράς δηλαδή την μορφή της χωρικής εξάρτησης.
- Η εκτίμηση kriging είναι ακριβής γραμμική μεταβολή, που σημαίνει πως σε κάθε σημείο που υπάρχει μέτρηση, η εκτίμηση είναι ίδια με την τιμή του δείγματος.

Για να μπορέσει να αξιολογηθεί η εκτίμηση που γίνεται χρησιμοποιείται η ακόλουθη

εξίσωση και προσδιορίζεται σύμφωνα με την τετραγωνική ρίζα της διασποράς του σφάλματος εκτίμησης.

$$\sigma_{E,SK}^2(\mathbf{u}) = \sigma_X^2 - \sum_{\alpha=1}^{n(\mathbf{u})} \sum_{\beta=1}^{n(\mathbf{u})} C_{u,\alpha} C_{\alpha,\beta}^{-1} C_{\beta,u} \quad (2.41)$$

Όπως φαίνεται και από την Σχέση 2.41, το σφάλμα αυξάνει αναλογικά προς την τυπική απόκλιση  $\sigma_X$ . Αν τα βάρη  $\lambda_\alpha$  έχουν θετικές τιμές, τότε η συσχέτιση  $\rho_{u,\alpha}$  τείνει στην μονάδα και το σφάλμα αυξάνεται όσο μειώνεται η απόσταση  $|\mathbf{u} - \mathbf{s}_\alpha|$  μεταξύ του σημείου εκτίμησης και των σημείων του δείγματος. [33, 25]

## 2.5.2 Κανονικό Kriging

Η πιο συχνά χρησιμοποιούμενη μέθοδος kriging είναι γνωστή ως Κανονικό kriging (Ordinary kriging), κυρίως γιατί δεν απαιτείται η εκτίμηση της μέσης τιμής του τυχαίου πεδίου. Όταν οι τιμές του δείγματος είναι συσχετισμένες η μέση τιμή δεν είναι απλά ο μέσος όρος των δειγματικών τιμών. Για να υπολογιστεί η μέση τιμή απαιτείται η γνώση της συνάρτησης συνδιασποράς. Αλλά για να υπολογιστεί η συνάρτηση συνδιασποράς πρέπει να είναι γνωστή η μέση τιμή. Χρησιμοποιώντας λοιπόν το κανονικό kriging αποφεύγεται το πρόβλημα αυτό. Στη πράξη, η δειγματοληψία μπορεί να στοχεύει σε περιοχές με χαμηλές ή υψηλές τιμές, οδηγώντας σε μεροληπτικές εκτιμήσεις της μέσης τιμής. [15, 33]

Η βασική εξίσωση του κανονικού kriging και η εκτίμηση του δίνονται παρακάτω, όπου η δεύτερη εξίσωση εκφράζει την συνθήκη αμεροληψίας.

$$\hat{X}(\mathbf{u}) = \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_\alpha X(\mathbf{s}_\alpha) \quad (2.42)$$

$$\sum_{\alpha=1}^{n(\mathbf{u})} \lambda_\alpha = 1 \quad (2.43)$$

Το ελάχιστο μέσο τετραγωνικό σφάλμα πρέπει να υπολογιστεί με τον περιορισμό της εξίσωσης 2.43 στις τιμές των βαρών.

Για τον υπολογισμό των γραμμικών βαρών εφαρμόζονται οι παρακάτω εξισώσεις:

$$\sum_{\beta=1}^{n(\mathbf{u})} \lambda_{\alpha} c_X(\mathbf{s}_{\alpha} - \mathbf{s}_{\beta}) + \mu = c_X(\mathbf{s}_{\alpha} - \mathbf{u}), \alpha = 1, \dots, n(\mathbf{u}) \quad (2.44)$$

$$\sum_{\alpha=1}^{n(\mathbf{u})} \lambda_{\alpha} = 1 \quad (2.45)$$

Ακόμα, το παραπάνω γραμμικό σύστημα μπορεί να παρουσιαστεί και ως γραμμικός πίνακας.

$$\begin{bmatrix} \sigma_X^2 & c_X(\mathbf{s}_1 - \mathbf{s}_2) & \dots & c_X(\mathbf{s}_1 - \mathbf{s}_n) & 1 \\ c_X(\mathbf{s}_2 - \mathbf{s}_1) & \sigma_X^2 & \dots & c_X(\mathbf{s}_2 - \mathbf{s}_n) & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ c_X(\mathbf{s}_n - \mathbf{s}_1) & c_X(\mathbf{s}_n - \mathbf{s}_2) & \dots & \sigma_X^2 & 1 \\ 1 & 1 & \dots & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ \mu \end{bmatrix} = \begin{bmatrix} c_X(\mathbf{s}_1 - \mathbf{u}) \\ c_X(\mathbf{s}_2 - \mathbf{u}) \\ \vdots \\ c_X(\mathbf{s}_n - \mathbf{u}) \\ 1 \end{bmatrix} \quad (2.46)$$

Το μέσο τετραγωνικό σφάλμα για το κανονικό kriging υπολογίζεται από την παρακάτω εξίσωση, με την τιμή του συντελεστή Lagrange, όπου είναι πάντα αρνητική:

$$\sigma_{E,OK}^2(\mathbf{u}) = \sigma_X^2 - \sum_{\alpha=1}^{n(\mathbf{u})} \lambda_{\alpha} c_X(\mathbf{u}) - \mu \quad (2.47)$$

Σαν συμπέρασμα, επειδή στην εκτίμηση του κανονικού kriging δεν συμπεριλαμβάνεται η μέση τιμή, η αβεβαιότητα είναι μεγαλύτερη σε σύγκριση με την εκτίμηση στο απλό kriging. Ακόμα, στο κανονικό kriging και το φαινόμενο κόκκου (nugget) επιδρά στο σφάλμα εκτίμησης. [33]

### 2.5.3 Αξιολόγηση των Εκτιμήσεων

Τα ακόλουθα στατιστικά μέτρα χρησιμοποιούνται συνήθως για την αξιολόγηση των επιδόσεων των προβλέψεων σε μια μελέτη. Υποθέτουμε ότι εκτελείται διασταυρωμένη επαλήθευση

leave-one-out (LOOCV), δηλαδή η μέθοδος της αφαίρεσης ενός σημείου εκ περιτροπής.

### Μέσο Σφάλμα

Το μέσο σφάλμα υπολογίζεται από την παρακάτω σχέση:

$$MAE = \frac{1}{N} \sum_{n=1}^N [\hat{x}(s_i) - x(s_i)] \quad (2.48)$$

Το μέσο σφάλμα υπολογίζει τη μεροληψία της εκτίμησης. Η μεροληψία ορίζεται ως η διαφορά μεταξύ της αναμενόμενης τιμής και της πραγματικής τιμής που μετρήθηκε. Στην περίπτωση που το υπόλοιπο είναι μηδέν θεωρείται ως τέλεια εκτίμηση.[27]

### Μέσο Απόλυτο Σφάλμα

Η εξίσωση που περιγράφει το Μέσο Απόλυτο Σφάλμα (MAE) φαίνεται παρακάτω:

$$MAE = \frac{1}{N} \sum_{n=1}^N |\hat{x}(s_i) - x(s_i)| \quad (2.49)$$

Το MAE ποσοτικοποιεί το μέγεθος (απόλυτη τιμή) των αποκλίσεων μεταξύ των εκτιμήσεων και των πραγματικών τιμών.[15]

### Μέσο Τετραγωνικό Σφάλμα

Το Μέσο Τετραγωνικό Σφάλμα (MSE) είναι το μέτρο το οποίο περιγράφει πόσο κοντά ένα γραμμικό μοντέλο ταιριάζει στα σημεία δεδομένων. Για κάθε σημείο δεδομένων υπολογίζεται η κάθετη απόσταση από το σημείο του άξονα y προς την καμπύλη και τετραγωνίζεται η τιμή. Όσο μεγαλύτερο είναι το MSE, τόσο χειρότερη είναι η προσαρμογή του μοντέλου. Η ρίζα του Μέσου Τετραγωνικού σφάλματος (RMSE) είναι απλώς η τετραγωνική ρίζα του μέσου τετραγωνικού σφάλματος. Είναι μια καλύτερη μέθοδος σύγκρισης από το συντελεστή

συσχέτισης. Υπολογίζεται από την εξίσωση:

$$RMSE = \sqrt{\frac{1}{N} \sum_{n=1}^N [\hat{x}(s_i) - x(s_i)]^2} \quad (2.50)$$

Η εκτίμηση του RMSE θεωρείται καλύτερη από αυτή του MAE καθώς αποφεύγεται η χρήση της απόλυτης τιμής, που είναι ανεπιθύμητη σε πολλούς μαθηματικούς υπολογισμούς. Σε εφαρμογές όπως αυτή που αφορά αυτήν την εργασία, τα μεγάλα σφάλματα των ελαχίστων τετραγώνων αποδεικνύεται ότι είναι πιο αποτελεσματικά για την βελτίωση της απόδοσης του μοντέλου.[5] Μια βασική διαφορά μεταξύ του RMSE και MAE είναι πως το πρώτο δίνει βαρύτητα στις ακραίες τιμές και στα μεγάλα σφάλματα, ενώ το MAE δίνει την ίδια βαρύτητα σε όλα τα σφάλματα.

### Συντελεστής Συσχέτισης του Pearson

Ο συντελεστής συσχέτισης χρησιμοποιείται πιο συχνά για την εκτίμηση μεταξύ δύο μεταβλητών και υπολογίζεται από την παρακάτω σχέση.

$$\rho = \frac{\sum_{n=1}^N (x_n^* - \bar{x}_n^*)(\hat{x}_n - \bar{\hat{x}}_n)}{\sqrt{\sum_{n=1}^N (x_n^* - \bar{x}_n^*)^2} \sqrt{\sum_{n=1}^N (\hat{x}_n - \bar{\hat{x}}_n)^2}} \quad (2.51)$$

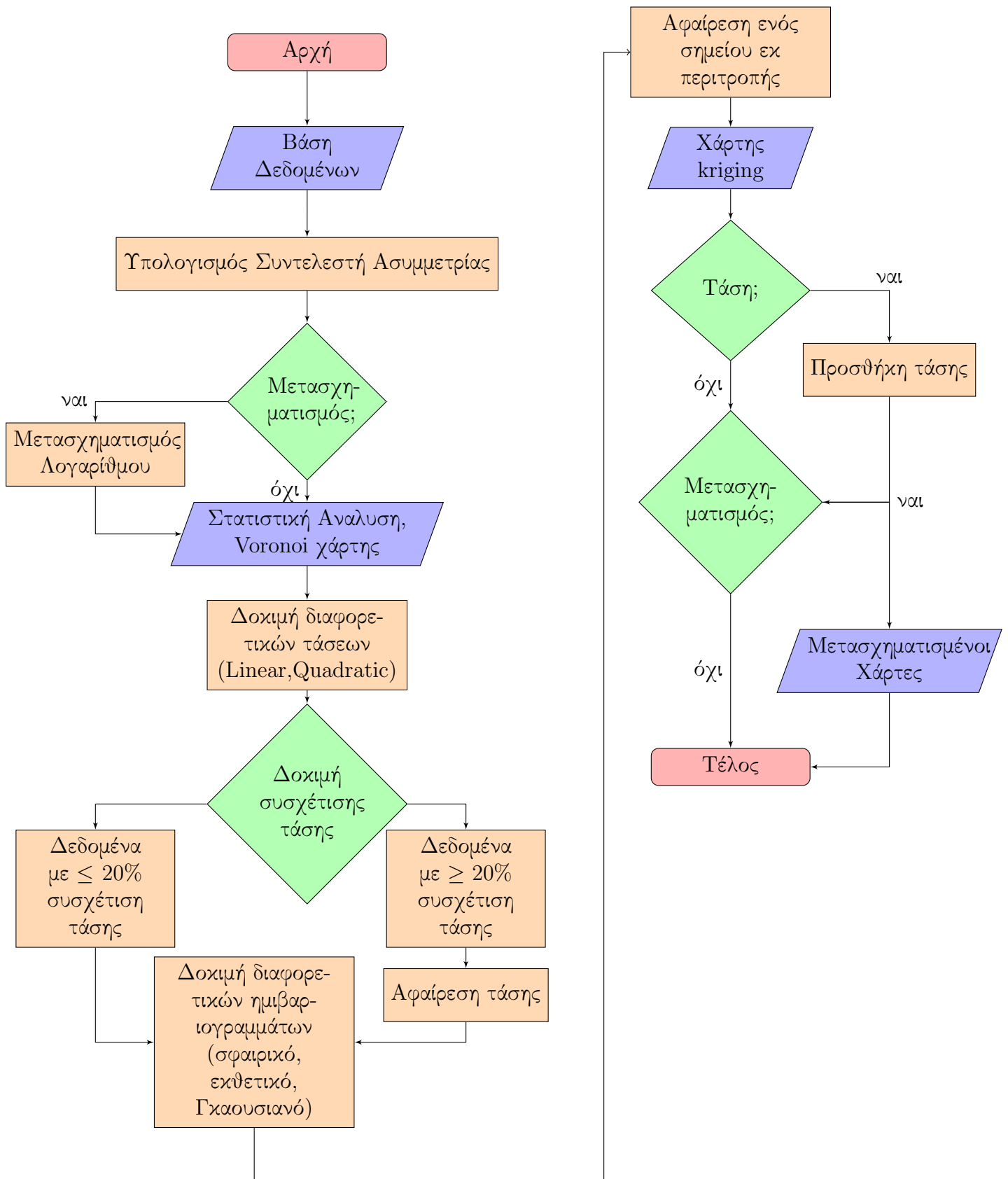
Η τιμή  $\rho$  μετρά τη στατιστική συσχέτιση ανάμεσα στα δεδομένα και τις εκτιμήσεις. Από την σχέση 2.51 ισχύει πως  $-1 \leq \rho \leq 1$ . Ο συντελεστής συσχέτισης  $\rho$  δίνει καλύτερα αποτελέσματα σε γραμμικές συσχετίσεις, ενώ αντίθετα σε μη γραμμικές σχέσεις δεν είναι αξιόπιστο. Η εξίσωση αυτή θα μπορούσε να παρουσιαστεί σε ένα διάγραμμα διασποράς. Με  $\rho = 1$  θα απεικονιζόταν μια ευθεία γραμμή με θετική κλίση, ενώ αν το  $\rho = -1$  θα ήταν μια ευθεία γραμμή με αρνητική κλίση. [25]



### 3 Γεωστατιστική Ανάλυση Δεδομένων

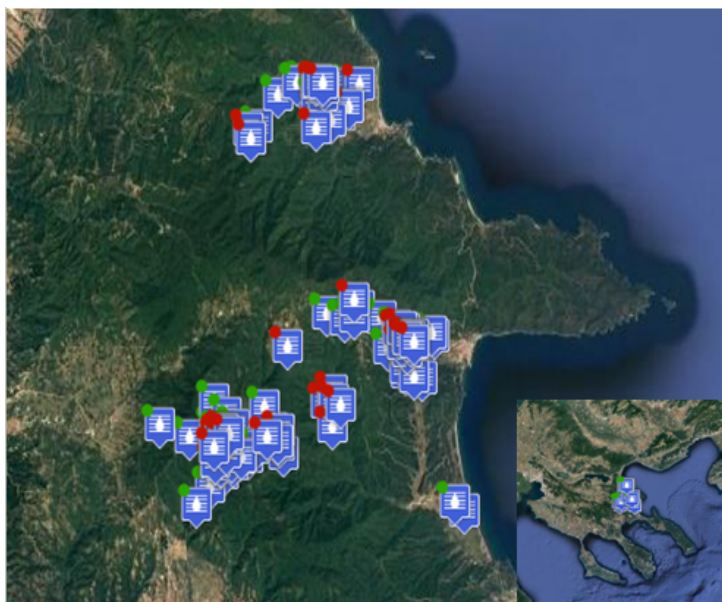
Σε αυτό το κεφάλαιο παρουσιάζεται η γεωστατιστική ανάλυση των δεδομένων. Τα αρχεία είναι πραγματικές διερευνητικές γεωτρήσεις του Ελληνικού Χρυσού, στα οποία είχαμε πρόσβαση μέσω της ιστοσελίδας τους με το πρόγραμμα περιβαλλοντικής παρακολούθησης[11]. Η ανάλυση των δεδομένων, κατά την εκπόνηση της παρούσας διπλωματικής, έγινε με τη χρήση της πλατφόρμας προγραμματισμού Matlab. Η ανάλυση χωρίστηκε σε τέσσερα βασικά στάδια. Τη στατιστική ανάλυση, την εύρεση και αφαίρεση τάσης, την βαριογραφία και τέλος την εφαρμογή της μεθόδου Ordinary Kriging (OK). Για μια ολοκληρωμένη εικόνα με τα βήματα της διαδικασίας που πραγματοποιήθηκαν στην συγκεκριμένη εργασία παρουσιάζεται το διάγραμμα ροής 3.1. Σε αυτό το κεφάλαιο εξηγούνται καθένα από τα βήματα αυτά.

Εικόνα 3.1: Διάγραμμα ροής



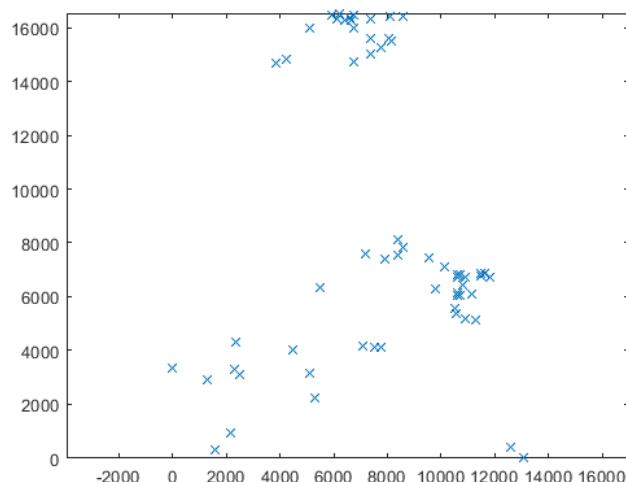
### 3.1 Περιγραφή των δεδομένων

Τα δεδομένα στα οποία αναφέρεται η συγκεκριμένη εργασία προέρχονται από τη περιβαλλοντική παρακολούθηση της εταιρίας «Ελληνικού Χρυσού» στη Χαλκιδική. Η περιοχή μελέτης περιλαμβάνει τρία μεταλλεία που διακρίνονται στο παρακάτω χάρτη 3.2: το μεταλλείο της Ολυμπιάδας στα βόρεια της περιοχής μελέτης, το μεταλλείο του Στρατωνίου ή Μαύρες πέτρες που βρίσκεται στο κέντρο και ανατολικά και το μεταλλείο των Σκουριών στα νοτιο-δυτικά της περιοχής. Στο σχήμα 3.3 παρουσιάζεται ο χάρτης με τις συντεταγμένες των σταθμών.



*Εικόνα 3.2: Χάρτης περιοχής μελέτης με τις θέσεις σταθμών παρακολούθησης υπογείων υδάτων*

Η εταιρία διαθέτει περίπου 135 σημεία παρακολούθησης για τα υπόγεια ύδατα. Για την συγκεκριμένη εργασία χρησιμοποιήθηκαν δεδομένα 62 ενεργών σημείων. Τα δεδομένα αναφέρονται σε μετρήσεις συγκέντρωσης χημικών στοιχείων που συλλέχθηκαν από το 1997 έως το 2021. Από κάθε γεώτρηση επιλέχθηκαν οι μετρήσεις των χημικών στοιχείων Cd και Mn, διότι οι στατιστικές τους ιδιότητες ήταν καταλληλότερες για την επιθυμητή γεωστατιστική ανάλυση.

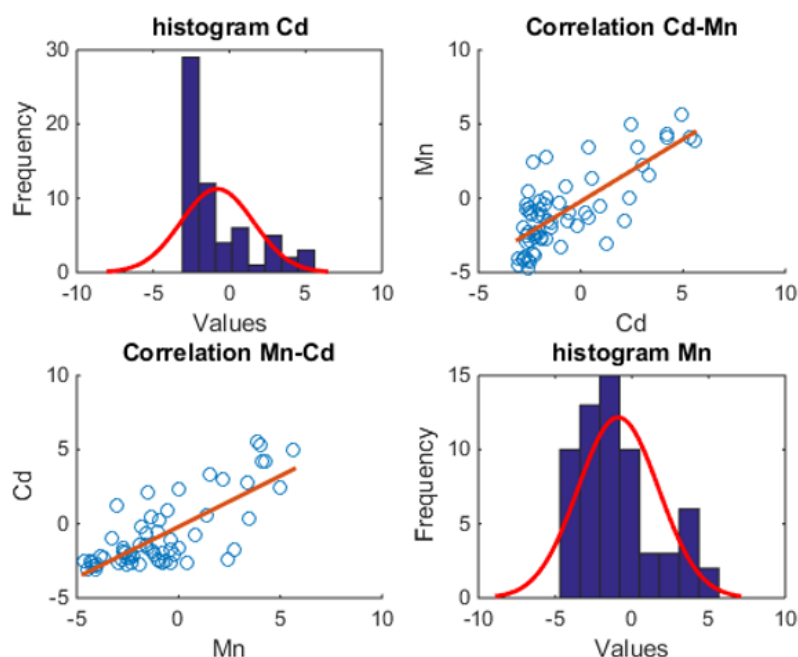


Εικόνα 3.3: Θέσεις σταθμών παρακολούθησης υπογείων υδάτων

Για την ανάλυση χρησιμοποιήθηκε ο χρονικός μέσος όρος της συγκεντρώσης σε κάθε σταθμό. Αυτός ο μέσος όρος σε κάθε θέση αποτελεί τα δεδομένα που χρησιμοποιήθηκαν στην ανάλυση που ακολουθεί.

## 3.2 Στατιστική Ανάλυση

Με τον όρο στατιστική ανάλυση εννοείται ο προσδιορισμός των παραμέτρων κατανομής πιθανότητας (βλ. 2.2) των δεδομένων και ο έλεγχος κανονικότητας. Τα αποτελέσματα των ιστογραμμάτων με τα δεδομένα των σταθμών δεν ταιριάζουν στην κανονική κατανομή. Για αυτό το λόγο, έγινε ο λογαριθμικός μετασχηματισμός  $\Phi$  σύμφωνα με την εξίσωση  $X^*(s) \equiv \log(X(s))$ , όπου με  $X^*$  συμβολίζεται το πεδίο των λογαρίθμων της συγκέντρωσης. Στο σχήμα 3.4 παρουσιάζονται τα ιστογράμματα κάθε στοιχείου όπως και η μεταξύ τους συσχέτιση, που υπολογίστηκε 54% (ισχυρή συσχέτιση).



Εικόνα 3.4: Συσχέτιση Cd-Mn με μετασχηματισμό λογαρίθμου

### 3.2.1 Παράμετροι Κατανομών Πιθανότητας

Οι παράμετροι κατανομής πιθανότητας των στοιχείων Cd και Mn εμφανίζονται στον παρακάτω πίνακα. Ακολουθούν τα ιστογράμματα στα σχήματα 3.5 και 3.6.

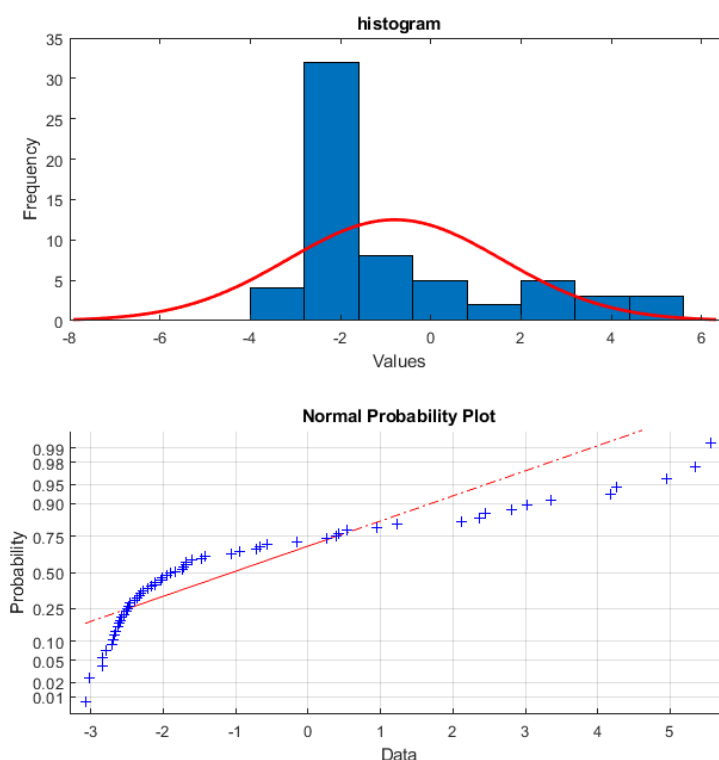
Πίνακας 3.1: Παράμετροι Κατανομών Πιθανότητας Cd-Mn με μετασχηματισμό λογαρίθμου

| Statistic              | Cd      | Mn      |
|------------------------|---------|---------|
| Μικρότερη Τιμή         | -3.0764 | -4.6721 |
| Μέγιστη Τιμή           | 5.5673  | 5.6682  |
| Μέση τιμή              | -0.7959 | -0.8834 |
| Διάμεσος               | -1.8638 | -1.2817 |
| Τυπική Απόκλιση        | 2.3736  | 2.6253  |
| Συντελεστή Ασυμμετρίας | 1.3020  | 0.7641  |
| Κύρτωση                | 3.5167  | 2.7944  |

Παρατηρείται πως λόγω του λογαρίθμου υπάρχουν τιμές μικρότερες του μηδέν. Όσον αφορά το στοιχείο Cd, από την διάμεσο γίνεται αντιληπτό πως οι περισσότερες τιμές και το μέσο της κατανομής βρίσκονται κοντά στο -2. Η μέση τιμή και η διάμεσος διαφέρουν αρκετά, άρα δεν υπάρχει συμμετρική κατανομή όπως διακρίνεται φανερά στο ιστογράμματα. Ο

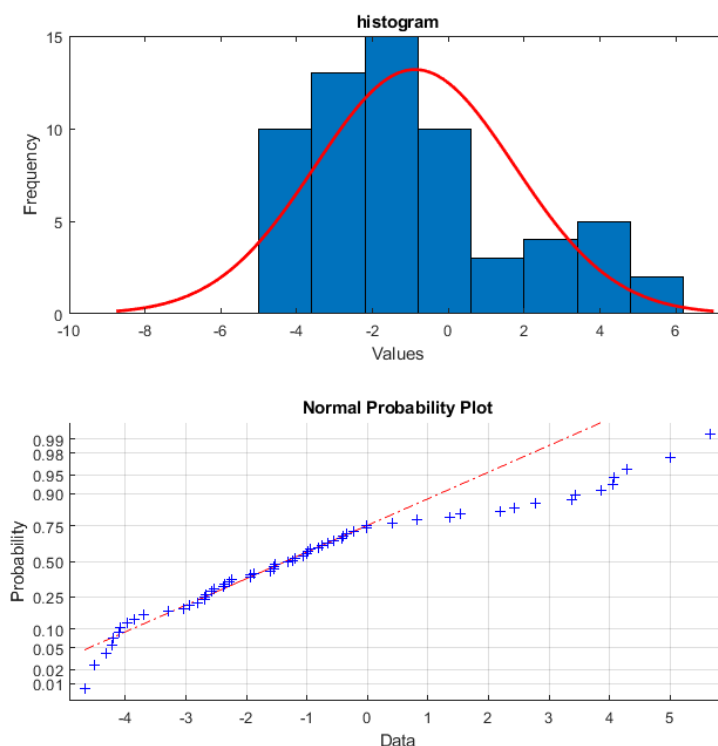
### Chapter 3. Γεωστατιστική Ανάλυση Δεδομένων

συντελεστής κύρτωσης είναι μεγαλύτερος από τρία. Αυτό φανερώνει πως η κύρτωση είναι λεπτοκυρτική και η πλειοψηφία των τιμών βρίσκονται στην μία άκρη και όχι στην κέντρο.



Εικόνα 3.5: Ιστόγραμμα και Norm Plot Cd με μετασχηματισμό λογαρίθμου

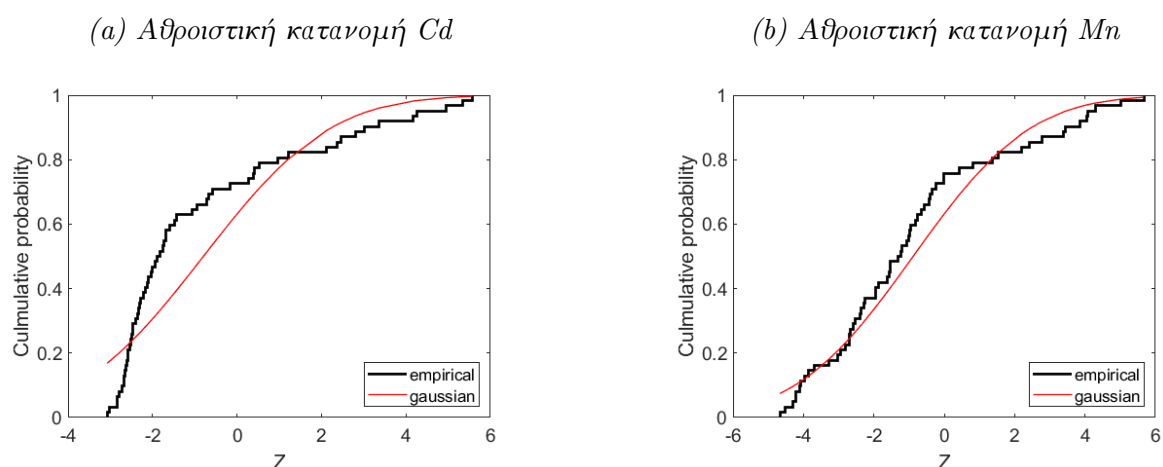
Ένας άλλος τρόπος που εξετάζεται πόσο απέχει μια κατανομή από την κανονική είναι μέσω του διαγράμματος κανονικής πιθανότητας (normal probability plot ή Normplot). Στο διάγραμμα αυτό, συγκρίνεται η αθροιστική κατανομή πιθανότητας των δεδομένων με την αθροιστική κατανομή πιθανότητας της κανονικής κατανομής. Όταν τα δεδομένα ακολουθούν την κανονική κατανομή, τότε τα σημεία εμφανίζονται κατά μήκος της γραμμής αναφοράς που αντιστοιχεί στην κανονική κατανομή. Διαφορετικές κατανομές εμφανίζουν απόκλιση. Παρατηρούμε στην Εικόνα 3.5 πως τα σημεία των δεδομένων δημιουργούν μια καμπυλωτή γραμμή γύρω από την κόκκινη γραμμή αναφοράς.



Εικόνα 3.6: Ιστόγραμμα και NormPlot Mn με μετασχηματισμό λογαρίθμου

Στην περίπτωση του Mn παρατηρούνται επίσης αρνητικές τιμές στις παραμέτρους λόγω λογαρίθμου. Τα δεδομένα όμως έχουν μεγαλύτερο εύρος τιμών με αποτέλεσμα οι τιμές της διαμέσου και της μέσης τιμής να είναι σχετικά κοντά και να υπάρχει καλύτερη συμμετρία στην κατανομή αυτή 3.6. Ο συντελεστής συμμετρίας είναι μικρότερος από ένα ενώ ο συντελεστής κύρτωσης παρουσιάζει πλατυκυρτική κατανομή αφού η τιμή του είναι μικρότερη από τρία.

Στην Εικόνα 3.7 παρουσιάζονται τα διαγράμματα αθροιστικής κατανομής με τον μετασχηματισμό του λογαρίθμου των στοιχείων Cd και Mn. Όπως φαίνεται στα διαγράμματα αυτά, η αθροιστική κατανομή του Mn πλησιάζει και ταιριάζει περισσότερο στην κανονική κατανομή σε σχέση με του Cd.



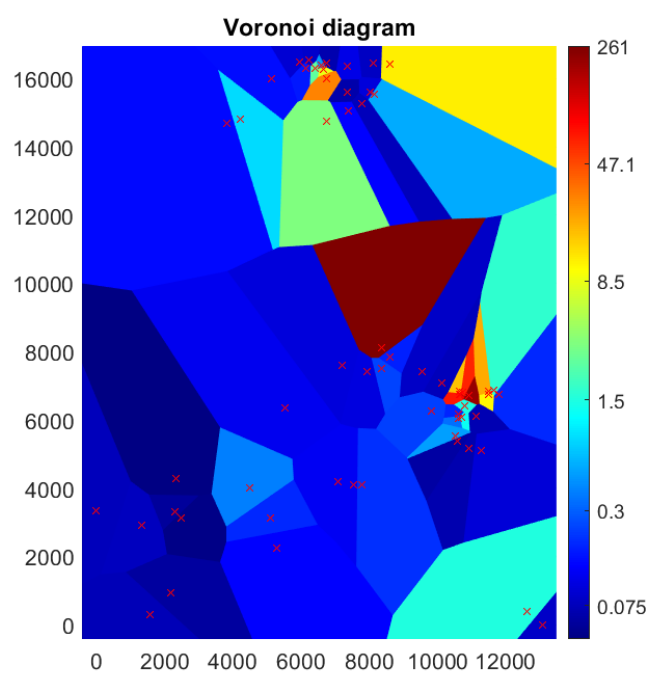
Εικόνα 3.7: Αθροιστική κατανομή λογαρίθμου

### 3.2.2 Voronoi

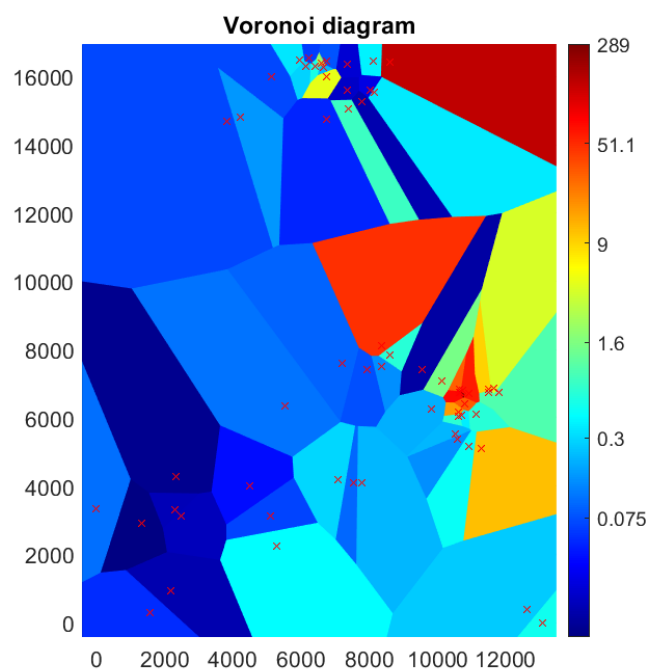
Οι θέσεις των δεδομένων στη περιοχή μελέτης και οι τιμές τους παρουσιάζονται μέσω των διαγραμμάτων Voronoi. Πρέπει να σημειωθεί πως σε αυτά τα διαγράμματα δεν εφαρμόστηκε κανένας μετασχηματισμός. Από το μέγεθος των πολυγώνων φαίνεται η πυκνότητα της δειγματοληψίας των μετρήσεων που εδώ είναι ακανόνιστη. Όπως φαίνεται στο διάγραμμα του Cd 3.8a οι τιμές της συγκέντρωσης στα σημεία μελέτης είναι γενικά χαμηλές με λίγες εξαιρέσεις. Όμως η τιμή της συγκέντρωσης μεταβάλλεται έντονα σε μικρές αποστάσεις γύρω από τα σημεία των υψηλών τιμών. Οι έντονες διακυμάνσεις των τιμών σε κοντινές αποστάσεις υποδεικνύει πως δεν υπάρχει στασιμότητα. Λόγω της μη στασιμότητας πρέπει να εκτιμηθεί μια εξίσωση τάσης που να περιγράφει επαρκώς αυτές τις υψηλές τιμές και να αφαιρεθεί. Με αυτό το τρόπο θα υποθέσουμε πως το πεδίο των διακυμάνσεων (βλ. 2.1.3) έχει μεγαλύτερη στατιστική ομοιογένεια. Στην περίπτωση του Mn 3.8b, το εύρος των τιμών, παρατηρείται λίγο μεγαλύτερο από του Cd. Φαίνεται από το διάγραμμα πως όπως και στο Cd οι περισσότερες τιμές είναι χαμηλές, αλλά υπάρχουν πάλι λίγες εξαιρέσεις. Υπάρχουν και σε αυτή την περίπτωση έντονες μεταβολές της τιμής σε μικρές αποστάσεις υποδηλώνοντας έλλειψη στασιμότητας. Όπως και στο Cd θα πρέπει να βρεθεί η τάση και να αφαιρεθεί.



(a) Voronoi διάγραμμα Cd (ppd)



(b) Voronoi διάγραμμα Mn (ppm)



Εικόνα 3.8: Voronoi διαγράμματα χωρίς μετασχηματισμό

### 3.2.3 Τάση

Η μαθηματική συνάρτηση που περιγράφει καλύτερα την τάση στο πεδίο είναι εξίσωση 2ου βαθμού της μορφής:

$$M(x) = a_0 + a_1X + a_2Y + a_3X^2 + a_4Y^2 + a_5XY$$

όπου οι γραμμικοί συντελεστές είναι:

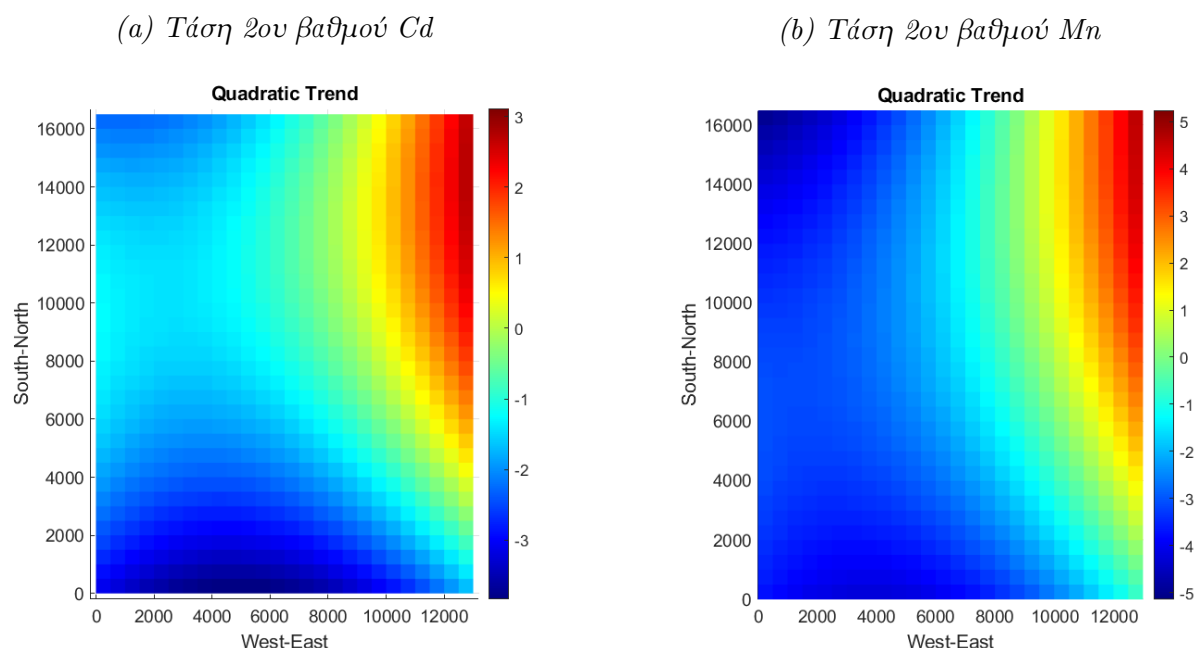
Πίνακας 3.2: Γραμμικοί Συντελεστές Τάσης με μετασχηματισμό λογαρίθμου

| coef  | Cd                     | Mn                      |
|-------|------------------------|-------------------------|
| $a_0$ | 3.0818                 | -3.6910                 |
| $a_1$ | $0.3254 \cdot 10^{-3}$ | $-0.3054 \cdot 10^{-3}$ |
| $a_2$ | $0.3973 \cdot 10^{-3}$ | $0.2183 \cdot 10^{-3}$  |
| $a_3$ | $0.3495 \cdot 10^{-7}$ | $0.4360 \cdot 10^{-7}$  |
| $a_4$ | $0.2154 \cdot 10^{-7}$ | $-0.1857 \cdot 10^{-7}$ |
| $a_5$ | $0.1743 \cdot 10^{-7}$ | $0.3264 \cdot 10^{-7}$  |

Ο Συντελεστής Συσχέτισης (correlation coefficient) με 2ου βαθμού τάση για το στοιχείο του Cd είναι ίσος με 45,70%, όπου θεωρείται αρκετά καλή συσχέτιση. Ενώ για το στοιχείο του Mn είναι ίσος με 63,76%, με ισχυρή συσχέτιση.

Στο διάγραμμα τάσης του Cd 3.9a απεικονίζονται βορειο-ανατολικά οι υψηλές τιμές, με ομαλή μείωση μέχρι το κέντρο. Αντίθετα, η εμφάνιση απότομων καμπυλών στα νοτιο-δυτικά αντιπροσωπεύει μια πιο γρήγορη μείωση της τιμής. Ομοίως, στο διάγραμμα τάσης του Mn 3.9b υπάρχουν οι υψηλές τιμές στα βορειο-ανατολικά του χάρτη και οι τιμές μειώνονται ομαλά προς τα δυτικά.

Το πεδίο της διακύμανσης  $X'$  προκύπτει από την αφαίρεση της τάσης από τα δεδομένα. Η διακύμανση θα χρησιμοποιηθεί στα επόμενα στάδια της ανάλυσης (βλ. 2.1.3).



Εικόνα 3.9: Διαγράμματα Τάσης με μετασχηματισμό λογαρίθμου

### 3.2.4 Βαριογραφία

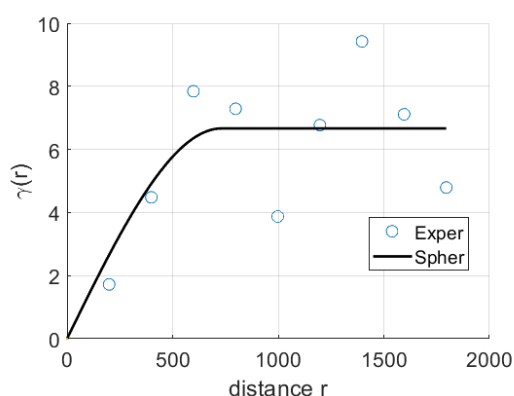
Συνεχίζοντας στο κομμάτι της βαριογραφίας χρησιμοποιούνται τα δεδομένα της διακύμανσης των λογαρίθμων. Η βαριογραφία εξυπηρετεί στην έρευνα και κατανόηση της χωρικής εξάρτησης των τιμών σε σχέση με την απόσταση. Σε πρώτο στάδιο, έγινε εμπειρικό ημιβαριόγραμμα για το κάθε στοιχείο. Για να εξεταστεί η χωρική συσχέτιση εξετάστηκαν διαφορετικά μοντέλα βαριογράμματος. Η επιλογή του καταλληλότερου μοντέλου έγινε με έλεγχο του τετραγωνικού σφάλματος προσαρμογής. Δηλαδή επιλέχθηκε το μοντέλο που έδινε την μικρότερη τετραγωνική απόκλιση από το εμπειρικό ημιβαριόγραμμα.

Στον Πίνακα 3.3 παρουσιάζονται οι παράμετροι βαριογράμματος για τα στοιχεία Cd και Mn.

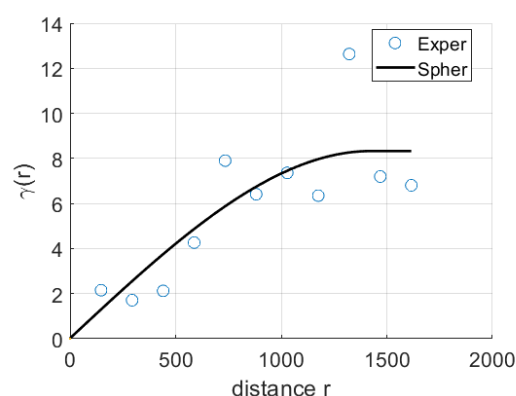
Πίνακας 3.3: Παράμετροι Βαριόγραμματος Cd και Mn:  $\sigma$  είναι η τυπική απόκλιση (Standard deviation),  $h$  είναι η ακτίνα συσχέτισης (Range) και  $c_0$  το φαινόμενο κόκκου (Nugget)

| Παράμετροι Βαριόγραμματος |                        |                     |
|---------------------------|------------------------|---------------------|
|                           | Cd                     | Mn                  |
| $\sigma^2$                | 6.6634                 | $0.0083 \cdot 10^3$ |
| $h(m)$                    | 731.4951               | $1.4246 \cdot 10^3$ |
| $c_0$                     | $1.6434 \cdot 10^{-4}$ | 0.0188              |

(a) Σφαιρικό Βαριόγραμμα Cd



(b) Σφαιρικό Βαριόγραμμα Mn



Εικόνα 3.10: Βαριόγραμμα Cd(ppb) και Mn(ppm). Οι κύκλοι (o) περιγράφουν το Εμπειρικό Βαριόγραμμα ενώ η γραμμή το σφαιρικό μοντέλο.

Το αποτέλεσμα της προσαρμογής στο βαριόγραμμα 3.10a φαίνεται ικανοποιητικό. Το φαινόμενο κόκκου είναι πολύ μικρό σε σχέση με τη διασπορά. Ακόμη, στο σφαιρικό βαριόγραμμα παρατηρείται πως η ζώνη επιρροής (range) είναι στα 731 μέτρα όπου σημαίνει πως οι τιμές των σημείων μέχρι 731 μέτρα είναι χωρικά εξαρτημένες. Το βαριόγραμμα του Mn 3.10b προκύπτει καλύτερο σε σχέση με αυτό του Cd, καθώς το σφαιρικό μοντέλο φαίνεται να ταιριάζει καλύτερα στο εμπειρικό βαριόγραμμα αλλά και επειδή η ακτίνα συσχέτισης (range) είναι μεγαλύτερη, στα 1425 μέτρα. Επίσης, και στην περίπτωση του Mn το φαινόμενο κόκκου είναι πολύ μικρό σε σχέση με τη διασπορά.

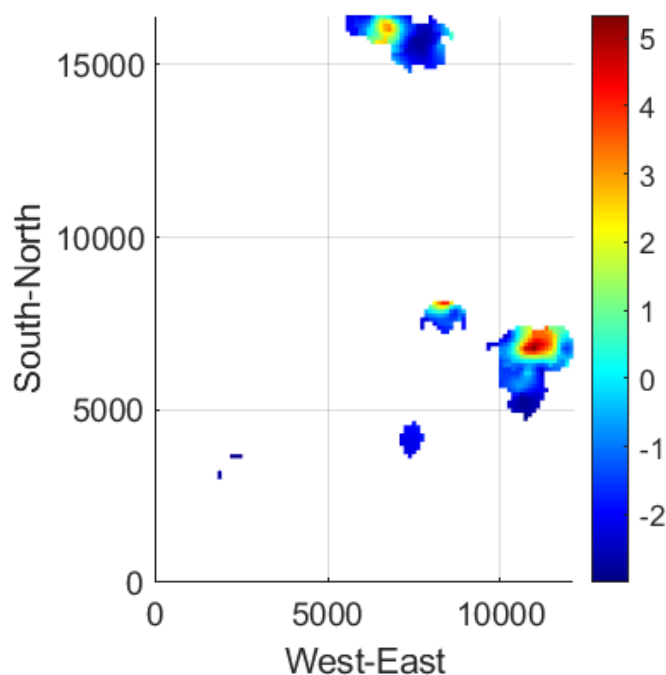
Άρα, με αυτές τις παραμέτρους είναι εφικτό να συνεχιστεί η ανάλυση και να ακολουθήσει η μέθοδος kriging.

### 3.2.5 Kriging

Η μέθοδος kriging (βλ. Κεφ. 2.5) στην εργασία αυτή βασίζεται στο σφαιρικό μοντέλο βαριογράμματος. Οι τιμές που χρησιμοποιούνται στο μοντέλο είναι οι παράμετροι των βαριογραμμάτων (βλ. 3.3), δηλαδή η τυπική απόκλιση  $\sigma$ , η ακτίνα συσχέτισης  $h$  και το φαινόμενο κόκκου  $c_0$ . Η χρήση του πίνακα βαριογράμματος σε αντίθεση με τον πίνακα συνδιασποράς μπορεί να γίνει λόγω της εξίσωσης 2.24.

Εφαρμόστηκε η μέθοδος OK σε κάναβο με 138x109 κελιά, όπου το κάθε κελί είναι 120 μέτρα και για τα δύο στοιχεία.

Οι γειτονιές στους χάρτες Kriging θα έπρεπε να δημιουργηθούν με την ακτίνα συσχέτισης των βαριογραμμάτων. Όμως επειδή αυτή η ακτίνα είναι μικρή και για τα δύο στοιχεία, τα αποτελέσματα δεν ήταν ικανοποιητικά. Οι χάρτες που προέκυπταν δεν έδιναν εκτίμηση σε μεγάλες περιοχές όπως φαίνεται στη παρακάτω Εικόνα 3.11.



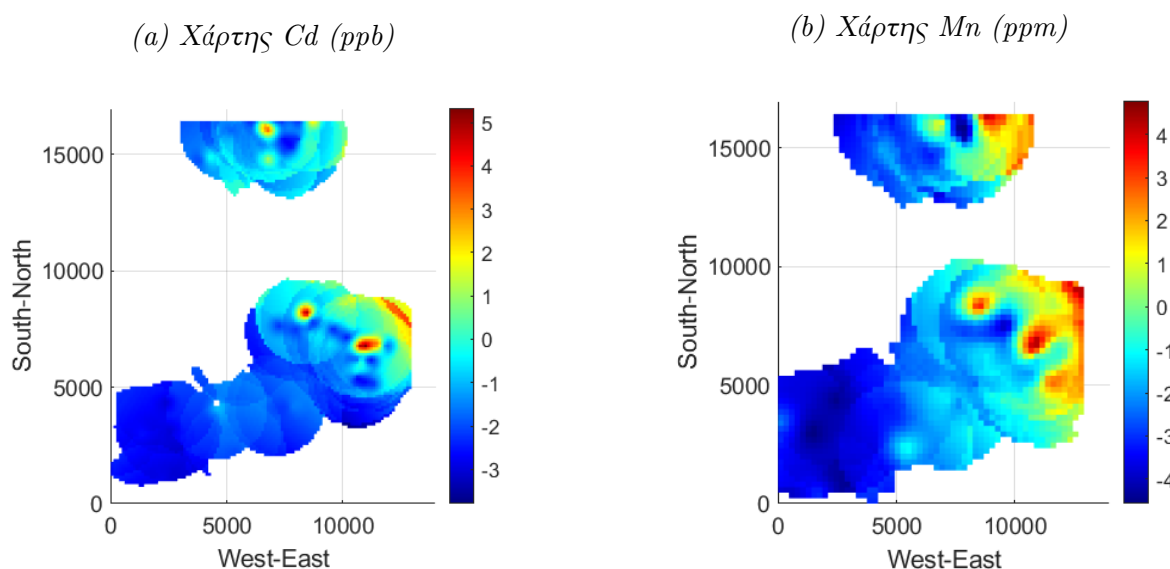
Εικόνα 3.11: Χάρτης περιοχής μελέτης με γειτονιά εκτίμησης όσο η ακτίνα συσχέτισης για το Cd.

Για το λόγο αυτό οι χάρτες υπολογίστηκαν με ακτίνα συσχέτισης 3 φορές μεγαλύτερη

για το Cd και 2 φορές μεγαλύτερη για το Mn. Η θεώρηση αυτή δεν θεωρείται σωστή αλλά μπορεί να χρησιμοποιηθεί στην περίπτωση που χρησιμοποιείται OK. Το OK μπορεί να εφαρμοστεί σε ασυσχέτιστες τιμές αλλά η αβεβαιότητα των αποτελεσμάτων θα είναι μεγάλη, όπως θα φανεί πιο κάτω στους χάρτες αβεβαιότητας ( $\sigma_{OK}^2$ ).

### Kriging στις διακυμάνσεις των λογαρίθμων

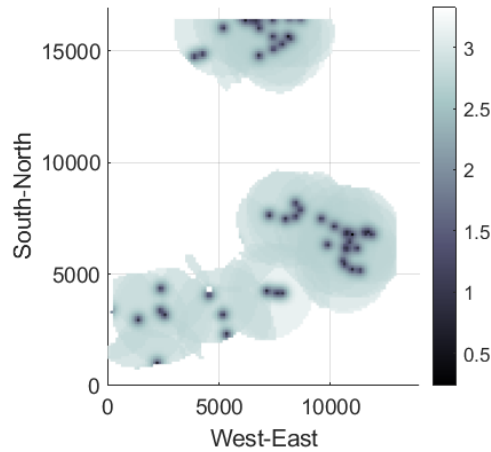
Εφαρμόστηκε OK στα δεδομένα της διακύμανσης των λογαρίθμων  $X^*(s)$ . Στους ακόλουθους λογαριθμικούς χάρτες Kriging 3.12 εντοπίζονται υψηλές τιμές περιεκτικότητας των στοιχείων Cd, Mn σε λίγα σημεία και συγκεκριμένα περισσότερο στο κέντρο και στα ανατολικά της περιοχής μελέτης. Το Cd παρουσιάζει λιγότερο υψηλές τιμές λογαρίθμου από το Mn. Μία παρατήρηση είναι πως παρόλο που αφαιρέθηκε η τάση υπάρχουν ακόμα έντονες και απότομες μεταβολές της εκτιμώμενης τιμής. Αυτές οι μεταβολές γίνονται σε μικρή απόσταση γύρω από τρία σημεία μόνο, υποδηλώνοντας πως το πεδίο των διακυμάνσεων δεν είναι στάσιμο.



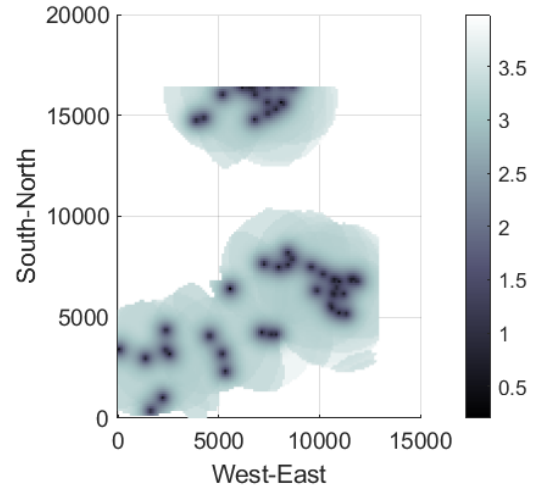
Εικόνα 3.12: Χάρτες εκτίμησης λογαριθμικού της συγκέντρωσης των στοιχείων.

Οι χάρτες τυπικής απόκλισης σφάλματος (Kriging Error Standard Deviation) 3.13 παρουσιάζουν πόσο μεταβάλλεται στο χώρο η αβεβαιότητα της εκτίμησης.

(a) Χάρτης τυπικής απόκλισης σφάλματος Kriging Cd



(b) Χάρτης τυπικής απόκλισης σφάλματος Kriging Mn



Εικόνα 3.13: Χάρτες τυπικής απόκλισης σφάλματος Kriging με μετασχηματισμό λογαρίθμου

Αυτοί οι χάρτες εμφανίζονται μόνο με την λογαριθμική μορφή καθώς θα ήταν λάθος να μετασχηματιστούν ως έχουν. Αυτό συμβαίνει γιατί κατά την αντιστροφή του λογαριθμικού μετασχηματισμού  $\Phi$ , γίνεται ύψωση. Δηλαδή,  $\hat{X}'(\mathbf{u}) = \exp(\hat{X}^*(\mathbf{u}))$ .

Αν οι εκτιμώμενες τιμές υψωθούν στο λογάριθμο αρχικά φαίνεται πως προστίθεται ή αφαιρείται η αβεβαιότητα αλλά στην πραγματικότητα λόγω του λογαρίθμου πολλαπλασιάζεται ή διαιρείται όπως φαίνεται στο παράδειγμα της εξίσωσης 3.1 .

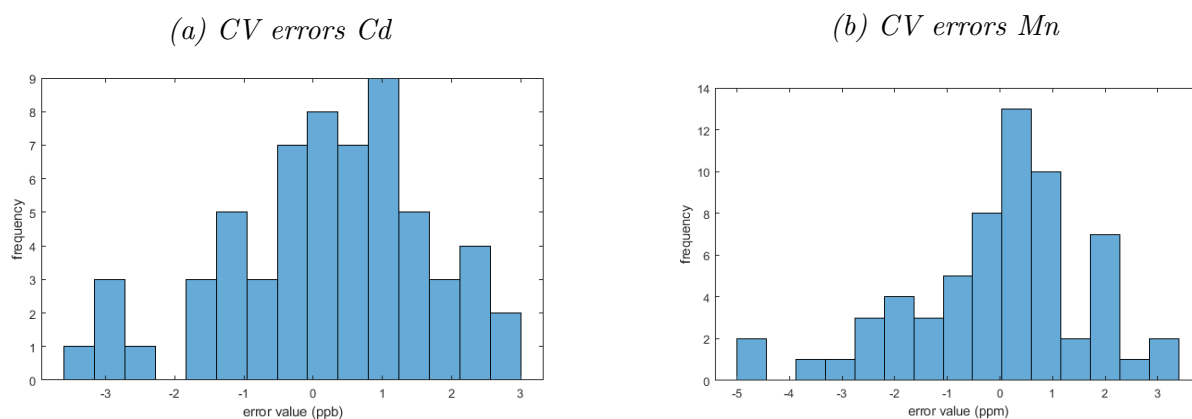
$$\exp [\hat{X}^*(\mathbf{u}) + \sigma_{OK}(\mathbf{u})] = \exp [\hat{X}^*(\mathbf{u})] \cdot \exp [\sigma_{OK}(\mathbf{u})] \quad (3.1)$$

Άρα, δεν είναι εφικτό να παρουσιαστούν αντίστοιχοι χάρτες. Η αβεβαιότητα  $\sigma_{OK}$  όμως μπορεί να δείξει πόσο μεγαλώνει η αβεβαιότητα μακριά από τις θέσεις των γεωτρήσεων. Παρατηρείται πως η αβεβαιότητα μεγαλώνει απότομα, καθώς από το 1-1,5 που είναι η τιμή κοντά στις γεωτρήσεις, αυξάνεται στο 3-3,5. Άρα, επειδή είναι λογάριθμος η αύξηση αυτή είναι πολύ μεγάλη.

Για την σύγκριση και την αξιολόγηση των τιμών της διακύμανσης με τις εκτιμώμενες τιμές (βλ. Κεφ. 2.5.3) παρουσιάζονται παρακάτω οι τιμές επαλήθευσης της διακύμανσης στα

### Chapter 3. Γεωστατιστική Ανάλυση Δεδομένων

ιστογράμματα 3.14 και στον Πίνακα 3.4. Λαμβάνοντας υπόψη τα διαγράμματα σφάλματος της διασταυρωμένης επαλήθευσης (CV Errors) , παρατηρείται ότι τα περισσότερα σφάλματα, τόσο στο Cd όσο και στο Mn, εμφανίζονται μεταξύ των τιμών -3 έως 3.



Εικόνα 3.14: CV errors με μετασχηματισμό λογαρίθμου

Στη συνέχεια, ακολουθεί ο Πίνακας 3.4 όπου εμφανίζονται οι στατιστικές τιμές επαλήθευσης της διακύμανσης.

Πίνακας 3.4: Τιμές επαλήθευσης στη διακύμανση, όπου MAE είναι το Μέσο Απόλυτο Σφάλμα, RMSE η ρίζα του Μέσου Τετραγωνικού Σφάλματος, MaxAE η τιμή του Μέγιστου Απόλυτου Σφάλματος και  $\rho$  ο συντελεστής συσχέτισης

|        | Cd(ppb) | Mn(ppm) |
|--------|---------|---------|
| MAE    | 1.2871  | 1.3155  |
| RMSE   | 1.6665  | 1.7015  |
| MaxAE  | 5.9343  | 4.9114  |
| $\rho$ | 60.87%  | 55.57%  |

Παρατηρείται πως οι τιμές του RMSE είναι και στα δύο στοιχεία περίπου ίσες με του MAE, ενώ το MaxAE και στις δύο περιπτώσεις είναι περίπου τέσσερις με πέντε φορές μεγαλύτερο του MAE. Είναι λογικό οι τιμές του MaxAE να είναι οι μέγιστες γιατί αντιπροσωπεύουν τις μεγαλύτερες απόλυτες τιμές του σφάλματος.

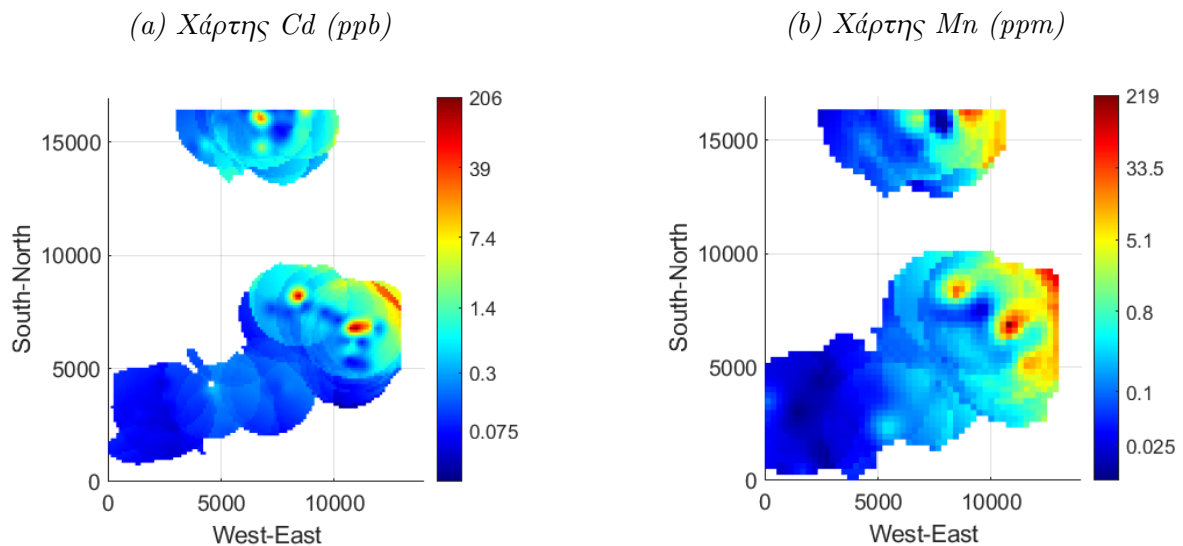


**Kriging στις εκτιμώμενες τιμές**

Οι χάρτες που παρουσιάστηκαν στην 3.12 είναι η χάρτες της διακύμανσης του λογαρίθμου των συγκεντρώσεων Cd και Mn. Για να επανέρθει η εκτίμηση στην αρχική κλίμακα που έχει φυσική σημασία, πρέπει να αντιστραφεί ο μετασχηματισμός. Πρώτα προστίθεται στην εκτίμηση της διακύμανσης η πολυωνυμική τάση. Στην συνέχεια αντιστρέφεται ο λογαριθμικός μετασχηματισμός. Η αντιστροφή φαίνεται στην εξίσωση 3.2

$$\hat{X}(\mathbf{u}) = \exp \left[ \hat{X}^*(\mathbf{u}) + M(\mathbf{u}) \right] \quad (3.2)$$

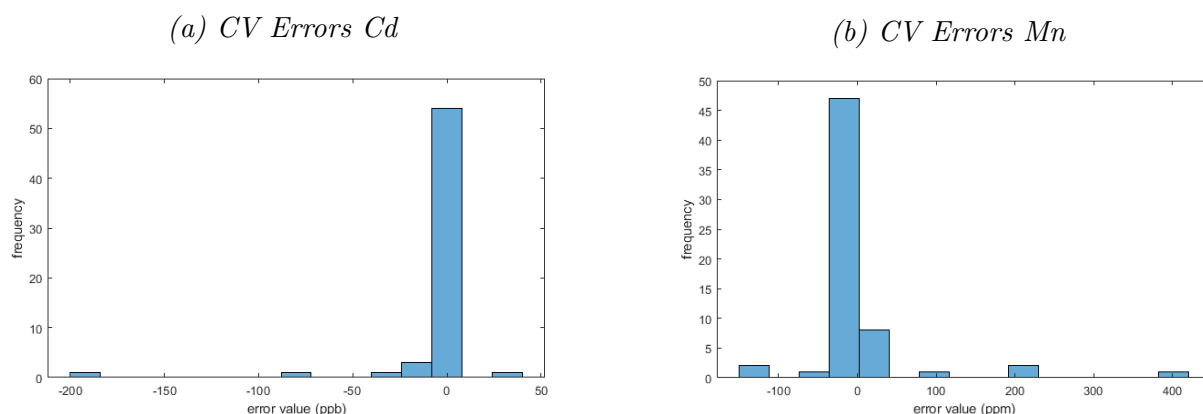
και οι χάρτες της εκτίμησης φαίνονται στα σχήματα 3.15.



Εικόνα 3.15: Χάρτες εκτίμησης συγκέντρωσης των στοιχείων (ppb, ppm)

Για την αξιολόγηση της εκτίμησης πραγματοποιήθηκε επαλήθευση με LOOCV στο πεδίο των εκτιμώμενων συγκεντρώσεων τόσο για το Cd όσο και για το Mn. Το ιστόγραμμα των σφαλμάτων LOOCV, μετά τον αντίστροφο μετασχηματισμό (βλ 3.2) φαίνεται στην Εικόνα 3.16 και για τα δύο στοιχεία.

Για να αξιολογηθούν τα αποτελέσματα της εκτίμησης, χρησιμοποιείται πάλι LOOCV μετά τον αντίστροφο μετασχηματισμό.



Εικόνα 3.16: CV Errors

Πίνακας 3.5: Τιμές επαλήθευσης στη διακύμανση και τάση

|        | Cd (ppb) | Mn (ppm) |
|--------|----------|----------|
| MAE    | 6.6388   | 21.8229  |
| RMSE   | 27.4962  | 70.2647  |
| MaxAE  | 194.0097 | 413.6846 |
| $\rho$ | 62.23%   | 47.40%   |

Εξετάζοντας τα ιστογράμματα 3.16 παρατηρείται πως τα περισσότερα σφάλματα έχουν χαμηλές τιμές. Όμως για το στοιχείο του Cd υπάρχει μία ακραία τιμή σφαλμάτων ενώ για το στοιχείο του Mn παρατηρούνται δύο ακραίες τιμές. Ένας περιορισμός της μεθόδου LOOCV είναι πως κατά την σειριακή αφαίρεση των δεδομένων, η εκτίμηση σημείων που έχουν πολύ διαφορετική τιμή από τους γείτονές τους, θα δώσει μεγάλα σφάλματα. Ο περιορισμός αυτός δημιουργεί τις ακραίες αποκλίσεις που αναφέρθηκαν. Από το 3.8 φαίνονται δύο περιοχές με πολύ υψηλή τιμή για το Cd το Mn στην περιοχή του Στρατωνίου και μια περιοχή στην Ολυμπιάδα. Η ύπαρξη τελμάτων απόθεσης αποβλήτων αιτιολογούν τις ακραίες τιμές.

Από τον Πίνακα 3.5 προκύπτει ότι και στα δύο στοιχεία Cd, Mn η τιμή του RMSE είναι σχεδόν τρεις με τέσσερις φορές υψηλότερη από το MAE. Η τιμή του MaxAE δείχνει την υψηλότερη απόλυτη τιμή (αρνητική ή θετική) και επηρεάζεται αποκλειστικά από την μέγιστη και ελάχιστη τιμή της εκτιμώμενης μεταβλητής, επομένως η παρουσία ακραίων τιμών σε κάθε στοιχείο δημιουργεί αυτή την απόκλιση. Είναι δηλαδή αναμενόμενες οι πολύ υψηλές τιμές για το MaxAE που φαίνονται στον πίνακα. Ο συντελεστής συσχέτισης επηρεάζεται έντονα

### Chapter 3. Γεωστατιστική Ανάλυση Δεδομένων

από τις ακραίες τιμές. Η ύπαρξη δύο ακραίων τιμών στα δεδομένα του Mn παρουσιάζει  $\rho = 0,467$  που είναι μέτρια συσχέτιση. Αν κατά την εφαρμογή της LOOCV δεν γίνει εκτίμηση στα δύο σημεία με τις ακραίες τιμές του Mn τα αποτελέσματα της επαλήθευσης βελτιώνονται σημαντικά. Ο συντελεστής συσχέτισης γίνεται  $\rho^- = 83,48 \%$  και το  $\text{RMSE}^- = 67 \text{ ppm}$ . Δηλαδή, το  $\text{RMSE}^-$  έχει παραπλήσια τιμή με το  $\text{RMSE}$ , ενώ ο συντελεστής  $\rho^-$  είναι σημαντικά μεγαλύτερος από τον συντελεστή  $\rho$ .

#### Συζήτηση αποτελεσμάτων

Η επιλογή της μεθόδου Ordinary Kriging αποδείχτηκε σωστή καθώς βοήθησε να παρουσιαστούν καλύτερα αποτελέσματα στους χάρτες εκτίμησης της περιοχή μελέτης. Οι χάρτες εκτίμησης 3.15 δείχνουν πως στο μεταλλείο των Σκουριών, οι συγκεντρώσεις του Cd και Mn έχουν πολύ χαμηλή εκτιμώμενη τιμή. Στο μεταλλείο της Ολυμπιάδας οι τιμές του Cd έχουν μεγάλο εύρος. Η συγκέντρωση είναι χαμηλή στο μεγαλύτερο μέρος του ορυχείου, αλλά στα βόρεια υπάρχει μια μικρή περιοχή με μεγαλύτερες τιμές. Παρόμοια, οι συγκεντρώσεις του Mn στο μεταλλείο της ολυμπιάδας είναι χαμηλές στο δυτικό τμήμα της περιοχής, λιγότερο χαμηλές στα Νοτιοανατολικά, ενώ στα ανατολικά παρουσιάζονται μεγάλες συγκεντρώσεις. Η συγκέντρωση του Cd στο μεταλλείο του Στρατωνίου επίσης παρουσιάζει μεγάλο εύρος τιμών. Παρατηρώντας στα διαγράμματα Voronoi 3.8 διακρίνονται στο μεταλλείο αυτό και για τα δύο στοιχεία πολύγωνα με υψηλές συγκεντρώσεις. Ενώ από τους χάρτες kriging 3.15 φαίνεται στα ανατολικά και σε δύο μικρές περιοχές στο εσωτερικό του μεταλλείου η συγκέντρωση να είναι υψηλή και στα δύο στοιχεία. Παρατηρείται ακόμη μια ακραία τιμή στα ανατολικά της περιοχής όπου πιθανόν οφείλεται στη τάση των δεδομένων όπως φαίνεται και στα διαγράμματα τάσης 3.9. Η εμφάνιση των ακραίων υψηλών συγκεντρώσεων στο μεταλλείο του Στρατωνίου ίσως σχετίζεται με το χώρο ξηρής απόθεσης της εταιρίας που ονομάζουν Κοκκινόλακκα. Σύμφωνα με την εταιρία, υπολλείματα προηγούμενων μεταλλευτικών δραστηριοτήτων έχουν μεταφερθεί στο χώρο αυτό. Στόχος τους είναι να καταργηθούν οι πολλοί διαφορετικοί χώροι απόθεσης και να συγκεντρωθούν τα απόβλητα σε ένα χώρο για την προστασία του περιβάλλοντος.

## 4 Συμπεράσματα - Προτάσεις

### 4.1 Συμπεράσματα

Η παρούσα διπλωματική εργασία εκπονήθηκε με σκοπό τη γεωστατιστική ανάλυση συγκέντρωσης ρυπαντικών φορτίων, Cd και Mn στα υπόγεια ύδατα των μεταλλείων Χαλκιδικής. Η ρύπανση των υπόγειων νερών με υψηλές συγκεντρώσεις Cd, Mn μπορεί να δημιουργήσει πολλά περιβαλλοντικά προβλήματα και σοβαρό κίνδυνο στη γύρω περιοχή. Η γεωστατιστική ανάλυση έγινε σε περιβάλλον προγραμματισμού MatLab σε δεδομένα συγκεντρώσεων που παραχώρησε η εταιρία «Ελληνικός Χρυσός». Οι σταθμοί παρακολούθησης υπογείων υδάτων κατανέμονται κυρίως στις τρεις περιοχές μεταλλευτικής δραστηριότητας: Το μεταλλείο Ολυμπιάδας, Στρατωνίου και Σκουριών.

Ο λογάριθμος των συγκεντρώσεων βελτίωσε την εκτίμηση σε σχέση με την εφαρμογή των μεθόδων χωρίς μετασχηματισμό καθώς οι λογαριθμισμένες τιμές ταίριαζαν περισσότερο στην κανονική κατανομή. Έγινε αφαίρεση πολυωνυμικής τάσης (2ου βαθμού) καθώς ο συντελεστής συσχέτισης μεταξύ της τάσης και του λογαρίθμου της συγκέντρωσης παρουσίαζε μέτρια (Cd) έως ισχυρή (Mn) συσχέτιση. Κατά την απολογαρίθμηση, μικρές αποκλίσεις στις εκτιμήσεις μπορεί να προκαλέσουν μεγάλες αλλαγές. Στην περίπτωση της Χαλκιδικής, ο λογαριθμικός μετασχηματισμός δεν δημιούργησε αυξημένα σφάλματα ή ακραίες τιμές κατά τον αντίστροφο μετασχηματισμό. Η πολυωνυμική τάση, όπως φαίνεται στα σχήματα της τάσης 3.9 και του Kriging 3.15, συνεισφέρει σε κάποιες ακραίες τιμές που παρουσιάζονται στα όρια της περιοχής μελέτης για το Στρατωνίου και την Ολυμπιάδα.

Μετά από έλεγχο της προσαρμογής διάφορων μοντέλων ημιβαριογράμματος με το εμπειρικό ημιβαριόγραμμα, επιλέχθηκε το σφαιρικό μοντέλο και για τα δύο στοιχεία. Το μοντέλο αυτό είχε την καλύτερη προσαρμογή με βάση την αθροιστική τετραγωνική απόκλιση.

## Chapter 4. Συμπεράσματα - Προτάσεις

Σύμφωνα με τα αποτελέσματα του LOOCV, τα καλύτερα και πιο ακριβή αποτελέσματα λαμβάνονται για το στοιχείο του Cd αν συμπεριληφθούν όλα τα δεδομένα. Η εκτίμηση για τις συγκεντρώσεις και των δύο στοιχείων ήταν ικανοποιητική, με εξαίρεση κάποιες ακραίες τιμές. Σε περίπτωση που δεν γίνει LOOCV στις ακραίες τιμές, το στοιχείο του Mn έχει καλύτερο συντελεστή συσχέτισης. Το RMSE όμως δεν διαφοροποιείται πολύ στην περίπτωση αυτή καθώς δεν επηρεάζεται τόσο από ακραίες τιμές. Το αποτέλεσμα αυτό υποδεικνύει πως δεν είναι καλή πρακτική να βασίζεται ο ερευνητής μόνο σε ένα μέτρο επαλήθευσης καθώς μπορεί να οδηγηθεί σε εσφαλμένα συμπεράσματα. Οι χάρτες kriging τόσο για το Cd όσο και για το Mn παρουσιάζουν ως επί το πλείστον χαμηλή συγκέντρωση εκτός από μερικές μικρές περιοχές με πολύ υψηλές τιμές. Σύμφωνα με την Ελληνική Νομοθεσία τα επιτρεπτά όρια για το Cd και το Mn είναι 5 ppb και 0,05 ppm αντίστοιχα. Στο μεταλλείο των Σκουριών, οι εκτιμήσεις της συγκέντρωσης και στα δύο στοιχεία είναι πολύ χαμηλές και δεν υπερβαίνουν τα όρια. Στο Μεταλλείο της Ολυμπιάδας δεν παρουσιάζονται μεγάλες συγκεντρώσεις Cd εκτός από μια μικρή περιοχή στο κέντρο του μεταλλείου. Αντίθετα, στο Mn, εμφανίζεται μεγαλύτερη συγκέντρωση που ξεπερνάει τα επιτρεπτά όρια στα ανατολικά του μεταλλείου. Για το Στρατώνι, το Cd παρουσιάζει πολύ υψηλές συγκεντρώσεις σε δύο περιοχές στο εσωτερικό του μεταλλείου και στο ανατολικό άκρο της περιοχής. Η συγκέντρωση του Mn εμφανίζει αντίστοιχη συμπεριφορά με αποτέλεσμα και στα δύο στοιχεία να υπερβαίνεται η οριακή τιμή. Η τάση εξηγεί τις αποκλίσεις στο ανατολικό άκρο της περιοχής μελέτης, όπως αναφέρθηκε προηγουμένως. Τα σημεία που παρουσιάζουν υψηλή περιεκτικότητα σε Cd και Mn στο εσωτερικό του Στρατωνίου και της Ολυμπιάδας οφείλονται στους χώρους εναπόθεσης των αποβλήτων από τις εξορυκτικές δραστηριότητες. Ένας τέτοιος χώρος απόθεσης από προηγούμενες εξορυκτικές δραστηριότητες είναι ο Κοκκινόλακας. Ο χώρος αυτός έχει δηλωθεί ως στόχος για ένα έργο αποκατάστασης.

Η εταιρία «Ελληνικός Χρυσός» έχει καταφέρει μέχρι σήμερα να μειώσει την ρύπανση στην περιοχή της Χαλκιδικής με προγράμματα περιβαλλοντικής διαχείρισης και τεχνολογίες που μειώνουν το περιβαλλοντικό αποτύπωμα. Είναι πολύ σημαντικό όλες οι μεταλλευτικές επιχειρήσεις να εκτιμούν την σοβαρότητα της ρύπανσης που μπορεί να προκαλούν και να προσπαθούν να την αποτρέψουν.

## 4.2 Προτάσεις για μελλοντική έρευνα

Στα πλαίσια μελλοντικής βελτίωσης της παρούσας έρευνας θα μπορούσαν να εφαρμοστούν περισσότερες μέθοδοι χωρικής ανάλυσης. Η χρήση άλλων μετασχηματισμών όπως Box-Cox, Yeo & Johnson ή Modulus πιθανόν να βελτιώνει περαιτέρω την κατανομή των μετασχηματισμένων δεδομένων [30]. Η σύγκριση και η επιλογή μεταξύ αυτών των μετασχηματισμών θεωρείται πρωτότυπη έρευνα, με ελάχιστες έρευνες στην βιβλιογραφία. Ακόμα, για την βελτίωση της εκτίμησης της αβεβαιότητας που υπάρχει στην διπλωματική αυτή θα μπορούσε να χρησιμοποιηθεί σειριακή προσομοίωση Gauss (SGS). Είναι γνωστό πως η μέθοδος SGS προτείνεται για την κατασκευή ενός μοντέλου όταν η εφαρμογή εστιάζει σε ακραίες τιμές εκτιμώμενων παραμέτρων και μικρά σφάλματα βαθμονόμησης, όπως σε μια προσομοίωση μεταφοράς ρύπων σε ετερογενείς υδροφορείς [16]. Τέλος, μια ακόμη ιδέα θα ήταν να συμπληρωθεί η μέθοδος του Ordinary kriging με την μέθοδο του Indicator kriging [6, 22] για την εκτίμηση των συγκεντρώσεων υπόγειων υδάτων σε επικίνδυνες περιοχές μέσω της αθροιστικής κατανομής που προκύπτει με την χρήση διαφορετικών επιπέδων αποκοπής (cutoff) στο Indicator Kriging.

## 5 Παράρτημα Α'

### 5.1 Κώδικας MatLab

Στο συγκεκριμένο παράρτημα παρουσιάζεται ο κώδικας της MatLab που χρησιμοποιήθηκε στην παρούσα διπλωματική.

```

5 - load Cdori
6
7 - k=MOCd;
8
9 - G=isnan(k);
10 - x=X(G==0);
11 - y=Y(G==0);
12 - Avek=log(k(G==0));
13
14 - x=x-min(x); y=y-min(y);
15 - min_x=min(x);
16 - min_y=min(y); max_x=max(x); max_y=max(y);
17 - [Xp,Yp] = meshgrid(min_x:120:max_x, min_y:120:max_y);
18
19 - figure(1)
20 - set(gca,'fontsize', 13)
21 - plot(x,y,'x')
22 - axis equal
23
24 % % ***** Histogram, Normplot *****
25 - figure(2)
26 - subplot(2,1,1);
27 - histogram(Avek,8)
28 - histfit(Avek)
29 - title('histogram')
30 - xlabel('Values')
31 - ylabel('Frequency')
32
33 - subplot(2,1,2);
34 - normplot(Avek)
35
36 - set(gcf, 'Position', [70, 70, 700, 700])
37
38 % ***** Data statistics *****
39 - Avek_min = min(Avek)
40 - Avek_max = max(Avek)
41 - Avek_mean = mean(Avek) % Mesh timh
42 - Avek_median = median(Avek) % Diamesos
43 - Avek_std = std(Avek) % typikh apoklish
44 - Avek_skewness = skewness(Avek) % Syntelesths asymetrias. 0 ~> symmetrikh
45 - Avek_kurtosis = kurtosis(Avek) % Kyrtwsh. 3 ~> kanonikh katanomh
46

```



```

48 % ***** AveCdF *****
49 - [F,z] = ecdf(Avek);
50
51 - figure(3)
52 - axes1 = axes('Parent',3);
53 - stairs(z,F,'k','LineWidth',2)
54
55 - hold on
56 - yl=cdf('norm',z,Avek_mean,Avek_std);
57 - plot(z,yl,'r','LineWidth',1)
58 - xlabel('Z')
59 - ylabel('Cumulative probability')
60 - set(axes1,'FontSize',14);
61 - legend('empirical','gaussian','Location','southeast')
62
63 - hold off
64 %% Voronoi polygons%
65 % W=Vorongrap(x, y,exp(Avek));
66
67 - megis=floor(nanmax(nanmax(exp(Avek))));
68 - elax=nanmin(nanmin(exp(Avek)));
69
70 - Tt=linspace(-3,log(megis),6);
71 - Tt2=round(exp(Tt)*10)/10;
72 - Tt2(1)=Tt2(2)/4;
73 - Tt=log(Tt2);
74 - for i=1:6
75 -     L1{i}=num2str(exp(Tt(i)));
76 - end
77
78 - W=Vorongrap(x, y,Avek);
79 - colorbar('YTick', Tt, 'YTickLabel', L1);
80 % ylim([0 4500000]);
81 % xlim([0 4500000]);
82 - title('Voronoi diagram')
83 - set(gcf, 'Position', [800, 70, 700, 700])
84
85 %% Linear Trend
86 - [Fluct,Mx,LinTren,QuaTren]=Detrend(x,y,Avek,2);
87
88 % ***** Prediction of Trend *****
89 - Xp=reshape(Xp,138*109,1);
90 - Yp=reshape(Yp,138*109,1);

```

```

91 - MAvek= [ones(length(Xp),1) Xp Yp Xp.^2 Yp.^2 Xp.*Yp] * QuaTren;
92 - MAvek1= [ones(length(Xp),1) Xp Yp] * LinTren;
93
94 - Xk = unique(Xp);
95 - Yk = unique(Yp);
96 - MAvek=reshape(MAvek,length(Yk),length(Xk));
97
98
99 % ***** Figure *****
100 - [Apo]=Krigrids( Xk, Yk, MAvek );
101
102 - title('Quadratic Trend')
103
104 - xlim([min(Xp)-200 max(Xp)+200]);
105 - ylim([min(Yp)-200 max(Yp)+200]);
106 - set(gcf, 'Position', [100, 10, 650, 650])
107 %
108 - disp('estimated Reserves (Mt)')
109 - Apo*1.15*10^6 / 10^6;
110
111 %% Variography
112
113 - f=pi/2; tole=pi/2;
114 - MDIS = 0.10;
115 - Lags=9;
116
117 % ***** Variogram model Fitting *****
118 - MODEL=3;
119 - Nc=Lags;
120 - Weighting=0;
121
122 - [Param, ERR]=variofit([x y Fluct], MODEL, Nc, 1, MDIS, Weighting, 1);
123
124 %% Kriging
125 % ***** Ordinary Kriging on Fluctuations *****
126
127 - NERAD=300/100;
128 - minneigh=2;
129 - [EKT, Xk, Yk, SF, PLG]=OrdKrig(Xp, Yp, x, y, Fluct, Param, MODEL, NERAD, minneigh);
130

```

```

131 % ***** Cross Validation *****
132
133 - [ECV, ~, ~, Err]=crossval2(x,y,Fluct,Param,MODEL,NERAD);
134 - ECV(ECV==0)=nan;
135
136 % The histogram of the Cross validation errors.
137 - figure
138 - IPERIERGO=(Fluct<log(inf));
139 - FLL=Fluct(IPERIERGO);
140 - histogram(ECV(IPERIERGO)-FLL,15)
141 %title('CV errors')
142 - xlabel('error value (ppm)')
143 - ylabel('frequency')
144 - set(gcf, 'Position', [70, 70, 700, 400])
145
146 % ***** Adding Trend back *****
147
148 - KREST=MAvek+EKT;
149 - MMAvek= [ones(length(x),1) x y x.^2 y.^2 x.*y] * QuaTren;
150
151 - disp('Diakimansi kai tasi')
152 - DnT=MMAvek(IPERIERGO)+ECV(IPERIERGO);
153
154 - AA=Avek(IPERIERGO);
155 - I=isnan(DnT);
156 - CC=DnT(I~=1);
157 - AA=AA(I~=1);
158
159 - CC=exp(CC);
160 - AA=exp(AA);
161 - KREST2=exp(KREST);
162
163 - E=CC-AA; %Error
164 - SqE=(CC-AA).^2; %Square Error
165 - MAE=mean(abs(E)) %Mean Absolute Error
166 - MSE=mean(SqE) %Mean Square Error
167 - RMSE=sqrt(MSE) %Root Mean Square Error
168 - MaxAE=max(abs(E)) %Maximum Absolute Error
169
170 - RCA= corrccoef(CC,AA) %correlation coefficient
171
172

```

```

171
172
173 - figure
174 - histogram(E,15)
175 - xlabel('error value (ppm)')
176 - ylabel('frequency')
177 - set(gcf, 'Position', [70, 70, 700, 400])
178
179 % ***** RESERVES ESTIMATION *****
180
181 - [Apo]=Krigrids( Xk, Yk, KREST, SF)
182
183 - megis=floor(nanmax(nanmax(KREST2)));
184 - elax=nanmin(nanmin(KREST2));
185
186 - TT=linspace(-3,log(megis),6);
187 - TT2=round(exp(TT)*10)/10;
188 - TT2(1)=TT2(2)/4;
189 - TT=log(TT2);
190 - for i=1:6
191 -     LL{i}=num2str(exp(TT(i)));
192 - end
193
194 - [Apo]=Krigrids( Xk, Yk, KREST);
195 - figure(12);
196 - colorbar('YTick', TT, 'YTickLabel', LL);
197 - ylim([0 17000]);
198 - xlim([0 14000]);
199
200
201
202

```

## Bibliography

- [1] In: (). URL: [https://environmental.hellas-gold.com/FEK/FEK\\_B\\_3322-30.12.2011.pdf](https://environmental.hellas-gold.com/FEK/FEK_B_3322-30.12.2011.pdf).
- [2] Marcel Abel. “Kriging Interpolation. Rstudio, QGIS and ArcMap Geostatistical Analysis. Groundwater Flow. Video introduction about the Geostatistical Analysis Course.” In: (2020). URL: [https://www.youtube.com/watch?v=0dAVkBo\\_hKY&list=PL5B4v-RS9DkWv\\_5lDo6IaxnZN3p4UrbT8](https://www.youtube.com/watch?v=0dAVkBo_hKY&list=PL5B4v-RS9DkWv_5lDo6IaxnZN3p4UrbT8).
- [3] Ashraf Ali. “Removal of Mn(II) from water using chemically modified banana peels as efficient adsorbent”. In: *Environmental Nanotechnology, Monitoring & Management* 7 (2017), pp. 57–63. ISSN: 2215-1532. DOI: <https://doi.org/10.1016/j.enmm.2016.12.004>. URL: <https://www.sciencedirect.com/science/article/pii/S2215153216300873>.
- [4] Franz Aurenhammer. “Voronoi diagrams—a survey of a fundamental geometric data structure”. In: *ACM Computing Surveys (CSUR)* 23.3 (1991), pp. 345–405.
- [5] Tianfeng Chai and Roland R Draxler. “Root mean square error (RMSE) or mean absolute error (MAE)”. In: *Geoscientific Model Development Discussions* 7.1 (2014), pp. 1525–1534.
- [6] J.P. Chiles and P. Delfiner. *Geostatistics: Modeling Spatial Uncertainty*. Hoboken, NJ, USA: Wiley, 2012.
- [7] National Research Council. *Hardrock Mining on Federal Lands*. The National Academies Press, 1999. DOI: 10.17226/9682. URL: <https://nap.nationalacademies.org/catalog/9682/hardrock-mining-on-federal-lands>.

## Bibliography

- [8] Carlos D Da Rosa, James S Lyon, and Philip M Hocker. *Golden dreams, poisoned streams: how reckless mining pollutes America's waters, and how we can stop it*. Mineral Policy Center, 1997.
- [9] Ali Akbar Daya and Hadi Bejari. "A comparative study between simple kriging and ordinary kriging for estimating and modeling the Cu concentration in Chehlkureh deposit, SE Iran". In: *Arabian journal of Geosciences* 8.8 (2015), pp. 6003–6020.
- [10] Esri. "Understanding a semivariogram: The range, sill, and nugget". In: (). URL: <https://pro.arcgis.com/en/pro-app/2.8/help/analysis/geostatistical-analyst/understanding-a-semivariogram-the-range-sill-and-nugget.htm>.
- [11] Hellas Gold. "Πρόγραμμα Περιβαλλοντικής Παρακολούθησης". In: (). URL: <https://environmental.hellas-gold.com/sdi.php?configfile=configWaterGround&long=23.807395697698&lat=40.528559512711&zoom=11>.
- [12] Anne-Monique Gounot. "Microbial oxidation and reduction of manganese: Consequences in groundwater and applications". In: *FEMS Microbiology Reviews* 14.4 (1994), pp. 339–349. ISSN: 0168-6445. URL: <https://www.sciencedirect.com/science/article/pii/0168644594900523>.
- [13] A Gouzinis et al. "Removal of Mn and simultaneous removal of NH<sub>3</sub>, Fe and Mn from potable water using a trickling filter". In: *Water Research* 32.8 (1998), pp. 2442–2450. ISSN: 0043-1354. DOI: [https://doi.org/10.1016/S0043-1354\(97\)00471-5](https://doi.org/10.1016/S0043-1354(97)00471-5). URL: <https://www.sciencedirect.com/science/article/pii/S0043135497004715>.
- [14] Tomislav Hengl, Gerard BM Heuvelink, and David G Rossiter. "About regression-kriging: From equations to case studies". In: *Computers & geosciences* 33.10 (2007), pp. 1301–1315.
- [15] Dionissios T. Hristopulos. *Random Fields for Spatial Data Modeling: A Primer for Scientists and Engineers*. Dordrecht, the Netherlands: Springer/Nature, 2020. DOI: 10.1007/978-94-024-1918-4.

## Bibliography

- [16] Cheng-Shin Jang and Chen-Wuing Liu. “Geostatistical analysis and conditional simulation for estimating the spatial variability of hydraulic conductivity in the Choushui River alluvial fan, Taiwan”. In: *Hydrological Processes* 18.7 (2004), pp. 1333–1350.
- [17] Wolfgang J Junk et al. “Current state of knowledge regarding the world’s wetlands and their future under global climate change: a synthesis”. In: *Aquatic sciences* 75.1 (2013), pp. 151–167.
- [18] MFGA Lloret-Cabot, Gordon A Fenton, and Michael A Hicks. “On the estimation of scale of fluctuation in geostatistics”. In: *Georisk: Assessment and management of risk for engineered systems and geohazards* 8.2 (2014), pp. 129–140.
- [19] Robert McClure and S Andrew. “More than a century of mining has left the West deeply scarred”. In: *The Seattle Post Intelligencer Publication* (2001), pp. 1–6.
- [20] A. Pavlides et al. “Spatial modeling of lignite energy reserves for exploitation planning and quality control”. In: *Energy* 93 (Dec. 2015), pp. 1906–1917. DOI: 10.1016/j.energy.2015.10.049.
- [21] Karlheinz Spitz and John Trudinger. *Mining and the environment: from ore to metal*. CRC Press, 2019.
- [22] Hans Wackernagel. *Multivariate geostatistics: an introduction with applications*. Springer Science & Business Media, 2003.
- [23] Μαντέλα Αργύρη. “Βαρέα Μέταλλα στην Υδρολογική Λεκάνη Του Κερίτη”. In: (2007).
- [24] Παπαδάκη Μαρία Βάσση Παρασκευή. “ΠΡΟΣΔΙΟΡΙΣΜΟΣ ΒΑΡΕΩΝ ΜΕΤΑΛΛΩΝ ΜΟΛΥΒΔΟΥ (Pb), ΜΑΓΓΑΝΙΟΥ (Mn), ΣΙΔΗΡΟΥ (Fe) και ΑΡΣΕΝΙΚΟΥ (As) ΣΕ ΠΟΣΙΜΟ ΝΕΡΟ ΤΟΥ ΝΟΜΟΥ ΤΡΙΚΑΛΩΝ”. In: (2013).
- [25] Παναγιώτα Γκάφα. “Γεωστατιστική ανάλυση αποθεμάτων γαιάνθρακα βάσει δεδομένων από την κομητεία Campbell των ΗΠΑ”. In: (2017). DOI: <https://doi.org/10.26233/heallink.tuc.67469>.

## Bibliography

- [26] Γενική Διεύθυνση Περιβάλλοντος Ευρωπαϊκή Επιτροπή. *Προστασία των υπόγειων υδάτων στην Ευρώπη*. Εκδόσεις Ευρωπαϊκής Ένωσης, 2008.
- [27] Χαρίκλεια-Παναγιώτης Θηβαίου. “Σύγκριση προϊόντων κατακρήμνισης από γεωστατιστικές και τηλεπισκοπικές μεθόδους για την περιοχή της δυτικής Αυστραλίας”. In: (2010), pp. 85–87. URL: <http://hdl.handle.net/11610/6181>.
- [28] Καθ. Αθανάσιος Λουκάς. “Υδρολογική Προσομοίωση και Πρόγνωση - Ενότητα 4: Γεωστατιστική”. In: *Εργαστήριο Υδρολογίας και Ανάλυσης Υδατικών Συστημάτων Τμήμα Πολιτικών Μηχανικών Πολυτεχνική Σχολή Θεσσαλίας* ().
- [29] Δ/νων Σύμβουλος ΕΧ Ομιλία Δ. Δημητριάδης. “VIII ΣΕΒΕ Export Summit, Awareness: A Key Factor to Success”. In: *HellasGoldSA* (2021). URL: <https://www.youtube.com/watch?v=CtWiG3vKTB4&t=378s>.
- [30] Ευστρατία Η. Χαριτίδου. “Επιλογή Οικογένειας Μετασχηματισμών Σε Μπεϋζιανά Στατιστικά Μοντέλα: Μεθοδολογία Και Εφαρμογές”. In: (2016). URL: [https://dspace.lib.ntua.gr/xmlui/bitstream/handle/123456789/44453/PhD\\_Efstratia\\_Charitidou\\_2016.pdf?sequence=1](https://dspace.lib.ntua.gr/xmlui/bitstream/handle/123456789/44453/PhD_Efstratia_Charitidou_2016.pdf?sequence=1).
- [31] Πηνελόπη Χατζηγεωργίου-Σταυράκη. “Impact of the heavy metals contamination in mining activity areas to the health of the local population”. In: (1998). DOI: 10.12681/eadd/12074.
- [32] Διονύσιος Θ. Χριστόπουλος. *Εισαγωγή στις Πιθανότητες και τη Στατιστική για Μηχανικούς*. 2016. URL: [https://www.eclass.tuc.gr/modules/document/file.php/MHXOP139/Probability\\_Statistics\\_2016.pdf](https://www.eclass.tuc.gr/modules/document/file.php/MHXOP139/Probability_Statistics_2016.pdf).
- [33] Διονύσιος Θ. Χριστόπουλος. *Εφαρμοσμένη Γεωστατιστική*. 2014. URL: <https://www.eclass.tuc.gr/modules/document/file.php/MHXOP178/Applied%20Geostats%202020.pdf>.