Elliptical Anisotropy Statistics on Spatial Data and Geostatistical Applications

Technical University of Crete

Manolis P. Petrakis

Department of Mineral Resources Engineering

Technical University of Crete

A thesis submitted for the degree of

Master of Sciences

21 December 2012

Advisory Committee:

D.T. Hristopulos (advisor), Dept. of Mineral Resources Engineering, TUC

Z.G. Agioutantis, Dept. of Mineral Resources Engineering, TUC

M.G. Galetakis, Dept. of Mineral Resources Engineering, TUC

Judgment is a tool to use on all subjects, and comes in everywhere. Therefore in the essays that I make of it here, I use every sort of occasion. If it is a subject I do not understand at all, even on that I essay my judgement, sounding the ford from a good distance; and then, finding it too deep for my height, I stick to the bank. And this acknowledgement that I cannot cross over is a token of its action, indeed one of those it is most proud of. Sometimes in a vain and nonexistent subject I try to see if it will find the wherewithal to give it body, prop it up, and support it. Sometimes I lead it to a noble and well-worn subject in which it has nothing original to discover, the road being so beaten that it can walk only in others' footsteps. There it plays its part by choosing the way that seems best to it, and of a thousand paths it says that this one or that was the most wisely chosen.

I take the first subject that chance offers. They are all equally good to me. And I never plan to develop them completely. For I do not see the whole of anything; nor do those who promise to show it to us. Of a hundred members and faces that each thing has, I take one, sometimes only to lick it, sometimes to brush the surface, sometimes to pinch it to the bone. I give it a stab, not as wide but as deep I know how. And most often I like to take them from some unaccustomed point of view. I would venture to treat some matter thoroughly, if I knew myself less well. Scattering a word here, there another, samples separated from their context, dispersed, without a plan and without a promise, I am not bound to make something of them or to adhere to them myself without varying when I please and giving myself up to doubt and uncertainty and my ruling quality, which is ignorance.

— MICHEL DE MONTAIGNE, ESSAYS, "OF DEMOCRITUS AND HERACLITUS" (1572–1580) de Montaigne [1958]

Περίληψη

Τα διδιάστατα χωρικά δεδομένα συχνά έχουν συναρτήσεις αυτο-συσχέτισης με ελλειπτικές ισοϋψείς καμπύλες, μια ιδιότητα γνω-στή ως στατιστική ανισοτροπία. Οι παράμετροι ανισοτροπίας πε-ριλαμβάνουν την κλίση της έλλειψης (ή γωνία προσανατολισμού) ϑ σε σχέση με το σύστημα συντεταγμένων, και τον λόγο R των κύριων μηκών συσχέτισης. Η εκτίμηση της ανισοτροπίας είναι σημαντική για εφαρμογές στις γεωεπιστήμες, στην επεξεργασία σήματος και εικόνας, καθώς και σε ιατρικές εφαρμογές. Οι δειγ-ματικές εκτιμήσεις των παραμέτρων ανισοτροπίας είναι χρήσιμες για τον ορισμό κατάλληλων χωρικών μοντέλων και για τη χωρική παρεμβολή σε σύνολα χωρικά διάσπαρτων δεδομένων.

Επί του παρόντος, η εκτίμηση της ανισοτροπίας γίνεται είτε με τη μέθοδο της μέγιστης πιθανοφάνειας, η οποία είναι υπολογιστικά απαιτητική, ή με την εξέταση εμπειρικών βαριογραμμάτων από εμπειρογνώμονες, η οποία μπορεί να οδηγήσει σε υποκειμενικά αποτελέσματα. Μη επιβλεπόμενες και υπολογιστικά αποτελεσμα-τικές μέθοδοι εκτίμησης της ανισοτροπίας είναι απαραίτητες για τη γεωστατιστική ανάλυση, καθώς και για συστήματα αυτοματο-ποιημένης παρατήρησης και έγκαιρης προειδοποίησης σε περι-πτώσεις περιβαλλοντικών απειλών.

Σε αρκετά επιστημονικά πεδία, περιλαμβανομένων και των γεω-επιστημών, είναι διαθέσιμο μόνο ένα δείγμα δεδομένων, εξαιτίας του κόστους των μετρήσεων ή άλλων περιορισμών. Επομένως, ένα πρόβλημα με πρακτικό ενδιαφέρον είναι η εκτίμηση των παραμέ-τρων ανισοτροπίας από ένα δείγμα, το οποίο μπορεί επίσης να έχει ακανόνιστη χωρική κατανομή. Πρόσφατα προτάθηκε μια υπολο-γιστικά αποτελεσματική εκτιμήτρια ανισοτροπίας για δείγματα με κανονικές ή ακανόνιστες χωρικές κατανομές, η οποία βασίζεται στην Ταυτότητα Εσσιανής της Συνδιακύμανσης του Swerling.

Σε αυτήν την εργασία, βασιζόμενοι στην Ταυτότητα Εσσιανής της Συνδιακύμανσης αναπτύσσεται —με αναλυτικούς υπολογισμούς— μια ρητή έκφραση για την συνάρτηση πυκνότητας πιθανότητας

$f_{\partial,R}(\hat{R}, \hat{\partial})$, η οποία να ισχύει για γκαουσιανά, στατιστικά ομοιο-γενή και παραγωγίσιμα τυχαία πεδία. Η συνάρτηση πυκνότητας πιθανότητας $f_{\partial,R}(\hat{R}, \hat{\partial})$ χαρακτηρίζει τη κατανομή των εκτιμήσεων ανισοτροπίας $(\hat{R}, \hat{\partial})$ της εκτιμήτριας. Στη συνέχεια, βασιζόμενοι στη συνάρτηση πυκνότητας πιθανότητας $f_{\partial,R}(\hat{R}, \hat{\partial})$, υπολογίζουμε μια προσεγγιστική έκφραση $f_{\partial,R}^{(0)}(\hat{R}, \hat{\partial})$ η οποία είναι ανεξάρτητη από την συνάρτηση αυτοσυσχέτισης και παρέχει συντηρητικές πε-ριοχές εμπιστοσύνης για τις εκτιμώμενες παραμέτρους. Επίσης προτείνουμε ένα στατιστικό έλεγχο ισοτροπίας βασιζόμενο στην προσεγγιστική συνάρτηση πυκνότητας πιθανότητας.

Η θεωρητική ανάλυση επαληθεύεται από προσομοιώσεις που πε-ριλαμβάνουν δεδομένα σε πλέγμα, αλλά και χωρικά διάσπαρτα δεδομένα. Παρουσιάζουμε τη χρήση των περιοχών εμπιστοσύνης σε δυο σενάρια με πραγματικά δεδομένα, προερχόμενα από λι-γνιτωρυχεία, και σε ένα σενάριο περιβαλλοντολογικού κινδύνου που περιλαμβάνει την εκπομπή ακτινοβολίας γάμμα εξαιτίας α-τυχήματος.

Τα αποτελέσματα αυτής της εργασίας μπορούν να χρησιμοποιη-θούν για να παρέχουν (α) μια αυτόνομη προσέγγιση της κατανο-μής των $(\hat{R}, \hat{\partial})$, (β) αρχικές τιμές ανισοτροπίας με υπολογιστικά αποτελεσματικό τρόπο για την περαιτέρω βελτίωση της εκτίμησης με τη μέθοδο της μέγιστης πιθανοφάνειας και (γ) μια χρήσιμη εκ των προτέρων (a priori) συνάρτηση πυκνότητας πιθανότητας για εκτίμηση ανισοτροπίας με τη μέθοδο Bayes.

Εν κατακλείδι, η εργασία αυτή παρουσιάζει ένα πλαίσιο για την αποτελεσματική ποσοτικοποίηση της στατιστικής σημαντικότητας αλλαγών στην ανισοτροπία χωρικών δεδομένων. Η ανίχνευση στατιστικά σημαντικής ανισοτροπίας σε προκαταρκτικά δεδομένα προερχόμενα από διερευνητικές γεωτρήσεις είναι σημαντική για τον σχεδιασμό και την εκμετάλλευση ορυχείων, επειδή επιτρέπει μεγαλύτερη ακρίβεια στην ανάλυση της χωρικής μεταβλητότητας του κοιτάσματος. Δεδομένου ότι μια σημαντική αλλαγή στην α-νισοτροπία μπορεί να συνεπάγεται σημαντική μεταβολή του φυ-σικού συστήματος, όπως για παράδειγμα έκλυση ρύπων στο πε-ριβάλλον λόγω ατυχήματος, προτείνουμε την ενσωμάτωση της α-νίχνευσης ανισοτροπίας σε συστήματα έγκαιρης προειδοποίησης. Όσον αφορά τις ιατρικές εφαρμογές, με βάση τα αποτελέσματα της παρούσας εργασίας, μπορεί να υλοποιηθεί ένας αλγόριθμος βασισμένος στην ανισοτροπία για την ανίχνευση ανωμαλιών σε δεδομένα από ψηφιακές μαστογραφίες.

Εκτενής Περίληψη

Στη παρούσα εργασία εξετάζονται οι στατιστικές ιδιότητες μιας νέ-
ας μεθόδου εκτίμησης στατιστικής ανισοτροπίας. Υπολογίζεται
αναλυτικά η μαθηματική έκφραση της κοινής συνάρτησης πυ-
κνότητας πιθανότητας των εκτιμήσεων στατιστικής ανισοτροπίας
και η έκφραση των περιοχών εμπιστοσύνης. Επίσης υπολογίζεται
αναλυτικά η μη-παραμετρική έκφραση της κοινής συνάρτησης
πιθανότητας και των περιοχών εμπιστοσύνης, όπως επίσης και έ-
νας μη-παραμετρικός έλεγχος στατιστικής ανισοτροπίας. Κατόπιν
παρουσιάζονται εφαρμογές σε συνθετικά και πραγματικά δεδομέ-
να.

Στις γεωεπιστήμες συχνά είναι διαθέσιμο μόνο ένα δείγμα χωρικά
ατάκτων δεδομένων, όπως για παράδειγμα ένα σύνολο μετρήσεων
συγκέντρωσης ενός ρύπου στην ατμόσφαιρα. Οι σημαντικότερες
αιτίες που περιορίζουν τη λήψη μετρήσεων κατά το δοκούν, είναι
το κόστος, αλλά και το ενδεχόμενο τα σημεία δειγματοληψίας να
είναι δυσπρόσιτα. Παρόμοιοι περιορισμοί υπάρχουν και σε άλλα
επιστημονικά πεδία, όπως στην επεξεργασία σήματος, ή σε ιατρι-
κές εφαρμογές. Η εκτίμηση των τιμών της μετρήσιμης ιδιότητας
στα σημεία που δεν υπάρχουν μετρήσεις, πραγματοποιείται με
χωρική παρεμβολή. Οι μέθοδοι χωρικής παρεμβολής μπορούν
να κατηγοριοποιηθούν σε δυο μεγάλες κλάσεις, τις *αιτιοκρατικές*,
και τις *στοχαστικές μεθόδους παρεμβολής*. Η οικογένεια μεθό-
δων γνωστή στις γεωεπιστήμες ως *kriging* βασίζεται στη θεωρία
των *τυχαίων πεδίων* και ανήκει στις *στοχαστικές* μεθόδους, οι ο-
ποίες έχουν το προσόν ότι προσφέρουν *εκτίμηση των σφαλμάτων
πρόβλεψης*.

Η θεωρία των τυχαίων πεδίων είναι το μαθηματικό εργαλείο που
χρησιμοποιείται για τη περιγραφή χωρικών δεδομένων και για την
ανάπτυξη στοχαστικών μεθόδων χωρικής παρεμβολής. Μια μέ-
τρηση της φυσικής ιδιότητας αντιστοιχεί σε μια κατάσταση (state)
του τυχαίου πεδίου (random field) $X(\mathbf{s})$ η οποία είναι μια πολυ-
διάστατη τυχαία μεταβλητή στο σημείο με *διάνυσμα θέσης* $\mathbf{s}$. Η
συνάρτηση αυτοσυσχέτισης (autocorrelation function) η οποία

περιγράφει την χωρική συσχέτιση του τυχαίου πεδίου. Δυο συνή-
θεις συναρτήσεις αυτοσυσχέτισης είναι η γκαουσιανή (Υποπαρά-
γραφος A.7.1) και η συνάρτηση Matérn (Υποπαράγραφος A.7.2).

Συχνά η συνάρτηση αυτοσυσχέτισης διδιάστατων δεδομένων πα-
ρουσιάζει ελλειπτικές ισοϋψείς καμπύλες, το οποίο πρακτικά ση-
μαίνει ότι η χωρική συσχέτιση έχει εξάρτηση από την κατεύθυνση.
Η ιδιότητα αυτή, γνωστή ως *στατιστική ανισοτροπία* (statistical an-
isotropy), περιγράφεται με δυο παραμέτρους: την κλίση της έλ-
λειψης ϑ και τον λόγο R (anisotropy ratio) των κύριων αξόνων της,
ο οποίος είναι ίσος και με τον λόγο των δυο κύριων μηκών συσχέ-
τισης (correlation lengths) ξ_1 και ξ_2, $R = \xi_2/\xi_1$ (βλ. Σχήμα 1.5).
Παραδείγματα ισοτροπικών και ανισοτροπικών *συνθετικών* τυχαί-
ων πεδίων με γκαουσιανή και Matérn συνάρτηση αυτοσυσχέτισης
παρουσιάζονται στα Σχήματα 2.1 και 2.2.

Παρόλο που είναι εμφανές ότι αρκετές φυσικές διεργασίες είναι
στατιστικά ανισότροπες, συχνά μοντελοποιούνται με ισοτροπικές
συναρτήσεις αυτοσυσχέτισης. Αγνοώντας τη στατιστική ανισοτρο-
πία, αποφεύγουμε την εκτίμηση δυο επιπλέον παραμέτρων, των
$\hat{R}$ και $\hat{\vartheta}$). Συνήθως η εκτίμηση των παραμέτρων της στατιστικής
ανισοτροπίας πραγματοποιείται με μεθόδους όπως η επισκόπη-
ση εμπειρικών βαριογραμμάτων, μια διαδικασία υποκειμενική,
γιατί υπεισέρχεσαι ο ανθρώπινος παράγοντας ο οποίος κρίνει αν
το εμπειρικό βαριόγραμμα είναι ισοτροπικό ή ανισοτροπικό, και
η υπολογιστικά απαιτητική μέθοδος της *μέγιστης πιθανοφάνειας*
(maximum likelihood). Επομένως οι παραπάνω μέθοδοι δεν είναι
ιδανικές για αυτοματοποιημένες εφαρμογές. Χρησιμοποιώντας ι-
σοτροπικά μοντέλα σε περιπτώσεις όπου είναι εμφανής η ύπαρξη
ανισοτροπίας στα δεδομένα, οδηγούμαστε σε εκτιμήσεις υποδεέ-
στερης ακρίβειας.

Μια νέα, υπολογιστικά αποτελεσματική μέθοδος εκτίμησης στα-
τιστικής ανισοτροπίας προτάθηκε πρόσφατα. Η μέθοδος αυτή
βασίζεται στην Ταυτότητα Εσσιανής της Συνδιακύμανσης (Covari-
ance Hessian Identity, Θεώρημα 2.1.1). Συνδέει με αναλυτικό
τρόπο τις παραμέτρους ανισοτροπίας με παραμέτρους που μπο-
ρούν να εκτιμηθούν από τις *χωρικές παραγώγους* του δείγματος
(Θεώρημα 2.1.2). Το θεώρημα απαιτεί να τηρούνται δύο προ-
ϋποθέσεις: Το δείγμα να είναι αρκετά μεγάλο ώστε να μπορεί να
χρησιμοποιηθεί το *Εργοδικό Θεώρημα* και να υπάρχουν οι χω-
ρικές παράγωγοι του πεδίου. Το Εργοδικό Θεώρημα επιτρέπει
την προσέγγιση μέσων τιμών που υπολογίζονται σε όλες τις κατα-

στάσεις του τυχαίου πεδίου, τις οποίες ποτέ δε διαθέτουμε στην πράξη, με μέσες τιμές υπολογισμένες από το μοναδικό δείγμα που συνήθως διαθέτουμε. Αν $|\mathcal{D}|$ είναι το εμβαδόν της περιοχής και $\bar{a}$ η μέση απόσταση μεταξύ των σημείων δειγματοληψίας οι δυο προϋποθέσεις συνοψίζονται ως εξής:

$$|\mathcal{D}| \gg \max(\xi_1, \xi_2) \quad \text{και} \quad \bar{a} \ll \min(\xi_1, \xi_2).$$

Δηλαδή στο δείγμα πρέπει να περιέχονται αρκετές συσχετισμένες υποπεριοχές, και η δειγματοληψία πρέπει να είναι αρκετά πυκνή ώστε να καταγράφεται επαρκώς η χωρική μεταβλητότητα του τυχαίου πεδίου, για τον αξιόπιστο υπολογισμό των χωρικών παραγώγων από το δείγμα.

Στη παρούσα εργασία μελετώνται οι *στατιστικές ιδιότητες* της προαναφερθείσας εκτιμήτριας στατιστικής ανισοτροπίας. Πιο συγκεκριμένα, υπολογίζονται με αναλυτικό τρόπο, η κοινή συνάρτηση πυκνότητας πιθανότητας (joint probability density function) $f_{\partial,R}(\hat{R}, \hat{\partial})$ των εκτιμήσεων ανισοτροπίας και οι αντίστοιχες περιοχές εμπιστοσύνης (confidence regions). Προκύπτει μια ρητή έκφραση της $f_{\partial,R}(\hat{R}, \hat{\partial})$ και των περιοχών εμπιστοσύνης, που όμως προϋποθέτουν τη γνώση της συνάρτησης αυτοσυσχέτισης. Γιαυτό, βασιζόμενοι στη συνάρτηση πυκνότητας πιθανότητας υπολογίζουμε μια προσεγγιστική έκφραση $f_{\partial,R}^{(0)}(\hat{R}, \hat{\partial})$ η οποία είναι *μη-παραμετρική* (non-parametric), δηλαδή δεν απαιτεί γνώση της συνάρτησης αυτοσυσχέτισης. Η $f_{\partial,R}^{(0)}(\hat{R}, \hat{\partial})$ εξαρτάται μονάχα από το μέγεθος του δείγματος N και από την εκτίμηση των παραμέτρων ανισοτροπίας.

Η διαδικασία που ακολουθήθηκε για τον υπολογισμό της $f_{\partial,R}(\hat{R}, \hat{\partial})$ μπορεί να συνοψιστεί στα εξής: Δοθέντων των αναλυτικών εκφράσεων (Θεώρημα 2.1.2) που συνδέουν τις παραμέτρους ανισοτροπίας με τις παραμέτρους Q_{ij} που μπορούν να εκτιμηθούν από το δείγμα, αν είναι γνωστή η κατανομή $f_{\hat{\mathbb{Q}}}(\hat{Q}_{11}, \hat{Q}_{22}, \hat{Q}_{12})$ των εκτιμήσεων $\hat{Q}_{ij}$, εκτελώντας τους απαραίτητους μετασχηματισμούς πιθανότητας (*Θεώρημα Jacobi*, Παράρτημα A.1), μπορεί να υπολογισθεί η κατανομή $f_{\partial,R}(\hat{R}, \hat{\partial})$ των εκτιμήσεων $(\hat{R}, \hat{\partial})$. Η ακολουθία των μετασχηματισμών απαιτεί ένα ενδιάμεσο βήμα, τον υπολογισμό της συνάρτησης πυκνότητας πιθανότητας $f_{\mathbf{q}}(\hat{q}_d, \hat{q}_o)$ των λόγων $(\hat{q}_d, \hat{q}_o)$, όπως και κάποιες προσεγγίσεις για να είναι εφικτός ο *αναλυτικός* υπολογισμός των κατανομών και για τον αναλυτικό υπολογισμό μη-παραμετρικών αποτελεσμάτων. Το Θεώρημα Jacobi εκφράζει τη διατήρηση πιθανότητας κατά την πραγματο-

ποίηση αλλαγών μεταβλητής. Η διατήρηση της πιθανότητας κατά την ακολουθία των μετασχηματισμών:

$$f_{\bar{\mathbb{Q}}}(\hat{Q}_{11}, \hat{Q}_{22}, \hat{Q}_{12}) \Rightarrow f_{\mathbf{q}}(\hat{q}_d, \hat{q}_o) \Rightarrow f_{\partial, R}(\hat{R}, \hat{\partial}),$$

αναπαρίσταται στο Σχήμα 2.8.

Επειδή οι δειγματικές εκτιμήσεις των Q_{ij} προκύπτουν από ένα ά-θροισμα (2.18) γινομένων χωρικών παραγώγων X_{ij} πάνω σε όλα τα σημεία δειγματοληψίας $\mathbf{s}_k$, μπορεί να εφαρμοσθεί το *Κεντρι-κό Οριακό Θεώρημα*, στην εκδοχή του για ελαφρά συσχετισμένες τυχαίες μεταβλητές, όπως είναι τα γινόμενα χωρικών παραγώγων X_{ij} (2.12) από τα οποία προκύπτουν τα Q_{ij}. Οι προσομοιώσεις σε υπολογιστή επαληθεύουν την υπόθεση αυτή, όπως φαίνεται στο Σχήμα 4.3. Επομένως οι εκτιμήσεις των Q_{ij} ακολουθούν μια γκα-ουσιανή κατανομή τριών μεταβλητών, για την οποία απαιτείται ο υπολογισμός του πίνακα συνδιασποράς $\mathbf{C}_{\bar{\mathbb{Q}}}$, όπως αυτός ορίζεται στο Λήμμα 2.2.1. Το καθένα από τα έξι στοιχεία του 3×3 συμ-μετρικού πίνακα $\mathbf{C}_{\bar{\mathbb{Q}}}$ υπολογίζονται αθροίζοντας πάνω σε όλα τα N^2 διανύσματα $\mathbf{r}_{nm} = \mathbf{s}_n - \mathbf{s}_m$, όπου $\mathbf{s}_n$ είναι τα διανύσματα θέ-σης και $n, m = 1, \dots, N$. Οι συναρτήσεις $C_{ij,kl}(\mathbf{r})$ του αθροίσματος εξαρτώνται από τις παραγώγους της συνάρτησης αυτοσυσχέτισης.

Η διαδοχή των αλλαγών μεταβλητών που προαναφέρθηκαν, πραγ-ματοποιείται με αναλυτικό τρόπο. Προκύπτει ότι η *παραμετρική* κατανομή εκτιμήσεων ανισοτροπίας $f_{\partial, R}(\hat{R}, \hat{\partial})$ έχει την εξής μορ-φή:

$$f_{\partial, R}(\hat{R}, \hat{\partial}; \mathbf{m}_{\mathbb{Q}}, \mathbf{C}_{\bar{\mathbb{Q}}}) = \frac{2\hat{R}\left|\hat{R}^2 - 1\right|}{\left(\hat{R}^2 \cos^2 \hat{\partial} + \sin^2 \hat{\partial}\right)^3} f_{\mathbf{q}}(\hat{\mathbf{q}}; \mathbf{m}_{\mathbb{Q}}, \mathbf{C}_{\bar{\mathbb{Q}}}),$$

όπου
$$f_{\mathbf{q}}(\hat{\mathbf{q}}; \mathbf{m}_{\mathbb{Q}}, \mathbf{C}_{\bar{\mathbb{Q}}}) = 2\sqrt{\pi}\, \tilde{K}\left(2\tilde{B}^2\, N^2 + 1\right) e^{N^2(\tilde{B}^2 - \tilde{C})},$$

(Θεώρημα 2.2.2). Η *παραμετρική* περιοχή εμπιστοσύνης δίδεται από την καμπύλη $\hat{R}, \hat{\partial}$ που ικανοποιούν την εξίσωση:

$$\left[\tilde{B}^2(\hat{\mathbf{q}}; \mathbf{m}_{\mathbb{Q}}, \mathbf{C}_{\bar{\mathbb{Q}}}) - \tilde{C}(\hat{\mathbf{q}}; \mathbf{m}_{\mathbb{Q}}, \mathbf{C}_{\bar{\mathbb{Q}}})\right] N^2 = \ln(1 - p),$$

(Πόρισμα 2.2.1), όπου p η στάθμη εμπιστοσύνης (confidence level), π.χ. $p = 0.95$. Οι συντελεστές $\tilde{B}, \tilde{K}, \tilde{A}$ εξαρτώνται από τη συνάρτηση αυτοσυσχέτισης.

Κρατώντας μόνο τον πρώτο όρο του αθροίσματος (2.18), που αν-

τιστοιχεί στα διανύσματα $\mathbf{r} = \mathbf{0}$, προκύπτει η *μη-παραμετρική* κατανομή $f_{\partial,R}^{(0)}(\hat{R}, \hat{\partial})$ (Θεώρημα 3.1.1) για την οποία *δεν* απαιτείται εκτίμηση παραμέτρων της συνάρτησης αυτοσυσχέτισης. Στα Σχήματα 3.1 και 3.2 παρουσιάζονται οι γραφικές παραστάσεις των μη-παραμετρικών κατανομών για διάφορες τιμές ανισοτροπίας και πλήθους δεδομένων N, οι οποίες είναι ασύμμετρες και μη-γκαουσιανές. Όπως είναι αναμενόμενο, για μεγάλο μέγεθος δείγματος, προκύπτουν στενότερες μη-παραμετρικές κατανομές και μικρότερες μη-παραμετρικές περιοχές εμπιστοσύνης, το οποίο σημαίνει ότι οι εκτιμήσεις ανισοτροπίας έχουν μικρότερη αβεβαιότητα για μεγάλα N.

Η *μη-παραμετρική* κατανομή εκτιμήσεων ανισοτροπίας έχει την εξής μορφή:

$$f_{\partial,R}^{(0)}(\hat{R}, \hat{\partial}; R, \partial, N) = \frac{2\hat{R}\left|\hat{R}^2 - 1\right|}{\left(\hat{R}^2 \cos^2 \hat{\partial} + \sin^2 \hat{\partial}\right)^3} f_{\mathbf{q}}^{(0)}(\hat{R}, \hat{\partial}; R, \partial, N),$$

όπου

$$\tilde{B}_0 = (2\tilde{A}_0)^{-1/2}\left[(R^2 - 1)(\hat{R}^2 - 1)\cos(2(\partial - \hat{\partial})) - (R^2 + 1)(\hat{R}^2 + 1)\right],$$
$$\tilde{K}_0 = (\pi\tilde{A}_0)^{-3/2}R^3\left[(\hat{R}^2 + 1) - (\hat{R}^2 - 1)\cos(2\hat{\partial})\right]^3,$$

όπου

$$\begin{aligned}
\tilde{A}_0 = {} & (\hat{R}^4 + 1)(3R^4 + 2R^2 + 3) + 2\hat{R}^2(R^2 - 1)^2 \\
& + (\hat{R}^2 - 1)^2(R^2 - 1)^2 \cos(4(\hat{\partial} - \partial)) \\
& \qquad - 4(\hat{R}^4 - 1)(R^4 - 1)\cos(2(\hat{\partial} - \partial)).
\end{aligned}$$

Η κατανομή $f_{\partial,R}^{(0)}(\hat{R}, \hat{\partial})$ εξαρτάται μονάχα από τις παραμέτρους ανισοτροπίας R, ∂ και το μέγεθος του δείγματος N. Η *μη-παραμετρική* περιοχή εμπιστοσύνης δίδεται από την καμπύλη που προκύπτει από τα $\hat{R}, \hat{\partial}$ που ικανοποιούν την εξίσωση:

$$\tilde{B}_0^2 - \frac{1}{2} = \frac{\ln(1 - p)}{N}$$

Η περιοχή εμπιστοσύνης εξαρτάται από το επίπεδο εμπιστοσύνης p και το μέγεθος του δείγματος N. Από τη μη-παραμετρική έκφραση των περιοχών εμπιστοσύνης συνεπάγεται ο ακόλουθος μη-παραμετρικός *στατιστικός έλεγχος ισοτροπίας*. Δοθέντος του

μεγέθους δείγματος N και του βαθμού εμπιστοσύνης p, προκύπτει το διάστημα ανισοτροπικών λόγων $(\hat{R}_-, \hat{R}_+)$ για το οποίο οι όποιες εκτιμήσεις ανισοτροπικού λόγου $\hat{R}$ μπορούν να θεωρηθούν σε στάθμη εμπιστοσύνης p ότι αντιστοιχούν σε στατιστικά ισοτροπικά πεδία. Δηλαδή αν το $\hat{R}$ ανήκει στο ακόλουθο διάστημα:

$$\hat{R} \in \left[\sqrt{\frac{1 - 2\sqrt{a_{p;N}(1 - a_{p;N})}}{1 - 2a_{p;N}}}, \sqrt{\frac{1 + 2\sqrt{a_{p;N}(1 - a_{p;N})}}{1 - 2a_{p;N}}} \right],$$

όπου $\ell_p = -2\ln(1 - p)$ και $a_{p;N} = \ell_p/N$, τότε το δείγμα είναι στατιστικά ισοτροπικό σε στάθμη εμπιστοσύνης p (Θεώρημα 3.2.1).

Στη συνέχεια, το θεωρητικό πλαίσιο αναφορικά με τις στατιστικές ιδιότητες των εκτιμήσεων ανισοτροπίας ελέγχεται με προσομοιώσεις Monte Carlo οι οποίες πραγματοποιήθηκαν σε υπολογιστικό περιβάλλον MATLAB$^{\circledR}$, και περιλαμβάνουν δεδομένα σε πλέγμα (Ενότητα 4.1), δεδομένα εκτός πλέγματος (Ενότητα 4.2), αλλά και δεδομένα προερχόμενα από πραγματικές μετρήσεις. Πιο συγκεκριμένα, μελετήθηκαν δεδομένα ενός σεναρίου περιβαλλοντικού κινδύνου αναφορικά με την εκπομπή ακτινοβολίας γάμμα εξαιτίας ατυχήματος (Ενότητα 4.3), αλλά και δεδομένα προερχόμενα από διερευνητικές γεωτρήσεις σε δυο λιγνιτωρυχεία της Ελλάδας, στο Αμύνταιο και στη Μαυροπηγή (Ενότητα 4.4). Σε όλες τις περιπτώσεις εξετάστηκε αν οι περιοχές εμπιστοσύνης (στη μη-παραμετρική και στη παραμετρική τους εκδοχή, όπου ήταν εφικτό), περιγράφουν επαρκώς τη μεταβλητότητα των εκτιμήσεων ανισοτροπίας.

Στη περίπτωση των συνθετικών δεδομένων, είναι γνωστές εκ των προτέρων οι συναρτήσεις αυτοσυσχέτισης, επομένως μπορεί να υπολογισθεί αριθμητικά ο πίνακας $\mathbf{C}_{\bar{\mathbb{Q}}}$ έτσι ώστε να υπολογισθούν οι παραμετρικές περιοχές εμπιστοσύνης (2.31). Ο υπολογισμός του πίνακα $\mathbf{C}_{\bar{\mathbb{Q}}}$ γίνεται προσεγγιστικά ώστε να αποφευχθεί ο υπολογισμός αθροισμάτων με μεγάλο πλήθος όρων. Τα αθροίσματα 2.21 προσεγγίζονται αθροίζοντας σε $N' \ll N^2$ όρους εντός μιας γειτονιάς στην οποία οι αθροιζόμενες συναρτήσεις έχουν σημαντική συνεισφορά στο άθροισμα, Εξίσωση (2.24) και Σχήμα 2.4.

Στο Σχήμα 4.1 απεικονίζονται τα αποτελέσματα για τυχαία πεδία με ισοτροπική γκαουσιανή συνάρτηση αυτοσυσχέτισης, στο Σχήμα 4.2 τα αποτελέσματα για τυχαία πεδία με ισοτροπική συνάρτηση αυτοσυσχέτισης Matérn, ενώ στο Σχήμα 4.4 παρουσιάζονται

τα αποτελέσματα για ανισοτροπικές γκαουσιανές και Matérn συναρτήσεις αυτοσυσχέτισης. Στα σχήματα οι εκτιμήσεις ανισοτροπίας αναπαριστώνται με γαλάζιους σταυρούς στο επίπεδο $(\hat{R}, \hat{\vartheta})$ και οι παραμετρικές περιοχές εμπιστοσύνης με πράσινες καμπύλες. Οι *κόκκινες* καμπύλες αντιστοιχούν στις μη-παραμετρικές περιοχές εμπιστοσύνης. Παρατηρούμε ότι οι παραμετρικές περιοχές εμπιστοσύνης περιγράφουν πολύ καλά τη μεταβλητότητα των εκτιμήσεων ανισοτροπίας, αλλά όπως αναφέραμε προηγουμένως, απαιτούν τη γνώση της συνάρτησης αυτοσυσχέτισης. Οι μη-παραμετρικές περιοχές εμπιστοσύνης είναι ευρύτερες και περικλείουν τις παραμετρικές περιοχές εμπιστοσύνης. Η έλλειψη γνώσης της συνάρτησης αυτοσυσχέτισης σημαίνει έλλειψη πληροφορίας, γιαυτό οι μη-παραμετρικές περιοχές εμπιστοσύνης είναι ευρύτερες από τις παραμετρικές περιοχές εμπιστοσύνης

Για τη δημιουργία συνθετικών *διάσπαρτων* δεδομένων, επιλέχθηκε τυχαία ένα μικρό ποσοστό σημείων από ένα πολύ πυκνό δείγμα σε κανονικό πλέγμα ενός ισοτροπικού τυχαίου πεδίου με αυτοσυσχέτιση Matérn. Στη συνέχεια επιλέχθηκαν τυχαία υποσύνολα σημείων και πραγματοποιήθηκαν εκτιμήσεις ανισοτροπίας. Τα αποτελέσματα αναπαριστώνται στο Σχήμα 4.6 και συγκρίνονται επιτυχώς με τη μη-παραμετρική περιοχή εμπιστοσύνης.

Στη συνέχεια εξετάσθηκαν εφαρμογές σε πραγματικά δεδομένα. Η πρώτη εφαρμογή αφορά τις μετρήσεις ακτινοβολίας γάμμα από ένα δίκτυο αισθητήρων στη Γερμανία οι οποίες χρησιμοποιήθηκαν στην άσκηση σύγκρισης μεθόδων χωρικής παρεμβολής SIC2004. Τα δεδομένα περιλαμβάνουν δυο σενάρια: Το σενάριο κανονικών μετρήσεων (normal, δηλαδή εντός ορίων ασφαλείας), και το σενάριο ενός υποτιθέμενου ατυχήματος (emergency scenario). Τα δεδομένα παρουσιάζονται στο Σχήμα 4.7. Υπολογίζοντας τις μη-παραμετρικές κατανομές εκτιμήσεων ανισοτροπίας για τα δυο σενάρια προκύπτει ότι είναι *στατιστικά σημαντικά διαφορετικές*, όπως φαίνεται και στο Σχήμα 4.8. Επίσης εξετάσθηκε αν οι εκτιμήσεις ανισοτροπίας για τα δέκα σύνολα κανονικών μετρήσεων από διαφορετικές χρονικές στιγμές, μπορούν να θεωρηθούν στατιστικά διαφορετικές. Στο Σχήμα 4.10 και οι δέκα εκτιμήσεις ανισοτροπίας κείτονται εντός της μη παραμετρικής περιοχής εμπιστοσύνης η οποία έχει υπολογιστεί χρησιμοποιώντας τις μέσες τιμές των $\hat{Q}_{ij}$. Αυτό σημαίνει ότι οι εκτιμώμενες παράμετροι ανισοτροπίας και των δέκα συνόλων κανονικών τιμών *δεν* είναι στατιστικά διαφορετικές.

Η δεύτερη εφαρμογή αφορά δεδομένα προερχόμενα από τα λι-
γνιτωρυχεία του Αμυνταίου και της Μαυροπηγής, που βρίσκον-
ται κοντά στη Πτολεμαΐδα (Σχήματα 4.11 και 4.12). Πιο συγ-
κεκριμένα, εξετάσθηκαν μετρήσεις συνολικού πάχους λιγνίτη α-
πό διερευνητικές γεωτρήσεις που πραγματοποιήθηκαν με σκοπό
την εκτίμηση των γεωλογικών αποθεμάτων λιγνίτη στις συγκεκρι-
μένες περιοχές. Πραγματοποιήθηκε εκτίμηση ανισοτροπίας και
υπολογίσθηκε η μη-παραμετρική κατανομή των εκτιμήσεων ανι-
σοτροπίας (Σχήματα 4.13 και 4.15. Όπως και στη περίπτωση των
συνθετικών διάσπαρτων δεδομένων, επιλέχθηκαν τυχαία 1000 υ-
ποσύνολα με 50% των αρχικών μετρήσεων ώστε να προκύψουν
τεχνητά πολλά σύνολα μετρήσεων του ίδιου φυσικού φαινομένου
και να διαπιστωθεί αν η μεταβλητότητα των εκτιμήσεων ανισο-
τροπίας περιγράφεται επαρκώς από τη μη-παραμετρική περιοχή
εμπιστοσύνης (Σχήματα 4.14 και 4.16). Επιπροσθέτως, υπολο-
γίζονται με αριθμητική ολοκλήρωση των μη-παραμετρικών κατα-
νομών οι κατανομές των $\hat{R}$ και $\hat{\vartheta}$ και συγκρίνονται με αυτές που
εκτιμήθηκαν από το σύνολο των 1000 εκτιμήσεων ανισοτροπίας
για κάθε ορυχείο. Και για τα δυο ορυχεία, η μη-παραμετρική συ-
νάρτηση πυκνότητας πιθανότητας περιγράφει σε πάρα πολύ καλό
βαθμό τη μεταβλητότητα των εκτιμήσεων ανισοτροπίας.

Συνοψίζοντας, η εργασία παρουσιάζει ένα πλήρες πλαίσιο αναφο-
ρικά με τις στατιστικές ιδιότητες εκτιμήσεων στατιστικής ανισοτρο-
πίας. Υπολογίσθηκε η γενική μαθηματική έκφραση της κοινής
συνάρτησης πυκνότητας πιθανότητας των εκτιμήσεων ανισοτρο-
πίας καθώς και η έκφραση της αντίστοιχης περιοχής εμπιστοσύ-
νης. Στη συνέχεια υπολογίσθηκε μια μη-παραμετρική έκφραση
της κοινής συνάρτησης πυκνότητας πιθανότητας η οποία οδηγεί
σε συντηρητικές περιοχές εμπιστοσύνης. Τα αποτελέσματα της
εργασίας μπορούν εύκολα να επεκταθούν ώστε να συμπεριλαμ-
βάνουν *λογαριθμοκανονικά* τυχαία πεδία. Επίσης υπολογίσθηκε
ένας μη-παραμετρικός στατιστικός έλεγχος ισοτροπίας. Διεξήχθη-
σαν προσομοιώσεις Monte Carlo για τον έλεγχο των αποτελεσμά-
των, ενώ παρουσιάσθηκαν εφαρμογές σε πραγματικά δεδομένα
από ένα σενάριο περιβαλλοντικού κινδύνου και δεδομένα από λι-
γνιτωρυχεία.

Σε μελλοντική έρευνα θα μπορούσε να διερευνηθεί η εφαρμογή
των αποτελεσμάτων σε *μη παραγωγίσιμα* τυχαία πεδία με τη χρή-
ση ενός πυρήνα εξομάλυνσης (smoothing kernel), σε *τριδιάστατα*
τυχαία πεδία, όπως και στα λεγόμενα *Σπαρτιάτικα* τυχαία πεδία.

Δεδομένου ότι οι μεταβολές ενός φυσικού συστήματος οδηγούν σε μεταβολές της ανισοτροπίας, είναι δυνατή η υλοποίηση αλγορίθμων ανίχνευσης στατιστικά σημαντικών αλλαγών ανισοτροπίας οι οποίοι μπορούν να ενσωματωθούν σε συστήματα έγκαιρης προειδοποίησης, ή ακόμα να χρησιμοποιηθούν για την ανίχνευση ανωμαλιών σε δεδομένα από ψηφιακές μαστογραφίες.

Abstract

Two-dimensional data often have autocovariance functions with elliptical equipotential contours, a property known as statistical anisotropy. The anisotropy parameters include the tilt of the ellipse (orientation angle) ∂ with respect to the coordinate system and the ratio R of the principal correlation lengths. Anisotropy estimation is important for applications in the geosciences, in signal and image processing, as well as in medical imaging. Sample estimates of anisotropy parameters are needed for defining suitable spatial models and for the interpolation of incomplete data sets.

Currently, anisotropy estimation is performed via the method of maximum likelihood, which is a computationally intensive procedure, or by inspection of empirical variograms by human experts, which often leads to subjective results. Unsupervised and computationally efficient methods of anisotropy estimation are thus needed for geostatistical analysis as well as for early-warning or automatic monitoring systems.

In several scientific disciplines, including the geosciences, only a single data sample is available, due to measurement costs or other limitations. Thus, a problem of practical interest is the estimation of anisotropy parameters from a single sample, which may also possess an irregular spatial distribution. A computationally efficient anisotropy estimator for single samples with regular or irregular spatial distributions, was recently proposed and is based on Swerling's Covariance Hessian Identity (CHI). The purpose of this thesis is to derive —by means of analytical calculations— an explicit expression for the sampling joint probability density $f_{\partial,R}(\hat{R}, \hat{\partial})$, which is valid for Gaussian, stationary and differentiable random fields. The sampling joint probability density $f_{\partial,R}(\hat{R}, \hat{\partial})$ characterizes the distribution of anisotropy statistics $(\hat{R}, \hat{\partial})$ of the CHI estimator. Based on the sampling joint probability density $f_{\partial,R}(\hat{R}, \hat{\partial})$, we derive an approximation $f_{\partial,R}^{(0)}(\hat{R}, \hat{\partial})$ that is independent of the autocovariance function and

provides conservative confidence regions for the anisotropy parameters $(\hat{R}, \hat{\partial})$. We also formulate a statistical test for isotropy based on the approximation $f_{\partial,R}^{(0)}(\hat{R}, \hat{\partial})$.

We validate the theoretical analysis by means of simulations involving data on square lattices as well as scattered data. We illustrate the use of confidence regions with two real-data case studies; lignite mining data and an environmental emergency scenario which involves the accidental release of gamma radiation are used.

The results of this research can be used to provide (i) a stand-alone approximate estimate of the $(\hat{R}, \hat{\partial})$ distribution (ii) computationally efficient initial values for maximum likelihood estimation, and (iii) a useful prior probability density function for Bayesian anisotropy inference.

Overall, this work provides a framework for efficiently quantifying the statistical significance of anisotropy variations. The identification of significant anisotropy in preliminary data from exploratory drill holes is important for the design and operation of mines, because it allows more accurate analysis of the spatial variability of the deposit. Since a significant anisotropy change may imply a significant change in the physical system, for example accidental release of a pollutant in the environment, we propose that early-warning systems incorporate anisotropy detection. Regarding medical applications, an anisotropy-based algorithm for detecting abnormalities in digital mammography data can be constructed based on the results obtained in this work.

Acknowledgements

The author acknowledges the contribution of the Public Power Corporation of Greece S.A. and Mr. Christos Roumpos for providing drill holes data for the Amynteon and Mavropigi lignite mines, and Andreas Pavlides for his hard work on providing quality data sets derived from the original raw measurements.

None of these could be possible without the help of my brother Eftichis Petrakis and our mutual friend, Tasos Sifalakis, who gave me the opportunity to meet and work with my advisor, Dionissis Hristopulos. Apart from being a supportive advisor, he provided an excellent working environment, despite the calamities the "economic crisis" has brought to all of us.

Since my roots are from the Physics department of the University of Crete, I wish to thank my professors who during my undergraduate years they always reminded us to focus on the *essence* of Physics. I remember the words of my advisor there, Grigoris Athanasiou, that *in order to get a feeling of a problem in hand, we should try to find a proper geometrical representation of it.* Also, I am mostly grateful to Stephanos Trachanas, for his inspiring lectures (my writing and teaching style is mostly influenced by his) and Petros Ditsas for his psychological support during my graduate years at the University of Crete.

And last, but not least, I thank my parents for their support throughout my study years, and my close friends Antonis and Stelios who are always tolerant on the quirks and peculiarities of my character.

Contents

List of Figures

List of Tables

Nomenclature

Roman Symbols

$O(\cdot)$	Big-O notation: If $f(x) \in O(g(x))$, f is bounded above by g (up to a constant factor) asymptotically
$\text{Cov}(X_1, X_2)$	Covariance of the random variables X_1 and X_2
$\tilde{C}_{\text{xx}}(\mathbf{k})$	Covariance spectral density, i.e., the Fourier transform of the covariance function $c_{\text{xx}}(\mathbf{r})$
$c_{\text{xx}}(\mathbf{r})$	(Auto)covariance function of the spatial random field $X(\mathbf{s})$
d	Number of spatial dimensions
$\det(\mathbf{M})$	Determinant of the matrix $\mathbf{M}$
$\dim(\mathbf{v})$	Dimension of vector $\mathbf{v}$
$\mathbb{E}[X]$	Expectation of the random variable X
ℓ_c	Integral scale factor (also known as *integral range*)
ℓ_p	Inverse of the χ^2 cumulative distribution function of probability p and 2 degrees of freedom
$\mathcal{F}$	The set of all observable events, $\mathcal{F} \subset \Omega$
$F^{-1}(p, \nu)$	Inverse of the chi-square (χ^2) cumulative distribution function of probability p and ν degrees of freedom
$f_X(x)$	Probability density function of the random variable X
$f_{X_1, X_2}(x_1, x_2)$	Joint probability density function of the random variables X_1 and X_2

H	Covariance Hessian matrix
J	Unit imaginary number $\sqrt{-1}$
k	Wavevector: If **s** belongs to the direct space, **k** belongs to the reciprocal (or Fourier) space
$K_\nu(\cdot)$	Modified Bessel function of the second kind of order ν
Kurt (X)	Excess kurtosis of X
m_x	Mean of the random variable X
N	Number of sample (data) points
$\mathbb{N}(\mathbf{s})$	Set of natural neighbors of **s**
p	Confidence level probability
$\mathcal{P}$	The probability associated with each event; $\mathcal{P}(\mathcal{F}) \in [0, 1]$
Q	Mean slope tensor (for random fields)
R	Anisotropy ratio
s	Position vector of a point belonging in $\mathcal{D}$
Skew (X)	Skewness of X
σ_x^2	Variance of the random variable X
$X(\mathbf{s})$	Scalar spatial random field

Greek Symbols

a	Lattice constant
χ^2	Chi-square distribution
$\Gamma(\cdot)$	Gamma function
ν	Order of the modified Bessel function; smoothness parameter of the Matérn covariance function; degrees of freedom of the χ^2 distribution
Ω	Sample state (ensemble) that includes all possible states (realizations) of the spatial random field

ω Field state

∂ Anisotropy orientation angle

ξ Correlation length

Superscripts

$*$ Sample, $X_k^* = X(\mathbf{s}_k, \omega^*)$ corresponding to a specific state ω^* of the field $X(\mathbf{s})$.

t Transpose of a vector or matrix

(0) First-term approximation

Other Symbols

$\mid \mathcal{D} \mid$ Enclosed area of the spatial domain $\mathcal{D} \subset \mathbb{R}^2$

$\partial_i X(\mathbf{s})$ Partial derivative of $X(\mathbf{s})$ with respect to the i-th component of $\mathbf{s}$

$\hat{X}$ Sample-based estimate (statistic) of the random variable X

$\| \mathbf{s} \|$ Euclidean norm of vector $\mathbf{s}$

$\blacksquare$ *Quod Erat Demonstrandum,* or *"which had to be demonstrated"* in Latin

Acronyms

CHI   Covariance Hessian Identity

CHM   Covariance Hessian Matrix

CLT   Central Limit Theorem

CR   Confidence Region

GIS   Geographic Information System

GKP   Gradient Kronecker Product

GSRF   Gaussian Spatial Random Field

iid   independent and identically distributed (referring to random variables)

INTAMAP Interoperability and Automated Mapping, a project funded by the European Commission (DG INFSO, FP6)

JPDF Joint Probability Density Function

PPC Public Power Corporation of Greece, S.A.

ppm parts per million

SIC Spatial Interpolation Comparison exercise, organized in 1997 and 2004 by the Radioactivity Environmental Monitoring Group (Institute for Environmental and Sustainability, Joint Research Centre, European Commission)

SRF Spatial Random Field

WSS Wide Sense Stationary random field

Chapter 1

Introduction

1.1 Modeling Spatial Processes in the Geosciences

Datasets collected to study the Earth usually come in the form of two- or three-dimensional scattered points to which attributes are attached. Unlike datasets from other scientific fields such as mechanical engineering where experiments are held in a controlled environment, *geoscientific* data often possess a highly irregular spatial distribution. For example, bathymetric data are collected at a high sampling rate along each ship's track, but there can be a very long distance between two ships' tracks. Also, geologic and oceanographic data, respectively, are gathered from boreholes and water columns; data are therefore usually abundant vertically but sparse horizontally Fisher *et al.* [2005]. In order to model, visualize and better understand such datasets, *spatial interpolation* is performed to estimate the value of an attribute at unsampled locations. *Spatiotemporal interpolation* is required for natural phenomena in which the temporal dimension is important and cannot be neglected. In this work we study the statistical properties of two-dimensional spatial datasets. Furthermore, we present applications on artificial (simulated) and real geoscientific data. In particular the real data involve radioactivity gamma dose rate measured over a network of fixed stations scattered over Europe, and lignite measurements from exploratory drill holes from two Greek lignite mines.

Geostatistics refer to the general class of statistical models and tools developed for statistical analysis of data originating from Geosciences, Chilès & Delfiner [1999]; Goovaerts [1997]; Isaaks & Srivastava [1990]; Journel & Huijbregts [2003]; Kitanidis [1997]; Wackernagel [1997]. Geostatistics
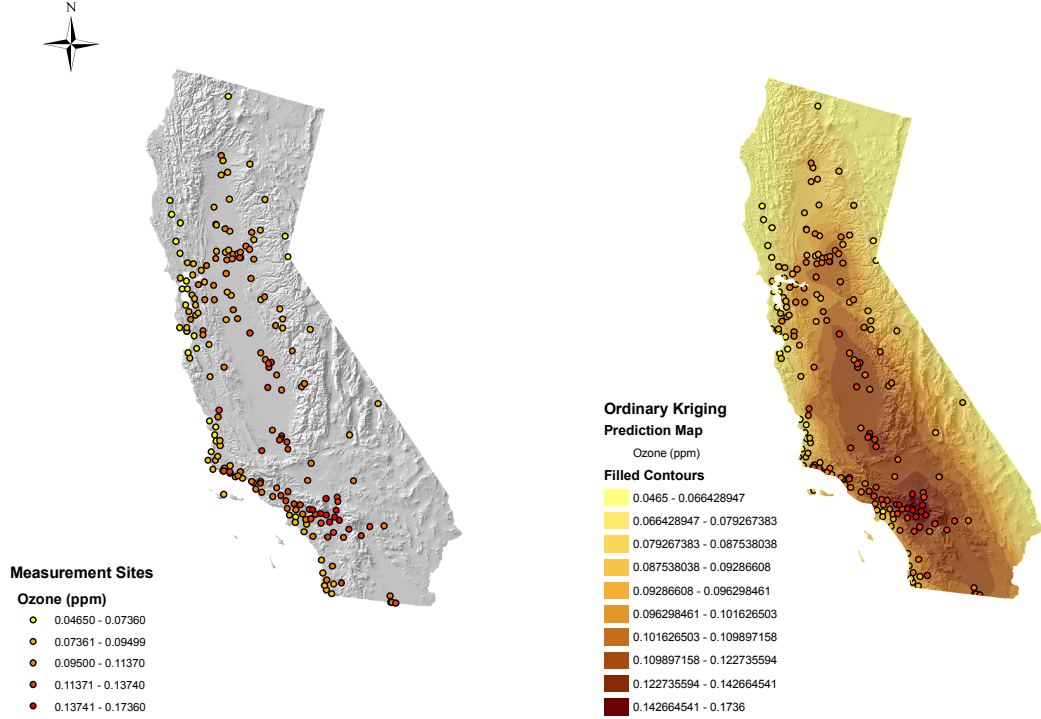
Figure 1.1: Example illustrating the spatial interpolation problem: Given a scattered dataset, for example ozone concentration measurements (in ppm) over California, estimations are required over the whole spatial domain, including locations where measurements are not available. On the left map, measurement values are represented as colored circles, where red colors denote high values. On the right map, ordinary kriging interpolation results are represented as filled contours.

relies on the theory of *random functions* or also known as *random fields*, in order to model the uncertainty associated with spatial estimation and simulation Christakos [1992]; Yaglom [1987]. An example illustrating the use of geostatistical tools for spatial interpolation of environmental variables is portrayed in Figures 1.1–1.3. These figures were produced using a demo version of the commercial[1] Geographic Information System (GIS) software ArcGIS® Geostatistical Analyst Johnston *et al.* [2003, 2008]; Krivoruchko [2008]. The example scenario involves a network of fixed ozone concentration measurement stations over California which provides the

---

[1]This example is for illustratory purposes only. The results of this work can be reproduced without the need of any commercial software. Most of the algorithms are already implemented in R, the open-source programming language for statistical computing, as a part of the INTAMAP project.

maximum value of ozone concentration in parts per million (ppm) over the past eight hours[2]. Given a set of measurements (ozone concentration, Fig. 1.1, left), sampled at several points in space, an estimation of the observable is required at locations where no measurements exist. Usually, the estimated values are presented in the form of a map over a specified spatial domain, (e.g., on the right in Fig. 1.1), along with an associated map displaying the error assessment of the estimates, if the error evaluation is possible. Contour maps, or probability maps of exceeding a threshold value can also be produced. Such maps, such as in Fig. 1.2, can serve as a visual aid for experts to assess the potential health hazards of unusual levels of ozone concentration, to decide issuing an alert, or to take even more extreme measures.

Many interpolation and approximation methods were developed in order to predict values of spatial phenomena at unsampled locations. For a review of the interpolation methods utilized in Geosciences, see Li & Heap [2008]. In general, the methods can be classified into two classes depending whether *randomness* is taken into account:

Deterministic spatial interpolation methods Randomness *is not* taken into account. There is no assessment of prediction errors. Examples of such interpolation methods are: nearest neighbors, triangular irregular network related interpolations, natural neighbors[3], inverse distance weighting, regression models, several spline-based models and Fourier series interpolation models.

Stochastic spatial interpolation methods Randomness *is* taken into account. Assuming random errors, these methods provide an assessment of prediction errors. This is the key advantage of the stochastic methods. Examples of stochastic (or geostatistical) interpolation methods are the several flavors of *kriging* and its extensions: Simple, Ordinary, Universal, Block, Indicator, Disjunctive kriging, kriging with an external drift, cokriging, etc.

The extensive variety of interpolation methods makes evident that a universal interpolation method cannot exist. Moreover, even for an ideal dataset (optimally-sampled, noise-free), there exists an infinite number of interpolation functions which pass through a given set of data points. The processes to be modeled are usually very complex, data are spatially heterogeneous

---

[2]This is an artificial scenario, created for demonstration purposes and based on real measurements during 1996. For data credits see Johnston *et al.* [2008].

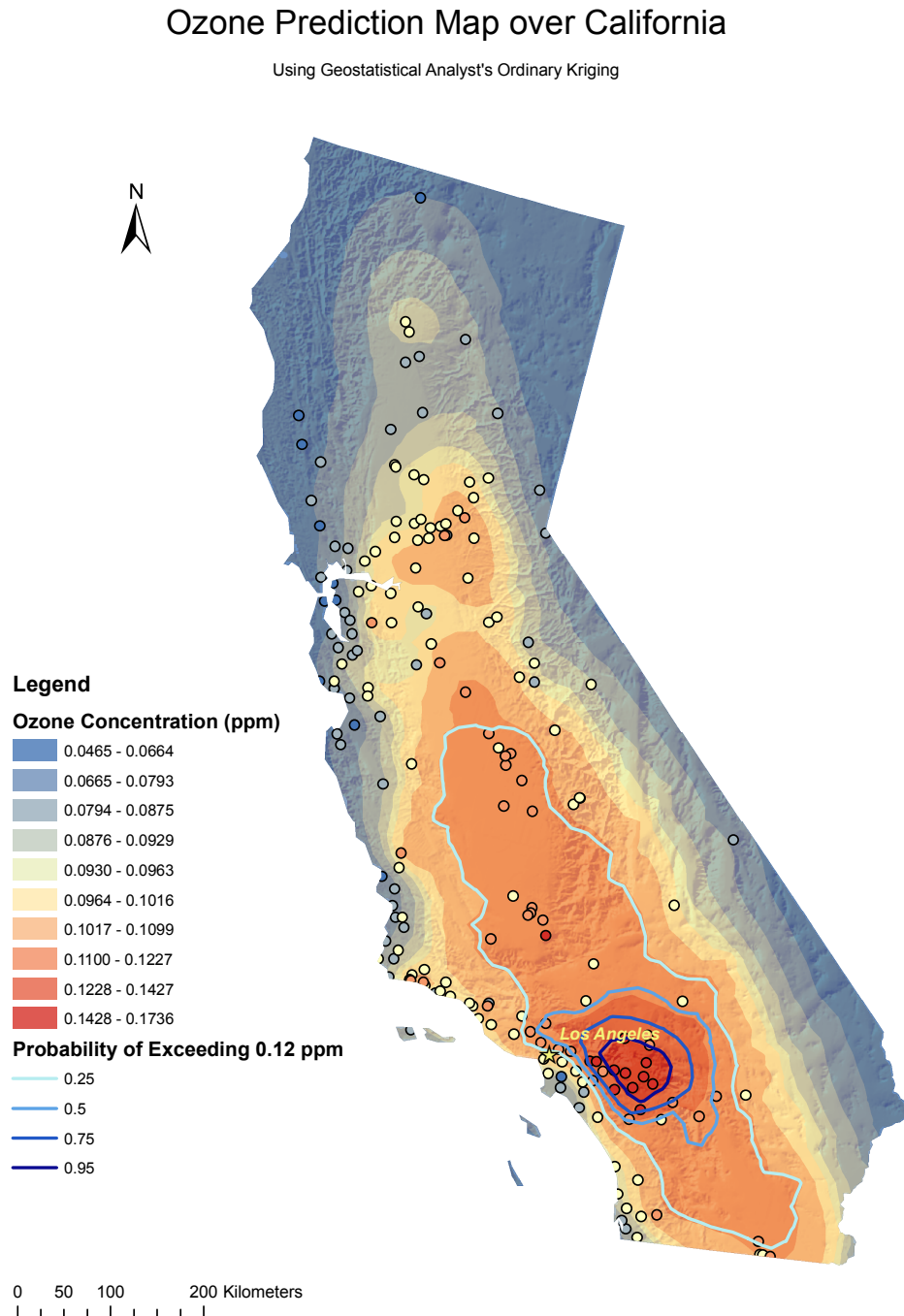[3]The natural neighbors interpolation method is introduced and used in Section 4.2.

Figure 1.2: Ozone concentration prediction over California. The map was obtained using the *ordinary kriging* prediction method as implemented in the ArcGIS® Geostatistical Analyst package.

Figure 1.3: *Ordinary kriging* standard error map associated with the prediction map of Fig. 1.2. Note that sparsely-sampled areas have larger error values.

and often based on less than optimal sampling configurations, and significant noise or discontinuities can be present. In addition, datasets can be very large ($N \simeq 10^3$–10^6), originating from various sources with different accuracies Mitas & Mitasova [1999]. An important criterion for adopting (or dismissing) a specific method is its consistency with the physical laws which the process obeys. Physical laws set constraints on the observables, and not every interpolation method respects such constraints; for example, nearest neighbors interpolation cannot be used for modeling spatially smooth natural processes, since it leads to discontinuous estimation surfaces. In order to take such constraints into account, prior knowledge of the underlying process is required. Usually such prior knowledge is unavailable and therefore some general assumptions should initially be made. Several other requirements address limitations on selecting an interpolation method. For example, we should take into account if a method produces artifacts e.g., unexpectedly high or low estimation values or fails to provide estimation values over a particular sub-domain. The computer processing time required in order to obtain an estimation is also an issue, e.g., the computation of a forecast regarding a spatiotemporal process at a future date, should finish earlier than that given date. Quoting Lewis Fry Richardson[4] from the early days of numerical weather prediction models: *"Perhaps some day in the dim future it will be possible to advance the computations faster than the weather advances. . . But that is a dream."* Lynch [2008]. Despite the advances in science and technology since Richardson's days, a good physical understanding of the process and good quality data can not be substituted by any complicated mathematical model, advanced algorithmic scheme, or the abundance of today's computer processing power.

1.2 Spatial Random Fields

Spatial random fields (SRFs), also known as spatial random functions Christakos [1992]; Yaglom [1987], are used in several other scientific and engineering disciplines that study spatially distributed processes, (e.g., image processing, theory of transport in heterogeneous media, wave propagation in random media). Natural processes which can be considered as

---

[4]Lewis Fry Richardson (1881–1953) was an English mathematician, physicist, meteorologist, psychologist and pacifist who pioneered modern mathematical techniques of weather forecasting, and the application of similar techniques to studying the causes of wars and how to prevent them. He is also noted for his pioneering work on fractals. Since 1997, the Lewis Fry Richardson Medal is been awarded by the European Geosciences Union for "exceptional contributions to nonlinear geophysics in general" Wikipedia contributors [2012a].

(a) Looking "up"   (b) Looking "down"   (c) Looking "inside"
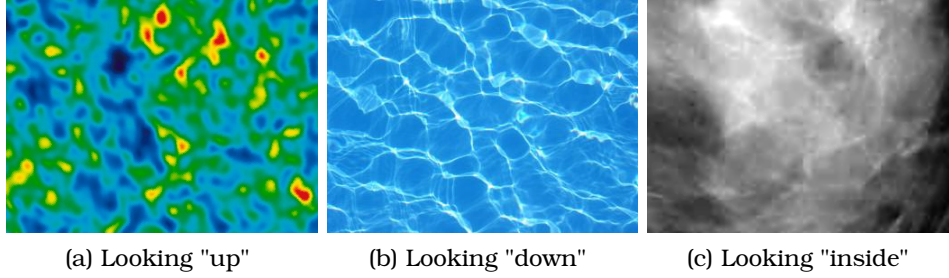
Figure 1.4: Random fields in nature: (a) From the primordial light where *everything* is bathed in, the *cosmic microwave background radiation* (source image: Hinshaw *et al.* [2009]), (b) the oceans of our planet, and (c), the "innerspace": a mammogram (source image: Richard & Bierme [2010]).

realizations of a spatial random field are illustrated in Figure 1.4. Spatial random fields with Gaussian joint probability density function are also used in machine learning, where they are known as *Gaussian processes* Rasmussen & Williams [2006].

For convenience, isotropic SRF models are often used, even though many real data sets display anisotropic patterns. *Physical anisotropy* implies different values of a specific variable along different directions and is expressed by means of tensor coefficients (e.g., electrical conductivity tensor). *Statistical anisotropy* characterizes scalar processes (e.g., values of gray-scale images, pollutant concentrations), the correlation range of which depends on the spatial direction. Herein we focus on statistical anisotropy, which implies SRFs with autocovariance functions that possess elliptical equipotential contours.

A problem of practical interest is the estimation of anisotropy parameters from a single available sample. Estimation of anisotropy parameters, is a topic of ongoing research activity in various signal and image processing applications Feng *et al.* [2008]; Jiang [2005]; Le Bihan *et al.* [2001]; Okada *et al.* [2005]; Olhede [2008]; Richard & Bierme [2010]; Wang & Leckie [2012]; Xu & Choi [2009], as well as in data assimilation Weaver & Mirouze [2012]. The characterization and measurement of anisotropy in biological materials is important for diagnostic and medical reasons Ranganathan *et al.* [2011]; Richard & Bierme [2010]. Improved methods for estimating anisotropy are necessary in various research fields.

Usually anisotropy estimates are obtained using maximum likelihood, which is a computationally demanding procedure. In the case of geoscientific applications, the choice between an isotropic or anisotropic model usually requires inspection of *empirical variograms* from a human expert, a

Rotate and Rescale

Anisotropic Field   Isotropic Field
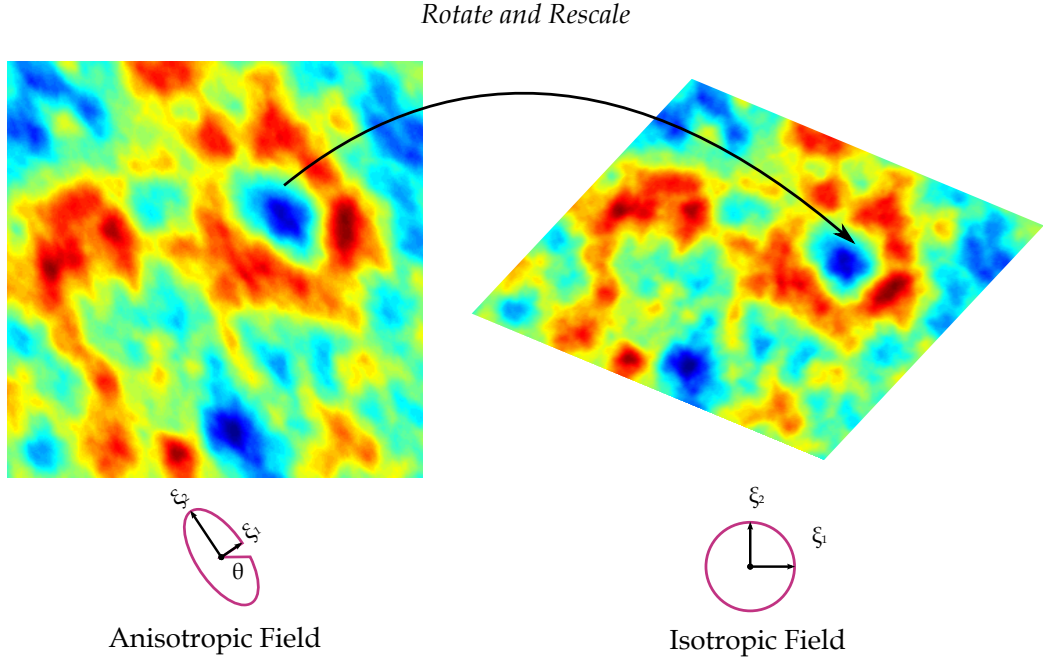
Figure 1.5: Anisotropy detection as a preprocessing step: Since the CHI anisotropy estimation and rotation/rescaling of initial data is computationally simple, anisotropy detection is a crucial preprocessing step since preprocessed data can be modeled with a simpler isotropic model.

procedure not suitable for automatic monitoring applications Pebesma *et al.* [2011], which can be subjective. The calculation of empirical variograms is also a computationally demanding procedure, especially when data size is large, e.g., data originating from GIS applications.

On the other hand, rotating and rescaling of anisotropic data is a computationally trivial preprocessing task which simplifies data representation and modeling. The preprocessed data can then be modeled with an isotropic spatial model with less parameters to be estimated. The anisotropy removal preprocessing procedure is depicted in Figure 1.5. Thus if anisotropy removal is going to be used as an initial preprocessing step, it should be fast and efficient. The importance of preprocessing data prior to modeling is stressed by paraphrasing a "silly" theorem from the neural networks community, *"modeling will always succeed, given the right preprocessor"*[5] Hertz *et al.* [1991]. The methods currently available for anisotropy estimation are computationally complex for being used as an initial preprocessing step

---

[5]The proof is simple: The preprocessor solves the initial problem and encodes the result into a simple form where a trivial model could reproduce the answer.

before the actual modeling. A method for providing anisotropy estimates without the computational cost of solving the *entire* modeling problem is necessary.

The *Covariance Hessian Identity (CHI)* method Chorti & Hristopulos [2008]; Hristopulos [2002] is a non-parametric, non-iterative method for obtaining semi-analytic estimates of SRF anisotropy parameters from two-dimensional data sets. Nevertheless, anisotropy estimates are *statistics*, i.e., random variables, the values of which fluctuate between different samples. The fluctuations in the two parameters, as we show below, are inter-dependent and non-Gaussian. If the joint distribution of the fluctuations is known, it is possible to evaluate whether deviations of anisotropy statistics between different data sets are statistically significant. Another practical question is whether an estimated anisotropy value actually means significant departure from isotropy; if not, simpler, isotropic autocovariance functions can be used for spatial modeling. Significant changes in anisotropy over time may suggest a crucial change in the underlying physical processes. For example, an accidental release of radioactivity may significantly alter the anisotropy of radioactivity patterns over the monitored area. Reliable and computationally fast detection of systematic changes in spatial distributions is crucial, especially for automatic monitoring systems Pebesma *et al.* [2011].

Answering questions as the above, requires mathematical expressions for the confidence regions of anisotropy statistics. In this thesis, we derive a non-parametric approximation of the *sampling Joint Probability Density Function (JPDF)* and *Confidence Region (CR)* of anisotropy statistics for *differentiable, Gaussian random fields*. We prove this expression using CHI, the Central Limit Theorem, Jacobi's multivariate transformation theorem, and a perturbation expansion. The term "non-parametric" implies that the approximation is independent of the SRF autocovariance function (henceforward, *covariance function* for simplicity). The non-parametric approximation is shown to yield a sampling JPDF which is more dispersed in parameter space than the exact JPDF. This implies a wider confidence region for the anisotropy statistics. Hence, if a sample is classified as isotropic at confidence level p based on the approximate JPDF, it is actually isotropic at $p' > p$. These estimates can be used as a prior in Bayesian inference Schmidt & O'Hagan [2003].

The rest of the thesis is structured as follows: In Chapter 2 we present essential definitions and an overview of CHI, on which the mathematical development of the joint PDF is based. Then, we derive a general expression for the JPDF, using the Central Limit Theorem and the conservation of probability under variable transformation. In addition, we obtain an equation

that determines the confidence regions of anisotropy statistics. In Chapter 3 we derive the non-parametric approximation of the JPDF and the corresponding confidence region expression and we formulate a non-parametric test for isotropy. In Chapter 4 we compare the non-parametric estimates of confidence regions with the results of numerical simulations, and we illustrate the application of confidence regions using real data regarding an environmental emergency scenario over Germany and data from two Greek lignite mines. Finally, in Chapter 5 we review the main results obtained in this work, we present our conclusions, and we outline directions for future research.

Anisotropy statistics

2.1 Anisotropic Random Fields and the Covariance Hessian Identity

IN the following, boldface symbols are used for vectors, matrices and tensors; the superscript "t" denotes the transpose of a vector or matrix. Let $\mathcal{D} \subset \mathbb{R}^2$ denote the spatial domain and $|\mathcal{D}|$ the enclosed area. The vector $\mathbf{s} \in \mathcal{D}$ denotes the position of a point in $\mathcal{D}$ and $\|\mathbf{s}\|$ denotes the Euclidean norm of $\mathbf{s}$. Let $X(\mathbf{s}, \omega)$ represent a scalar SRF on the probability space $(\Omega, \mathcal{F}, \mathcal{P})$ [1]. The state index ω determines the field state and will be suppressed in the following for the sake of brevity. The SRF $X(\mathbf{s})$ represents a scalar variable, e.g., temperature, dose rates, or grey-scale intensity levels of a digital image. The events in $\mathcal{F}$ comprise the measured SRF realization(s) or *sample states(s)*. $\mathbb{E}[\cdot]$ denotes the expectation over the ensemble of states, and the operator

$$\text{Cov}(Z_1, Z_2) = \mathbb{E}[Z_1 Z_2] - \mathbb{E}[Z_1]\mathbb{E}[Z_2], \tag{2.1}$$

denotes the covariance of the random variables Z_1 and Z_2.

We will focus on Gaussian SRFs (GSRFs) that possess normal joint probability density functions with variance σ_x^2. *Wide-sense stationarity* will be assumed, i.e., that the mean $m_x = \mathbb{E}[X(\mathbf{s})]$ is constant, and the *covariance function*

$$c_{xx}(\mathbf{r}) = \mathbb{E}[X(\mathbf{s}) X(\mathbf{s} + \mathbf{r})] - m_x^2 \tag{2.2}$$

---

[1] Ω denotes the sample space (ensemble) that includes all the possible states (realizations) of the SRF, $\mathcal{F}$ is the set of all observable events, $\mathcal{F} \subset \Omega$, and $P(\mathcal{F}) \in [0, 1]$ is the probability associated with each event.

is independent of $\mathbf{s}$. Let $\bar{\mathbf{e}}_i$ denote the unit vector in the spatial direction i, ($i = 1, 2$). If the partial derivatives $\partial^2 c_{xx}(\mathbf{r})/\partial r_i^2$ in the orthogonal directions $i = 1, 2$ exist at $\mathbf{r} = (0, 0)$, the SRF is differentiable in the mean square sense (m.s.s) for every $\mathbf{s} \in \mathcal{D}$, i.e.,

$$\lim_{h \to 0} \mathbb{E}\left[\left\|\partial_i X(\mathbf{s}) - \frac{X(\mathbf{s} + h\bar{\mathbf{e}}_i) - X(\mathbf{s})}{h}\right\|^2\right] = 0, \quad \forall i = 1, 2. \tag{2.3}$$

For Gaussian SRFs, mean square differentiability practically implies that the derivatives of the sample states exist almost surely Yaglom [1987]. We will focus on SRFs that possess at least continuous first-order derivatives in the mean square sense. In addition, we will assume *short-range* correlations, namely that the covariance falls off sufficiently fast as $\|\mathbf{r}\| \to \infty$, so that the *correlation area* $V_c := \frac{1}{\sigma_x^2} \int d\mathbf{r}\, c_{xx}(\mathbf{r})$ is finite.

In general, a state (realization) of the SRF can be decomposed into a deterministic trend $m_x(\mathbf{s})$, a fluctuation $\chi_\lambda(\mathbf{s})$, and a noise term $\eta(\mathbf{s})$ according to Goovaerts [1997]

$$X(\mathbf{s}) = m_x(\mathbf{s}) + \chi_\lambda(\mathbf{s}) + \eta(\mathbf{s}). \tag{2.4}$$

The trend represents large-scale variations of the field, corresponding to the ensemble average Kanevski & Maignan [2004]. The fluctuation models faster variations, which may appear as quasi-random changes, and have a cut-off scale λ determined by the spatial resolution of the experiment. Any fluctuations that occur at smaller scales can not be resolved, and they are incorporated with various other random perturbations in the noise term, which is modeled as a zero-mean white Gaussian SRF. In the following, we assume that the trend is removed from the initial SRF, and we will use the symbol $X(\mathbf{s})$ for the residual, i.e., the fluctuation and the noise components.

The spatial dependence of the SRF fluctuations is determined by means of the covariance function or the structure function (semivariogram). In many cases, one can assume that the fluctuation is a *Wide Sense Stationary (WSS) SRF, or a SRF with WSS increments* Christakos [1992]; Papoulis & Pillai [2002]; Yaglom [1987]. For WSS SRFs the *structure function* (semivariogram) $\gamma_x(\mathbf{r})$, defined as

$$\gamma_x(\mathbf{r}) = \frac{1}{2}\mathbb{E}\left[X(\mathbf{s}) - X(\mathbf{s} + \mathbf{r})\right]^2, \tag{2.5}$$

contains the same information as the covariance. In practice the structure function is often estimated instead of the covariance Yaglom [1987]. The structure function has a sill that is equal to the variance σ_x^2 of the SRF, and the following relation holds Jupp *et al.* [1988]; Papoulis & Pillai [2002],

$$\gamma_x(\mathbf{r}) = \sigma_x^2 - c_{xx}(\mathbf{r}). \tag{2.6}$$

An SRF $X(\mathbf{s})$ has second-order stationary increments if the SRF $\Psi(\mathbf{s}) = X(\mathbf{s}) - X(\mathbf{s}_0)$ is WSS. Such an SRF is also called *intrinsic* Stein [2001]. The category of intrinsic SRF's includes fractional Brownian motion Mandelbrot [1968]. In this case, the translation invariance of the covariance is lost. The structure function is still purely a function of the space lag, but it increases without bound. One can investigate anisotropy in the intrinsic SRFs by focusing on their increments, which are WSS.

The fluctuation SRF is *isotropic* if the covariance function depends purely on the magnitude of the distance vector, i.e., $c_{\text{xx}}(\mathbf{r}) = c_{\text{xx}}(\|\mathbf{r}\|)$, where $\|\mathbf{r}\|$ is the Euclidean norm of the vector $\mathbf{r}$. An *anisotropic* covariance function can be expressed as $c_{\text{xx}}(\mathbf{r}) = c_{\text{xx}}(\|\mathbf{r}\|_A)$, where $\|\mathbf{r}\|_A$ is an anisotropic norm. In two dimensions, if

$$\mathbf{r} = (r_1, r_2)^t \text{ and } \mathbf{A} = \begin{bmatrix} A_1 & A_{12} \\ A_{12} & A_2 \end{bmatrix},$$

then the square of the anisotropic norm is

$$\|\mathbf{r}\|_A^2 = \mathbf{r}^t \mathbf{A} \mathbf{r} = A_1 r_1^2 + A_2 r_2^2 + 2 A_{12} r_1 r_2. \tag{2.7}$$

The matrix A reflects the scaling and rotation effect of the anisotropy. The elements of $\mathbf{A}$ are variables of the anisotropy parameters. Specifically,

$$A_1 = \left(\frac{\cos \partial}{\xi_1}\right)^2 + \left(\frac{\sin \partial}{\xi_2}\right)^2 \tag{2.8a}$$

$$A_2 = \left(\frac{\cos \partial}{\xi_2}\right)^2 + \left(\frac{\sin \partial}{\xi_1}\right)^2 \tag{2.8b}$$

$$A_{12} = \left(\frac{1}{\xi_1^2} - \frac{1}{\xi_2^2}\right) \sin \partial \cos \partial. \tag{2.8c}$$

The anisotropy parameters involve the *orientation (rotation) angle*, ∂ which determines the orientation of the principal axes, and the *anisotropic ratio*

$$R = \frac{\xi_2}{\xi_1} \tag{2.9}$$

which represents the ratio of the principal correlation lengths ξ_2 over ξ_1 (chosen arbitrarily). In the coordinate system of the principal axes $\mathbf{r}'$, the matrix $\mathbf{A}$ is *diagonal*. The correlation lengths ξ_i, $i = 1, 2$ of $X(\mathbf{s})$ determine the local rate of change of the covariance function along the principal directions, i.e.,

$$\xi_i^{-2} = -\frac{a}{\sigma_{\text{x}}^2} \left.\frac{\partial^2 c_{\text{xx}}(\mathbf{r}')}{\partial r_i'^2}\right|_{\mathbf{r}'=\mathbf{0}}, \tag{2.10}$$

where $a > 0$ is an $O(1)$ constant.

Realizations of isotropic and anisotropic Gaussian random fields generated by computer simulations using the *Fourier Filtering Method* Pardo-Igúzquiza & Chica-Olmo [1993] are shown in Figures 2.1 and 2.2. The covariance function is plotted along the horizontal (x) and vertical (y) axis and compared with sample-based estimates. Estimates of the covariance function can be calculated from a single realization of a RF $X(\mathbf{s})$ by means of the *ergodic theorem* which states that under certain mild conditions, ensemble expectations can be replaced by spatial averages Yaglom [1987]. The sample-based estimates of the covariance function in Figs. 2.1 and 2.2 were obtained utilizing the function $\texttt{xcorr}$ of MATLAB$^{\circledR}$. Deviations are due to finite size of the sample domain which makes the replacement of expectations to spatial averages to be approximate.

Furthermore, we assume that the *sample*, X_k^* comprises the values $X_k^* = X(\mathbf{s}_k, \omega^*)$ of the field $X(\mathbf{s})$ for a specific state ω^*, where $\mathbf{s}_k$, $k = 1, \ldots, N$ are sampling locations. We will denote sample-based estimates by the "hat" symbol, e.g., $\hat{R}$ represents the estimate of the anisotropic ratio R.

The *Covariance Hessian Matrix* $\mathbf{H}$ (CHM) of a stationary, at least once differentiable, SRF $X(\mathbf{s})$ is defined as follows

$$H_{ij}(\mathbf{r}) = -\frac{\partial^2 c_{\mathrm{xx}}(\mathbf{r})}{\partial r_i\, \partial r_j}, \quad i, j = 1, 2. \tag{2.11}$$

Let $X_{ij} = \partial_i X(\mathbf{s})\, \partial_j X(\mathbf{s})$, $i = 1, 2$ be the *Gradient Kronecker Product (GKP)* tensor; $\mathbf{X}$ is a symmetric second-rank tensor. The expectation of GKP, henceforward called the *mean slope tensor* and denoted by $\mathbf{Q}$, is defined as follows

$$Q_{ij} = \mathbb{E}\left[\partial_i X(\mathbf{s})\, \partial_j X(\mathbf{s})\right]. \tag{2.12}$$

The mean slope tensor is also known as the *matrix of spectral moments*, and plays a key role in determining the local maxima and excursion sets of random fields Adler [1981]. Swerling has proved the following *Covariance Hessian Identity* (CHI) Swerling [1962]:

Theorem 2.1.1 (Swerling's Covariance Hessian Identity)**.** *Let $X(\mathbf{s})$ be a statistically stationary GSRF with a covariance function that admits partial derivatives $\partial^2 c_{\mathrm{xx}}(\mathbf{r})/\partial r_i^2$ in the orthogonal directions $i = 1, 2$ at $\mathbf{r} = (0, 0)$. Then, the mean slope tensor is connected to the covariance Hessian matrix as follows:*

$$\mathbf{Q} = \mathbf{H}(\mathbf{r})|_{\mathbf{r}=\mathbf{0}}\,. \tag{2.13}$$

The mean slope tensor $\mathbf{Q}$ is nonnegative definite since it is the covariance matrix of the random vector $[\partial_1 X(\mathbf{s}), \partial_2 X(\mathbf{s})]^t$: Based on the *Cauchy-Schwarz*

(a) Gaussian random field with isotropic Gaussian covariance.

(b) Normalized (theoretical and sample) covariance function along x and y axes.

(c) Gaussian random Field with isotropic Matérn covariance.

(d) Normalized (theoretical and sample) covariance function along x and y axes.
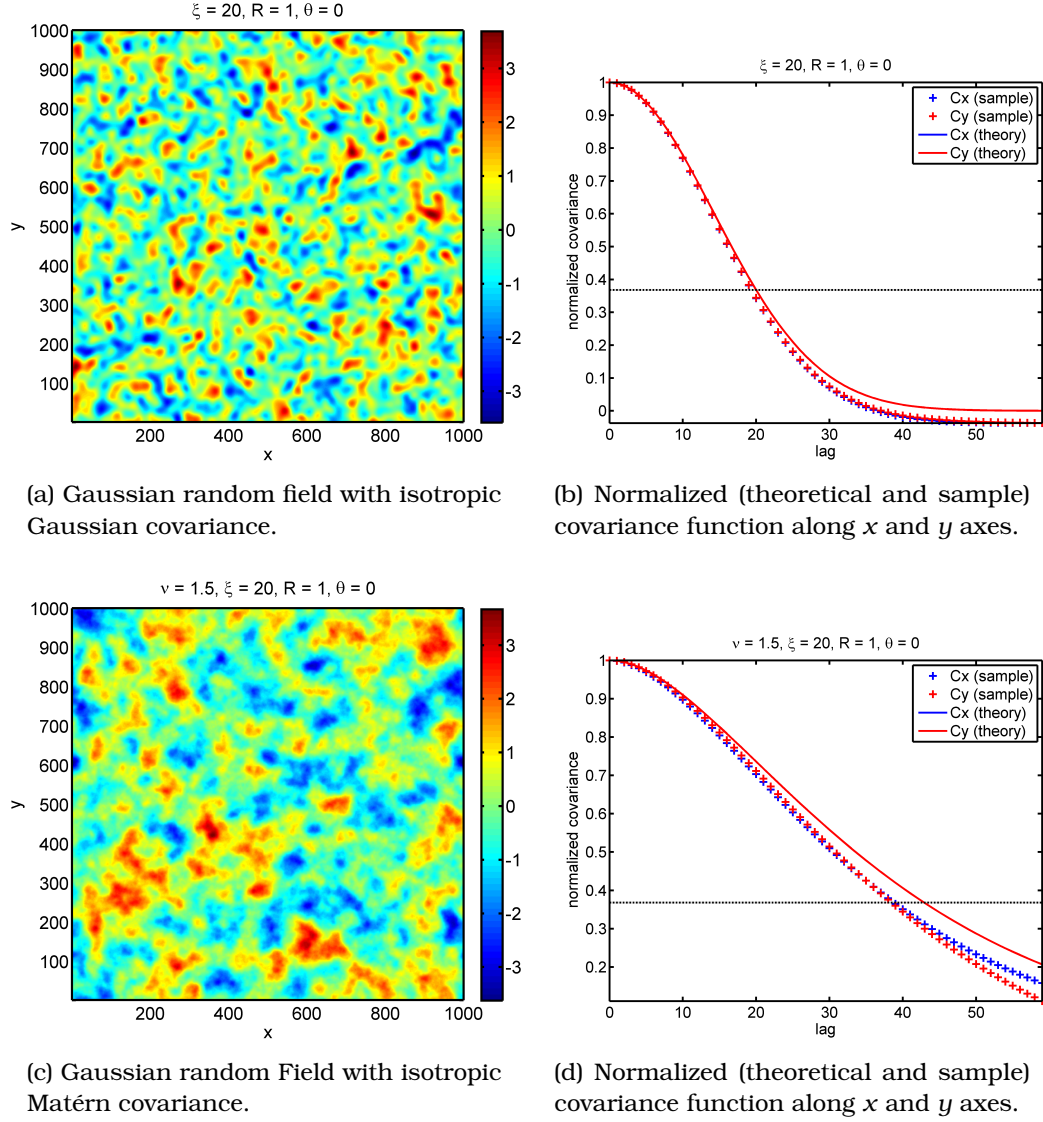
Figure 2.1: Examples of Gaussian random fields with isotropic correlation functions: A realization of a zero-mean, unit-variance Gaussian random field with correlation length $\xi = 20$ over a 1000×1000 square grid and (a) Gaussian, (c) Matérn covariance ($\nu = 1.5$). Theoretical and estimated covariance functions across the x (horizontal) and y (vertical) directions for the (b) Gaussian and (d) Matérn covariance. Deviations from theory are due to finite-size effects. The horizontal line in (b) and (d) is the value of the normalized isotropic *Gaussian* correlation function at $r = \xi$.

(a) Gaussian random field with aniso-tropic Gaussian covariance).

(b) Normalized (theoretical and sample) covariance function along x and y axes.

(c) Gaussian random field with aniso-tropic Matérn covariance.

(d) Normalized (theoretical and sample) covariance function along x and y axes.

Figure 2.2: Examples of Gaussian random fields with anisotropic correlation functions: A realization of a zero-mean, unit-variance Gaussian random field with correlation length $\xi = 10$ over a 1000×1000 square grid and (a) Gaussian covariance with $R = 2$, $\partial = 30°$ and (c) Matérn covariance with $\nu = 1.5$, $R = 0.4$ and $\partial = -20°$. Theoretical and estimated covariance functions across the x (horizontal) and y (vertical) directions for the (b) Gaussian and (d) Matérn covariance. Deviations from theory are due to finite-size effects. The horizontal line in (b) and (d) is the value of the normalized isotropic *Gaussian* correlation function at $r = \xi$. Note that ξ is smaller than of Fig. 2.1.

16

inequality, $(\mathbb{E}\left[\partial_1 X(\mathbf{s})\,\partial_2 X(\mathbf{s})\right])^2 \leq \mathbb{E}\left[\{\partial_1 X(\mathbf{s})\}^2\right]\mathbb{E}\left[\{\partial_2 X(\mathbf{s})\}^2\right]$, it follows that $\det(\mathbf{Q}) \geq 0$.

CHI is valid in any number of spatial dimensions. In two dimensions, the anisotropy parameters (R, ϑ) satisfy the following theorem Chorti & Hristopulos [2008]:

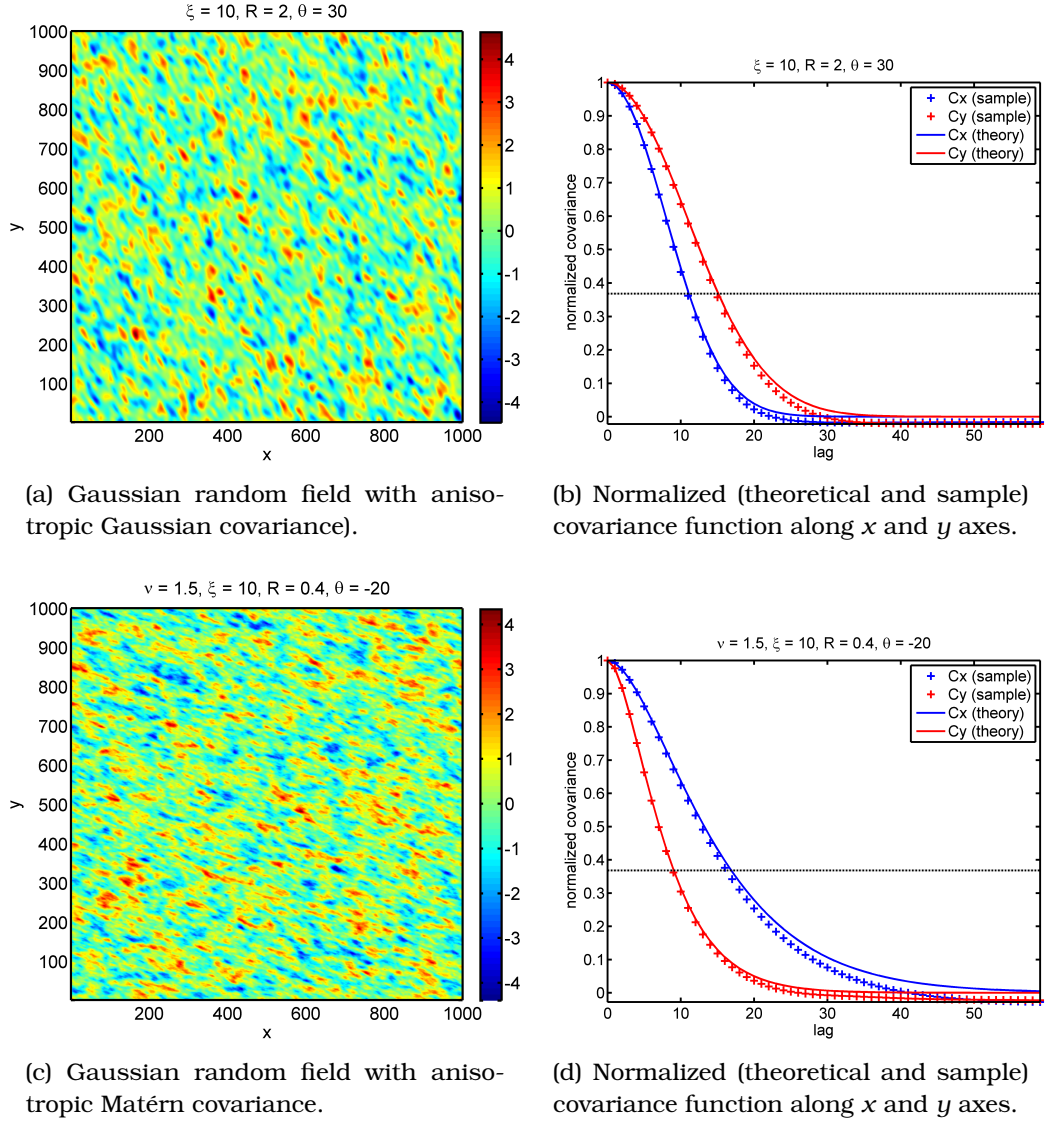**Theorem 2.1.2.** *Let $X(\mathbf{s})$ be a statistically stationary and anisotropic GSRF respecting the conditions of 2.1.1, and ξ_i, $i = 1, 2$ represent the principal correlation lengths of $X(\mathbf{s})$. We define the anisotropy ratio $R = \xi_2/\xi_1$, and the orientation (rotation) angle, ϑ, as the angle between the horizontal axis of the coordinate system and the first principal axis of the SRF (arbitrarily defined). We also define the slope tensor ratios of the elements Q_{ij}, $i, j = 1, 2$:*

$$q_{\mathrm{d}} \doteq \frac{Q_{22}}{Q_{11}} = \frac{1 + R^2 \tan^2 \vartheta}{R^2 + \tan^2 \vartheta}, \qquad (2.14)$$

$$q_{\mathrm{o}} \doteq \frac{Q_{12}}{Q_{11}} = \frac{\tan \vartheta (R^2 - 1)}{R^2 + \tan^2 \vartheta}. \qquad (2.15)$$

Then, the anisotropy parameters are given by

$$\vartheta = \frac{1}{2}\tan^{-1}\left(\frac{2q_{\mathrm{o}}}{1 - q_{\mathrm{d}}}\right), \qquad (2.16)$$

$$R = \left[1 + \frac{1 - q_{\mathrm{d}}}{q_{\mathrm{d}} - (1 + q_{\mathrm{d}})\cos^2 \vartheta}\right]^{-1/2}. \qquad (2.17)$$

Proof. The proof, which is based on Theorem 2.1.1, is shown in Chorti & Hristopulos [2008]. The only difference is that therein $R = R_{2(1)} = \xi_1/\xi_2$ is used, while above we defined $R = \xi_2/\xi_1$ in (2.9). The results from Chorti & Hristopulos [2008] apply by means of the transformation $R \to 1/R$. ∎

Equations (2.14)–(2.15) are invariant under the pair of transformations $\tan \vartheta \to -(\tan \vartheta)^{-1}$, that is, $\vartheta \to \vartheta \pm \pi/2$, and $R \to 1/R$. By restricting the parameter space to $R \in [0, \infty)$ and $\vartheta \in [-\pi/4, \pi/4)$, or equivalently $R \in [1, \infty)$ and $\vartheta \in [-\pi/2, \pi/2)$, the pair $(\hat{R}, \hat{\vartheta})$ satisfying (2.14)–(2.15) for given $(\hat{q}_d, \hat{q}_o)$ is unique, thus ensuring that the transformation $(\hat{q}_d, \hat{q}_o) \to (\hat{R}, \hat{\vartheta})$ is one-to-one. Theorem 2.1.2 permits estimating the anisotropy parameters if the mean slope tensor can be estimated from the data Chorti & Hristopulos [2008]; Hristopulos [2002].

2.2 Sampling Joint Probability Density Function of Anisotropy Statistics

The sampling values of Q_{ij} will be denoted by the random variable $\hat{Q}_{ij}$. In the following, the hat over a mathematical symbol denote *random variables* that are sample-based estimates (statistics). We propose the spatially averaged *slope tensor estimate* $\hat{Q}_{ij}$, $i, j = 1, 2$:

$$\hat{Q}_{ij} := \frac{1}{N} \sum_{k=1}^{N} X_{ij}(\mathbf{s}_k) = \frac{1}{N} \sum_{k=1}^{N} \partial_i X(\mathbf{s}_k)\, \partial_j X(\mathbf{s}_k). \tag{2.18}$$

Equation (2.12) involve the expectation of the GKP elements, which are random functions. By invoking the *ergodic theorem*, ensemble expectations can be replaced by spatial averages (2.18) over the domain of a single realization. Additionally, the field derivatives $\partial_i X(\mathbf{s})$, $i = 1, 2$ are replaced by their estimates, $\hat{\partial}_i(\mathbf{s})$, $i = 1, 2$. In practice, estimates of the field derivatives are obtained by applying a simple interpolation method (in the case of scattered data) and using the centered-difference derivative approximation on the resulting interpolated surface. The JPDF of the mean slope tensor estimate $\hat{\mathbf{Q}}$ can be approximated by using the Central Limit Theorem considering that (2.18) involve averages of the random variables $\hat{Q}_{ij}$. Since the slope tensor estimates are not *independent* random variables, the extension of the classical CLT for short-range correlations is applied.

For each sample of the SRF, in general a different estimate of the tensor $\hat{Q}_{ij}$ is obtained, leading through application of (2.16) and (2.17) to different estimates of $\hat{R}$ and $\hat{\partial}$. Hence, a probability distribution is obtained for $\hat{R}$ and $\hat{\partial}$. In this section we will calculate the joint PDF of $\hat{R}$ and $\hat{\partial}$ from an ensemble of states. The calculation of $f_{\partial,R}(\hat{R}, \hat{\partial})$ is based on the application of the classical *Central Limit Theorem* (CLT). As stated below, this implies $N \gg 1$, a condition that is satisfied in most applications of interest. The derivation of $f_{\partial,R}(\hat{R}, \hat{\partial})$ involves several technical steps that include the variable transformations from $\hat{\mathbf{Q}}$ to $(\hat{q}_d, \hat{q}_o)$ and finally to $(\hat{R}, \hat{\partial})$ and the respective transformations of the PDFs. In order to maintain focus in the main text, some of the proofs are relegated to Appendices. In Appendix A.1, Jacobi's theorems related to the transformation of probability distributions under a variable transformation are reviewed Papoulis & Pillai [2002]. In Appendix A.2, we calculate the univariate PDFs of the gradient components and the GKP tensor elements. The covariance matrix of the GKP tensor elements is derived in Appendix A.3. The JPDF of the GKP tensor elements is derived in Appendix A.4 based on the Central Limit Theorem. Appendix A.5

presents the proof of Lemma 2.2.3, which determines the bivariate JPDF of the average slope tensor ratios, $\hat{q}_o$, and $\hat{q}_o$. Finally, Appendix A.6 proves the explicit relation for the JPDF of the anisotropy statistics, (2.30), as formulated in Theorem 2.2.2.

2.2.1 Central Limit Theorem and Joint Probability Density Function of Slope Tensor Estimates $\hat{\mathbf{Q}}$

The classical CLT for scalar random variables is discussed in Feller [1971]; Gnedenko & Kolmogorov [1954]; Levy [1954]. An extension of the classical CLT applies to vector random variables Anderson [1984]:

Theorem 2.2.1 (Multivariate CLT). *Let us assume N independent and identically distributed vector variables $\mathbf{Z}_k$, $k = 1, \ldots, N$ with mean $\mathbf{m}$ and covariance matrix $\mathbf{C}_{ZZ}$. The random vector $\bar{\mathbf{Z}} = (\mathbf{Z}_1 + \ldots \mathbf{Z}_N)/N$ is asymptotically (i.e., for $N \to \infty$) normally distributed with mean $\mathbf{m}$ and covariance matrix $\mathbf{C}_{ZZ}/N$.*

The same theorem also applies to *correlated random variables*, i.e., random fields $\mathbf{Z}(\mathbf{s})$. The main requirements are stationarity and finite range of correlations. Stationarity ensures that the correlations between a pair of points only depend on the distance but not the location of the pair. If the size of the correlated areas is defined as $V_c := \max_{i,j}\left((\sigma_{Z_i}\sigma_{Z_j})^{-1} \int d\mathbf{r}\, c_{Z_i Z_j}(\mathbf{r})\right)$, the requirement is that V_c be finite and $|\mathcal{D}|/V_c \gg 1$. This follows by a straightforward extension of the scalar case Bouchaud & Georges [1990]. An intuitive explanation of the constraint on the correlation range in the scalar case is as follows: consider that $\mathcal{D}$ comprises "blobs" of size $\propto V_c$. The point at the center of the blob is correlated with other points inside the blob and uncorrelated with points outside the blob. Hence, one can think of the blobs as representing independent random variables, viz. Fig. 2.3. Assuming a uniform distribution of sampling points, each blob contains, on average, a number of points $N_b \approx N V_c/\mathcal{D}$. Thus, the effective number of "independent units" is $N_{\text{eff}} = N/N_b \approx |\mathcal{D}|/V_c$. If V_c is finite and $N_{\text{eff}} \gg 1$, CLT applies to $\hat{\mathbf{q}}$ given by (2.18). Note that $N_{\text{eff}} \gg 1$ also implies $N \gg 1$, since $N = N_{\text{eff}} \times N_b$.

2.2.2 Joint Probability Density Function of Slope Tensor Elements

We define the vector $\bar{\mathbb{Q}} \equiv (\hat{Q}_{11}, \hat{Q}_{22}, \hat{Q}_{12})^t$ that comprises the *independent components of the slope tensor statistics*. According to (2.18), $\hat{Q}_{ij} = \frac{1}{N} \sum_{k=1}^{N} X_{ij}(\mathbf{s}_k)$.
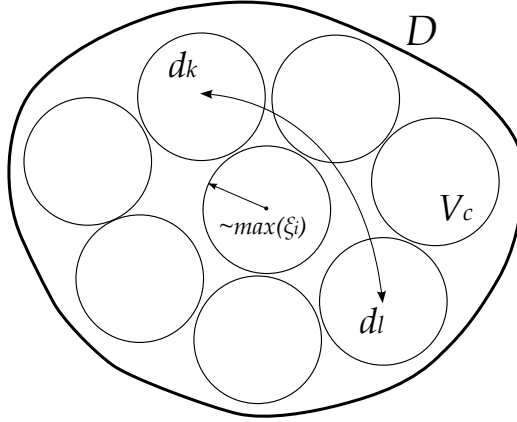
Figure 2.3: The sample's spatial domain $\mathcal{D}$ can be divided into small blobs d_i of approximate size V_c, which can be assumed to be approximately uncorrelated. The sum (2.18) is performed over points belonging to approximately uncorrelated blobs e.g., d_k and d_l. Thus the PDF of the sum will be approximately Gaussian, according to the Central Limit Theorem.

The univariate PDFs of the $X_{ij}(\mathbf{s}_k)$, are derived in Appendix A.2.

Lemma 2.2.1 (Covariance matrix $\mathbf{C}_{\hat{\mathbb{Q}}}$). *For a statistically stationary, and anisotropic GSRF, sampled at points $\mathbf{s}_1 \ldots \mathbf{s}_N$, the covariance matrix $\mathbf{C}_{\hat{\mathbb{Q}}}$ is defined by $C_{ij,kl} \equiv \mathrm{Cov}\left(\hat{Q}_{ij}, \hat{Q}_{kl}\right)$ for $i, j, k, l = 1, 2$, i.e.,*

$$\mathbf{C}_{\hat{\mathbb{Q}}} = \begin{bmatrix} C_{11,11} & C_{11,22} & C_{11,12} \\ C_{22,11} & C_{22,22} & C_{22,12} \\ C_{12,11} & C_{12,22} & C_{12,12} \end{bmatrix}. \tag{2.19}$$

The matrix (2.19) is symmetric. If we define the tensor covariance function

$$C_{ij,kl}(\mathbf{r}) = H_{ik}(\mathbf{r})H_{jl}(\mathbf{r}) + H_{il}(\mathbf{r})H_{jk}(\mathbf{r}), \tag{2.20}$$

then the six independent elements (upper triangular entries) of $\mathbf{C}_{\hat{\mathbb{Q}}}$ are given by the following series

$$C_{ij,kl} = \frac{1}{N^2} \sum_{\mathbf{r}_{nm}} C_{ij,kl}(\mathbf{r}_{nm})$$

$$= \frac{1}{N}\left[Q_{ik}\,Q_{jl} + Q_{il}\,Q_{jk}\right] + \frac{1}{N^2} \sum_{\mathbf{r}_{nm} \neq \mathbf{0}} C_{ij,kl}(\mathbf{r}_{nm}), \tag{2.21}$$

where $\mathbf{r}_{nm} = \mathbf{s}_n - \mathbf{s}_m$ $(n, m = 1, \ldots, N)$ is the lag vector between two locations $\mathbf{s}_n$ and $\mathbf{s}_m$.

Proof. The proof is given in Appendix A.3. ∎

The leading term $\frac{1}{N}\left[Q_{ik}\,Q_{jl} + Q_{il}\,Q_{jk}\right]$ in (2.21) leads to the non-parametric approximation of $f_{\partial,R}(\hat{R}, \hat{\partial})$ as shown below. The sums over $\mathbf{r}_{nm} \neq \mathbf{0}$ contribute parametric corrections that involve the covariance function. By definition, $\mathbf{C}_{\bar{\mathbb{Q}}}$ is a symmetric matrix, namely $C_{ij,kl} = C_{kl,ij}$.

Lemma 2.2.2 (Joint PDF of $\hat{\mathbf{Q}}$). *Let a statistically stationary and anisotropic GSRF $X(\mathbf{s})$ with covariance $c_{xx}(\mathbf{r})$ that is short-ranged and its spectral density satisfies $\tilde{C}_{xx}(\mathbf{k}) \underset{k\to\infty}{\sim} o(k^{-3-\epsilon})$ where $k = \|\mathbf{k}\|$ and $\epsilon > 0$. Then, the joint PDF of $\hat{Q}_{11}$, $\hat{Q}_{12}$, and $\hat{Q}_{22}$ tends to the following trivariate Gaussian as $N \to \infty, |\mathcal{D}| \to \infty$*

$$f_{\bar{\mathbb{Q}}}(\bar{\mathbb{Q}}; \mathbf{m}_{\mathbb{Q}}, \mathbf{C}_{\bar{\mathbb{Q}}}) = \frac{1}{(2\pi)^{3/2}\,\det(\mathbf{C}_{\bar{\mathbb{Q}}})^{1/2}}\,e^{-\frac{1}{2}(\bar{\mathbb{Q}}-\mathbf{m}_{\mathbb{Q}})^t \mathbf{C}_{\mathbb{Q}}^{-1}(\bar{\mathbb{Q}}-\mathbf{m}_{\mathbb{Q}})}, \qquad (2.22)$$

where the ensemble mean $\mathbf{m}_{\mathbb{Q}}$ is given by

$$\mathbf{m}_{\mathbb{Q}}{}^t = \frac{1}{N}\sum_{k=1}^{N} \mathbb{E}\left[(X_{11}(\mathbf{s}_k), X_{22}(\mathbf{s}_k), X_{12}(\mathbf{s}_k))\right] = (Q_{11}, Q_{22}, Q_{12}), \qquad (2.23)$$

and the covariance matrix $\mathbf{C}_{\bar{\mathbb{Q}}}$ is defined by (2.19) and (2.21).

Proof. The proof is based on Theorem 2.2.1 and presented in Appendix A.4. The condition $\tilde{C}_{xx}(\mathbf{k}) \underset{k\to\infty}{\sim} O(k^{-3-\epsilon})$, $\epsilon > 0$ implies that $\forall k \to \infty, \exists \epsilon > 0, C_\infty > 0$, such that $\tilde{C}_{xx}(\mathbf{k}) \leq C_\infty/k^{3+\epsilon}$. This condition is not particularly restrictive, since it is satisfied by most finite-range, differentiable covariance functions, including the Gaussian, Matérn with $\nu > 1$ and Spartan covariance models. ∎

The normal probability plots of $\hat{Q}_{11}$, $\hat{Q}_{12}$, $\hat{Q}_{22}$ of simulated data, shown in Fig. 4.3, confirm the asymptotic normality of the univariate PDFs in agreement with the CLT.

2.2.3 Numerical Approximation of the Covariance Matrix $\mathbf{C}_{\bar{\mathbb{Q}}}$

A numerical approximation of the covariance of the slope tensor elements (2.21) require knowledge of the correlation function and the summation of the series (2.21) over *all* lag vectors. As we will show in the following,
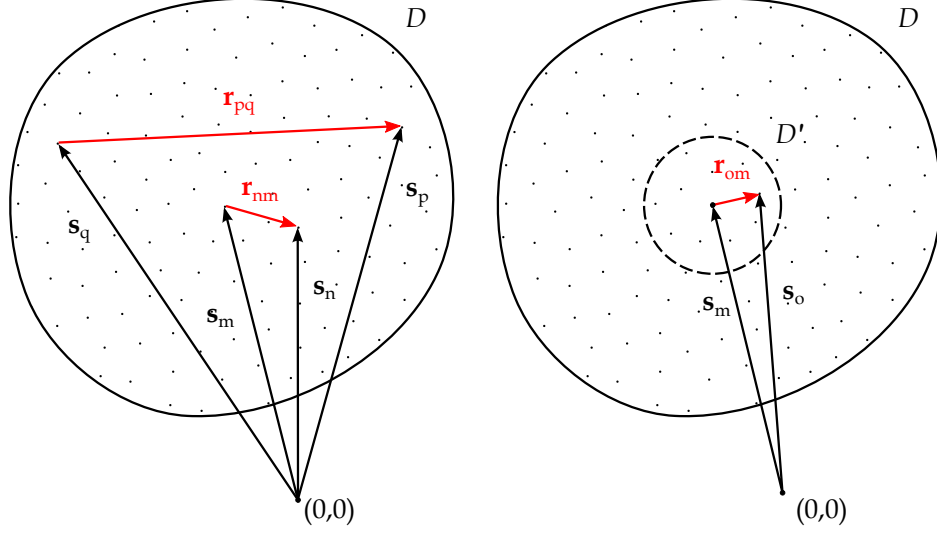
Figure 2.4: Approximation of the sum (2.21). The lag vector $\mathbf{r}_{pq}$ (left) has a minor contribution to the sum, in contrast to vectors with smaller lengths, e.g., $\mathbf{r}_{nm}$. In order to avoid summing over N^2 lag vectors, we can keep one position vector fixed, i.e., $\mathbf{s}_m$ and let the second vector i.e., $\mathbf{s}_o$ vary within a smaller domain $\mathcal{D}'$ with $N' \ll N^2$ points (right). The sum can be approximated by multiplying the partial sum with N.

the full summation can be avoided, especially in the case of dense regular grids, since an approximation can be obtained by performing the sum in a smaller set of lag vectors and generalizing the result to the whole domain $\mathcal{D}$.

Consider the points $\mathbf{s}_i$ belonging to a discretization of the domain $\mathcal{D}$, and $i = 1 \ldots N$. The lag vectors between two points in $\mathcal{D}$ also lie within $\mathcal{D}$. Equation (2.21) is a sum of the function $C_{ij,kl}(\mathbf{r})$ evaluated over N^2 lag vectors $\mathbf{r}$. If the dominant contribution of $C_{ij,kl}(\mathbf{r})$ to the sum is due to a subdomain $\mathcal{D}' \subset \mathcal{D}$ centered at $\mathbf{r} = \mathbf{0} \in \mathcal{D}'$ then the sum can be performed only within $\mathcal{D}'$. Neglecting edge effects, the result can be used for an approximation to the full sum as follows:

$$\underbrace{\sum_{\mathbf{r}_{nm} \in \mathcal{D}} C_{ij,kl}(\mathbf{r}_{nm})}_{N^2 \text{ terms}} \simeq N \times \underbrace{\sum_{\mathbf{r}_{nm} \in \mathcal{D}'} C_{ij,kl}(\mathbf{r}_{nm})}_{N' \text{ terms}} \tag{2.24}$$

The right-hand side sum is over $N' \ll N^2$ lag vectors; thus less computation time is required. The summation approximation is depicted schematically in Figure 2.4: The lag vector $\mathbf{r}_{pq}$ has length $\left\| \mathbf{r}_{pq} \right\| \gg \xi$ while $\mathbf{r}_{nm}$ has length

comparable (or smaller) to ξ. Instead of performing the sum over all the N^2 lag vectors, we choose a fixed position vector, e.g., $\mathbf{s}_m$ and perform the sum over the N' lag vectors contained in the smaller domain $\mathcal{D}'$. One of these lag vectors is $\mathbf{r}_{om}$. The approximation (2.24) requires full knowledge of the covariance function while the derivation of the non-parametric JPDF requires the restriction to evaluate (2.21) over the zero-length lag vectors only.

A good approximation is obtained if the domain $\mathcal{D}'$ is a domain where $C_{ij,kl}(\mathbf{r})$ is non-vanishing. In most cases, a domain with radius $\sim 3\xi$ is sufficient. For example, consider a Gaussian correlation function with unit variance. Figure 2.5a shows plots of $C_{ij,kl}(\mathbf{r})$ for an isotropic Gaussian covariance function of unit correlation length, and Figure 2.5b demonstrates plots of $C_{ij,kl}(\mathbf{r})$ for an anisotropic Gaussian correlation function having $\xi_1 = 1$, $\xi_2 = 2$ and anisotropy angle $\partial = 30°$. In both cases, a square domain with side $\sim 3\xi$ or $\sim 3\max(\xi_1, \xi_2)$ is appropriate for the summation approximation.

The accuracy of approximating the covariance matrix (2.19) (summation of the series (2.21)), by keeping the first term in the sums (2.21), is examined for the 1-D Gaussian covariance case, where the covariance matrix reduces to only one element, $C_{11,11} = 2/N^2 \times \sum_{n=1}^{N} H_{11}(r_n)^2$. In Figure 2.6, we plot the ratio of the first-term approximation $C_{11,11}^{(0)} = 2/N \times Q_{11}^2$ of the covariance to the covariance $C_{11,11}$ as a function of ξ/a where a is the lattice constant and L/ξ where L is the size of the 1D lattice. The ratio $C_{11,11}^{(0)}/C_{11,11}$ depends strongly on ξ/a and weakly on L/ξ. The first-term approximation becomes accurate for $\xi/a \lesssim 3$. The non-monotonic behavior of $C_{11,11}^{(0)}/C_{11,11}$ is due to the non-monotonic behavior of the square of the second derivative of the Gaussian correlation function $H_{11}(r)^2$, which is plotted in Fig. 2.7.

2.2.4 Probability Density Function of Slope Tensor Ratios

Albeit $\hat{\mathbf{Q}}$ involves three independent elements, the anisotropic parameters are determined from two slope tensor ratios, c.f. (2.14)–(2.15). Next, we derive the JPDF of the slope tensor ratios $f_{\mathbf{q}}(\hat{\mathbf{q}}; \mathbf{m}_{\mathbb{Q}}, \mathbf{C}_{\bar{\mathbb{Q}}})$ where $\hat{\mathbf{q}} = (\hat{q}_d, \hat{q}_o)^t$ and $\hat{q}_d$, $\hat{q}_o$ are the slope tensor ratio estimates as defined in Theorem 2.1.2, from the joint PDF of $\hat{\mathbf{Q}}$.

Lemma 2.2.3 (PDF of slope tensor ratios). *For a statistically stationary and anisotropic GSRF $X(\mathbf{s})$ with a covariance $c_{xx}(\mathbf{r})$ that satisfies the conditions of Lemma 2.2.2, the PDF $f_{\mathbf{q}}(\hat{\mathbf{q}}; \mathbf{m}_{\mathbb{Q}}, \mathbf{C}_{\bar{\mathbb{Q}}})$ of the slope tensor ratios $\hat{q}_d$, $\hat{q}_o$ is given*
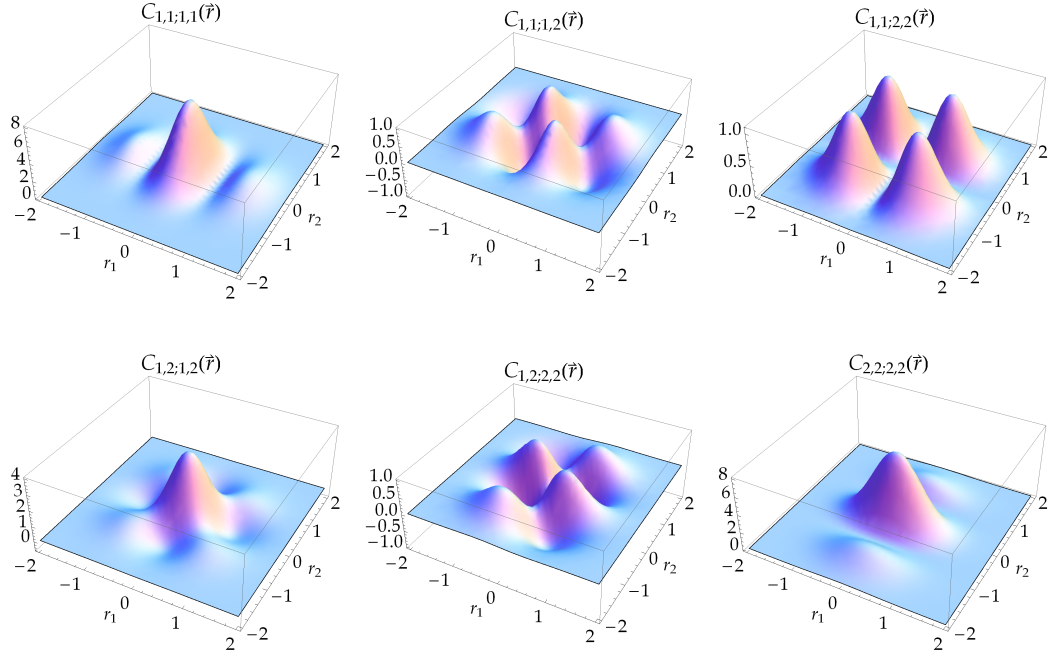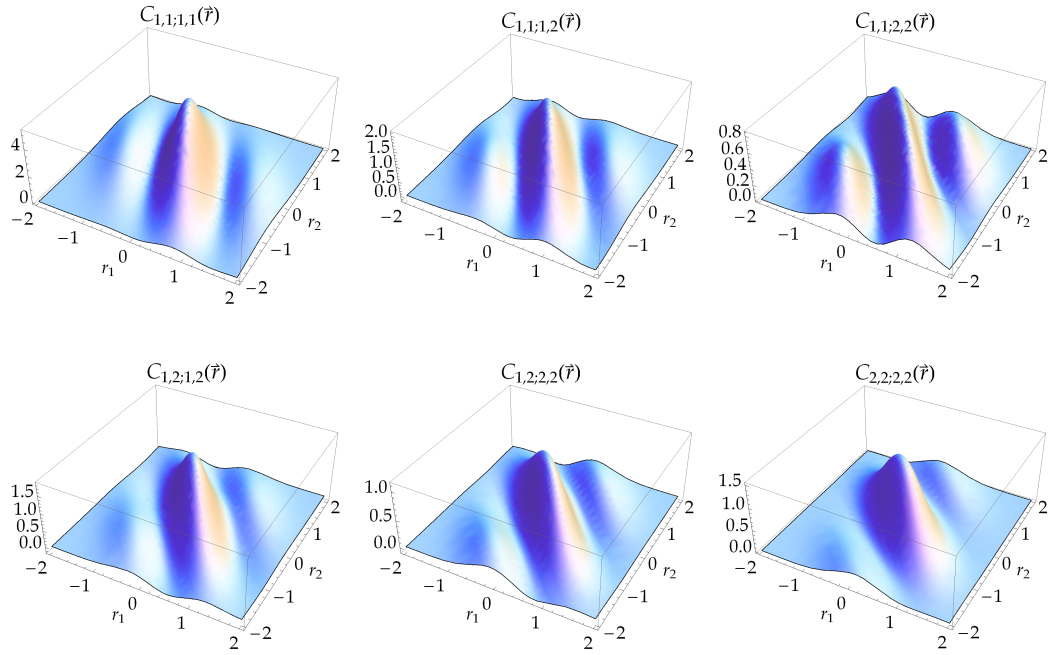
(a) Isotropic: $\xi = 1$, $\sigma_x^2 = 1$.

(b) Anisotropic: $\xi_1 = 1$, $\xi_2 = 2$, $\vartheta = 30°$, $\sigma_x^2 = 1$.

Figure 2.5: Plots of $C_{ij;kl}(\mathbf{r})$ for an (a) isotropic and (b) anisotropic Gaussian correlation function. In both cases $C_{ij;kl}(\mathbf{r})$ is essentially non-vanishing within a square of side $\sim 2\xi$ or $\sim 3\max(\xi_1, \xi_2)$.
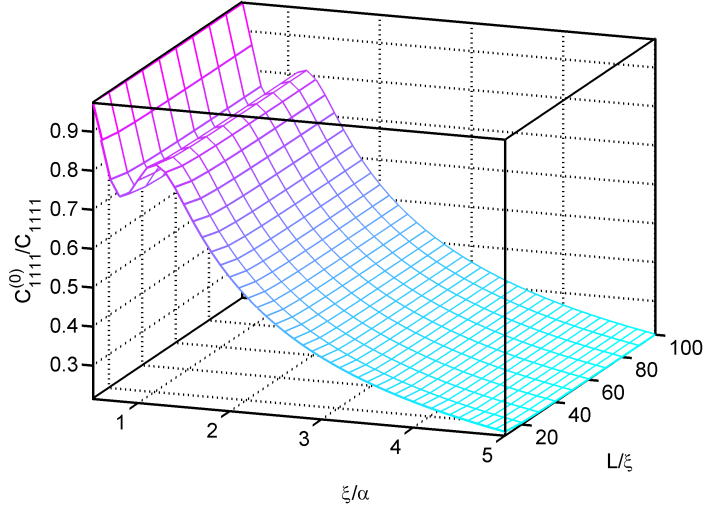
Figure 2.6: Numerical examination of the accuracy of the first-term approximation $C_{11,11}^{(0)}$ of the covariance $C_{11,11}$, for a 1D Gaussian correlation function. The ratio $C_{11,11}^{(0)}/C_{11,11}$ is plotted against ξ/a where a is the lattice constant and L/ξ where L is the size of the lattice. The approximation strongly depends on ξ/a and becomes accurate for $\xi/a \lesssim 3$. The non-monotonic decrease is due to the behavior of $H_{11}(r)^2$, depicted in Fig. 2.7.
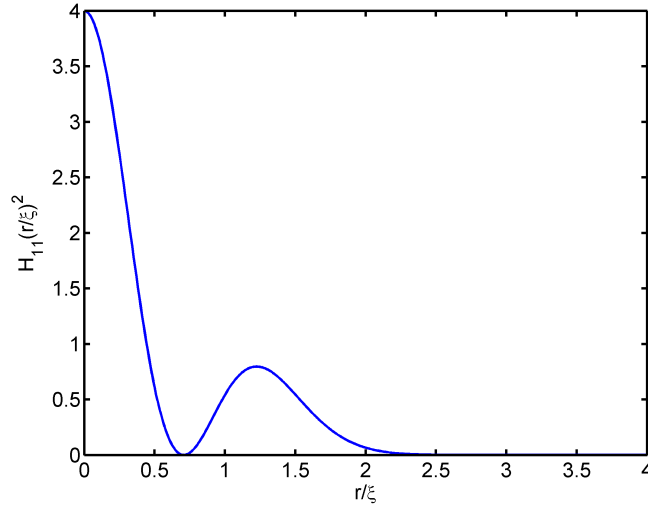
Figure 2.7: The non-monotonic behavior of the square of the second derivative of the Gaussian correlation function $H_{11}(r/\xi)^2$, as a function of the ratio r/ξ.

25

by the expression

$$f_{\mathbf{q}}(\hat{\mathbf{q}}; \mathbf{m}_{\mathbb{Q}}, \mathbf{C}_{\bar{\mathbb{Q}}}) = \frac{K}{8A^{\frac{5}{2}}} \, e^{-\frac{C}{2}} \left[\sqrt{2\pi} \, e^{\frac{B^2}{8A}} (B^2 + 4A) \, \mathrm{erfc}\left(\frac{B}{2\sqrt{2A}} \right) - 4B\sqrt{A} \right], \quad (2.25)$$

where $\mathrm{erfc}(\cdot)$ *is the complementary error function,*

$$\mathrm{erfc}(x) \doteq \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} \, dt.$$

A and B are functions of the random vector $\hat{\mathbf{q}}$ *as well as* $\mathbf{C}_{\bar{\mathbb{Q}}}$ *and* $\mathbf{m}_{\mathbb{Q}}$, *while C and K are functions of* $\mathbf{C}_{\bar{\mathbb{Q}}}$ *and* $\mathbf{m}_{\mathbb{Q}}$ *only. These functions are given below.*

$$A(\hat{\mathbf{q}}, \mathbf{C}_{\bar{\mathbb{Q}}}) = \hat{\mathbf{q}}^t \, \mathbf{C}_{\bar{\mathbb{Q}}}^{-1} \, \hat{\mathbf{q}}, \tag{2.26a}$$

$$B(\hat{\mathbf{q}}, \mathbf{m}_{\mathbb{Q}}, \mathbf{C}_{\bar{\mathbb{Q}}}) = -2 \, \mathbf{m}_{\mathbb{Q}}^t \, \mathbf{C}_{\bar{\mathbb{Q}}}^{-1} \, \hat{\mathbf{q}}, \tag{2.26b}$$

$$C(\mathbf{m}_{\mathbb{Q}}, \mathbf{C}_{\bar{\mathbb{Q}}}) = \mathbf{m}_{\mathbb{Q}}^t \, \mathbf{C}_{\bar{\mathbb{Q}}}^{-1} \, \mathbf{m}_{\mathbb{Q}}, \tag{2.26c}$$

$$K(\mathbf{C}_{\bar{\mathbb{Q}}}) = (2\pi)^{-3/2} \, [\det(\mathbf{C}_{\bar{\mathbb{Q}}})]^{-1/2}. \tag{2.26d}$$

Proof. Theorem A.1.2 is employed for the transformation of the JPDF $f_{\bar{\mathbb{Q}}}$ of the random vector $\bar{\mathbb{Q}}$, which is a trivariate Gaussian (2.22), to the JPDF $f_{\mathbf{q}}$ of the random vector $\mathbf{q}$. Details are given in Appendix A.5. ∎

The joint PDF of $\mathbf{q}$ given by (2.25) is quite different from the Gaussian form of the joint PDF of $\bar{\mathbb{Q}}$.

2.2.4.1 Dimensional Analysis

Since $\hat{q}_d, \hat{q}_o$ are dimensionless, so is $f_{\mathbf{q}}$; the coefficients A, B, C, K, nevertheless, have dimensions due to their dependence on $\mathbf{C}_{\bar{\mathbb{Q}}}$ and $\mathbf{m}_{\mathbb{Q}}$. In addition, (2.25) implicitly depends on N through $\mathbf{C}_{\bar{\mathbb{Q}}}$, c.f. (2.21). Below, we express (2.25) in terms of dimensionless coefficients and N, instead of the A, B, C, K.

First, let $c_{\mathrm{xx}}(\mathbf{r}) = \sigma_{\mathrm{x}}^2 \, \rho(\mathbf{r}; \vec{p})$ where $\vec{p} = (\xi_1, R, \vartheta)$ and $-1 \leq \rho(\mathbf{r}; \vec{p}) \leq 1$ is the dimensionless autocorrelation function. Based on (2.21) and the definition of the covariance Hessian matrix (2.11), it follows that $\mathbf{C}_{\bar{\mathbb{Q}}}^{-1} = N^2 \, \xi_1^4 \, \sigma_{\mathrm{x}}^{-4} \, \mathbb{P}(\hat{\mathbf{q}}; R, \vartheta)$, where $\mathbb{P}(\hat{\mathbf{q}}; R, \vartheta)$ is a dimensionless matrix that depends on the functional form of $\rho(\mathbf{r}; \vec{p})$. Similarly, $A = N^2 \, \xi_1^4 \, \sigma_{\mathrm{x}}^{-4} \, \mathcal{A}(\hat{\mathbf{q}}; R, \vartheta)$, $B = N^2 \, \xi_1^2 \, \sigma_{\mathrm{x}}^{-2} \, \mathcal{B}(\hat{\mathbf{q}}; R, \vartheta)$, $C = N^2 \, \mathcal{C}(\hat{\mathbf{q}}; R, \vartheta)$, and $K = N^3 \, \xi^6 \, \sigma_{\mathrm{x}}^{-6} \, \mathcal{K}(\hat{\mathbf{q}}; R, \vartheta)$, where the scalar coefficients $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{K}$ depend implicitly on $\rho(\mathbf{r})$. Based on the above scaling relations, we propose the following transformations that

involve the *dimensionless* functions $\tilde{B}(\hat{\mathbf{q}}; R, \partial)$, $\tilde{C}(\hat{\mathbf{q}}; R, \partial)$, $\tilde{K}(\hat{\mathbf{q}}; R, \partial)$, which are independent of N, σ_x^2 and ξ_1:

$$B = 2N \sqrt{2A}\tilde{B}, \qquad (2.27a)$$

$$C = 2\, N^2\, \tilde{C}, \qquad (2.27b)$$

$$K = \sqrt{2}A^{3/2}\, \tilde{K}. \qquad (2.27c)$$

Then, the slope-tensor-ratios PDF (2.25) is expressed as follows:

$$f_{\mathbf{q}}(\hat{\mathbf{q}}; \mathbf{m}_{\mathbb{Q}}, \mathbf{C}_{\bar{\mathbb{Q}}}) = \tilde{K}\, e^{-\tilde{C}N^2} \left[\sqrt{\pi}\, e^{\tilde{B}N^2}(2\tilde{B}^2\, N^2 + 1)\, \mathrm{erfc}(\tilde{B}\,N) - 2\,\tilde{B}\,N \right]. \qquad (2.28)$$

The JPDF and the resulting confidence regions depend on R, ∂ and the functional form of $\rho(\mathbf{r}; \bar{p})$.

2.2.4.2 Asymptotic Probability Density Function Limit

Since $A > 0$ and $B < 0$, $\tilde{B} < 0$, the argument of $\mathrm{erfc}(\tilde{B}\,N)$ in (2.28) is negative. Therefore, for $N \to \infty$, and $x = \tilde{B}\,N$, it follows that $\mathrm{erfc}(x) \sim 2 + e^{-x^2}\left(\pi^{-1/2}x^{-1} + O(x^{-2})\right)$ Abramowitz & Stegun [1970]. Hence, to leading-order in N, (2.28) is approximated as follows:

$$f_{\mathbf{q}}(\hat{\mathbf{q}}; \mathbf{m}_{\mathbb{Q}}, \mathbf{C}_{\bar{\mathbb{Q}}}) = 2\,\sqrt{\pi}\,\tilde{K}\left(2\tilde{B}^2\, N^2 + 1\right)\, e^{N^2(\tilde{B}^2 - \tilde{C})}. \qquad (2.29)$$

Numerical comparisons show that the absolute relative error between the exact, (2.28), and the approximate, (2.29), JPDF is less than $\sim 10^{-9}$ for $N = 50$ and $\sim 10^{-6}$ even for $N = 30$. The numerical computations are based on the non-parametric approximation of $\mathbf{C}_{\bar{\mathbb{Q}}}$ derived in Section 3.1.

2.2.5 Joint Probability Density Function of Anisotropy Statistics

Theorem 2.2.2 (Joint PDF of anisotropy statistics)**.** *For a statistically stationary and anisotropic GSRF $X(\mathbf{s})$ with a covariance $c_{xx}(\mathbf{r})$ that satisfies the conditions of Lemma 2.2.2, the JPDF of the statistics $\hat{R}$ and $\hat{\partial}$ is given by the following equation:*

$$f_{\partial, R}(\hat{R}, \hat{\partial}; \mathbf{m}_{\mathbb{Q}}, \mathbf{C}_{\bar{\mathbb{Q}}}) = \frac{2\hat{R}|\hat{R}^2 - 1|}{\left(\hat{R}^2 \cos^2 \hat{\partial} + \sin^2 \hat{\partial}\right)^3}\, f_{\mathbf{q}}(\hat{\mathbf{q}}; \mathbf{m}_{\mathbb{Q}}, \mathbf{C}_{\bar{\mathbb{Q}}}), \qquad (2.30)$$

where $f_{\mathbf{q}}(\hat{\mathbf{q}}; \mathbf{m}_{\mathbb{Q}}, \mathbf{C}_{\bar{\mathbb{Q}}})$ is given by (2.28).

Proof. The sequence of probability transformations concludes with the final transformation of $(\hat{R}, \hat{\partial})$ to $(\hat{R}, \hat{\partial})$, to obtain the JPDF $f_{\partial,R}$ of $(\hat{R}, \hat{\partial})$. Details are given in Appendix A.6. ∎

The function $f_{\partial,R}(\hat{R}, \hat{\partial})$ is clearly non-Gaussian; it depends on $\mathbf{m}_{\mathbb{Q}}$ and $\mathbf{C}_{\bar{\mathbb{Q}}}$ via $f_{\mathbf{q}}(\hat{\mathbf{q}}; \mathbf{m}_{\mathbb{Q}}, \mathbf{C}_{\bar{\mathbb{Q}}})$, while $\hat{\mathbf{q}}$ is expressed in terms of $(\hat{R}, \hat{\partial})$ using (2.14) and (2.15). If degrees are used, instead of radians, $f_{\partial,R}(\hat{R}, \hat{\partial})$ should be multiplied by $\pi/180$, which is the absolute value of the Jacobian determinant of the transformation from radians to degrees.

2.2.6   Confidence Regions

Herein, we derive expressions for confidence regions of the anisotropy statistics $(\hat{R}, \hat{\partial})$. Confidence regions are used instead of intervals due to the asymmetry of $f_{\partial,R}(\hat{R}, \hat{\partial})$. The confidence region for a probability level $p \in [0, 1]$ is the "volume" of parameter space which contains the sampled values of the statistics with probability p. for the same probability level p, the confidence volume $\mathcal{E}$ of $\bar{\mathbb{Q}}$ lie in $[0, \infty) \times [0, \infty) \times (-\infty, \infty)$; the confidence region $\mathcal{C}'$ of $\hat{\mathbf{q}}$ lie in $[0, \infty) \times (-\infty, \infty)$ and the confidence region $\mathcal{C}''$ of $(\hat{R}, \hat{\partial})$ lie in $[0, \infty) \times [-\pi/4, \pi/4)$.

The confidence region is defined by the following equivalent equations:

$$p = \begin{cases} \int_{\mathcal{E}} d\bar{\mathbb{Q}} f_{\bar{\mathbb{Q}}}(\hat{Q}_{11}, \hat{Q}_{22}, \hat{Q}_{12}; \mathbf{m}_{\mathbb{Q}}, \mathbf{C}_{\bar{\mathbb{Q}}}), & \mathcal{E} \subset [0, \infty) \times [0, \infty) \times (-\infty, \infty) \\ \int_{\mathcal{C}'} d\hat{q}_d d\hat{q}_o f_{\mathbf{q}}(\hat{q}_d, \hat{q}_o; \mathbf{m}_{\mathbb{Q}}, \mathbf{C}_{\bar{\mathbb{Q}}}), & \mathcal{C}' \subset [0, \infty) \times (-\infty, \infty) \\ \int_{\mathcal{C}''} d\hat{R} d\hat{\partial} f_{\partial,R}(\hat{R}, \hat{\partial}; \mathbf{m}_{\mathbb{Q}}, \mathbf{C}_{\bar{\mathbb{Q}}}). & \mathcal{C}'' \subset [0, \infty) \times [-\pi/4, \pi/4) \end{cases}$$

The above equations represent the evolution of the confidence region under the variable transformations $\bar{\mathbb{Q}} \to \hat{\mathbf{q}} \to (\hat{R}, \hat{\partial})$ as shown schematically in Fig. 2.8.

Corollary 2.2.1 (Parametric equation of confidence regions)**.** *For a statistically stationary and anisotropic GSRF $X(\mathbf{s})$ with a covariance $c_{xx}(\mathbf{r})$ that satisfies the conditions of Lemma 2.2.2, the confidence region corresponding to level $p \in (0, 1)$ in $(\hat{R}, \hat{\partial})$-space is given by the parametric equation*

$$\left[\tilde{B}^2(\hat{\mathbf{q}}; \mathbf{m}_{\mathbb{Q}}, \mathbf{C}_{\bar{\mathbb{Q}}}) - \tilde{C}(\hat{\mathbf{q}}; \mathbf{m}_{\mathbb{Q}}, \mathbf{C}_{\bar{\mathbb{Q}}}) \right] N^2 = \ln(1 - p), \tag{2.31}$$

where $\tilde{B}, \tilde{C}$ are given respectively by (2.27a)–(2.27b) and $\hat{\mathbf{q}}$ translates into
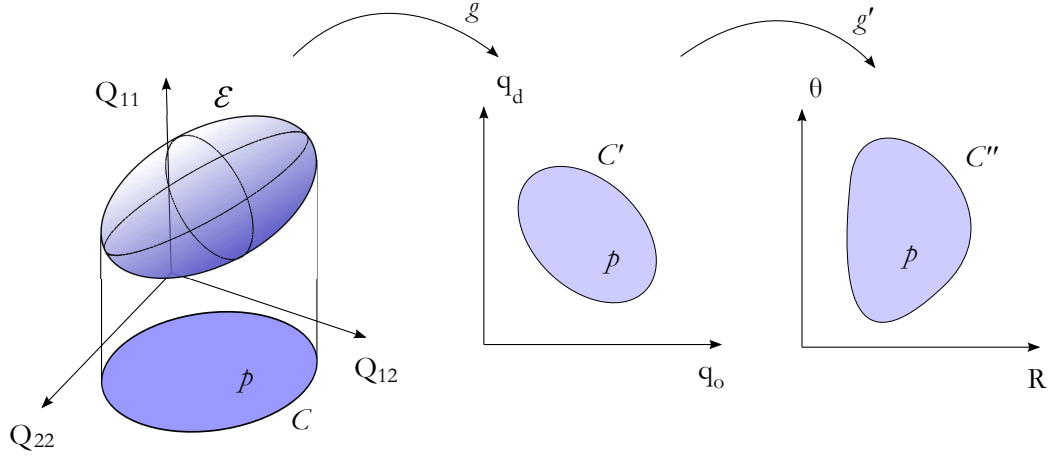
Figure 2.8: Schematic illustrating the transformation of confidence regions at level ℓ_p across different coordinate systems.

$(\hat{R}, \hat{\partial})$ *by means of*

$$\hat{\partial} = \frac{1}{2} \tan^{-1}\left(\frac{2\hat{q}_o}{1 - \hat{q}_d}\right), \tag{2.32}$$

$$\hat{R} = \left[1 + \frac{1 - \hat{q}_d}{\hat{q}_d - (1 + \hat{q}_d)\cos^2\hat{\partial}}\right]^{-1/2}, \tag{2.33}$$

which are the equivalent of (2.14) *and* (2.15) *for sample values.*

Proof. The JPDF $f_{\vec{\mathbb{Q}}}$ is given by the trivariate Gaussian (2.22) with mean $\mathbf{m}_{\mathbb{Q}}$ and covariance $\mathbf{C}_{\vec{\mathbb{Q}}}$. Hence, the confidence region of $\vec{\mathbb{Q}}$ is an ellipsoid in $[0, \infty) \times [0, \infty) \times (-\infty, \infty)$ whose surface satisfies the equation

$$(\vec{\mathbb{Q}} - \mathbf{m}_{\mathbb{Q}})^t \, \mathbf{C}_{\vec{\mathbb{Q}}}^{-1} \, (\vec{\mathbb{Q}} - \mathbf{m}_{\mathbb{Q}}) = \ell_p, \tag{2.34}$$

where $\ell_p = F^{-1}(\chi^2 = p, 2)$ is the *inverse of the chi-square cumulative distribution function* with $\nu = 2$ degrees of freedom Siotani [1964]. Under the transformation $\vec{\mathbb{Q}} \to \hat{\mathbf{q}}$, the ellipsoid is projected onto an ellipse, and following the transformation $\hat{\mathbf{q}} \to (\hat{R}, \hat{\partial})$ into an asymmetric convex curve, as shown schematically in Fig. 2.8. Based on (A.22), the equation of the corresponding ellipsoid in $(u, \hat{q}_d, \hat{q}_o)$-space is given by

$$A(\hat{\mathbf{q}}, \mathbf{C}_{\vec{\mathbb{Q}}}) \, u^2 + B(\hat{\mathbf{q}}, \mathbf{m}_{\mathbb{Q}}, \mathbf{C}_{\vec{\mathbb{Q}}}) \, u + C(\mathbf{m}_{\mathbb{Q}}, \mathbf{C}_{\vec{\mathbb{Q}}}) - \ell_p = 0,$$

where the coefficients A, B, C are given by (2.26). For the quadratic equation to have a unique real solution $u = \hat{Q}_{11}$ for any $\hat{\mathbf{q}}$, the discriminant should

29

vanish, i.e.,

$$B^2(\hat{\mathbf{q}}; \mathbf{m}_{\mathbb{Q}}, \mathbf{C}_{\bar{\mathbb{Q}}}) - 4A(\hat{\mathbf{q}}; \mathbf{C}_{\bar{\mathbb{Q}}}) \left[C(\mathbf{m}_{\mathbb{Q}}, \mathbf{C}_{\bar{\mathbb{Q}}}) - \ell_p \right] = 0. \qquad (2.35)$$

We can verify, using (2.26), that (2.35) represents an ellipse in the space of $\hat{\mathbf{q}}$, i.e.,

$$\hat{\mathbf{q}}^t \left[(\mathbf{C}_{\bar{\mathbb{Q}}}^{-1} \mathbf{m}_{\mathbb{Q}}) (\mathbf{C}_{\bar{\mathbb{Q}}}^{-1} \mathbf{m}_{\mathbb{Q}})^t - \left(\mathbf{m}_{\mathbb{Q}}^t \mathbf{C}_{\bar{\mathbb{Q}}}^{-1} \mathbf{m}_{\mathbb{Q}} - \ell_p \right) \mathbf{C}_{\bar{\mathbb{Q}}}^{-1} \right] \hat{\mathbf{q}} = 0.$$

Using the dimensionless scaling functions (2.27) in (2.35) we obtain the parametric equation $2(\tilde{C} - \tilde{B}^2)N^2 = \ell_p$, where by definition $F(\ell_p, \nu = 2) = p$; since $F(x, \nu = 2) = 1 - \exp(-x/2)$ it follows that $\ell_p = -2\ln(1 - p)$ finally leading to (2.31). ∎

Non-parametric Anisotropy Statistics

3.1 Non-parametric Joint Probability Density Function and Confidence Region

The expressions for the JPDF, $f_{\hat{\partial},R}(\hat{R}, \hat{\partial}; \mathbf{m}_{\mathbb{Q}}, \mathbf{C}_{\bar{\mathbb{Q}}})$, and the confidence regions of the anisotropy statistics derived above depend on the matrix $\mathbf{C}_{\bar{\mathbb{Q}}}$, given by (2.21). This matrix involves a series that does not, in general, admit a closed form expression. An approximate, analytical expression can be derived by keeping only the leading term in the covariance matrix series (2.21). The truncation is justified if $c_{xx}(\mathbf{r})$ has short-range correlations, which implies that $H_{ij}(\mathbf{r})$ decays fast for $\|\mathbf{r}\| \gg \max(\xi_1, \xi_2)$. Then, the *leading-order approximation of* $\mathbf{C}_{\bar{\mathbb{Q}}}$ is given by the zero order term in the series expansions i.e.,

$$\mathbf{C}_{\bar{\mathbb{Q}}} \approx \frac{2}{N} \begin{bmatrix} \mathcal{Q}_{11}^2 & \mathcal{Q}_{12}^2 & \mathcal{Q}_{11}\mathcal{Q}_{12} \\ \mathcal{Q}_{12}^2 & \mathcal{Q}_{22}^2 & \mathcal{Q}_{12}\mathcal{Q}_{22} \\ \mathcal{Q}_{11}\mathcal{Q}_{12} & \mathcal{Q}_{12}\mathcal{Q}_{22} & \frac{1}{2}(\mathcal{Q}_{12}^2 + \mathcal{Q}_{11}\mathcal{Q}_{22}) \end{bmatrix} = \mathbf{C}_{\bar{\mathbb{Q}}}^{(0)}. \tag{3.1}$$

Since $\mathbf{C}_{\bar{\mathbb{Q}}}^{(0)}$ does not account for correlations, we expect that it will lead to a joint PDF with higher uncertainty, and hence wider confidence regions, than the true PDF. We confirmed this hypothesis by means of numerical simulations (see Section 4.1). In addition, $\mathbf{C}_{\bar{\mathbb{Q}}}^{(0)}$, as evidenced in (2.21), has an $1/N$ scaling prefactor in contrast with $1/N^2$ for the truncated terms. To accommodate this sample-size dependence, the scaling relations (2.27) are accordingly modified below.

Theorem 3.1.1 (Non-parametric JPDF). *For a statistically stationary and anisotropic GSRF $X(\mathbf{s})$ with covariance $c_{xx}(\mathbf{r})$, anisotropy parameters (R, ∂), sampled at N points and satisfies the conditions of Lemma 2.2.2, the non-parametric approximation of the JPDF for the anisotropy statistics $(\hat{R}, \hat{\partial})$ is expressed as*

$$f_{\partial,R}^{(0)}(\hat{R}, \hat{\partial}; R, \partial, N) = \left|\det(\mathbf{J}_{\partial,R})\right| f_{\mathbf{q}}^{(0)}(\hat{R}, \hat{\partial}; R, \partial, N), \tag{3.2a}$$

where

$$f_{\mathbf{q}}^{(0)}(\hat{R}, \hat{\partial}; R, \partial, N) = 2\sqrt{\pi}\,\tilde{K}_0\left(2\tilde{B}_0^2 N + 1\right) e^{N(\tilde{B}_0^2 - 1/2)}. \tag{3.2b}$$

The coefficients $\tilde{B}_0$ and $\tilde{K}_0$ are given by the following expressions

$$\tilde{B}_0 = (2\tilde{A}_0)^{-1/2}\left[(R^2 - 1)(\hat{R}^2 - 1)\cos(2(\partial - \hat{\partial})) - (R^2 + 1)(\hat{R}^2 + 1)\right], \tag{3.3a}$$

$$\tilde{K}_0 = (\pi\tilde{A}_0)^{-3/2} R^3 \left[(\hat{R}^2 + 1) - (\hat{R}^2 - 1)\cos(2\hat{\partial})\right]^3, \tag{3.3b}$$

whereas $\tilde{A}_0$ is given by

$$\tilde{A}_0 = (\hat{R}^2 - 1)^2(R^2 - 1)^2\cos(4(\hat{\partial} - \partial)) - 4(\hat{R}^4 - 1)(R^4 - 1)\cos(2(\hat{\partial} - \partial)) \tag{3.3c}$$

$$+ (\hat{R}^4 + 1)(3R^4 + 2R^2 + 3) + 2\hat{R}^2(R^2 - 1)^2.$$

Proof. Equations (3.3a)-(3.3c) are derived by replacing $\mathbf{C}_{\bar{\mathbb{Q}}}$ with $\mathbf{C}_{\bar{\mathbb{Q}}}^{(0)}$, defined by (3.1), in (2.26). Also the asymptotic result (2.29) of Lemma 2.2.3 is modified accordingly to accommodate the $1/N$ scaling prefactor of $\mathbf{C}_{\bar{\mathbb{Q}}}^{(0)}$ and obtain the non-parametric JPDF $f_{\mathbf{q}}^{(0)}$. The $\tilde{B}_0$ and $\tilde{K}_0$ are obtained from the following dimensionless scaling functions (2.27a)-(2.27c)

$$B = 2\sqrt{N}\,\sqrt{2A}\,\tilde{B}_0,$$
$$C = 2N\,\tilde{C}_0,$$
$$K = \sqrt{2}A^{3/2}\tilde{K}_0,$$

and from the B, C, and K given respectively by (2.27). In the non-parametric approximation, the coefficient $\tilde{C}$ in the exponent of (2.29) is reduced to $\tilde{C}_0 = 1/2$ in the exponent of (3.2b). ∎

Figures 3.1 and 3.2 demonstrate representative plots of the non-parametric JPDF based on (3.2). Note the bimodal structure of the JPDF for $N = 100$

(a) $N = 100$ (b) $N = 500$
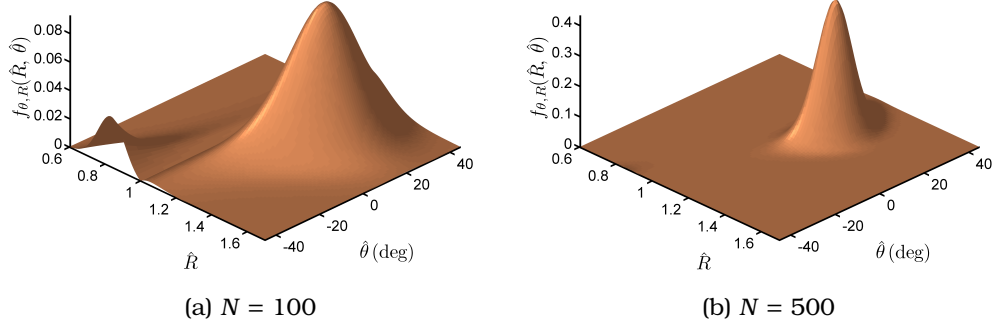
Figure 3.1: Non-parametric JPDF for $R = 1.2$, $\partial = 20°$ and $N = 100, 500$ based on (3.2).

(a) $N = 100$ (b) $N = 500$
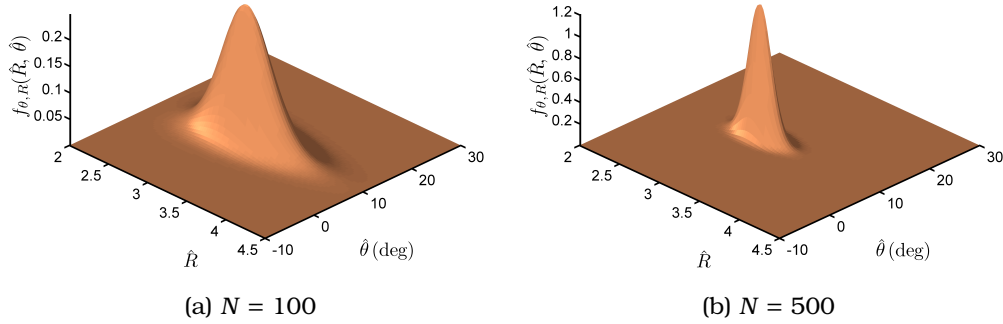
Figure 3.2: Non-parametric JPDF for $R = 3$, $\partial = 10°$ and $N = 100, 500$ based on (3.2).

in Fig. 3.1 with one mode at $R = 1.2$ and the other (smaller) at $R \approx 0.8$. This is due to the considerable spread of $\hat{\partial}$, which results from the relatively small number of sampling points, and the degeneracy of the solution, i.e., the fact that the combination (R, ∂) is equivalent to $(1/R, \partial - \pi/2)$; hence, there is an extended degenerate peak at $(0.83, -70°)$, part of which is folded back into the primary domain. On the other hand, for $R = 3$ the smaller dispersion of $\hat{\partial}$ leads to a single mode even for $N = 100$. Anisotropy angle uncertainty increases as $R \to 1$ as expected.

Corollary 3.1.1 (Non-parametric confidence region)**.** *For a statistically stationary and anisotropic GSRF $X(\mathbf{s})$ with a covariance $c_{xx}(\mathbf{r})$ that satisfies the conditions of Lemma 2.2.2, the non-parametric equation of the confidence region corresponding to level p is given by*

$$\tilde{B}_0^2 - \frac{1}{2} = \frac{\ln(1 - p)}{N}. \tag{3.4}$$

Proof. The equation is derived by applying the modified scaling relations of Theorem (3.1.1) to the general result (2.35) of Theorem (2.2.1). ∎

3.2 Statistical Test of Isotropy

Theorem 3.2.1 (Isotropic ratio)**.** *Let $X(\mathbf{s})$ be a statistically isotropic GSRF $(R = 1)$ with correlation length ξ sampled at N points. In addition, assume that (i) $|\mathcal{D}| \gg \xi^2$ and (ii) $N \gg 1$. The sample values of the statistic $\hat{R}$ are contained with probability p in the following interval:*

$$\hat{R} \in \left[\sqrt{\frac{1 - 2\sqrt{a_{p;N}(1 - a_{p;N})}}{1 - 2a_{p;N}}}, \ \sqrt{\frac{1 + 2\sqrt{a_{p;N}(1 - a_{p;N})}}{1 - 2a_{p;N}}} \right], \qquad (3.5)$$

where $\ell_p = F^{-1}(\chi^2 = p, 2) = -2\ln(1 - p)$ is the inverse of the chi square cumulative distribution function with two degrees of freedom and $a_{p;N} = \ell_p/N$.

Proof. $|\mathcal{D}| \gg \xi^2$ and $N \gg 1$ enforce the asymptotic conditions of Corollary 3.1.1. For $R = 1$ the trigonometric terms in $\tilde{B}_0$, i.e., in (3.3a) vanish, showing explicitly that the confidence region is independent of $\hat{\partial}$. Plugging the resulting (3.3a) in (3.4), the following parametric equation is obtained: $N(\hat{R}^2 - 1)^2 - 2\ell_p(\hat{R}^4 + 1) = 0$. The constraint $N > 2\ell_p$ ensures that the roots are real numbers. This is satisfied for $N \gg 1$; for example, for $p = 0.95$ it implies $N > 12$. The two admissible (positive) roots of the quartic equation are given by (3.5). ∎

Equation (3.5) is independent of the covariance function and thus provides a non-parametric approximation of the confidence region for $\hat{R}$. The JPDF (3.2) is independent of ∂ and $\hat{\partial}$ for $R = 1$. The dependence on N of the 95% confidence interval for $\hat{R}$ is shown in Fig. 3.3. The PDF, $f_R(\hat{R})$, of $\hat{R}$ for $R = 1$ and $N = 100$ is shown in Fig. 3.4, including the 95% confidence region predicted by (3.5). Note that the PDF vanishes, instead of peaking, at $\hat{R} = 1$. This is not an artifact of the non-parametric approximation, since the complete PDF (2.30) also vanishes at $\hat{R} = 1$. This is due to the root of the Jacobian (A.23) at $\hat{R} = 1$. Isotropy corresponds to a single point $(1, 0)$ in the $(\hat{q}_d, \hat{q}_o)$-space which in turn is mapped to the straight line $\hat{R} = 1$ in the $(\hat{R}, \hat{\partial})$-space and $\hat{\partial}$ can take any value in $[-\pi/4, \pi/4)$. The vanishing of the density at $\hat{R} = 1$ is also borne out in numerical simulations that do not use the Jacobian (see Figure 4.6d below). This counter-intuitive behavior of $f_R(\hat{R})$, emphasizes the usefulness of Theorem 3.2.1 as a test for isotropy.

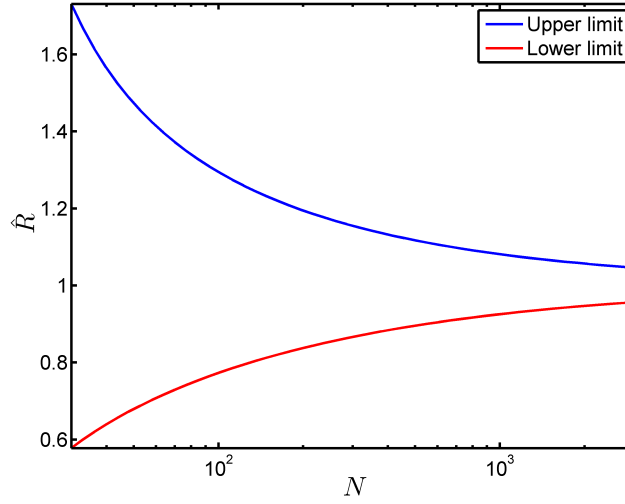Figure 3.3: Evolution of the lower and upper limits of the 95% confidence interval for the isotropic case ($R = 1$) versus the number of sampling points N based on (3.5).
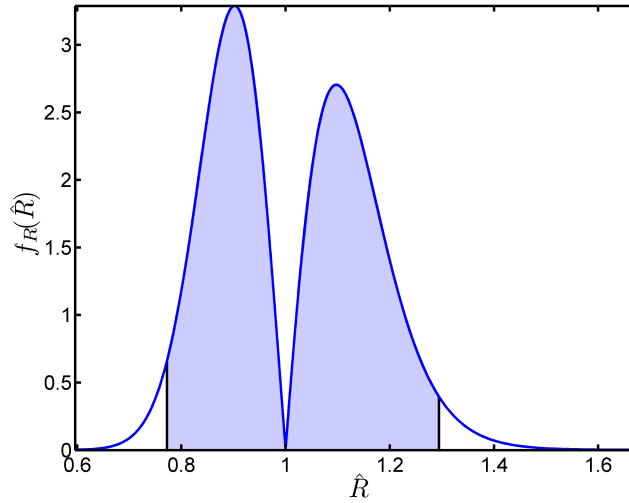

Figure 3.4: Non-parametric PDF, $f_R(\hat{R})$, for the isotropic case ($R = 1$) and $N = 100$ sampling points based on (3.2). The shaded area represents the 95% confidence interval, $(0.77, 1.29)$, of $\hat{R}$ based on (3.5).

35

Applications

IN the following, values of statistics derived from a sample of $N \gg 1$ points will be denoted by a star superscript, e.g., $\hat{Q}_{ij}^*, \hat{R}^*, \hat{\partial}^*$. Average values of the statistics over M different samples will be denoted by a bar over the respective symbol, i.e., $\overline{Q}_{ij}$. Estimation of $\hat{Q}_{ij}^*$ is based on discretized partial derivative operators, $\hat{\partial}X_i(\mathbf{s}_k)$, where $i = 1, 2$. Discretization introduces errors that increase as the sampling pattern becomes sparser. A "good" sampling pattern is characterized by a typical distance $\hat{a}$ between nearest neighbors which is approximately uniform (ideally, a regular lattice pattern is best) and $\hat{a} \ll \min(\xi_1, \xi_2)$, where ξ_1, ξ_2 are the principal correlation lengths. Different possibilities for $\hat{\partial}X_i(\mathbf{s}_k)$ are investigated in Chorti & Hristopulos [2008]. The parameters $\mathbf{m}_\mathbb{Q}$ and $\mathbf{C}_{\bar{\mathbb{Q}}}$ in (2.30) are unknown since they represent ensemble properties. For simulated data $\mathbf{m}_\mathbb{Q}$ and $\mathbf{C}_{\bar{\mathbb{Q}}}$ are replaced by the averages of the sample estimates, $\mathbf{m}_\mathbb{Q} \approx (\overline{Q}_{11}, \overline{Q}_{22}, \overline{Q}_{12})^t$ and $\mathbf{C}_{\bar{\mathbb{Q}}} \approx \mathbf{C}_{\overline{\mathbf{Q}}}$. In the non-parametric approximation, $\mathbf{C}_{\bar{\mathbb{Q}}}^{(0)}$ is obtained from (3.1) by replacing Q_{ij} with $\overline{Q}_{ij}$.

4.1 Simulated Lattice Random Fields

We generate multiple SRF realizations with specified (R, ∂) to validate the expression for the confidence region of the anisotropy statistics (3.4). The simulations are conducted on 100×100 square grids with lattice constant $a = 1$, by means of the Fourier Filtering Method Lantuéjoul [2002]; Pardo-Igúzquiza & Chica-Olmo [1993]. We use Gaussian, $c_{xx}(\mathbf{r}) = \sigma_x^2 \exp(-r^2/\xi^2)$, and Matérn, $c_{xx}(\mathbf{r}) = \sigma_x^2 \, 2^{1-\nu} \, r^\nu \, \Gamma(\nu)^{-1} \, \xi^{-\nu} K_\nu \, (r/\xi)$, covariance functions[1]. In

---

[1]These expressions correspond to the isotropic case.

the Gaussian case, the range of correlations is controlled by ξ but in the Matérn case by both ξ and the *smoothness parameter* ν; the latter adjusts the differentiability of the random field between two extremes obtained for $\nu = 1$ (non-differentiable exponential function) and $\nu \to \infty$ (infinitely differentiable Gaussian function). For given ξ, the field becomes smoother by increasing ν. To compensate for this effect, we use *rescaled correlation lengths* $\tilde{\xi} = \xi/\ell_c$ where ℓ_c is the integral scale factor. For Matérn correlations in $d = 2$ it follows that $\ell_c = 2\sqrt{\pi\nu}$, while for Gaussian correlations $\ell_c = \pi$, Hristopulos & Žukovič [2011].

The non-parametric confidence region obtained by (3.4), is compared with the CHI anisotropy estimates obtained from $N_s = 1000$ SRF samples. For each sample, we estimate the expectation (2.12) by means of the spatial average (2.18). Then, we obtain estimates $(\hat{R}^*, \hat{\partial}^*)$ of the anisotropy parameters by applying (2.16)–(2.17). We assume that (R, ∂) are unknown *a priori*, and we estimate them based on $\overline{R}$, $\overline{\partial}$. The latter are obtained by calculating the average slope tensor

$$\overline{\mathbf{Q}} = \frac{1}{N_s}\sum_{i=1}^{N_s}\hat{\mathbf{Q}}_i^*, \tag{4.1}$$

over the available realizations and then applying Theorem 2.1.2. The use of $\overline{\mathbf{Q}}$ helps to reduce biases due to finite grid size and the approximation of derivatives by means of finite differences Chorti & Hristopulos [2008]. In addition, it compensates for potential deviations of the simulated SRF from the target anisotropy values due to the Fourier filtering method.

In Figs. 4.1, 4.2 and 4.4 we compare the cloud of $(\hat{R}^*, \hat{\partial}^*)$ estimates obtained from each realization with the non-parametric confidence region obtained by (3.4). The latter is denoted by the solid lines that contain the cloud. The estimated anisotropy vector, based on $\overline{\mathbf{Q}}$, is denoted by a small circle inside the cloud. In addition, we include the parametric confidence region given by (2.31); $\mathbf{C}_{\bar{\mathbb{Q}}}$ is estimated by numerical summation of the series (2.21) using a square window function of side $\sim 3\tilde{\xi}$ around a grid point $\mathbf{s}_n$ at the center of the grid, as explained in subsection 2.2.3. This restricts the sum over $\mathbf{s}_m$ to correlated neighbors. The summation over all $\mathbf{s}_n$ is approximated by multiplying the result with the grid size[2]. So long as (i) the window area exceeds $[2\max(\tilde{\xi}_1, \tilde{\xi}_2)]^2$ and (ii) $a < \min(\tilde{\xi}_1, \tilde{\xi}_2)$, the truncated approximation is stable for different box sizes and shapes

---

[2]This approximation introduces errors in the summands from points near the grid boundaries; however, the fraction of boundary points for an $L \times L$ grid varies as $O(1/L)$ and is thus negligible for large grids.

(e.g., rectangular window). Fig. 4.1 investigates the *isotropic* case ($R = 1$) while the anisotropic case $R = 1.5, \vartheta = -30°$ is considered in Fig. 4.4. The non-parametric confidence region is more extended than the cloud and also encloses the parametric confidence region (contour lines inside the cloud). This result is justified, since the non-parametric approximation excludes higher-order correlated terms that reduce the uncertainty in $\mathbf{C}_{\bar{\mathbb{Q}}}$. The assumption of slope tensor normality, which was based on the application of CLT in Lemma 2.2.2, is graphically confirmed by the normal probability plots in Figs. 4.3a–4.3c.

We conducted numerical experiments (not shown here) for several values of the ratio $\tilde{\xi}/a$ and various N in order to confirm that the non-parametric JPDF is more extended in parameter space more than the actual JPDF. When $\tilde{\xi}/a \to 0$, i.e., as the simulated SRF tends to random uncorrelated noise, the scatter cloud of the anisotropy estimates expands and tends to fill the non-parametric confidence region. On the other hand, as $\tilde{\xi}/a$ increases, i.e., for dense sampling of the SRF, the scatter cloud shrinks inside the corrected confidence region. These observations confirm that the non-parametric approximation encloses the actual confidence region. This is expected from an information-theoretic viewpoint, since the additional information incorporated in the covariance function should lead to less uncertainty (i.e., a tighter confidence region) than the non-parametric approximation which discards covariance terms for non-zero distances.

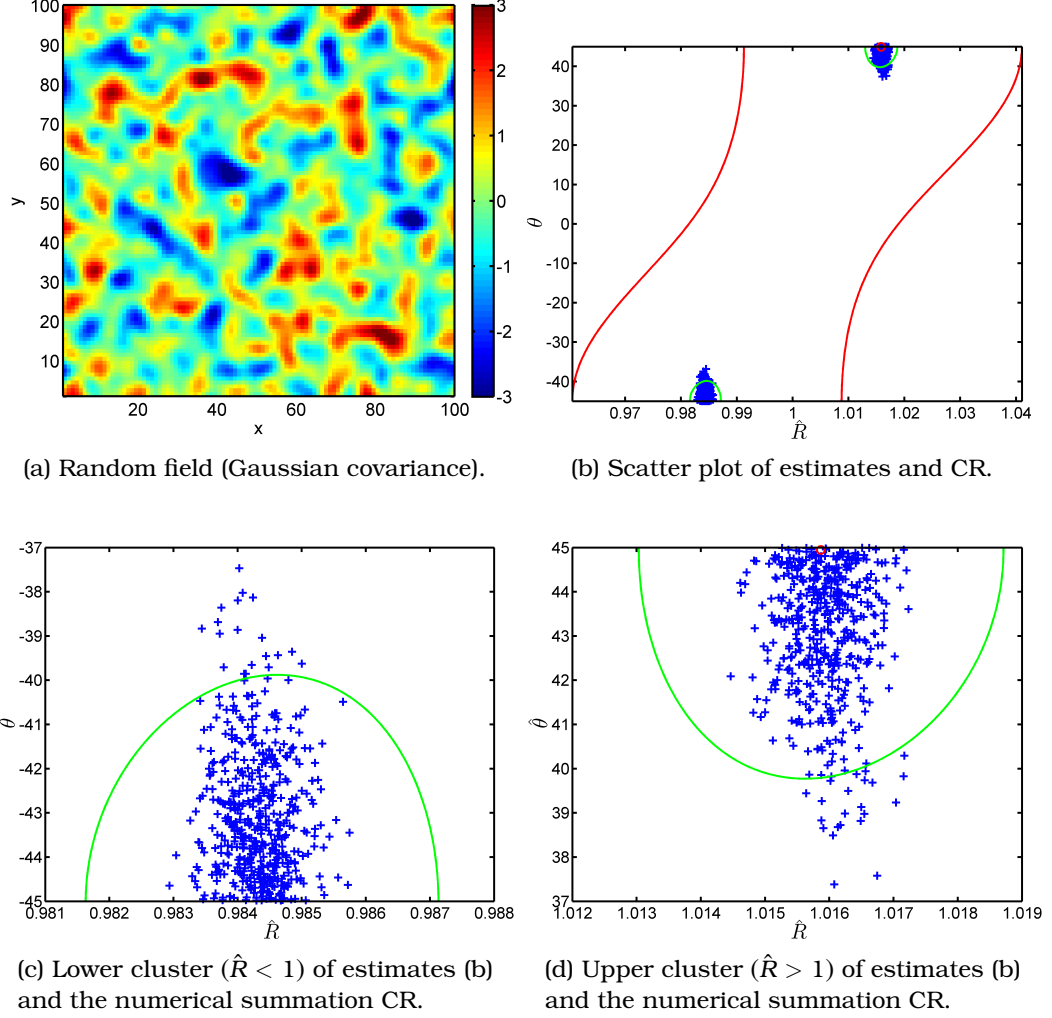(a) Random field (Gaussian covariance).

(b) Scatter plot of estimates and CR.

(c) Lower cluster ($\hat{R} < 1$) of estimates (b) and the numerical summation CR.

(d) Upper cluster ($\hat{R} > 1$) of estimates (b) and the numerical summation CR.

Figure 4.1: (a) A realization of an isotropic SRF with Gaussian correlations ($\tilde{\xi} = 4$) over an 100×100 square grid. (b) Scatter plot of anisotropy estimates (crosses) and non-parametric confidence region (outer contours) for 1000 realizations of isotropic RF with Gaussian correlations. Confidence regions are based on anisotropy parameters obtained from the mean slope tensor elements. Non-parametric confidence regions (outer contours) are obtained by (3.4). Tighter confidence regions (inner contours) are calculated numerically using the appropriate covariance function and incorporate the covariance-dependent terms in (2.21). (c)–(d): The clusters of the scatter plot (b) in detail, depicting the estimates (crosses), numerical summation CR (contour) and the anisotropy parameters (circle, upper border of (b) and (d)) used for the calculation of the CR. Confidence interval for isotropy is $(\hat{R}_-, \hat{R}_+) = (0.975, 1.025)$.
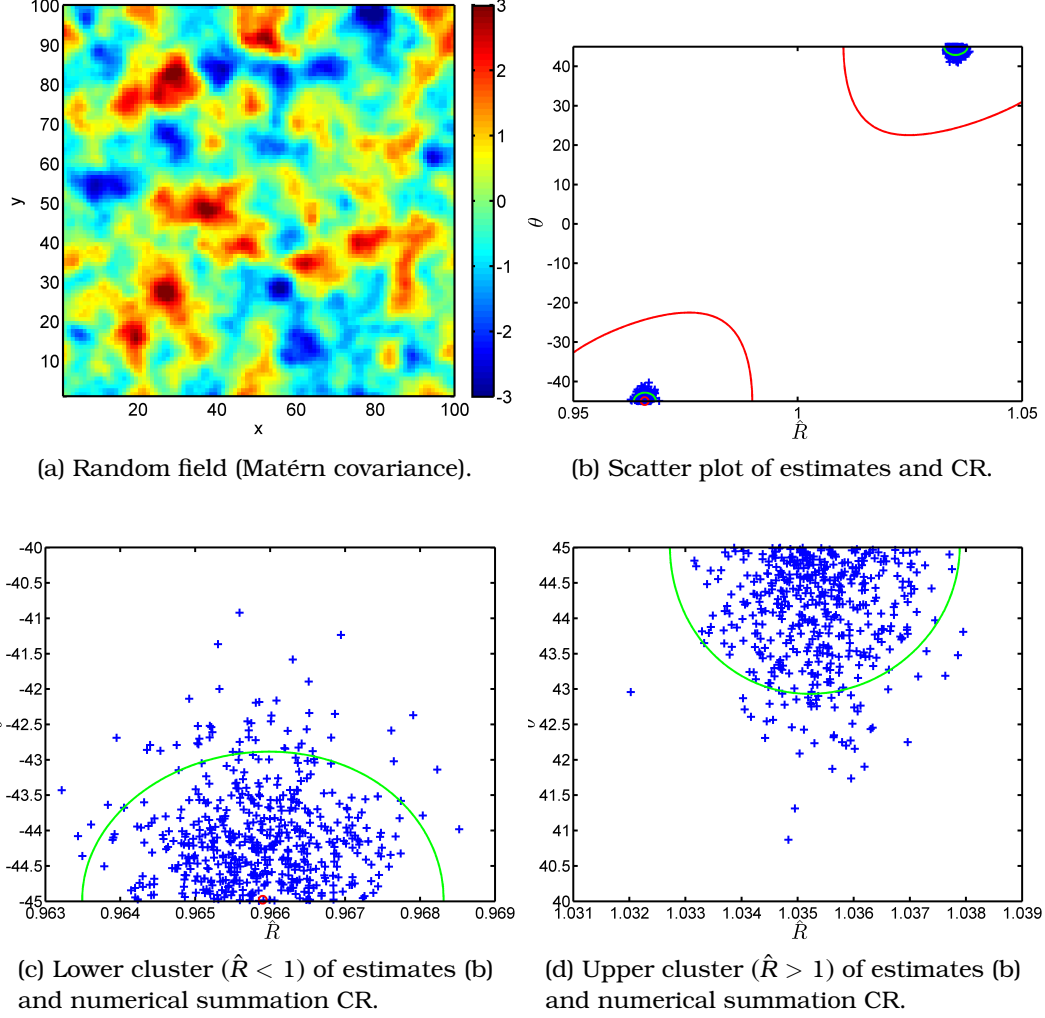
(a) Random field (Matérn covariance).

(b) Scatter plot of estimates and CR.

(c) Lower cluster ($\hat{R} < 1$) of estimates (b) and numerical summation CR.

(d) Upper cluster ($\hat{R} > 1$) of estimates (b) and numerical summation CR.

Figure 4.2: (a) A realization of an isotropic SRF with Matérn correlations ($\tilde{\xi} = 2.51$, $\nu = 2$) over an 100×100 square grid. (b) Scatter plot of anisotropy estimates (crosses) and non-parametric confidence region (outer contours) for 1000 realizations of isotropic random field with Matérn correlations. Confidence regions are based on anisotropy parameters obtained from the mean slope tensor elements. Non-parametric confidence regions (outer contours) are obtained by (3.4). Tighter confidence regions (inner contours) are calculated numerically using the appropriate covariance function and incorporate the covariance-dependent terms in (2.21). (c)–(d): The clusters of the scatter plot (b) in detail, depicting the estimates (crosses), numerical summation CR (contour) and the anisotropy parameters (circle, bottom border of (c) and (d)) used for the calculation of the CR. Confidence interval for isotropy is $(\hat{R}_-, \hat{R}_+) = (0.975, 1.025)$.

41
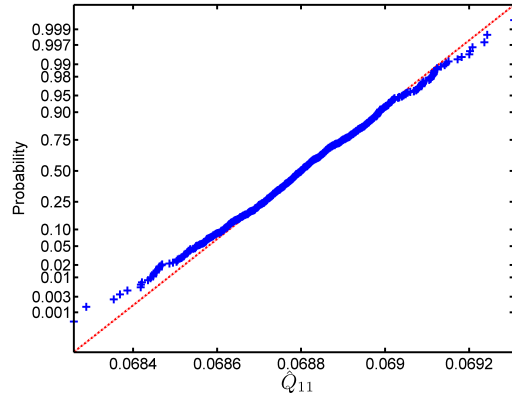
(a) $\hat{Q}_{11}$
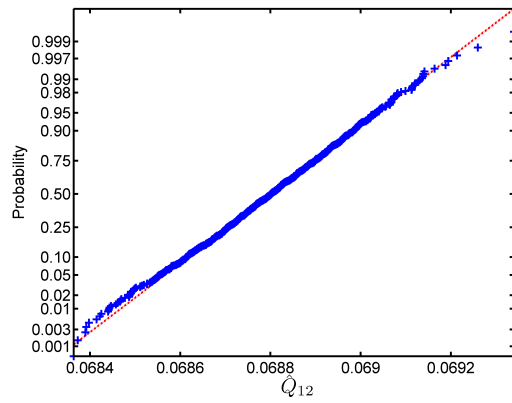
(b) $\hat{Q}_{12}$
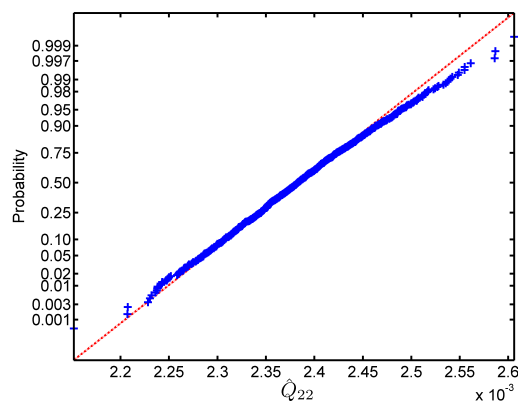
(c) $\hat{Q}_{22}$

Figure 4.3: Normal probability plots for the slope tensor estimates (Matérn case) justify the use of CLT in Lemma 2.2.2.

(a) Scatter plot and CR (Gaussian covariance).

(b) Random field (Gaussian covariance).

(c) Scatter plot and CR (Matérn covariance).

(d) Random field (Matérn covariance).
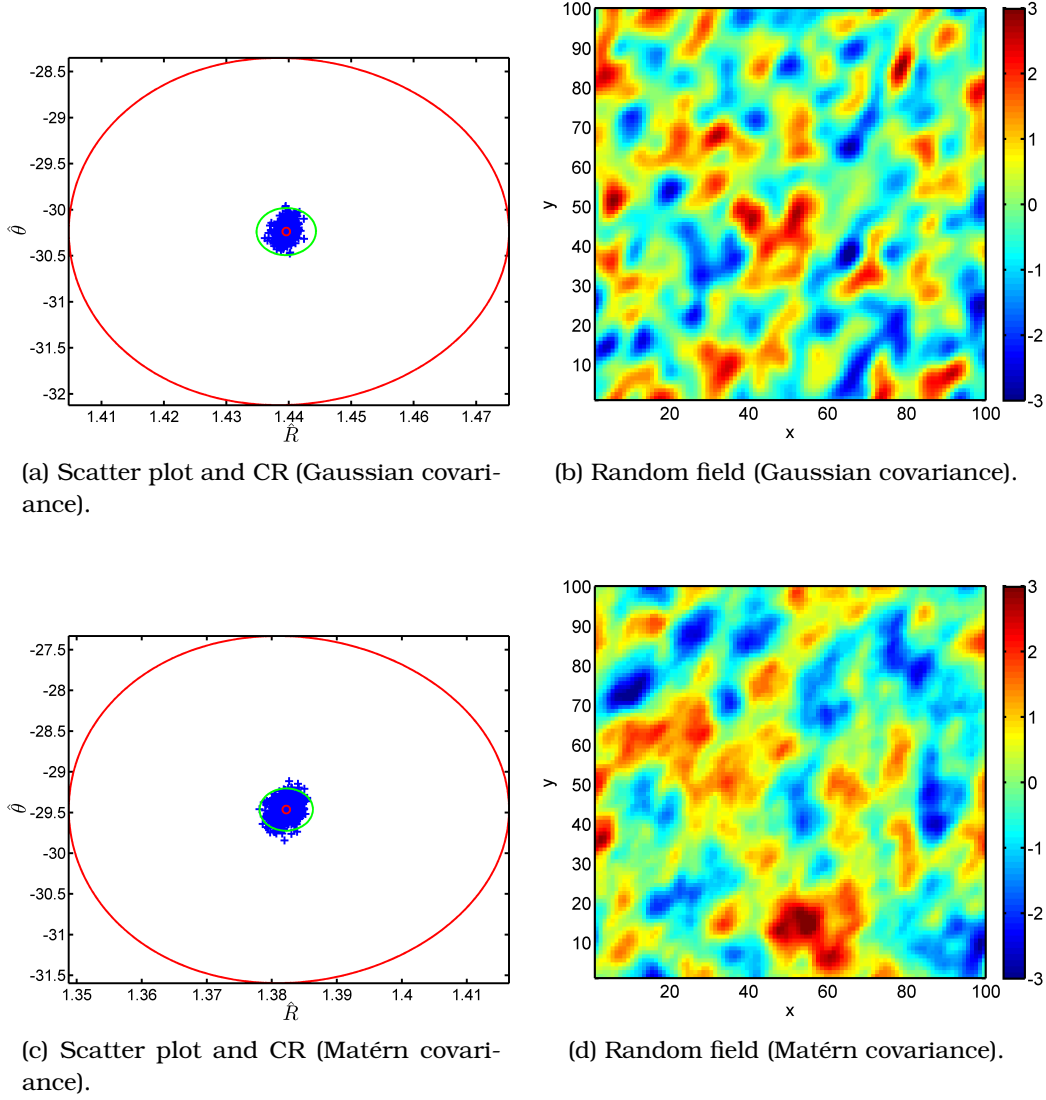
Figure 4.4: (a) and (c): Scatter plot of anisotropy estimates and non-parametric confidence region (outer contour) for 1000 realizations of anisotropic random field with Gaussian ($\tilde{\xi} = 4$) and Matérn ($\tilde{\xi} = 2.51$, $\nu = 2$) correlations, both having $R = 1.5$, $\vartheta = -30°$. A single realization for each field is shown in (b) and (d). Tighter confidence region (inner contour) is numerically calculated and incorporates the covariance-dependent terms in (2.21).

43

4.2   Simulated Scattered Data

We simulate scattered data using the following method: First, a realization of an GSRF is generated on a regular grid. Then, we randomly choose a fraction of the grid points to mimic scattered data of a real process. Next, we generate several random subsamples from the scattered data set and perform anisotropy estimation for each subsample. The subsamples should respect the conditions specified in the opening paragraph of Section **??**. The procedure is depicted in Figure 4.6. Figure 4.6a demonstrates an GSRF realization on a 1000×1000 grid, generated by means of the Fourier Filtering Method. The zero-mean, unit-variance isotropic GSRF has a Matérn covariance with $\nu = 2$ and $\xi = 15$. A randomly extracted set of scattered 2000 points is shown in Figure 4.6b.

The subsampling procedure should not be confused with the *spatial nonparametric bootstrap* procedure, which involves generating new samples, called bootstrap samples, by resampling from the actual data, and computing estimates for these new samples. Then, the distribution of these bootstrap estimates serves as a proxy for the actual distribution of the data estimates, so that statistical inference, such as the construction of confidence intervals, can be performed Loh [2008]. As a simpler alternative to spatial bootstrap, we decided to subsample the original data sets into smaller subsets in such a manner to artificially increase the available realizations of the same random field describing the actual natural process. The percentage of original data points used should be large enough to retain the spatial properties of the original dataset and keep the interpolation/estimation errors to a minimum. On the other hand, the percentage should be kept as small as possible, in order to be able to obtain more subsamples and provide a clear picture of the behavior of the anisotropy estimator on the $(\hat{R}, \hat{\partial})$-plane. The aim is to confirm that the behavior is captured by the analytical result obtained in Section 3.1. It was found experimentally that keeping 50% of the original datasets regarding the scattered data scenarios was a good compromise. For the interpolation, we employed the natural neighbor interpolation method Fisher *et al.* [2005] as implemented in MATLAB®. The choice was made for computation speed and for the interpolation surface smoothness properties discussed below. Details of the impact of different interpolators on anisotropy estimation can be found in Chorti & Hristopulos [2008].

The natural[3] neighbor interpolation method is a deterministic spatial

---

[3]Not to be confused with the *nearest* neighbor interpolation; a well-known method of spatial interpolation. Nearest neighbor interpolation does not produce smooth interpolation surfaces, thus it is inappropriate for the CHI method.
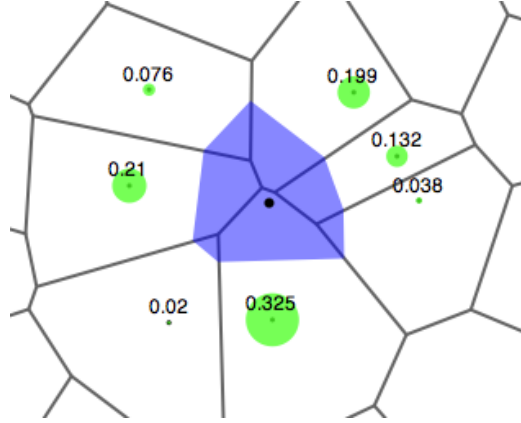
Figure 4.5: Natural neighbor interpolation. The colored circles, which represent the interpolating weights, w_i, are generated using the ratio of the shaded area to that of the cell area of the surrounding points. The shaded area is due to the insertion of the point to be interpolated into the Voronoi tessellation. From Wikipedia contributors [2012b].

interpolation method which is based on the Voronoi tessellation of a discrete set of spatial points Sibson [1981]. The interpolant fits the data at the nodes exactly, is local, and guarantees continuity in first and second derivatives, except at the nodes Boissonnat & Cazals [2000]; Sambridge *et al.* [1995]. The discontinuities at the nodes is not an issue in practice, since the interpolation grid is unlikely to include a node. While these properties hold in any dimension, so far efficient procedures for performing natural neighbor interpolation exist only in 2D. Being a deterministic interpolation method, an error estimate is not available. Also, since it provides smooth interpolation surfaces, it is suitable for random field gradient estimates required by the CHI method. It has been shown that the method behaves consistently for different sample distributions, and since it does not require user-defined parameters, it is appropriate for automatic mapping applications Fisher *et al.* [2005]; Ledoux & Gold [2008].

The natural neighbor interpolation concerns a weighted summation where the selection of neighbors and their corresponding weight calculation is based on Voronoi tessellation. For each $\mathbf{s} \in \mathcal{D}$, the estimate (interpolant) is calculated as the weighted sum over the set $\mathcal{N}(\mathbf{s})$ of neighbors of $\mathbf{s}$

$$\hat{X}(\mathbf{s}) = \sum_{\mathbf{s}_i \in \mathcal{N}(\mathbf{s})} w_i \, X(\mathbf{s}_i),$$

where $\hat{X}(\mathbf{s})$ is the interpolant at $\mathbf{s}$, w_i are the weights and $X(\mathbf{s}_i)$ are the

known data values at $\mathbf{s}_i \in \mathcal{D}$. The neighbors $\mathbf{s}_i$ used in the estimation are selected using the adjacency relationships of the Voronoi diagram. For each $\mathbf{s}$ there exist a set of neighbors $\mathcal{N}(\mathbf{s})$ that both surround and are close to $\mathbf{s}$. The weights are calculated by finding how much of each of the surrounding areas is 'stolen' when inserting $\mathbf{s}$ into the tessellation. The weight calculation is schematically depicted in Figure 4.5.

Figure 4.6c demonstrates a specific subsample of the scattered data set, containing 1000 points. The depicted field corresponds to the interpolation of the 1000 points used to estimate $(\hat{R}^*, \hat{\vartheta}^*)$. Figure 4.6d demonstrates the non-parametric confidence region (contour lines) along with the scatter cloud of anisotropy estimates, generated by randomly selecting 1000 subsamples of 1000 points from the scattered data set. The absence of estimates near the line $\hat{R} = 1$ is apparent and is expected since the JPDF should vanish at $\hat{R} = 1$; viz. Fig 3.4 and the accompanied discussion in Section 3.2. Deviations from the non-parametric CR are due to the sparseness of the subsamples and edge effects, which both introduces interpolation artifacts. Such artifacts can be seen in Fig. 4.6c at the right side ($x \simeq 0, y \simeq 450$–650) and bottom side ($x \simeq 500$–$700, y \simeq 0$) and since they are elongated and oriented either horizontally or vertically, they introduce anisotropy estimates with either high or low anisotropy ratios and anisotropy angles near zero. This explains the few anisotropy estimates in the scatter plot of Fig. 4.6d which lie around $\hat{\vartheta} = 0°$ and does not appear to follow the calculated non-parametric CR. The confidence interval for isotropy is $(\hat{R}_-, \hat{R}_+) = (0.925, 1.081)$ as derived from (3.5) by inserting $N = 1000$. With the exception of the few anisotropy estimates in the right side of Fig. 4.6d, the anisotropy estimates cloud is contained within the calculated isotropy confidence interval.

(a) Realization of isotropic Matérn SRF over a $10^3 \times 10^3$ square grid.

(b) Scattered data locations (2000 random points) and interpolated field (for visualization purposes).

(c) Subsample realization: 1000 random locations (from those in Fig. 4.6b) and interpolated field (used for anisotropy estimation).

(d) Anisotropy estimates via subsampling and non-parametric confidence region for $N = 1000$.
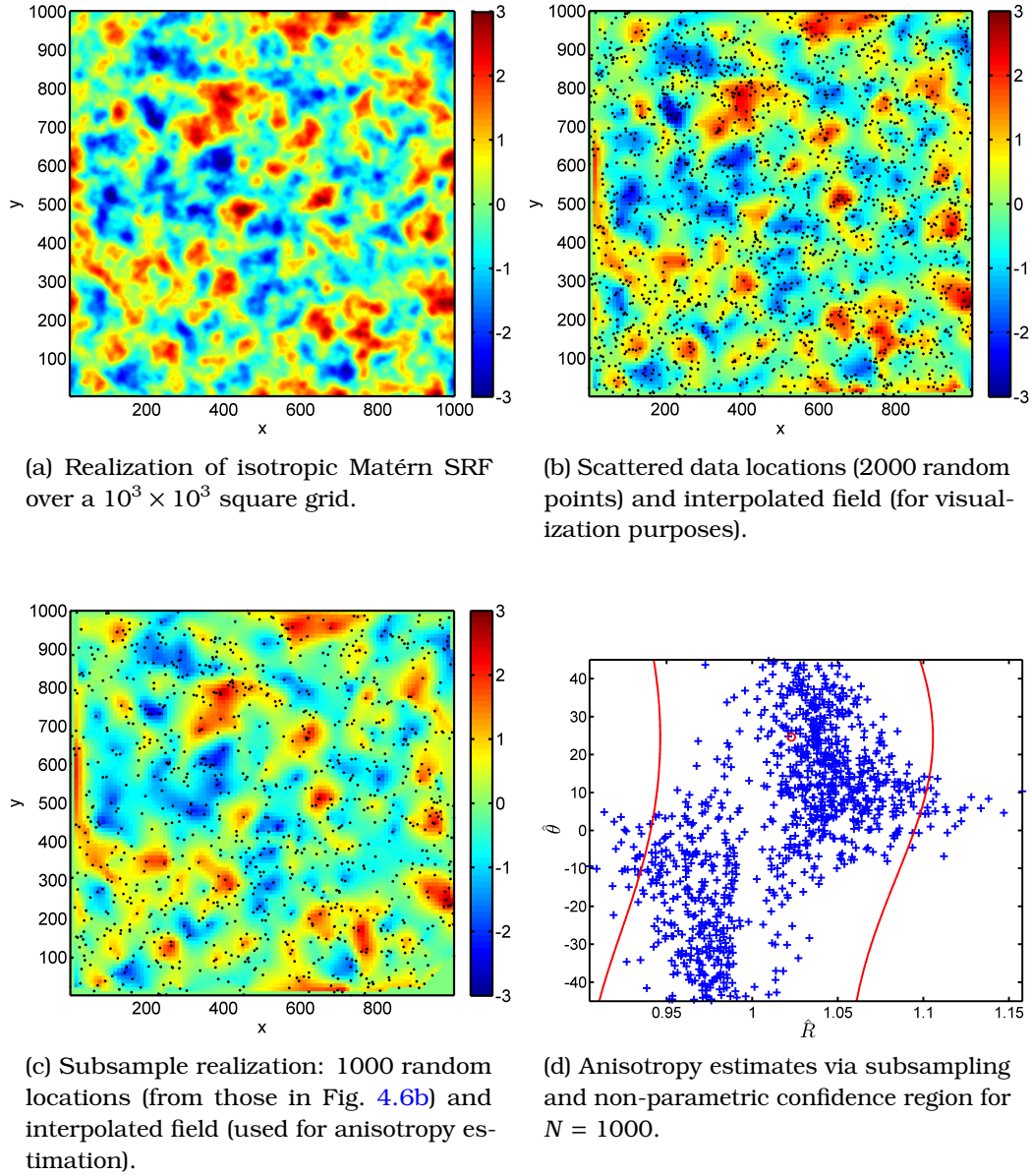
Figure 4.6: Non-parametric confidence region estimation for scattered data. (a) Realization of isotropic Matérn SRF with $\xi = 15$, $\nu = 2$ on a $10^3 \times 10^3$ square grid. (b) Locations of random sample (2000 points) and interpolated field. (c) Specific subsample realization and interpolated field used for anisotropy estimation. (d) Anisotropy estimates generated from 1000 random subsamples of 1000 points; curve corresponds to 95% non-parametric confidence region calculated with $N = 1000$ and anisotropy parameters estimated from the mean slope tensor elements. Confidence interval for isotropy is $(\hat{R}_-, \hat{R}_+) = (0.925, 1.081)$.

Table 4.1: Summary statistics of radioactivity dose rate exhaustive data sets (nSv/h) and CHI anisotropy estimates. Abbreviations: min: minimum sample value; max: maximum sample value; std: sample standard deviation; skew: sample skewness coefficient[4]; kurt: sample excess kurtosis coefficient.

$N = 1008$	min	mean	med	max	std	skew	kurt	$\hat{R}^*$	$\hat{\partial}^*$
Normal	57.0	97.7	98.6	180.0	19.6	0.4	0.6	1.18	7.36
Emergency	57.0	106.1	98.9	1528.2	92.5	11.3	144.1	0.45	-0.75

4.3 Environmental Emergency Scenario

Application of the method described in Section 4.2 to environmental emergency scenario follows. The data represent daily means of radioactivity gamma dose rates over part of the Federal Republic of Germany, and they were provided by the German automatic radioactivity monitoring network Dubois & Galmarini [2005, 2006]. The exhaustive dataset contains gamma dose rate measurements from 1008 fixed monitoring locations. The rates are measured in nanosieverts per hour (nSv/h). The *normal* data set corresponds to typical background radioactivity measurements ($\approx$ 100 nSv/h). The *emergency* data includes a simulated local release of radioactivity from the south-west corner of the monitored area that results in five stations reporting dose rates around 10 times above the background (exceeding 1000 nSv/h). A measurement station in the German network triggers an alarm if the dose rate measured exceeds 300 nSv/h. Natural background radiation in Germany is 40–240 nSv/h. With a dose rate exceeding 100 μSv/h it is necessary to take protective measures (e.g., to shelter indoors) Cornford [2006].

Figure 4.7 demonstrates the exhaustive normal and emergency datasets, using the natural neighbor interpolation method, while Table 4.1 summarizes their statistics. The two rightmost columns of Table 4.1 show the anisotropy parameters estimated with the CHI method. The normal set data follows the Gaussian distribution (graph not shown here), and thus has skewness and excess kurtosis coefficients close to zero. Data normalization was initially performed and had negligible impact in the quality of the anisotropy estimation. For that reason we performed our analysis of the SIC2004 data on the original data.

---

[4] The *skewness* of a random variable Z is defined as $\text{Skew}(Z) = \mathbb{E}\left[(Z - \mu_Z)\right]^3 / \sigma_Z^3$. The *excess kurtosis* as $\text{Kurt}(Z) = \mathbb{E}\left[(Z - \mu_Z)\right]^4 / \sigma_Z^4 - 3$. A symmetric distribution has skewness equal to zero. A distribution with negative skew has a longer left tail while a positive-
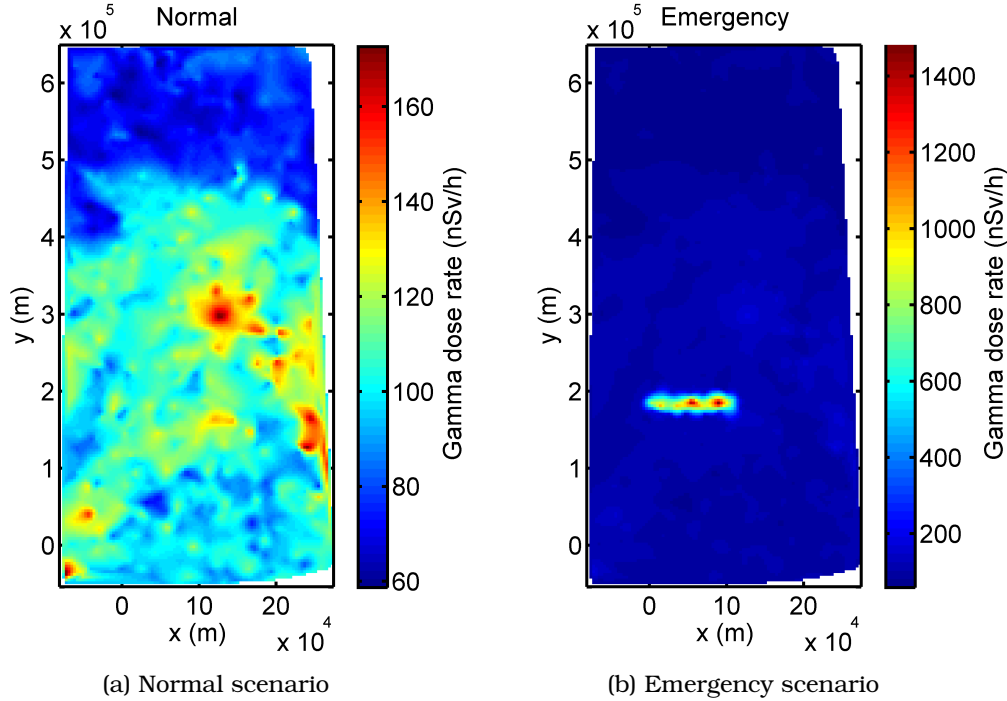
(a) Normal scenario   (b) Emergency scenario

Figure 4.7: Visualization of the SIC2004 exhaustive datasets of mean daily gamma dose rates using the natural neighbor interpolation method obtained from the 1008 measurement locations.

We calculated the JPDF and the 95% confidence regions of the anisotropy statistics based on the anisotropy parameters estimated by CHI, i.e., performing natural neighbor interpolation and anisotropy estimation using CHI in both exhaustive data sets and then plotting the non-parametric confidence regions and JPDFs. The results are shown in Fig. 4.8. There is no overlap of the two density functions, and the contours corresponding to the 95% confidence regions do not intersect, which suggests a statistically significant difference of the anisotropy parameters between the background and the emergency data. Since the computation of the anisotropy estimates is very fast, our method provides a straightforward indicator for a significant physical change in a system.

So far we examined the case where an automatic early warning system is expected to distinguish between normal (background) and abnormal

---

skew distribution has a longer right tail. The Gaussian distribution has *excess* kurtosis equal to zero. Distributions with negative excess kurtosis are said to be *platykurtic* while distributions with positive excess kurtosis are said to be *leptokurtic*.
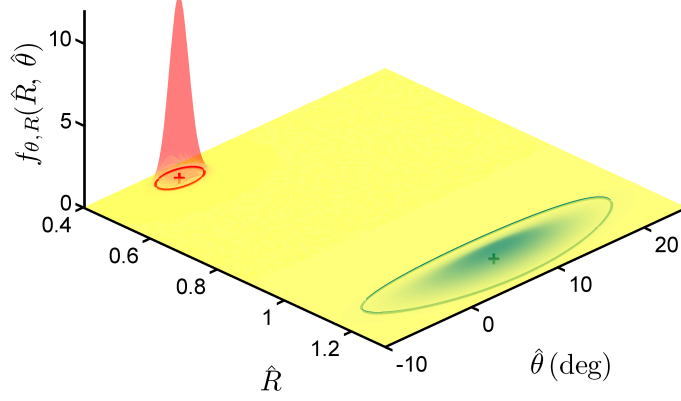
49

Figure 4.8: Joint probability density functions (surfaces) and confidence 95% regions (solid contours) for the exhaustive radioactivity dose rate data sets: normal data set (bottom right) and emergency data set (far left).

(emergency) situations. In the following we will investigate a case in which different measurements are classified as normal and should be classified as such. As a part of the SIC2004 interpolation exercise, ten *additional* data sets of mean daily values measured at 200 monitoring stations were provided for algorithm training purposes. Each set was randomly drawn from a different month of the year 2003. No specific date details are publicly available, according to the rules of the contest Pebesma *et al.* [2012]. The ten datasets (numbered 1–10) involve normal (background) gamma dose rates. In Figure 4.9 a natural neighbor interpolation of the first of the additional normal datasets is shown. Note that each of the additional datasets has only 200 data points, in contrast to the exhaustive datasets 4.7, each having 1008 data points. We calculated the anisotropy parameters for each of the 10 datasets. Figure 4.10 demonstrates the results along with the confidence region calculated using the anisotropy parameters estimated from the mean slope tensor elements as in 4.2. All estimates lie within the 95% CR thus there is no statistically significant change in the anisotropy parameters of the ten measurements according to the statistical properties of the anisotropy estimator we examine here. On the other hand, note that the anisotropy estimates differ from the normal case of the exhaustive dataset plotted in Fig. 4.8 because the number of data points is $\approx 1/5$ of the exhaustive data sets, affecting the quality of the interpolation used for anisotropy estimation, as it can be testified by comparing Figures 4.7a and 4.9.
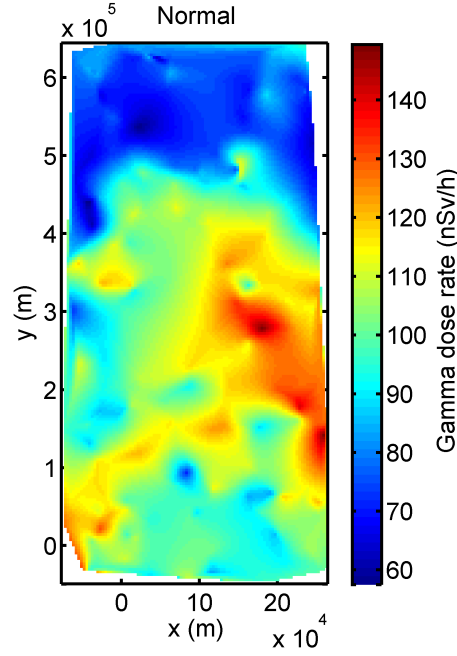
Figure 4.9: Visualization of the first of the additional ten normal SIC2004 datasets of mean daily gamma dose rates containing $N = 200$ points. The natural neighbor interpolation method was used.
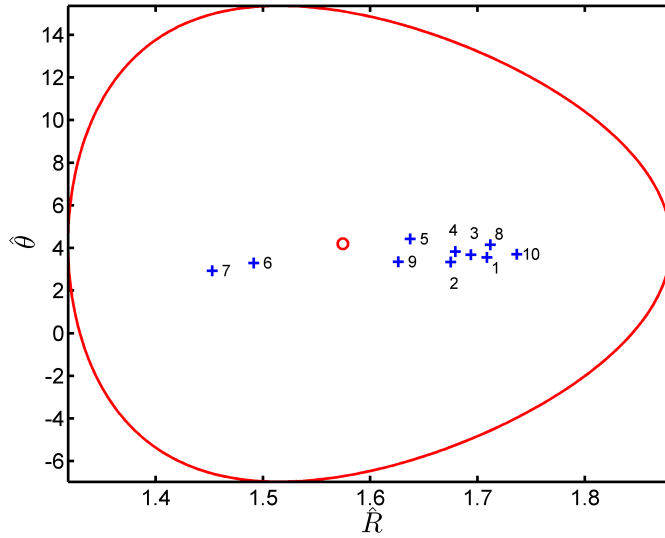

Figure 4.10: Anisotropy estimates (crosses) for the 10 background dose rate sets (number correspond to each dataset) and 95% confidence region calculated using the mean slope tensor elements. Each dataset has $N = 200$ points. All estimates lie within the CR thus there is no statistically significant change in anisotropy.

51

Table 4.2: Summary statistics of the Amynteon and Mavropigi lignite thickness data sets (units: meters) and CHI anisotropy estimates.

Mine	N	min	mean	med	max	std	skew	kurt	$\hat{R}^*$	$\hat{\partial}^*$
Amynteon	644	0.49	14.12	13.40	54.86	8.69	0.82	1.24	1.16	-33.5
Mavropigi	416	0.36	17.38	13.83	73.62	13.21	1.03	0.94	0.83	-21.6

4.4 Geologic Reserves Data

In this section, we investigate the anisotropy statistics of geologic reserve estimates from drill measurements obtained from two lignite mines in Greece. Also we will test the performance of the non-parametric CR using the subsampling method introduced in subsection 4.2. Additionally, the sample histograms and kernel-based estimates of the anisotropy parameters PDFs will be compared to the theoretical PDFs obtained via numerical integration of the corresponding non-parametric JPDF.

The measurements are obtained from exploratory drill holes made in two lignite mines in Greece, Amynteon and Mavropigi and are kindly provided by the Public Power Corporation (PPC) of Greece S.A. Both mines are located in Ptolemaida, a town in the prefecture of Kozani, Greece and operated by the Greek PPC. Figure 4.11 demonstrates a satellite map of the region where the mines and the thermoelectric power plants are located. A schematic of the area is shown in Figure 4.12 (in Greek, courtesy of the Greek PPC). While both datasets contain measurements of several variables regarding lignite quality, we focus on lignite thickness measurements aggregated per drill hole.

Table 4.2 summarizes the statistics of both mine datasets. Both datasets are non-Gaussian and asymmetric, as confirmed by the sample kurtosis and skewness coefficients. A natural neighbor interpolation of the lignite thickness measurements and the drill hole locations are depicted in Figures 4.13a and 4.15a. The drill hole locations are also shown. The displayed ellipse has orientation and semimajor axes ratio according to the anisotropy parameter estimates. The plotted lengths of the ellipse semimajor axes *does not* correspond to the correlation lengths for each principal direction. Coordinates shifted to zero mean (for each direction) and divided by a factor of 1000, whereas *z*-values (lignite thickness, in meters) left unscaled. In Figures 4.13b and 4.15b the JPDF corresponding to the estimated anisotropy parameters for each mine is also plotted.

The anisotropy estimates obtained via the subsampling procedure are depicted in Figures 4.14 and 4.16 along with the corresponding non-

parametric CR curves, calculated using the mean slope tensor elements as in Section 4.2. Furthermore, the univariate PDFs of $\hat{R}$ and $\hat{\partial}$ derived via numerical integration of the theoretical JPDF are compared to the sample PDFs estimated from the subsampling anisotropy estimates. The sample histograms, probability density estimates and theoretical PDFs are depicted in Figures 4.14b and 4.14c for the Amynteon dataset and in Figures 4.16b and 4.16c for the Mavropigi dataset. The probability density estimates are calculated[5] performing kernel density estimation Bowman & Azzilini [1997]; Parzen [1962] on the subsamples anisotropy parameter estimates using the *Epanechnikov* kernel function Epanechnikov [1969]. The kernel window size is $0.4°$ for the kernel estimates of $f_{\partial}(\hat{\partial})$ and to 0.04 for the kernel estimates of $f_R(\hat{R})$. There is good agreement between the sample PDFs and those obtained from integrating the non-parametric JPDF.

For the Amynteon dataset, consisting of $N = 644$ data points, the isotropy test (3.5) provides $(\hat{R}_-, \hat{R}_+) = (0.91, 1.10)$, while the estimated anisotropy ratio is $\hat{R} = 1.16$. For the Mavropigi dataset, with $N = 416$, the isotropy test provides $(\hat{R}_-, \hat{R}_+) = (0.89, 1.13)$, while the estimated anisotropy ratio is $\hat{R} = 0.83$. Both anisotropy ratio estimates lie outside the corresponding isotropy confidence region thus both datasets can *not* be considered isotropic at 95% confidence level, according to the statistical test (3.5). The isotropy hypothesis is rejected in both cases, thus anisotropic models may be more appropriate for modeling the mining datasets examined here.

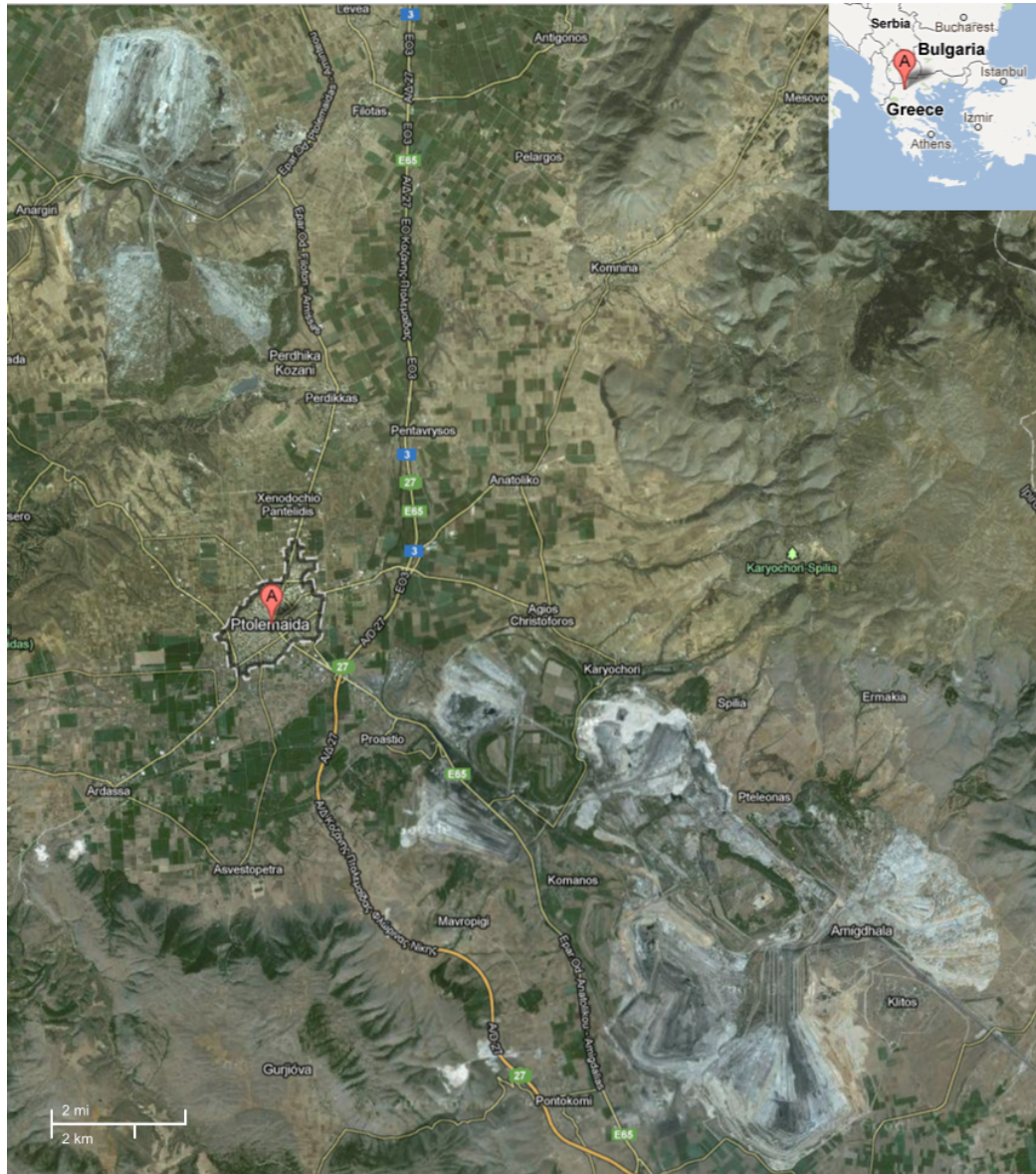[5]The MATLAB$^{®}$ function `ksdensity` was used.

Figure 4.11: Lignite mines in the area near the city of Ptolemaida (red marker) located at Kozani, Greece. Source images retrieved from *Google Maps*: http://goo.gl/maps/Sp80d
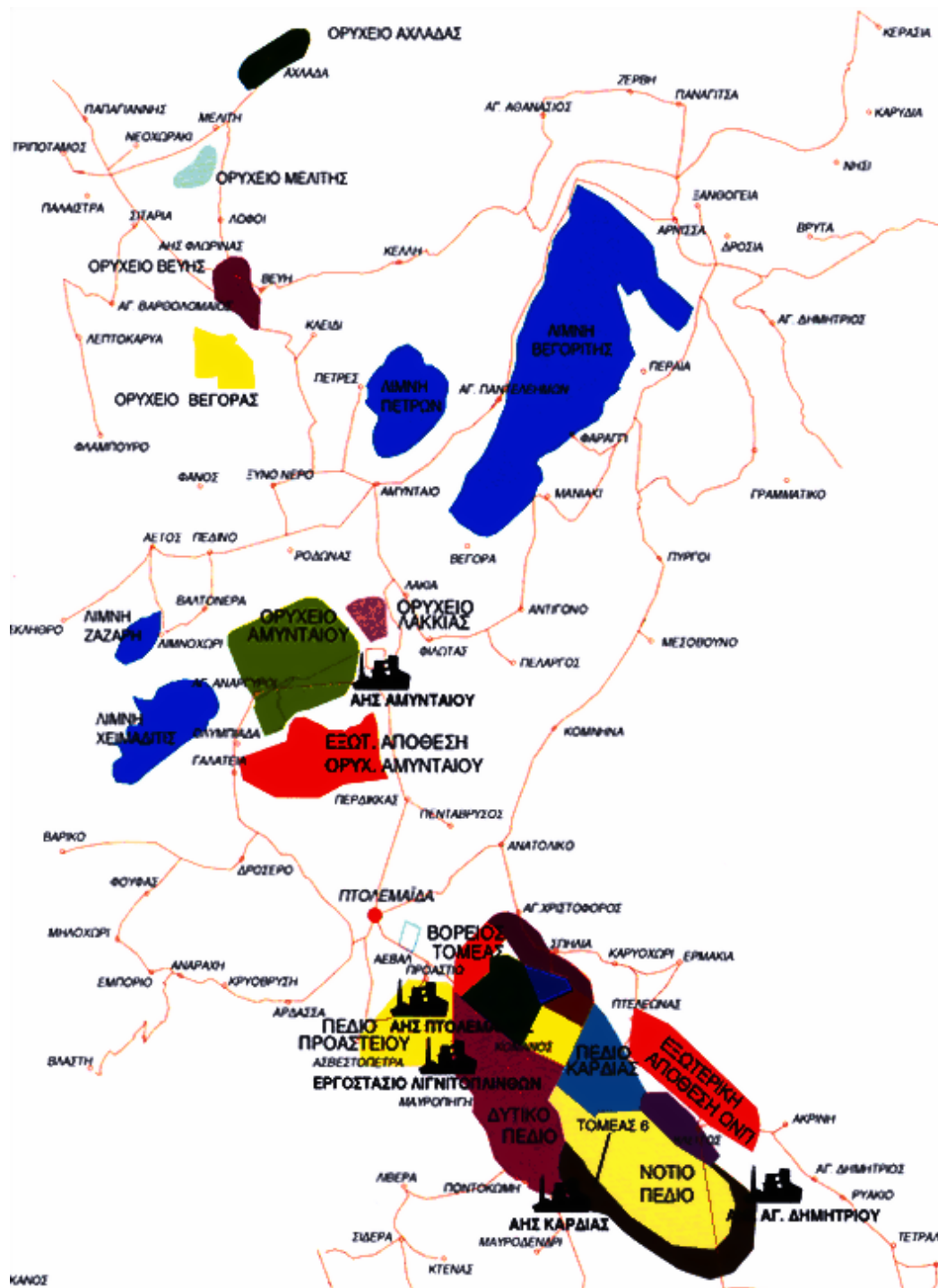
Figure 4.12: Lignite mines and thermoelectric power plants near the city of Ptolemaida, in the Kozani prefecture, Greece. Chart retrieved from the Greek Public Power Corporation website.
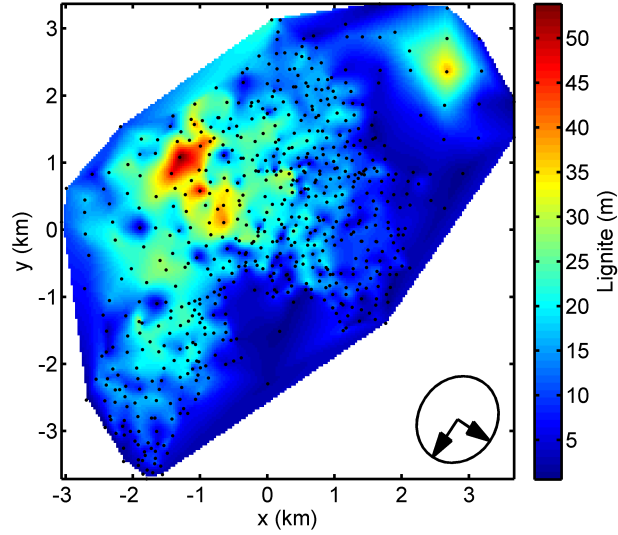
(a) Amynteon mine
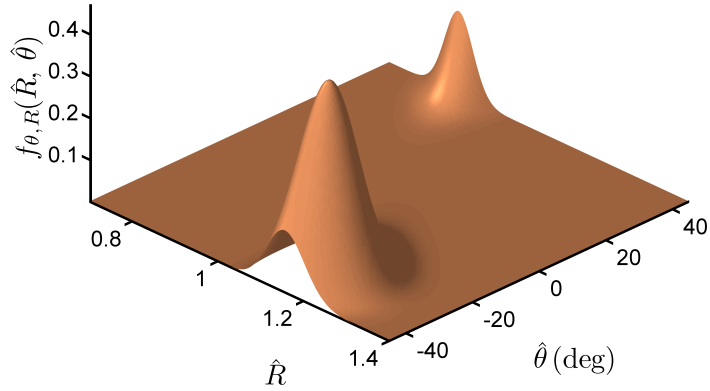
(b) JPDF

Figure 4.13: (a) Visualization of the Amynteon lignite data using the natural neighbor interpolation. The black dots are the drill hole locations. The ellipse axes length ratio and orientation represent the anisotropy estimates; the axes lengths *does not* correspond to correlation lengths. (b) The nonparametric JPDF calculated for the anisotropy estimates $(\hat{R}, \hat{\partial}) = (1.16, -33.5°)$ and $N = 644$.
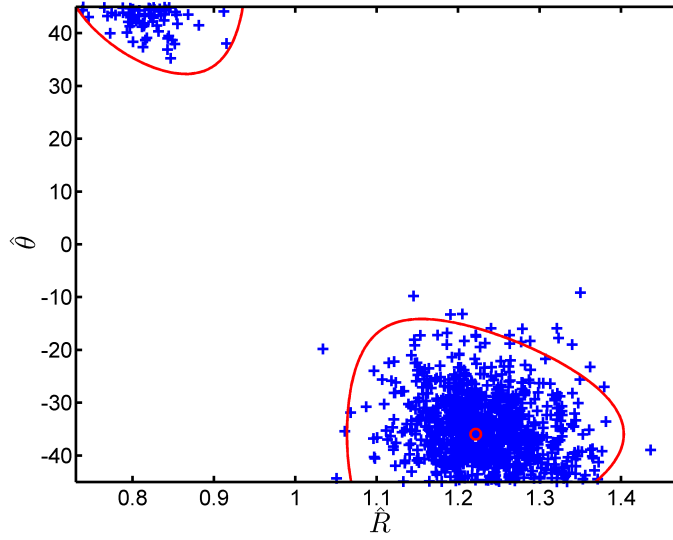
(a) Anisotropy estimates via subsampling

(b) $f_R(\hat{R})$

(c) $f_{\hat{\partial}}(\hat{\partial})$

Figure 4.14: Testing the validity of the non-parametric confidence region estimation via subsampling, for the Amynteon dataset. (a) Anisotropy estimates (crosses) using 1000 random subsamples each containing 50% (322) of the original data points. The solid line depicts the theoretical 95% confidence region calculated in the same manner as in 4.1. (b) and (c): Comparison of the univariate theoretical PDFs derived using numerical integration of the analytical JPDF, to the sample histograms and the kernel density estimations of the PDFs, calculated from the anisotropy estimates of (a).

57

(a) Mavropigi mine

(b) JPDF

Figure 4.15: (a) Visualization of the Mavropigi lignite data using the natural neighbor interpolation. The black dots are the drill hole locations. The ellipse axes length ratio and orientation represent the anisotropy estimates; the axes lengths *does not* correspond to correlation lengths. (b) The nonparametric JPDF calculated for the anisotropy estimates $(\hat{R}, \hat{\partial}) = (0.83, -21.6°)$ and $N = 416$.
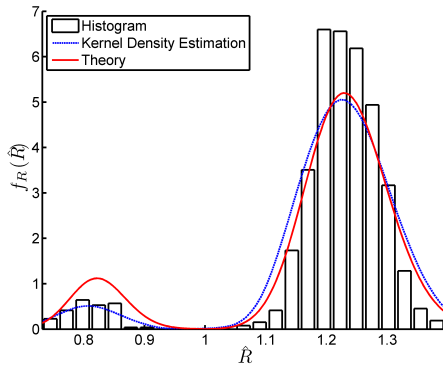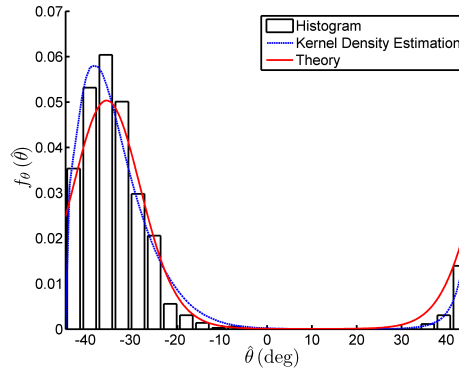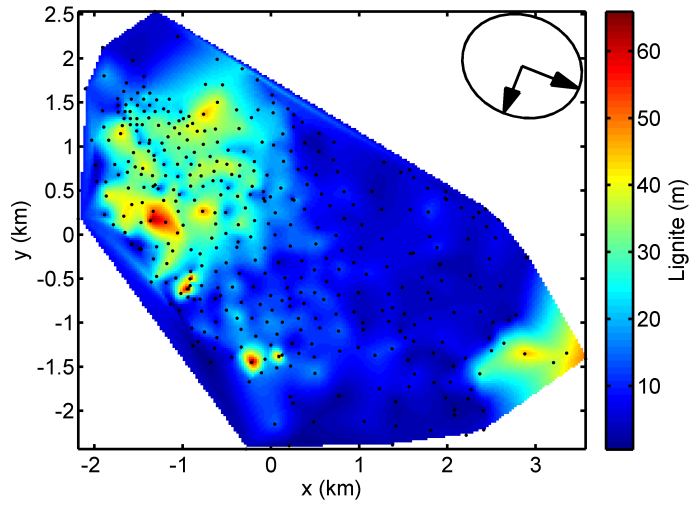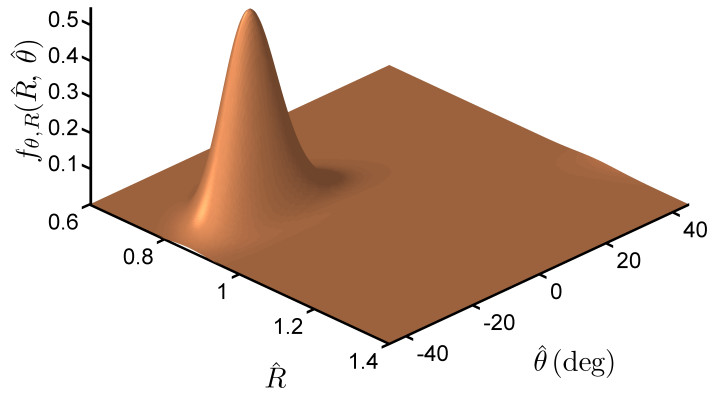
58
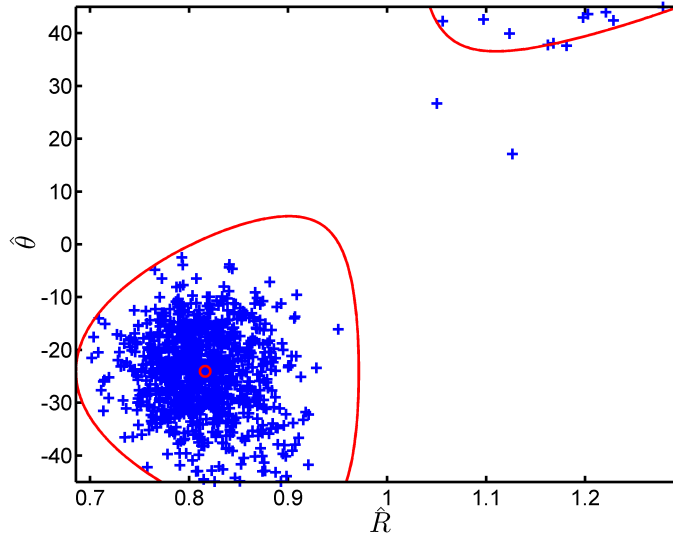
(a) Anisotropy estimates via subsampling
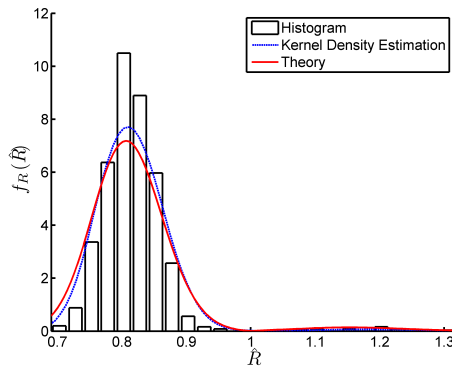
(b) $f_R(\hat{R})$
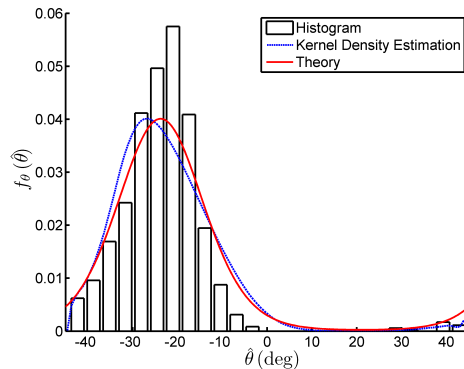
(c) $f_{\hat{\theta}}(\hat{\theta})$

Figure 4.16: Testing the validity of the non-parametric confidence region estimation via subsampling, for the Mavropigi dataset. (a) Anisotropy estimates (crosses) using 1000 random subsamples each containing 50% (208) of the original data points. The solid line depicts the theoretical 95% confidence region calculated in the same manner as in 4.1. (b) and (c): Comparison of the univariate theoretical PDFs derived using numerical integration of the analytical JPDF, to the sample histograms and the kernel density estimations of the PDFs, calculated from the anisotropy estimates of (a).

— RANDALL MUNROE, XKCD

Chapter 5
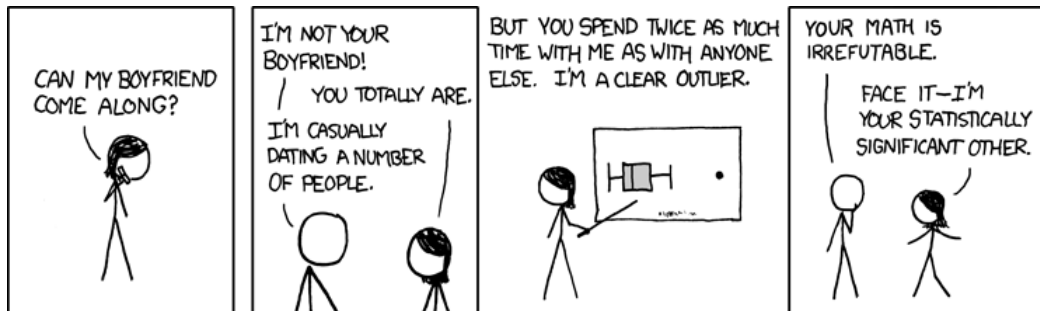
Discussion and Conclusions

5.1 Discussion and Conclusions

THIS work focuses on the estimation of anisotropy by means of the Covariance Hessian Identity (CHI) in two-dimensional digital data that are either scattered or supported on a grid. We derive explicit expressions for the joint PDF of anisotropy statistics, equation (2.30), and for the corresponding anisotropy confidence regions at any confidence level, equation (2.31). The main assumptions used are (i) that the data are drawn from a jointly Gaussian, stationary and differentiable random field (ii) that the covariance function is short-ranged We also derive a non-parametric approximation for the joint PDF of the anisotropy statistics, which can be used if the covariance function is unknown *a priori*, or to avoid numerical calculations required for estimating the covariance. The non-parametric approximation is given by (3.2). The corresponding equation for the non-parametric approximation of the confidence region is given by (3.4).

Practical application of the results of this research requires the estimation of anisotropy statistics using CHI (or other methods). Accurate estimation based on CHI requires a large sample size, $N \gg 1$, and a sample domain that is large with respect to the correlation area. The latter may be difficult to satisfy for data with large anisotropy ($R \gg 1$ or $R \ll 1$). In such cases, the CHI estimate tends to underestimate the actual anisotropy.

We illustrate the application of the joint PDF and the confidence regions with simulated and real data. The results of this research can be used to identify significant deviations in anisotropy between data sets, e.g., due to structural differences or major changes in the underlying physical process. The computational cost is minimal, since the corresponding expressions are analytical. The major contribution to the computational cost is the

estimation of partial derivatives of the random field. The user can freely choose a more accurate field derivatives estimator, at the cost of additional computational time.

In the realistic scenarios examined in this work we demonstrated that the anisotropy statistics follow the proposed CR, by utilizing a subsampling method to artificially generate realizations of the same underlying physical process. In practical situations, the anisotropy estimates accompanied by the analytical results regarding their statistics can be helpful for model selection. For example, the anisotropy test can serve as an aid to choose between a simpler isotropic model, and a more complex (since two more parameters are required) anisotropic model. On the other hand, statistical significance does not necessarily imply interpolation performance. While under the proposed isotropy test a small dataset is considered isotropic, this should not prohibit the expert from choosing an anisotropic model if the additional complexity leads to interpolation performance. The non-parametric JPDF which is available in closed form, can be used as Bayesian prior if the modeling is going to be performed by following the Bayesian paradigm.

5.2   Publications and Presentations

During the study of the statistical properties of the CHI estimator, two presentations in international conferences were made Hristopulos *et al.* [2008, 2009]. The resulting isotropy test was published (without the proof details) in Spiliopoulos *et al.* [2011]. The theoretical framework and most of the application scenarios that are presented in this work, was submitted at the Cornell University Library electronic archive and distribution server for research articles (arXiv.org) and is publicly available Petrakis & Hristopulos [2012].

5.3   Future Work

Straightforward extension of this work is possible for jointly lognormal data along the lines of Chorti & Hristopulos [2008]. Application on non-differentiable random fields is possible if an angle-preserving smoothing kernel is first applied to the data. The same ideas, albeit with increased mathematical complication, can be applied in three dimensions as well.
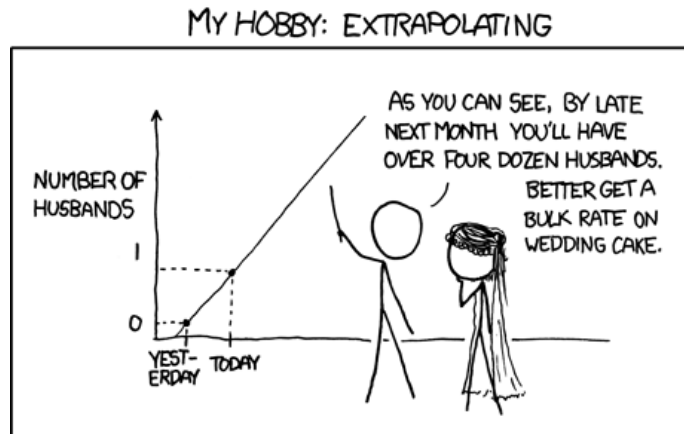
— Randall Munroe, xkcd

Appendices

Appendix A

Mathematical Details

The mathematical details are included in this Appendix.

A.1 Jacobi's Theorems

We use Jacobi's theorems e.g., Papoulis & Pillai [2002]; Soong [2004] to determine the PDF under the change-of-variables transformations $\vec{\mathbb{Q}} \rightarrow \hat{\mathbf{q}} \rightarrow (\hat{R}, \hat{\partial})$.

Theorem A.1.1 (Jacobi's univariate theorem)**.** *Let Z be a continuous random variable and $Y = g(Z)$, where $g(Z)$ is a continuous function in Z. If $y = g(z)$ admits at most a countable number of roots $z_j = g_j^{-1}(y)$, $j = 1, \ldots, r$, then*

$$f_Y = \sum_{j=1}^{r} f_Z(g_j^{-1}(y)) \left| \frac{dg_j^{-1}(y)}{dy} \right|.$$

In the case of a multivariate variable transformation, Jacobi's theorem becomes:

Theorem A.1.2 (Jacobi's multivariate theorem)**.** *Let $\mathbf{Z}$ and $\mathbf{Y}$ be two n-dimensional continuous random vectors with components $(Z_1, \ldots, Z_n)$ and $(Y_1, \ldots, Y_n)$, respectively. The transformation $\mathbf{Y} = \mathbf{g}(\mathbf{Z})$ represents the set of equations $y_l = g_l(\mathbf{Z})$, $l = 1, \ldots, n$. Assume that the functions g_l are continuous and possess continuous partial derivatives with respect to each of their arguments.*

1. *If the g_l define one-to-one mappings, unique inverse functions g_l^{-1} such that $\mathbf{Z} = \mathbf{g}^{-1}(\mathbf{Y})$ exist. Then, the transformation of the JPDF $f_{\mathbf{Z}}$ to $f_{\mathbf{Y}}$ is*

accomplished by means of

$$f_{\mathbf{Y}} = f_{\mathbf{Z}}\left(\mathbf{g}^{-1}(\mathbf{Y})\right)|\det(\mathbf{J})|,$$

where $\mathbf{J}$ *is the Jacobian of the transformation:*

$$\mathbf{J} = \frac{\partial(g_1^{-1}, \ldots, g_n^{-1})}{\partial(y_1, \ldots, y_n)}.$$

2. *If* $\mathbf{Y} = \mathbf{g}(\mathbf{Z})$ *admits at most a countable number of roots* $\mathbf{Z}_j = \mathbf{g}_j^{-1}(\mathbf{y})$, $j = 1, \ldots, r$, *then*

$$f_{\mathbf{Y}} = \sum_{j=1}^{r} f_{\mathbf{Z}}\left(\mathbf{g}_j^{-1}(\mathbf{y})\right)|\det(\mathbf{J}_j)|,$$

where $\mathbf{J}_j$ *is the Jacobian corresponding to the j-th root, defined by*

$$\mathbf{J}_j = \frac{\partial(g_{j,1}^{-1}, \ldots, g_{j,n}^{-1})}{\partial(y_1, \ldots, y_n)}.$$

Theorem A.1.2 also applies if $\dim(\mathbf{Y}) = m < n$. Then, the m-dimensional vector $\mathbf{Y}$ is augmented by an $(n - m)$-dimensional vector $\mathbf{Z}' = \mathbf{h}(\mathbf{Z})$ where $\mathbf{h}(\cdot)$ is a simple function with continuous partial derivatives. The $n - m$ dummy variables in $\mathbf{Z}'$ are eliminated from the JPDF of $\mathbf{Y}$ by integration.

A.2 Probability Density Function of Gradient Product Tensor

Below, we obtain the PDF of the gradient components $X_i = \partial_i X(\mathbf{s})$ for $i = 1, 2$. Let us define by $f_{X_i}(\partial_i X = z)$ the PDF of the gradient component $\partial_i X$, and by $f_{X_{ij}}(X_{ij} = y)$ the PDF of the gradient product $\partial_i X(\mathbf{s})\partial_j X(\mathbf{s})$. The following theorem holds Yaglom [1987]:

Theorem A.2.1 (PDF of field gradient). *For a Gaussian, differentiable and stationary SRF* $X(\mathbf{s})$, *the gradient component* $\partial_i X(\mathbf{s})$ *is a zero-mean Gaussian SRF with covariance function given by the following expression:*

$$\mathbb{E}\left[\partial_i X(\mathbf{s})\,\partial_j X(\mathbf{s} + \mathbf{r})\right] = -\frac{\partial^2 c_{\mathrm{xx}}(\mathbf{r})}{\partial r_i \partial r_j}. \tag{A.1}$$

In light of (2.13) and (A.1) the *variance* of $\partial_i X(\mathbf{s})$ is given by $\mathrm{Var}\left(\partial_i X(\mathbf{s})\right) =$

Q_{ii}. Hence, the univariate PDF f_{X_i} is given by

$$f_{X_i}(\partial_i X = z) = \frac{1}{\sqrt{2\pi Q_{ii}}}\, e^{-z^2/2Q_{ii}}. \tag{A.2}$$

A.2.1 Probability Density Function of Diagonal Elements of Gradient Product Tensor

Theorem A.2.2 (PDF of diagonal elements of gradient product tensor). *Let $X(\mathbf{s})$ be a Gaussian, statistically stationary SRF that admits (in the mean square sense) all partial derivatives of second-order (at least). Then, the PDF of X_{ii} is given by the chi square (χ_ν^2) distribution with one degree of freedom ($\nu = 1$):*

$$f_{X_{ii}}(X_{ii} = y) = \frac{\exp(-\frac{y}{2Q_{ii}})}{\sqrt{2\pi Q_{ii} y}}. \tag{A.3}$$

Proof. Let us define $Z = \partial_i X(\mathbf{s})$, ($i = 1, 2$) and $Y = X_{ii}$. The equation $y = g(z) = z^2$ admits two real roots for $y > 0$, i.e., $y_{1,2} = \pm\sqrt{z}$, but no real roots for $y < 0$. Since $dg^{-1}/dy = \pm 1/(2\sqrt{y})$, by applying Theorem A.1.1, we obtain:

$$f_{X_{ii}}(X_{ii} = y) = \frac{1}{2\sqrt{y}}\left[f_{X_i}(\sqrt{y}) + f_{X_i}(-\sqrt{y})\right], \ y > 0. \tag{A.4}$$

Equation (A.3) follows from the above and (A.2). The standard density of the χ_1^2 distribution is obtained from (A.3) via the replacement $y' = y/Q_{ii}$. ∎

The mean and variance of X_{ii} are thus obtained by

$$\mathbb{E}\left[X_{ii}\right] = Q_{ii}, \tag{A.5}$$

$$\mathrm{Var}\left(X_{ii}\right) = 2(Q_{ii})^2. \tag{A.6}$$

A.2.2 Probability Density Function of Non-diagonal Elements of Gradient Product Tensor

Theorem A.2.3 (PDF of non-diagonal elements of the gradient product tensor). *Let $X(\mathbf{s})$ be an SRF that satisfies the conditions of Theorem A.2.2. Then, the PDF of $X_{12} = \partial_1 X(\mathbf{s})\, \partial_2 X(\mathbf{s})$ is given by:*

$$f_{X_{12}}(X_{12} = y) = \frac{1}{\pi\sqrt{\det(\mathbf{Q})}}\exp\left(y\,\frac{Q_{12}}{\det(\mathbf{Q})}\right)K_0\left(|y|\,\frac{\sqrt{Q_{11}\,Q_{22}}}{\det(\mathbf{Q})}\right), \tag{A.7}$$

where $\mathbf{Q}$ is the slope tensor, $\det(\mathbf{Q})$ is its determinant, and K_0 is the modified Bessel function of the second kind and order zero. Since $Q_{12} = Q_{21}$, it holds that $f_{X_{21}}(X_{21} = y) = f_{X_{12}}(X_{12} = y)$.

Proof. We define the random vectors $\mathbf{Z}^t = (\partial_1 X(\mathbf{s}), \partial_2 X(\mathbf{s})) = (X_1, X_2)$, $\mathbf{Y}^t = (X_1 X_2, X_2)$. The Jacobian determinant for the transformation $\mathbf{Z} \to \mathbf{Y}$ and its absolute value are given by

$$\det(\mathbf{J}) = \begin{vmatrix} y_2^{-1} & -y_1 \, y_2^{-2} \\ 0 & 1 \end{vmatrix} = \frac{1}{y_2} \Rightarrow |\det(\mathbf{J})| = \frac{1}{|y_2|}.$$

Thus, according to Theorem A.1.2, $f_{X_{12}}$ is given by

$$f_{X_{12}}(X_{12} = y_1) = \int_{-\infty}^{\infty} dy_2 \, f_{zz}\left(y_1/y_2, y_2\right) \frac{1}{|y_2|}, \tag{A.8}$$

where f_{zz} is the bivariate Gaussian PDF of the two-component fluctuation (z_1, z_2), given by

$$f_{zz}(z_1, z_2) = \frac{1}{2\pi \, \det(\mathbf{Q})} \, e^{-\frac{1}{2} \mathbf{Z}^t \mathbf{Q}^{-1} \mathbf{Z}}, \tag{A.9}$$

Then, the integration in (A.8) can be evaluated using the integral (3.471.9) [Gradshteyn & Ryzhik, 2007, p. 368] and the transformation $y_2^2 \to x$, which lead to (A.7). ∎

The mean of X_{12} is given by (2.12). The variance is obtained using the normality of the $\partial_1 X(\mathbf{s})$ and $\partial_2 X(\mathbf{s})$ distributions by applying the Isserlis-Wick moment factorization theorem Isserlis [1918]; Wick [1950]:

$$\mathbb{E}\left[X_{12}\right] = Q_{12}, \tag{A.10}$$

$$\mathrm{Var}\left(X_{12}\right) = Q_{11} \, Q_{22} + Q_{12}^2. \tag{A.11}$$

For $|x| \to \infty$, $K_0(|x|) \approx \sqrt{\frac{\pi}{2|x|}} e^{-|x|}$ Abramowitz & Stegun [1970]. Hence, based on (A.7), $f_{X_{12}}$ decays for large $|X_{12}|$ as

$$f_{X_{12}} \approx \frac{1}{\sqrt{2\pi}(Q_{11} \, Q_{22})^{1/4}} \frac{1}{\sqrt{|X_{12}|}} \, \exp\left[\frac{-|X_{12}|}{\sqrt{Q_{11} \, Q_{22}} + \mathrm{sign}(X_{12}) \, Q_{12}}\right], \tag{A.12}$$

where $\mathrm{sign}(x) = 1, x > 0 \ \wedge \ \mathrm{sign}(x) = -1, x < 0$.

Asymptotic convergence of $f_{X_{12}}$ to zero for $|X_{12}| \to \infty$ requires that the denominator of the exponent be positive. Q_{11} and Q_{22} are positive by definition. Regardless of the sign of Q_{12}, the positive definiteness of $\mathbf{Q}$ (see Theorem 2.1.1), implies that $\sqrt{Q_{11} \, Q_{22}} > \pm Q_{12}$. Hence, the denominator is

indeed positive. The asymptotic dependence of $f_{X_{12}}$ matches that of the χ_1^2 distribution.

A.3 Proof of Lemma 2.2.1

Proof. Using the definition (2.18) we obtain:

$$C_{ij,kl} = \text{Cov}\left(\frac{1}{N}\sum_{n=1}^{N} X_{ij}(\mathbf{s}_n), \frac{1}{N}\sum_{m=1}^{N} X_{kl}(\mathbf{s}_m)\right) = \frac{1}{N^2}\sum_{n,m} \text{Cov}\left(X_{ij}(\mathbf{s}_n), X_{kl}(\mathbf{s}_m)\right).$$
(A.13)

Due to the translation invariance of $X(\mathbf{s})$, the double series in (A.13) is reduced to a single series over all (N^2) lag vectors $\mathbf{r}_{nm} = \mathbf{s}_n - \mathbf{s}_m$ ($n, m = 1, \ldots, N$), i.e.,

$$
\begin{aligned}
C_{ij,kl} &= \frac{1}{N^2}\sum_{\mathbf{r}_{nm}} \text{Cov}\left(X_{ij}(\mathbf{s}_0), X_{kl}(\mathbf{s}_0 + \mathbf{r}_{nm})\right) \\
&= \frac{1}{N}\text{Cov}\left(X_{ij}(\mathbf{s}_0), X_{kl}(\mathbf{s}_0)\right) + \frac{1}{N^2}\sum_{\mathbf{r}_{nm}\neq\mathbf{0}} \text{Cov}\left(X_{ij}(\mathbf{s}_0), X_{kl}(\mathbf{s}_0 + \mathbf{r}_{nm})\right) \\
&= \frac{1}{N}\text{Cov}\left(X_{ij}(\mathbf{0}), X_{kl}(\mathbf{0})\right) + \frac{1}{N^2}\sum_{\mathbf{r}_{nm}\neq\mathbf{0}} \text{Cov}\left(X_{ij}(\mathbf{0}), X_{kl}(\mathbf{r}_{nm})\right).
\end{aligned}
$$
(A.14)

The third line in (A.14) takes advantage of the stationarity of $X(\mathbf{s})$.

Covariance of the gradient product tensor: Let $\mathbf{r}$ denote any lag vector (including $\mathbf{r} = \mathbf{0}$) between two sampling points. Based on the definition of the covariance function it follows that

$$\text{Cov}\left(X_{ij}(\mathbf{0}), X_{kl}(\mathbf{r})\right) = \mathbb{E}\left[X_{ij}(\mathbf{0})X_{kl}(\mathbf{r})\right] - \mathbb{E}\left[X_{ij}(\mathbf{0})\right]\mathbb{E}\left[X_{kl}(\mathbf{r})\right].$$
(A.15)

Note that $\mathbb{E}\left[X_{ij}(\mathbf{0})X_{kl}(\mathbf{r})\right] = \mathbb{E}\left[\partial_i X(\mathbf{0})\,\partial_j X(\mathbf{0})\,\partial_k X(\mathbf{r})\,\partial_l X(\mathbf{r})\right]$. According to Theorem A.2.1, the gradient fields are Gaussian SRFs. Hence, $\mathbb{E}\left[X_{ij}(\mathbf{0})\,X_{kl}(\mathbf{r})\right]$ can be calculated using the moment factorization property of multivariate normal distributions Isserlis [1918]; Wick [1950]:

$$
\begin{aligned}
\mathbb{E}\left[X_{ij}(\mathbf{0})\,X_{kl}(\mathbf{r})\right] &= \mathbb{E}\left[\partial_i X(\mathbf{0})\,\partial_j X(\mathbf{0})\right]\mathbb{E}\left[\partial_k X(\mathbf{r})\,\partial_l X(\mathbf{r})\right] \\
&+ \mathbb{E}\left[\partial_i X(\mathbf{0})\,\partial_k X(\mathbf{r})\right]\mathbb{E}\left[\partial_j X(\mathbf{0})\,\partial_l X(\mathbf{r})\right] + \mathbb{E}\left[\partial_i X(\mathbf{0})\,\partial_l X(\mathbf{r})\right]\mathbb{E}\left[\partial_j X(\mathbf{0})\,\partial_k X(\mathbf{r})\right] \\
&= H_{ij}(\mathbf{0})\,H_{kl}(\mathbf{0}) + H_{ik}(\mathbf{r})\,H_{jl}(\mathbf{r}) + H_{il}(\mathbf{r})\,H_{jk}(\mathbf{r}).
\end{aligned}
$$
(A.16)

The last equality follows from the definition of the covariance function, Theorem A.2.1, and the definition (2.11). The second term on the right-hand side of (A.15) is simply

$$\mathbb{E}\left[X_{ij}(\mathbf{0})\right]\mathbb{E}\left[X_{kl}(\mathbf{r})\right] = \mathbb{E}\left[\partial_i X(\mathbf{0})\,\partial_j X(\mathbf{0})\right]\mathbb{E}\left[\partial_k X(\mathbf{r})\,\partial_l X(\mathbf{r})\right] = H_{ij}(\mathbf{0})\,H_{kl}(\mathbf{0}).$$

(A.17)

Thus, in light of (A.16) and (A.17), equation (A.15) becomes

$$\mathrm{Cov}\left(X_{ij}(\mathbf{0}), X_{kl}(\mathbf{r})\right) = H_{ik}(\mathbf{r})H_{jl}(\mathbf{r}) + H_{il}(\mathbf{r})H_{jk}(\mathbf{r}).$$

(A.18)

Using (A.18) in (A.14), and Theorem 2.1.1 to express the zero-lag component of the covariance Hessian matrix, equation (2.21) is obtained.

∎

A.4 Proof of Theorem 2.2.1

Proof. The $X_{ij}(\mathbf{s}_k)$ are stationary GSRFs by virtue of the stationarity of the GSRF $X(\mathbf{s})$. Hence, $\phi_{ijkl}(\mathbf{r}) := \mathrm{Cov}\left(X_{ij}(\mathbf{s}), X_{kl}(\mathbf{s}+\mathbf{r})\right) = \mathrm{Cov}\left(X_{ij}(\mathbf{0}), X_{kl}(\mathbf{r})\right)$. Based on (A.18), $\phi_{ijkl}(\mathbf{r}) = H_{ik}(\mathbf{r})H_{jl}(\mathbf{r}) + H_{il}(\mathbf{r})H_{jk}(\mathbf{r})$. The range of the GSRF $X_{ij}(\mathbf{s}_k)$ is determined by the integral

$$V_c = \max_{i,j,k,l}\left(\frac{1}{\phi_{ijkl}(\mathbf{0})}\int d\mathbf{r}\,\phi_{ijkl}(\mathbf{r}), \quad \left|\phi_{ijkl}(\mathbf{0}) \neq 0\right.\right).$$

Based on (A.18), $\phi_{ijkl}(\mathbf{0}) = Q_{ij}\,Q_{kl} + Q_{il}\,Q_{jk}$ and thus $\phi_{ijkl}(\mathbf{0})$ has a finite value for finite correlation lengths.

To calculate $\int_{\mathcal{D}} d\mathbf{r}\,\phi_{ijkl}(\mathbf{r})$, we assume that $|\mathcal{D}| \to \infty$ and express the integral in terms of the Fourier transform of $c_{xx}(\mathbf{r})$. Any permissible covariance function $c_{xx}(\mathbf{r})$ admits the following pair of transformations, where $\tilde{C}_{xx}(\mathbf{k})$ is the *covariance spectral density*:

$$c_{xx}(\mathbf{r}) = \frac{1}{(2\pi)^2}\int d\mathbf{k}\,e^{j\mathbf{k}\cdot\mathbf{r}}\tilde{C}_{xx}(\mathbf{k})$$

(A.19)

$$\tilde{C}_{xx}(\mathbf{k}) = \int d\mathbf{r}\,e^{-j\mathbf{k}\cdot\mathbf{r}}\,c_{xx}(\mathbf{r}).$$

(A.20)

Based on the above, we obtain $H_{ij}(\mathbf{r}) = \frac{1}{(2\pi)^2}\int d\mathbf{k}\,k_i k_j\,e^{j\mathbf{k}\cdot\mathbf{r}}\tilde{C}_{xx}(\mathbf{k})$, and thus

$$\int d\mathbf{r}\,\phi_{ijkl}(\mathbf{r}) = \frac{1}{(2\pi)^2}\int d\mathbf{k}\,k_i k_j k_k k_l[\tilde{C}_{xx}(\mathbf{k})]^2.$$

In the above, $J = \sqrt{-1}$, $\mathbf{k} \cdot \mathbf{r} = k_1 r_1 + k_2 r_2$ is the inner vector product, and $\int d\mathbf{k} = \int_{-\infty}^{\infty} dk_1 \int_{-\infty}^{\infty} dk_2$ or $\int d\mathbf{k} = \int_0^{\infty} k\, dk \int_0^{2\pi} d\phi$ in polar coordinates. The completeness of the plane-wave basis was used, i.e., $\int d\mathbf{r} e^{J(\mathbf{k}+\mathbf{k}')\cdot\mathbf{r}} = (2\pi)^2 \delta(\mathbf{k} + \mathbf{k}')$, where $\delta(\cdot)$ is the Dirac delta function. Existence of this integral requires integrability of $\tilde{C}_{xx}(\mathbf{k})$ at $k = 0$ and at $k \to \infty$. Since $c_{xx}(\mathbf{r})$ is short-ranged, the integral $\int d\mathbf{r}\, c_{xx}(\mathbf{r}) = \tilde{C}_{xx}(\mathbf{0})$ is finite, and thus the limit $k = 0$ is well-behaved. At the limit $k \to \infty$, the integral converges (using polar coordinates) if $[\tilde{C}_{xx}(\mathbf{k})]^2$ falls off faster than $k^{-6-2\epsilon}$, where $\epsilon > 0$. This ensures that the covariance $\phi_{ijkl}(\mathbf{r})$ is short-ranged. Thus, Theorem 2.2.1 applies to the vector random variable $\mathbf{Z}_k = ((X_{11}(\mathbf{s}_k), X_{22}(\mathbf{s}_k), X_{12}(\mathbf{s}_k))^t$ leading to (2.22). ∎

A.5 Proof of Lemma 2.2.3

Proof. We use Theorem A.1.2 with $\mathbf{Z} \to \vec{\mathbb{Q}}$, $\mathbf{Y} \to (\hat{q}_d, \hat{q}_o)^t$. Since $\dim(\mathbf{Y}) = 2 < \dim(\mathbf{Z}) = 3$, we append to $\mathbf{Y}$ the dummy variable $u = \hat{Q}_{11} \geq 0$. Using definitions (2.14) and (2.15), the absolute value of the Jacobian determinant for the transformation $(\hat{Q}_{11}, \hat{Q}_{22}, \hat{Q}_{12}) \to (\hat{q}_d, \hat{q}_o, u)$ is

$$\mathbf{J_q} = \frac{\partial(\hat{Q}_{11}, \hat{Q}_{22}, \hat{Q}_{12})}{\partial(\hat{q}_d, \hat{q}_o, u)} \Rightarrow \left|\det(\mathbf{J_q})\right| = u^2.$$

The dummy variable u is integrated according to Theorem A.1.2, leading to

$$f_{\mathbf{q}}(\hat{\mathbf{q}}; \mathbf{m}_{\mathbb{Q}}, \mathbf{C}_{\bar{\mathbb{Q}}}) = \int_0^{\infty} f_{\bar{\mathbb{Q}}}(u, \hat{q}_o u, \hat{q}_d u; \mathbf{m}_{\mathbb{Q}}, \mathbf{C}_{\bar{\mathbb{Q}}})\, u^2\, du. \qquad (A.21)$$

In terms of $\hat{q}_d$ and $\hat{q}_o$, the exponent of the trivariate PDF (2.22) is transformed as follows

$$(\vec{\mathbb{Q}} - \mathbf{m}_{\mathbb{Q}})^t\, \mathbf{C}_{\bar{\mathbb{Q}}}^{-1}\, (\vec{\mathbb{Q}} - \mathbf{m}_{\mathbb{Q}}) = A(\hat{\mathbf{q}}, \mathbf{C}_{\bar{\mathbb{Q}}})\, u^2 + B(\hat{\mathbf{q}}, \mathbf{m}_{\mathbb{Q}}, \mathbf{C}_{\bar{\mathbb{Q}}})\, u + C(\mathbf{m}_{\mathbb{Q}}, \mathbf{C}_{\bar{\mathbb{Q}}}). \quad (A.22)$$

By virtue of the above, the integral (A.21) is expressed as follows (suppressing the dependence of A, B, C, K for brevity):

$$f_{\mathbf{q}}(\hat{\mathbf{q}}; \mathbf{m}_{\mathbb{Q}}, \mathbf{C}_{\bar{\mathbb{Q}}}) = K \int_0^{\infty} u^2\, e^{-\frac{1}{2}(Au^2 + Bu + C)}\, du.$$

73

According to (2.26a), $A > 0$ because $\mathbf{C}_{\bar{\mathbb{Q}}}$ is a covariance matrix, and hence $\mathbf{C}_{\bar{\mathbb{Q}}}$ as well as $\mathbf{C}_{\bar{\mathbb{Q}}}^{-1}$ are positive definite[1]. Thus, the Gaussian integral above exists and its value is given by (2.25). ∎

A.6 Proof of Theorem 2.2.2

Proof. Equation (2.30) follows from the transformation $(\hat{q}_d, \hat{q}_o) \to (\hat{R}, \hat{\partial})$ with Jacobian matrix $\mathbf{J}_{\partial,R}$. According to A.1.2 the transformed PDF is given by

$$f_{\partial,R}(\hat{R}, \hat{\partial}; \mathbf{m}_{\mathbb{Q}}, \mathbf{C}_{\bar{\mathbb{Q}}}) = f_{\mathbf{q}}(\hat{\mathbf{q}}; \mathbf{m}_{\mathbb{Q}}, \mathbf{C}_{\bar{\mathbb{Q}}}) \left|\det(\mathbf{J}_{\partial,R})\right|$$

where $\det(\mathbf{J}_{\partial,R})$ is given by

$$\det(\mathbf{J}_{\partial,R}) = \begin{vmatrix} \dfrac{\partial \hat{q}_d}{\partial \hat{\partial}} & \dfrac{\partial \hat{q}_d}{\partial \hat{R}} \\ \dfrac{\partial \hat{q}_o}{\partial \hat{\partial}} & \dfrac{\partial \hat{q}_o}{\partial \hat{R}} \end{vmatrix} = \frac{2\hat{R}\left(\hat{R}^2 - 1\right)}{\left(\hat{R}^2 \cos^2 \hat{\partial} + \sin^2 \hat{\partial}\right)^3}. \tag{A.23}$$

Under the restriction of the parameter space to $R \in [0, \infty)$ and $\partial \in [-\pi/4, \pi/4)$, or equivalently $R \in [1, \infty)$ and $\partial \in [-\pi/2, \pi/2)$, the transformation $(\hat{q}_d, \hat{q}_o) \to (\hat{R}, \hat{\partial})$ is one-to-one with a Jacobian determinant given by (A.23). Finally, based on Theorems A.1.2 and 2.2.3, $f_{\partial,R}(\hat{R}, \hat{\partial})$ is given by (2.30). ∎

A.7 Some Properties of Common Covariance Functions

There is a plethora of covariance functions used in Geostatistics. In the following we will define the notion of *integral scale factor* and demonstrate some basic properties of two well-known functions, namely the Gaussian and the Whittle - Matérn (or simply Matérn) covariance functions (or models), which are examined in this work.

The spatial variability of a random field is controlled by the correlation length ξ in both the Gaussian and the Matérn model. The latter has an additional parameter, ν which is the *smoothness coefficient* and controls the differentiability of the random field. Specifically, as ν increases, the corresponding random field is smoother. While a second parameter adds

---

[1]A square $p \times p$ matrix $\mathbf{M}$ is positive-definite, denoted by $\mathbf{M} > 0$, if $\mathbf{x}^t \mathbf{M} \mathbf{x} > 0 \ \forall \ p \times 1$ vectors $\mathbf{x} \neq \mathbf{0}$; then, $\mathbf{M}^{-1} > 0$.

more flexibility, the spatial variability of a Matérn model is a function of both ξ and v. Even in the case of models with only one parameter, the same value of correlation length across different models, does not guarantee random fields with the same (or comparable) spatial variability.

In order to be able to compare the spatial variability between different isotropic covariance functions, the *integral scale factor* (or *integral range*) Hristopulos & Žukovič [2011] in d-dimensions is defined by the equation

$$\ell_c \doteq \left[\frac{\int d^d \mathbf{r}\, c_{xx}(r)}{c_{xx}(0)} \right]^{1/d} = \sigma_x^{-2/d}\, \tilde{c}_{xx}(0)^{1/d}. \tag{A.24}$$

A.7.1 The Gaussian Model

The isotropic Gaussian covariance function is

$$c_{xx}(\mathbf{r}) = \sigma_x^2 e^{-(r/\xi)^2}, \tag{A.25}$$

where $r = \|\mathbf{r}\|$. The two-dimensional covariance spectral representation is given by

$$\tilde{c}_{xx}(\mathbf{k}) = \pi \xi^2 \sigma_x^2 e^{-(k\xi/2)^2} \tag{A.26}$$

The first and second derivatives are

$$\frac{dc_{xx}(r)}{dr} = -2(\sigma_x/\xi)^2 r e^{-(r/\xi)^2}, \tag{A.27}$$

$$\frac{d^2 c_{xx}(r)}{dr^2} = -2(\sigma_x/\xi)^2 (1 - 2(r/\xi)^2) e^{-(r/\xi)^2}. \tag{A.28}$$

Specifically the second derivative at zero lag is

$$\left. \frac{d^2 c_{xx}(r)}{dr^2} \right|_{r=0} = -2(\sigma_x/\xi)^2. \tag{A.29}$$

The integral scale factor of the Gaussian model in 2-D is $\ell_c = \pi$.

A.7.2 The Whittle - Matérn Model

The isotropic Whittle - Matérn covariance function is given by

$$c_{xx}(\mathbf{r}) = \sigma_x^2 \frac{2^{1-v} r^v}{\Gamma(v)\xi^v} K_v\left(\frac{r}{\xi}\right), \tag{A.30}$$

where $K_\nu(\cdot)$ is the *modified Bessel function of the second kind* of order ν, $\Gamma(\cdot)$ is the *Gamma* function. For $\nu \to \infty$ the Gaussian model is recovered, while for $\nu \to 1/2$ the exponential model is obtained. The Matérn model corresponds to a *non-differentiable* field for $\nu \leq 1$ in $d = 2$.

The d-dimensional covariance spectral density is given by

$$\tilde{c}_{xx}(k) = \sigma_x^2 \frac{\Gamma(\nu + d/2)}{\pi^{d/2}\Gamma(\nu)} \frac{(2\pi\xi)^d}{(1 + (k\xi)^2)^{\nu+d/2}}. \tag{A.31}$$

Given the recurrence relation Abramowitz & Stegun [1970]; Gradshteyn & Ryzhik [2007],

$$\frac{d}{dz}(z^\nu K_\nu(z)) = -z^\nu K_{\nu-1}(z),$$

the first and second derivatives of the Matérn covariance function are

$$\frac{dc_{xx}(r)}{dr} = -\sigma_x^2 \frac{2^{1-\nu}r^\nu}{\Gamma(\nu)\xi^{\nu+1}} K_{\nu-1}\left(\frac{r}{\xi}\right), \tag{A.32}$$

$$\frac{d^2c_{xx}(r)}{dr^2} = \sigma_x^2 \frac{2^{1-\nu}r^{\nu-1}}{\Gamma(\nu)\xi^{\nu+2}}\left(rK_{\nu-2}\left(\frac{r}{\xi}\right) - \xi K_{\nu-1}\left(\frac{r}{\xi}\right)\right). \tag{A.33}$$

Given the limiting form for small arguments of the modified Bessel function of the second kind Gradshteyn & Ryzhik [2007],

$$K_\nu(z) \to \Gamma(\nu)2^{\nu-1}z^{-\nu}, \quad \text{for } z \to 0 \text{ and } \nu > 0,$$

the second derivative at zero lag is

$$\left.\frac{d^2c_{xx}(r)}{dr^2}\right|_{r=0} = -\frac{\sigma_x^2}{2(\nu - 1)\xi^2}. \tag{A.34}$$

The integral scale factor for the Matérn model in d-dimensions is $\ell_c = 2\pi^{1/2}\xi\left[\frac{\Gamma(\nu+d/2)}{\Gamma(\nu)}\right]^{1/d}$. For $d = 2$, $\ell_c = 2\sqrt{\pi\nu}$.
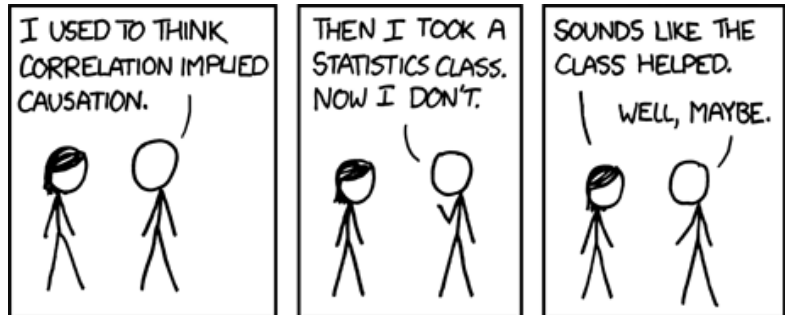
— RANDALL MUNROE, XKCD

Appendix B

Selected MATLAB® Scripts

This Appendix provides some of the MATLAB® scripts and functions used in this work. For a complete package please e-mail the author.

GNU General Public License

The following scripts are free software: you can redistribute them and/or modify them under the terms of the GNU General Public License as published by the Free Software Foundation, version 3 of the License, or any later version.

The following scripts are distributed in the hope that they will be useful, but WITHOUT ANY WARRANTY; without even the implied warranty of MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the GNU General Public License for more details:

www.gnu.org/licenses/gpl-3.0.html

B.1 Non-parametric Joint Probability Density Function

```matlab
1  function [p, cr] = jpdfNonParametric(Rhat, thetahat, Qij, N)
2
3  % Non — Parametric Joint Probability and Confidence Region for …
        estimates of
4  % R and theta.
5  %
6  %   p is the probability calculated over the domain [Rhat, …
        thetahat], provided
7  %   by a meshgrid command.
8  %
9  %   cr is the contour function to be plotted for a given …
        confidence level.
10 %
11 %   The vector Qij should contain the slope tensor elements in …
        the following
12 %   manner: Qij = [Q11, Q22, Q12]
13 %
14 %   N is the number if samples.
15 %
16 %   As of 22 December 2011
17 %
18 %
19 % This program is free software: you can redistribute
20 % it and/or modify it under the terms of the GNU General
21 % Public License as published by the Free Software
22 % Foundation, either version 3 of the License, or any
23 % later version.
24 %
25 % This program is distributed in the hope that it
26 % will be useful, but  WITHOUT ANY WARRANTY;
27 % without even the implied warranty of
28 % MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE.
29 % See the GNU General Public License for more details,
30 % <http://www.gnu.org/licenses/>
31
32 % Degree to radian conversion
33 thetahat = thetahat * pi/180 ;
34
35 Q11 = Qij(1); Q22 = Qij(2); Q12 = Qij(3);
36
37 % Non-Parametric CQ
38 CQ = 2/N * [  Q11^2      Q12^2          Q11*Q12             ;
39                Q12^2      Q22^2          Q12*Q22             ;
```

```matlab
40                  Q11*Q12    Q12*Q22   1/2*(Q12^2 + Q11*Q22) ] ;
41
42 % Slope tensor ratios calculated from Rhat, thetahat (which …
       are meshgrids)
43 qd = ( 1./Rhat.^2 + tan(thetahat).^2 ) ./ ( 1 + …
       (tan(thetahat)./Rhat).^2 ) ;
44 qo = tan(thetahat).*(1 – 1./Rhat.^2) ./ ( 1 + …
       (tan(thetahat)./Rhat).^2 ) ;
45 q = [ones(numel(thetahat), 1)  qd(:)  qo(:)]' ;
46
47 v = [Q11 ; Q22 ; Q12] ;
48
49 % Invert CQ
50 CQinv = inv(CQ) ;
51
52 % Preallocation for A, B
53 A = zeros(1, length(q)) ;
54 B = A ;
55
56 % Calculate A, B
57 for i=1:length(q)
58     A(i) = q(:,i)' * CQinv * q(:,i) ;
59     B(i) = –2*v'   * CQinv * q(:,i) ;
60 end
61 A = reshape(A, size(qd)) ;
62 B = reshape(B, size(qd)) ;
63
64 % Calculate C and K
65 C = v' * CQinv * v ;
66 K = (2*pi)^(–3/2) * sqrt( det(CQinv) ) ;
67
68 % Reparametrization
69 b = B./(2*sqrt(2*A)) ;
70 c = C/2 ;
71 k = K./(sqrt(2)*A.^(3/2)) ;
72
73 % Clean–up
74 clear A B C K
75
76 % fq
77 fq = 2*sqrt(pi) * k.*exp(b.^2–c) .* (2*b.^2+1) ;
78
79 % Absolute value of the Jacobian for the transformation to R, …
       theta
80 absJ = 2*Rhat.*abs(Rhat.^2–1)./((Rhat.*cos(thetahat)).^2 + …
       sin(thetahat).^2).^3 ;
81
82 % JPDF. Since we are using degrees, the Jacobian is pi/180
83 p = pi/180 .* fq .* absJ ;
```

79

```
84
85  % Confidence region (confidence level probability)
86  cr = 1 − exp(b.^2 − c) ;
```

B.2 Subsampling Procedure for Mining Data

```matlab
% Runs the subsampling procedure using mines data.
%
% As of 17 July 2012
%
% This program is free software: you can redistribute
% it and/or modify it under the terms of the GNU General
% Public License as published by the Free Software
% Foundation, either version 3 of the License, or any
% later version.
%
% This program is distributed in the hope that it
% will be useful, but  WITHOUT ANY WARRANTY;
% without even the implied warranty of
% MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE.
% See the GNU General Public License for more details,
% <http://www.gnu.org/licenses/>

% Choose a mine. Mavropigi has duplicates.
%mine = 'Amydaio';
mine = 'Mavropigi';

% Load data
[x, y, l] = loadLigniteData(mine) ;

% Artificial data
% x = data(:,1) ; y = data(:,2) ; l = data(:,3);

% SIC2004 data
%x = normalData(:,1) ; y = normalData(:,2) ; l = …
    normalData(:,3);

disp('——————————————————————————————————') ;
disp(['Mine :' mine]) ;
dataSize = length(x) ; disp(['N = ' num2str(dataSize)]);

% Descriptive statistics of z-values (lignite)
disp('——————Descriptive Statistics————————') ;
minl = min(l) ; disp(['min : ' num2str(minl)]) ;
meanl = mean(l); disp(['mean : ', num2str(meanl)]);
medianl = median(l); disp(['median : ',num2str(medianl)]) ;
stdl = std(l); disp(['std dev : ', num2str(stdl)]) ;
skewl = skewness(l) ; disp(['skewness : ', num2str(skewl)]) ;
kurtl = kurtosis(l)-3 ; disp(['excess kurtosis : ', …
    num2str(kurtl)]) ;
maxl = max(l); disp(['max : ', num2str(maxl)]);
```

```matlab
44
45  % Histogram / Weibull plot
46  % figure ; hist(l, 30) ;
47  % figure ; weibplot(l) ;
48
49  % Box-Cox ?
50  % [l, lambda] = boxcox(l); lambda
51
52  % Data normalization
53  unitlength = 1000 ; lscale = 1;
54  x = (x-mean(x))/unitlength ; y=(y-mean(y))/unitlength ; l = …
        l/lscale;
55  xmin = min(x); xmax = max(x); xr = xmax - xmin;
56  ymin = min(y); ymax = max(y); yr = ymax - ymin;
57
58  % Interpolation meshgrid
59  [xint yint] = meshgrid(linspace(xmin, xmax, 200), …
        linspace(ymin, ymax, 200));
60
61  % Interpolate using natural neighbor method; contains NaNs
62  lint = interpolateScatteredData(x', y', l', xint, yint, …
        'natural');
63
64  % Estimate anisotropy
65  est = estimateAnisotropy(xint,yint, ...
66      interpolateScatteredData(x', y', l', xint, yint, …
            'natural', mean(l)));
67  Rest = est(1); thetaest = est(2);
68  disp('------------Anisotropy--------------');
69  disp([Rest, thetaest]);
70
71  % Isotropy test
72  disp('-----------Isotropy test------------');
73  disp(isotropyConfidenceInterval(dataSize)) ;
74
75  % Draw anisotropy ellipse
76  scale = 0.1;
77  if strcmp(mine, 'Mavropigi')
78      center = [xmin + 7/8 * xr , ymin + 7/8 * yr];
79  else
80      center = [xmin + 7/8 * xr , ymin + 1/8 * yr];
81  end
82  semimajor = scale * xr / Rest;
83
84  % Ellipse points for visualizing estimated anisotropy
85  ellp = calculateEllipse(center(1), center(2), semimajor, …
        scale*xr, thetaest ,37);
86
87  % Plot data and anisotropy ellipse
```

```matlab
88  figure ;
89  hold on
90  box on
91  pcolor(xint, yint, lint) ; shading flat ;
92  t = colorbar('peer',gca);
93  set(get(t,'ylabel'), 'String', 'Lignite (m)');
94  plot(x,y,'.k') ;
95  plot(ellp(:,1), ellp(:,2), '-k') ;
96  arrow(center, [ellp(1,1) ,ellp(1,2)]) ;
97  arrow(center, [ellp(28,1), ellp(28,2)]) ;
98  axis equal tight
99  xlabel('x (km)') ; ylabel('y (km)') ;
100 hold off
101 figbolden;
102 print('-dpng', '-r400', [lower(mine),'.png']);
103
104 %break;
105
106 % Do subsampling and plot results
107 runs = 1000 ;
108
109 % Use 50% of the data
110 samples = round(0.5 * dataSize) ;
111 disp(['Samples used: ', num2str(samples)]);
112
113 % Scatter plot compared to theoretical confidence region
114 plotConfidenceIntervalScattered([x y l], runs, samples, [0.5 …
        1.5], [-45 45]);
115
116 %Isotropy Test for sparse dataset
117 disp('Isotropy test for sparse set') ;
118 disp(isotropyConfidenceInterval(samples)) ;
```


— RANDALL MUNROE, XKCD

83

References

ABRAMOWITZ, M. & STEGUN, I.A. (1970). *Handbook of Mathematical Functions*. Dover books on mathematics, Dover Publications, 1st edn. 27, 70, 76

ADLER, R.J. (1981). *The Geometry of Random Fields*. Wiley, New York, 1st edn. 14

ANDERSON, T.W. (1984). *An Introduction to Multivariate Statistical Analysis*. Wiley, New York, 3rd edn. 19

BOISSONNAT, J.D. & CAZALS, F. (2000). Smooth surface reconstruction via natural neighbour interpolation of distance functions. In *Proceedings of the sixteenth annual symposium on Computational geometry*, SCG '00, 223–232, ACM, New York, NY, USA. 45

BOUCHAUD, J..P. & GEORGES, A. (1990). Anomalous diffusion in disordered media: Statistical mechanisms, models and physical applications. *Physics Reports*, **195**, 127–293. 19

BOWMAN, A.W. & AZZILINI, A. (1997). *Applied Smoothing Techniques for Data Analysis: The Kernel Approach with S-Plus Illustrations*. Oxford Science Publications, Oxford University Press, USA. 53

CHILÈS, J. & DELFINER, P. (1999). *Geostatistics: Modeling Spatial Uncertainty*. Wiley series in probability and statistics, Wiley. 1

CHORTI, A. & HRISTOPULOS, D.T. (2008). Nonparametric identification of anisotropic (elliptic) correlations in spatially distributed data sets. *IEEE Transactions on Signal Processing*, **56**, 4738–4751. 9, 17, 37, 38, 44, 62

CHRISTAKOS, G. (1992). *Random Field Models in Earth Sciences*. Academic Press. 2, 6, 12

REFERENCES

CORNFORD, D. (2006). Dose and dose rates. http://wiki.intamap.org/index.php, [Online; accessed 24-July-2012]. 48

DE MONTAIGNE, M. (1958). *The complete essays of Montaigne*, vol. 1. Stanford University Press, translated by Donald M. Frame. i

DUBOIS, G. & GALMARINI, S. (2005). Introduction to the spatial interpolation comparison (sic) 2004 exercise and presentation of the datasets. *Applied GIS*, **1**, 9–1. 48

DUBOIS, G. & GALMARINI, S. (2006). Spatial interpolation comparison (SIC) 2004: introduction to the exercise and overview of results. In G. Dubois, ed., *Automatic Mapping Algorithms for Routine and Emergency Monitoring*, vol. EUR-21595-EN, 7–18, Office for Official Publications of the European Communities, Luxembourg, European Communities. 48

EPANECHNIKOV, V. (1969). Non-parametric estimation of a multivariate probability density. *Theory of Probability & Its Applications*, **14**, 153–158. 53

FELLER, W. (1971). *An Introduction to Probability Theory*. Wiley, New York. 19

FENG, L., HOTZ, I., HAMANN, B. & JOY, K.I. (2008). Anisotropic noise samples. *IEEE Transactions on Visualization and Computer Graphics*, **14**, 342–354. 7

FISHER, P., LEDOUX, H. & GOLD, C. (2005). An efficient natural neighbour interpolation algorithm for geoscientific modelling. In *Developments in Spatial Data Handling*, 97–108, Springer Berlin Heidelberg. 1, 44, 45

GNEDENKO, B.V. & KOLMOGOROV, A.N. (1954). *Limit Distributions for Sums of Independent Random Variables*. Addison Wesley, Reading, MA. 19

GOOVAERTS, P. (1997). *Geostatistics for Natural Resources Evaluation*. Applied Geostatistics Series, Oxford University Press. 1, 12

GRADSHTEYN, I. & RYZHIK, I. (2007). *Table of Integrals, Series, and Products*. Elsevier, New York, seventh edn. 70, 76

HERTZ, J., PALMER, R.G. & KROGH, A.S. (1991). *Introduction to the Theory of Neural Computation*. Perseus Publishing, 1st edn. 8

HINSHAW, G., WEILAND, J..L., HILL, R..S., ODEGARD, N., LARSON, D., BENNETT, C..L., DUNKLEY, J., GOLD, B., GREASON, M..R., JAROSIK, N., KOMATSU, E., NOLTA, M..R., PAGE, L., SPERGEL, D..N., WOLLACK, E., HALPERN, M., KOGUT, A., LIMON, M., MEYER, S..S., TUCKER, G..S. & WRIGHT, E..L. (2009). Five-year wilkinson microwave anisotropy probe (wmap) observations: Data processing, sky maps, and basic results. *The Astrophysical Journal Supplement Series*, **180**, 225–245. 7

HRISTOPULOS, D. (2002). New anisotropic covariance models and estimation of anisotropic parameters based on the covariance tensor identity. *Stochastic Environmental Research and Risk Assessment*, **16**, 43–62. 9, 17

HRISTOPULOS, D. & ŽUKOVIČ, M. (2011). Relationships between correlation lengths and integral scales for covariance models with more than two parameters. *Stochastic Environmental Research and Risk Assessment*, **25**, 11–19. 38, 75

HRISTOPULOS, D.T., CHORTI, A., SPILIOPOULOS, G. & PETRAKIS, M.P. (2008). Systematic detection of anisotropy in spatial data obtained from environmental monitoring networks. In *EGU General Assembly 2008, Vienna*. 62

HRISTOPULOS, D.T., PETRAKIS, M.P. & CHORTI, A. (2009). Non-parametric estimation of geometric anisotropy from environmental sensor network measurements. In *StatGIS2009, 17-19 June 2009, Milos, Greece*. 62

ISAAKS, E. & SRIVASTAVA, R. (1990). *An Introduction to Applied Geostatistics*. Oxford University Press, USA. 1

ISSERLIS, L. (1918). On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables. *Biometrika*, **12**, 134–139. 70, 71

JIANG, X. (2005). On orientation and anisotropy estimation for online fingerprint authentication. *IEEE Transactions on Signal Processing*, **53**, 4038 – 4049. 7

JOHNSTON, K., VER HOEF, J.M., KRIVORUCHKO, K. & LUCAS, N. (2003). Using ArcGIS Geostatistical Analyst. http://dusk.geo.orst.edu/gis/geostat_analyst.pdf, [Online; accessed 9-September-2012]. 2

JOHNSTON, K., VER HOEF, J.M., KRIVORUCHKO, K., LUCAS, N. & MAGRI, A. (2008). ArcGIS Geostatistical Analyst tutorial. http://webhelp.esri.com/arcgisdesktop/9.3/pdf/

REFERENCES

Geostatistical_Analyst_Tutorial.pdf, [Online; accessed 9-September-2012]. 2, 3

JOURNEL, A. & HUIJBREGTS, C. (2003). *Mining Geostatistics*. Blackburn Press. 1

JUPP, D., STRAHLER, A. & WOODCOCK, C. (1988). Autocorrelation and regularization in digital images. I. Basic theory. *IEEE Transactions on Geoscience and Remote Sensing*, **26**, 463–473. 12

KANEVSKI, M. & MAIGNAN, M. (2004). *Analysis and Modelling of Spatial Environmental Data*. EPFL Press. 12

KITANIDIS, P. (1997). *Introduction to Geostatistics: Applications in Hydrogeology*. Stanford-Cambridge Program, Cambridge University Press. 1

KRIVORUCHKO, K. (2008). Introduction to modelling spatial processes using Geostatistical Analyst. http://www.esri.com/library/whitepapers/pdfs/intro-modeling.pdf, [Online; accessed 9-September-2012]. 2

LANTUÉJOUL, C. (2002). *Geostatistical Simulation: Models and Algorithms*. Springer, New York. 37

LE BIHAN, D., MANGIN, J., POUPON, C., CLARK, C., PAPPATA, S., MOLKO, N. & CHABRIAT, H. (2001). Diffusion tensor imaging: concepts and applications. *Journal of Magnetic Resonance Imaging*, **13**, 534–546. 7

LEDOUX, H. & GOLD, C. (2008). Modelling three-dimensional geoscientific fields with the voronoi diagram and its dual. *International Journal of Geographical Information Science*, **22**, 547–574. 45

LEVY, P. (1954). *Théorie de l' Addition des Variables Aléatoires*. Gauthier Villars, Paris. 19

LI, J. & HEAP, A.D. (2008). *A review of spatial interpolation methods for environmental scientists*, vol. 23. Geoscience Australia. 3

LOH, J. (2008). A valid and fast spatial bootstrap for correlation functions. *The Astrophysical Journal*, **681**, 726. 44

LYNCH, P. (2008). The origins of computer weather prediction and climate modeling. *Journal of Computational Physics*, **227**, 3431 – 3444. 6

MANDELBROT, B. (1968). Fractional Brownian motions, fractional noises and applications. *SIAM Review*, **10**, 422–437. 13

MITAS, L. & MITASOVA, H. (1999). Spatial interpolation. In P. Longley, M.F. Goodchild, D.J. Maguire & D.W. Rhind, eds., *Geographical Information Systems: Principles, Techniques, Management and Applications*, vol. 1, chap. 34, 481–492, Wiley. 6

MUNROE, R. (2012). xkcd: A webcomic of romance, sarcasm, math, and language. http://xkcd.com/, [Online; accessed 12-December-2012]. 30, 60, 63, 76, 83

OKADA, K., COMANICIU, D. & KRISHNAN, A. (2005). Robust anisotropic Gaussian fitting for volumetric characterization of pulmonary nodules in multislice CT. *IEEE Transactions on Medical Electronics*, **24**, 409–423. 7

OLHEDE, S.C. (2008). Localisation of geometric anisotropy. *IEEE Transactions on Signal Processing*, **56**, 2133–2137. 7

PAPOULIS, A. & PILLAI, S.U. (2002). *Probability, Random Variables and Stochastic Processes*. McGraw-Hill, New York. 12, 18, 67

PARDO-IGÚZQUIZA, E. & CHICA-OLMO, M. (1993). The Fourier integral method: an efficient spectral method for simulation of random fields. *Mathematical Geology*, **25**, 177–217. 14, 37

PARZEN, E. (1962). On estimation of a probability density function and mode. *The annals of mathematical statistics*, **33**, 1065–1076. 53

PEBESMA, E., CORNFORD, D., DUBOIS, G., HEUVELINK, G., HRISTOPOULOS, D., PILZ, J., STOEHLKER, U., MORIN, G. & SKOIEN, J. (2011). Intamap: the design and implementation of an interoperable automated interpolation web service. *Computers and Geosciences*, **37**, 343–352. 8, 9

PEBESMA, E. *et al.* (2012). Package gstat: spatial and spatio-temporal geostatistical modelling, prediction and simulation. http://cran.r-project.org/web/packages/gstat/gstat.pdf, [Online; accessed 24-July-2012]. 50

PETRAKIS, M.P. & HRISTOPULOS, D.T. (2012). Elliptical anisotropy statistics of two-dimensional differentiable gaussian random fields: Joint probability density function and confidence regions. http://arxiv.org/abs/1203.5010. 62

REFERENCES

RANGANATHAN, S.I., OSTOJA-STARZEWSKI, M. & FERRARI, M. (2011). Quantifying the anisotropy in biological materials. *Journal of Applied Mechanics*, **78**, 064501. 7

RASMUSSEN, C.E. & WILLIAMS, C.K.I. (2006). *Gaussian Processes for Machine Learning*. MIT Press, Boston, MA. 7

RICHARD, F. & BIERME, H. (2010). Statistical tests of anisotropy for fractional Brownian textures. application to full-field digital mammography. *Journal of Mathematical Imaging and Vision*, **36**, 227–240. 7

SAMBRIDGE, M., BRAUN, J. & McQUEEN, H. (1995). Geophysical parametrization and interpolation of irregular data using natural neighbours. *Geophysical Journal International*, **122**, 837–857. 45

SCHMIDT, A.M. & O'HAGAN, A. (2003). Bayesian inference for non-stationary spatial covariance structure via spatial deformations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **65**, 743–758. 9

SIBSON, R. (1981). A brief description of natural neighbour interpolation. In *Interpreting Multivariate Data*, Wiley Series in Probability and Mathematical Statistics, 21–36, John Wiley & Sons, Chichester. 45

SIOTANI, M. (1964). Tolerance regions for a multivariate normal population. *Annals of the Institute of Statistical Mathematics*, **16**, 135–153. 29

SOONG, T. (2004). *Fundamentals of Probability and Statistics for Engineers*. John Wiley & Sons. 67

SPILIOPOULOS, I., HRISTOPULOS, D.T., PETRAKIS, M.P. & CHORTI, A. (2011). A multigrid method for the estimation of geometric anisotropy in environmental data from sensor networks. *Computers and Geosciences*, **37**, 320–330. 62

STEIN, M.L. (2001). Local stationarity and simulation of self-affine intrinsic random functions. *IEEE Transactions on Information Theory*, **47**, 1385–1390. 13

SWERLING, P. (1962). Statistical properties of the contours of random surfaces. *IRE Transactions on Information Theory*, **IT-8**, 315–321. 14

WACKERNAGEL, H. (1997). *Multivariate Geostatistics*. Springer Verlag, Berlin. 1

WANG, L. & LECKIE, C.A. (2012). Improved Gaussian process classification via feature space rotation. *Neurocomputing*, **83**, 89–97. 7

WEAVER, A.T. & MIROUZE, I. (2012). On the diffusion equation and its application to isotropic and anisotropic correlation modelling in variational assimilation. *Quarterly Journal of the Royal Meteorological Society*, **n/a**, n/a. 7

WICK, G.C. (1950). The evaluation of the collision matrix. *Physical Review*, **80**, 268–272. 70, 71

WIKIPEDIA CONTRIBUTORS (2012a). Lewis Fry Richardson. Retrieved from http://en.wikipedia.org/wiki/Lewis_Fry_Richardson, [Online; accessed 9-September-2012]. 6

WIKIPEDIA CONTRIBUTORS (2012b). Natural neighbor. Retrieved from http://en.wikipedia.org/wiki/Natural_neighbor, [Online; accessed 24-July-2012]. 45

XU, Y. & CHOI, J. (2009). Mobile sensor networks for learning anisotropic Gaussian processes. In *American Control Conference, 2009. ACC '09*, 5049 – 5054. 7

YAGLOM, A.M. (1987). *Correlation Theory of Stationary and Related Random Functions I*. Springer Verlag, New York. 2, 6, 12, 14, 68