

DISSERTATION

NONPARAMETRIC TESTS OF SPATIAL ISOTROPY AND A  
CALIBRATION-CAPTURE-RECAPTURE MODEL

Submitted by

Zachary D. Weller

Department of Statistics

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Spring 2017

Doctoral Committee:

Advisor: Jennifer A. Hoeting

Dan Cooley

Mevin Hooten

Jason Ahola

Copyright by Zachary D. Weller 2017

All Rights Reserved

## ABSTRACT

### NONPARAMETRIC TESTS OF SPATIAL ISOTROPY AND A CALIBRATION-CAPTURE-RECAPTURE MODEL

In this dissertation we present applied, theoretical, and methodological advances in the statistical analysis of spatially-referenced and capture-recapture data. An important step in modeling spatially referenced data is choosing the spatial covariance function. Due to the development of a variety of covariance models, practitioners are faced with a myriad of choices for the covariance function. One of these choices is whether or not the covariance function is isotropic. Isotropy means that the covariance function depends only the distance between observations in space and not their relative direction. Part I of this dissertation focuses on nonparametric hypothesis tests of spatial isotropy.

Statisticians have developed diagnostics, including graphical techniques and hypothesis tests, to assist in determining if an assumption of isotropy is adequate. Nonparametric tests of isotropy are one subset of these diagnostic methods, and while the theory for several nonparametric tests has been developed, the efficacy of these methods in practice is less understood. To begin part I of this dissertation, we develop a comprehensive review of nonparametric hypothesis tests of isotropy for spatially-referenced data. Our review provides informative graphics and insight about how nonparametric tests fit into the bigger picture of modeling spatial data and considerations for choosing a test of isotropy. An extensive simulation study offers comparisons of method performance and recommendations for test implementation. Our review also gives rise to a number of open research questions.

In the second section of part I, we develop and demonstrate software that implements several of the tests. Because the tests were not available in software, we created the R package **spTest**, which implements a number of nonparametric tests of isotropy. The package is open source and available on the Comprehensive R Archive Network (CRAN). We provide a detailed demonstration of how to use **spTest** for testing isotropy on two spatially-referenced

data sets. We offer insights into test limitations and how the tests can be used in conjunction with graphical techniques to evaluate isotropy properties.

To conclude our work with spatially-referenced data in part I, we develop a new non-parametric test of spatial isotropy using the spectral representation of the spatial covariance function. Our new test overcomes some of the short-comings of other nonparametric tests. We develop theory that describes the distribution of our test statistic and explore the efficacy of our test via simulations and applications. We also note several difficulties in implementing the test, explore remedies to these difficulties, and propose several areas of future work.

Finally, in part II of this dissertation, we shift our focus away from spatially-referenced data to capture-recapture data. Our capture-recapture work is motivated by methane concentration data collected by new mobile sensing technology. Because this technology is still in its infancy, there is a need to develop algorithms to extract meaningful information from the data. We develop a new Bayesian hierarchical capture-recapture model which we call the calibration-capture-recapture (CCR) model. We use our model and methane data to estimate the number and emission rate of methane sources within an urban sampling region. We apply our CCR model to methane data collected in two U.S. cities. Our new CCR model provides a framework to draw inference from data collected by mobile sensing technologies. The methodology for our capture-recapture model is useful in other capture-recapture settings, and the results of our model are important for informing climate change and infrastructure discussions.



## ACKNOWLEDGEMENTS

I am grateful for the support, advice, and expertise of many people during the completion of this work. First, I thank God for giving me the ability and perseverance to complete this dissertation. Next, I am grateful for the support of my family. My success is a direct reflection of the time my parents, Jim and Jackie, invested in my upbringing and the work ethic they instilled in me at a young age. I am also appreciative of the support from my siblings: Grant, Savana, Macy, and Simon.

I owe a great deal of gratitude to my advisor, Dr. Jennifer Hoeting, for believing in my abilities as a statistician. She has been instrumental in my development as a researcher, and I cannot thank her enough for her time, patience, feedback, and willingness to look out for students' best interests. I am very fortunate to call Jennifer a mentor and a friend.

I am also grateful for my relationships with other CSU faculty. My committee of Dan Cooley, Mevin Hooten, and Jason Ahola have provided constructive feedback in developing this dissertation. The Hooten-Hoeting lab provided constructive feedback on the work in Chapter 3. Piotr Kokoszka was instrumental in improving the work in Chapter 4. My collaborator Joe von Fischer has supported me through funding and the opportunity to go into the field to observe the data collection process. I am very thankful for Mark Dahlke's support and willingness to listen during the tough times. His encouragement and teaching mentorship helped keep me motivated during the completion of this work.

I have also been supported by faculty outside of CSU. I would like to thank Peter Guttorp and Alexandra Schmidt for organizing the 2014 Pan-American Study Institute (PASI) on spatial and spatial-temporal statistics in Buzios, Brazil. My experience at PASI inspired the work in Chapters 2-4. Michael Stein has been helpful by providing feedback on the work in Chapter 4.

I have been fortunate to develop numerous friendships during my time as a graduate student. I am particularly thankful for my relationships with fellow graduate students David Russo,

Ryan Haunfelder, Wade Herndon, Greg Ellison, Luke Muzynoski, Ryan Hicks, Erin Schliep, Ben Goldman, Frank Marrs, Ben Prytherch, Henry Scharf, Josh Hewitt, and Sarah Koehler. I am happy to call Travis Engle and Mike Callahan great friends outside of the CSU Statistics department. I am extremely grateful for the support of Darin and Anita Fossum, my Fort Collins parents, and Jim, Ruth, and Lisa Herickhoff, my Fort Collins family.

My graduate work was supported in part by the network for Statistical Methods for Atmospheric and Oceanic Sciences (STATMOS) under NSF Grant DMS-1106862. The work in Chapter 5 was supported by the Environmental Defense Fund (EDF). I would like to thank Adam Gaylord, Sam Chamberlain, and Duck Keun Yang for managing the methane data and providing useful summaries.

## DEDICATION

*To those who believed in me.*

## TABLE OF CONTENTS

<b>Abstract . . . . .</b>	<b>ii</b>
<b>Acknowledgements . . . . .</b>	<b>iv</b>
<b>Dedication . . . . .</b>	<b>vi</b>
<b>List of Tables . . . . .</b>	<b>ix</b>
<b>List of Figures . . . . .</b>	<b>x</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Motivation and Outline . . . . .	1
1.2 Spatial Statistics . . . . .	2
<b>2 A Review of Nonparametric Tests of Isotropy</b>	<b>8</b>
2.1 Introduction . . . . .	8
2.2 Additional Background and Definitions . . . . .	11
2.3 Tests of Isotropy and Symmetry . . . . .	13
2.4 Simulation Study . . . . .	23
2.5 Recommendations . . . . .	29
2.6 Discussion . . . . .	37
2.7 Simulation Study Details and Further Results . . . . .	39
<b>3 spTest: an R Package Implementing Nonparametric Tests of Isotropy</b>	<b>47</b>
3.1 Introduction . . . . .	47
3.2 Nonparametric Covariance Functions and Estimation . . . . .	48
3.3 Nonparametric Tests of Isotropy . . . . .	50
3.4 Nonparametric Tests Implemented in spTest . . . . .	52
3.5 Applications: Using spTest to Check for Anisotropy . . . . .	54

3.6	Discussion . . . . .	75
<b>4</b>	<b>A Spectral Domain Nonparametric Test of Spatial Isotropy</b>	<b>77</b>
4.1	Introduction . . . . .	77
4.2	Background and Definitions . . . . .	79
4.3	Periodogram . . . . .	82
4.4	Test Statistic . . . . .	89
4.5	Simulation Study . . . . .	95
4.6	Application: WRFG Temperatures . . . . .	104
4.7	Discussion . . . . .	105
4.8	Aliasing in the Spectral Domain . . . . .	106
<b>5</b>	<b>A Calibration-Capture-Recapture Model</b>	<b>121</b>
5.1	Introduction . . . . .	121
5.2	Data Collection and Processing . . . . .	124
5.3	A Calibration-Capture-Recapture Model . . . . .	128
5.4	MCMC Algorithm . . . . .	138
5.5	Prior Distributions and Data Augmentation . . . . .	150
5.6	Application: A Tale of Two Cities . . . . .	152
5.7	Discussion . . . . .	160
<b>6</b>	<b>Conclusion and Future Work</b>	<b>163</b>
6.1	Future Work: Nonparametric Tests of Spatial Isotropy . . . . .	163
6.2	Future Work: Calibration-Capture-Recapture . . . . .	165
6.3	Conclusions . . . . .	167
	<b>References</b>	<b>168</b>

## LIST OF TABLES

<b>Chapter 2</b>	8
Table 2.1 Properties of nonparametric tests of isotropy	17
Table 2.2 Test implementation, part 1	18
Table 2.3 Test implementation, part 2	24
Table 2.4 General recommendations for test implementation	32
Table 2.5 Empirical size and power, gridded sampling locations	43
Table 2.6 Empirical size and power, non-gridded sampling locations	44
Table 2.7 Effects of changing the lag set.	44
Table 2.8 Effects of changing the window/block size	46
Table 2.9 Effects of changing bandwidth	46
<b>Chapter 3</b>	47
Table 3.1 Nonparametric tests of isotropy available in the R package <code>spTest</code>	53
<b>Chapter 4</b>	77
Table 4.1 Comparison of test statistic performance.	97
Table 4.2 Comparison of testing methods.	98
Table 4.3 Empirical power as a function of $n$ .	101
Table 4.4 Computation time comparison.	102
<b>Chapter 5</b>	121
Table 5.1 A summary of the calibration-capture-recapture (CCR) model.	137
Table 5.2 Prior distributions and parameter values for our CCR model.	150
Table 5.3 The number of augmented sources for each city by sampling attempt.	151
Table 5.4 Counts of observed methane sources in two U.S. cities.	152
Table 5.5 Posterior estimates from the CCR model.	153
Table 5.6 Posterior estimates of the calibration parameters.	154

## LIST OF FIGURES

<b>Chapter 1</b>	1
Figure 1.1 Heat maps of isotropic and anisotropic processes	5
<b>Chapter 2</b>	8
Figure 2.1 Isotropy modeling flow chart	10
Figure 2.2 Empirical size and power, gridded sampling locations	30
Figure 2.3 Empirical size and power, non-gridded sampling locations.	31
Figure 2.4 Spatial sampling design considerations for choosing a test	34
Figure 2.5 Correlogram and contours of equal correlation	41
Figure 2.6 Lag sets used for the simulations	45
<b>Chapter 3</b>	47
Figure 3.1 Heat map of WRFG temperatures in latitude and longitude coordinates	57
Figure 3.2 Graphical assessments of isotropy in WRFG temperature data	59
Figure 3.3 Graphical assessments of isotropy in residuals from the WRFG data	63
Figure 3.4 Quilt plots showing the locations of the weather stations in Colorado	66
Figure 3.5 Results from the model relating log(precipitation) and elevation	67
Figure 3.6 A graphical assessment of isotropy in residuals from precipitation data	68
Figure 3.7 Quilt plot of the studentized residuals from precipitation data	72
<b>Chapter 4</b>	77
Figure 4.1 Isotropic Matérn covariance and spectral density functions.	81
Figure 4.2 Plot of Fourier frequencies.	90
Figure 4.3 Empirical power of spectral hypothesis tests as a function of $\theta$ .	99
Figure 4.4 Periodogram values as a function of geometric anisotropy.	100
Figure 4.5 Heat maps of $25 \times 20$ subset of WRFG data.	105
Figure 4.6 Example of aliasing.	106
Figure 4.7 Matérn spectral density and aliased spectral density functions I.	108

Figure 4.8	Matérn spectral density and aliased spectral density functions II. . . .	109
Figure 4.9	Sampling locations from full grid and two sub grids. . . . .	112
Figure 4.10	Fourier frequencies from full grid and two sub grids. . . . .	113
Figure 4.11	Comparison of periodogram values from the full and subgrids. . . . .	115
Figure 4.12	Check for aliasing in the WRFG residuals. . . . .	116
Figure 4.13	Covariance parameter estimates using the Whittle approximation I. . .	119
Figure 4.14	Covariance parameter estimates using the Whittle approximation II. .	120
<b>Chapter 5</b>	. . . . .	121
Figure 5.1	Cartoon illustrating detection of methane sources with GSV vehicles.	127
Figure 5.2	Relationship between $\log(\text{release rate})$ and $\log(\text{excess } CH_4)$ . . . . .	129
Figure 5.3	Estimated detection functions in two U.S. cities. . . . .	154
Figure 5.4	Posterior distributions for the estimated number of undetected sources.	155
Figure 5.5	Estimated number of undetected sources in City B by sampling attempt.	156
Figure 5.6	Estimated number of undetected sources in City A by sampling attempt.	156
Figure 5.7	Posterior distribution of total emissions. . . . .	157
Figure 5.8	Posterior distributions of individual emission rates. . . . .	158
Figure 5.9	Estimated cumulative emissions by cumulative ranked sources. . . . .	159



## CHAPTER 1

### INTRODUCTION

#### 1.1 Motivation and Outline

An important step in modeling spatially-referenced data is appropriately specifying the second-order properties of the random field. Due the emergence of large spatially referenced data sets, statisticians have recently developed a plethora of new spatial models for describing second-order structure. A scientist developing a model for spatial data has a number of options regarding the nature of the dependence between observations. One of these options is deciding whether or not the dependence between observations depends on direction, or, in other words, whether or not the spatial covariance function is isotropic. Isotropy implies that spatial dependence is a function of only the distance and not the direction of the spatial separation between sampling locations.

A researcher may use graphical techniques, such as directional sample semivariograms, to determine whether an assumption of isotropy holds. These graphical diagnostics can be difficult to assess, subject to personal interpretation, and potentially misleading as they typically do not include a measure of uncertainty. A hypothesis test can serve as an alternative to graphical techniques and be less subjective. To avoid potential misspecification of the spatial covariance function, statisticians have developed a number of nonparametric tests of isotropy (e.g., Guan et al., 2004; Maity and Sherman, 2012). While these nonparametric tests have sound theoretical development, there is less understanding regarding the use and performance of the methods for applied problems. Additionally, these methods were not widely available in software. In Chapters 2 and 3, we explore the use of these methods in practice and develop and demonstrate software that implements several of the investigated tests.

Through our work, we provide clear guidelines and examples of how to use these methods to assist practitioners in deciding whether or not an assumption of isotropy is reasonable.

An exploration of existing nonparametric tests of spatial isotropy reveals several practical challenges for each of the methods. In Chapter 4 we develop a new nonparametric test of isotropy that overcomes some of these challenges. Our new method uses the spectral representation of the spatial covariance function. We explore the challenges of aliasing when using the spectral representation, and develop a diagnostic to aid in determining whether or not aliasing is a concern. Our new test and diagnostic provide new tools for assessing anisotropy and aliasing concerns in spatial data.

Finally, we change our focus in Chapter 5 and develop a new capture-recapture model, which we call calibration-capture-recapture (CCR), to estimate the number of methane sources in an urban sampling area. We apply our model to data collected by atmospheric methane analyzers mounted on Google Street View (GSV) cars. Because this mobile detection technology has only recently been deployed, our model is one of the earliest statistical models used to analyze methane data collected by mobile detection platforms. Our model couples calibration and capture-recapture models to estimate the number of methane sources, the probability of source detection, and source emission rate.

In the remainder of this chapter, we provide notation and definitions for spatial statistics used in Chapters 2-4.

## 1.2 Spatial Statistics

In this section we provide background and definitions on geostatistics. For additional background, see Schabenberger and Gotway (2004). Let  $\{Y(\mathbf{s}) : \mathbf{s} \in \mathcal{D} \subseteq \mathbb{R}^d\}$  be a set of random variables with a well-defined joint probability distribution and indexed by spatial location  $\mathbf{s}$ . We call the set  $\{Y(\mathbf{s})\}$  a spatial stochastic process or random field (RF). For the remainder of the dissertation, we assume the common case of  $d = 2$ , although many of the results hold for the more general case of  $d \geq 2$ .

Let  $\{\mathbf{s}_1, \dots, \mathbf{s}_n\} \in \mathcal{D}$  be the finite set of locations at which the RF is observed, providing the random vector  $\mathbf{Y} = (Y(\mathbf{s}_1), \dots, Y(\mathbf{s}_n))^\top$ . The sampling locations may follow one of several spatial sampling designs, for example, gridded locations, randomly and uniformly distributed locations, a cluster design, or any other general design. Note that there is a distinction between a lattice process and a geostatistical process observed on a grid (e.g., Fuentes and Reich, 2010; Schabenberger and Gotway, 2004, pg. 6-10).

It is often of interest to infer the effect of covariates on the RF, deduce dependence structure, and/or predict  $Y$  with associated uncertainty at new locations. To achieve these goals, we must specify the distributional properties of the spatial process. A common assumption is that the finite-dimensional joint distribution of  $\{Y(\mathbf{s}) : \mathbf{s} \in \mathcal{D} \subseteq \mathbb{R}^2\}$  is multivariate normal, in which case we call the RF a Gaussian random field (GRF).

A RF is strictly stationary if its distribution is invariant under coordinate translation, thus

$$\mathbb{P}(Y(\mathbf{s}_1) < y_1, \dots, Y(\mathbf{s}_k) < y_k) = \mathbb{P}(Y(\mathbf{s}_1 + \mathbf{h}) < y_1, \dots, Y(\mathbf{s}_k + \mathbf{h}) < y_k) \quad (1)$$

for all spatial lags  $\mathbf{h} = (h_1, h_2)^\top \in \mathbb{R}^2$  and  $k \in \mathbb{Z}^+$ . A RF is weakly, or second-order, stationary if

$$\begin{aligned} \mathbb{E}(Y(\mathbf{s})) &= \mathbb{E}(Y(\mathbf{s} + \mathbf{h})) = \mu \\ \text{and} \end{aligned} \quad (2)$$

$$\text{Cov}(Y(\mathbf{s}), Y(\mathbf{s} + \mathbf{h})) = \text{Cov}(Y(\mathbf{0}), Y(\mathbf{h})) = C(\mathbf{h}),$$

where  $C(\mathbf{h}) < \infty$  is called the covariance function or covariogram. In this dissertation we focus on second-order properties of the random field; thus, hereafter we assume  $\mu = 0$ . A RF that is weakly stationary (2) has a constant mean, and the covariance between values at two spatial locations depends only on the spatial lag. Absolute coordinates are irrelevant for the

mean, variance, and covariance of a weakly stationary RF. A RF is intrinsically stationary if

$$\begin{aligned} E(Y(\mathbf{s} + \mathbf{h}) - Y(\mathbf{h})) &= 0 \\ \text{and} \end{aligned} \tag{3}$$

$$\text{Var}(Y(\mathbf{s} + \mathbf{h}) - Y(\mathbf{s})) = 2\boldsymbol{\gamma}(\mathbf{h}),$$

where  $2\boldsymbol{\gamma}(\mathbf{h})$  is called the variogram function or variogram, and  $\boldsymbol{\gamma}(\mathbf{h}) = \frac{1}{2}\text{Var}(Y(\mathbf{s} + \mathbf{h}) - Y(\mathbf{s}))$  is the semivariogram function or semivariogram. Note that intrinsic stationarity is defined in terms of the increments  $Y(\mathbf{s} + \mathbf{h}) - Y(\mathbf{s})$ .

The relationship between the different types of stationarity is given by

$$\text{strict} \implies \text{weak} \implies \text{intrinsic}. \tag{4}$$

In general the converse of (4) is not true; however, the converse of (4) holds under the assumption that the RF is a GRF. From (2) and (3) it follows that  $\boldsymbol{\gamma}(\mathbf{h}) = C(\mathbf{0}) - C(\mathbf{h})$  when the RF is weakly stationary. This relationship implies that the second-order properties of a weakly stationary RF can be viewed from the perspective of either the (semi)variogram or covariogram.

A common simplifying assumption on the spatial dependence structure is that it is isotropic.

**Definition 1.** *A second-order stationary spatial process is isotropic if the semivariogram,  $\boldsymbol{\gamma}(\mathbf{h})$  (or equivalently, the covariance function  $C(\mathbf{h})$ ) of the spatial process depends on the lag vector  $\mathbf{h}$  only through its Euclidean length,  $h = \|\mathbf{h}\|$  (i.e.,  $\boldsymbol{\gamma}(\mathbf{h}) = \boldsymbol{\gamma}_0(h)$  for some function  $\boldsymbol{\gamma}_0(\cdot)$  of a univariate argument).*

Isotropy implies that the dependence between any two observations depends only on the distance between their sampling locations and not on their relative orientation. A spatial process that is not isotropic is called anisotropic. Anisotropy is often broadly categorized

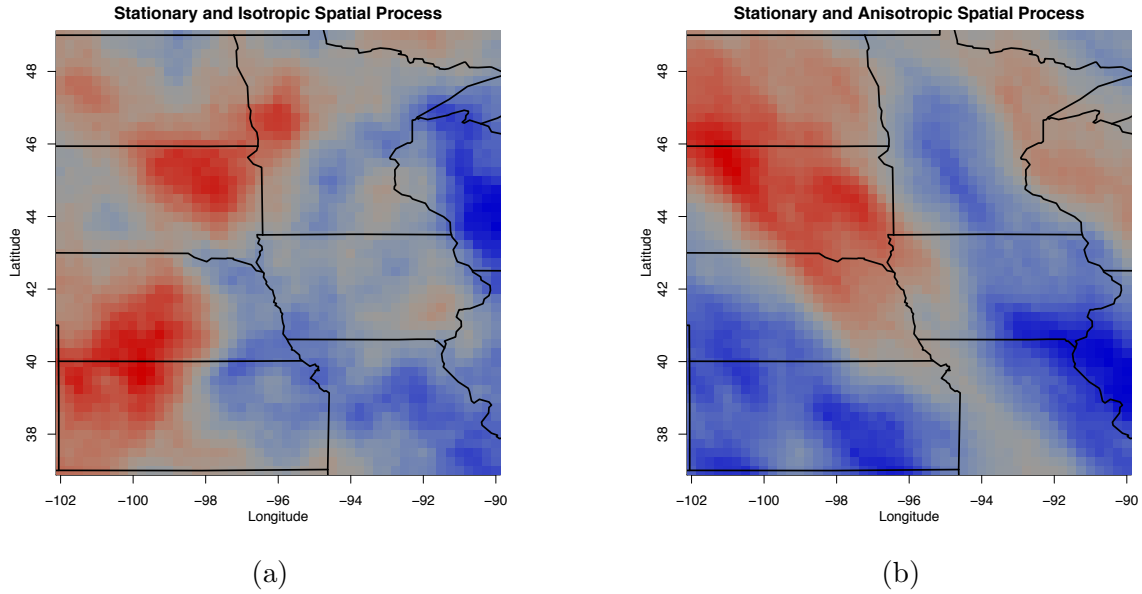


Figure 1.1: Heat maps displaying simulated data from two stationary random fields (RFs) over a section of the midwestern United States. Figure 1.1a displays data simulated from an isotropic covariance function. Figure 1.1b shows data simulated from a geometrically anisotropic covariance function. The existence of anisotropy in Figure 1.1b is evidenced by bands of similar values running in the northwest to southeast directions. These bands indicate that spatial dependence is stronger in the NW-SE direction than the SW-NE direction.

as either geometric or zonal (Zimmerman, 1993). In practice, if a process is assumed to be anisotropic, it is often assumed to be geometrically anisotropic due to its precise formal and functional definition (Ecker and Gelfand, 1999). Geometric anisotropy is governed by a scaling parameter,  $R > 1$ , and rotation parameter,  $0 \leq \theta < 2\pi$ , and implies the anisotropy can be corrected via a linear transformation of the lag vector or, equivalently, the sampling locations (Cressie, 1993, pg. 64). Geometric anisotropy implies that the range, the distance at which correlation between observations equals zero, of the spatial process is direction dependent. Figure 1.1 displays data simulated from two stationary RFs, one that is isotropic and one that is geometrically anisotropic. Although there is some disagreement on the definition (e.g., see Isaaks and Srivastava (1989, pg. 377,385) and Journel and Huijbregts (1978, pg. 181)), most authors agree that a common element of zonal anisotropy is a direction dependent sill. For a second-order stationary RF, the sill is defined as  $C(\mathbf{0})$ , the value of the covariance function at the origin. We refer the interested reader to Zimmerman (1993) for more discussion on geometric, zonal, and other types of anisotropies.

The aforementioned property of isotropy is defined in terms of examining the spatial random process in the spatial domain, where second-order properties depend on the spatial separation,  $\mathbf{h}$ . Alternatively, a spatial random process and its covariance function can be represented in the spectral domain using Fourier analysis. For the purposes of investigating second-order properties, we are interested in the spectral representation of the covariance function, called the spectral density function and denoted  $f(\boldsymbol{\omega})$ , where  $\boldsymbol{\omega} = (\omega_1, \omega_2)$ . Under certain conditions and assumptions (Fuentes and Reich, 2010, pg. 62), the spectral density function is given by

$$f(\boldsymbol{\omega}) = \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} \exp(-i\boldsymbol{\omega}^T \mathbf{h}) C(\mathbf{h}) d\mathbf{h}, \quad (5)$$

so that the covariance function,  $C(\mathbf{h})$ , and the spectral density function,  $f(\boldsymbol{\omega})$ , form a Fourier transform pair.

Properties of the covariance function imply properties of the spectral density. Consequently, second-order properties of a second-order stationary RF can be explored via either the covariance function or the spectral density function. For example, if the covariance function is isotropic, then the spectral density (5) depends on  $\boldsymbol{\omega}$  only through its length,  $\omega = ||\boldsymbol{\omega}||$ , and we can write  $f(\boldsymbol{\omega}) = f_0(\omega)$ , where  $f_0(\cdot)$  is called the isotropic spectral density (Fuentes, 2013). We provide more details on the spectral representation and its properties in Chapter 4.

## CHAPTER 2

### A REVIEW OF NONPARAMETRIC TESTS OF ISOTROPY

#### 2.1 Introduction

Early spatial models relied on the simplifying assumptions that the covariance function is stationary and isotropic. With the emergence of new sources of spatial data, for instance, remote sensing via satellite, climate model output, or environmental monitoring, a variety of methods and models have been developed that relax these assumptions. In the case of anisotropy, there are a number of methods for modeling both zonal anisotropy (Journel and Huijbregts, 1978, pg. 179-184; Ecker and Gelfand, 2003; Schabenberger and Gotway, 2004, pg. 152; Banerjee et al., 2014, pg. 31) and geometric anisotropy (Borgman and Chao, 1994; Ecker and Gelfand, 1999). Rapid growth of computing power has allowed the implementation and estimation of these models.

When modeling a spatial process, the specification of the covariance function will have an effect on kriging and parameter estimates and the associated uncertainty (Cressie, 1993, pg. 127-135). Sherman (2011, pg. 87-90) and Guan et al. (2004) use numerical examples to demonstrate the adverse implications of incorrectly specifying isotropy properties on kriging estimates. Given the variety of choices available regarding the properties of the covariance function (e.g., parametric forms, isotropy, stationarity) and the effect these choices can have on inference, practitioners may seek methods to inform the selection of an appropriate covariance model.

A number of graphical diagnostics have been proposed to evaluate isotropy properties in observed data. Perhaps the most commonly used methods are directional semivariograms and rose diagrams (Matheron, 1961; Isaaks and Srivastava, 1989, pg. 149-154). Banerjee



et al. (2014, pg. 38-40) suggest using an empirical semivariogram contour plot to assess isotropy as a more informative method than directional sample semivariograms. Another technique involves comparing empirical estimates of the covariance at different directional lags to assess symmetry for data on gridded sampling locations (Modjeska and Rawlings, 1983). One caveat of the aforementioned methods is that they can be challenging to assess, are open to subjective interpretations, and can be misleading (Guan et al., 2004) because they typically do not include a measure of uncertainty. Experienced statisticians may have intuition about the interpretation and reliability of these diagnostics, but a novice user may wish to evaluate assumptions via a hypothesis test.

Statistical hypothesis tests of second-order properties can be used to supplement and reinforce intuition about graphical diagnostics and can be more objective. Like the graphical techniques, hypothesis tests have their own caveats; for example, a parametric test of isotropy demands specification of the covariance function. A nonparametric method for testing isotropy avoids the potential problems of misspecification of the covariance function and the requirement of model estimation under both the null and alternative hypotheses, which can be computationally expensive for large datasets. Furthermore, nonparametric methods do not require the common assumption of geometric anisotropy. However, in abandoning the parametric assumptions about the covariance function, implementing a test of isotropy presents several challenges (see Section 2.5). A nonparametric test of isotropy or symmetry can serve as another form of exploratory data analysis that supplements graphical techniques and informs the choice of an appropriate nonparametric or parametric model. Figure 2.1 illustrates the process for assessing and modeling isotropy properties.

In this chapter we review nonparametric hypothesis tests developed to test the assumptions of symmetry and isotropy in spatial processes. We summarize tests in both the spatial and spectral domain and provide tables that enable convenient comparisons of test properties. A simulation study evaluates the empirical size and power of several of the methods and enables a direct comparison of method performance. We provide recommendations regarding

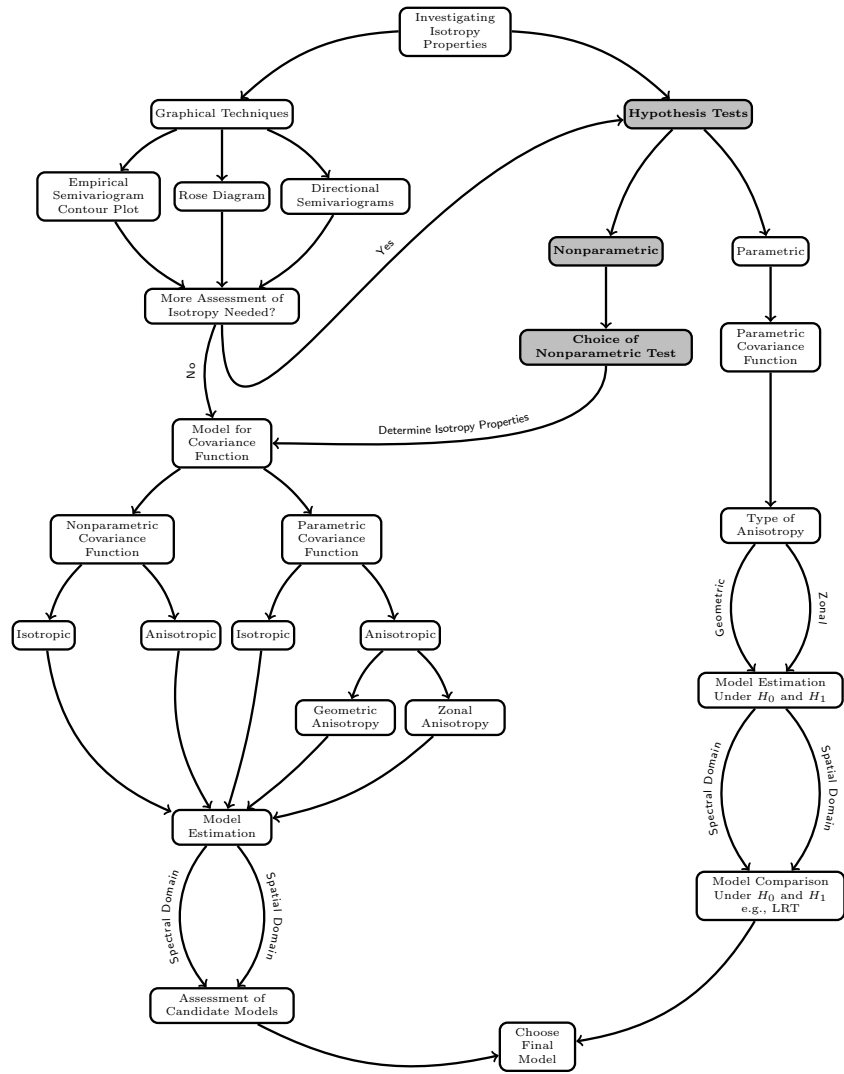


Figure 2.1: A flow chart illustrating the process of assessing and modeling isotropy in spatial data. The gray boxes indicate the focus of this chapter.

choices for test implementation and investigate the effects of these choices via simulations. Finally, we include graphics that demonstrate considerations for choosing a nonparametric test and illustrate the process of determining isotropy properties.

The remainder of this chapter is organized as follows: Section 2.2 establishes some additional definitions; Section 2.3 details the various nonparametric hypothesis tests of isotropy and symmetry and includes Tables 2.1-2.4 which facilitate comparison between tests as well as test selection for users; Section 2.4 describes the simulation study comparing the various methods and displays some of the results; Section 2.5 provides recommendations for test implementation; and Section 2.6 proposes new avenues of future research and concludes the chapter. Additional details on the simulation study and additional simulation results are available in the Section 2.7.

## 2.2 Additional Background and Definitions

The classical, moment-based estimator of the semivariogram (Matheron, 1962) is

$$\hat{\gamma}(\mathbf{h}) = \frac{1}{2|\mathcal{D}(\mathbf{h})|} \sum [Y(\mathbf{s}) - Y(\mathbf{s} + \mathbf{h})]^2, \quad (6)$$

where the sum is over  $\mathcal{D}(\mathbf{h}) = \{\mathbf{s} : \mathbf{s}, \mathbf{s} + \mathbf{h} \in \mathcal{D}\}$  and  $|\mathcal{D}(\mathbf{h})|$  is the number of elements in  $\mathcal{D}(\mathbf{h})$ . The set  $\mathcal{D}(\mathbf{h})$  is the set of sampling location pairs that are separated by spatial lag  $\mathbf{h}$ . This nonparametric estimator of second-order properties is used as the point estimator in several of the nonparametric tests of isotropy.

In addition to isotropy, symmetry is another directional property of the covariance (semivariogram) function, which is often used to describe the spatial variation of processes on a grid. We discuss symmetry properties here as they are a subset of isotropy, and methods for testing isotropy can often be used to test symmetry. The following definitions come from Lu and Zimmerman (2005) and Scaccia and Martin (2005) where the notation  $C(h_1, h_2)$  denotes

the covariance between random variables located  $h_1$  columns and  $h_2$  rows apart on the grid, denoted  $\mathcal{L}^2$ .

**Definition 2.** *A second-order stationary spatial process on a grid is reflection or axially symmetric if  $C(h_1, h_2) = C(-h_1, h_2)$  for all  $(h_1, h_2) \in \mathcal{L}^2$ .*

**Definition 3.** *A second-order stationary spatial process on a grid is diagonally or laterally symmetric if  $C(h_1, h_2) = C(h_2, h_1)$  for all  $(h_1, h_2) \in \mathcal{L}^2$ .*

**Definition 4.** *A second-order stationary spatial process on a grid is completely symmetric if it is both reflection and laterally symmetric (i.e.,  $C(h_1, h_2) = C(-h_1, h_2) = C(h_2, h_1) = C(-h_2, h_1)$  for all  $(h_1, h_2) \in \mathcal{L}^2$ ).*

Complete symmetry is a weaker property than isotropy. Isotropy requires that  $C(h_1, h_2)$  depends only on  $\sqrt{h_1^2 + h_2^2}$  for all  $h_1, h_2$ . The relationship between these properties is given by:

$$\text{isotropy} \implies \text{complete symmetry} \implies \begin{matrix} \text{reflection symmetry} \\ \text{diagonal symmetry} \end{matrix}. \quad (7)$$

Thus, rejecting a null hypothesis of reflection symmetry implies evidence against the assumptions of reflection symmetry, complete symmetry, and isotropy. However, failure to reject a null hypothesis of reflection symmetry does not imply an assumption of complete symmetry or isotropy is appropriate.

As we mentioned in Chapter 1, properties of the covariance function imply properties of the spectral density. Test statistics quantifying second-order properties can be constructed using the periodogram, an estimator of the spectral density (5) and denoted by  $I(\cdot)$ . For a real-valued spatial process observed on a rectangular grid  $\mathbb{Z}^2 \subset \mathbb{R}^2$ , a moment-based

periodogram used to estimate (5) is

$$I(\omega_1, \omega_2) = \frac{1}{(2\pi)^2} \sum_{h_1=-n_1+1}^{n_1-1} \sum_{h_2=-n_2+1}^{n_2-1} \hat{C}(h_1, h_2) \cos(h_1\omega_1 + h_2\omega_2), \quad (8)$$

where  $n_1$  and  $n_2$  denote the number of rows and columns of the grid and  $\hat{C}(h_1, h_2)$  is a non-parametric estimator of the covariance function. It is important to note that, depending on whether the underlying process is discrete or continuous, (8) is an estimator of different quantities. See Fuentes and Reich (2010) for more details on this distinction. In practice, the periodogram (8) is used to estimate the spectral density at the Fourier or harmonic frequencies. The frequency  $\boldsymbol{\omega} = (\omega_1, \omega_2)$  is a Fourier or harmonic frequency if  $\omega_j$  is a multiple of  $2\pi/n_j$ ,  $j = 1, 2$ . Furthermore, the set of frequencies is limited to  $\{\omega_j = 2\pi k_j/n_j, k_j = 0, \pm 1, \pm 2, \dots, \pm n_j^*\}$ , where  $n_j^*$  is  $(n_j - 1)/2$  if  $n_j$  is odd and  $n_j/2 - 1$  if  $n_j$  is even.

## 2.3 Tests of Isotropy and Symmetry

### 2.3.1 Brief History

Matheron (1961) developed one of the earliest hypothesis tests of isotropy when he used normality of sample variogram estimators to construct a  $\chi^2$  test for anisotropy in mineral deposit data. Cabana (1987) tested for geometric anisotropy using level curves of random fields. Vecchia (1988) and Baczkowski and Mardia (1990) developed tests for isotropy assuming a parametric covariance function. Baczkowski (1990) also proposed a randomization test for isotropy. Despite these early works, little work on testing isotropy was published during the 1990s, although the PhD dissertation work of Lu (1994) would eventually have a noteworthy impact on the literature. Then, in the 2000s, a number of nonparametric tests of second-order properties emerged. Some of the developments used estimates of the variogram or covariogram to test symmetry and isotropy properties (Lu and Zimmerman, 2001; Guan, 2003; Guan et al., 2004, 2007; Maity and Sherman, 2012). These works generally borrowed

ideas from two bodies of literature: (a) theory on the distributional and asymptotic properties of semivariogram estimators (e.g., Baczkowski and Mardia, 1987; Cressie, 1993, pg. 69-47; Hall and Patil, 1994) and (b) subsampling techniques to estimate the variance of statistics derived from spatial data (e.g., Possolo, 1991; Politis and Sherman, 2001; Sherman, 1996; Lahiri, 2003; Lahiri and Zhu, 2006). Other nonparametric methods used the spectral domain to test isotropy and symmetry (Scaccia and Martin, 2002, 2005; Lu and Zimmerman, 2005; Fuentes, 2005). These works generally extended ideas used in the time series literature (e.g., Priestley and Rao, 1969; Priestley, 1981) to the spatial case. Methods for testing isotropy and symmetry in both the spatial and spectral domains, under the assumption of a parametric covariance function, have also been developed recently (Stein et al., 2004; Haskard, 2007; Fuentes, 2007; Matsuda and Yajima, 2009; Scaccia and Martin, 2011).

### **2.3.2 Nonparametric Methods in the Spatial Domain**

A popular approach to testing second-order properties was pioneered in the works of Lu (1994) and Lu and Zimmerman (2001) who leveraged the joint asymptotic normality of the sample variogram computed at different spatial lags. The subsequent works of Guan et al. (2004, 2007) and Maity and Sherman (2012) built upon these ideas and are the primary focus of this subsection. Lu (1994) details methods for testing axial symmetry. While Lu and Zimmerman (2001), Guan et al. (2004), and Maity and Sherman (2012) focus on testing isotropy, these methods can also be used to test symmetry. Finally, Bowman and Crujeiras (2013) detail a more computational approach for testing isotropy. Both Li et al. (2007, 2008b) and Jun and Genton (2012) use an approach analogous to the methods from Lu and Zimmerman (2001), Guan et al. (2004, 2007), and Maity and Sherman (2012) but for spatiotemporal data. Table 2.1 summarizes test properties discussed in this section and Section 2.3.3.

Nonparametric tests for anisotropy in the spatial domain are based on a null hypothesis that is an approximation to isotropy. Under the null hypothesis that the RF is isotropic,

it follows that the values of the semivariogram  $\gamma(\cdot)$  evaluated at any two spatial lags that have the same norm are equal, regardless of the direction of the lags. To fully specify the most general null hypothesis of isotropy, theoretically, one would need to compare variogram values for an infinite set of lags. In practice a small number of lags are specified. Then it is possible to test a hypothesis consisting of a set of linear contrasts of the form

$$H_0 : \mathbf{A}\gamma(\cdot) = \mathbf{0} \quad (9)$$

as a proxy for the null hypothesis of isotropy, where  $\mathbf{A}$  is a full row rank matrix (Lu and Zimmerman, 2001). For example, a set of lags, denoted  $\mathbf{\Lambda}$ , commonly used in practice for gridded sampling locations with unit spacing is

$$\mathbf{\Lambda} = \{\mathbf{h}_1 = (1, 0), \mathbf{h}_2 = (0, 1), \mathbf{h}_3 = (1, 1), \mathbf{h}_4 = (-1, 1)\}, \quad (10)$$

and the corresponding  $\mathbf{A}$  matrix under  $H_0 : \mathbf{A}\gamma(\mathbf{\Lambda}) = \mathbf{0}$  is

$$\mathbf{A} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}. \quad (11)$$

One of the first steps in detecting potential anisotropy is the choice of lags, as the test results will only hold for the particular set of lags considered (Guan et al., 2004). While this choice is subjective, there are several considerations and useful guidelines for determining the set of lags (see Section 2.5).

For nonparametric tests of symmetry, the null hypothesis of symmetry using (9) can be expressed by a countable set of contrasts for a process observed on a grid. Tests of symmetry will be subject to similar practical considerations as tests of isotropy, and practitioners testing symmetry properties will need to choose a small set of lags and form a hypothesis that is a surrogate for symmetry. For example, testing reflection symmetry of a process observed

on the integer grid would require comparing estimates of  $C(\cdot)$  evaluated at the lag pairs  $\{(1, 0), (-1, 0)\}, \{(2, 0), (-2, 0)\}, \{(1, 1), (-1, 1)\}$ , etc.

The tests in Lu and Zimmerman (2001), Guan et al. (2004, 2007), and Maity and Sherman (2012) involve estimating either the semivariogram  $\gamma(\cdot)$  or covariogram  $C(\cdot)$  at the set of chosen lags  $\mathbf{A}$ . Denoting the set of point estimates of the semivariogram/covariance function at the chosen lags as  $\hat{\mathbf{G}}_n$ , the true values as  $\mathbf{G}$ , and normalizing constant  $a_n$ , a central result for all three methods is that

$$a_n(\hat{\mathbf{G}}_n - \mathbf{G}) \xrightarrow[n \rightarrow \infty]{d} \text{MVN}(\mathbf{0}, \mathbf{\Sigma}), \quad (12)$$

under increasing domain asymptotics and mild moment and mixing conditions on the RF. Using the  $\mathbf{A}$  matrix, an estimate of the variance covariance matrix,  $\hat{\mathbf{\Sigma}}$ , and  $\hat{\mathbf{G}}_n$ , a quadratic form is constructed, and a p-value can be obtained from an asymptotic  $\chi^2$  distribution with degrees of freedom given by the row rank of  $\mathbf{A}$ . The primary differences between these works are the assumed distribution of sampling locations, the shape of the sampling domain, and the estimation of  $\mathbf{G}$  and  $\mathbf{\Sigma}$ . These differences are important when choosing a test that is appropriate for a particular set of data (see Tables 2.1 and 2.2 and Figure 2.4 for more information about these differences).

Maity and Sherman (2012) developed a test with the fewest restrictions on the shape of the sampling region and distribution of sampling locations. Their test can be used when the sampling region is any convex subset in  $\mathbb{R}^d$  and the distribution of sampling locations in the region follows any general spatial sampling design. The test in Guan et al. (2004) also allows convex subsets in  $\mathbb{R}^d$  and is developed for gridded and non-gridded sampling locations but requires non-gridded sampling locations to be uniformly distributed on the domain (i.e., generated by a homogenous Poisson process). The Poisson assumption is dropped in Guan et al. (2007). Lu and Zimmerman (2001) require the domain to be rectangular and the observations to lie on a grid.



Table 2.1: Properties of nonparametric tests of isotropy. “Sym” indicates whether or not the test can be used to test symmetry properties. “Domain” refers to the domain used to represent the RF (spatial or spectral), “Test Stat Based On” lists the nonparametric estimator used to construct the test statistic “Dist’n” gives the limiting asymptotic distribution of the test statistic, and “GP” denotes whether the test requires data to be generated from a Gaussian process.

	<b>Hypothesis Test Properties</b>						
<b>Test Method</b>	Isotropy	Sym	Domain	Test Stat Based On	Asymptotics	Dist’n	GP
Lu and Zimmerman (2001)	yes	yes	spatial	semivariogram	inc domain	$\chi^2$	yes
Guan et al. (2004, 2007)	yes	yes	spatial	(kernel) <sup>a</sup> variogram	inc domain	$\chi^2$ <sup>b</sup>	no
Scaccia and Martin (2002, 2005)	partial	yes	spectral	periodogram	inc domain	$Z, t$	no
Lu and Zimmerman (2005)	partial	yes	spectral	periodogram	inc domain	$\chi^2, F$	no
Fuentes (2005)	partial	no	spectral	spatial periodogram	mixed	$\chi^2$	yes
Maity and Sherman (2012)	yes	yes	spatial	kernel covariogram	inc domain	$\chi^2$	no
Bowman and Crujeiras (2013)	yes	no	spatial	variogram	inc domain	$\chi^2$ <sup>c</sup>	yes
Van Hala et al. (2014)	yes	yes	spectral	empirical likelihood	mixed	$\chi^2$	no

<sup>a</sup>for gridded sampling locations, the estimator in (6) is used while a kernel semivariogram estimator is used for non-gridded sampling locations

<sup>b</sup>p-values may need to be approximated using finite sample adjustments

<sup>c</sup> distribution is approximately  $\chi^2$

Table 2.2: Test implementation, part 1. “Subsamp” defines whether spatial subsampling procedures are needed to perform the test, “S&P Sim” denotes whether or not the author(s) of the method provide a simulation of test size and power (See also Table 2.4).

	<b>Hypothesis Test Implementation</b>				
<b>Test Method</b>	Sampling Domain Shape	Sampling Design	Subsamp	S&P Sim	Software
Lu and Zimmerman (2001)	rectangular in $\mathbb{R}^2$	grid	no	yes <sup>a</sup>	no
Guan et al. (2004, 2007)	convex subsets in $\mathbb{R}^d$	grid/unif <sup>b</sup> /non-unif <sup>c</sup>	yes	yes <sup>a</sup>	R pkg <b>spTest</b>
Scaccia and Martin (2002, 2005)	rectangular in $\mathbb{R}^2$	grid	no	yes <sup>a</sup>	no
Lu and Zimmerman (2005)	rectangular in $\mathbb{R}^2$	grid	no	yes	R pkg <b>spTest</b>
Fuentes (2005)	rectangular in $\mathbb{R}^d$	grid	no	yes <sup>a</sup>	no
Maity and Sherman (2012)	convex subsets in $\mathbb{R}^d$	non-unif <sup>c</sup>	yes	yes <sup>a</sup>	R pkg <b>spTest</b>
Bowman and Crujeiras (2013)	convex subsets in $\mathbb{R}^d$	unif <sup>b</sup>	no	yes <sup>a</sup>	R pkg <b>sm</b>
Van Hala et al. (2014)	subsets in $\mathbb{R}^d$	non-unif <sup>c</sup>	no	yes <sup>a</sup>	no

<sup>a</sup> simulated data are Gaussian only

<sup>b</sup> sampling locations must be generated by homogeneous Poisson process, i.e. uniformly distributed on the domain

<sup>c</sup> sampling locations can be generated by any general sampling design

Another difference between methods is the form of the nonparametric estimator of  $\mathbf{G}$ . In Lu and Zimmerman (2001),  $\hat{\mathbf{G}}_n$  is computed using the log of the classical sample semi-variogram estimator (6). Guan et al. (2004, 2007) also use the estimator in (6) for gridded sampling locations, but use a kernel estimator of  $\gamma(\mathbf{h})$  for non-gridded locations. Maity and Sherman (2012) use a kernel estimator of the covariance function. When smoothing over spatial lags in  $\mathbb{R}^2$ , the Nadaraya-Watson (Nadaraya, 1964; Watson, 1964) product kernel estimator is typically adopted, independently smoothing over horizontal and vertical lags. Common choices for the kernel are the Epanechnikov or truncated Gaussian kernels. The kernel estimators also require the selection of a bandwidth parameter,  $w$ . Choosing an appropriate bandwidth is one of the challenges of implementing the tests for non-gridded sampling locations, and the conclusion of the test has the potential to be sensitive to the choice of the bandwidth parameter (see Section 2.5 for recommendations on bandwidth selection).

Nonparametric tests in the spatial domain also vary in the estimation of  $\Sigma$ , the asymptotic variance-covariance of  $\hat{\mathbf{G}}_n$  in (12). Lu and Zimmerman (2001) use a plug-in estimator that requires the choice of a parameter,  $m$ , that truncates the sum used in estimation. Spatial resampling methods are another approach used to estimate  $\Sigma$ . The method used for spatial resampling and properties of estimators computed from spatial resampling will depend on the underlying spatial sampling design (Lahiri, 2003, pg. 281). Guan et al. (2004, 2007) use a moving window approach, creating overlapping subblocks that cover the sampling region. Maity and Sherman (2012) employ the grid based block bootstrap (GBBB) (Lahiri and Zhu, 2006). The GBBB approach divides the spatial domain into regions, then replaces each region by sampling (with replacement) a block of the sampling domain having the same shape and volume as the region, creating a spatial permutation of blocks of sampling locations. When using the resampling methods, the user must choose the window or block size and the conclusion of the test has the potential to change based on the chosen size. Irregularly shaped sampling domains can pose a challenge in using the subsampling methods. For example, for an irregularly shaped sampling domain, many incomplete blocks may complicate the

subsampling procedure. We summarize guidelines for choosing the window/block size in Section 2.5.

Another approach to testing isotropy in the spatial domain is given by Bowman and Crujeiras (2013) who take a more empirical and computationally-intensive approach. Their methods are available in the R software (R Core Team, 2015) package `sm` (Bowman and Azzalini, 2014). One caveat of using the `sm` package is that the methods are computationally expensive, even for moderate sample sizes. For example, a test of isotropy on 300 uniformly distributed sampling locations on a  $10 \times 16$  sampling domain took approximately 9.5 minutes where the methods from Guan et al. (2004) took 1.6 seconds using a laptop with 8 GB of memory and a 2 GHz Intel Core i7 processor. Because of the computational costs, we do not consider this method further.

### 2.3.3 Nonparametric Methods in the Spectral Domain

For gridded sampling locations, nonparametric spectral methods have been developed for testing symmetry (Scaccia and Martin, 2002, 2005; Lu and Zimmerman, 2005) and stationarity (Fuentes, 2005), but none are designed with a primary goal of testing isotropy. This may, in part, be due to the aliasing that occurs when estimating the spectral density of a continuous process observed on a discrete set of locations (Fuentes and Reich, 2010). Additionally, due to the difficulties presented by non-gridded sampling locations, historically there have been fewer developments using spectral methods for non-gridded sampling locations than for gridded (or lattice) data, but this is an area that has received more attention recently (see, e.g., Fuentes, 2007; Matsuda and Yajima, 2009; Bandyopadhyay et al., 2015). Despite the challenges, Van Hala et al. (2014) have proposed a nonparametric, empirical likelihood approach to test isotropy and separability for non-gridded sampling locations.

The primary motivation for using the spectral domain over the spatial domain are simpler asymptotics in the spectral domain. Unlike estimates of the variogram or covariogram at different spatial lags, estimates of the spectral density at different frequencies via the

periodogram are asymptotically independent under certain conditions (Pagano, 1971; Schabenberger and Gotway, 2004, pg. 78,194). Additionally, in practice, tests of symmetry in the spectral domain are generally not subject to as many choices (e.g., spatial lag set, bandwidth, block size) as those in the spatial domain.

Analogous to testing isotropy in the spatial domain by using a finite set of spatial lags, tests of symmetry in the spectral domain typically involve estimating and comparing the spectral density (5) at a finite set of the Fourier frequencies. For example, axial symmetry (2) of the covariance function implies axial symmetry of the spectral density,  $f(\omega_1, \omega_2) = f(-\omega_1, \omega_2)$ , which can be evaluated by comparing  $I(\omega_1, \omega_2)$  to  $I(-\omega_1, \omega_2)$  at a finite set of frequencies. Similarly, the null hypothesis of isotropy can be approximated by comparing periodogram estimates (8) at a set of distinct frequencies with the same norm (Fuentes, 2005). Although most of the current spectral methods are not directly designed to test isotropy, the hypothesis of complete symmetry can be used to reject the assumption of isotropy due to (7). However, certain types of anisotropy may not be detected by these tests. For example, a geometrically anisotropic process having the major axis of the ellipses of equicorrelation parallel to the  $x$ -axis is axially symmetric, and the anisotropy wouldn't be detected by a test of axial symmetry.

Scaccia and Martin (2002, 2005) use the periodogram (8) to develop a test for axial symmetry. They propose three test statistics that are a function of the periodogram values. The first test statistic uses the average of the difference in the log of the periodogram values,  $\log I(\omega_1, \omega_2) - \log I(\omega_1, -\omega_2)$ . The second and third test statistics use the average of standardized periodogram differences,  $[I(\omega_1, \omega_2) - I(\omega_1, -\omega_2)]/[I(\omega_1, \omega_2) + I(\omega_1, -\omega_2)]$ . These test statistics are shown to asymptotically follow a standard normal or  $t$  distribution via the Central Limit Theorem, and the corresponding distributions are used to obtain a p-value.

Lu and Zimmerman (2005) also use the periodogram as an estimator of the spectral density to test properties of axial and complete symmetry of processes on the integer grid,  $\mathbb{Z}^2$ . They use the asymptotic distribution of the periodogram to construct two potential test

statistics. Both test statistics leverage the fact that, under certain conditions and at certain frequencies,

$$\frac{2I(\omega_1, \omega_2)}{f(\omega_1, \omega_2)} \xrightarrow[n_1, n_2 \rightarrow \infty]{\text{iid}} \chi_2^2, \quad (13)$$

where  $f()$  is the spectral density of the (non-continuous) process on the grid. Under the null hypothesis of axial or complete symmetry, (13) implies that ratios of periodogram values at different frequencies follow an  $F(2, 2)$  distribution. The preferred test statistic produces a p-value via a Cramér-von Mises (CvM) goodness of fit test using the appropriate set of periodogram ratios. Because rejecting a hypothesis of axial symmetry implies rejecting a hypothesis of complete symmetry, Lu and Zimmerman (2005) recommend a two-stage procedure for testing complete symmetry. At the first stage, they test the hypothesis of axial symmetry, and if the null hypothesis is not rejected, they test the hypothesis of complete symmetry. To control the overall type-I error rate at  $\alpha$ , the tests at each stage can be performed using a significance level of  $\alpha/2$ .

Leveraging the asymptotic independence of the periodogram at different frequencies, Van Hala et al. (2014) propose a spatial frequency domain empirical likelihood (SFDEL) approach that can be used for inference about spatial covariance structure. One of the applications of this method is testing isotropy. An advantage of this method over other frequency domain approaches is that it can be used for non-gridded sampling locations. To implement the test, the user must select the set of lags and, because the sampling locations are not gridded, the number and spacing of frequencies. Van Hala et al. (2014) offer some guidelines for these choices based on the simulations and theoretical considerations (e.g., the frequencies need to be asymptotically distant). After these choices are made, Van Hala et al. (2014) maximize an empirical likelihood under a moment constraint corresponding to isotropy and show that the log-ratio of the constrained and unconstrained maximizer asymptotically follows a  $\chi^2$  distribution. The SFDEL method relies on the asymptotic independence of the

periodogram values, and the smallest sample size used in simulations was  $n = 600$ . Thus, it is not clear how the method will perform for smaller sample sizes.

Fuentes (2005) introduces a nonparametric, spatially varying spectral density to represent nonstationary spatial processes. While the method can be used to test the assumption of isotropy, the test requires a large sample size on a fine grid. For this reason, and also because the test was primarily designed to test the assumption of stationarity, we do not consider it further.

## 2.4 Simulation Study

We designed a simulation study to compare the empirical size, power, and computational costs for four of the methods. For gridded sampling locations, we compare Lu and Zimmerman (2005)[hereafter, LZ] to Guan et al. (2004)[hereafter denoted as GSC or GSC-g when we are specifically referring to the test when applied to gridded sampling locations]. For uniformly distributed sampling locations we compare Maity and Sherman (2012)[MS] to Guan et al. (2004, 2007)[GSC-u for the method used for uniformly distributed sampling locations].

We performed the tests on the same realizations of the RF. Data are simulated on rectangular grids or rectangular sampling domains because they are more realistic than square domains and simulations on rectangular domains were not previously demonstrated. We simulate Gaussian RFs with mean zero and exponential covariance functions with no nugget, a sill equal to one, and effective range values corresponding to short, medium, and long range dependence. The effective range is the range at which spatial correlation equals 0.05. We introduce varying degrees of geometric anisotropy via coordinate transformations governed by a rotation parameter  $\theta$  and scaling parameter  $R$  that define the ellipses of equicorrelation (see Figure 2.5 in Section 2.7). The parameter  $\theta$  quantifies the angle between the major axis of the ellipse and the  $x$ -axis (counter-clockwise rotation) while  $R$  gives the ratio of the major and minor axes of the ellipse. We also performed simulations that investigate the effect of the lag set, block size, and bandwidth. Although some simulations are given in the original

Table 2.3: Test implementation, part 2. This table continues the list of choices and considerations for implementing a given test. “Samp Size (S/A)” indicates the minimum sample sizes used in simulations (S) and applications (A) provided by the author(s) of the method.

	<b>Hypothesis Test Implementation</b>		
<b>Test Method</b>	Choices	Other Considerations	Samp Size (S/A)
Lu and Zimmerman (2001)	spatial lag set, truncation parameter	optimal truncation parameter	100/112
Guan et al. (2004) gridded design	spatial lag set, window size	optimal window size, edge effects, finite sample adjustment	400/289
Guan et al. (2004) uniform design Guan et al. (2007) non-uniform design	spatial lag set, kernel function, bandwidth parameter, window size	optimal bandwidth & window size, edge effects, finite sample adjustment	400/289 500/584
Scaccia and Martin (2002, 2005)	test statistic	requires gridded sampling locations; designed to test symmetry	121/–



Lu and Zimmerman (2005)	test statistic	requires gridded sampling locations; two-stage testing procedure, designed to test symmetry; relies on asymptotic independence	100/–
Fuentes (2005)	kernel function, bandwidth parameters, frequency set, spatial knots	requires fine grid; designed to test stationarity	5175/5175
Maity and Sherman (2012)	lag set $\mathbf{\Lambda}$ , kernel function, bandwidth parameter, subblock size, number of bootstrap samples	optimal bandwidth & block size	350/584
Bowman and Crujeiras (2013)	bandwidth parameter	computationally intensive	49/148

Van Hala et al. (2014)	lag set, number and spacing of frequencies	optimal number and spacing of frequencies, relies on asymptotic independence	600/–
------------------------	---	---	-------

works, our simulations serve to provide a direct comparison of the effects of changing these values and provide further insight into how to choose them. See Section 2.7 for additional simulation details and results.

Figures 2.2 and 2.3 illustrate a subset of the simulation results comparing empirical size, power, and computational time (full results in Section 2.7, Tables 2.5 and 2.6). These simulations indicate that nonparametric tests for anisotropy have higher power for gridded (Table 2.5 and Figure 2.2) than for non-gridded (Table 2.6 and Figure 2.3) sampling designs. In both comparisons the methods from GSC have favorable empirical power over the competitor with a comparable empirical size. As the effective range increases, both empirical size and power tend to increase for the methods from GSC, but they tend to decrease for MS. GSC-g and LZ have similar computation time, while MS is much more computationally expensive than GSC-u. This difference is due to the bootstrapping required by MS.

Unsurprisingly, as the strength of anisotropy increases (measured by  $R$ ), power increases for all the methods. For a geometrically anisotropic process, the major and minor axes of anisotropy are orthogonal. In comparing the effect of the orientation of isotropy ( $\theta$ ) on the methods, it is important to note that, when  $\theta = 0$ , the major axis of the ellipse defining the geometric anisotropy is parallel to the  $x$ -axis and corresponds to a spatial process that is axially symmetric but not completely symmetric. When  $\theta = 3\pi/8$  the major axis of the ellipse forms a 67.5-degree angle with the  $x$ -axis, and the spatial process is neither axially nor completely symmetric (see Figure 2.5 in Section 2.7 for contours of equal correlation used in the simulation). The original works generally only simulate data from a geometrically anisotropic process with the major axis of anisotropy forming a 45-degree angle with the  $x$ -axis; hence, our simulation study more carefully explores the effect of changing the orientation of geometric anisotropy. The methods from GSC exhibit higher power when  $\theta = 0$  than when  $\theta = 3\pi/8$ . This is due to the fact that the lag set,  $\mathbf{\Lambda}$ , from (10) used for the tests contains a pair of spatial lags that are parallel to the major and minor axes of anisotropy when  $\theta = 0$ , indicating that an informed choice of spatial lags improves the test's ability to detect

anisotropy. This same result does not hold for MS. It is unclear whether this behavior is observed because the method uses the covariogram rather than the semivariogram, the GBBB rather than moving window approach for estimating  $\Sigma$ , or perhaps both. The simulation results indicate that the LZ test has low empirical power; however, this method was developed to test symmetry properties on square grids, and the choice of a rectangular grid for our simulation study does not allow for a large number of periodogram ordinates for the second stage of the procedure for testing the complete symmetry hypothesis.

Results from simulations that investigate the effects of changing the lag set, the block size, and the bandwidth for non-gridded sampling locations are displayed in Tables 2.7-2.9, respectively, in Section 2.7. For both GSC-u and MS, the lag set in (10) provided an empirical size close to the nominal level. Using more lags or longer lags decreased the size and power for GSC-u. This may be due to the additional uncertainty induced by estimating the covariance between the semivariance at more lags and the larger variance of semivariance estimates at longer lags. For MS the longer lags lead to an inflated size and more lags decreased the power. In this case, the GBBB may not be capturing the uncertainty in covariance estimates at longer lags with the chosen block size. The MS test was not overly sensitive to block size with larger blocks leading to slightly higher power. MS found that an overly large block size was adverse for test size. For GSC-u the small and normal sized windows performed at nominal size levels with comparable power while larger windows were detrimental to test size and power. For GSC-u, we find that choosing a large window tends to lead to overestimation of the asymptotic variance-covariance matrix due to fewer blocks being used to re-estimate the semivariance. Finally, the results investigating the bandwidth selection for GSC-u indicate that choosing an overly large bandwidth inflates test size while choosing too small a bandwidth deflates test size and power. However, the results also indicate that, for the small sample size, test size and power are less negatively affected when approximating the p-value via the finite sample adjustment.

Weller (2016c) demonstrates applications of several of these methods on two real data sets. The R package `spTest` (Weller, 2016d) implements the tests in LZ, GSC, and MS for rectangular grids and sampling regions and is available on the Comprehensive R Archive Network (CRAN). The R scripts for reproducing simulation results are available online: <https://sites.google.com/site/zachdweller/research>.

## 2.5 Recommendations

Based on the simulation results, we offer recommendations for implementation of nonparametric tests of isotropy. The flow chart in Figure 2.1 along with Figure 2.4 summarize the steps in the process. Tables 2.1-2.4 compare the tests. Table 2.4 summarizes the recommendations provided below.

In choosing a nonparametric test for isotropy, the distribution of sampling locations on the sampling domain is perhaps the most important consideration. Data on a grid simplifies estimation because the semivariogram or covariogram can be estimated at spatial lags that are exactly observed separating pairs of sampling locations. A grid also allows the option of using easily implemented tests in the spectral domain.

Sample size requirements for the asymptotic properties of tests using the spatial domain to approximately hold will depend on the dependence structure of the random field. GSC note that convergence of their test statistic is slow in the case of gridded sampling locations and obtain an approximate p-value via subsampling rather than the asymptotic  $\chi^2$  distribution. Tests using the spectral domain rely on the asymptotic independence of periodogram values, and correlation in finite samples can lead to an inflated test size (LZ). Based on our simulations, we recommend the sample size be at least 150 for gridded sampling locations and at least 300 for non-gridded sampling locations. However, power tends to be low when the sample size is small and/or the anisotropy is weak (Figures 2.2 and 2.3).

We focus on implementation of the methods that use the spatial domain for the remainder of this section. We discuss the choice of lags, block size, and bandwidth for the tests in GSC

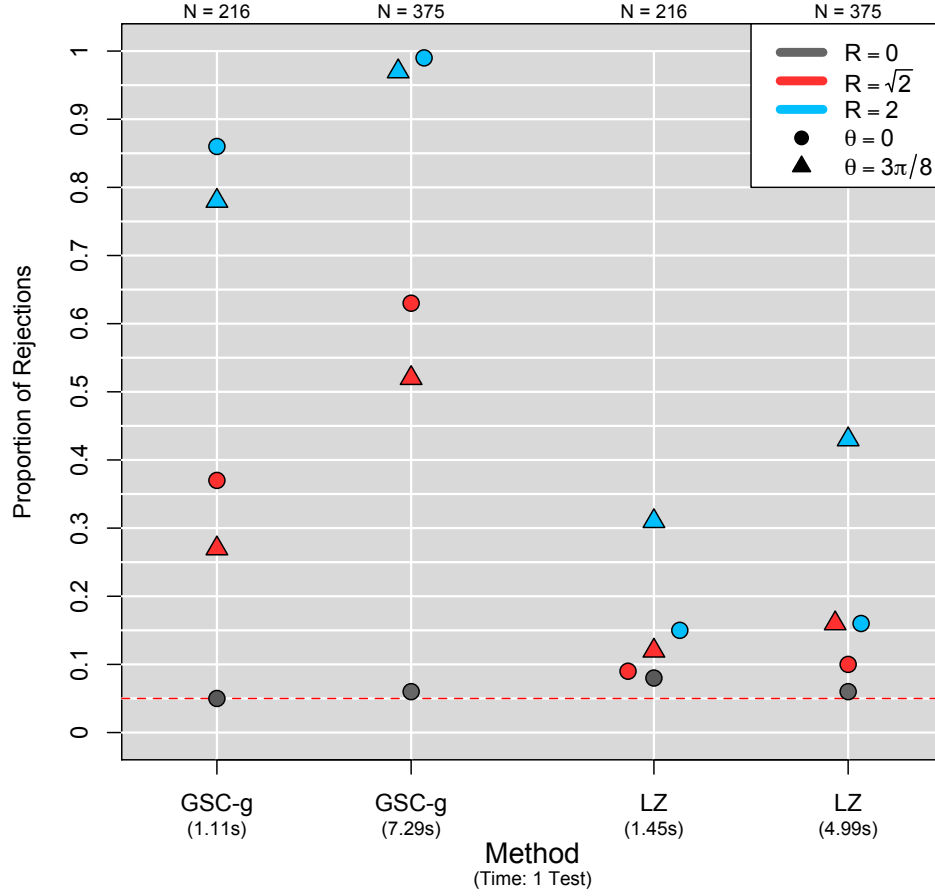


Figure 2.2: Empirical size and power for Guan et al. (2004) [GSC-g] and Lu and Zimmerman (2005) [LZ] for 1000 realizations of a mean 0 GRF with gridded sampling locations using a nominal level of  $\alpha = 0.05$ . Colors and shapes indicate the type of anisotropy. Gray points correspond to the isotropic case. The results correspond to a “medium” effective range. Computational time for each method is also displayed.

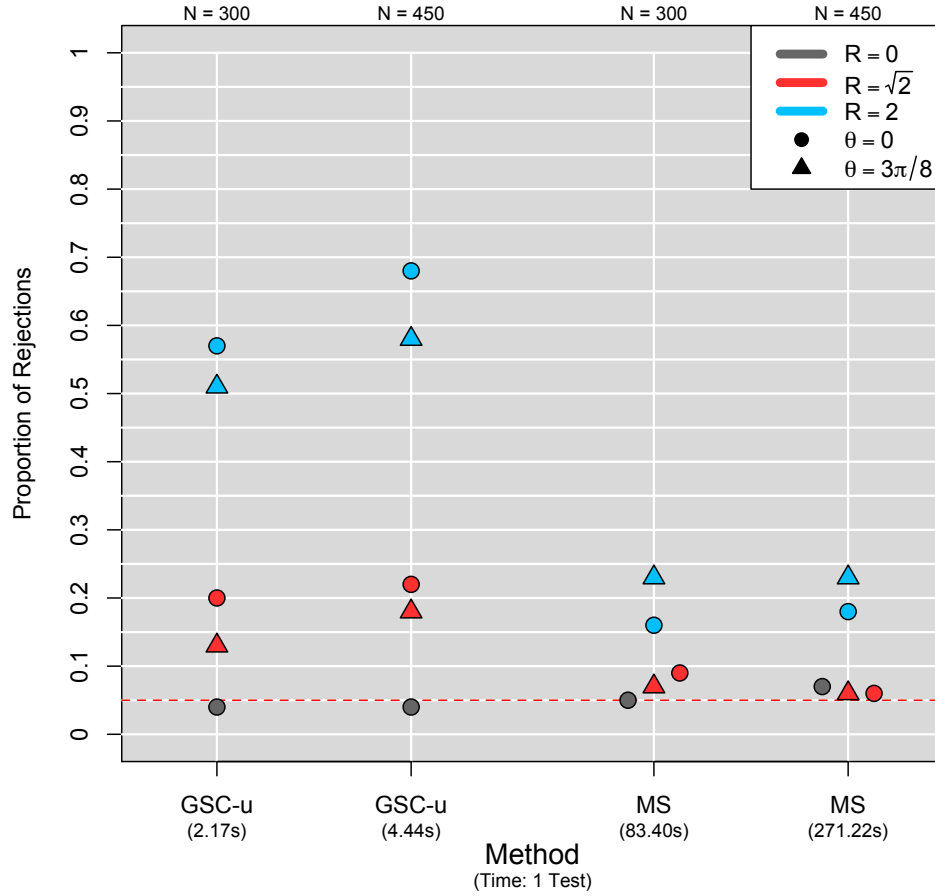


Figure 2.3: Empirical size and power for Guan et al. (2004) [denoted GSC-u] and Maity and Sherman (2012) [denoted MS] for 1000 realizations of a mean 0 GRF with uniformly distributed sampling locations using a nominal level of  $\alpha = 0.05$ . Colors and shapes indicate the type of anisotropy. Gray points correspond to the isotropic case. The results correspond to a “medium” effective range. Computational time for each method is also displayed.

Table 2.4: General Recommendations for Test Implementation. This table contains a list of general recommendations for test implementation. These guidelines will not apply in all situations and will vary based on a variety of factors including, but not limited to, the sample size, density of sampling locations, and scale of the problem. See additional discussion in Section 2.5.

Test Method	Hypothesis Test Choices				
	Lag Set <sup>a</sup>	Block Size	Bandwidth	P-value	min. $n$
Guan et al. (2004) gridded design	Length: short  Orientation: Eqn (10)	$n_b < n^{1/2}$	n/a	finite sample adjustment	150
Guan et al. (2004, 2007) uniform design		$n_b \lesssim n^{1/2}$	$0.6 < w < 0.9^b$	finite sample adjustment when $n < 500$ , asymptotic $\chi^2$ when $n \geq 500$	300
Maity and Sherman (2012)	Number: 4 (2 pairs)	$n_b \gtrsim n^{1/2}$	empirical	asymptotic $\chi^2$	300

<sup>a</sup> Prior knowledge, if available, should be used to inform the choice of lags.

<sup>b</sup> Our simulations suggest these bandwidth values are reasonable when using a Gaussian kernel with truncation parameter of 1.5.



and MS. Due to the large number of choices required to implement the tests (e.g., block size, bandwidth, kernel function, subsampling method), features of the random field (e.g., sill, range), and properties of the sampling design (e.g., density of sampling locations, shape of sampling domain), the recommendations we offer will not apply in all situations. The numerous moving parts make it challenging to develop general recommendations, especially when choosing a bandwidth.

When determining the lag set,  $\mathbf{\Lambda}$ , for use in (9), the user needs to select

- (a) the norm of the lags (e.g., Euclidean distance),
- (b) the orientation (direction) of the lags, and
- (c) the number of lags.

Regarding (a), short lags are preferred because estimates of the spatial dependence at large lags tend to be more variable than estimates at shorter lags. Additionally, empirical and theoretical evidence (Lu and Zimmerman, 2001) indicates that values of  $\gamma(\cdot)$  in two different directions generally exhibit the largest difference at a lag less than the effective range, the distance beyond which pairs of observations can be assumed to be independent. Finally, there is mathematical support that correctly specifying the covariance function at short lags is important for spatial prediction (Stein, 1988). Considering (b), if the process is anisotropic, the ideal choice of  $\mathbf{\Lambda}$  and  $\mathbf{A}$  produces contrasting lags with the same norm but oriented in the direction of weakest and strongest spatial correlation. Typically, the directions of weakest and strongest spatial correlation will be orthogonal and thus, lags contrasted using the  $\mathbf{A}$  matrix should also be orthogonal. Prior information, if available, about the underlying physical/biological process giving rise to the data can also be used to inform the orientation of the lags (Guan et al., 2004). If no prior information about potential anisotropy is available, lags oriented in the same directions as those in (10) are a good starting set. In regards to (c), detecting certain types of anisotropy requires a sufficient number of lags but using a large

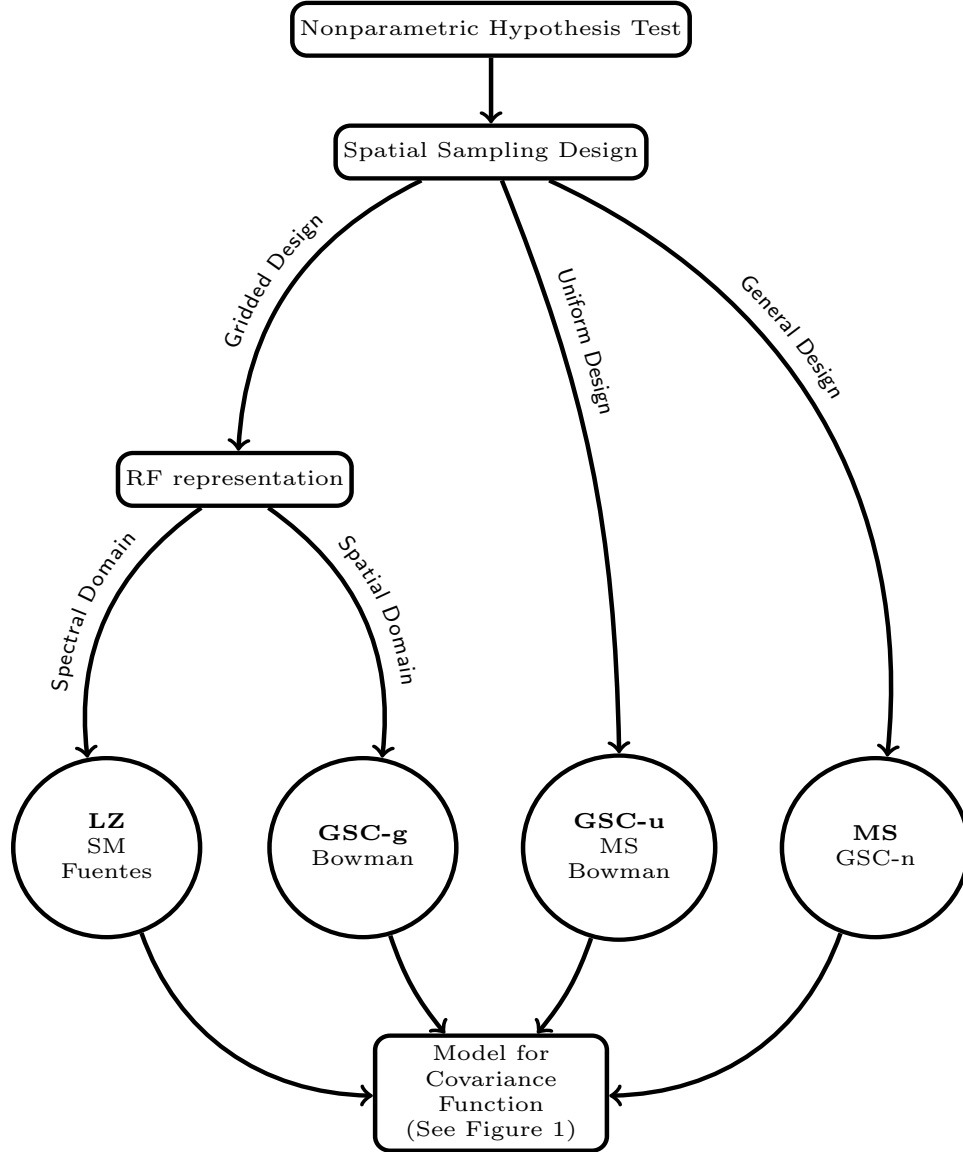


Figure 2.4: Spatial sampling design considerations for choosing a nonparametric hypothesis test of isotropy, including LZ = Lu and Zimmerman (2005); SM = Scaccia and Martin (2005); GSC-g = Guan et al. (2004) for gridded sampling locations; GSC-u = Guan et al. (2004) for uniformly distributed sampling locations; MS = Maity and Sherman (2012); GSC-n = Guan et al. (2007) for non-uniform sampling locations. The method we recommended for testing isotropy in each situation is given in bold.

number of lags requires a large number of observations (Guan et al., 2004). As a general guideline, we suggest using four lags to construct two contrasts.

Several tests require selection of a window or block size to estimate the variance-covariance matrix. The moving window from GSC creates overlapping subblocks of data by sliding the window over a grid placed on the region. Each of these subblocks are used to re-estimate the semivariance. The block size from MS defines the size of resampled blocks when implementing the GBBB. The GBBB permutes resampled blocks to create a new realization of the process over the entire domain. Choosing the window size in GSC requires balancing two competing goals. First, the moving window should be large enough to create subblocks that are representative of the dependence structure for the entire RF. Second, the window should be small enough to allow for a sufficient number of subblocks to re-estimate the semivariance, as these values are used to obtain an estimate of the asymptotic variance-covariance. A window that is too large or too small can potentially lead to under- or over-estimation of the asymptotic variance-covariance. For GSC-u, the windows must be large enough to ensure enough pairs of sampling locations are in each subblock to compute an estimate of the semivariance without having to over-smooth. For gridded sampling locations, GSC demonstrate good empirical size and power by using moving windows with size of only  $2 \times 2$ . However, they find slow convergence to the asymptotic  $\chi^2$  distribution, and a p-value is instead computed by approximating the distribution of the test statistic by computing its value for each of the subblocks. Hence, approximating the p-value to two decimal places will require at least 100 subblocks over the sampling region. This may not be possible in practice. For example, a  $12 \times 12$  grid of sampling locations with moving windows of size  $2 \times 2$  results in only 90 subblocks when correcting for edge effects. The challenge of choosing the block size in MS is subject to similar considerations as the window size in GSC. The p-value for both tests will change when performing the test with different window or block sizes, and the user may decide to perform the test with different block sizes (e.g., MS). There are a number of works on resampling spatial data to obtain an estimate of the variance of a spatial statistic

(e.g., Sherman, 1996; Politis and Sherman, 2001; Lahiri, 2003; Lahiri and Zhu, 2006), but they do not directly consider variance estimation in the case of a nonparametric estimate of the semivariogram/covariogram. Denoting the number of points per block as  $n_b$ , Sherman (1996) proposes choosing the block size such that  $n_b \approx cn^{1/2}$  for a constant,  $c$ , when the spatial dependence does not exhibit a large range. In a number of different applications of spatial subsampling,  $c$  is typically chosen to be between 0.5 and 2 (Politis and Sherman, 2001; Guan et al., 2004, 2006). Based on our simulations, we find acceptable empirical size and power for GSC-g using small windows and approximating the p-value with the finite sample adjustment. Thus, we recommend setting  $n_b < n^{1/2}$  for GSC-g. For example, we used windows with size  $3 \times 2$  and  $5 \times 3$  for sampling domains of  $18 \times 12$  and  $25 \times 15$ , respectively. In the case of uniformly distributed sampling locations (see Table 2.8 in Section 2.7), the empirical size and power from GSC-u was negatively affected by a large moving window size; hence, we recommend setting  $c = 1$  and choosing  $n_b \lesssim n^{1/2}$ . For the MS test, a small block size negatively affected the empirical size and power; thus, we recommend choosing  $n_b \gtrsim n^{1/2}$  for this test.

Between the choices of a lag set, block size, and bandwidth, choosing an appropriate bandwidth to smooth over observed spatial lags for non-gridded sampling locations is the most challenging. For GSC-u the user needs to choose the form of the smoothing kernel as well as the bandwidth for both the entire grid and the subblocks while MS use an Epanechnikov kernel and empirical bandwidth based on a user-specified tuning parameter. If the selected bandwidth is too large then over-smoothing occurs. In over-smoothing, there is very little filtering of the lag distance and direction. The lack of filtering produces similar estimates of the spatial dependence at lags with different directions and distances. If the selected bandwidth is too small, then there is very little smoothing and estimates of the spatial dependence are based on a small number of pairs of sampling locations and thus highly variable. Considering the aforementioned effects of the bandwidth, the bandwidth should decrease as  $n$  increases under the usual increasing domain asymptotics. For example,

simulations (not included) indicated a bandwidth of  $w = 0.65$  maintains nominal size when  $n = 950$ , but leads to deflated test size and power when  $n = 400$  on a smaller domain. García-Soidán et al. (2004), García-Soidán (2007), and Kim and Park (2012) develop theoretically optimal bandwidths for nonparametric semivariogram estimation, but these works are not applicable here because they focus on the isotropic case and require an estimate of the second derivative of the semivariogram. We have found that the empirical bandwidth used by MS tends to produce nominal size (see Table 2.6). For GSC-u we find the most consistent results with a bandwidth in the range of  $0.60 < w < 0.90$  when using a normal kernel truncated at 1.5, but these values will change when a different truncation value or kernel function are employed. For small sample sizes ( $n < 500$ ), our simulations demonstrate that test size and power are less affected by the choice of bandwidth when the p-value is approximated using a finite sample adjustment, indicating poor convergence to the asymptotic  $\chi^2$  distribution. Thus, the user should consider using the finite sample adjustment for non-gridded sampling locations when the sample size is small and there are at least 100 subblocks. While it is challenging to choose a bandwidth for GSC-u and the p-value of the test is sensitive to this parameter, GSC-u exhibits nominal size and has substantially higher power than MS when an appropriate bandwidth is selected.

## 2.6 Discussion

There is a volume of work on tests for isotropy in other areas of spatial statistics. Methods for detecting anisotropy in spatial point process data have been developed (e.g., Schabenberger and Gotway (2004, pg. 200-205), Guan (2003), Guan et al. (2006), and Nicolis et al. (2010)). For multivariate spatial data, Jona-Lasinio (2001) proposed a test for isotropy. Gneiting et al. (2007) provided a review of potential second-order assumptions and models for spatiotemporal geostatistical data, and a number of tests for second-order properties of spatiotemporal data have been developed (e.g., Fuentes (2006), Li et al. (2007), Park and Fuentes (2008), Shao and Li (2009), Jun and Genton (2012)). Li et al. (2008a) constructed

a test of the covariance structure for multivariate spatiotemporal data. Tests for isotropy have also been developed in the computer science literature (e.g., Molina and Feito, 2002; Chorti and Hristopulos, 2008; Spiliopoulos et al., 2011; Thon et al., 2015).

Appropriately specifying the second-order properties of the random field is an important step in modeling spatial data, and a number of models have been developed to capture anisotropy in spatial processes. Graphical tools, such as directional sample semivariograms, are commonly used to evaluate the assumption of isotropy, but these diagnostics can be misleading and open to subjective interpretation. We have presented and reviewed a number of procedures that can be used to more objectively test hypotheses of isotropy and symmetry without assuming a parametric form for the covariance function. These tests may be helpful for a novice user deciding on an appropriate spatial model. In abandoning parametric assumptions, these hypothesis testing procedures are subject and sensitive to choices regarding smoothing parameters, subsampling procedures, and finite sample adjustments. The test that is most appropriate for a set of data will largely depend on the sampling design. Additionally, there are trade-offs between the empirical power demonstrated by the tests and the number of choices user must make to implement the tests (e.g., between Guan et al. (2004) and Maity and Sherman (2012)). We have offered recommendations regarding the various choices of method and their implementation and have made the tests available in the **spTest** software. Because of the sensitivity of the tests to the various choices, we believe that graphical techniques and nonparametric hypothesis tests should be used in a complementary role. Graphical techniques can provide an initial indication of isotropy properties and inform sensible choices for a hypothesis test (e.g., in choosing the spatial lag set), while hypothesis tests can affirm intuition about graphical techniques.

## 2.7 Simulation Study Details and Further Results

We define the isotropic exponential covariance function as

$$C(h) = \begin{cases} \sigma^2 \exp(-\phi h) & \text{if } h > 0, \\ \tau^2 + \sigma^2 & \text{otherwise} \end{cases} \quad (14)$$

where  $h = \|\mathbf{s}_i - \mathbf{s}_j\|$  is the distance between sites  $\mathbf{s}_i$  and  $\mathbf{s}_j$  (Irvine et al., 2007). The corresponding semivariogram is  $\gamma(h) = (\tau^2 + \sigma^2) - \sigma^2 \exp(-\phi h)$ , where  $\tau^2$  is the nugget,  $\tau^2 + \sigma^2$  is the sill, and the effective range,  $\xi$ , the distance beyond which the correlation between observations is less than 0.05, is

$$\xi = \frac{-1}{\phi} \log \left( 0.05 \frac{\tau^2 + \sigma^2}{\sigma^2} \right).$$

Simulations in Section 2.4 were performed using the exponential covariance function (14) with a partial sill of one ( $\sigma^2 = 1$ ) and no nugget ( $\tau^2 = 0$ ). We also performed simulations using different nugget values (results not included). Introducing a nugget lead to decreased empirical test size and power. For the no nugget simulations, we chose the effective range,  $\xi$ , for isotropic processes to be 3, 6, and 12 corresponding to short, medium, and long range dependence. We introduce geometric anisotropy by transforming the sampling locations according to a scaling parameter,  $R$ , and a rotation parameter,  $\theta$ . Given an  $(R, \theta)$  pair, the coordinates  $(x, y)$  are transformed to the anisotropic coordinates,  $(x_a, y_a)$  via

$$(x_a, y_a) = (x, y) \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{R} \end{bmatrix}.$$

A realization from the anisotropic process is then created by simulating using the distance matrix from the transformed coordinates and placing the observed values at their corresponding untransformed sampling locations. Figure 2.5 shows the isotropic exponential cor-

relogram corresponding to  $\tau^2 = 1$  and  $\xi = 6$  and contours of equicorrelation corresponding to the  $(R, \theta)$  values used in the simulation study. Note that  $R = 0$  and  $\theta = 0$  corresponds to an isotropic process and a larger value of  $R$  corresponds to a more anisotropic process.

For the simulations comparing the GSC-g and LZ tests in Table 2.5, data were simulated on a subset of the integer grid,  $\mathbb{Z}^2$ . The p-values for the GSC-g test were approximated using a finite sample statistic (Guan et al., 2004), and we used the lag set in (10) and  $\mathbf{A}$  matrix in (11). For the results involving the LZ test, a test of complete symmetry was performed as an approximation to the null hypothesis of isotropy. The p-values for the LZ test were obtained using the CvM\* statistic. A nominal level of  $\alpha = 0.05$  was maintained by first testing reflection symmetry at  $\alpha = 0.025$  then testing complete symmetry at  $\alpha = 0.025$  if the hypothesis of reflection symmetry was not rejected. For the GSC-g test, the moving window dimensions were  $3 \times 2$  (width, height) and  $5 \times 3$  for the parent grids of  $18 \times 12$  and  $25 \times 15$ , respectively.

For the simulations in Table 2.6 comparing the GSC-u and MS tests, we simulated data at random, uniformly distributed sampling locations on  $10 \times 16$  and  $10 \times 20$  sampling domains. The lag set,  $\mathbf{A}$ , used for both tests is given in (10) with  $\mathbf{A}$  matrix (11), and the p-values for both methods were obtained using the asymptotic  $\chi^2_2$  distribution. For semivariogram estimates in GSC-u, we use independent (product) Gaussian (normal) kernels with a truncation parameter of 1.5. The bandwidth for the Gaussian kernel for smoothing over lags on the entire field and on moving windows was chosen as  $w = 0.75$ . We used the empirical bandwidth and the product Epanechnikov kernel given in Maity and Sherman (2012) to implement the MS test. For both tests, a grid with spacing of 1 was laid on the sampling region. Using this grid, the moving window dimensions for the GSC-u test were  $4 \times 2$  and the block size for the MS test were  $4 \times 2$ . For the MS test,  $B = 100$  resamples using the GBBB were used to estimate the asymptotic variance-covariance matrix.

For the results in Tables 2.7 - 2.9, we simulated mean 0, Gaussian RFs with exponential covariance function with no nugget, a sill of one, and medium effective range ( $\xi = 6$ ).



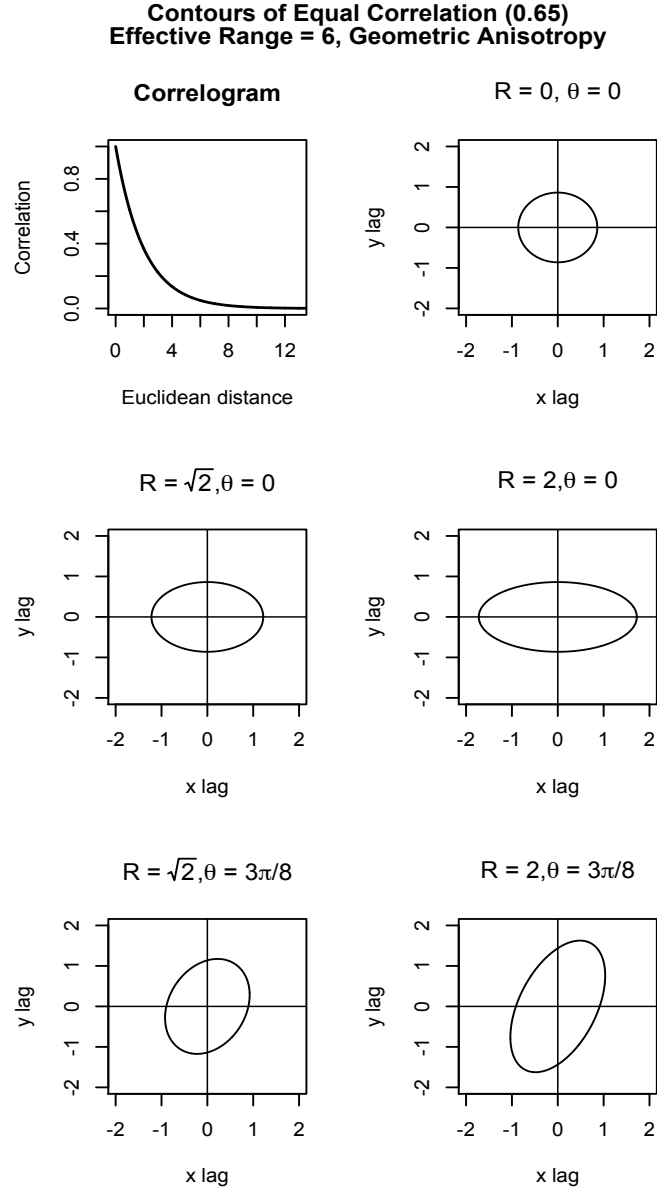


Figure 2.5: Correlogram and contours of equal correlation for the covariance models used in the simulation study. Note that  $R = 0, \theta = 0$  corresponds to an isotropic covariance function.

Sampling locations were generated randomly and uniformly over a  $16 \times 10$  sampling domain. We use the lag set and  $\mathbf{A}$  matrix from 10 and 11, respectively, unless otherwise noted. All tests were performed using a nominal level of  $\alpha = 0.05$ . For the GSC-u tests, we use product Gaussian kernels with a truncation parameter of 1.5. For the MS tests, we use the default product Epanechnikov kernels with empirical bandwidth specified in Maity and Sherman (2012).

The simulation results in Table 2.7 demonstrate the effects of changing the set of lags for the GSC-u and MS tests. For these simulations, the lag set labeled “regular” corresponds to the lag set given in (17). The lag set labeled “long” represents the lags in (17) multiplied by 2.5. Finally, the lag set labeled “more” stands for the lags in (17) with the additional pair of lags  $\{\mathbf{h}_5 = (1.132, 0.469)', \mathbf{h}_6 = (-0.469, 1.132)'\}$ . The lags  $\mathbf{h}_5$  and  $\mathbf{h}_6$  are a pair of lags that create approximate  $22.5^\circ$  and  $112.5^\circ$  angles, respectively, with the  $x$ -axis (counter-clockwise rotation) and have Euclidean length of approximately 1.22. These were chosen to supplement the lag pairs  $(\mathbf{h}_1, \mathbf{h}_2)$  which have unit length and create  $0^\circ$  and  $90^\circ$  angles with the  $x$ -axis and  $(\mathbf{h}_3, \mathbf{h}_4)$  which have length  $\sqrt{2} \approx 1.41$  and create  $45^\circ$  and  $135^\circ$  angles with the  $x$ -axis. The lag sets are plotted in Figure 2.6. The  $\mathbf{A}$  matrix for the “more” lag set was constructed as in (11), where orthogonal lags are contrasted. The p-values were calculated using the asymptotic  $\chi^2$  distribution with degrees of freedom based on the number of pairs of lags contrasted. For the GSC-u method, we used a bandwidth of 0.75. The moving window dimensions were  $4 \times 2$ . For the MS method, we chose block dimensions of  $4 \times 2$  and used  $B = 75$  resamples using the GBBB to estimate the asymptotic variance-covariance matrix. Table 2.8 demonstrates the effects of changing the block size for the GSC-u and MS tests. For these simulations, the labels “small”, “medium”, and “large” correspond to moving windows/blocks of size  $3 \times 2$ ,  $4 \times 2$ , and  $5 \times 3$ , respectively. Because we simulated  $n = 300$  uniformly distributed sampling locations on a  $16 \times 10$  domain, we expect 1.875 sampling locations per unit area. Thus, we expect  $n_b = 11.3$ , 15, and 28.1 points per block for the small, medium, and large block sizes, respectively. We find that the methods tend to have nominal size when  $n_b \approx n^{1/2} = 17.3$ .

Table 2.5: Empirical size and power for Guan et al. (2004) [denoted GSC-g] and Lu and Zimmerman (2005) [denoted LZ] for 1000 realizations of a mean 0 GRF with gridded sampling locations using a nominal level of  $\alpha = 0.05$ . Computational time for each method is also included.

(a) Sample size of  $n = 216$  gridded sampling locations.

18 cols $\times$ 12 rows grid					
			effective range		
$R$	$\theta$	Method	3	6	12
0	0	GSC-g	0.04	0.05	0.05
		LZ	0.05	0.08	0.05
$\sqrt{2}$	0	GSC-g	0.29	0.37	0.36
		LZ	0.07	0.09	0.09
2	0	GSC-g	0.82	0.86	0.87
		LZ	0.17	0.15	0.15
$\sqrt{2}$	$\frac{3\pi}{8}$	GSC-g	0.25	0.27	0.29
		LZ	0.10	0.12	0.12
2	$\frac{3\pi}{8}$	GSC-g	0.75	0.78	0.80
		LZ	0.29	0.31	0.31
Computational Time for 1 Test					
GSC-g				1.11 seconds	
LZ				1.45 seconds	

(b) Sample size of  $n = 375$  gridded sampling locations.

25 cols  $\times$  15 rows grid

			effective range		
$R$	$\theta$	Method	3	6	12
0	0	GSC-g	0.05	0.06	0.06
		LZ	0.06	0.06	0.08
$\sqrt{2}$	0	GSC-g	0.59	0.63	0.63
		LZ	0.08	0.10	0.09
2	0	GSC-g	0.98	0.99	0.98
		LZ	0.17	0.15	0.15
$\sqrt{2}$	$\frac{3\pi}{8}$	GSC-g	0.51	0.52	0.54
		LZ	0.13	0.16	0.17
2	$\frac{3\pi}{8}$	GSC-g	0.96	0.97	0.98
		LZ	0.40	0.43	0.46

Computational Time for 1 Test	
GSC-g	7.29 seconds
LZ	4.99 seconds

For both tests, we used the lags in (17), and the blocks are defined by a grid with spacing 0.5 placed on the sampling region (i.e., a  $4 \times 2$  window is achieved by setting the window dimensions to  $8 \times 4$  in the `spTest` software). We performed the tests using a nominal level of  $\alpha = 0.05$ , and the p-values were calculated using the asymptotic  $\chi^2$  distribution. For the GSC-u method, we used a bandwidth of 0.75. For the MS method, we used  $B = 100$  resamples using the GBBB to estimate the asymptotic variance-covariance matrix. Finally, Table 2.9 demonstrates the effects of changing the bandwidth for the GSC-u test. We use bandwidths of  $w = 0.65, 0.75$ , and  $0.85$ . The p-values are calculated using both the asymptotic  $\chi^2$  distribution and using a finite sample adjustment similar to the one used by Guan et al. (2004) for gridded sampling locations.

Table 2.6: Empirical size and power for Guan et al. (2004) [denoted GSC-u] and Maity and Sherman (2012) [denoted MS] for 1000 realizations of a mean 0 GRF with uniformly distributed sampling locations using a nominal level of  $\alpha = 0.05$ . Computational time for each method is also included.

(a) Sample size of  $n = 300$  uniformly distributed sampling locations.      (b) Sample size of  $n = 450$  uniformly distributed sampling locations.

10 height $\times$ 16 width domain					
			effective range		
$R$	$\theta$	Method	3	6	12
0	0	GSC-u	0.02	0.03	0.04
		MS	0.04	0.04	0.04
$\sqrt{2}$	0	GSC-u	0.10	0.19	0.19
		MS	0.08	0.06	0.05
2	0	GSC-u	0.43	0.50	0.53
		MS	0.21	0.13	0.13
$\sqrt{2}$	$\frac{3\pi}{8}$	GSC-u	0.10	0.16	0.16
		MS	0.11	0.06	0.05
2	$\frac{3\pi}{8}$	GSC-u	0.37	0.48	0.51
		MS	0.27	0.18	0.16
Computational Time for 1 Test					
GSC-u				2.17 seconds	
MS				83.40 seconds	

10 height $\times$ 20 width domain					
			effective range		
$R$	$\theta$	Method	3	6	12
0	0	GSC-u	0.02	0.03	0.05
		MS	0.04	0.04	0.02
$\sqrt{2}$	0	GSC-u	0.13	0.22	0.24
		MS	0.10	0.07	0.05
2	0	GSC-u	0.57	0.65	0.69
		MS	0.29	0.19	0.16
$\sqrt{2}$	$\frac{3\pi}{8}$	GSC-u	0.13	0.18	0.23
		MS	0.13	0.09	0.06
2	$\frac{3\pi}{8}$	GSC-u	0.53	0.64	0.67
		MS	0.38	0.26	0.21
Computational Time for 1 Test					
GSC-u				4.44 seconds	
MS				162.35 seconds	

Table 2.7: Effects of changing the lag set. Empirical size and power for Guan et al. (2004) [denoted GSC-u] and Maity and Sherman (2012) [MS] for 500 realizations of a mean 0 GRF with  $n = 400$  uniformly distributed sampling locations. The label “regular” corresponds to the lag set in (17), while “long” represents using longer lags, and “more” denotes using more lags (see Figure 2.6).

16 width $\times$ 10 height domain					
			Lag Set		
$R$	$\theta$	Method	regular	long	more
0	0	GSC-u	0.03	0.02	0.03
		MS	0.02	0.18	0.03
$\sqrt{2}$	$\frac{3\pi}{8}$	GSC-u	0.18	0.07	0.13
		MS	0.10	0.23	0.08
2	$\frac{3\pi}{8}$	GSC-u	0.56	0.19	0.51
		MS	0.24	0.38	0.22

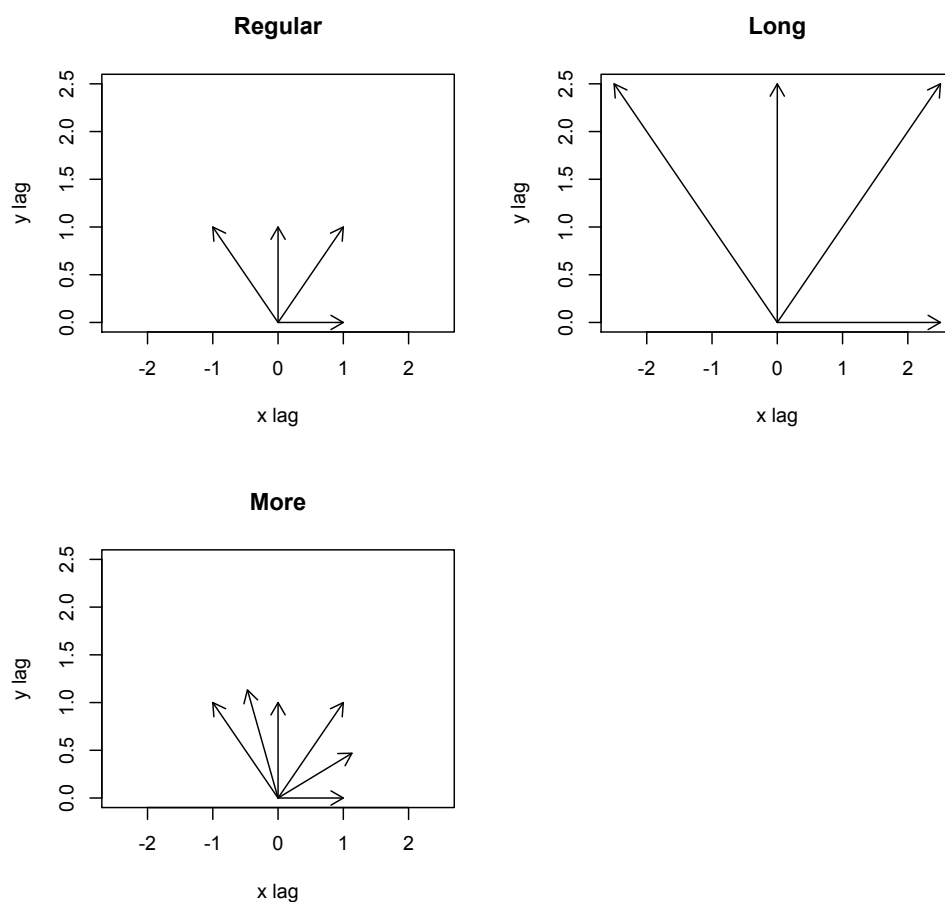


Figure 2.6: The lag sets used for the simulations in Table 2.7. These simulations were designed to explore the effect of changing the spatial lag sets on test size and power. “Regular” denotes the usual set of spatial lags used to test isotropy. “Long” denotes a set of spatial lags with greater Euclidean length than those in the “regular” set. “More” denotes a set of spatial lags that cover more directions than the “regular” set.

Table 2.8: Effects of changing the window/block size. Empirical size and power for Guan et al. (2004) [denoted GSC-u] and Maity and Sherman (2012) [MS] for 500 realizations of a mean 0 GRF with  $n = 300$  uniformly distributed sampling locations. The label “medium” corresponds to the window/block size of  $4 \times 2$ , while “small” represents using a smaller window, and “large” denotes using a larger window.

**16 width  $\times$  10 height domain**

			Window/Block Size		
$R$	$\theta$	Method	small	medium	large
0	0	GSC-u	0.05	0.04	0.01
		MS	0.03	0.04	0.04
$\sqrt{2}$	0	GSC-u	0.17	0.13	0.04
		MS	0.06	0.07	0.08
2	0	GSC-u	0.58	0.53	0.23
		MS	0.20	0.20	0.22

Table 2.9: Effects of changing bandwidth. “Bdwth” denotes the bandwidth used in the Gaussian smoothing kernel. Empirical size and power for Guan et al. (2004) [denoted GSC-u] for 500 realizations of a mean 0 GRF with  $n = 400$  uniformly distributed sampling locations using a nominal level of  $\alpha = 0.05$ .

(a) P-value: asymptotic  $\chi^2$  distribution

**16 width  $\times$  10 height domain**

			Effective Range		
$R$	$\theta$	Bdwth	3	6	12
0	0	0.65	0.00	0.00	0.01
		0.75	0.02	0.04	0.06
		0.85	0.07	0.13	0.14
$\sqrt{2}$	0	0.65	0.02	0.05	0.08
		0.75	0.12	0.20	0.23
		0.85	0.26	0.34	0.38
2	0	0.65	0.18	0.24	0.29
		0.75	0.51	0.60	0.62
		0.85	0.67	0.72	0.75
$\sqrt{2}$	$\frac{3\pi}{8}$	0.65	0.03	0.04	0.03
		0.75	0.14	0.17	0.21
		0.85	0.24	0.32	0.40
2	$\frac{3\pi}{8}$	0.65	0.16	0.23	0.27
		0.75	0.46	0.55	0.57
		0.85	0.63	0.74	0.74

(b) P-value: finite sample

**16 width  $\times$  10 height domain**

			Effective Range		
$R$	$\theta$	Bdwth	3	6	12
0	0	0.65	0.01	0.03	0.05
		0.75	0.03	0.07	0.08
		0.85	0.06	0.09	0.11
$\sqrt{2}$	0	0.65	0.09	0.17	0.19
		0.75	0.18	0.27	0.32
		0.85	0.24	0.31	0.37
2	0	0.65	0.42	0.51	0.50
		0.75	0.63	0.65	0.68
		0.85	0.65	0.71	0.73
$\sqrt{2}$	$\frac{3\pi}{8}$	0.65	0.11	0.11	0.16
		0.75	0.19	0.20	0.30
		0.85	0.23	0.27	0.37
2	$\frac{3\pi}{8}$	0.65	0.36	0.45	0.46
		0.75	0.53	0.61	0.61
		0.85	0.58	0.68	0.67

## CHAPTER 3

# SPTEST: AN R PACKAGE IMPLEMENTING NONPARAMETRIC TESTS OF ISOTROPY

### 3.1 Introduction

An important step in modeling a spatial process is choosing the form of the covariance function. The form of the covariance function will have an effect on kriging as well as parameter estimates and the associated uncertainty (Cressie, 1993, pg. 127-135). A common simplifying assumption about the spatial covariance function is that it is isotropic, that is, the dependence between sampling locations depends only on the distance between locations and not on their relative orientation. This assumption may not always be reasonable; for example, wind may lead to directional dependence in environmental monitoring data. Misspecification of the second order properties can lead to misleading inference. Sherman (2011, pg. 87-90) and Guan et al. (2004) summarize the effects of incorrectly specifying isotropy properties on kriging estimates through numerical examples. In order to choose an appropriate model, a statistician must first assess the nature of the spatial variation of his or her data. To check for anisotropy (directional dependence) in spatially-referenced data, a number of graphical techniques are available, such as directional sample variograms or rose diagrams (Matheron, 1961; Isaaks and Srivastava, 1989, pg. 149-154). Spatial statisticians may have intuition about the interpretation and reliability of these diagnostics, but a less experienced user may desire evaluation of assumptions via a hypothesis test.

A number of nonparametric tests of isotropy, which avoid the choice of a parametric covariance function, have been developed (see, e.g., Lu and Zimmerman, 2001; Guan et al., 2004; Lu and Zimmerman, 2005; Maity and Sherman, 2012). In Weller and Hoeting (2016),

we provide a review of the different nonparametric methods available for testing isotropy and symmetry properties, including an extensive simulation study. In the current chapter, we aim to showcase the functionality of our **R** (R Core Team, 2015) package **spTest** (Weller, 2016d), which implements several of the aforementioned methods. We use two real data examples to illustrate how the nonparametric hypothesis tests available in **spTest** can be used to assess isotropy properties in spatially-referenced data. The examples also demonstrate how graphical techniques and hypothesis tests can be used in a complementary role. The remainder of this chapter is organized as follows: Sections 3.2 and 3.3 establish additional notation and background information; Section 3.4 describes the functionality of the **spTest** package; Section 3.5 demonstrates how to use the functions in **spTest** in conjunction with graphical techniques on two different data sets; Section 3.6 concludes the chapter with a discussion.

## 3.2 Nonparametric Covariance Functions and Estimation

When modeling a RF, a typical assumption is that the spatial covariance  $C(\mathbf{h})$  or semi-variogram  $\gamma(\mathbf{h})$  function can be described by a parametric model. A number of methods are available for estimating the parameters of these models, for example, maximum likelihood or least squares (Cressie, 1993, pg. 90-97). Parametric models ensure that the covariance function is valid and provide parameters that can be interpreted as describing characteristics of the random field, such as the range of dependence or the direction of anisotropy (Schabenberger and Gotway, 2004, pg. 141-152). The methods in **spTest** are nonparametric tests of isotropy, which circumvent the choice of a parametric form for the covariance (semivariogram) function. A nonparametric test of isotropy escapes potential misspecification of a parametric covariance function and avoids the potential for having to estimate the covariance function twice (e.g., under the null and alternative hypothesis for a likelihood-ratio test).

A nonparametric test of isotropy requires a nonparametric estimator of second order properties of the RF. We discuss nonparametric estimation of the semivariogram function



here and note that similar techniques can be used for nonparametric estimation of the covariance function. The classical moment-based estimator of the semivariogram (Matheron, 1962) is the sample semivariogram given by

$$\hat{\gamma}(\mathbf{h}) = \frac{1}{2|\mathcal{D}(\mathbf{h})|} \sum [Y(\mathbf{s}) - Y(\mathbf{s} + \mathbf{h})]^2, \quad (15)$$

where the sum is over  $\mathcal{D}(\mathbf{h}) = \{\mathbf{s} : \mathbf{s}, \mathbf{s} + \mathbf{h} \in \mathcal{D}\}$  and  $|\mathcal{D}(\mathbf{h})|$  is the number of elements in  $\mathcal{D}(\mathbf{h})$ . The set  $\mathcal{D}(\mathbf{h})$  is the set of sampling location pairs that are separated by spatial lag  $\mathbf{h}$ . There are two important modifications to the estimator in (15) that are pertinent to the methods described in this chapter. First, for non-gridded sampling locations, the estimator needs to be modified to account for the fact that very few or no pairs of locations will be separated by a specific spatial lag,  $\mathbf{h}$ . One solution to this challenge is to specify a distance tolerance,  $\epsilon$ , such that lags having length  $\|\mathbf{h}\| \pm \epsilon$  are included in estimating the semivariogram at lag  $\mathbf{h}$ . Second, directional sample semivariograms can be estimated by using only observations that are separated by spatial lags in a specific direction. For example, to investigate potential anisotropy, we can compare sample semivariograms between the horizontal and vertical directions. For non-gridded sampling locations, very few pairs of locations will lie at a specific distance and directional lag, so we need to allow for both a distance and a directional tolerance when estimating the semivariogram. A common method for doing this is by using a product kernel smoother that smoothes over both the horizontal ( $h_1$ ) and vertical ( $h_2$ ) components of the spatial lag  $\mathbf{h} = (h_1, h_2)^\top$ .

Spatial RFs and their second order properties can also be expressed in the spectral (or frequency) domain using Fourier transforms. The spectral representation of RFs and their second order properties provides alternative methods for testing second order properties. Here we focus only on the methods in the spatial domain and refer the interested reader to Weller and Hoeting (2016), Fuentes and Reich (2010), and/or Chapter 4. We note that that,

in addition to the methods from the spatial domain, the nonparametric spectral methods from Lu and Zimmerman (2005) are also implemented in `spTest`.

### 3.3 Nonparametric Tests of Isotropy

Lu (1994) and Lu and Zimmerman (2001) pioneered a popular approach to testing second-order properties when they used the joint asymptotic normality of the sample semivariogram computed at different spatial lags to evaluate symmetry and isotropy properties. The subsequent works of Guan et al. (2004, 2007) and Maity and Sherman (2012) built upon these ideas and are the primary methods programmed in `spTest`. Here we give an overview of the tests in Guan et al. (2004) and Maity and Sherman (2012).

Under the null hypothesis that the RF is isotropic, it follows that the values of  $\gamma(\cdot)$  evaluated at any two spatial lags that have the same norm are equal, independent of the direction of the lags. For example, under the assumption of isotropy,  $\gamma((-6, 0)) = \gamma((\sqrt{3}, \sqrt{3}))$ . To completely specify the null hypothesis of isotropy, theoretically, one would need to compare semivariogram values for an infinite set of lags. In practice, a small number of lags are specified. Then it is possible to test a hypothesis consisting of a set of linear contrasts of the form

$$H_0 : \mathbf{A}\gamma(\cdot) = \mathbf{0} \tag{16}$$

as a proxy for the null hypothesis of isotropy, where  $\mathbf{A}$  is a full row rank matrix (Lu and Zimmerman, 2001). For example, a set of lags, denoted  $\mathbf{\Lambda}$ , commonly used in practice on

gridded sampling locations with unit spacing is

$$\mathbf{\Lambda} = \{\mathbf{h}_1 = (1, 0), \mathbf{h}_2 = (0, 1), \mathbf{h}_3 = (1, 1), \mathbf{h}_4 = (-1, 1)\}, \quad (17)$$

and the corresponding  $\mathbf{A}$  matrix under  $H_0 : \mathbf{A}\boldsymbol{\gamma}(\mathbf{\Lambda}) = \mathbf{0}$  is

$$\mathbf{A} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}. \quad (18)$$

In other words, using (17) and (18) under the hypothesis (16), we contrast the semivariogram values at lags  $\mathbf{h}_1 = (1, 0)$  and  $\mathbf{h}_2 = (0, 1)$ , and likewise, contrast semivariogram values at lags  $\mathbf{h}_3 = (1, 1)$  and  $\mathbf{h}_4 = (-1, 1)$ . An important step in detecting anisotropy is the choice of lags,  $\mathbf{\Lambda}$ , as the test results will only hold for the set of lags considered (Guan et al., 2004). While this choice is somewhat subjective, there are several considerations for determining the set of lags. First, in terms of Euclidean distance, short lags should be chosen. Short lags are preferred because estimates of the semivariogram at long lags tend to be more variable. Second, pairs of orthogonal lags should be contrasted because, for an anisotropic process, the directions of strongest and weakest spatial correlation will typically be orthogonal. For other considerations and more detailed guidelines regarding the choice of lags, see Weller and Hoeting (2016), Lu and Zimmerman (2001), and Guan et al. (2004).

The tests in Guan et al. (2004) and Maity and Sherman (2012) require estimating the semivariogram and covariance function values, respectively, at the set of chosen lags. We denote the vector of point estimates of the semivariogram/covariance function at the chosen lags as  $\hat{\mathbf{G}}_n$ , the true values as  $\mathbf{G}$ , and the asymptotic variance-covariance matrix of  $\hat{\mathbf{G}}_n$  as

$\Sigma$ . Under increasing domain asymptotics, a central result for both methods is that

$$a_n(\widehat{\mathbf{G}}_n - \mathbf{G}) \xrightarrow[n \rightarrow \infty]{d} MVN(\mathbf{0}, \Sigma), \quad (19)$$

where  $a_n$  is a normalizing constant. This result holds assuming stationarity of the RF and mild conditions on the RF's mixing and moments. See Section 3.6 for additional discussion regarding the assumptions for the tests. The test statistic is a quadratic form

$$TS = b_n^2(\mathbf{A}\widehat{\mathbf{G}}_n)^\top (\mathbf{A}\widehat{\Sigma}\mathbf{A}^\top)^{-1}(\mathbf{A}\widehat{\mathbf{G}}_n), \quad (20)$$

where  $\widehat{\Sigma}$  is an estimate of the asymptotic variance-covariance matrix and  $b_n$  is a normalizing constant. A  $p$  value can be obtained from the asymptotic  $\chi^2$  distribution with degrees of freedom given by the row rank of  $\mathbf{A}$ . Important differences between these works regard the distribution of sampling locations, shape of the sampling domain, and estimation of  $\mathbf{G}$  and  $\Sigma$ . Table 3.1 outlines the primary differences between these methods; we refer the interested reader to Weller and Hoeting (2016) for more details.

### 3.4 Nonparametric Tests Implemented in `spTest`

The R package `spTest` includes functions for implementing the tests developed in Guan et al. (2004), Lu and Zimmerman (2005), and Maity and Sherman (2012). The `spTest` functions for implementing these tests are listed in Table 3.1. For example, the test from Guan et al. (2004) for data observed at non-gridded, but uniformly distributed, sampling locations is implemented in the function `GuanTestUnif`, which takes the following arguments:

```
GuanTestUnif(spdata, lagmat, A, df, h = 0.7, kernel = "norm",
  truncation = 1.5, xlims, ylims, grid.spacing = c(1, 1),
  window.dims = c(2, 2), subblock.h, sig.est.finite = T).
```

Table 3.1: Nonparametric tests of isotropy available in the R package **spTest**, including GSC = Guan et al. (2004), LZ = Lu and Zimmerman (2005), and MS = Maity and Sherman (2012). The column “**spTest Func’n**” lists the name of the function used to implement the test in the **spTest** package. “Design” indicates the spatial sampling design for which the test is valid. “Estimator” describes the method used to estimate second order properties.

	<b>Hypothesis Test Properties</b>			
<b>Method</b>	<b>spTest Func’n</b>	<b>Design</b>	<b>Estimator</b>	<b>Primary User Choices</b>
GSC	<b>GuanTestGrid</b>	grid	sample semi-variogram	lag set, window size
LZ	<b>LuTest</b>	grid	periodogram	—
GSC	<b>GuanTestUnif</b>	uniform <sup>a</sup>	kernel sample semivariogram	lag set, window size, bandwidth
MS	<b>MaityTest</b>	any	kernel sample covariogram	lag set, block size

<sup>a</sup> sampling locations must be uniformly distributed on the domain

There are several necessary inputs. The matrix **spdata** includes the coordinates of sampling locations and the corresponding data values. The spatial lags used to estimate the semivariogram, denoted  $\mathbf{\Lambda}$ , are specified in the matrix **lagmat**. The matrix  $\mathbf{A}$  in (20) is specified by **A** and provides the contrasts of the semivariogram estimates, and its row rank is indicated by the parameter **df** (the degrees of freedom for the asymptotic  $\chi^2$  distribution). The values **h** and **kernel** provide the bandwidth (smoothing) parameter and form of the kernel smoother, respectively, used to smooth over spatial lags when estimating the semivariogram. If a normal smoothing kernel is used, then the **truncation** parameter indicates where to truncate the normal kernel (i.e., zero weight for spatial lags larger than this value). The parameters **xlims** and **ylims** give the horizontal and vertical limits of the sampling region (a rectangular sampling region is assumed). When performing a nonparametric tests of isotropy for non-gridded sampling locations, we must place a grid on the sampling domain and choosing a moving window or block size to estimate  $\mathbf{\Sigma}$  in (19) (see Section 3.5.2). A grid is placed over the sampling region according the width and height specified by **grid.spacing**. The dimensions of the moving window, given in the units of

the underlying grid, are determined by the values in `window.dims`. The bandwidth of the smoothing kernel used to estimate the semivariogram on the subblocks of data created by the moving window is indicated by `subblock.h` and a finite sample adjustment to the estimate of the asymptotic variance-covariance matrix is made by setting `sig.est.finte = T`. For more information about the different arguments and guidelines on how to choose them, see Weller and Hoeting (2016), the `spTest` manual, and the original works (Guan et al., 2004, 2007; Maity and Sherman, 2012).

### 3.5 Applications: Using `spTest` to Check for Anisotropy

We demonstrate the functionality of the `spTest` package on two data sets: the first containing data at gridded sampling locations; the second containing data collected via a non-gridded sampling design. For more details on the functions and examples using simulated data, see the `spTest` manual. The `spTest` package can be used independently of other packages built for analyzing spatial data, but it works nicely with two other packages loaded into R: `fields` (Nychka et al., 2014) and `geoR` (Ribeiro Jr. and Diggle, 2001). We also load the `splines` (R Core Team, 2015), `MASS` (Venables and Ripley, 2002), and `rgdal` (Bivand et al., 2016) packages, which we use to estimate mean functions, compute studentized residuals, and calculate map projection coordinates, respectively.

```
R> library("spTest")
R> library("fields")
R> library("geoR")
R> library("splines")
R> library("MASS")
R> library("rgdal")
```

For the two examples given below, we use graphical diagnostics and the hypothesis tests implemented in `spTest` to determine whether or not an assumption of isotropy is reasonable

for spatially-referenced data. The general strategy will be to first do exploratory data analysis (EDA) of the original data and create a model for the mean of the spatial process using appropriate covariates. After estimating a model for the mean, we extract residuals and again use EDA to check for remaining spatial dependence and utilize graphical diagnostics and hypothesis tests to investigate potential anisotropy. For brevity, we have not included the full version of EDA code and plots; instead, we include only the most relevant to demonstrating the functionality of the `spTest` package. The complete version of the code is available on `github` (Weller, 2016b).

### 3.5.1 Gridded Sampling Locations

The gridded data used in this section come from the North American Regional Climate Change Assessment Program [NARCCAP] (Mearns et al., 2009). The data set `WRFG` in `spTest` includes coordinates and a 24-year average of yearly average temperatures from runs of the Weather Research and Forecasting - Grell configuration (WRFG) regional climate model (RCM) using boundary conditions from the National Centers for Environmental Prediction (NCEP). The original data are available on the NARCCAP website and the R code used to create the yearly averages is available on `github` (Weller, 2016a). The data set contains both latitude and longitude and universal transverse mercator (UTM) coordinates. The UTM coordinates specify the regular grid for 14,606 grid boxes along with average temperature at surface at each grid box. Figure 3.1 displays a heat map of all of the data and was created using the `image.plot` function from the `fields` package. Due to computational considerations and because the methods in `spTest` assume stationarity, for our analysis we use a  $20 \times 20$  subset of the grid boxes defined by the UTM coordinates over the central United States (see Figures 3.1 and 3.2a).

```
R> data("WRFG")
R> coords <- expand.grid(WRFG$xc, WRFG$yc)
R> sub <- which(coords[, 1] > 2900000 & coords[, 1] < 3950000 &
```

```

+     coords[, 2] > 1200000 & coords[, 2] < 2250000)
R> coords.ll <- cbind((WRFG$lon - 360)[sub], WRFG$lat[sub])
R> image.plot(WRFG$lon - 360, WRFG$lat, WRFG$WRFG.NCEP.tas,
+     col = two.colors(n = 256, start = "blue3", end = "red3",
+     middle = "gray60"), legend.lab = "Temp (K)",
+     legend.cex = 0.8, legend.line = 2.2, xlab = "Longitude",
+     ylab = "Latitude", main = "Mean WRFG-NCEP Temperatures")
R> world(add = T)
R> left <- seq(1, 400, by = 20)
R> right <- seq(20, 400, by = 20)
R> for (i in 2:20) {
+     segments(coords.ll[i - 1, 1], coords.ll[i - 1, 2],
+     coords.ll[i, 1], coords.ll[i, 2], lwd = 2)
+     segments(coords.ll[left[i - 1], 1], coords.ll[left[i -
+     1], 2], coords.ll[left[i], 1], coords.ll[left[i],
+     2], lwd = 2)
+     segments(coords.ll[right[i - 1], 1], coords.ll[right[i -
+     1], 2], coords.ll[right[i], 1], coords.ll[right[i],
+     2], lwd = 2)
+     j <- i + 380
+     segments(coords.ll[j - 1, 1], coords.ll[j - 1, 2],
+     coords.ll[j, 1], coords.ll[j, 2], lwd = 2)
+ }

```

To investigate potential anisotropy in the relevant subset of these data, we can examine two graphical diagnostics: a heat map and directional sample semivariograms. We use the function `variog4` from the `geoR` package to estimate directional semivariograms to visually assess isotropy properties.



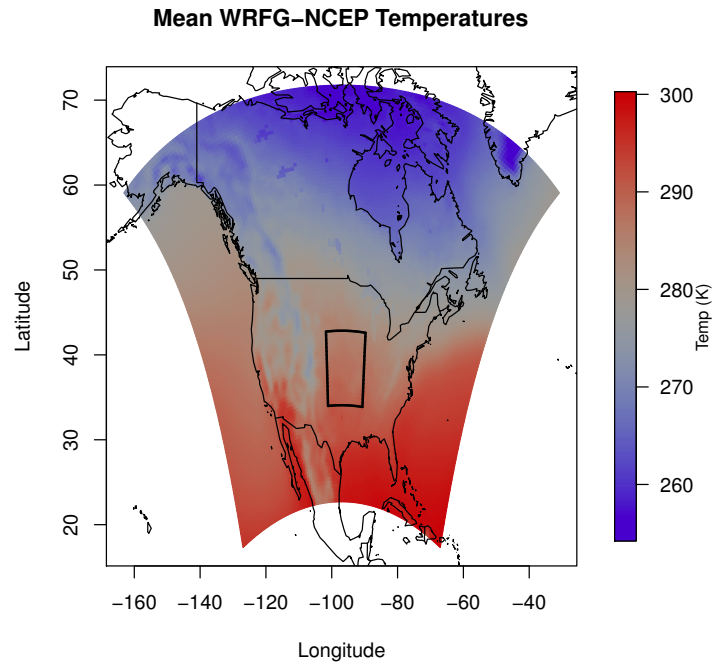


Figure 3.1: Heat map of temperatures in latitude and longitude coordinates. The temperature values are the 24 year average of WRFG-NCEP yearly average temperature. The black box indicates the relevant subset of data used for the example in Section 3.5.1, and this subset is displayed in Figure 3.2a.

```

R> tas <- c(WRFG$WRFG.NCEP.tas)[sub]
R> x.coord <- unique(coords[sub, 1])
R> y.coord <- unique(coords[sub, 2])
R> nx <- length(x.coord)
R> ny <- length(y.coord)
R> tas.mat <- matrix(tas, nrow = nx, ncol = ny, byrow = F)
R> image.plot(x.coord, y.coord, tas.mat, col = two.colors(n = 256,
+       start = "blue3", end = "red3", middle = "gray60"),
+       legend.lab = "Temp (K)", legend.cex = 0.8, legend.line = 2.2,
+       ylab = "Northing", xlab = "Easting",
+       main = "Subset of Temperatures")

R> tas.geodat <- as.geodata(cbind(coords[sub, 1], coords[sub,
+       2], tas))
R> plot(variog4(tas.geodat), xlab = "distance (meters)",
+       ylab = "estimated semivariogram")
R> title("Directional Sample Semivariograms")

```

The heat map in Figure 3.2a indicates that the spatial process is anisotropic, having a stronger spatial dependence in the horizontal direction than the vertical direction. Intuitively, northing coordinates are an important factor in determining average temperature, and we need to include its effect in a model for these data. We also notice non-linear trends in temperature as a function of easting coordinates in Figure 3.2a. Thus, the anisotropy can be attributed, at least in part, to the fact that we have not modeled important covariates related to the process. The directional sample semivariograms in Figure 3.2b reaffirm the notion that the data exhibit anisotropy as the 90° sample semivariogram appears much different than the other three. Before modeling the effects of northing and easting coordinates, we

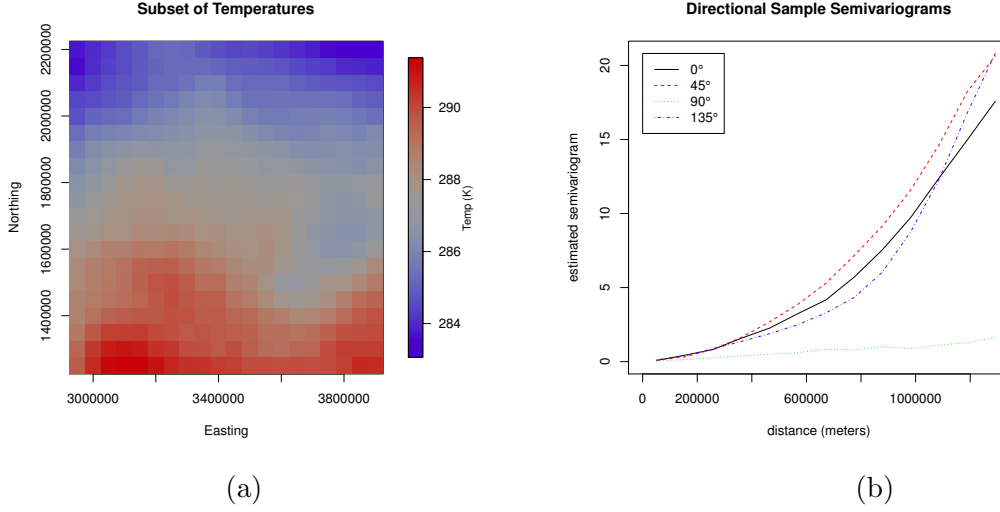


Figure 3.2: Graphical assessments of isotropy in the  $20 \times 20$  subset of WRFG temperature data. Because northing coordinates have not been accounted for, the heat map (Figure 3.2a) indicates that the dependence between observations is stronger in the east-west direction than the north-south direction. The directional dependence is also evidenced by the differences between the directional sample semivariograms (Figure 3.2b).

use the `GuanTestGrid` function from `spTest` to affirm our understanding that these data exhibit anisotropy.

Necessary conditions for the asymptotic properties of the nonparametric tests to hold are typically met when the data are Gaussian (see Section 3.6). A quantile-quantile (QQ) plot (not shown) of the relevant subset of WRFG temperatures indicates that a Gaussian assumption is reasonable. To implement the nonparametric test in Guan et al. (2004) via the function `GuanTestGrid`, we need to specify the spatial lags that will be used to test for differences in the semivariogram. For this test we choose the lag set (17) and use the matrix  $\mathbf{A}$  in (18) to contrast the semivariogram estimates. With the first row of  $\mathbf{A}$  and the first two entries of  $\mathbf{A}$ , we are contrasting the estimated dependence structure in the  $0^\circ(\mathbf{h}_1)$  and  $90^\circ(\mathbf{h}_2)$  directions for data separated by one horizontal or vertical sampling location. The second row of  $\mathbf{A}$  and second two entries of  $\mathbf{A}$  contrast the estimated dependence structure in the  $45^\circ(\mathbf{h}_3)$  and  $135^\circ(\mathbf{h}_4)$  directions for data separated by one diagonal sampling location. Because the grid spacing between sampling locations is 50,000 meters, we set the the scaling parameter

`delta = 50,000`. To create subblocks of data used to estimate  $\Sigma$  in (20), we choose a moving window with a size of  $4 \times 4$  grid cells. The moving window dimensions should be chosen so that the window has the same shape (i.e., square or rectangle) and orientation as the sampling domain. To maximize the amount of data used to estimate  $\Sigma$ , the dimensions of the window should evenly divide the number of columns and rows, respectively, of the entire region. The window dimensions should also be compatible with the spatial lags in  $\Lambda$ . For example, if sampling locations are on the integer grid  $\mathbb{Z}^2$ , a window with dimensions of  $2 \times 2$  grid cells cannot be used to estimate the variability of the semivariogram at a lag with Euclidean distance longer than  $\sqrt{2}$ , the maximum distance between locations in the moving window. For this example there are 20 rows and columns, and we are using lags with spacings of one or two sampling locations; hence, window dimensions of  $2 \times 2$  or  $4 \times 4$  grid cells are reasonable choices. We run the test using window dimensions of  $4 \times 4$  grid cells via the following code, suppressing some of the output for brevity.

```
R> my.delta <- 50000
R> mylags <- rbind(c(1, 0), c(0, 1), c(1, 1), c(-1, 1))
R> myA <- rbind(c(1, -1, 0, 0), c(0, 0, 1, -1))
R> tr <- GuanTestGrid(spdata = tas.geodat, delta = my.delta,
+   lagmat = mylags, A = myA, df = 2, window.dims = c(4,
+   4), pt.est.edge = TRUE, sig.est.edge = TRUE,
+   sig.est.finite = TRUE)
R> tr$alternative <- NULL
R> tr$sigma.hat <- NULL
R> print(tr)
```

Test of isotropy from Guan et. al. (2004) for gridded  
sampling locations using the sample semivariogram.

```

data:  tas.geodat

Chi-sq = 34.063, df = 2, p-value = 4.012e-08

p-value (finite adj.) < 2.2e-16, number of subblocks: 240


sample estimates: (lag value)

      (1,0)      (0,1)      (1,1)      (-1,1)
0.02887917 0.07927363 0.09644188 0.11504961

```

As we suspected, the results of the hypothesis test ( $p$  value  $< 0.05$ ) indicate that the data exhibit anisotropy. We note that for gridded data, Guan et al. (2004) recommend using the  $p$  value computed via a finite sample correction. The function `GuanTestGrid`, and other functions in `spTest`, return a  $p$  value(s) for the test and information used in computing the  $p$  value, such as the point estimates ( $\hat{\mathbf{G}}_n$ ), estimates of the asymptotic variance-covariance matrix ( $\hat{\Sigma}$ ), the number of subblocks used to estimate  $\Sigma$ , and other information about the estimation process. We note that the point estimates for the directional semivariograms are slightly different between the functions from the `spTest` and `geoR` packages due to different kernel methods used in estimation.

As previously mentioned, we need to model the effects of northing and easting UTM coordinates on average temperature. We fit temperature as a nonparametric additive function of both the northing and easting coordinates via least-squares using cubic splines. The cubic splines can be specified using the function `ns` from the `splines` package and the least squares fit is computed via the `lm` function.

```

R> m1 <- lm(tas ~ ns(coords[sub, 1], df = 3) + ns(coords[sub,
+      2], df = 3))
R> summary(m1)

```

After removing the mean effects of the coordinates, we can check for any remaining (unaccounted for) spatial dependence and evidence of anisotropy in the residuals using graphical

diagnostics and a hypothesis test. A QQ plot of the studentized residuals (not shown) indicates that a Gaussian assumption is reasonable.

```
R> resid.mat <- matrix(studres(m1), nrow = nx, ncol = ny,
+   byrow = F)
R> image.plot(x.coord, y.coord, resid.mat, col = two.colors(n = 256,
+   start = "blue3", end = "red3", middle = "gray60"),
+   xlab = "Easting", ylab = "Northing")
R> title("Heat Map of Studentized Residuals")

R> resid.geo <- as.geodata(as.matrix(cbind(coords[sub, 1:2],
+   studres(m1))))
R> plot(variog4(resid.geo), xlab = "distance (meters)",
+   ylab = "estimated semivariogram")
R> title("Directional Sample Semivariograms")
```

The clusters of similar values in the heat map of Figure 3.3a, and the increase, followed by a leveling off, of the semivariogram values as distance increases in the directional sample semivariograms in Figure 3.3b indicate that the residuals are still spatially dependent. However, the plots in Figure 3.3 do not clearly illustrate whether or not the residuals exhibit anisotropy. There appears to be directional dependence along the NW to SE direction in the northern parts of the heatmap (Figure 3.3a). The directional sample semivariograms do not appear to be different until the distance is greater than 200,000 meters. Semivariogram estimates at large distances can be unreliable because there are fewer pairs of sampling locations to estimate this value than at short distances. Likewise, directional semivariograms are less reliable than a uni-directional semivariogram because fewer pairs of sampling locations are used at each distance for directional estimation. The unreliability of the sample semivariograms at the larger distances, coupled with the lack of a measure of uncertainty, make it difficult to determine whether or not an assumption of isotropy is reasonable using a plot

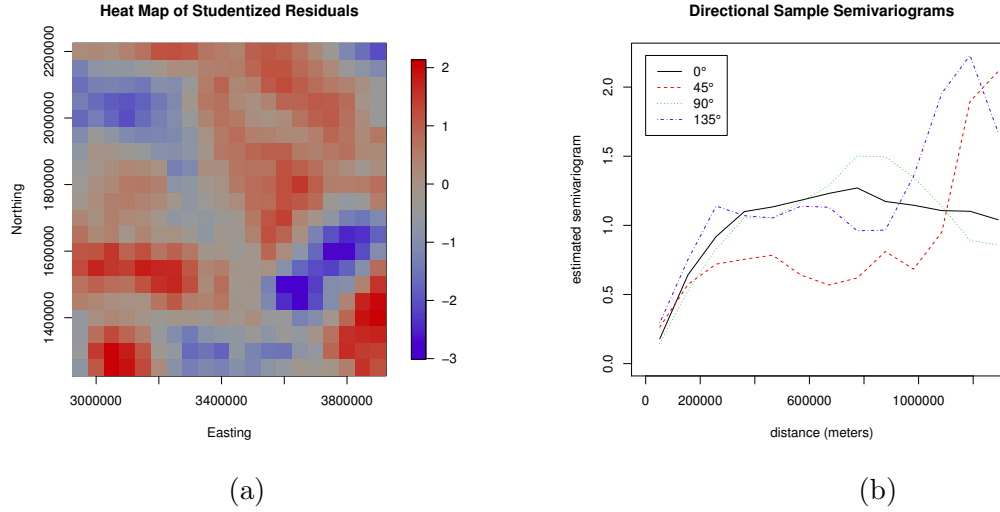


Figure 3.3: Graphical assessments of isotropy in the studentized residuals from the WRFG temperature data. The appearance of elongated areas of similar residual values in the heat map (Figure 3.3a) indicates that the process may be anisotropic. The directional semivariograms (Figure 3.3b) do not appear to exhibit differences, indicating that the process is isotropic. A nonparametric test of isotropy can assist in determining whether or not an assumption of isotropy is reasonable.

of the directional sample semivariograms. To gain more insight into the isotropy properties, we perform a nonparametric hypothesis test of isotropy using the residuals with the same choices for  $\mathbf{\Lambda}$ ,  $\mathbf{A}$ , and the window dimensions.

```
R> tr <- GuanTestGrid(spdata = resid.geo, delta = my.delta,
+   lagmat = mylags, A = myA, df = 2, window.dims = c(4,
+   4))
R> tr$p.value.finite

p.value.finite
0.2
```

Here the residuals do not provide evidence for anisotropy ( $p$  value  $> 0.05$ ). These results suggest that it may be appropriate to choose an isotropic covariance function to model the residuals. However, it is important to note that we have not included the effect of other potentially influential covariates such as elevation or water cover in the model for

temperature. Additionally, although we examined a  $20 \times 20$  subset of the data, the grid boxes still cover a large geographic region of the U.S., and thus an assumption of stationarity, which is needed for the asymptotic properties of the hypothesis test to hold, may not be reasonable.

### 3.5.2 Non-gridded Sampling Locations

The non-gridded data set used in this section describes monthly surface meteorology in a region of the state of Colorado and comes from the National Center for Atmospheric Research (NCAR). The data are available in the R package `fields`. For this example, our variable of interest is the log (mm) of the 30-year average of average yearly precipitation at 344 station locations during the time period 1968-1997.

Like the temperature data, our goal will be to model the mean effect of covariates and check for spatial dependence and potential anisotropy in the residuals. The first two steps of the analysis are to compute the yearly precipitation averages and convert the latitude/longitude coordinates to UTM coordinates. We divide the UTM coordinates by 100,000 so that distances are measured in hundreds of kilometers. Scaling the coordinates eases the choice of tuning parameters for the test in Guan et al. (2004). To meet the Gaussian assumption, we take the log transform of the average precipitation. Figure 3.4 displays quilt plots of the log precipitation values and the elevation of the stations.

```
R> data("COmonthlyMet")
R> sub30 <- CO.ppt[74:103, , ]
R> nstations <- 376
R> years <- 1968:1997
R> nyears <- length(years)
R> yr.avg <- matrix(data = NA, nrow = nstations, ncol = nyears)
R> for (i in 1:nyears) {
+   yr.dat <- sub30[i, , ]
+   yr.avg[, i] <- apply(yr.dat, 2, mean, na.rm = T)
```



```

+ }
R> avg30 <- apply(yr.avg, 1, mean, na.rm = T)
R> CO.loc.utm <- project(as.matrix(CO.loc),
+       "+proj=utm +zone=13 ellps=WGS84")/1e+05
R> quilt.plot(CO.loc.utm, log(avg30), col = two.colors(n = 256,
+       start = "blue3", end = "red3", middle = "gray60"),
+       legend.lab = "Precip (log mm)", legend.cex = 0.8,
+       legend.line = 2.2, xlab = "Easting", ylab = "Northing",
+       main = "Quilt Plot of log(precip)")
R> mp <- map("state", region = c("colorado", "wyoming",
+       "nebraska", "utah", "new mexico", "oklahoma"), plot = F)
R> states <- project(cbind(mp$x, mp$y),
+       "+proj=utm +zone=13 ellps=WGS84")/1e+05
R> points(states[, 1], states[, 2], type = "l", lwd = 1.5)

R> quilt.plot(CO.loc.utm, CO.elev, col = two.colors(n = 256,
+       start = "blue3", end = "red3", middle = "gray60"),
+       legend.lab = "Elevation (meters)", legend.cex = 0.8,
+       legend.line = 2.7, xlab = "Easting", ylab = "Northing",
+       main = "Quilt Plot of Elevation")
R> points(states[, 1], states[, 2], type = "l", lwd = 1.5)

```

Colorado has two distinct geographic regions: the mountainous region in the west and the plains region in the east. Figure 3.4b illustrates these two regions, and we can begin to notice a possible relationship between elevation and average precipitation. We explore the potential relationship between log precipitation and elevation using scatter plots (see Figure 3.5a).

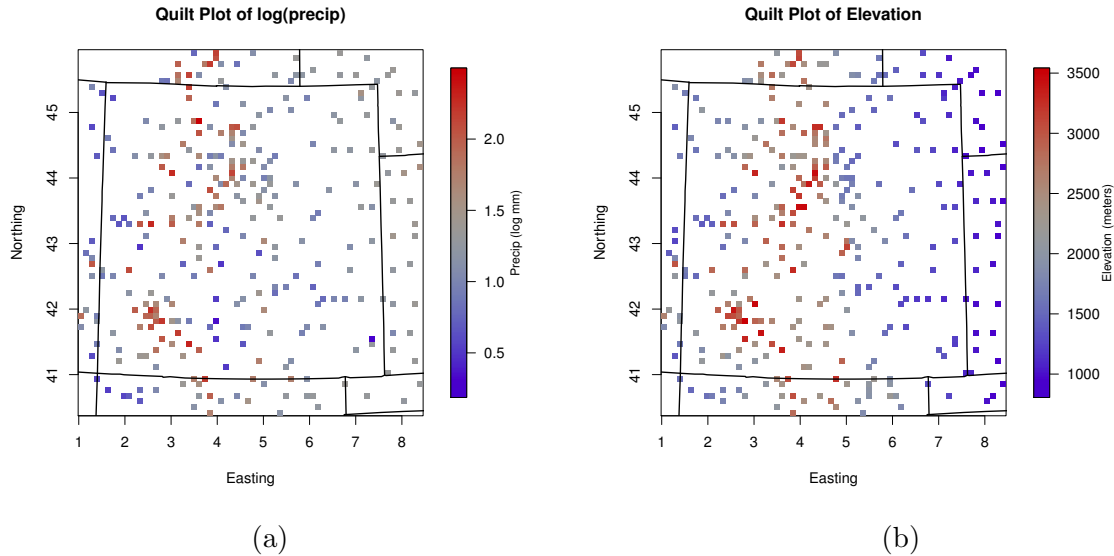


Figure 3.4: Quilt plots showing the locations of the weather stations in Colorado and surrounding region along with the log of average yearly precipitation (3.4a) and elevation (3.4b) at each station.

```
R> plot(CO.elev, log(avg30), xlab = "Elevation (meters)",
+       ylab = "Precip (log mm)",
+       main = "Scatter of log(precip) vs. Elevation")
R> m1 <- lm(log(avg30) ~ ns(CO.elev, df = 3))
R> summary(m1)
R> fits <- m1$fitted.values
R> bad <- is.na(avg30)
R> x <- CO.elev[which(!bad)]
R> lines(sort(x), fits[order(x)], lwd = 3, col = "red")

R> qqnorm(studres(m1))
R> abline(0, 1)
```

We fit a cubic smoothing spline via least squares to model the relationship between  $\log(\text{precipitation})$  and elevation. The estimate is shown in Figure 3.5a, and a QQ plot of residuals in Figure 3.5b indicates that a Gaussian assumption is reasonable. We will use the

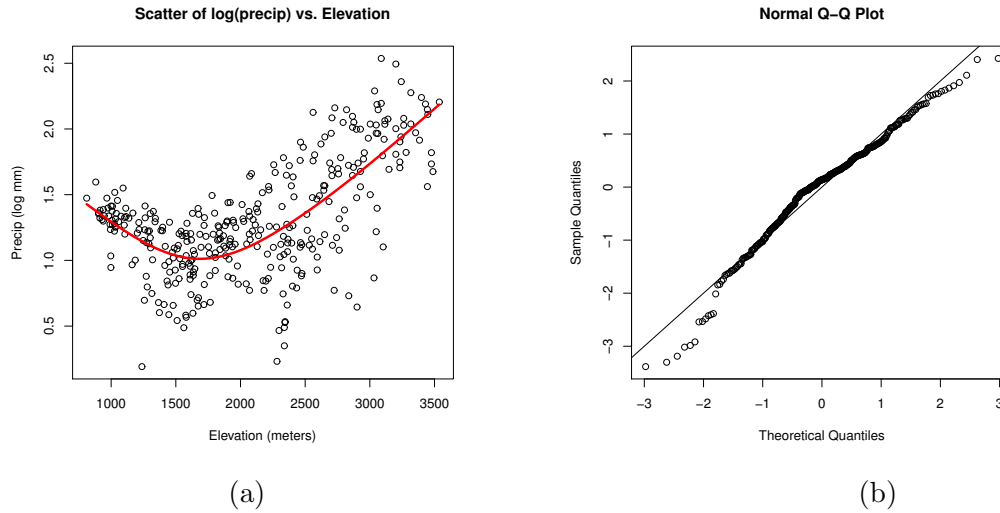


Figure 3.5: Results from the model relating  $\log(\text{precipitation})$  and elevation. Figure 3.5a displays the nonparametric fit relating elevation to  $\log(\text{precipitation})$ . Figure 3.5b shows the QQ plot of studentized residuals from the nonparametric fit.

residuals from this model to check for remaining spatial dependence and potential anisotropy. We use `variog4` to estimate directional sample semivariograms.

```
R> precip.resid <- cbind(CO.loc.utm[which(!bad), ], 1,
+      CO.loc.utm[which(!bad), ], 2, studres(m1))
R> precip.geo <- as.geodata(precip.resid)
R> plot(variog4(precip.geo), xlab = "distance (100s of km)",
+      ylab = "estimated semivariogram", legend = F)
R> legend("bottomright", legend = c(expression(0 * degree),
+      expression(45 * degree), expression(90 * degree),
+      expression(135 * degree)), col = 1:4, lty = 1:4)
R> title("Directional Sample Semivariograms")
```

The increase, followed by a leveling off, of the semivariogram values as distance increases in Figure 3.6 indicates that there is spatial dependence remaining in the data. We notice that the  $0^\circ$  semivariogram appears to be slightly different than the other three, but there is no measure of uncertainty, so we cannot determine if the differences are statistically signifi-

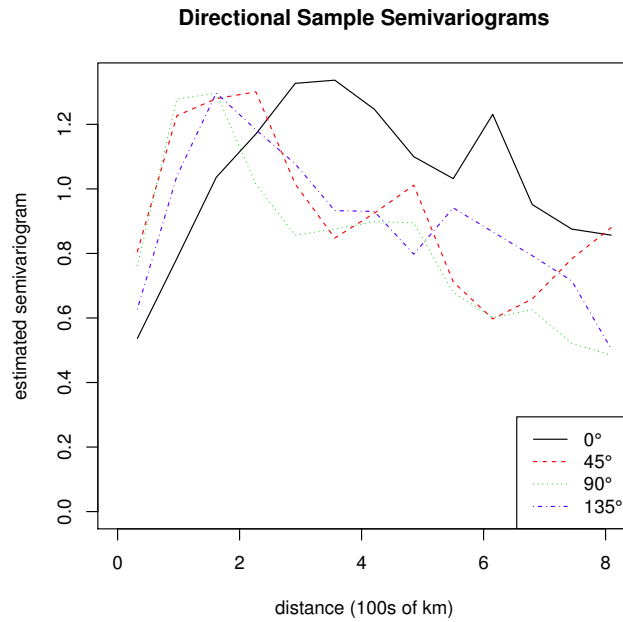


Figure 3.6: A graphical assessment of isotropy in the studentized residuals from the model relating  $\log(\text{precipitation})$  to elevation generated by using the `geoR` function `variog4`. The directional sample semivariogram in the  $0^\circ$  direction appears to be different from the other three at distances less than 2. Because there is no measure of uncertainty, it can be difficult to determine whether or not an assumption of isotropy is reasonable.

cant. The sample semivariograms also suggest the presence of a large amount of small scale variation, often called a nugget effect, in the data. This can be seen by noting the ratio of semivariogram values at the shortest observed distance to the semivariogram values at the longest reliable distance. The large jumps and decrease in the estimated semivariogram values in Figure 3.6 indicate that semivariogram estimates become unreliable beyond a distance of two. The sample semivariogram values in Figure 3.6 are approximately 0.7 at the shortest distance and approximately 1.1 at a distance of two. This suggests that approximately  $63\% \approx (0.7/1.1)100\%$  of the variability in the data is due to small scale variation. We note the apparently large nugget effect because this small scale variation is detrimental to the size and power of nonparametric tests of isotropy (Weller and Hoeting, 2016). Despite the small scale variation, we will proceed with nonparametric hypothesis tests to assist in determining if an assumption of isotropy is reasonable.

There are two procedures for testing isotropy in non-gridded data available in `spTest`: Guan et al. (2004) and Maity and Sherman (2012). To choose between these two, we need to decide whether or not it is reasonable to assume that sampling locations are uniformly distributed on the sampling domain. The methods for non-gridded data from Guan et al. (2004) rely on the assumption that sampling locations are uniformly distributed while Maity and Sherman (2012) can be used on any general sampling design. To check this assumption, we can turn to methods from the spatial point process literature to perform a test of complete spatial randomness (CSR) (i.e., a uniform spatial distribution) for the sampling locations. Methods for testing CSR are available in the R package `spatstat` (Baddeley and Turner, 2005). For brevity, we do not display the results of the CSR test here, but note that they do not provide evidence against the assumption of CSR for these sampling locations so either test of isotropy can be used.

For both Guan et al. (2004) and Maity and Sherman (2012), we need to choose the lag set,  $\mathbf{A}$ , and the contrast matrix,  $\mathbf{A}$ . Because semivariogram estimates appear to be unreliable at distances greater than two, we should choose lags having Euclidean distance less than this

distance. We choose the lag set

$$\mathbf{A} = \{\mathbf{h}_1 = (0.60, 0), \mathbf{h}_2 = (0, 0.60), \mathbf{h}_3 = (0.45, 0.45), \mathbf{h}_4 = (-0.45, 0.45)\},$$

and again we use the matrix  $\mathbf{A}$  in (18). The set  $\mathbf{A}$  corresponds to lags in the  $0^\circ, 90^\circ, 45^\circ, 135^\circ$  directions, respectively, having Euclidean distances of  $\|\mathbf{h}_1\| = \|\mathbf{h}_2\| = 0.60$  (60 km) and  $\|\mathbf{h}_3\| = \|\mathbf{h}_4\| \approx 0.64$  (64 km).

```
R> mylags <- rbind(c(0.60, 0), c(0, 0.60), c(0.45, 0.45), c(-0.45, 0.45))
R> myA <- rbind(c(1, -1, 0, 0), c(0, 0, 1, -1))
```

The next step in implementing the methods from Guan et al. (2004) and Maity and Sherman (2012) is to determine the size of the moving windows and the block size, respectively, used to estimate the asymptotic variance-covariance matrix,  $\mathbf{\Sigma}$ . The moving window is shifted over the sampling region, creating subblocks of data used to estimate  $\mathbf{\Sigma}$ . Likewise, for the test in Maity and Sherman (2012), the block size is used to implement the grid-based block bootstrap [GBBB] (Lahiri and Zhu, 2006).

There are two steps in determining the appropriate window/block size for non-gridded sampling locations. First, we place a grid over the sampling domain; second, we specify scaling parameters that will define the window/block size in terms of that grid. We should complete this two step process while keeping three goals in mind: (1) the number of sampling locations per window/block, denoted  $n_b$ , should be approximately  $\sqrt{n}$  (Weller and Hoeting, 2016); (2) the windows/blocks should have the same orientation (i.e., square or rectangle) as the entire sampling domain; and (3) the scaling parameters should be compatible with the dimensions of the underlying grid.

For the Colorado precipitation data, recall that one unit of distance equals 100 km. The dimensions of the sampling region are approximately  $7.3 \times 5.5$  (width  $\times$  height), providing a total area of 40.15. For  $n = 344$  uniformly distributed sampling locations, we expect approximately  $344/40.15 = 8.6$  sampling locations per unit area. Recalling goal (1), we seek

to construct windows/blocks with  $n_b \approx \sqrt{344} = 18.5$  sampling locations, or equivalently, windows/blocks with an area of approximately  $18.5/8.6 = 2.15$ . Goal (2) indicates we want to create rectangular windows/blocks with slightly larger width than height, and (3) says that if our grid divides the  $x$ -axis into 12 grid cells, then the scaling parameter defining the width of the window/block should be 3 or 4 because those numbers evenly divide 12. For the CO precipitation data, if we choose our grid to divide the  $x$ -axis into 16 cells and the  $y$ -axis into 12 cells, we have a grid with  $(x, y)$  spacing of roughly  $(7.3/16, 5.5/12) \approx (0.46, 0.46)$ . The resulting grid is plotted in Figure 3.7. Then, choosing our scaling parameters to be  $4 \times 3$ , we have windows/blocks with dimensions of approximately  $(4)(0.46) \times (3)(0.46) = 1.84 \times 1.38$  and area of  $(1.84)(1.38) \approx 2.54$ , or equivalently with an expected number of points per block of  $n_b = (2.54)(8.6) \approx 21.8$ .

```
R> quilt.plot(precip.resid[, 1:2], precip.resid[, 3],
+           col = two.colors(n = 256,
+           start = "blue3", end = "red3", middle = "gray60"),
+           xlab = "Longitude", ylab = "Latitude", xlim = c(0.75,
+           8.65), ylim = c(40.1, 46.2))
R> title("Quilt Plot of Residuals and Grid Used for Subsampling")
R> tol <- 0.02
R> my.xlims <- c(min(precip.resid[, 1]) - tol, max(precip.resid[,
+           1]) + tol)
R> my.ylims <- c(min(precip.resid[, 2]) - tol, max(precip.resid[,
+           2]) + tol)
R> xlen <- my.xlims[2] - my.xlims[1]
R> ylen <- my.ylims[2] - my.ylims[1]
R> my.grid.spacing <- c(xlen/16, ylen/12)
R> xgrid <- seq(my.xlims[1], my.xlims[2], by = my.grid.spacing[1])
R> ygrid <- seq(my.ylims[1], my.ylims[2], by = my.grid.spacing[2])
```

**Quilt Plot of Residuals and Grid Used for Subsampling**

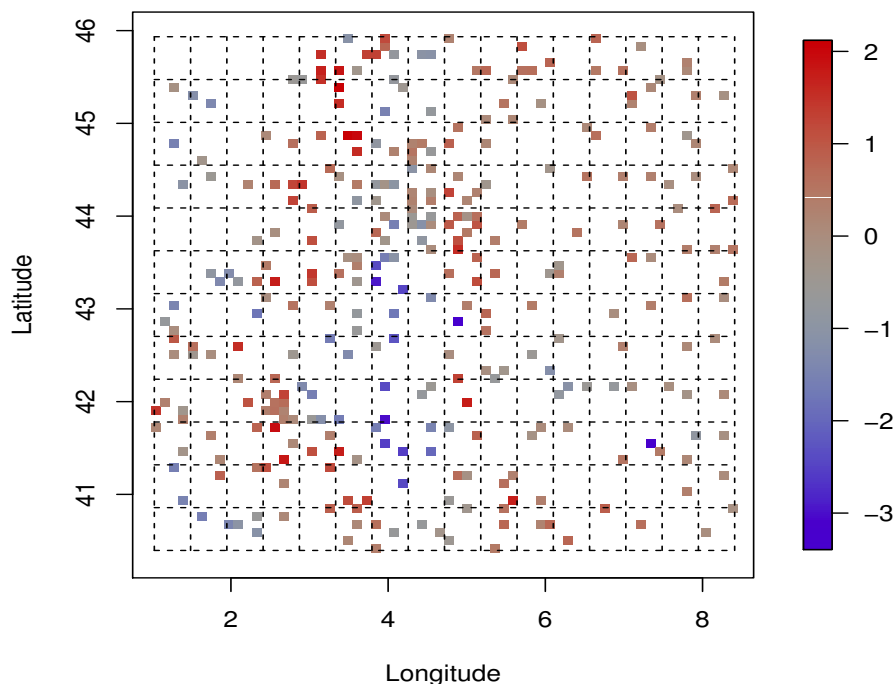


Figure 3.7: Quilt plot of the studentized residuals from the model relating elevation to  $\log(\text{precipitation})$  for the weather station locations. The grid placed on the region that is used to define the moving windows in Guan et al. (2004) and block size in Maity and Sherman (2012) is also shown. Because the sampling locations are not gridded, it can be difficult to assess isotropy properties via a quilt plot.

```
R> segments(x0 = xgrid, y0 = min(my.ylims), y1 = max(my.ylims),
+          lty = 2)
R> segments(x0 = min(my.xlims), y0 = ygrid, x1 = max(my.xlims),
+          lty = 2)
```

For the functions `GuanTestUnif` and `MaityTest`, the upper and lower limits of the sampling region in the  $x$  and  $y$  directions are given by the `xlims` and `ylims` arguments. Note that the values defining the upper and lower limits should be slightly larger than the minimum and maximum observed  $x$  and  $y$  coordinates. The horizontal and vertical spacing, respectively, of the grid laid on the sampling region is defined by the two values in `grid.spacing`.



The horizontal and vertical scaling parameters that define the size of the moving windows in `GuanTestUnif` and blocks in `MaityTest` in terms of the underlying grid are given by the `window.dims` and `block.dims` arguments, respectively. We recommend using visualizations of different grid choices and algebraic calculations, as done above, to assist in choosing a grid and the window/block dimensions. When the scaling parameters defining the moving window or block dimensions are not compatible with the number of rows or columns of gridded sampling locations or the dimensions of the grid laid on the sampling region for non-gridded locations, the functions in `spTest` will print a warning message because they do not currently handle partial (incomplete) blocks. Likewise, if the chosen window or block dimensions for non-gridded sampling locations creates (sub)blocks of data with few or no sampling locations, the functions `GuanTestUnif` and `MaityTest` will discard (sub)blocks that do not have enough sampling locations and print a warning message. The  $p$  value of the hypothesis test will be sensitive to the choice of moving window and block dimensions. See Weller and Hoeting (2016) and the original works (Guan et al., 2004; Maity and Sherman, 2012) for more recommendations on choosing these values.

The next step for implementing the test in Guan et al. (2004) is choosing the smoothing (bandwidth) parameters for smoothing over lags on the entire domain and within each sub-block created by the moving windows. The smoothing parameters should be chosen based on the number and density of the sampling locations with larger values of the smoothing parameter inducing higher levels of smoothing, i.e., allowing a greater distance and direction tolerance. In our experience, smoothing parameter values between 0.6 and 0.9 tend to produce reasonable results when using a standard normal Gaussian smoothing kernel truncated at 1.5. However, the  $p$  value of the hypothesis test will change with the bandwidth. For this example, we choose a bandwidth of `h = 0.70` for smoothing over lags on the entire domain, and a bandwidth of `subblock.h = 0.85` for smoothing over lags on the subblocks of data created by the moving window. Choosing a larger bandwidth for the subblocks equates to allowing for a larger lag distance and direction tolerance, which is needed for subblocks that

have few sampling locations. We also use the default Gaussian smoothing kernel (`kernel = "norm"`) truncated at 1.5 (`truncation = 1.5`). Because the sample size is less than 500, we use a finite sample adjustment to approximate the  $p$  value (Guan et al., 2004; Weller and Hoeting, 2016).

Finally, for the test in Maity and Sherman (2012) we need to choose the number of bootstrap resamples that will be used in the GBBB procedure to estimate  $\Sigma$ . We recommend using at least 50 bootstrap samples; however, the bootstrapping procedure is computationally intensive. We choose `nBoot = 100` bootstrap samples for our example, and we note that the number of bootstraps does not affect the precision of the  $p$  value, which is computed via the asymptotic  $\chi^2$  distribution. Having determined values for the different options, we can now perform the hypothesis tests. For reproducibility of the bootstrap in the `MaityTest` function, we set the random seed.

```
R> myh <- 0.7
R> myh.sb <- 0.85
R> tr.guan <- GuanTestUnif(spdata = precip.resid, lagmat = mylags,
+   A = myA, df = 2, h = myh, kernel = "norm", truncation = 1.5,
+   xlims = my.xlims, ylims = my.ylims, grid.spacing = my.grid.spacing,
+   window.dims = c(4, 3), subblock.h = myh.sb)
R> tr.guan$p.value.finite

p.value.finite
0.04615385

R> set.seed(2016)
R> tr.maity <- MaityTest(spdata = precip.resid, lagmat = mylags,
+   A = myA, df = 2, xlims = my.xlims, ylims = my.ylims,
+   grid.spacing = my.grid.spacing, block.dims = c(4,
```

```
+          3), nBoot = 100)
R> tr.maity$p.value

p.value.chisq
0.021859
```

For both of the tests, the data provide evidence in favor of anisotropy ( $p$  value  $< 0.05$ ). Thus, an isotropic model may be appropriate for modeling the residuals. Additionally, the apparent anisotropy may also be present due to unaccounted for covariates (e.g., northing/easting coordinates).

### 3.6 Discussion

Choosing a covariance function is an important step in modeling spatially-referenced data and a variety of choices for the covariance function are available (e.g., anisotropy, nonstationarity, parametric forms). The R package `spTest` implements several nonparametric tests for checking isotropy properties which avoid specifying a parametric form for the covariance function. Weller and Hoeting (2016) perform a simulation study comparing the empirical size and power of the methods for different degrees of anisotropy. They find that methods from Guan et al. (2004) tend to outperform the competitor for gridded and non-gridded data.

One concern regarding the methods in `spTest` is that they tend to have low power when the anisotropy is weak and the data are not gridded (Weller and Hoeting, 2016; Guan et al., 2004; Maity and Sherman, 2012). A second concern is that the results of the tests are potentially sensitive to user choices, for example, the moving window size and bandwidth in the method from Guan et al. (2004). The optimal choices for these values is still an open question. Weller and Hoeting (2016) offer further recommendations for how to choose the user defined values, such as the window size and bandwidth, based on simulated data. Finally, as noted earlier, the size and power of the methods are adversely affected by the

presence of small scale variation (nugget effect). Because of these concerns, we recommend using the nonparametric methods in conjunction with graphical techniques.

An implicit assumption of the methods discussed in this chapter is ergodicity of the spatial process, an assumption that is difficult to verify (Cressie, 1993, pg. 57-58). However, there are two important assumptions which practitioners should consider. The first is an assumption of strict stationarity (1). While this assumption is also difficult to check, it follows from assuming the RF is weakly stationary and Gaussian (4). The assumption of Gaussian data lies at the heart of many spatial analyses (Gelfand and Schliep, 2016) and is easily checked with a QQ plot. The assumption of weak stationarity may be questionable for spatial data over large geographic regions and methods have been developed for testing this assumption (see, e.g., Corstanje et al., 2008; Fuentes, 2005; Jun and Genton, 2012; Bandyopadhyay and Rao, 2015). The second assumption required is a mixing condition that states the dependence between observations goes to 0 at large distances (Hall and Patil, 1994; Sherman and Carlstein, 1994). In the case of a stationary GRF, this condition is met when the covariance function  $C(\mathbf{h})$  is 0 for sufficiently large  $\|\mathbf{h}\|$ , which also implies ergodicity (Cressie, 1993, pg. 58). One way to check this assumption in practice is to look for a leveling off of the sample covariogram or semivariogram values as distance increases, indicating that the data are nearly independent at large distances.

After determining whether or not an assumption of isotropy is reasonable, we can choose a parametric or nonparametric model for the covariance function. Weller and Hoeting (2016) further illustrate the role of nonparametric tests of isotropy in the process of modeling spatially-referenced data. We have demonstrated how graphical techniques and the functions available in the R package **spTest** can be used in a complementary role to check for anisotropy.

## CHAPTER 4

# A SPECTRAL DOMAIN NONPARAMETRIC TEST OF SPATIAL ISOTROPY

### 4.1 Introduction

Modeling spatially-referenced data typically requires choosing a covariance function to describe spatial dependence. Due to numerous recent developments, there are a wide variety of covariance models available. A common simplifying assumption is that the covariance function is isotropic. Isotropy implies that the dependence between any two observations depends only on the Euclidean distance between their sampling locations and not on their relative orientation. Graphical techniques, such as directional sample variograms, are typically used to check the assumption of isotropy, but their interpretation is subjective and open to interpretation. Thus, there is a need for more objective assessments of this assumption.

We develop a nonparametric hypothesis test of the assumption of isotropy using the spectral representation of the covariance function. While spectral methods have been used for some time in the time series context, there have been a number of more recent developments on the use of spectral methods for the analysis of spatial random fields (RFs) (see, e.g., Fuentes, 2001, 2007). By using a nonparametric test, we avoid the problems of choosing a covariance function from the plethora of available models, misspecification of the covariance function, and the need to estimate the covariance function under the null and alternative hypotheses.

There are several methods that use nonparametric estimation of the spectral density, the Fourier transform of the covariance function, to test second order properties. For example, Lu and Zimmerman (2005) and Scaccia and Martin (2005, 2002) test symmetry properties

of the covariance function for lattice processes. Fuentes (2005) develops a nonparametric, spatially varying spectral density to test for nonstationarity. Van Hala et al. (2014) use a spatial frequency domain empirical likelihood (SFDEL) to test isotropy and separability of the covariance function. None of these methods were developed with the primary goal of testing isotropy. Thus, while many of the concepts that are used for our test are not new, their combined use for testing isotropy has not been previously exploited or explored. Our method is the first to use the spectral representation of the covariance function with the primary goal of testing the assumption of isotropy.

Guan et al. (2004), Maity and Sherman (2012), and others have developed nonparametric tests of isotropy in the spatial domain. Two drawbacks of these methods are that they require the choice of tuning parameters, and they can be computationally intensive (Weller and Hoeting, 2016). For these spatial domain tests, the tuning parameters can affect test results and their optimality remains an open question. As a result, there is a need to develop methods that are tuning-parameter free and less computationally expensive. The spectral method we develop addresses both of these issues as it avoids the need to choose tuning parameters and is less computationally intensive than methods in the spatial domain. For a review of nonparametric methods for testing isotropy and symmetry, see Weller and Hoeting (2016).

The rest of this chapter is organized as follows: in Section 4.2 we outline important background on spectral methods and provide the motivating theorem for our test; Section 4.3 describes the periodogram, the nonparametric estimator of the spectral density, and its properties; Section 4.4 develops the proposed test statistics and describes associated challenges; Sections 4.5 and 4.6 offer a simulation study and application of the new method, respectively; we discuss the new test in Section 4.7. Finally, we discuss the problem of aliasing, develop a graphical diagnostic for aliasing, and explore the effects of aliasing on covariance estimation in Section 4.8.

## 4.2 Background and Definitions

First we provide a review of notation, definitions, and results regarding spatial RFs and their spectral representation. We follow the introductions given by Schabenberger and Gotway (2004) and Fuentes and Reich (2010); see these sources for additional background. Below we assume dimension  $d = 2$ , although many of the results hold for the more general case of  $d \geq 2$ . When working in the spectral domain, one important distinction is whether the spatial process is a geostatistical (continuous) process or lattice (discrete) process. We focus on the case of a geostatistical process observed at gridded sampling locations.

### 4.2.1 Spectral Representation of the Covariance Function

Let  $\{Y(\mathbf{s}) : \mathbf{s} \in \mathcal{D} \subseteq \mathbb{R}^2\}$  be a mean square continuous, second order (weakly) stationary, mean 0, RF with covariance function  $C(\mathbf{h})$ ,  $\mathbf{h} = (h_1, h_2)^\top$ . For the purposes of investigating the second order properties, we are interested in the spectral representation of the covariance function of the RF, called the spectral density function and denoted  $f(\boldsymbol{\omega})$ ,  $\boldsymbol{\omega} = (\omega_1, \omega_2)^\top$ . To ensure the existence of the spectral density, we assume  $C(\mathbf{h}) \rightarrow 0$  as  $\mathbf{h} \rightarrow \infty$ , and  $C(\mathbf{h})$  is absolutely integrable, i.e.,  $\int_{\mathbb{R}^2} |C(\mathbf{h})| d\mathbf{h} < \infty$ . The spectral representation of the covariance function  $C(\mathbf{h})$  in  $\mathbb{R}^2$  is given by

$$C(\mathbf{h}) = \int_{\mathbb{R}^2} \exp(i\mathbf{h}^\top \boldsymbol{\omega}) f(\boldsymbol{\omega}) d\boldsymbol{\omega}, \quad (21)$$

and the spectral density function is given by

$$f(\boldsymbol{\omega}) = \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} \exp(-i\boldsymbol{\omega}^\top \mathbf{h}) C(\mathbf{h}) d\mathbf{h}, \quad (22)$$

so that  $C(\mathbf{h})$  and  $f(\boldsymbol{\omega})$  form a Fourier transform pair, (21) and (22). Properties of the covariance function will imply properties of the spectral density. Namely, if the covariance function is isotropic, then the spectral density function is isotropic. Thus, second order

properties of a weakly stationary RF can be explored via either the covariance function or the spectral density function.

When modeling a spatial RF, a common simplifying assumption about the covariance function is that it is isotropic.

**Definition 5.** *A second order stationary spatial process is isotropic if the covariance function  $C(\mathbf{h})$  of the spatial process depends on the lag vector  $\mathbf{h}$  only through its Euclidean length,  $\|\mathbf{h}\|$ , i.e.,  $C(\mathbf{h}) = C(h_1, h_2) = C_0(h)$ , where  $h = \|\mathbf{h}\|$  and  $C_0(\cdot)$  is a valid covariance function taking a univariate argument.*

A spatial process that is not isotropic is called anisotropic.

As one example, the isotropic Matérn covariance function is

$$C(\mathbf{h}) = C_0(h) = \frac{\pi\phi}{2^{\nu-1}\Gamma(\nu+1)\alpha^{2\nu}}(\alpha h)^\nu \mathcal{K}_\nu(\alpha h) \quad (23)$$

with parameters  $(\nu, \alpha, \phi)$  all greater than 0, and where  $\mathcal{K}_\nu$  is a modified Bessel function of the third kind. The parameter  $\alpha$  governs the range of spatial correlation,  $\nu$  determines the smoothness of the process, and  $\phi$  is proportional to the process variance times the range to the  $2\nu^{th}$  power (Fuentes and Reich, 2010). The corresponding isotropic Matérn spectral density function is given by

$$f(\boldsymbol{\omega}) = f(\omega_1, \omega_2) = f_0(\omega) = \phi(\alpha^2 + \omega^2)^{-\nu-1}, \quad (24)$$

where  $\omega = \|\boldsymbol{\omega}\|$  and  $f_0(\cdot)$  is a spectral density function taking a univariate argument. Two examples of the Matérn covariance and spectral density function are plotted in Figure 4.1. In Section 4.5, we simulate data from RFs with isotropic (23) and anisotropic Matérn covariance functions.

Due to the one to one relationship between the covariance function and spectral density function in (21) and (22), we can use the spectral density as an alternative to the covariance



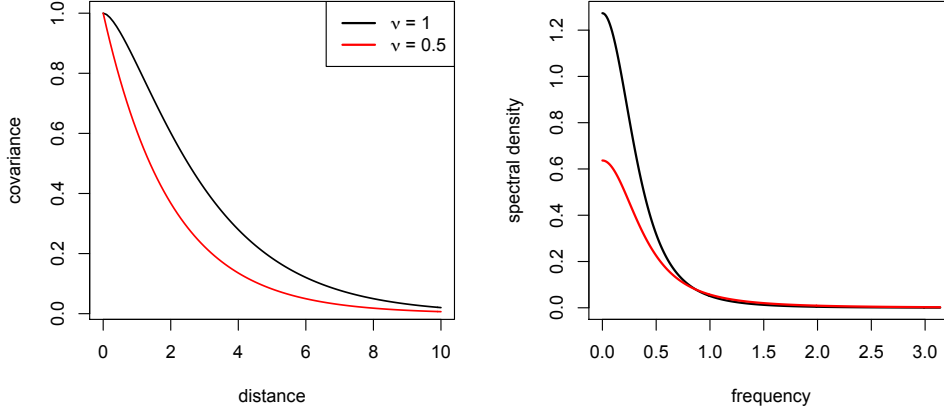


Figure 4.1: Isotropic Matérn covariance functions (left) and spectral density functions (right). The red functions correspond to a Matérn covariance function with smoothness  $\nu = 0.5$  (i.e., an exponential covariance function) while the black functions correspond to  $\nu = 1$ .

function for exploring directional dependence properties (see, e.g., Lu and Zimmerman, 2005; Scaccia and Martin, 2005, 2002). To develop a test of isotropy in the spectral domain, we rely on the following theorem:

**Theorem 1.** (Adler, 1981; Yaglom, 1987; Fuentes, 2005, 2013) *Let  $\mathbf{Y}$  be a real, second order stationary random field in  $\mathbb{R}^2$  with valid, continuous covariance function  $C(\mathbf{h})$ ,  $\mathbf{h} = (h_1, h_2)^\top$ . Assume that  $\int_{\mathbb{R}^2} |C(\mathbf{h})|^2 d\mathbf{h} < \infty$ . Then the covariance function  $C(\mathbf{h})$  is isotropic,  $C(\mathbf{h}) = C_0(h)$ ,  $h = \|\mathbf{h}\| = (h_1^2 + h_2^2)^{1/2}$ , if and only if the spectral density function is isotropic,  $f(\boldsymbol{\omega}) = f_0(\omega)$ ,  $\omega = \|\boldsymbol{\omega}\| = (\omega_1^2 + \omega_2^2)^{1/2}$ .*

Theorem 1 implies that we can investigate the isotropy of the covariance function by examining the isotropy of the spectral density function. To perform this investigation via a hypothesis test, we need an estimator of the spectral density function (22).

### 4.3 Periodogram

In this section, we describe the periodogram, an estimator of the spectral density function (22) and consider its asymptotic properties. We observe the continuous process  $\mathbf{Y}$  at a finite set of locations. We assume the sampling locations form a rectangular grid. If we have  $n$  gridded sampling locations with row and column spacing  $\Delta = (\delta_1, \delta_2)^\top$ , then, due to the aliasing phenomenon, we restrict estimation of the spectral density to frequencies within the interval  $\omega \in [-\pi/\delta_1, \pi/\delta_1) \times [-\pi/\delta_2, \pi/\delta_2)$ . The aliasing phenomenon refers to the fact that it is not possible to distinguish oscillations that differ by an integer multiple of  $2\pi$  when observations occur at a discrete set of points. More generally, aliasing means that it is not possible to recover a continuous signal from a finite set of observation locations (see Fuentes and Reich (2010, pg. 58-62) for more details). We can then write the spectral density of the lattice process in terms of the spectral density of the continuous process as

$$f_{\Delta}(\omega) = \sum_{\mathbf{z} \in \mathbb{Z}^2} f\left(\omega + \frac{2\pi\mathbf{z}}{\Delta}\right) = \sum_{z_1 \in \mathbb{Z}} \sum_{z_2 \in \mathbb{Z}} f\left(\omega_1 + \frac{2\pi z_1}{\delta_1}, \omega_2 + \frac{2\pi z_2}{\delta_2}\right). \quad (25)$$

The periodogram, denoted  $I(\omega)$ , is a nonparametric estimator of the aliased spectral density (25) and can be computed by taking the Fourier transform of the sample covariance,

$$I(\omega) = \frac{\delta_1 \delta_2}{(2\pi)^2} \sum_{h_1=-n_1+1}^{n_1-1} \sum_{h_2=-n_2+1}^{n_2-1} \widehat{C}(\Delta \circ \mathbf{h}) \exp(-i(\Delta \circ \mathbf{s})^\top \omega), \quad (26)$$

where  $\mathbf{v} \circ \mathbf{w} = (v_1 w_1, v_2 w_2)$  is the Hadamard product. The sample covariance  $\widehat{C}(\Delta \circ \mathbf{h})$  is defined as

$$\widehat{C}(\Delta \circ \mathbf{h}) = \frac{1}{n} \sum_{s_1=1}^{n_1} \sum_{s_2=1}^{n_2} Y(\Delta \circ \mathbf{s}) Y(\Delta \circ (\mathbf{s} + \mathbf{h})). \quad (27)$$

where  $n = n_1 \times n_2$ , the number of rows ( $n_1$ ) times the number of columns ( $n_2$ ).

The asymptotic properties of the spatial periodogram  $I(\omega)$  depend on a number of factors, including (1) the frequency  $\omega$ , (2) the shape of the sampling domain, (3) whether or not the

spatial process is discrete or continuous and (4) for a geostatistical process, the type of spatial asymptotics used (Bandyopadhyay and Lahiri, 2009). Accordingly, there is a rich literature on the asymptotic properties of Fourier transforms (e.g., Brillinger, 1970, 1974; Pagano, 1971; Guyon, 1982; Stein, 1995; Crujeiras, 2006; Lim and Stein, 2008; Bandyopadhyay and Lahiri, 2009). Finite sample properties of the periodogram have also been derived (e.g., see Kim and Fuentes, 2000; Fuentes and Smith, 2001; Crujeiras, 2006; Porcu et al., 2009). Next, we provide a brief description of different spatial asymptotic schemes and discuss the asymptotic properties of  $I(\boldsymbol{\omega})$  relevant for our test of isotropy.

Under increasing domain asymptotics, the size of the sampling region grows without bound, but the spacing between sampling locations does not change. This asymptotic scheme is typically used for discrete (lattice) spatial processes (e.g., see Brillinger, 1981). Infill, or fixed-domain, asymptotics assume that the size of the sampling domain is fixed and the spacing between sampling locations goes to zero. Infill asymptotics are often used when considering continuous (geostatistical) spatial processes (e.g., see Stein, 1995; Zhang, 2004). Finally, under mixed increasing domain and infill asymptotics, the size of the sampling domain grows without bound and the spacing between sampling locations goes to zero (e.g., see Fuentes, 2002).

Fuentes (2001, 2002) and Fuentes and Reich (2010) provide theoretical moment properties of the periodogram (26) under different asymptotic schemes. Under increasing domain asymptotics, the aliasing problem does not diminish as the sample size increases so that the periodogram (26) is always estimating the aliased spectral density (25). Under fixed domain asymptotics, we gain more information about the covariance function at short and medium distance spatial lags than at the longest lags as the sample size increases. Because the periodogram is computed as the Fourier transform of the sample covariance function (27) over all observed spatial lags, neither increasing domain or fixed domain asymptotics alone is satisfactory for considering periodogram asymptotics. As a result, we follow Fuentes (2002) and adopt a mixed asymptotic scheme to ensure an accurate nonparametric estimate

of the spectral density function (22). Like Fuentes (2002), we also assume a Gaussian RF. Under the mixed scheme, the asymptotic expected value and variance of  $I(\boldsymbol{\omega})$  are  $f(\boldsymbol{\omega})$  and  $f^2(\boldsymbol{\omega})$ , respectively. Periodogram values at fixed, distinct frequencies are also asymptotically uncorrelated. This asymptotically uncorrelated property of the periodogram makes the spectral domain attractive for doing computations such as fitting parametric models (Fuentes and Reich, 2010). In summary, under mixed asymptotics, the periodogram has the following asymptotic properties:

$$\lim_{n \rightarrow \infty} \begin{cases} \mathbb{E}(I(\boldsymbol{\omega})) = f(\boldsymbol{\omega}) \text{ for } \boldsymbol{\omega} \in [-\pi/\Delta, \pi/\Delta]^2, \\ \text{Var}(I(\boldsymbol{\omega})) = f^2(\boldsymbol{\omega}), \text{ and} \\ \text{Cov}(I(\boldsymbol{\omega}_1), I(\boldsymbol{\omega}_2)) = 0 \text{ for } \boldsymbol{\omega}_1 \neq \boldsymbol{\omega}_2. \end{cases} \quad (28)$$

These results indicate that the periodogram is an asymptotically unbiased, but inconsistent estimator of the periodogram. In both the time series and spatial settings, the typical approach to developing a consistent estimator of the spectral density at frequency  $\boldsymbol{\omega}$  is to smooth (e.g., average) the periodogram values in a small frequency interval containing  $\boldsymbol{\omega}$  (e.g., see Brockwell and Davis, 2009, pg. 124-125). In our development of a test statistic for testing isotropy, we choose not to smooth the periodogram for two reasons. First, if the spatial process is anisotropic, smoothing may result in some loss of the directional signal in the periodogram. Second, smoothing introduces dependence between periodogram values at distinct, but nearby, frequencies. We rely on the result that periodogram values at distinct frequencies are approximately uncorrelated in deriving the distribution of our test statistic. Results of our test may vary when using the smoothed periodogram, and its use for our test is an area of future work.

Finally, we also want to know the asymptotic distribution of the periodogram under the mixed asymptotic scheme. To this end, we prove the following new result which gives the asymptotic distribution of the periodogram.

**Theorem 2.** Let  $\{Y(\mathbf{s}) : \mathbf{s} \in \mathbb{R}^2\}$  be a mean 0, weakly stationary random field with covariance function  $C(\mathbf{h})$ , where the covariance function satisfies  $|C(\mathbf{h})| \rightarrow 0$  as  $\|\mathbf{h}\| \rightarrow 0$  and  $\int_{\mathbb{R}^2} |C(\mathbf{h})| d\mathbf{h} < \infty$ . Let  $f(\boldsymbol{\omega})$  be the spectral density associated with  $C(\mathbf{h})$  satisfying  $\int_{\mathbb{R}^2} |f(\boldsymbol{\omega})| d\boldsymbol{\omega} < \infty$ . Assume that  $Y$  satisfies conditions (A.1) and (A.2) of Bandyopadhyay and Lahiri (2009, pg. 228), and is observed at  $n$  sampling locations on a  $n_1 \times n_2$  rectangular grid ( $n = n_1 n_2$ ). Let  $\delta = \delta_1 = \delta_2$  denotes the spacing between rows and columns. Under the mixed asymptotic scheme (Fuentes, 2002), assume that  $\delta \rightarrow 0$ ,  $n_1 \rightarrow \infty$ ,  $n_2 \rightarrow \infty$ ,  $n_1/n_2 \rightarrow c$  for a constant  $c > 0$ ,  $\delta n_1 \rightarrow \infty$ , and  $\delta n_2 \rightarrow \infty$ .

Under these conditions and for frequencies  $\boldsymbol{\omega}_n \rightarrow \boldsymbol{\omega} \in \mathbb{R}^2 \setminus \mathbf{0}$ , the asymptotic distribution of the periodogram is given by

$$2 \frac{I(\boldsymbol{\omega}_n)}{f(\boldsymbol{\omega})} \xrightarrow{D} \chi_2^2.$$

*Proof.* Because  $|C(\mathbf{h})| \rightarrow 0$  as  $\|\mathbf{h}\| \rightarrow 0$  and  $\int_{\mathbb{R}^2} |C(\mathbf{h})| d\mathbf{h} < \infty$ , the spectral density function  $f(\boldsymbol{\omega})$  exists. Because  $\int_{\mathbb{R}^2} |C(\mathbf{h})| d\mathbf{h} < \infty$ , we can write  $F(\boldsymbol{\omega}) = F(\omega_1, \omega_2) = \int_{-\infty}^{\omega_1} \int_{-\infty}^{\omega_2} f(v_1, v_2) dv_2 dv_1$ . Therefore, by the Fundamental Theorem of Lebesgue Calculus, the spectral distribution function  $F(\boldsymbol{\omega})$  is absolutely continuous. By Theorem 7 of Rosenblatt (1985, pg. 73), the absolute continuity of the spectral distribution function implies that the process  $Y$  is strongly mixing. We assume this strong mixing satisfies conditions (A.1) and (A.2) of Bandyopadhyay and Lahiri (2009, pg. 228), which place a moment restriction on  $Y$  and state that the dependence between observations goes to 0 as their spatial separation gets large. Finally, by defining the mixed asymptotic scheme as,  $\delta \rightarrow 0$ ,  $n_1 \rightarrow \infty$ ,  $n_2 \rightarrow \infty$ ,  $n_1/n_2 \rightarrow c$  for a constant  $c > 0$ ,  $\delta n_1 \rightarrow \infty$ , and  $\delta n_2 \rightarrow \infty$ , we satisfy condition (A.3) of Bandyopadhyay and Lahiri (2009, pg. 237) which states that the the volume of the sampling domain must grow at a faster rate than the rate at which the distance between sampling locations decreases. Namely, we let the set  $\mathcal{Z}^2$  (Bandyopadhyay and Lahiri, 2009, pg. 226) be the integer grid  $\mathbb{Z}^2$ ,  $\eta_n = \delta$  and  $\lambda_n = \delta n$  so that  $\eta_n \lambda_n = \delta \delta n = \delta^2 n_1 n_2 \rightarrow \infty$ .

For  $\boldsymbol{\omega} \in \mathbb{R}^2$ , define the cosine and sine transforms, respectively, of spatial data at  $n$  sampling locations as

$$C_n(\boldsymbol{\omega}) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \cos(\boldsymbol{\omega}^\top \mathbf{s}_i) Y(\mathbf{s}_i)$$

and,

$$S_n(\boldsymbol{\omega}) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \sin(\boldsymbol{\omega}^\top \mathbf{s}_i) Y(\mathbf{s}_i).$$

With conditions (A.1)-(A.3) satisfied, Theorem 3.5 of Bandyopadhyay and Lahiri (2009, pg. 238) states the following:

For  $r \in \mathbb{Z}^+$ , let  $\{\boldsymbol{\omega}_{1n}\}, \dots, \{\boldsymbol{\omega}_{rn}\}$  be frequencies of the form  $\boldsymbol{\omega}_{jn} = \frac{2\pi \mathbf{k}_{jn}}{\delta n}$  for  $\mathbf{k}_{jn} \in \mathbb{Z}^2 \setminus \mathbf{0}$  such that  $\boldsymbol{\omega}_{jn} \neq \mathbf{0} \forall jn$ ,  $\boldsymbol{\omega}_{jn} \rightarrow \boldsymbol{\omega}_j \in \mathbb{R}^2 \setminus \mathbf{0}$ , and  $\boldsymbol{\omega}_j \pm \boldsymbol{\omega}_k \neq \mathbf{0} \forall 1 \leq j \neq k \leq r$ . Then

$$\delta[C_n(\boldsymbol{\omega}_{1n}), S_n(\boldsymbol{\omega}_{1n}), \dots, C_n(\boldsymbol{\omega}_{rn}), S_n(\boldsymbol{\omega}_{rn})]^\top \xrightarrow{D} N(\mathbf{0}, \Sigma)$$

where  $\Sigma$  is block diagonal with  $r$  blocks of the form  $\left(\frac{(2\pi)^2}{2} f(\boldsymbol{\omega}_l)\right) \mathbf{I}_2$  for  $l = 1, 2, \dots, r$ .

Letting  $\boldsymbol{\omega} = \boldsymbol{\omega}_{jn}$ , the periodogram can be written as follows (e.g., see Fuentes and Reich, 2010, pg. 68):

$$\begin{aligned}
I_n(\boldsymbol{\omega}) &= \frac{\delta^2}{(2\pi)^2 n} \left| \sum_{i=1}^n Y(\mathbf{s}_i) \exp \{ -i s_i^\top \boldsymbol{\omega} \} \right|^2 \\
&= \frac{\delta^2}{(2\pi)^2 n} \left| \sum_{i=1}^n \cos(\boldsymbol{\omega}^\top \mathbf{s}_i) Y(\mathbf{s}_i) - i \sin(\boldsymbol{\omega}^\top \mathbf{s}_i) Y(\mathbf{s}_i) \right|^2 \\
&= \left| \left( \frac{\delta}{2\pi\sqrt{n}} \right) \sum_{i=1}^n \cos(\boldsymbol{\omega}^\top \mathbf{s}_i) Y(\mathbf{s}_i) - i \sin(\boldsymbol{\omega}^\top \mathbf{s}_i) Y(\mathbf{s}_i) \right|^2 \\
&= \left| \frac{\delta C_n(\boldsymbol{\omega})}{2\pi} + i \frac{\delta S_n(\boldsymbol{\omega})}{2\pi} \right|^2 \\
&= \left( \frac{\delta C_n(\boldsymbol{\omega})}{2\pi} \right)^2 + \left( \frac{\delta S_n(\boldsymbol{\omega})}{2\pi} \right)^2 \\
&= \left( \frac{\sqrt{\frac{(2\pi)^2}{2} f(\boldsymbol{\omega}_j)}}{2\pi} \frac{\delta C_n(\boldsymbol{\omega})}{\sqrt{\frac{(2\pi)^2}{2} f(\boldsymbol{\omega}_j)}} \right)^2 + \left( \frac{\sqrt{\frac{(2\pi)^2}{2} f(\boldsymbol{\omega}_j)}}{2\pi} \frac{\delta S_n(\boldsymbol{\omega})}{\sqrt{\frac{(2\pi)^2}{2} f(\boldsymbol{\omega}_j)}} \right)^2 \\
&= \left( \sqrt{\frac{f(\boldsymbol{\omega}_j)}{2}} \frac{\delta C_n(\boldsymbol{\omega})}{\sqrt{\frac{(2\pi)^2}{2} f(\boldsymbol{\omega}_j)}} \right)^2 + \left( \sqrt{\frac{f(\boldsymbol{\omega}_j)}{2}} \frac{\delta S_n(\boldsymbol{\omega})}{\sqrt{\frac{(2\pi)^2}{2} f(\boldsymbol{\omega}_j)}} \right)^2 \\
&= \frac{f(\boldsymbol{\omega}_j)}{2} \left\{ \left( \frac{\delta C_n(\boldsymbol{\omega})}{\sqrt{\frac{(2\pi)^2}{2} f(\boldsymbol{\omega})}} \right)^2 + \left( \frac{\delta S_n(\boldsymbol{\omega})}{\sqrt{\frac{(2\pi)^2}{2} f(\boldsymbol{\omega}_j)}} \right)^2 \right\}.
\end{aligned}$$

Rearranging, we can write,

$$\frac{2I_n(\boldsymbol{\omega})}{f(\boldsymbol{\omega}_j)} = \left( \frac{\delta C_n(\boldsymbol{\omega})}{\sqrt{\frac{(2\pi)^2}{2} f(\boldsymbol{\omega}_j)}} \right)^2 + \left( \frac{\delta S_n(\boldsymbol{\omega})}{\sqrt{\frac{(2\pi)^2}{2} f(\boldsymbol{\omega}_j)}} \right)^2.$$

From Theorem 3.5 of Bandyopadhyay and Lahiri (2009, pg. 238), we have that

$$Z_c = \frac{\delta C_n(\boldsymbol{\omega})}{\sqrt{\frac{(2\pi)^2}{2} f(\boldsymbol{\omega}_j)}} \xrightarrow{D} N(0, 1), \quad Z_s = \frac{\delta S_n(\boldsymbol{\omega})}{\sqrt{\frac{(2\pi)^2}{2} f(\boldsymbol{\omega}_j)}} \xrightarrow{D} N(0, 1)$$

and  $Z_c$  and  $Z_s$  are asymptotically independent. Because  $g(x) = x^2$  is a continuous function, it follows that

$$\frac{2I_n(\boldsymbol{\omega})}{f(\boldsymbol{\omega}_j)} = (Z_c)^2 + (Z_s)^2 \xrightarrow{D} \chi_2^2,$$

because  $\phi_{Z_c^2 + Z_s^2}(t) = \phi_{Z_c^2}(t)\phi_{Z_s^2}(t) \rightarrow \phi_{\chi_1^2}(t)\phi_{\chi_1^2}(t) = \phi_{\chi_2^2}(t)$ , where  $\phi_X(t)$  is the characteristic function of random variable  $X$ . ■

Note that an assumption that the RF is a Gaussian RF is not required for Theorem 2. However, the mixing conditions of Bandyopadhyay and Lahiri (2009) are satisfied in the case that the RF is a Gaussian RF with  $C(\mathbf{h}) = 0$  for sufficiently large  $\|\mathbf{h}\|$  (Maity and Sherman, 2012).

While the periodogram has attractive asymptotic properties, aliasing may thwart its use for testing isotropy. Isotropy of the spectral density of the continuous (geostatistical) process,  $f(\boldsymbol{\omega})$ , does not imply isotropy of the aliased spectral density,  $f_{\Delta}(\boldsymbol{\omega})$ . In order to use spectral methods for a continuous process, the spacing between sampling locations must be small enough “so that higher frequencies make only negligible contribution to the total power of the process” (Fuentes, 2002). Similar to Fuentes (2002) and Crujeiras et al. (2010), we assume that aliasing is not a concern so that  $f(\boldsymbol{\omega}) \approx f_{\Delta}(\boldsymbol{\omega})$ . We explore the impact of aliasing on covariance estimation and develop a graphical diagnostic to assess whether or not aliasing is a concern in Section 4.8.



## 4.4 Test Statistic

### 4.4.1 Constructing a Test Statistic

In this section we develop potential test statistics to test the hypotheses,

$$\begin{aligned} H_0: C(\mathbf{h}) &= C_0(\|\mathbf{h}\|); & (\text{covariance function is isotropic}) \\ H_1: C(\mathbf{h}) &\neq C_0(\|\mathbf{h}\|); & (\text{covariance function is anisotropic}). \end{aligned} \tag{29}$$

By Theorem (1), we can rewrite (29) in terms of the spectral density function,

$$\begin{aligned} H_0: f(\boldsymbol{\omega}) &= f_0(\|\boldsymbol{\omega}\|); & (\text{spectral density function is isotropic}) \\ H_1: f(\boldsymbol{\omega}) &\neq f_0(\|\boldsymbol{\omega}\|); & (\text{spectral density function is anisotropic}). \end{aligned} \tag{30}$$

In developing a test statistic for isotropy, we follow steps similar to those used in Scaccia and Martin (2005) who used the periodogram to test symmetry and separability of lattice processes. Under the null hypothesis of isotropy, we have that  $f(\boldsymbol{\omega}_i) = f(\boldsymbol{\omega}_j)$  when  $\|\boldsymbol{\omega}_i\| = \|\boldsymbol{\omega}_j\|$ . To estimate differences in the spectral density at different frequencies with the same norm, we consider the log ratio of periodogram values  $D_k$  where

$$D_k = \log \left( \frac{I(\boldsymbol{\omega}_i)}{I(\boldsymbol{\omega}_j)} \right) = \log I(\boldsymbol{\omega}_i) - \log I(\boldsymbol{\omega}_j) \text{ where } \|\boldsymbol{\omega}_i\| = \|\boldsymbol{\omega}_j\|.$$

Using the log of the periodogram approximately stabilizes the variance and reduces the non-normality of the periodogram values (Scaccia and Martin, 2005). Under the null hypothesis (30) of isotropy, the expected log ratio equals 0:

$$E(D_k) = E[\log I(\boldsymbol{\omega}_i)] - E[\log I(\boldsymbol{\omega}_j)] = 0. \tag{31}$$

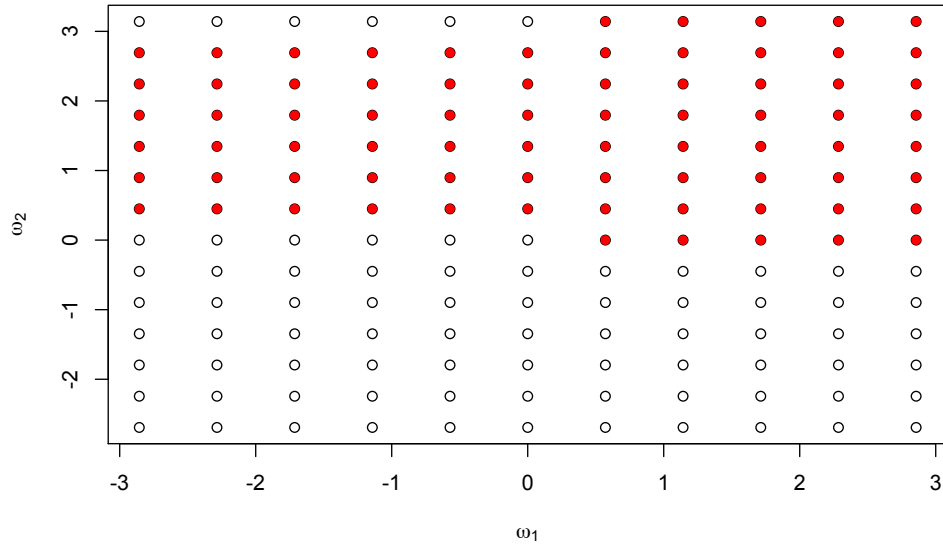


Figure 4.2: The circles ( $\circ$ ) denote the set of Fourier frequencies  $\mathcal{F}$  (32) for data observed on a  $14 \times 11$  (rows  $\times$  columns) subset of  $\mathbb{Z}^2$ . The red circles indicate the frequencies in the set  $\mathcal{K}$  (33) that are used to test isotropy. There are  $K = 30$  pairs of frequencies with the same norm in  $\mathcal{K}$ .

We wish to use the mean of the  $D_k$  values to compute a test statistic. To construct the  $D_k$ 's, we begin by estimating the periodogram (26) at the set of Fourier frequencies,

$$\mathcal{F} = \left\{ \boldsymbol{\omega} = (\omega_1, \omega_2) : \omega_1 = \frac{2\pi}{\delta_1 n_1} k_1, \omega_2 = \frac{2\pi}{\delta_2 n_2} k_2 \right\} \quad (32)$$

where

$$k_1 \in \left\{ -\left\lfloor \frac{n_1-1}{2} \right\rfloor, -\left\lfloor \frac{n_1-2}{2} \right\rfloor, \dots, 0, \dots, \left\lfloor \frac{n_1-2}{2} \right\rfloor, \left\lfloor \frac{n_1-1}{2} \right\rfloor \right\},$$

$$k_2 \in \left\{ -\left\lfloor \frac{n_2-1}{2} \right\rfloor, -\left\lfloor \frac{n_2-2}{2} \right\rfloor, \dots, 0, \dots, \left\lfloor \frac{n_2-2}{2} \right\rfloor, \left\lfloor \frac{n_2-1}{2} \right\rfloor \right\},$$

and  $\delta_1, \delta_2$  are the spacings between the rows and columns, respectively. Figure 4.2 displays the set of Fourier frequencies  $\mathcal{F}$  (32) for data observed on a  $14 \times 11$  (rows  $\times$  columns) subset of  $\mathbb{Z}^2$  (i.e.,  $\delta_1 = \delta_2 = 1$ ). Because the periodogram is an even function, it follows that  $I(\omega_1, \omega_2) = I(-\omega_1, -\omega_2)$ , and  $I(\omega_1, \pi) = I(-\omega_1, \pi)$ . As a result, only a subset of the periodogram values will be useful for testing isotropy (Lu and Zimmerman, 2005). Specifically,  $D_k$  is computed at the set of relevant frequencies  $\mathcal{K}$ , where

$$\mathcal{K} = \mathcal{F} \setminus \left( \{(\omega_1, \omega_2) : \omega_2 < 0\} \cup \{(\omega_1, \omega_2) : \omega_1 \leq 0 \text{ and } \omega_2 = 0 \text{ or } \pi\} \right). \quad (33)$$

Figure 4.2 displays the set  $\mathcal{K}$  for data on a  $14 \times 11$  subset of  $\mathbb{Z}^2$ . In Figure 4.2 there are  $K = 30$  pairs of frequencies in the set  $\mathcal{K}$  with the same norm. If the spacing between sampling locations in the row and columns is the same ( $\delta_1 = \delta_2$ ), then due to the construction of the Fourier frequencies (32), the number of pairs of frequencies having the same norm  $K$  will depend on the number of common factors between the number of rows  $n_1$  and number of columns  $n_2$ . More common factors between  $n_1$  and  $n_2$  implies a larger  $K$ .

Let  $\bar{D} = \left( \sum_{k=1}^K D_k \right) / K$ , where  $K$  is the number of pairs of Fourier frequencies in  $\mathcal{K}$  that have the same norm. It follows from (31) that, under the null hypothesis of isotropy (30),

$E(\bar{D}) = 0$ . Recall, periodogram values are approximately uncorrelated (28). As a result,

$$Z = \frac{\bar{D} - 0}{\sigma_{\bar{D}}} \xrightarrow{D} N(0, 1) \quad (34)$$

by the Central Limit Theorem.

#### 4.4.2 Test Statistic Variance Approximations

To use the test statistic (34) in practice, we need an estimator of  $\sigma_{\bar{D}}^2$ , where

$$\sigma_{\bar{D}}^2 = \text{Var}(\bar{D}) = \text{Var}\left(\frac{\sum_{k=1}^K D_k}{K}\right) = \frac{1}{K^2} \left[ \sum_{k=1}^K \text{Var}(D_k) + 2 \sum_{1 \leq k < l \leq K} \text{Cov}(D_k, D_l) \right]. \quad (35)$$

We can rewrite  $\text{Var}(D_k)$  in (35) as:

$$\begin{aligned} \text{Var}(D_k) &= \text{Var}[\log I(\boldsymbol{\omega}_i) - \log I(\boldsymbol{\omega}_j)] \\ &= \text{Var}[\log I(\boldsymbol{\omega}_i)] + \text{Var}[\log I(\boldsymbol{\omega}_j)] - 2\text{Cov}[\log I(\boldsymbol{\omega}_i), \log I(\boldsymbol{\omega}_j)]. \end{aligned} \quad (36)$$

Assuming the asymptotic property of uncorrelated periodogram values (28),

$$\text{Cov}[\log I(\boldsymbol{\omega}_i), \log I(\boldsymbol{\omega}_j)] = 0, \text{ for } \boldsymbol{\omega}_i \neq \boldsymbol{\omega}_j. \quad (37)$$

Next we explore two approximations of  $\text{Var}[\log I(\boldsymbol{\omega}_i)]$  in (36). The first approach is to use the asymptotic properties (28) of the periodogram and a first order Taylor series:

$$\begin{aligned} \text{Var}(\log I(\boldsymbol{\omega})) &\approx \left[ \frac{1}{E(I(\boldsymbol{\omega}))} \right]^2 \text{Var}(I(\boldsymbol{\omega})) \\ &= \left[ \frac{1}{f(\boldsymbol{\omega})} \right]^2 (f(\boldsymbol{\omega}))^2 = 1. \end{aligned} \quad (38)$$

The second approach to approximate  $\text{Var}[\log I(\boldsymbol{\omega})]$  is to consider the asymptotic distribution of  $\log(I(\boldsymbol{\omega})/f(\boldsymbol{\omega}))$ . Recall, Theorem 2 states that under mixed increasing domain and infill

asymptotics,

$$\frac{I(\boldsymbol{\omega})}{f(\boldsymbol{\omega})} \xrightarrow{D} \frac{1}{2}\chi_2^2 \text{ for } \boldsymbol{\omega} \in \mathbb{R}^2 \setminus \mathbf{0}.$$

where  $\xrightarrow{D}$  represents convergence in distribution. Equivalently, we can write

$$\frac{I(\boldsymbol{\omega})}{f(\boldsymbol{\omega})} \xrightarrow{D} \text{Exp}(1) \text{ for } \boldsymbol{\omega} \in \mathbb{R}^2 \setminus \mathbf{0}$$

because a  $\chi_2^2/2$  distribution is an exponential distribution with parameter 1. By the continuous mapping theorem, if  $X_n \xrightarrow{D} X$  and  $g: \mathbb{R} \rightarrow \mathbb{R}$  is a continuous function, then  $g(X_n) \xrightarrow{D} g(X)$ . If we let  $g(x) = -\log x$ , which is a continuous function for  $x > 0$ , then

$$-\log\left(\frac{I(\boldsymbol{\omega})}{f(\boldsymbol{\omega})}\right) \xrightarrow{D} -\log(\text{Exp}(1)) = \text{Gumbel}(0,1). \quad (39)$$

for  $\boldsymbol{\omega} \in \mathbb{R}^2 \setminus \mathbf{0}$ . If  $X \sim \text{Gumbel}(0,1)$ , then  $\text{Var}(X) = \pi^2/6$ . Therefore, it follows from (39) that

$$\text{Var}(\log f(\boldsymbol{\omega}) - \log I(\boldsymbol{\omega})) = \text{Var}(\log I(\boldsymbol{\omega})) \approx \pi^2/6. \quad (40)$$

To complete our estimate of  $\text{Var}(\bar{D})$  we need to explore  $\text{Cov}(D_k, D_l)$  in (35), where

$$\text{Cov}(D_k, D_l) = E(D_k D_l) - E(D_k)E(D_l).$$

Recall that  $E(D_k) = 0$  under the null hypothesis of isotropy (30). Next

$$\begin{aligned} E(D_k D_l) &= E[(\log I(\boldsymbol{\omega}_{k_1}) - \log I(\boldsymbol{\omega}_{k_2}))(\log I(\boldsymbol{\omega}_{l_1}) - \log I(\boldsymbol{\omega}_{l_2}))] \\ &= E[\log I(\boldsymbol{\omega}_{k_1}) \log I(\boldsymbol{\omega}_{l_1}) - \log I(\boldsymbol{\omega}_{k_2}) \log I(\boldsymbol{\omega}_{l_1}) \\ &\quad - \log I(\boldsymbol{\omega}_{k_1}) \log I(\boldsymbol{\omega}_{l_2}) + \log I(\boldsymbol{\omega}_{k_2}) \log I(\boldsymbol{\omega}_{l_2})] \end{aligned}$$

where  $\|\boldsymbol{\omega}_{k_1}\| = \|\boldsymbol{\omega}_{k_2}\| \equiv \|\boldsymbol{\omega}_k\|$  and  $\|\boldsymbol{\omega}_{l_1}\| = \|\boldsymbol{\omega}_{l_2}\| \equiv \|\boldsymbol{\omega}_l\|$ . There are two cases to consider: (1)  $\|\boldsymbol{\omega}_k\| \neq \|\boldsymbol{\omega}_l\|$  and (2)  $\|\boldsymbol{\omega}_k\| = \|\boldsymbol{\omega}_l\|$ . In case (1) we can use the asymptotic

independence of periodogram values to obtain  $E(D_k D_l) = E(D_k)E(D_l) = 0$ . In case (2) all of the frequencies have the same norm. Therefore, under  $H_0$ , all the periodogram values will have the same expectation so that  $E(D_k D_l) = 0$ . Thus,  $\text{Cov}(D_k, D_l) = 0$ , and we ignore the second summand in (35).

Combining the above results, we have two approximations for  $\text{Var}(\bar{D})$ . The first is given by first plugging the results from (38) and (37) into (36), then substituting that result into (35). This gives the following variance approximation,

$$\text{Taylor series: } \sigma_D^2 \approx \frac{2}{K}. \quad (41)$$

The second variance approximation is obtained by first plugging (40) and (37) into (36), then substituting that result into (35),

$$\text{asymptotic distribution: } \sigma_D^2 \approx \frac{\pi^2/3}{K} \approx \frac{3.29}{K}. \quad (42)$$

It is important to note that when deriving expressions for both summands in (35), we have used the assumption that periodogram values are approximately uncorrelated (28). Uncorrelated periodogram values is an asymptotic property and, as we discuss further in Section 4.5.3 below, may not hold for finite samples.

### 4.4.3 Three Test Statistics

Given (34) and the variance approximations in (41) and (42), consider three test statistics for testing isotropy:

$$Z_1 = \frac{\bar{D} - 0}{\sqrt{2/K}}, \quad (43)$$

$$Z_2 = \frac{\bar{D} - 0}{\sqrt{\pi^2/(3K)}}, \text{ and} \quad (44)$$

$$t = \frac{\bar{D} - 0}{S_{D_k}/\sqrt{K}}. \quad (45)$$

Here (43) and (44) follow from the variance approximations (41) and (42), respectively. In (45)  $S_{D_k}$  is the sample standard deviation of the  $D_k$ 's. Assuming the  $D_k$ 's are approximately uncorrelated,  $t_{test}$  follows a Student  $t$  distribution with  $K - 1$  degrees of freedom. We explore the size and power of these test statistics in Section 4.5.

## 4.5 Simulation Study

### 4.5.1 Set Up

We perform a simulation study to evaluate the performance of our spectral domain non-parametric test of spatial isotropy. The simulation study allows us to evaluate our three test statistics and compare our method with other tests of spatial covariance properties. We compare the empirical size, power, and computational cost of our method with the test of complete symmetry from Lu and Zimmerman (2005) and the test of isotropy from Guan et al. (2004). We perform the tests on the same realizations of the RF. Except where noted, we simulate mean 0, Gaussian RFs on rectangular subsets of  $\mathbb{Z}^2$ . We use a Matérn covariance function (23) with the sill ( $\sigma^2$ ) and smoothness ( $\nu$ ) equal to one. We vary  $\alpha^{-1}$  so that the effective range, the distance at which spatial correlation equals 0.05, of an isotropic process corresponded to short ( $\alpha^{-1} = 0.75$ ), medium ( $\alpha^{-1} = 1.52$ ), and long range ( $\alpha^{-1} = 3.1$ )

dependence. The parameter  $\phi$  in (23) is given by

$$\phi = \frac{\sigma^2 \alpha^{2\nu} \Gamma(\nu + 1)}{\pi \Gamma(\nu)}.$$

We simulate data from isotropic and geometrically anisotropic covariance functions. Following Weller and Hoeting (2016) we introduce geometric anisotropy via coordinate transformations. Geometric anisotropy is governed by a rotation parameter  $\theta$  and scaling parameter  $R$  that define the ellipses of equicorrelation (see the contour figures embedded in Table 4.1). The parameter  $\theta, 0 \leq \theta < \pi$ , quantifies the angle between the major axis of the ellipse and the  $x$ -axis (counter-clockwise rotation) while  $R > 1$  gives the ratio of the major and minor axes of the ellipse. Note that  $R = 1$  and  $\theta = 0$  corresponds to the isotropic case. As  $R$  becomes larger, the spatial process becomes more anisotropic (i.e., the spatial dependence becomes stronger in the direction of  $\theta$  than in the direction perpendicular to  $\theta$ ).






#### 4.5.2 Results

Table 4.1 displays the empirical size and power for the three test statistics proposed in Section 4.4.2. Test statistics  $Z_2$  (44) and  $t$  (45) have nominal size for the tests we considered. Test statistic  $Z_1$  (43) has inflated Type I errors and is thus anti-conservative. These results indicate that the Taylor series variance (41) is a poor approximation. An examination of the simulated  $\bar{D}$  values indicated that the empirical variance of  $\bar{D}$  was much larger than the approximation in (41). This poor approximation leads to an inflated test size when using the test statistic  $Z_1$ .

The results in Table 4.1 indicate that  $Z_1$  has higher empirical power than  $Z_2$  and  $t$ , although this is at the expense of an inflated size. Test statistics  $Z_2$  and  $t$  have very similar power in practice. As expected, when  $\theta = 3\pi/8$ , all three test statistics display larger power as the process becomes more anisotropic (larger  $R$ ). Surprisingly, the results indicate that all three test statistics have low empirical power when  $\theta = 0$ . We further explore the power of








Table 4.1: Comparison of test statistic performance. We compare the size and power of our three test statistics: (43), (44), (45). Empirical size and power are based on 1000 realizations of a random field with a Matérn covariance function and nominal size of  $\alpha = 0.05$ .

18 cols $\times$ 12 rows grid							
					effective range		
	Contour	$R$	$\theta$	Test Stat	3	6	12
Size		0	0	$Z_1$	0.14	0.12	0.12
				$Z_2$	0.05	0.05	0.04
				$t_{test}$	0.05	0.05	0.05
Power		$\sqrt{2}$	0	$Z_1$	0.12	0.11	0.11
				$Z_2$	0.05	0.04	0.04
				$t_{test}$	0.05	0.04	0.06
		2	0	$Z_1$	0.12	0.10	0.08
				$Z_2$	0.04	0.04	0.03
				$t_{test}$	0.04	0.05	0.04
		$\sqrt{2}$	$\frac{3\pi}{8}$	$Z_1$	0.35	0.45	0.43
				$Z_2$	0.19	0.28	0.29
				$t_{test}$	0.18	0.29	0.31
		2	$\frac{3\pi}{8}$	$Z_1$	0.71	0.85	0.83
				$Z_2$	0.54	0.71	0.71
				$t_{test}$	0.51	0.70	0.72

the test statistic  $t$  as a function of  $\theta$  below. Based on the size and power results in Table 4.1, we recommend using either test statistic  $Z_2$  (44) or  $t$  (45) to test spatial isotropy.

Table 4.2 explores the size and power of our test, using test statistic  $t$  (44), compared to the nonparametric tests from Lu and Zimmerman (2005) and Guan et al. (2004). The data for this table were simulated on an  $18 \times 12$  subset of  $\mathbb{Z}^2$ . For this grid there were enough frequencies to perform the test of reflection and complete symmetry in Lu and Zimmerman (2005). Reflection and complete symmetry are covariance function properties for lattice (discrete) processes. The relationship between isotropy and the different types of symmetry

Table 4.2: Comparison of methods. We compare the size and power of our three methods for testing second order properties. The value  $t_{test}$  (45) is our method for testing isotropy using the spectral domain. LZ denotes the test of complete symmetry from Lu and Zimmerman (2005) using spectral methods. GSC denotes the test of isotropy from Guan et al. (2004) in the spatial domain. Empirical size and power are based on 1000 realizations of a GRF with a Matérn covariance function and nominal size of  $\alpha = 0.05$ .

18 cols $\times$ 12 rows grid							
	Contour	$R$	$\theta$	Test Stat	effective range		
					3	6	12
Size		0	0	$t_{test}$	0.05	0.05	0.05
				LZ	0.05	0.06	0.06
				GSC	0.04	0.07	0.13
Power		$\sqrt{2}$	0	$t_{test}$	0.05	0.04	0.06
				LZ	0.05	0.06	0.06
				GSC	0.64	0.76	0.75
		2	0	$t_{test}$	0.04	0.05	0.04
				LZ	0.08	0.08	0.08
				GSC	1.00	1.00	1.00
		$\sqrt{2}$	$\frac{3\pi}{8}$	$t_{test}$	0.18	0.29	0.31
				LZ	0.16	0.25	0.27
				GSC	0.46	0.65	0.66
		2	$\frac{3\pi}{8}$	$t_{test}$	0.51	0.70	0.72
				LZ	0.51	0.67	0.68
				GSC	0.99	1.00	0.99

is

$$\text{isotropy} \implies \text{complete symmetry} \implies \text{reflection symmetry}. \quad (46)$$

If the covariance function of a lattice process is isotropic, then it is completely and reflection symmetric. As a result, the test in Lu and Zimmerman (2005) can be used to reject a null hypothesis of isotropy. However, the test in Lu and Zimmerman (2005) is testing a more narrow hypothesis (symmetry) than our method (isotropy). Because of the relationship (46), failing to reject a null hypothesis of reflection and/or complete symmetry does not imply that an assumption of isotropy is appropriate. Similarly, rejecting a null hypothesis of reflection symmetry implies rejecting the null hypothesis of complete symmetry, but failing to reject a null hypothesis of reflection symmetry does not imply that an assumption of complete

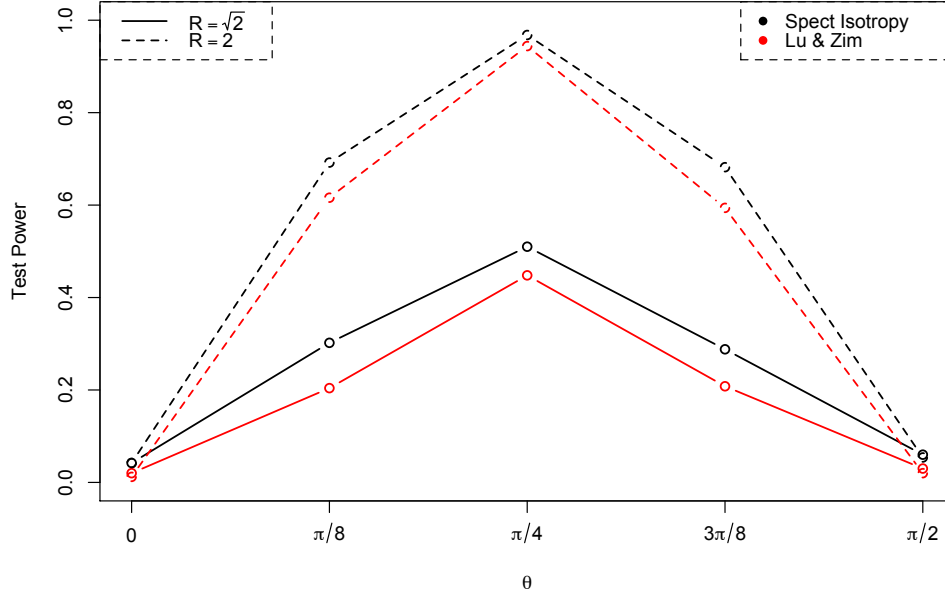


Figure 4.3: Empirical power of spectral hypothesis tests as a function of  $\theta$  for data on a  $20 \times 14$  subset of  $\mathbb{Z}^2$ . The black points correspond to a test of isotropy using  $t_{test}$  (45) while the red points correspond to a test of reflection symmetry from Lu and Zimmerman (2005). Empirical power is based on 500 realizations of a GRF with a Matérn covariance function and nominal size of  $\alpha = 0.05$ .

symmetry is appropriate. As a result, Lu and Zimmerman (2005) use a two-stage approach to test the assumption of complete symmetry: by first testing reflection symmetry then testing complete symmetry. To control the overall Type I error rate of a test of complete symmetry at level  $\alpha$ , each of the tests is performed at a level of  $\alpha/2$ . The results in Table 4.2 indicate that the size and power of the two spectral methods ( $t_{test}$  and Lu and Zimmerman (2005)) are similar. While the spatial method of Guan et al. (2004) exhibits an inflated size as the effective range increases, the tests using the spectral domain maintain nominal size. The power for the spatial method (Guan et al., 2004) is larger than both spectral methods.

Figure 4.3 compares the power of our test to the test from Lu and Zimmerman (2005) as a function of the angle of geometric anisotropy,  $\theta$ . For these simulations, we used a  $20 \times 14$  subset of  $\mathbb{Z}^2$ . These grid dimensions do not allow us to perform the test of complete symmetry

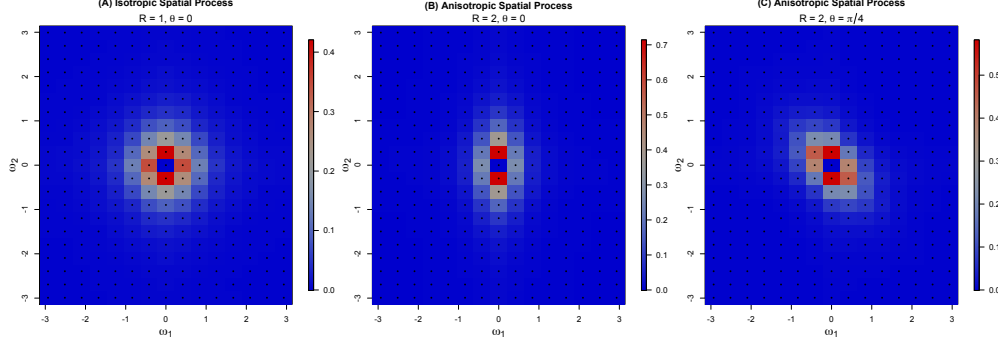


Figure 4.4: Periodogram values as a function of geometric anisotropy. The plots display the mean periodogram values (26) for an isotropic process (A) and two geometrically anisotropic processes, (B) and (C). The mean values are based on 500 realizations of a GRF with Matérn covariance function on a  $21 \times 15$  subset of  $\mathbb{Z}^2$ . The black points in the plots are the Fourier frequencies. The power for spectral tests of directional dependence is larger when  $\theta = \pi/4$  because there are more pairs of frequencies with the same norm having different periodogram values than when  $\theta = 0$ . This tends to lead to larger values of  $|\bar{D}|$  and the test statistic, therefore increasing power. Figure 4.3 displays empirical power as a function of  $\theta$ .

from Lu and Zimmerman (2005), which was designed for square grids. As a result, only the test of reflection symmetry was performed. Under this scenario, our method has higher power than Lu and Zimmerman (2005). Both methods, which use the spectral domain, have little to no power when the direction of strongest correlation is parallel to the  $x$  or  $y$  axis ( $\theta = 0, \pi/2$ ) and the highest power when the direction of strongest correlation creates a 45-degree angle with the  $x$  and  $y$  axis ( $\theta = \pi/4$ ). The change in power as a function of  $\theta$  occurs because the periodogram measures the directional dependence present in the spatial domain (see Figure 4.4) through the estimates of the spectral density at different frequencies. When  $\theta = \pi/4$  there are many Fourier frequencies that have the same norm but different periodogram values (Figure 4.4C). These many different periodogram values tend to lead to larger values of the statistic  $|\bar{D}|$ , a larger test statistic, e.g., (45), and therefore larger power. When  $\theta = 0$  or  $\pi/2$ , there are relatively fewer frequencies with the same norm having different periodogram values. This tends to lead to smaller values of the statistic  $|\bar{D}|$ , a smaller test statistic, and therefore smaller power.

The results in Table 4.3 explore the power of our method as a function of the sample size  $n$ . To emulate mixed increasing domain and infill asymptotics, we increased the domain size and reduced the spacing between sampling locations as we increased the sample size. The dimensions and spacings of the grids used in the simulation were as follows: (1)  $n = 104$  on a  $18 \times 13$  grid with spacing  $\Delta = (2, 2)$ , (2)  $n = 459$  on a  $27 \times 17$  grid with spacing  $\Delta = (1, 1)$ , and (3)  $n = 2109$  on a  $57 \times 37$  grid with spacing  $\Delta = (0.5, 0.5)$ . Under these three scenarios, the distance between sampling locations decreases ( $\delta_1 = \delta_2 = 2, 1, 0.5$ ) while the size of the domain increases (maximum distances of 27.8, 30.5, and 33.3, respectively). For these simulations, we set  $\nu = 1$ ,  $\sigma^2 = 1$ , and  $\alpha^{-1} = 1.53$ . As expected, the empirical power increases as a function of  $n$ . Similar to the results in Figure 4.3, the power increases as  $\theta$  goes from 0 to  $\pi/4$ .

The computation time for our method is significantly less than the time needed for the tests in Lu and Zimmerman (2005) and Guan et al. (2004) (Table 4.4). The larger computation time for Lu and Zimmerman (2005) is due to the fact that the distribution of the test statistic needs to be simulated to compute a  $p$  value. Guan et al. (2004) requires spatial subsampling, which increases computation time. For a large data set, our test will offer a substantial computational savings over other methods.

Table 4.3: Empirical power as a function of  $n$ . Empirical power is based on 1000 realizations of a GRF with a Matérn covariance function and nominal size of  $\alpha = 0.05$ . We used the  $t$  test statistic (45) and mixed increasing domain/infill spatial asymptotics.



Contour	$R$	$\theta$	$n$	Power
	1.7	$\frac{7\pi}{8}$	104	0.15
			459	0.72
			2109	0.88
	1.7	$\frac{\pi}{4}$	104	0.29
			459	0.94
			2109	1.00

Table 4.4: Computation time for tests of second order properties for a  $25 \times 16$  grid of sampling locations.

Method	$t_{test}$	LZ	GSC
<b>Computation Time</b>	0.06 s	5.48 s	6.09 s

#### 4.5.3 Variance Estimation and Finite Sample Correlation

In simulation studies we found that the empirical variance of  $\bar{D}$  was not always well approximated by the expressions in the denominator of the test statistics (43), (44), and (45). The asymptotic assumption that  $\text{Cov}(D_k, D_l) = 0$  is not always adequate in practice, and the  $D_k$  values can have high positive correlation. The positive correlation increases the empirical variance (35) of  $\bar{D}$  thus inflating test size. Positive correlation between the  $D_k$  values becomes non-negligible as the number of frequencies with the same norm increases. Because we estimate the spectral density at the Fourier frequencies  $\mathcal{F}$  and the Fourier frequencies are a function of the number of rows and columns of gridded data (see (32)), the number of frequencies with the same norm will be a function of the number of rows and columns of data. Square grids with equal spacing between rows and columns allow for the maximum number of Fourier frequencies having the same norm. For example, if there are  $n = 400$  sampling locations, a  $20 \times 20$  grid of  $\mathbb{Z}^2$  creates  $K = 291$  pairs of frequencies in  $\mathcal{K}$  with the same norm while a  $25 \times 16$  grid creates  $K = 84$ .

We explored several methods to account for or reduce the large finite sample correlations that occur for square grids. First we derived the approximate finite sample covariance

(correlation) via a Taylor series expansion of  $\text{Cov}(D_k, D_l)$ ,

$$\begin{aligned}
\text{Cov}(D_k, D_l) &= E[(\log I(\boldsymbol{\omega}_{k_1}) - \log I(\boldsymbol{\omega}_{k_2}))(\log I(\boldsymbol{\omega}_{l_1}) - \log I(\boldsymbol{\omega}_{l_2}))] \\
&= E[\log I(\boldsymbol{\omega}_{k_1}) \log I(\boldsymbol{\omega}_{l_1}) - \log I(\boldsymbol{\omega}_{k_2}) \log I(\boldsymbol{\omega}_{l_1}) \\
&\quad - \log I(\boldsymbol{\omega}_{k_1}) \log I(\boldsymbol{\omega}_{l_2}) + \log I(\boldsymbol{\omega}_{k_2}) \log I(\boldsymbol{\omega}_{l_2})] \\
&\approx \log f(\boldsymbol{\omega}_{k_1}) \log f(\boldsymbol{\omega}_{l_1}) - \frac{1}{2} \{\log f(\boldsymbol{\omega}_{l_1}) + \log f(\boldsymbol{\omega}_{k_1})\} \\
&\quad - \left( \log f(\boldsymbol{\omega}_{k_2}) \log f(\boldsymbol{\omega}_{l_1}) - \frac{1}{2} \{\log f(\boldsymbol{\omega}_{l_1}) + \log f(\boldsymbol{\omega}_{k_2})\} \right) \\
&\quad - \left( \log f(\boldsymbol{\omega}_{k_1}) \log f(\boldsymbol{\omega}_{l_2}) - \frac{1}{2} \{\log f(\boldsymbol{\omega}_{l_2}) + \log f(\boldsymbol{\omega}_{k_1})\} \right) \\
&\quad + \log f(\boldsymbol{\omega}_{k_2}) \log f(\boldsymbol{\omega}_{l_2}) - \frac{1}{2} \{\log f(\boldsymbol{\omega}_{l_2}) + \log f(\boldsymbol{\omega}_{k_2})\}.
\end{aligned} \tag{47}$$

The Taylor series expansion (47) gives an expression that is a function of the spectral density function,  $f(\boldsymbol{\omega})$ . Similarly, expressions for the finite sample covariance of  $I(\boldsymbol{\omega}_i)$  and  $I(\boldsymbol{\omega}_j)$  have been derived (Kim and Fuentes, 2000; Crujeiras, 2006; Fuentes and Smith, 2001) but also depend on the value of the spectral density function, which is unknown. Because our test is nonparametric, we do not want to assume and/or estimate a parametric spectral density function.

As a second attempt to reduce finite sample correlation and improve our approximation to  $\text{Var}(\bar{D})$ , we used a subset of the Fourier frequencies to compute our test statistic (45) for data on square grids. We used two different subsets of  $\mathcal{K}$  for testing isotropy: (1) use every other Fourier frequency, or (2) use only frequencies that have norm less than half of the maximum norm. Simulations indicated that reducing the number of frequencies for square grids did not adequately reduce the positive correlation (inflated size) and resulted in a loss of the anisotropic signal (reduced power).

Finally, we attempted to adjust  $K$  to account for effective sample size. Simulations indicated that, even with the correct adjustment to  $K$ , the method for testing symmetry from Lu and Zimmerman (2005), which was designed for square grids, had larger power than

our test. We believe that this is due to the fact that Lu and Zimmerman (2005) test the more specific hypothesis of symmetry, a property implied by isotropy. Square grids provide a large number of pairs of frequencies for testing symmetry. As a result, for square grids we recommend using Lu and Zimmerman (2005).

## 4.6 Application: WRFG Temperatures

We apply our method to gridded temperature data from the North American Regional Climate Change Assessment Program [NARCAPP] (Mearns et al., 2009). The data set WRFG in the R package `spTest` (Weller, 2016d) includes coordinates and the 24-year average of yearly average temperature from runs of the Weather Research and Forecasting - Grell configuration (WRFG) regional climate model (RCM) using boundary conditions from the National Centers for Environmental Prediction. We focus on a  $25 \times 20$  subset of these data over the central United States (Figure 4.5). We perform a similar analysis to Weller (2016c) who analyzed a  $20 \times 20$  subset of these data.

Before testing for directional dependence, we first model the effects of northing and easting coordinates UTM coordinates on average temperature. We fit the mean temperature as a nonparametric, additive function of both the easting and northing coordinates via least-squares using cubic splines. After removing the trends in temperature, we use the graphical diagnostic from Section 4.8.3 to check for aliasing. After determining that aliasing is not a concern (see Section 4.8.3), we test for directional dependence in the studentized residuals (see Figure 4.5) using our spectral method  $t_{test}$  and the spatial method from Guan et al. (2004). For our spectral test, there are  $K = 122$  frequencies with the same norm. The log periodogram values at these frequencies provided  $\bar{D} = -0.11$  and  $t_{test} = -0.66$ . Thus, based on our spectral test, there is not enough evidence to reject an assumption of isotropy ( $p$  value = 0.51). Likewise, the spatial test from Guan et al. (2004) indicated a failure to reject the isotropy assumption ( $p$  value = 0.09). For this example, the first advantage of our method is the avoidance of choosing the tuning parameters needed for Guan et al. (2004), such as



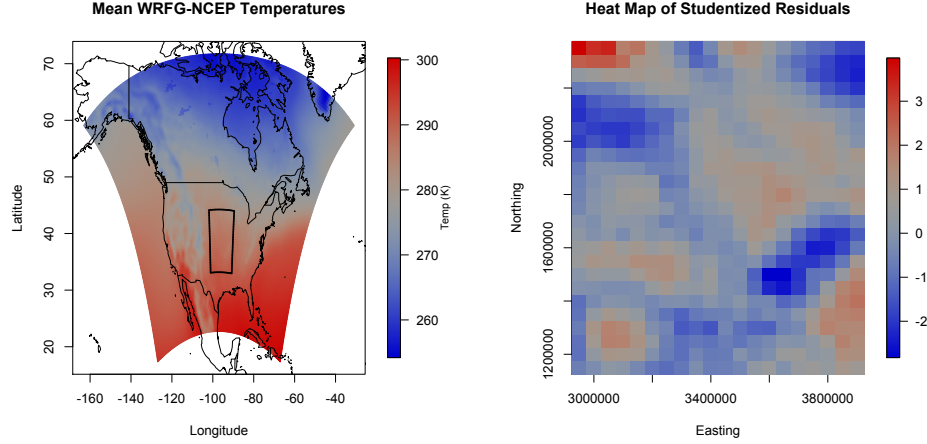


Figure 4.5: Heat maps displaying the  $25 \times 20$  subset of WRFG data (left) and the studentized residuals (right) used to test for directional dependence in Section 4.6.

the spatial lag set and the moving window size. The second advantage is the computation time: 0.08 seconds for our method and 36.6 second for Guan et al. (2004).

## 4.7 Discussion

While spectral methods are not commonly used for spatially-referenced data, we developed a nonparametric test of spatial isotropy using the spectral representation of the covariance function. Use of the spectral domain presents several challenges, including variance approximations and correlation of nonparametric estimators. We were able to address and overcome these challenges to develop a test statistic and approximate its variance (Section 4.4) in order to test for anisotropy. Finally, we explored the size and power of our method via simulation studies and applied the method to a data set of temperatures over the central U.S.

Our test has several advantages over other tests of isotropy. Our test avoids potential misspecification of a parametric covariance function. Additionally, we avoid the need to estimate covariance parameters under the null and alternative hypotheses (e.g., for a likelihood ratio test), which is computationally demanding for large data sets. Our test has larger power for rectangular grids than the nonparametric spectral methods of Lu and Zimmerman

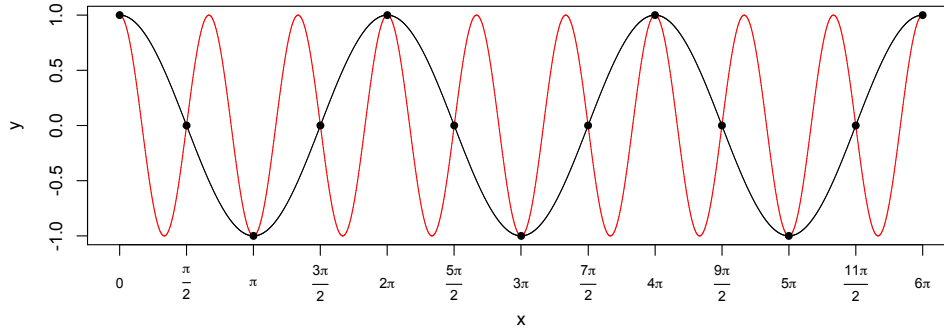


Figure 4.6: Example of aliasing. Plots of  $y = \cos(x)$  [black line] and  $y = \cos(3x)$  [red line] for  $0 \leq x \leq 6\pi$ . If we observe  $y$  at the set of locations  $x \in \{0, \pi/2, \pi, \dots, 11\pi/2, 6\pi\}$ , the aliasing effect refers to the fact that we cannot distinguish between the frequencies 1 and 3.

(2005). Nonparametric tests of isotropy in the spatial domain require user choices, such as lag set and window size (e.g., Guan et al., 2004). Nonparametric spatial domain tests are potentially sensitive to these choices and their optimality remains an open question (Weller and Hoeting, 2016). Thus, an attractive feature of our spectral test is avoidance of these user choices. Likewise, our test does not require spatial subsampling, which reduces the computation time. This computational savings increases with sample size because the number of spatial subsamples increases with sample size.

## 4.8 Aliasing in the Spectral Domain

### 4.8.1 Introduction

The aliasing phenomenon described in Section 4.3 can present a major challenge when using spectral methods for geostatistical processes. The aliasing phenomenon refers to the fact that it is not possible to recover a continuous signal from a discrete set of observations (Fuentes and Reich, 2010). Figure 4.6 displays the aliasing phenomenon using the cos

function. Recall, the aliased spectral density  $f_{\Delta}(\omega)$  is given by

$$f_{\Delta}(\omega) = \sum_{\mathbf{z} \in \mathbb{Z}^2} f\left(\omega + \frac{2\pi\mathbf{z}}{\Delta}\right) = \sum_{z_1 \in \mathbb{Z}} \sum_{z_2 \in \mathbb{Z}} f\left(\omega_1 + \frac{2\pi z_1}{\delta_1}, \omega_2 + \frac{2\pi z_2}{\delta_2}\right) \quad (48)$$

where  $-\pi/\Delta < \omega < \pi/\Delta$  and  $\Delta = (\delta_1, \delta_2)$  is the row and column spacing between sampling locations. When using the spectral domain to estimate the spatial covariance function, the use of the aliased spectral density creates an undesirable computational burden because (48) must be evaluated at every iteration of a maximization procedure. Our test of isotropy relies on the assumption that aliasing is not a concern so that  $f_{\Delta}(\omega) \approx f(\omega)$  (see Section 4.3). The causes and effects of aliasing for spatial data are not well understood. When developing spectral methods for geostatistical processes, the concerns raised by aliasing are usually avoided by either adopting a mixed asymptotic framework (Fuentes, 2002) or assuming that spacing between sampling locations is sufficiently small (Crujeiras and Fernández-Casal, 2010). For a finite sample where the covariance function is unknown, it is unclear when aliasing is a concern and what effects it may have on estimation. As a result we seek to understand when aliasing is a concern, i.e., when  $f_{\Delta}(\omega)$  is a poor approximation of  $f(\omega)$ , how to check for aliasing, and the effects of aliasing on covariance estimation. We propose a graphical diagnostic to detect aliasing and explore the effect of aliasing on estimation when using the Whittle approximation.

#### 4.8.2 What Causes Aliasing?

Because we cannot distinguish between frequencies differing by a multiple of  $2\pi/\Delta$  for a finite sample, the spectrum of the sample is restricted to the principle interval  $-\pi/\Delta < \omega < \pi/\Delta$ . Frequencies outside of this interval have an alias in this interval and are therefore “folded” down into the principle interval (Fuentes and Reich, 2010). For any finite sample of a continuous process with non-zero spacing between locations ( $\Delta > \mathbf{0}$ ), the aliased spectral density (48) will be different than the unaliased spectral density (22).

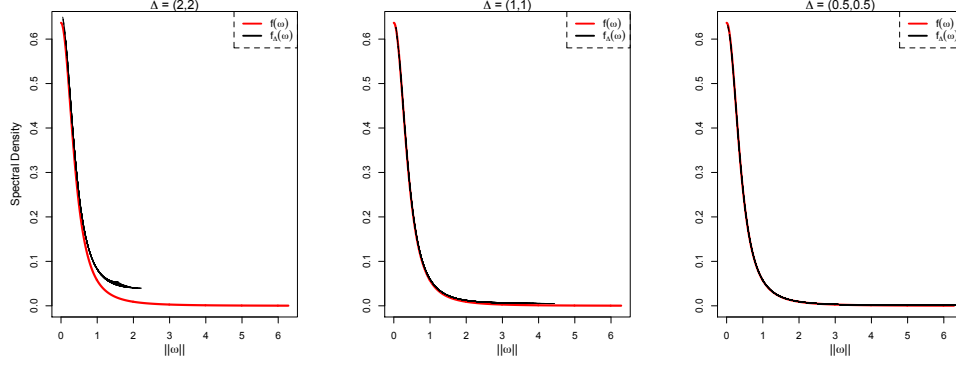


Figure 4.7: The Matérn spectral density (24) and aliased spectral density (48) for different values of  $\Delta$ , the spacing between sampling locations. For this example, we set  $\sigma^2 = 1$ ,  $\nu = 0.5$ , and  $\alpha^{-1} = 2$ . With these parameters, the difference between  $f_{\Delta}(\omega)$  and  $f(\omega)$  becomes negligible, indicating the aliasing is not a concern, once the spacing between sampling location is less than 1.

Another way to understand and motivate the aliasing concern is the concept of the power of the spatial process. The total power of the spatial process is the integral of the spectral density and is equal to the variance of the process (Schabenberger and Gotway, 2004, pg. 70-71). The spectral density describes how the power is distributed across frequencies, and the integral

$$\int_c^d \int_a^b f(\omega) d\omega_1 d\omega_2$$

is the contribution to the total power from frequencies  $\omega_1 \in (a, b) \times \omega_2 \in (c, d)$ . In order for  $f_{\Delta}(\omega) \approx f(\omega)$ , the frequencies  $\pi/\Delta$  must be high enough that the power from frequencies outside of the principle interval make a negligible contribution to the total power of the process (Fuentes and Reich, 2010). In other words, aliasing is not a concern if either the spacing between sampling locations is sufficiently small and/or the generating process is such that high frequencies contribute little to the total power. Next, we examine how the spacing between sampling locations and the characteristics of the covariance function affect the principle interval and power at high frequencies.

To demonstrate the aliasing phenomenon, we explore the difference between the unaliased ( $f(\omega)$ ) and aliased ( $f_{\Delta}(\omega)$ ) spectral density functions as a function of the spacing between

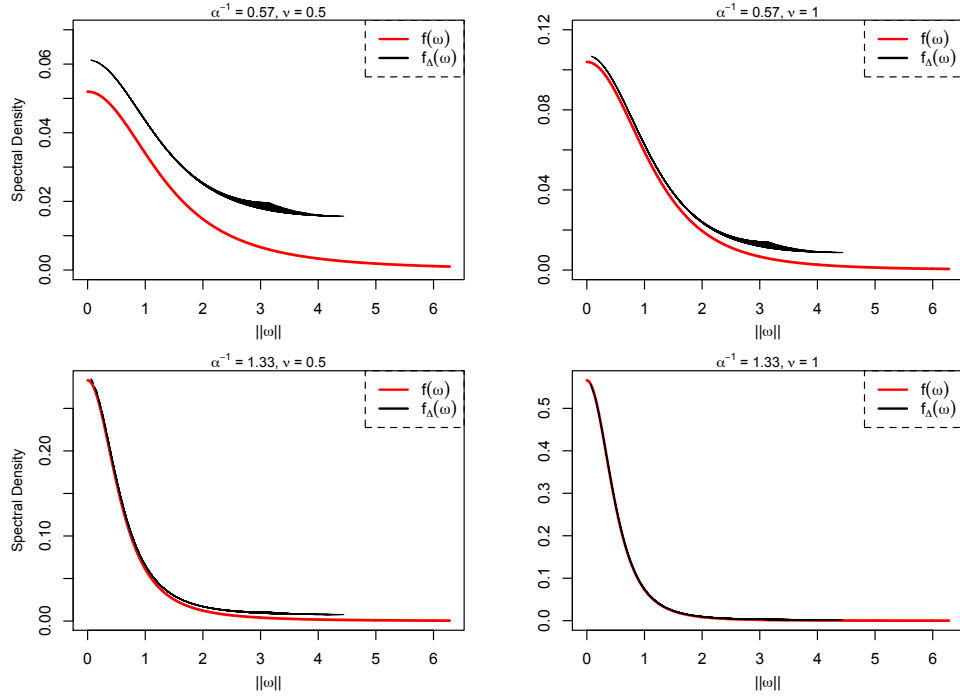


Figure 4.8: Plots of the Matérn spectral density (24) and aliased spectral density (48) for different values of the range ( $\alpha^{-1}$ ) and smoothness ( $\nu$ ) parameters. In addition to the spacing between sampling locations, the range and smoothness of the generating covariance function will also determine when aliasing is a concern, i.e., when  $f_{\Delta}(\omega)$  is a poor approximation of  $f(\omega)$ . As the range and smoothness increase, there is less mass at high frequencies; therefore, the aliased spectral density  $f_{\Delta}$  begins to look more like the unaliased spectral density  $f$ . For all four plots, the variance ( $\sigma^2$ ) is 1 and the sampling locations are assumed to be on the integer grid  $\mathbb{Z}^2$  (i.e.,  $\Delta = (1, 1)$ ).

sampling locations and the generating spatial covariance function. Figure 4.7 shows the difference between the Matérn aliased spectral density (48) and unaliased spectral density (24) as a function of the spacing between sampling locations  $\Delta$ . For this plot, we set  $\sigma^2 = 1$ ,  $\nu = 0.5$ , and  $\alpha^{-1} = 2$ , corresponding to an exponential covariance function with an effective range of approximately 5.8. We truncate the sum in (48) to the first 2,601 terms,  $z_1, z_2 \in \{-25, -24, \dots, 24, 25\}$ . In Figure 4.7 we can see that the difference between  $f(\omega)$  and  $f_{\Delta}(\omega)$  becomes negligible as the spacing between sampling locations decreases, as expected. Figure 4.8 displays the aliased spectral density for a Matérn covariance function using different values of the range ( $\alpha^{-1}$ ) and smoothness ( $\nu$ ) parameters. We assume  $\delta = \delta_1 = \delta_2 = 1$  and again truncate the sum in (48) to the first 2,601 terms. Based on these plots, we can begin to see when aliasing may be a concern. In particular, the aliased spectral density is a poor approximation of the spectral density when the range and/or smoothness parameters are small. In this case, there is more power (larger spectral density) at high frequencies, which is folded down into the principle interval. This causes a disagreement between the aliased and unaliased spectral density.

Thus, large spacing between sampling locations, a small range, small smoothness, or a combination there of will imply that aliasing is a concern. In practice the covariance function generating the spatial process is unknown; therefore, we cannot compare the spectral density to the aliased spectral density as in Figures 4.7 and 4.8. Likewise, the spacing between sampling locations is relative to the effective range, which is unknown. As a result, we develop a graphical technique using observed data to help determine whether or not aliasing is a concern.

### 4.8.3 A Graphical Diagnostic for Aliasing

In this section, we develop a graphical diagnostic to assess when aliasing may be a concern, i.e., when  $f_{\Delta}(\omega)$  is a poor approximation of  $f(\omega)$ . We check for aliasing by comparing periodogram values estimated using the full grid of sampling locations and sub grids of

sampling locations. If the estimated periodogram values on the full and sub grid are similar, it suggests that the spacing between sampling locations is small enough so that aliasing is not a concern. Because the Fourier frequencies (32) used to estimate the spectral density are a function of the distance between sampling locations and the sample size, the challenge in comparing periodogram values on the full and sub grids lies in matching Fourier frequencies that are “close”.

Here we describe one way to compare periodogram values from a full and sub grid, using an example. Suppose we have a parent grid that is a  $21 \times 21$  subset of  $\mathbb{Z}^2$  ( $\Delta = (1, 1)$ ). From this parent grid, we want to create square sub grids using as many of the sampling locations as possible. To this end, we create  $11 \times 11$  and  $10 \times 10$  sub grids with spacing  $\Delta = (2, 2)$ . Figure 4.9 plots the parent grid and two sub grids. Because the spacing between sampling locations is different from the parent grid and the number of rows and columns varies between the two sub grids, the set of Fourier frequencies (32) for the sub grids will be different than those from the parent grid. To check for aliasing, we want to compare estimates of the spectral density from the full and sub grids at frequencies that are “close”. The definition of close frequencies will depend on the problem (i.e., the number of sampling locations and grid spacing). Therefore, a visual inspection of the frequencies (e.g., see Figure 4.10) can help inform how they should be matched. For our example, we match sub grid frequencies with parent frequencies within 0.1 (i.e., sub grid frequencies  $(\omega_1^s, \omega_2^s)$  are close to parent frequencies  $(\omega_1^p, \omega_2^p)$  if  $\omega_i^p - 0.1 \leq \omega_i^s \leq \omega_i^p + 0.1$  for  $i = 1, 2$ ). Due to the symmetry of the periodogram (see Section 4.4), only frequencies with  $\omega_2 \geq 0$  are matched. Figure 4.10 displays the Fourier frequencies corresponding to the full and subgrids and the region where Fourier frequencies were matched to compare periodogram values.

We simulate mean 0, GRFs using two isotropic Matérn covariance functions (23) corresponding to a scenario where aliasing may be a concern and one where it may not. The first covariance function has a small range ( $\alpha^{-1} = 0.57$ ) and smoothness ( $\nu = 0.5$ ). The top left plot of Figure 4.8 displays the spectral density  $f(\omega)$  and aliased spectral density  $f_{\Delta}(\omega)$

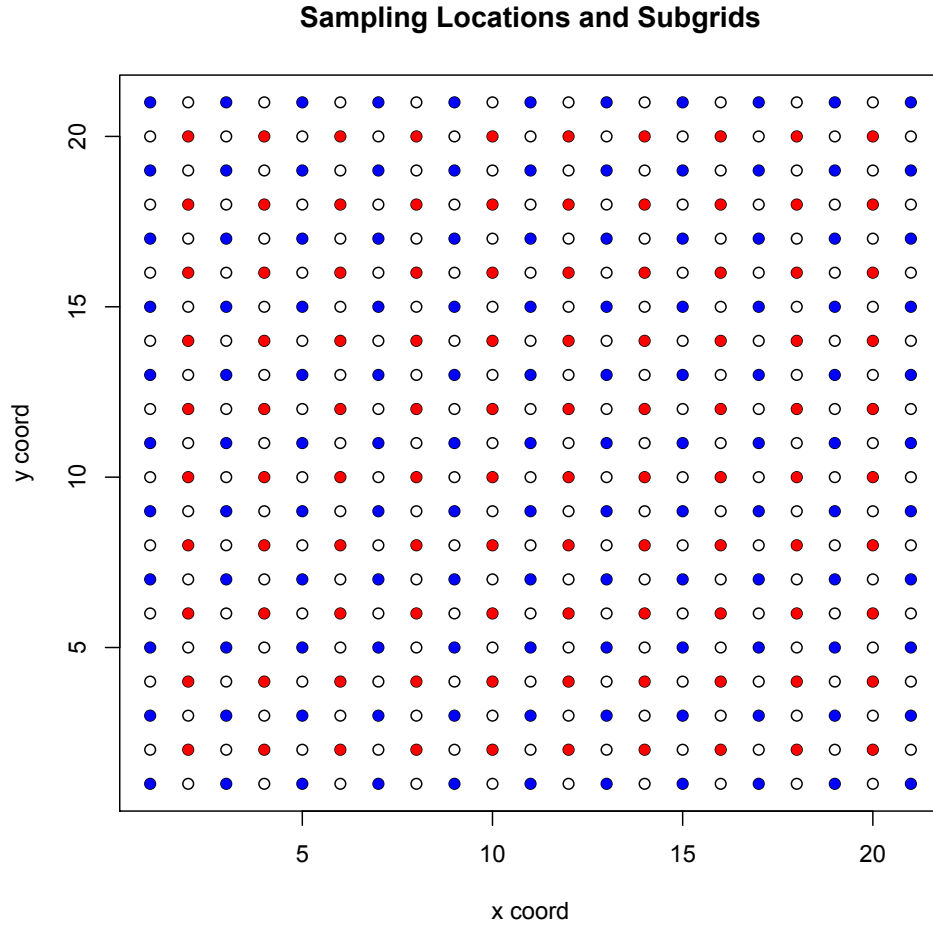


Figure 4.9: Plot of the sampling locations from the full grid and two subgrids. The black circles correspond to the sampling locations from the full  $21 \times 21$  grid of sampling locations with spacing  $\Delta = (1, 1)$ . The blue and red circles correspond to the sampling locations for an  $11 \times 11$  and  $10 \times 10$  subgrid, respectively, with spacing  $\Delta = (2, 2)$ .



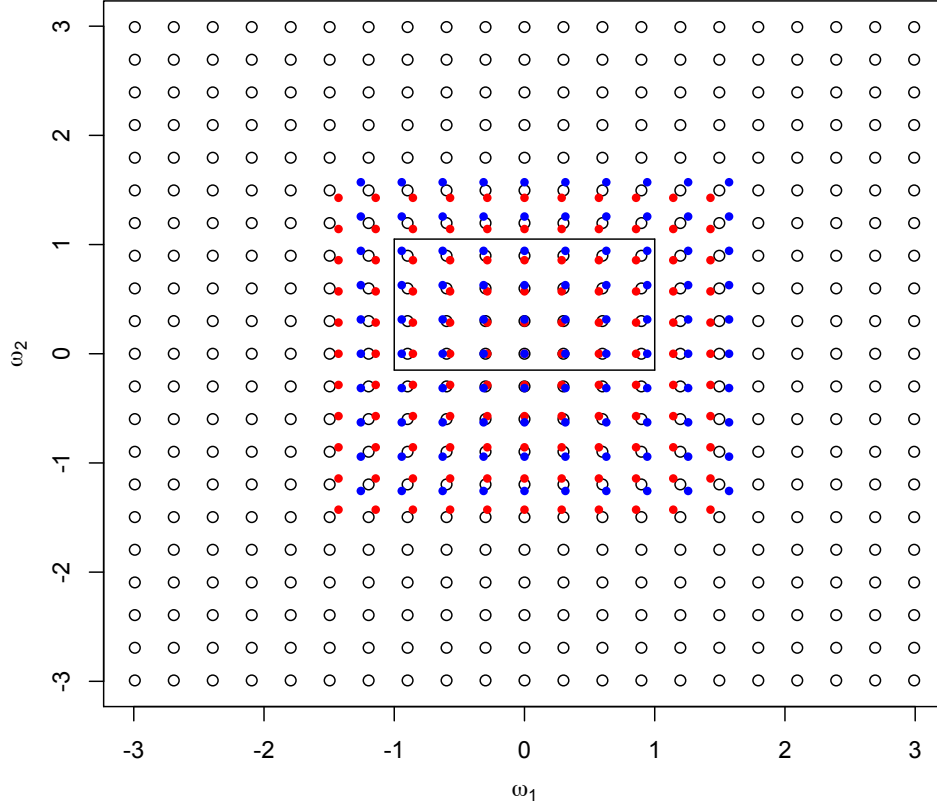


Figure 4.10: Plot of the Fourier frequencies from the full grid and two subgrids. The black circles correspond to the Fourier frequencies from the full  $21 \times 21$  grid of sampling locations with spacing  $\Delta = (1, 1)$ . The blue and red circles correspond to the Fourier frequencies for the  $11 \times 11$  and  $10 \times 10$  sub grids, respectively, with spacing  $\Delta = (2, 2)$ . To check for aliasing, we compare periodogram values at the Fourier frequencies for the full and sub grid. Frequencies within the black box in the middle of the plot are matched, and the periodogram values are compared at these frequencies.

corresponding to this covariance function. This plot suggests that aliasing is a concern because there is a noticeable difference between the spectral density and the aliased spectral density. Specifically, the aliased spectral density over-estimates the spectral density function. The second covariance function has a relatively large range ( $\alpha^{-1} = 1.33$ ) and smoothness ( $\nu = 1.5$ ). The bottom right plot of Figure 4.8 displays the spectral density  $f(\omega)$  and aliased spectral density  $f_{\Delta}(\omega)$  corresponding to this covariance function. This plot suggest that aliasing is a not a concern because the spectral density function and aliased spectral density function are nearly identical. Of course, we would not know the true spectral density in practice. However, we hope to use our graphical diagnostic to reach the same conclusions about aliasing concerns that are readily apparent in Figure 4.8.

For each covariance function, we simulate three realizations of the random field. After computing the periodogram values for the parent grid and two subgrids, we match the Fourier frequencies from the parent grid and two subgrids and plot the periodogram from the full grid versus the two subgrids. These plots can be seen Figure 4.11 for the three realizations. When the range and smoothness are small, there is more power (larger spectral density) at high frequencies that is folded down into the principle interval. This causes a disagreement in periodogram values between the parent and subgrids. More specifically, the periodogram values from the subgrid, which has larger spacing between sampling locations, tends to be larger than the periodogram values from the full grid. This disagreement is an indication that aliasing may be a concern. On the contrary, the periodogram values from the full and subgrids tend to agree when the range and smoothness are large, indicating the aliasing is not a concern.

Interpreting these diagnostic plots is analogous to interpreting a normal quantile-quantile (QQ) plot: an understanding of the appropriate conclusion takes repetition and the conclusion may be open to personal interpretation. The correlation between the periodogram values from the full and sub grids provides a supplement to the graphical diagnostic. A strong positive correlation between the values indicates that aliasing is not a concern. For

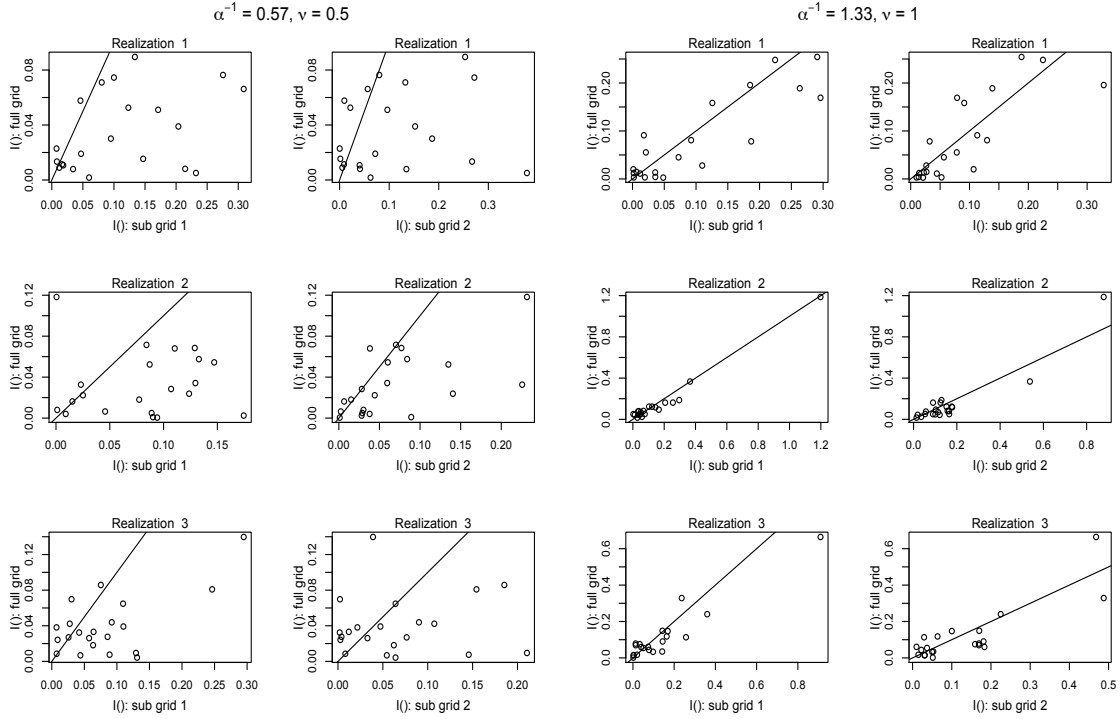


Figure 4.11: Comparison of periodogram values from the full and subgrids for 3 realizations of two random fields. The two columns on the left represent periodogram estimates from a RF with small range ( $\alpha^{-1}$ ) and smoothness ( $\nu$ ) parameters. For these columns, we see substantial disagreement between periodogram values from the full and sub grids, suggesting that aliasing is a concern. The two columns on the right represent periodogram estimates from a RF with relatively large range ( $\alpha^{-1}$ ) and smoothness ( $\nu$ ). For these columns, we tend to see agreement between periodogram values from the full and subgrids, suggesting that aliasing may not be a concern (i.e.,  $f_{\Delta}(\omega) \approx f(\omega)$ ).

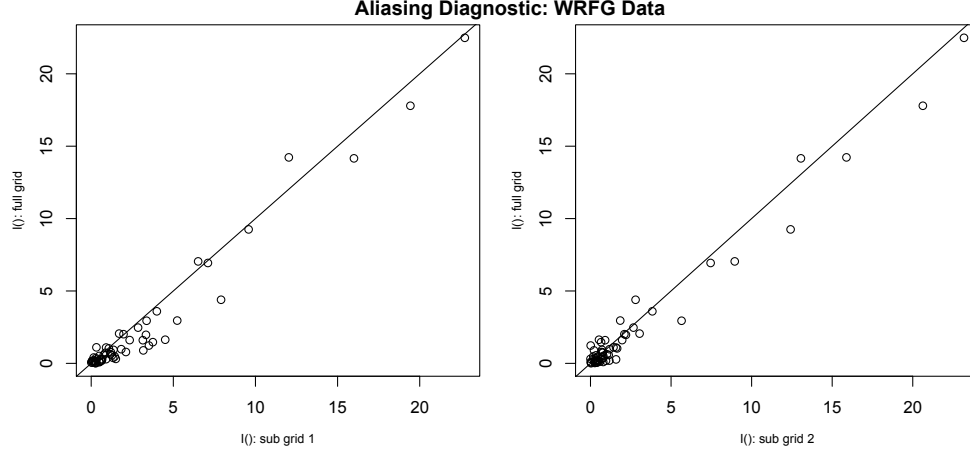


Figure 4.12: Check for aliasing in the WRFG residuals. We compare periodogram values estimated using the residuals from the WRFG data on the full grid and two sub grids of sampling locations. Based on the similarity of the periodogram values from the full grid and two sub grids, we can conclude that aliasing is not a concern for these data. As a result, our test of isotropy will not be affected by aliasing.

example, the correlations from the right two columns in Figure 4.11 tended to be greater 0.90. The correlations from the left two columns in Figure 4.11 were typically less than 0.5, but were sometimes as large as 0.8, indicating that the correlation alone is not sufficient for determining whether or not aliasing is a concern. Additional knowledge about the underlying spatial process can also supplement our graphical check for aliasing. Specifically, we have seen that a small range and/or smoothness can lead to aliasing concerns. Thus, a heat map (like Figure 4.5) or plot of the empirical covariogram, which give an indication of the range and smoothness, can supplement our graphical check for aliasing, much like a histogram can supplement a QQ plot.

We constructed our graphical diagnostic for aliasing using the residuals from the WRFG temperature data in Section 4.6. The plots are displayed in Figure 4.12 and indicate that aliasing is not concern for these data. As a result, we can assume that  $f_{\Delta}(\omega) \approx f(\omega)$ , which is required for our spectral test of spatial anisotropy.

#### 4.8.4 Effects of Aliasing: Whittle Approximation

One of the most well known spectral methods for spatial data is the Whittle approximation. In this section we explore the effects of aliasing on the estimates given by the Whittle approximation. An approximation to the Gaussian negative log likelihood (Whittle, 1954) is given by

$$\frac{n}{(2\pi)^2} \sum_{\omega \in \mathcal{F}} \log f(\omega) + I_n(\omega) f(\omega)^{-1} \quad (49)$$

where  $\mathcal{F}$  is the set of Fourier frequencies (32),  $f(\omega)$  is the spectral density (22),  $n = n_1 n_2$  is the sample size, and  $I_n(\omega)$  is the periodogram (26). Fuentes and Reich (2010, pg. 76) used the Whittle approximation for estimating covariance parameters and explored the difference in estimation when using both  $f_{\Delta}(\omega)$  and  $f(\omega)$  for a set of observed data. For other data sets, however, it is unclear what effect using  $f_{\Delta}(\omega)$  rather than  $f(\omega)$  in (49) will have on parameter estimates. An understanding of these effects is important for accurate parameter estimation and computational savings.

To explore this effect, we simulated 300 realizations of a mean 0 Gaussian random field and estimated the covariance parameters via the Whittle approximation (49). We limited our number of simulations to 300 because of the large computation time needed when using the aliased spectral density in (49). For each realization, there were  $n = 441$  sampling locations on a  $21 \times 21$  subset of  $\mathbb{Z}^2(\Delta = (1, 1))$ . We simulated RFs with Matérn covariance functions with  $\sigma^2 = 1$ . We varied the range ( $\alpha^{-1}$ ) and smoothness ( $\nu$ ) parameters to emulate a scenario where aliasing is a concern and one where it is not.

For the first scenario, we set  $\alpha^{-1} = 0.57$  and  $\nu = 0.5$ , and a plot of the spectral density and aliased spectral density with these parameters and unit spacing between sampling locations is given in the top left plot of Figure 4.8. Our graphical diagnostic for aliasing under this scenario is plotted in the first two columns of Figure 4.11. From both of these plots, it is evident that aliasing is a concern. For the second scenario, we set  $\alpha^{-1} = 1.33$  and  $\nu = 1.0$ . A plot of the spectral density and aliased spectral density with these parameters and unit

spacing between sampling locations is given in the bottom right plot of Figure 4.8. Our graphical diagnostic for aliasing under this scenario is plotted in the last two columns of Figure 4.11. Based on these plots, it appears that aliasing is not a concern.

During estimation, we constrained the smoothness parameter  $0.5 \leq \nu \leq 2$  because smoothness parameters outside of this range are typically considered unrealistic for geostatistical data. The results of the parameter estimates using both  $f(\omega)$  and  $f_{\Delta}(\omega)$  in the Whittle approximation are given in Figures 4.13 and 4.14. Figure 4.13 corresponds to the first scenario, where aliasing is a concern. Figure 4.14 corresponds to the second scenario, where aliasing is not a concern.

When aliasing is a concern, the results in Figure 4.13 indicate that using  $f_{\Delta}(\omega)$  over  $f(\omega)$  in (49) results in better estimation of variance  $\sigma^2$  and range  $\alpha^{-1}$  parameters. When we use the original spectral density  $f(\omega)$ , we notice bias in the variance and range parameters. It is important to note that Guyon (1982) proves Whittle estimates are biased, but this bias can be corrected via tapering (Dahlhaus and Künsch, 1987). Estimation of the smoothness  $\nu$  is better when using  $f(\omega)$  than  $f_{\Delta}(\omega)$ , but in many applications, the smoothness  $\nu$  is often treated as a known, fixed constant.

When aliasing is not a concern, the results in Figure 4.14 show that estimation of the variance and range parameters are similar whether  $f_{\Delta}(\omega)$  or  $f(\omega)$  is used in the Whittle approximation. For the smoothness parameter, use of  $f(\omega)$  results in a large downward estimation bias while using  $f_{\Delta}(\omega)$  creates a relatively slight upward bias. Additional simulations using different parameter values and spacings between sampling locations gave similar results to those displayed in Figures 4.13 and 4.14 but are not included here.

Considering the results from Figures 4.13 and 4.14, it seems that we should always use the aliased spectral density  $f_{\Delta}(\omega)$  in lieu of the original spectral density  $f(\omega)$  when performing estimation via the Whittle approximation (49). However, using  $f_{\Delta}(\omega)$  comes with a large increase in computational cost. For example, under our first scenario with  $n = 441$  sampling

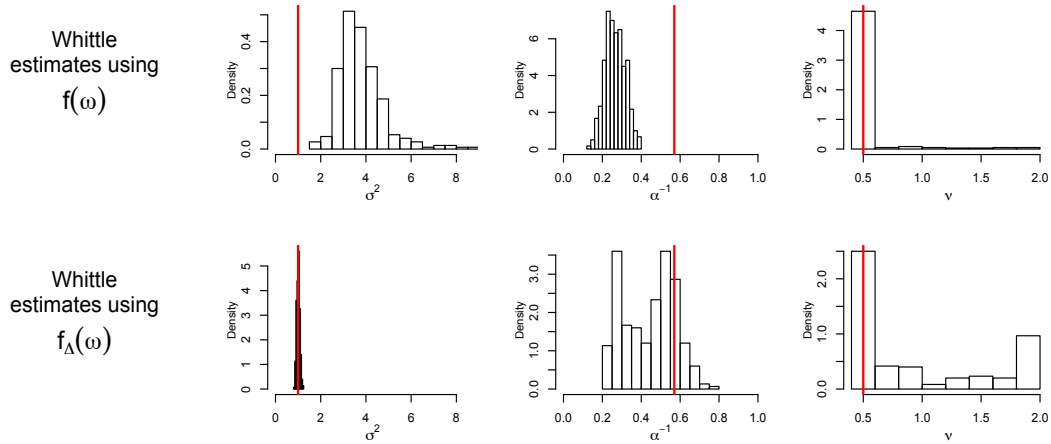


Figure 4.13: Covariance parameter estimates using the Whittle approximation when aliasing is a concern. The true covariance parameters are  $\sigma^2 = 1$ ,  $\alpha^{-1} = 0.57$ , and  $\nu = 0.5$ . The first row corresponds to using the original spectral density  $f(\omega)$  for estimation in the Whittle approximation (49), while the second row corresponds to using the aliased spectral density  $f_{\Delta}(\omega)$  for estimation in (49). Because aliasing is a concern in this scenario, estimation of the variance  $\sigma^2$  and range  $\alpha^{-1}$  tends to improve when using  $f_{\Delta}(\omega)$  rather than  $f(\omega)$ .

locations, it took 0.05 seconds to compute Whittle estimates using  $f(\omega)$  but 115.0 seconds to compute estimates using  $f_{\Delta}(\omega)$ .

#### 4.8.5 Discussion: Aliasing

When using spectral methods for geostatistical spatial data, the aliasing phenomenon is typically ignored and not well understood. In addition to a large spacing between sampling locations, we have explored other factors that can cause aliasing to be a concern, such as a small range or smoothness of the underlying spatial process. These developments elucidate when aliasing is a potential concern. We also developed a graphical diagnostic to help assess whether or not aliasing is a concern based on observed data. Combined with some understanding of the underlying spatial process, this diagnostic can help inform whether or not aliasing will have an effect on inference that uses the spectral domain (e.g., our test of isotropy in Section 4.4 or the Whittle approximation (49)). Aliasing concerns are sometimes

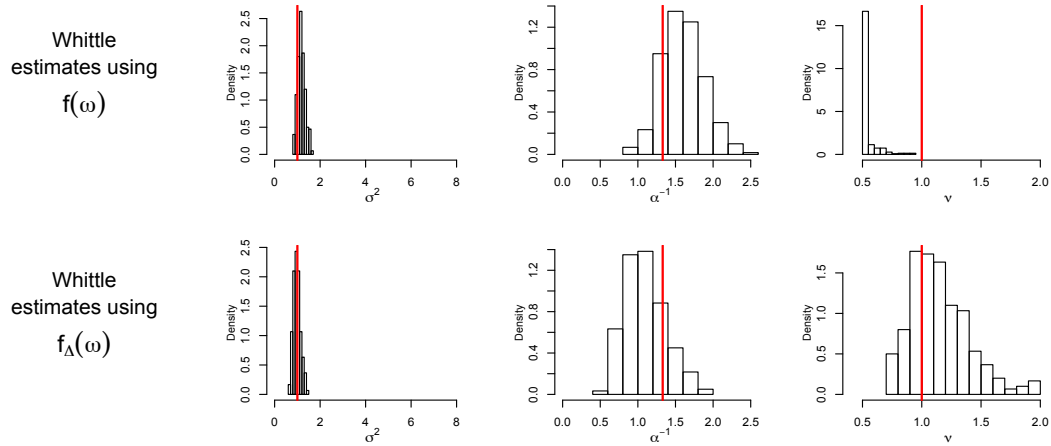


Figure 4.14: Covariance parameter estimates using the Whittle approximation when aliasing is not a concern. The true covariance parameters are  $\sigma^2 = 1$ ,  $\alpha^{-1} = 1.33$ , and  $\nu = 1.0$ . The first row corresponds to using the original spectral density  $f(\omega)$  for estimation in the Whittle approximation (49), while the second row corresponds to using the aliased spectral density  $f_{\Delta}(\omega)$  for estimation in (49). Because aliasing is not a concern in this scenario, estimation of the variance  $\sigma^2$  and range  $\alpha^{-1}$  are similar whether or not we use  $f_{\Delta}(\omega)$  or  $f(\omega)$ .

avoided by assuming that the spacing between sampling locations is sufficiently small, but our simulations in Section 4.8.4 suggest that the assumption that aliasing is not a concern should be given more careful consideration.



**A CALIBRATION-CAPTURE-RECAPTURE MODEL****5.1 Introduction**

The greenhouse gas methane ( $CH_4$ ) has a larger global warming potential than  $CO_2$  in the first two decades after its release and is considered “the second most important anthropogenic greenhouse gas” (Boucher et al., 2009). Methane is the primary component of natural gas. In several urban areas of the United States, pipelines used to deliver natural gas have deteriorated and become susceptible to developing leaks. Identifying the locations of methane emitting sources, such as natural gas leaks, and quantifying their emission rate is essential for reducing greenhouse gas emissions and understanding the condition of underlying infrastructure. Locating methane sources and estimating their magnitude is traditionally time-consuming, but recently mobile methane detection technologies have been developed and employed. By measuring atmospheric methane concentrations at a high frequency with these mobile technologies, scientists and utility companies hope to identify methane source locations and quantify their magnitude.

In an effort to locate and quantify the magnitude of methane emitting sources, such as natural gas leaks, the Environmental Defense Fund (EDF) has initiated a survey of atmospheric methane concentrations in designated sampling regions within several urban areas (von Fischer et al., 2017). As part of this survey, EDF has partnered with the Google Street View (GSV) program to place atmosphere analyzers on GSV vehicles. The analyzers on these vehicles have enabled the collection of precision methane concentrations and their associated locations on a fine spatial-temporal scale over large metropolitan areas. Because this

technology is still in its infancy, there is a need to develop algorithms to extract meaningful information from the data.

In this chapter we present a model to estimate the total number of methane sources within a sampling region and the emission rate of each of those sources using data collected by mobile detection platforms. Drivers of the GSV vehicles were instructed to drive every street within designated sampling areas at least two times. Due to the logistics of driving in urban areas, some stretches of roadway were sampled only twice, while others were sampled over 10 times. Detection of methane sources is imperfect due to variables such as wind, emission rate, and source proximity to the GSV vehicle. Thus, there are sources within a sampling region that are never detected by the GSV car’s instrumentation. This imperfect detection, and the aforementioned repeated sampling, lead us to develop a capture-recapture model to estimate the total number of methane sources.

Capture-recapture models have been widely developed and applied to estimate animal population size, survival, and other quantities (see, e.g., Nichols, 1992). Over a series of sampling attempts, individuals in the population are either observed (i.e., captured) or unobserved. Letting  $t \in \{1, 2, \dots, T\}$  denote the sampling attempts and  $i \in \{1, 2, \dots, N\}$  indicate individuals, each individual produces a capture history  $\mathbf{Y}_i = (Y_{i1}, Y_{i2}, \dots, Y_{iT})^\top$ , where  $Y_{it} = 1$  if individual  $i$  is captured on sampling attempt  $t$  and 0 otherwise. Typically, a subset individuals will never be captured. Thus, as in our study, we obtain capture histories on  $n \leq N$  individuals (i.e., on individuals with  $\sum_{j=1}^t Y_{ij} \geq 1$ ). A common goal of a capture-recapture model is to estimate the total population size,  $N$ .

One of our goals is to estimate the total number of methane sources  $N$  in a sampled region. The use of a capture-recapture model for data collected by a mobile detection platform presents several unique challenges. First, unlike traditional capture-recapture data, individual methane sources are never “marked” or directly observed by our sampling mechanism. Instead, atmospheric  $CH_4$  concentrations measured by instruments on the GSV vehicles are indicators of a source’s presence. Second, a source’s emission rate, which affects the probabil-

ity of detection, is never directly observed. Additionally, the measured  $CH_4$  concentrations, a proxy for source emission rate, varies by sampling attempt due to environmental variables such as wind. Finally, the number of sampling attempts varies by individual for our methane sources.

In this chapter we develop a new capture-recapture model, which we call the calibration-capture-recapture (CCR) model, to analyze data collected by mobile methane detection platforms. Our CCR model addresses the aforementioned challenges and leverages data from controlled methane release experiments and data collected by GSV vehicles to provide inference for a number of desired quantities, including the total number of sources and source emission rate. The CCR model is an integrated population model because it leverages data from multiple sources to provide inference for characteristics of a population of methane sources (e.g., see Schaub and Abadi, 2011; Ruiz-Gutierrez et al., 2016). In this chapter our notation and terminology focuses on the methane monitoring system. However, as we describe further in Section 5.7, our CCR model can be used for other applications, such as analyzing data collected by other mobile detection platforms or for more traditional animal capture-recapture studies.

The remainder of the chapter is organized as follows. We provide details regarding data collection and processing and the controlled methane release experiments in Section 5.2. In Section 5.3, we provide background information on calibration and capture-recapture models and develop our CCR model. Details on our Markov chain Monte Carlo (MCMC) sampling algorithm are located in Sections 5.4 and 5.5. In Section 5.6 we apply our CCR model to data collected in two U.S. cities. Because there is a non-disclosure agreement between the EDF and the utility companies servicing cities where sampling occurs, we redact the city names in Section 5.6. Finally, Section 5.7 concludes with a discussion and presents avenues of future research.

## 5.2 Data Collection and Processing

In this section we briefly discuss data collection and processing. We used GSV vehicle data from controlled methane release experiments and collected in two U.S. cities. The data processing decisions we present here are based on the work and advice of our subject-area collaborators. More information regarding the instrumentation and study design are provided in von Fischer et al. (2017).

For the field data, atmospheric  $CH_4$  concentrations were collected approximately every 0.5 seconds as GSV cars drove roadways in designated sampling areas. At a typical driving speed of 40 km/h, methane concentrations were recorded at approximately 5.5 meter spatial intervals. While ambient atmospheric  $CH_4$  concentrations average around 1.8-2.2 parts per million (ppm), air analyzers on GSV vehicles measured concentrations down to 0.001 ppm. For each methane measurement, instrumentation also recorded the GSV car’s location via global positioning system (GPS). Methane concentration readings are not instantaneous because it takes time for air to move from the intake, located on the car’s bumper, into the measuring instrument, located inside the car. As a result,  $CH_4$  measurements were adjusted after sampling based on vehicle speed to correct the misalignment between  $CH_4$  measurements and GPS locations. A small number of the methane observations ( $< 1\%$ ) were discarded due to GPS error or because the GSV car was traveling too fast. Although drivers of GSV vehicles were instructed to drive every street in the sampling region at least two times, the practicalities of navigating the sampling region resulted in some stretches of roadway being sampled twice while others were driven as many as ten or more times. Recording  $CH_4$  concentrations on this fine temporal and large spatial scale leads to a large quantity of data. As one example, instruments in a GSV car recorded over 1.6 million observations during sampling in one U.S. city.

From the raw methane concentration measurements, we identified locations with elevated methane levels. Elevated methane concentrations are evidence that a methane emitting

source is close to the GSV vehicle. Because background methane concentrations change on a spatial and temporal scale, we followed von Fischer et al. (2017) and defined a baseline  $CH_4$  concentration for each methane concentration measurement by the GSV car. For a given  $CH_4$  measurement, the baseline is defined as the average of all methane readings that occurred within a five minute window and within 400 meters of the given reading. An elevated methane reading was defined as any concentration that was greater than 110% of the baseline value. Using these elevated readings, we defined a source detection as at least four consecutive elevated methane concentration readings as the GSV car travels throughout the sampling region. Using four consecutive elevated readings aids in the removal of methane sources that are not of interest for our analysis. For example, transient methane sources such as vehicle emissions may only produce one or two elevated concentration readings as they pass by the GSV car.

For a series of elevated readings (i.e., a detection), we used the GPS location associated with the maximum  $CH_4$  concentration as the estimated source location. Therefore, each estimated source location is a set of GPS coordinates and an associated  $CH_4$  concentration. After the GSV car had driven the entire sampling region two or more times, we processed the raw methane concentration data from a sampling region using the aforementioned methods. After this processing, we had a spatial map of estimated source locations and associated  $CH_4$  concentration. Source locations from different sampling attempts that are spatially close are likely to be detections of the same methane source. Environmental variables, such as wind, can move methane gas plumes away from their origin. Thus, a methane emitting source's estimated location and measured  $CH_4$  concentration can change from sampling attempt to sampling attempt. As a result, we aggregated estimated source locations that are spatially close into source sites. We created source sites by grouping together estimated source locations that are within 20 meters of each other. For source site  $i$ , let  $j = 1, \dots, m_i$  denote the number of detections at that site (i.e., the number of source locations aggregated to create the site). The site centroid, an estimate of the source's origin, was estimated using a

weighted spatial average of the estimated source locations composing that site. The weights were based on the  $CH_4$  concentration at each of the source locations. Locations with higher  $CH_4$  concentrations, which indicate the GSV vehicle is likely in close proximity to the source, were given larger weight. Specifically, the weight associated with estimated source location  $j$  at source site  $i$  is

$$w_j = \frac{(CH_4)_j}{\sum_{j=1}^{m_i} (CH_4)_j}.$$

The centroid of site  $i$  is the weighted spatial average

$$(\text{lon}, \text{lat})_i = \sum_{j=1}^{m_i} (w_j)(\text{lon}, \text{lat})_j,$$

where “lon” and “lat” denote longitude and latitude coordinates. Finally, the source site is defined as a circle with a radius of 20 m around the site centroid.

A sampling attempt, or drive-by  $t$ , is defined as the GSV car driving through the source site. The total number of sampling attempts (number of times the GSV car drove through the site), denoted  $d$ , varied by site. We let  $i_d$  denote the  $i^{th}$  site in the region sampled  $d$  times. For each drive-by of site  $i_d$ , we let  $Y_{i_d t} = 1$  if the source was detected (i.e., elevated readings occurred) and  $Y_{i_d t} = 0$  if the source was not detected. The number of detections of source  $i_d$  is

$$m_{i_d} = \sum_{t=1}^d Y_{i_d t}.$$

Figure 5.1 displays a cartoon that illustrates the sampling and source detection process. As  $d$  increases for a given roadway, sites along that roadway are more likely to be detected. Note that the set of sites sampled  $d$  times does not need to be spatially connected and will typically be dispersed across road segments in spatially disjoint regions.

Finally, we used additional criteria to aid in the removal of methane sources that are not of interest for our analysis. We removed any site where  $\sum_{t=1}^d Y_{i_d t} = 1$  and  $d \geq 5$ . In other words, any site where only 1 detection occurred over 5 or more sampling attempts. Again,

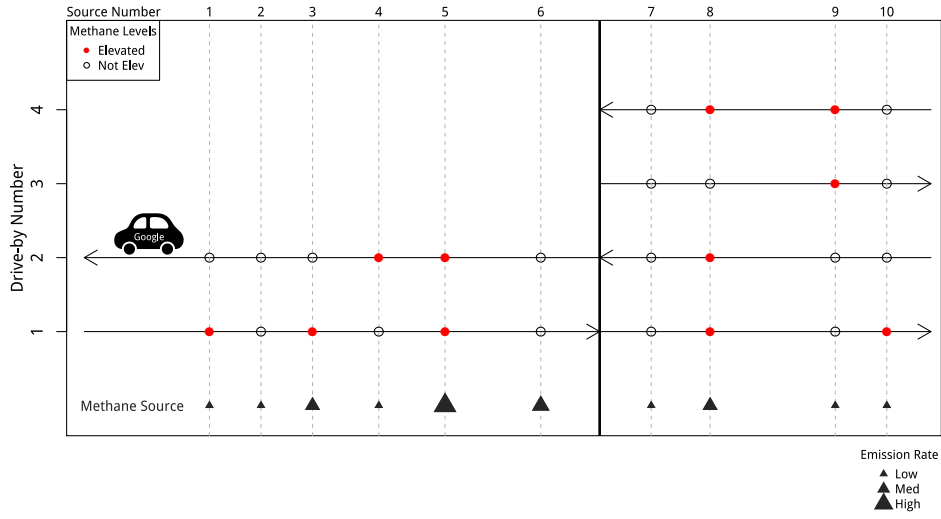


Figure 5.1: Cartoon illustrating the process of detecting methane sources using Google Street View (GSV) vehicles. For each drive-by of a methane source, instrumentation on the GSV vehicle measures either elevated methane levels (i.e., detects the source) or background methane levels (i.e., does not detect the source). Because they emit larger, more concentrated plumes of methane gas, sources with large emission rates have a higher probability of detection than sources with small emission rates. Due to the logistics of driving in an urban area, the number of sampling attempts varies by individual. Some sources, like source numbers 2, 6, and 7 in this cartoon, are never detected by the GSV car instruments.

this criteria aids in the removal of transient methane sources such as vehicle emissions. We also removed any site where unusually large plumes of elevated methane levels were detected. An unusually large plume is defined as observing elevated readings over a distance of greater than 160 meters. We used this criteria to remove larger, non-amendable methane sources such as wetlands or landfills.

After creating source sites and removing sites that are not of interest, we had  $n$  observed source sites. Let  $\mathcal{D}$  denote the set of unique sampling attempts across the  $n$  sites. For our study,  $\mathcal{D} = \{2, 3, 4, 5, 6, 7, 8, 9, 10\}$ , implying that sources were detected in areas sampled 2-10 times. Let  $n_d, d \in \mathcal{D}$  denote the number of sources detected in the region sampled  $d$  times and  $i_d = 1, 2, \dots, n_d$  denote individuals within the region sampled  $d$  times.

### 5.3 A Calibration-Capture-Recapture Model

In this section we propose a new integrated population model that we call calibration-capture-recapture (CCR). We begin this section by describing the controlled methane release experiments and a calibration model for estimating the source emission rate. We then discuss a capture-recapture appropriate for our data and the issue of estimation of the probability of source detection. Next we describe parameter-expanded data augmentation and how it is used in the CCR model. We conclude this section by combining the calibration and capture-recapture models to fully specify the CCR model for the mobile methane detection application.

#### 5.3.1 Calibration and Controlled Releases

In this section we describe statistical calibration and the controlled methane release experiments. Statistical calibration and the particular case of univariate linear statistical calibration, which we use here, have been widely studied (see Osborne (1991) for a review). The basic idea behind univariate linear statistical calibration is to first perform a calibration experiment where  $C$  pairs of observations  $(x_c, Y_c)$  are used to estimate the linear relationship



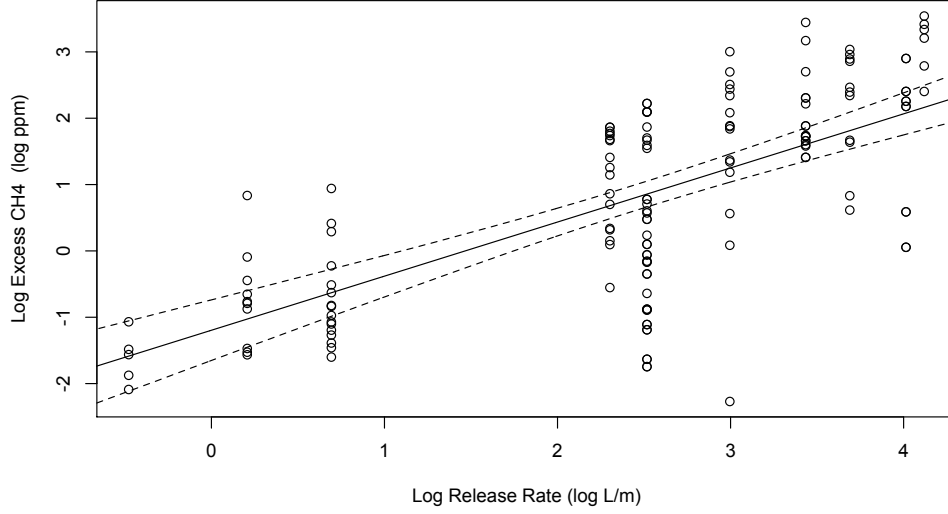


Figure 5.2: Scatter plot of the relationship between  $\log(\text{release rate})$  and  $\log(\text{excess } CH_4)$  for the controlled methane releases. Because the distance from the GSV car to the methane emitting source is unknown in the field, we do not include the known distance from the controlled release experiments in this plot. The solid line is the estimated simple linear regression relationship using the CCR model applied to City B. The dashed lines display the 95% posterior interval for the regression line.

between  $x$  and  $Y$ . Typically,  $x_c$  is a precise measurement that is either expensive or difficult to obtain and  $Y_c$  is a relatively inexpensive or easy to obtain measurement that serves as a surrogate for the desired measure  $x_c$ . Next, we obtain new values of the surrogate,  $Y_j^{(new)}$ ,  $j = 1, \dots, m$ , from an unknown value of the precisely measured variable  $X^{(new)}$ . Finally, using the estimated relationship based on the calibration data, we desire inference for the unobserved value  $X^{(new)}$ .

In our setting, methane concentrations measured by the GSV car's instrumentation are a surrogate for a source's emission rate. Measuring a source's emission rate is labor and time intensive and potentially dangerous based on traffic conditions. Conversely, measuring atmospheric methane concentrations with mobile detection platforms is relatively quick and automated.

To understand the relationship between a methane source’s emission rate and the measured methane levels, von Fischer et al. (2017) performed controlled methane releases and drove GSV vehicles through the resulting plumes. Using the controlled release data, we can develop a calibration model to estimate a new source’s unknown emission rate using measured  $CH_4$  levels. During the controlled release experiments, experimenters varied the distance of the GSV vehicle from the source, the rate of the methane release, and the surrounding conditions (e.g., different types of streets). The range of known release rates varied from 0.62 L/min to 61.5 L/min. This range was chosen to reflect the range of possible leak sizes that might be observed in the field, and a total of 10 different release rates were chosen for testing. For each known release rate, scientists drove the GSV car through the resulting plume and recorded atmospheric methane concentrations. Although the number of passes varied for each emission rate, multiple passes of each known emission rate provided a total of  $C = 150$  replications. An analysis of the controlled release data indicated that the maximum  $CH_4$  on each drive-by was the best proxy for the emission rate. During field sampling the maximum  $CH_4$  is affected by unobserved variables such as the GSV car’s distance from source, wind direction, and source surface expression. Surface expression refers to the nature of how methane gas is released into the atmosphere. For example, gas may be moving from soil to atmosphere through several cracks in the roadway or migrating to escape the soil at the edge of pavement. Because the maximum  $CH_4$  will also depend on the background methane levels, we define the excess  $CH_4$  as the maximum  $CH_4$  observed minus the  $CH_4$  baseline when a source is detected (i.e., when a series of elevated readings occurs during a drive-by),

$$X_c = \text{excess } CH_4 = (\max CH_4 - \text{baseline } CH_4).$$

for  $c = 1, 2, \dots, C$ . The relationship between the known emission rate, denoted  $v_c$ , and observed excess  $CH_4$   $X_c$  for the controlled releases is displayed on the log scale in Fig-

ure 5.2. Based on the controlled release study, we use simple linear regression to model the relationship between the known log emission rate ( $\log v$ ) and log excess methane ( $\log X$ ).

When a source is detected ( $Y_{idt} = 1$ ) by the GSV instrumentation in the field, we calculate the excess methane  $X$ . We aim to estimate the source’s latent, unknown methane emission rate  $V$  using these excess methane values and the data from the controlled release experiments. To achieve this goal we employ a Bayesian calibration model as follows. From the controlled release experiments, we observe  $C$  pairs of data  $(v_c, X_c)$ . We model the relationship between the known release rate  $v$  and excess methane  $X$  on the log scale as

$$\log X_c = \beta_0 + \beta_1 \log v_c + \epsilon_c \quad (50)$$

for  $c = 1, \dots, C$ , where  $\epsilon_c \sim N(0, \sigma^2)$ . Let  $m_{id} = \sum_{t=1}^d Y_{idt}$  denote the number of detections and  $V_{id}$  the unknown emission rate for methane source  $i_d$ . For each detection of source  $i_d$ , we assume

$$\log X_{idj} = \beta_0 + \beta_1 \log V_{id} + \epsilon_{idj} \quad (51)$$

where  $j = 1, \dots, m_{id}$ , and  $\epsilon_{idj} \sim N(0, \sigma^2)$ . To account for uncertainty in source emission rate, we will couple this calibration model with our capture-recapture model.

Our focus is on developing a capture-recapture model to estimate the number of unobserved sources while accounting for uncertainty in source emission rate. Therefore, we do not consider more complicated emission rate estimation schemes here. We provide further discussion on the challenges associated with emission rate estimation and suggested improvements in Section 5.7.

### 5.3.2 Capture-Recapture and Detection Probability

Here we describe the capture-recapture model appropriate for our mobile methane detection application. The earliest capture-recapture models made the simplifying assumption that capture probabilities are the same for each individual and sampling attempt. Models

with this simplifying assumption are often referred to as model  $M_0$  in the capture-recapture literature (e.g., Otis et al., 1978). A number of models have been developed that relax this assumption (Amstrup et al., 2010). One particular class of capture-recapture models, referred to as model  $M_h$ , accounts for heterogeneity in detection probability between individuals in the population (Pollock, 2002). For example, an individual’s detection probability may be modeled as a random effect (e.g., Coull and Agresti, 1999) or as a function of an individual’s covariate(s) such as weight and/or sex (e.g., Royle, 2009).

We use capture-recapture model  $M_h$  to estimate the number of methane sources within a designated sampling region, and we justify its use here. Because the controlled releases were performed with the primary goal of calibrating an emission rate estimation algorithm, experimenters drove the vehicle downwind of the methane source during the controlled releases. As a result of driving downwind, sampling instruments typically recorded elevated methane levels, but the probability of elevated readings (detections) decreased as the known emission rate decreased. To further understand the stochastic nature of detecting methane sources in real urban conditions, von Fischer et al. (2017) conducted extensive sampling in one U.S. city over a 9 month period. During this sampling, 138 sources were detected, each of which were driven past an average of 20 times. The probability of detection (elevated readings) of the sources varied by estimated source emission rate: 0.35 for small sources (0-6 L/min), 0.63 for medium sources (6-42 L/min), and 0.74 for large sources (42+ L/min).

Based on the aforementioned controlled methane release experiments and data from known methane sources, a source’s emission rate affects the probability of detection (i.e., the probability of being “captured” by the GSV car’s instrumentation). Specifically, sources that have a large emission rate have a larger probability of detection than sources with a small emission rate because they tend to create larger, more highly concentrated plumes of methane gas. We model the detection probability of source  $i_d$  as function of the estimated

source emission rate:

$$\text{logit}(p_{i_d}) = \alpha_0 + \alpha_1 V_{i_d}. \quad (52)$$

In a standard capture-recapture model  $M_h$ , detection probabilities are modeled as a function of an individual’s known, observed covariate(s). We treat the emission rate as latent because we do not observe it directly and instead observe a surrogate, excess  $CH_4$ , for the emission rate. To estimate the unobserved emission rate, we used the source’s empirical detection probability, observed excess  $CH_4$  value(s), the controlled release data, and our calibration model in (50) and (51).

### 5.3.3 Parameter-Expanded Data Augmentation (PX-DA)

In this section we describe parameter-expanded data augmentation and how we use it to develop and estimate our calibration-capture-recapture (CCR) model. We used a hierarchical model to combine the calibration and capture-recapture models to create a CCR model. Our model provides estimates and uncertainty for the number of methane sources and sampling region-wide methane emissions while accounting for undetected methane sources. The model also provides estimates and uncertainty for the detection probability and individual sources’ emission rate.

We estimate the CCR model parameters using Bayesian parameter-expanded data augmentation (PX-DA) (Liu and Wu, 1999). For capture-recapture modeling, data augmentation involves adding a large number of all-zero encounters,  $\mathbf{Y}_i = \mathbf{0}^\top$ , to the observed capture data. These all-zero encounters represent a combination of unobserved methane sources and “artificial” sources added for ease of computation. In the capture-recapture setting, adding all-zero encounters equates to model expansion (Royle and Dorazio, 2012). This model expansion as a result of data augmentation is referred to as parameter-expanded data augmentation (Liu and Wu, 1999) and distinguished from data augmentation in the sense of Tanner and Wong (1987). Specifically, adding a large number of all-zero encounters represents an zero-inflation

of the the observed data. After augmenting the data with all-zero encounters, we have a  $K$  (observed plus augmented) methane sources. We can think of the  $K$  methane sources as a super-population of which there is a subset of sources available for capture (Royle et al., 2007). We introduce the partially observed, latent, binary variable  $Z_i$  which represents membership in the capture-able population ( $Z_i = 1$  if individual  $i$  is in the capture-able population). By assuming that  $Z_i|\psi \sim \text{Bernoulli}(\psi)$ , the parameter  $1 - \psi$  represents the proportion of zero-inflation (e.g., the proportion of all-zero encounters which are not members of the capture-able population). Thus, we have augmented our data and expanded our model by introducing all-zero encounter histories, the indicator variable  $Z_i$ , and the zero-inflation parameter  $\psi$ .

Royle et al. (2007) first applied PX-DA to analyze capture-recapture data, and Royle and Dorazio (2012) offers further insight into the computational benefits of PX-DA coupled with Bayesian analysis for capture-recapture models. We choose to perform estimation using PX-DA and Bayesian methods because of their computational convenience. In particular, because our model contains latent, individual parameters for a large number of sources and several conditional distributions yield convenient forms for Gibbs sampling, PX-DA and a Bayesian hierarchical model readily yield inference for a number of quantities of interest.

#### 5.3.4 Calibration-Capture-Recapture for Mobile Detection

Here we specify the CCR model for our mobile methane detection application. Unlike a typical capture-recapture setting, the number of sampling attempts,  $d$ , will vary by individual in our model. For example, some source sites were sampled two times while others were sampled as many as ten times. As a result, the way we augment the data needs to account for the varying sampling attempts.

Recall,  $Y_{idt} = 1$  denotes that source  $i_d$  was detected on drive-by  $t$ . The set  $\mathcal{D}$  denotes the set of sampling attempt values. Note that  $\mathcal{D}$  can be any discrete subset of  $\mathbb{Z}^+$  and the values in  $\mathcal{D}$  do not need to be consecutive. Likewise, the notation  $i_d$  denotes the  $i^{th}$  methane

source detected in the region sampled  $d$  times. From our data, we obtain detection histories for  $n$  observed methane sources. For example, if we had  $\mathcal{D} = \{2, 3, 4, 5\}$ , we can write

$$n = n_2 + n_3 + n_4 + n_5,$$

where  $n_d$  is the number of detected (observed) individuals in areas sampled  $d$  times. Because detection of methane sources is imperfect, some individuals (source sites) are never observed. One of our goals is to estimate the total number of methane sources,  $N$ , for a closed population of sources. Our model accounts for the fact that the number of undetected sources will decrease with the number of sampling attempts (drives).

Because we do not know the number of undetected sources, we augment the data with  $K_n = K - n$  all-zero encounter histories (i.e.,  $\mathbf{Y}_{i_d} = \mathbf{0}^\top, i_d \in \{n_d + 1, n_d + 2, \dots, K_{n_d}\}$ ) and introduce the partially observed latent variable  $Z_{i_d}$  that indicates population membership. Due to the variable number of sampling attempts, when we augment the data with all-zero encounters, we also generate the number of sampling attempts for the augmented individuals. As a result, our total  $K_n$  augmented sources (i.e., all-zero encounter histories) can be classified based on sampling attempt. Continuing the example where  $\mathcal{D} = \{2, 3, 4, 5\}$ ,

$$K_n = K_{n_2} + K_{n_3} + K_{n_4} + K_{n_5},$$

is the total number of augmented sources and  $K_{n_d}$  denotes the number of augmented individuals in the region sampled  $d$  times. Likewise,

$$K = K_n + n = \sum_{d \in \mathcal{D}} (K_{n_d} + n_d) = K_2 + K_3 + K_4 + K_5$$

is the super population size, and  $K_d = K_{n_d} + n_d$  is the super population size in the region sampled  $d$  times. Because we expect that the number of unobserved individuals will decrease as the number of sampling attempts increase, we decrease the number of augmented indi-

viduals as the number of sampling attempts increases. For example, we add more missing individuals to areas that have been sampled twice than to areas that have been sampled four times ( $K_{n_2} > K_{n_4}$ ).

Adding all-zero encounters amounts to a zero-inflation of the detection data. We add a large number of all-zero encounters, and to estimate  $N$ , we estimate what proportion of these all-zero encounters are members of our population. Let  $Z_{i_d}$  be a partially observed, latent variable indicating population membership:  $Z_{i_d} = 1$  if a member of the super-population is in the population of sources that are available for sampling and  $Z_{i_d} = 0$  otherwise. We observe  $Z_{i_d} = 1$  for  $i_d \in \{1, 2, \dots, n_d\}, d \in \mathcal{D}$ . We model the methane source detections as

$$Y_{i_d t} = \begin{cases} 0 & \text{when } Z_{i_d} = 0, \\ \begin{cases} 0 & \text{w.p. } 1 - p_{i_d} \\ 1 & \text{w.p. } p_{i_d} \end{cases} & \text{when } Z_{i_d} = 1 \end{cases} \quad (53)$$

for  $t \in \{1, 2, \dots, d\}$ , where  $p_{i_d}$  is the probability of detecting source  $i_d$  and given in (52). The model statement in (53) corresponds to capture-recapture model  $M_h$  because the probability of detection varies by individual. We can rewrite (53) as a Binomial process,

$$m_{i_d} = \sum_{t=1}^d Y_{i_d t} \sim \text{Binom}(d, Z_{i_d} p_{i_d}).$$

We model the partially observed population membership variables  $Z_{i_d}$  as

$$Z_{i_d} \sim \text{Bernoulli}(\psi_d) \quad (54)$$

for  $d \in \mathcal{D}$  and  $i_d \in \{1, 2, \dots, K_d\}$ . The parameter  $1 - \psi_d$  represents the proportion of artificial zeroes (i.e., zero inflation) added to the region sampled  $d$  times due to parameter-expanded data augmentation. For capture-recapture models, parameter-expanded data augmentation changes the problem of estimating  $N$  to estimating  $\psi$  (Royle et al., 2007) and offers com-



Table 5.1: A summary of the calibration-capture-recapture (CCR) model. The calibration model is specified by the top two sections on the left. The capture-recapture model is described in the right side column. The distribution for the population of emission rates links the two models.

SLR Calibration Model	C-R Model with PX-DA
<u>Controlled Release Data:</u>	<u>Source detections:</u>
$\log(X_c) = \beta_0 + \beta_1 \log(v_c) + \epsilon_c$ $\epsilon_c \sim N(0, \sigma^2)$	$m_{i_d} = \sum_{t=1}^d Y_{i_d t} \sim \text{Binom}(d, Z_{i_d} p_{i_d}).$
<u>Excess <math>CH_4</math> when source is detected:</u>	<u>Detection probability function:</u>
$\log(X_{i_d j}) = \beta_0 + \beta_1 \log V_{i_d} + \epsilon_{i_d j}$ $\epsilon_{i_d j} \sim N(0, \sigma^2)$	$\text{logit}(p_{i_d}) = \alpha_0 + \alpha_1 V_{i_d}.$
<b>Model Link</b>	<u>Parameter-Expanded Data Augmentation:</u>
<u>Population of Emission Rates:</u>	$Z_{i_d} \sim \text{Bern}(\psi_d).$
$V_{i_d} \sim \text{logNormal}(\mu, \tau^2)$	

putational benefits (e.g., the dimension of the data does not change with changing  $N$ ). An estimate of the number of undetected methane sources in the region sampled  $d$  times is given by  $\hat{N}_d = \sum_{i_d=1}^{K_d} Z_{i_d}$  and the estimated total number of methane sources is given by  $\hat{N} = \sum_{d \in \mathcal{D}} \hat{N}_d$ .

To couple our calibration and capture-recapture model, we need to specify the distribution of the unobserved, latent source emission rates  $V_{i_d}$ . We assume that the source emission rates follow a logNormal distribution,

$$V_{i_d} \mid \mu, \sigma^2 \sim \text{logNormal}(\mu, \tau^2). \quad (55)$$

We choose a logNormal distribution for the population of source emission rates because a logNormal distribution has a heavy tail, and we expect that the majority of methane sources within a sampling region will have small emission rates while a handful of sources will have

very large emission rates. Note that under the logNormal assumption, the population mean emission rate is given by  $E(V_{i_d}) = \exp\{\mu + \tau^2/2\}$ . Because a source’s detection probability is a function of emission rate, our model will leverage both an observed source’s empirical detection probability and measured excess  $CH_4$  levels to infer its emission rate.

Table 5.1 summarizes the calibration-capture-recapture (CCR) model. To complete a Bayesian specification of our model, we choose prior distributions for our parameters of interest. To this end, we assume prior distributions for  $\beta_0, \beta_1, 1/\sigma^2, \alpha_0, \alpha_1, \mu, 1/\tau^2$ , and  $\psi_d, d \in \mathcal{D}$ . We derive the necessary posterior distributions and describe the MCMC sampling algorithm for parameter estimation in Section 5.4.

In addition to estimating the total number of methane sources in the sampled region, our model provides inference for the emission rate of sources detected by the GSV car’s instrumentation. Each source that is estimated to be a member of the population of sources, but went undetected by our sampling mechanism, will also have an associated estimated emission rate. Intuitively, sources that went undetected are more likely to be sources with a small emission rate rather than a large one. By estimating the emission rate of unobserved sources, we can estimate the total emissions within a sampling region while accounting for the undetected sources.

## 5.4 MCMC Algorithm

In this section we describe the posterior distributions and MCMC sampling algorithm used to obtain posterior draws from the calibration-capture-recapture model displayed in Table 5.1. We let  $\mathbf{X}_m$  denote the excess  $CH_4$  values observed when a source is detected, and  $s$  denotes the  $s^{th}$  iteration of the sampler. For a variable  $W$ ,  $lW$  denotes  $\log W$ . The joint posterior distribution of the unknown parameters and latent variables given the data can be

written as

$$\begin{aligned}
[\boldsymbol{\beta}, \sigma^2, \mathbf{V}, \mu, \tau^2, \mathbf{Z}, \boldsymbol{\psi}, \boldsymbol{\alpha} | \mathbf{X}_c, \mathbf{X}_m, \mathbf{Y}] &\propto [\mathbf{X}_c, \mathbf{X}_m, \mathbf{Y} | \boldsymbol{\beta}, \sigma^2, \mathbf{V}, \mathbf{Z}, \boldsymbol{\alpha}] \\
&[\boldsymbol{\beta}] [\sigma^2] [\mathbf{V} | \mu, \tau] [\mu] [\tau^2] [\boldsymbol{\alpha}] [\mathbf{Z} | \boldsymbol{\psi}] [\boldsymbol{\psi}] \\
&\propto [\mathbf{X}_c | \boldsymbol{\beta}, \sigma^2] [\mathbf{X}_m | \boldsymbol{\beta}, \sigma^2, \mathbf{V}, \mathbf{Y}] [\mathbf{Y} | \mathbf{V}, \mathbf{Z}, \boldsymbol{\alpha}] \\
&[\boldsymbol{\beta}] [\sigma^2] [\mathbf{V} | \mu, \tau] [\mu] [\tau^2] [\boldsymbol{\alpha}] [\mathbf{Z} | \boldsymbol{\psi}] [\boldsymbol{\psi}].
\end{aligned} \tag{56}$$

The likelihood corresponding to the controlled release (calibration) data (50) is

$$[\mathbf{X}_c | \boldsymbol{\beta}, \sigma^2] \propto \prod_{c=1}^C \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ \frac{-1}{2\sigma^2} (lX_c - \beta_0 - \beta_1 v_c)^2 \right\}. \tag{57}$$

The likelihood for the process of observing excess  $CH_4$  when a source is detected (51) is

$$[\mathbf{X}_m | \boldsymbol{\beta}, \sigma^2, \mathbf{V}, \mathbf{Y}] \propto \prod_{d \in \mathcal{D}} \prod_{i_d=1}^{K_d} \left( \prod_{j=1}^{\sum_{t=1}^d Y_{i_d t}} \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ \frac{-1}{2\sigma^2} (lX_{i_d j} - \beta_0 - \beta_1 lV_{i_d})^2 \right\} \right)^{I(\sum_{t=1}^d Y_{i_d t} > 0)}. \tag{58}$$

Finally, the likelihood corresponding to the detection (53) process is

$$[\mathbf{Y} | \mathbf{V}, \mathbf{Z}, \boldsymbol{\alpha}] \propto \prod_{d \in \mathcal{D}} \prod_{i_d=1}^{K_d} (Z_{i_d} p(V_{i_d}))^{\sum_{t=1}^d Y_{i_d t}} (1 - Z_{i_d} p(V_{i_d}))^{d - \sum_{t=1}^d Y_{i_d t}}. \tag{59}$$

We use Gibbs sampling to estimate (56) with Metropolis-Hastings steps when the conditional posterior distribution is not available in closed form. We outline the MCMC algorithm here:

1. Update the calibration parameters,  $\beta_0, \beta_1, \sigma^2$ , from their complete conditional distributions using a Gibbs step.
2. Update the detection parameters,  $\alpha_0$  and  $\alpha_1$ , using a Metropolis-Hastings step to sample from each complete conditional distribution.

3. Update the latent source emission rates,  $V_{i_d}$ . The complete conditional distribution for the source emission rate depends on whether or not the source was detected by the GSV car's instruments ( $m_{i_d} > 0$ ) and whether or not it is considered a member of the population ( $Z_{i_d} = 1$ ). These characteristics determine whether a Gibbs step or a Metropolis-Hasting step are required to sample from the complete conditional distribution.
4. Update the source emission rate population parameters,  $\mu$  and  $\tau^2$ , from their complete conditional distributions using Gibbs steps.
5. Update the unobserved, latent population membership indicators  $Z_{i_d}$  using their complete conditional distribution and Gibbs step.
6. Update the zero-inflation parameters,  $\psi_d$ , by sampling from their complete conditional distributions using Gibbs steps.

The details of the MCMC algorithm and conditional distributions are given below. Let  $\gamma^2 = 1/\sigma^2$  and  $\lambda^2 = 1/\tau^2$ . We let  $m_{i_d} = \sum_{t=1}^d Y_{i_d t}$  be the number of detections of source  $i_d$ . For the majority of the steps, we suppress the iteration counter,  $(s - 1)$ , that is used to denote parameter estimates from the previous iteration. The prior distributions are specified in Section 5.5.

**Step (1):** Draw  $\beta_0$ . Assuming a  $N(\mu_{\beta_0}, 1/\xi_{\beta_0}^2)$  prior distribution, the posterior is given by

$$\begin{aligned} [\beta_0|\cdot] &\propto \exp \left\{ \frac{-\gamma^2}{2} \sum_{c=1}^C (lX_c - \beta_0 - \beta_1 l v_c)^2 \right\} \\ &\exp \left\{ \frac{-\gamma^2}{2} \sum_{d \in \mathcal{D}} \sum_{i_d=1}^{n_d} \sum_{j=1}^{m_{i_d}} (lX_{i_d j} - \beta_0 - \beta_1 l V_{i_d})^2 \right\} \\ &\exp \left\{ \frac{-\xi_{\beta_0}^2}{2} (\beta_0^2 - 2\beta_0 \mu_{\beta_0}) \right\}. \end{aligned}$$

Combining the above, the posterior distribution of  $\beta_0$  is a normal distribution,

$$\beta_0|\cdot \sim N(\theta_{\beta_0}, \eta_{\beta_0}) \text{ where}$$

$$\theta_{\beta_0} = \frac{\gamma^2 \sum_{c=1}^C (lX_c - \beta_1 l v_c) + \gamma^2 \sum_{d \in \mathcal{D}} \sum_{i_d=1}^{n_d} \sum_{j=1}^{m_{i_d}} (lX_{i_d j} - \beta_1 l V_{i_d}) + \xi_{\beta_0}^2 \mu_{\beta_0}}{\gamma^2(C + M) + \xi_{\beta_0}^2}$$

and

$$\eta_{\beta_0} = \frac{1}{\gamma^2(C + M) + \xi_{\beta_0}^2},$$

where  $C$  is the number of replicates from the controlled release experiments and  $M = \sum_{d \in \mathcal{D}} \sum_{i_d=1}^{n_d} m_{i_d}$  is the total number of source detections.

**Step (2):** Draw  $\beta_1$ . Assuming a  $N(\mu_{\beta_1}, 1/\xi_{\beta_1}^2)$  prior distribution, the posterior distribution is

$$\begin{aligned} [\beta_1|\cdot] &\propto \exp \left\{ \frac{-\gamma^2}{2} \sum_{c=1}^C (lX_c - \beta_0 - \beta_1 l v_c)^2 \right\} \\ &\exp \left\{ \frac{-\gamma^2}{2} \sum_{d \in \mathcal{D}} \sum_{i_d=1}^{n_d} \sum_{j=1}^{m_{i_d}} (lX_{i_d j} - \beta_0 - \beta_1 l V_{i_d})^2 \right\} \\ &\exp \left\{ \frac{-\xi_{\beta_1}^2}{2} (\beta_1^2 - 2\beta_1 \mu_{\beta_1}) \right\}. \end{aligned}$$

After some algebra, we can show that the posterior distribution of  $\beta_1$  is a normal distribution,

$$\beta_1|\cdot \sim N(\theta_{\beta_1}, \eta_{\beta_1}) \text{ where}$$

$$\theta = \frac{\gamma^2 \sum_{c=1}^C l v_c (lX_c - \beta_0) + \gamma^2 \sum_{d \in \mathcal{D}} \sum_{i_d=1}^{n_d} l V_{i_d} \sum_{j=1}^{m_{i_d}} (lX_{i_d j} - \beta_0) + \xi_{\beta_1}^2 \mu_{\beta_1}}{\gamma^2 \sum_{c=1}^C l v_c^2 + \gamma^2 \sum_{d \in \mathcal{D}} \sum_{i_d=1}^{n_d} m_{i_d} l V_{i_d}^2 + \xi_{\beta_1}^2}$$

and

$$\eta_{\beta_1} = \frac{1}{\gamma^2 \sum_{c=1}^C l v_c^2 + \gamma^2 \sum_{d \in \mathcal{D}} \sum_{i_d=1}^{n_d} m_{i_d} l V_{i_d}^2 + \xi_{\beta_1}^2}.$$

**Step (3):** Sample  $\gamma^2 = 1/\sigma^2$ . Assuming a  $\text{Gamma}(a_\gamma, b_\gamma)$  prior distribution, the posterior is given by

$$\begin{aligned} [\gamma^2|\cdot] &\propto (\gamma^2)^{\frac{C+M}{2}} \exp \left\{ \frac{-\gamma^2}{2} \left( \sum_{c=1}^C (lX_c - \beta_0 - \beta_1 l v_c)^2 + \sum_{d \in \mathcal{D}} \sum_{i_d=1}^{n_d} \sum_{j=1}^{m_{i_d}} (lX_{i_d j} - \beta_0 - \beta_1 l V_{i_d})^2 \right) \right\} \\ &(\gamma^2)^{a_\gamma-1} \exp \{ -\gamma^2 b_\gamma \}. \end{aligned}$$

The posterior distribution of  $\gamma^2$  is a Gamma distribution:

$$\gamma^2 | \cdot \sim \text{Gamma}(\theta_\gamma, \eta_\gamma)$$

where

$$\theta_\gamma = a_\gamma + \frac{C + M}{2},$$

and

$$\eta_\gamma = b_\gamma + \frac{1}{2} \left( \sum_{c=1}^C (lX_c - \beta_0 - \beta_1 l v_c)^2 + \sum_{d \in \mathcal{D}} \sum_{i_d=1}^{n_d} \sum_{j=1}^{m_{i_d}} (lX_{i_d j} - \beta_0 - \beta_1 l V_{i_d})^2 \right).$$

**Step (4):** Draw  $\alpha_0$  and  $\alpha_1$ , conditional on the  $\mathbf{V}$  and  $Z_{i_d} = 1$ . We assume a normal prior distribution for each parameter. The conditional posterior distributions are a product of Binomial probabilities multiplied by the prior. We sample the detection probability parameters using a random walk Metropolis step, where we propose values from the Normal

prior distributions. Recall that  $m_{i_d} = \sum_{t=1}^d Y_{i_d t}$  is the number of detections of source  $i_d$ .

$$\begin{aligned}
[\alpha_0, \alpha_1 | \cdot] &\propto (\xi_{\alpha_0}^2)^{\frac{1}{2}} \exp \left\{ \frac{-\xi_{\alpha_0}^2}{2} (\alpha_0 - \mu_{\alpha_0})^2 \right\} (\xi_{\alpha_1}^2)^{\frac{1}{2}} \exp \left\{ \frac{-\xi_{\alpha_1}^2}{2} (\alpha_1 - \mu_{\alpha_1})^2 \right\} \\
&\prod_{d \in \mathcal{D}} \prod_{i_d=1}^{K_d} \left( Z_{i_d} \frac{\exp(\alpha_0 + \alpha_1 V_{i_d})}{1 + \exp(\alpha_0 + \alpha_1 V_{i_d})} \right)^{\sum_{t=1}^d Y_{i_d t}} \left( 1 - Z_{i_d} \frac{\exp(\alpha_0 + \alpha_1 V_{i_d})}{1 + \exp(\alpha_0 + \alpha_1 V_{i_d})} \right)^{d - \sum_{t=1}^d Y_{i_d t}} \\
&= (\xi_{\alpha_0}^2)^{\frac{1}{2}} \exp \left\{ \frac{-\xi_{\alpha_0}^2}{2} (\alpha_0 - \mu_{\alpha_0})^2 \right\} (\xi_{\alpha_1}^2)^{\frac{1}{2}} \exp \left\{ \frac{-\xi_{\alpha_1}^2}{2} (\alpha_1 - \mu_{\alpha_1})^2 \right\} \\
&\prod_{d \in \mathcal{D}} \prod_{\{i_d: Z_{i_d}=1\}} \left( \frac{\exp(\alpha_0 + \alpha_1 V_{i_d})}{1 + \exp(\alpha_0 + \alpha_1 V_{i_d})} \right)^{m_{i_d}} \left( 1 - \frac{\exp(\alpha_0 + \alpha_1 V_{i_d})}{1 + \exp(\alpha_0 + \alpha_1 V_{i_d})} \right)^{d - m_{i_d}} \\
&= (\xi_{\alpha_0}^2)^{\frac{1}{2}} \exp \left\{ \frac{-\xi_{\alpha_0}^2}{2} (\alpha_0 - \mu_{\alpha_0})^2 \right\} (\xi_{\alpha_1}^2)^{\frac{1}{2}} \exp \left\{ \frac{-\xi_{\alpha_1}^2}{2} (\alpha_1 - \mu_{\alpha_1})^2 \right\} \\
&\prod_{d \in \mathcal{D}} \prod_{\{i_d: Z_{i_d}=1\}} (p_{i_d})^{m_{i_d}} (1 - p_{i_d})^{d - m_{i_d}} \\
&= (\xi_{\alpha_0}^2)^{\frac{1}{2}} \exp \left\{ \frac{-\xi_{\alpha_0}^2}{2} (\alpha_0 - \mu_{\alpha_0})^2 \right\} (\xi_{\alpha_1}^2)^{\frac{1}{2}} \exp \left\{ \frac{-\xi_{\alpha_1}^2}{2} (\alpha_1 - \mu_{\alpha_1})^2 \right\} \\
&\exp \left\{ \log \left( \prod_{d \in \mathcal{D}} \prod_{\{i_d: Z_{i_d}=1\}} (p_{i_d})^{m_{i_d}} (1 - p_{i_d})^{d - m_{i_d}} \right) \right\} \\
&= (\xi_{\alpha_0}^2)^{\frac{1}{2}} \exp \left\{ \frac{-\xi_{\alpha_0}^2}{2} (\alpha_0 - \mu_{\alpha_0})^2 \right\} (\xi_{\alpha_1}^2)^{\frac{1}{2}} \exp \left\{ \frac{-\xi_{\alpha_1}^2}{2} (\alpha_1 - \mu_{\alpha_1})^2 \right\} \\
&\exp \left\{ \sum_{d \in \mathcal{D}} \sum_{\{i_d: Z_{i_d}=1\}} m_{i_d} \log p_{i_d} + (d - m_{i_d}) \log(1 - p_{i_d}) \right\}.
\end{aligned}$$

At iteration  $s$  of the sampler, we propose  $\alpha_0^*$  from a  $N(\alpha_0^{(s-1)}, 1/\xi_{\alpha_0}^2)$  distribution. The Metropolis ratio for sampling  $\alpha_0$  at iteration  $s$  is

$$R(\alpha_0^{(s-1)}, \alpha_0^*) = \exp \left\{ \frac{\sum_{d \in \mathcal{D}} \sum_{\{i_d: Z_{i_d}=1\}} m_{i_d} \log p_{i_d}^* + (d - m_{i_d}) \log(1 - p_{i_d}^*)}{\sum_{d \in \mathcal{D}} \sum_{\{i_d: Z_{i_d}=1\}} m_{i_d} \log p_{i_d}^{(s-1)} + (d - m_{i_d}) \log(1 - p_{i_d}^{(s-1)})} \right\},$$

where

$$p_{i_d}^* = \frac{\exp(\alpha_0^* + \alpha_1^{(s-1)} V_{i_d})}{1 + \exp(\alpha_0^* + \alpha_1^{(s-1)} V_{i_d})},$$

and  $p_{i_d}^{(s-1)}$  is defined analogously. The Metropolis ratio is similar for  $\alpha_1$ .



**Step (5):** Sample  $V_{i_d}$ . The posterior distribution depends on the index  $i_d$  and is given by

$$\begin{aligned}
[V_{i_d} | \cdot] &\propto \left( \prod_{j=1}^{m_{i_d}} \exp \left\{ \frac{-\gamma^2}{2} (lX_j - \beta_0 - \beta_1 lV_{i_d})^2 \right\} \right)^{I(m_{i_d} > 0)} \\
&\quad (Z_{i_d} p(V_{i_d}))^{\sum_{t=1}^d Y_{i_d t}} (1 - Z_{i_d} p(V_{i_d}))^{d - \sum_{t=1}^d Y_{i_d t}} \\
&\quad (\lambda^2)^{1/2} \exp \left\{ \frac{-\lambda^2}{2} (lV_{i_d} - \mu)^2 \right\}.
\end{aligned} \tag{60}$$

There are three cases to consider, which are based on the values of  $m_{i_d}$  and  $Z_{i_d}$ . We outline the posterior distribution of  $V_{i_d}$  for the three cases below.

Case (1): Source  $i_d$  is detected by the GSV car ( $m_{i_d} > 0 \implies Z_{i_d} = 1$ ). In this case we have  $m_{i_d}$  measurements of excess  $CH_4$ ,  $X_{i_d j}$ , that we can use to estimate the emission rate of source  $i_d$ . The posterior is given above in (60). In this case, we can combine the first and

third lines of (60) to obtain:

$$\begin{aligned}
[lV_{i_d}|\cdot] &\propto \exp \left\{ \frac{-\gamma^2}{2} \left( m_{i_d}\beta_1^2 lV_{i_d}^2 - 2lV_{i_d}\beta_1 \sum_{j=1}^{m_{i_d}} (lX_{i_dj} - \beta_0) \right) \right\} \\
&\quad \exp \left\{ \frac{-\lambda^2}{2} (lV_{i_d}^2 - 2\mu lV_{i_d}) \right\} \\
&\propto \exp \left\{ \left( \frac{-\gamma^2}{2} m_{i_d}\beta_1^2 + \frac{-\lambda^2}{2} \right) lV_{i_d}^2 + \left( \gamma^2\beta_1 \sum_{j=1}^{m_{i_d}} (lX_{i_dj} - \beta_0) + \lambda^2\mu \right) lV_{i_d} \right\}
\end{aligned}$$

This implies the following conditional posterior,

$$\log V_{i_d} \sim N(\theta_V, \eta_V)$$

where

$$\begin{aligned}
\theta_V &= \frac{\gamma^2\beta_1 \sum_{j=1}^{m_{i_d}} (lX_j - \beta_0) + \lambda^2\mu}{\gamma^2 m_{i_d}\beta_1^2 + \lambda^2}, \text{ and} \\
\eta_V &= \frac{1}{\gamma^2 m_{i_d}\beta_1^2 + \lambda^2}.
\end{aligned}$$

We can think of this conditional posterior as the posterior distribution of  $V_{i_d}$  given calibration data and the observed excess  $CH_4$  values. We can now rewrite the full posterior from Case

1 as follows:

$$\begin{aligned} [V_{i_d}|\cdot] &\propto \frac{1}{V_{i_d}\sqrt{2\pi\eta_V}} \exp\left\{\frac{-1}{2\eta_V}(\log V_{i_d} - \theta_V)^2\right\} (p(V_{i_d}))^{m_{i_d}} (1 - p(V_{i_d}))^{d-m_{i_d}} \\ &\equiv f(V_{i_d}). \end{aligned}$$

We propose a new value  $V_{i_d}^*$  by drawing from a  $\text{logNormal}(\log V_{i_d}^{(s-1)}, \sigma^{2(s-1)})$  distribution.

This implies the following Metropolis ratio:

$$\begin{aligned} R(V_{i_d}^{(s-1)}, V_{i_d}^*) &= \frac{f(V_{i_d}^*) \frac{1}{V_{i_d}^{(s-1)}\sqrt{2\pi\sigma^2}} \exp\left\{\frac{-1}{2\sigma^2}(\log V_{i_d}^{(s-1)} - \log V_{i_d}^*)^2\right\}}{f(V_{i_d}^{(s-1)}) \frac{1}{V_{i_d}^*\sqrt{2\pi\sigma^2}} \exp\left\{\frac{-1}{2\sigma^2}(\log V_{i_d}^* - \log V_{i_d}^{(s-1)})^2\right\}} \\ &= \frac{\exp\left\{\frac{-1}{2\eta_V}(\log V_{i_d}^* - \theta_V)^2\right\} (p(V_{i_d}^*))^{m_{i_d}} (1 - p(V_{i_d}^*))^{d-m_{i_d}}}{\exp\left\{\frac{-1}{2\eta_V}(\log V_{i_d}^{(s-1)} - \theta_V)^2\right\} (p(V_{i_d}^{(s-1)}))^{m_{i_d}} (1 - p(V_{i_d}^{(s-1)}))^{d-m_{i_d}}} \end{aligned}$$

We can rewrite the ratio quantity as

$$R(V_{i_d}^{(s-1)}, V_{i_d}^*) = R = \exp\{\log R\}.$$

Suppressing the subscripts and letting  $p^* = p(V_{i_d}^*)$ ,  $p^{(s-1)} = p(V_{i_d}^{(s-1)})$ , and  $m = m_{i_d}$ , we can write

$$\begin{aligned} \log R &= \frac{-1}{2\eta_V}(lV^* - \theta_v)^2 + m \log p^* + (d - m) \log(1 - p^*) \\ &\quad - \left[ \frac{-1}{2\eta_V}(lV^{(s-1)} - \theta_v)^2 + m \log p^{(s-1)} + (d - m) \log(1 - p^{(s-1)}) \right] \end{aligned}$$

Case (2): Source  $i_d$  is an augmented source that is estimated to be a member of the population ( $Z_{i_d} = 1$ ) due to imperfect detection but was never observed by the GSV car

$(m_{i_d} = 0)$ . The posterior distribution (60) becomes

$$[V_{i_d}|\cdot] \propto (1 - p(V_{i_d}))^d \frac{(\lambda^2)^{1/2}}{V_{i_d}} \exp \left\{ \frac{-\lambda^2}{2} (lV_{i_d} - \mu)^2 \right\}.$$

In this case we propose  $V_{i_d}^*$  by drawing from the  $\text{logNormal}(\mu^{(s-1)}, \tau^{2(s-1)})$  prior distribution.

The Metropolis ratio becomes

$$R(V_{i_d}^{(s-1)}, V_{i_d}^*) = \frac{(1 - p^*)^d}{(1 - p^{(s-1)})^d}.$$

Case (3): Source  $i_d$  is an augmented source that is not estimated to be a population member ( $Z_{i_d} = 0$  and  $m_{i_d} = 0$ ). The posterior distribution (60) simplifies to

$$[V_{i_d}|\cdot] \propto (\lambda^2)^{1/2} \exp \left\{ \frac{-\lambda^2}{2} (lV_{i_d} - \mu)^2 \right\}.$$

In this case we propose and automatically accept values from the  $\text{logNormal}(\mu^{(s-1)}, \tau^{2(s-1)})$  prior distribution.

**Step (6):** Update  $\mu$ . We assume a  $\text{Normal}(\mu_0, 1/\xi_\mu^2)$  prior distribution giving the following posterior:

$$[\mu|\cdot] \propto \left( \prod_{d \in \mathcal{D}} \prod_{i_d=1}^{K_d} \exp \left\{ \frac{-\lambda^2}{2} (lV_{i_d} - \mu)^2 \right\} \right) \exp \left\{ \frac{-\xi_\mu^2}{2} (\mu - \mu_0)^2 \right\}.$$

Some algebra yields a Normal posterior distribution,  $\mu|\cdot \sim \text{N}(\theta_\mu, \eta_\mu)$ , where

$$\theta_\mu = \frac{\lambda^2 \sum_{d \in \mathcal{D}} \sum_{i_d=1}^{K_d} lV_{i_d} + \xi_\mu^2 \mu_0}{\lambda^2 K + \xi_\mu^2} \text{ and } \eta_\mu = \frac{1}{\lambda^2 K + \xi_\mu^2}.$$

**Step (7):** Update  $\lambda = 1/\tau^2$ . We assume a  $\text{Gamma}(a_\lambda, b_\lambda)$  prior. The posterior distribution is

$$[\lambda^2|\cdot] \propto \prod_{d \in \mathcal{D}} \prod_{i_d=1}^{K_d} (\lambda^2)^{1/2} \exp \left\{ \frac{-\lambda^2}{2} (lV_{i_d} - \mu)^2 \right\} \\ (\lambda^2)^{a_\lambda-1} \exp\{-\lambda^2 b_\lambda\}.$$

Thus the posterior is a Gamma distribution,  $\lambda^2|\cdot \sim \text{Gamma}(\theta_\lambda, \eta_\lambda)$  where

$$\theta_\lambda = a_\lambda + \frac{K}{2} \text{ and } \eta_\lambda = b_\lambda + \frac{1}{2} \sum_{d \in \mathcal{D}} \sum_{i_d=1}^{K_d} (lV_{i_d} - \mu)^2.$$

**Step (8):** Sample  $Z_{i_d}$  for  $d \in \mathcal{D}$ ,  $i_d = n_d + 1, n_d + 2, \dots, K_d$  (i.e., when  $m_{i_d} = \sum_{t=1}^d Y_{i_d t} = 0$ ). The posterior distribution is a Bernoulli:

$$Z_{i_d}|\cdot \sim \text{Bern}(\phi_{i_d})$$

where

$$\phi_{i_d} = \frac{\psi_d (1 - p_{i_d})^d}{\psi_d (1 - p_{i_d})^d + (1 - \psi_d)}.$$

Recall we can compute the estimated population size using the  $Z_{i_d}$  values,  $\hat{N} = \sum_{d \in \mathcal{D}} \sum_{i_d=1}^{K_d} Z_{i_d}$ .

We can also compute the estimated total number of sources in the region sampled  $d$  times

Table 5.2: Prior distributions and parameter values for our CCR model.

Parameter	Prior	Parameter Values
$\beta_0$	$N(\mu_{\beta_0}, 1/\xi_{\beta_0}^2)$	$\mu_{\beta_0} = 0, 1/\xi_{\beta_0}^2 = 10$
$\beta_1$	$N(\mu_{\beta_1}, 1/\xi_{\beta_1}^2)$	$\mu_{\beta_1} = 0, 1/\xi_{\beta_1}^2 = 10$
$\gamma^2 = 1/\sigma^2$	$\text{Gamma}(a_\gamma, b_\gamma)$	$a_\gamma = 0.01, b_\gamma = 0.01$
$\alpha_0$	$N(\mu_{\alpha_0}, 1/\xi_{\alpha_0}^2)$	$\mu_{\alpha_0} = 0, 1/\xi_{\alpha_0}^2 = 0.05$
$\alpha_1$	$N(\mu_{\alpha_1}, 1/\xi_{\alpha_1}^2)$	$\mu_{\alpha_1} = 0, 1/\xi_{\alpha_1}^2 = 0.03$
$\mu$	$N(\mu_0, 1/\xi_\mu)$	$\mu_0 = 0, 1/\xi_\mu = 5$
$\lambda^2 = 1/\tau^2$	$\text{Gamma}(a_\gamma, b_\gamma)$	$a_\gamma = 0.01, b_\gamma = 0.01$
$\psi_d, d \in \mathcal{D}$	$\text{Beta}(a_{\psi_d}, b_{\psi_d})$	$a_{\psi_d} = 0.0000001, b_{\psi_d} = 0.01 \forall d \in \mathcal{D}$

( $\hat{N}_d = \sum_{i_d=1}^{K_d} Z_{i_d}$ ) and number of undetected sources in the region sampled  $d$  times,  $\hat{N}_{n_d} = \hat{N}_d - n_d$ .

**Step (9):** Update  $\psi_d$  for  $d \in \mathcal{D}$ . Assuming a  $\text{Beta}(a_\psi, b_\psi)$  prior, the conditional posterior is a Beta distribution,

$$\psi_d | \cdot \sim \text{Beta}(\theta_\psi, \eta_\psi)$$

where  $\theta_\psi = a_\psi + N_d$ ,  $\eta_\psi = b_\psi + K_d - N_d$ , and  $N_d = \sum_{i_d=1}^{K_d} Z_{i_d}$ .

## 5.5 Prior Distributions and Data Augmentation

In this section we specify the prior distributions and their associated parameter values used in our analysis of two U.S. cities described in Section 5.6. We also discuss starting values and the number of augmented sources. The MCMC sampling algorithm is described in Section 5.4.

Table 5.2 displays the prior distributions and prior distribution parameter values for our CCR model. We chose normal priors for all detection coefficients and for the mean of the log emission rates,  $\mu$ . For  $\beta_0, \beta_1$ , and  $\mu$  we chose a large variance to make the prior non-informative. For the detection coefficients  $\alpha_0$  and  $\alpha_1$ , we choose the prior variance based on the acceptance probability of proposed values during random walk Metropolis sampling. These variances are relatively small because the detection function and estimated emission

Table 5.3: The number of augmented sources (i.e., all-zero detection histories) for each city, classified by sampling attempt. Because the number of undetected sources decreases with sampling effort, we decrease the number of augmented sources as sampling effort increases.

City	$K_{n_2}$	$K_{n_3}$	$K_{n_4}$	$K_{n_5}$	$K_{n_6}$	$K_{n_7}$	$K_{n_8}$	$K_{n_9}$	$K_{n_{10}}$	Total
A	85	60	40	20	20	20	20	20	20	305
B	40	30	25	15	15	10	10	10	10	165

rate values must agree with the observed detections. For example, if  $p_{i_d} = 1$  then the number of detections  $m_{i_d}$  must equal the number of sampling attempts  $d$ . We found that smaller variances produce better mixing than a large variance, which also lead to exceptionally low acceptance rates. We use non-informative gamma priors for the precision (inverse variance) parameters. The prior parameters for  $\psi_d$  are chosen to approximate a scale prior on  $N_d$  (Link, 2013).

For each city, we also had to choose the number of augmented sources. The number of augmented sources needs to be large enough to have little effect on the posterior distribution (i.e., chosen so that  $N \ll K$ ). However, adding too many augmented sources increases computation time. Because we expect the number augmented sources to decrease with sampling effort, we decrease the number of augmented sources as sampling effort increases. We add more augmented sources in City A than City B because empirical detection rates were lower in City A relative to City B. The empirical detection rate is based on the number of observed source sites in each city and defined as the total number of detections divided by the total number of sampling attempts from the observed sites. Table 5.3 displays the counts of augmented sources by sampling attempt for each of the two cities.

We describe the starting values for our MCMC sampling algorithm. We computed starting values for  $\beta_0, \beta_1$ , and  $\sigma^2$  using the controlled release data. Namely, we computed starting values for  $\beta_0$  and  $\beta_1$  using least squares. The mean square error from the least squares fit is the starting value for  $\sigma^2$ . For each of the  $n$  detected sources, the starting values for the emission rate  $V_{i_d}$  were computed by averaging the source's log excess  $CH_4$  ( $X_{i_dj}$ ) values,

Table 5.4: The number of observed methane sources  $n$  by sampling attempt in two U.S. cities. We let  $n_d$  denote the number of sources detected in the region sampled  $d$  times.

Drives	2	3	4	5	6	7	8	9	10	Total, $n$
City A	22	20	23	6	18	5	9	10	6	119
City B	21	32	25	23	20	23	11	7	8	170

substituting the mean into the least squares regression equation, and solving for  $V_{i_d}$ . Using the set of initial emission rate ( $V_{i_d}$ ) values, we calculated the starting values for  $\mu$  and  $\tau^2$  using the mean and variance of the  $\log(V_{i_d})$  values, respectively. We imputed initial emission rates for the augmented sources by drawing from a logNormal distribution with the initial  $\mu$  and  $\tau^2$  parameters. We used random starting values for  $\alpha_0, \alpha_1$ , the unobserved (augmented)  $Z_{i_d}$ , and for  $\psi_d, d \in \mathcal{D}$ . We generated starting values for  $\alpha_0$  and  $\alpha_1$  by drawing from a  $N(0, 1)$  distribution. For each augmented source, we set  $Z_{i_d} = 1$  with probability 0.15. The starting values for  $\psi_d$  were drawn from a  $\text{Beta}(5, 5)$  distribution.

We used standard diagnostics to assess convergence of our MCMC algorithm. Specifically, we checked ACF and trace plots and computed Gelman-Rubin statistics. After discarding burn-in and thinning the chains, these diagnostics all indicated adequate convergence.

## 5.6 Application: A Tale of Two Cities

In this section we apply our calibration-capture-recapture model to methane data collected from two U.S. cities. Because of our non-disclosure agreement, we call these two cities A and B. City A has an accelerated natural gas pipeline replacement program. City B does not have an accelerated pipeline replacement program and has older infrastructure relative to City A. Methane concentrations were collected within designated sampling regions in each city. In City A GSV vehicles sampled approximately 887 miles of roadway two or more times. In City B approximately 1251 miles of roadway were sampled two or more times. In both cities the majority of sampled roadway was driven between 2 and 7 times. A very small



Table 5.5: Posterior estimates from our CCR model. “Post. Med.” denotes the posterior median and “95% CI” denotes the 95% posterior credible interval based on the highest posterior density.

	City A		City B	
Parameter	Post. Med.	95% CI	Post. Med.	95% CI
Total Number of Sources, $N$	161	(138, 187)	176	(172, 182)
Total Emissions (L/min), $\sum V_i$	123	(51, 205)	693	(465, 927)
Detection Parameter, $\alpha_0$	-1.64	(-2.47, -0.99)	-0.05	(-0.39, 0.26)
Detection Parameter, $\alpha_1$	1.28	(0.44, 2.70)	0.23	(0.11, 0.39)
Avg Emission Rate, $e^{\mu+\tau^2/2}$	0.76	(0.31, 1.28)	3.91	(2.53, 5.62)
sd(Log Emission Rate), $\tau$	0.76	(0.45, 1.10)	1.07	(0.81, 1.38)

amount of roadway in each city was sampled more than 10 times, and sources are not likely to go undetected after 10 sampling attempts. For these reasons we restrict our attention to areas sampled between 2 and 10 times, inclusive, so that  $\mathcal{D} = \{2, 3, 4, 5, 6, 7, 8, 9, 10\}$ .

Within the sampled regions,  $n = 119$  and 175 sources were detected in Cities A and B, respectively (Table 5.4). Using the observed excess  $CH_4$  values and a least squares fit of the controlled release data, the average source emission rate in City A is an estimated 2.15 L/min while in City B it is an estimated 6.07 L/min. Empirical source detection was much higher in City B than City A. In City A there were 734 drive-bys of the 119 sources, and the sources were detected on 282 of these sampling attempts. In City B there were 911 drive-bys of the 170 sources which resulted in 576 detections. Based on these numbers, the empirical detection probabilities, ignoring source emission rate, are 0.38 and 0.63 for Cities A and B, respectively. We did not estimate a common detection function (i.e., common  $\alpha_0$  and  $\alpha_1$ ) between the two cities because we believe the relationship between emission rate and detection varies by city due to local conditions (e.g., location of pipeline infrastructure, urban geography).

We estimated our Bayesian calibration-capture-recapture model via MCMC. See Sections 5.4 and 5.5 for more details on prior distributions, diagnostics, and the MCMC algorithm. We ran the chains for 1,000,000 iterations using R software (R Core Team, 2015), discarded

Table 5.6: Posterior estimates of the calibration parameters from our CCR model. “Post. Med.” denotes the posterior median and “95% CI” denotes the 95% posterior credible interval based on the highest posterior density.

	City A		City B	
Parameter	Post. Med.	95% CI	Post. Med.	95% CI
$\beta_0$ , Calibration Parameter	-1.18	(-1.67, -0.72)	-1.20	(-1.65, -0.74)
$\beta_1$ , Calibration Parameter	0.81	(0.63, 0.98)	0.82	(0.66, 0.99)
$\sigma$ SD, Calibration Parameter	1.28	(1.18, 1.40)	1.21	(1.13, 1.29)

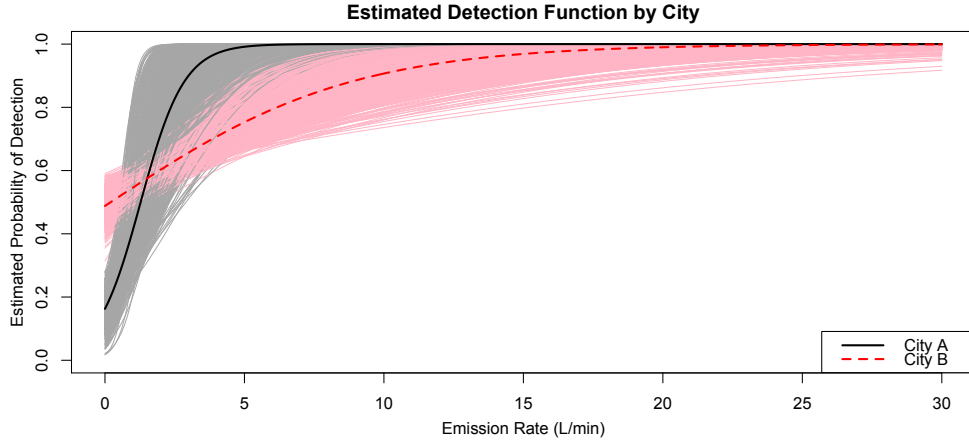


Figure 5.3: The estimated detection functions from two U.S. cities. The black and red lines are the estimated detection functions based on the posterior medians of  $\alpha_0$  and  $\alpha_1$  for Cities A and B, respectively. As expected, the estimated probability of detection increases with emission rate. The gray and pink lines represent the uncertainty in the detection function based on posterior draws of  $\alpha_0$  and  $\alpha_1$ .

the first 10,000 samples for burn-in, and thinned the chain by 25, resulting in 39,600 samples from the posterior distributions.

The posterior estimates and associated uncertainty of the calibration parameters  $\beta_0$ ,  $\beta_1$ , and  $\sigma^2$  were similar between the two cities. This is because these estimates are largely driven by the calibration data which are the same for both cities. It is reassuring that these estimates are similar even though they were estimated using separate models (Table 5.6). Based on the 95% highest posterior density intervals of the detection coefficient  $\alpha_1$ , we can conclude that there is a significant positive relationship between source emission rate and

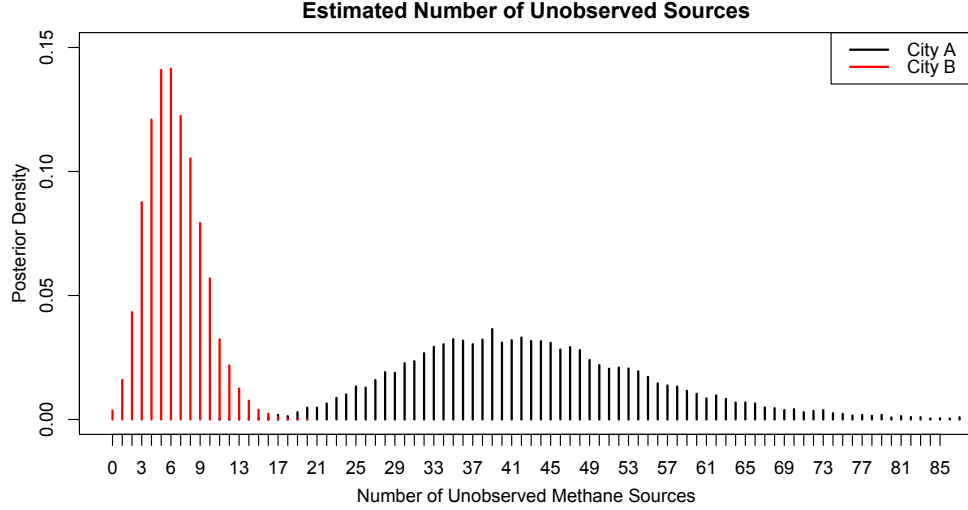


Figure 5.4: The posterior distributions for the estimated number of undetected sources within the sampled regions of Cities A and B. In City A an estimated 42 sources went undetected while only 6 went undetected in City B. Despite having an estimated fewer number of total sources, City B has a significantly larger estimate of total emissions due to the presence of sources with large emission rates.

probability of detection in both U.S. cities (Table 5.5). Figure 5.3 displays the posterior estimates of the detection function and associated uncertainty. In City A the detection of small sources is lower than in City B, but the probability of detection increases more quickly as a function of source emission rate than in City B. As a result of the low detection of small sources in City A, most of the undetected sources are estimated as small sources.

Figure 5.4 displays the posterior distribution for the number of unobserved sources in the two cities. In City A, an estimated 42 sources went undetected by the GSV car while there are just 6 estimated undetected sources in City B. The difference in the estimated number of sources reflects the difference in the empirical detection rate between the two cities. Because detection is a function of emission rate, this difference also reflects the difference in the distributions of source emission rates for the two cities. Recall that roadways within the designated sampling area were not sampled equally, and during data augmentation, we include an assignment of sampling effort to augmented sources. As a result, we can break down the undetected sources based on sampling effort. Figure 5.5 displays the posterior

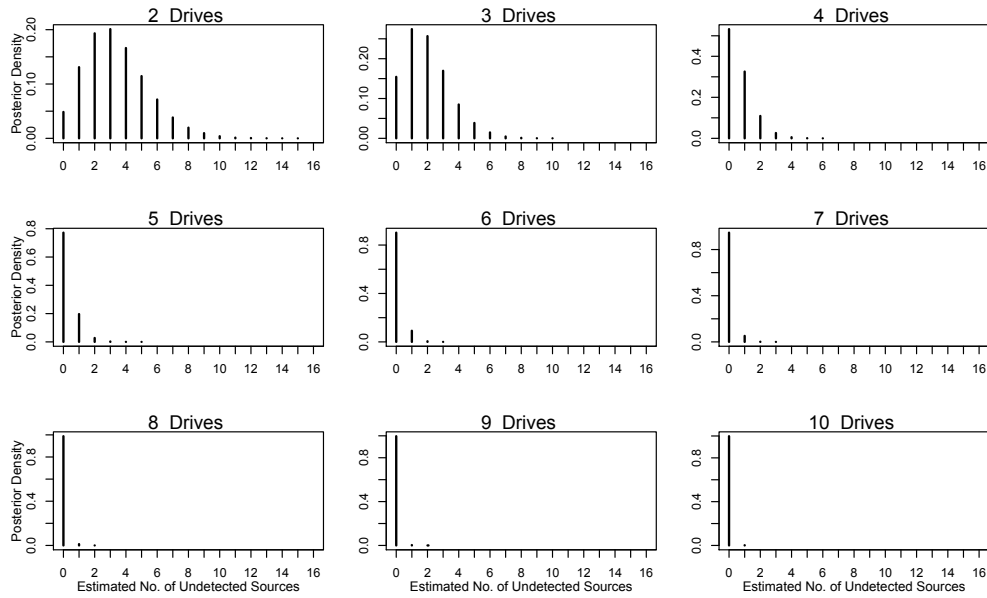


Figure 5.5: The estimated number of undetected sources in City B, classified by sampling attempt. As expected, the estimated number of undetected sources decreases as sampling attempts increases. We see a particularly sharp change in the number of undetected sources between 5 and 6 sampling attempts.

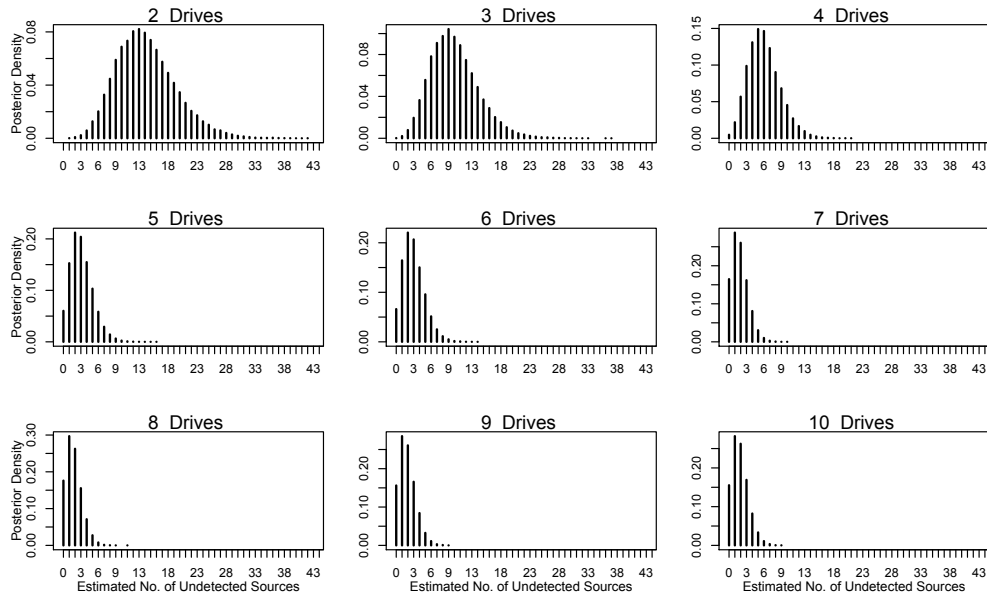


Figure 5.6: The estimated number of undetected sources in City A, classified by sampling attempt. As expected, the estimated number of undetected sources decreases as sampling attempts increases.

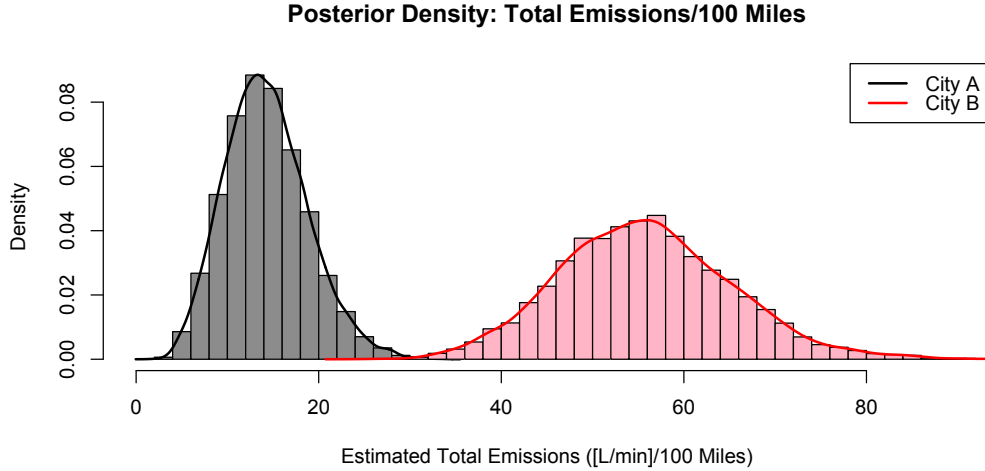


Figure 5.7: Total emissions per 100 miles of roadway sampled. The posterior distribution for the total emissions (L/min) per 100 miles of roadway from the sampled regions in two U.S. cities. These emission totals account for both detected and undetected methane sources.

distribution of undetected sources in City B by sampling attempt. As expected, the number of undetected sources decreases with sampling effort. There is a particularly large decrease in the number of undetected sources when the number of sampling attempts is greater than 6 in City B. The same plot for City A is included in Figure 5.6.

The rate of detected sources is similar between the two cities: 13.4 sources/100 miles in City A and 13.6 sources/100 miles in City B. Based on the model results, which account for undetected sources, these rates are an estimated 18.2 (15.6, 21.1) sources/100 miles in City A and 14.2 (13.7, 14.5) sources/100 miles in City B. The estimated difference in these rates is primarily because City A had a larger number of sources with a small emission rate, which tend to have a small empirical detection probability, than City B. As a result of the lower detection probability, the model indicates that there are more undetected, but small, sources in City A. This is evidenced by the estimate of the total emissions per 100 miles in each city. Despite the cities having a similar number of estimated total sources, the estimated total emission rate in City B, 55.4 (37.2, 74.1) L/min/100 miles is approximately 4 times greater than the estimated total emission rate in City A, 13.9 (5.8, 23.1) L/min. This is

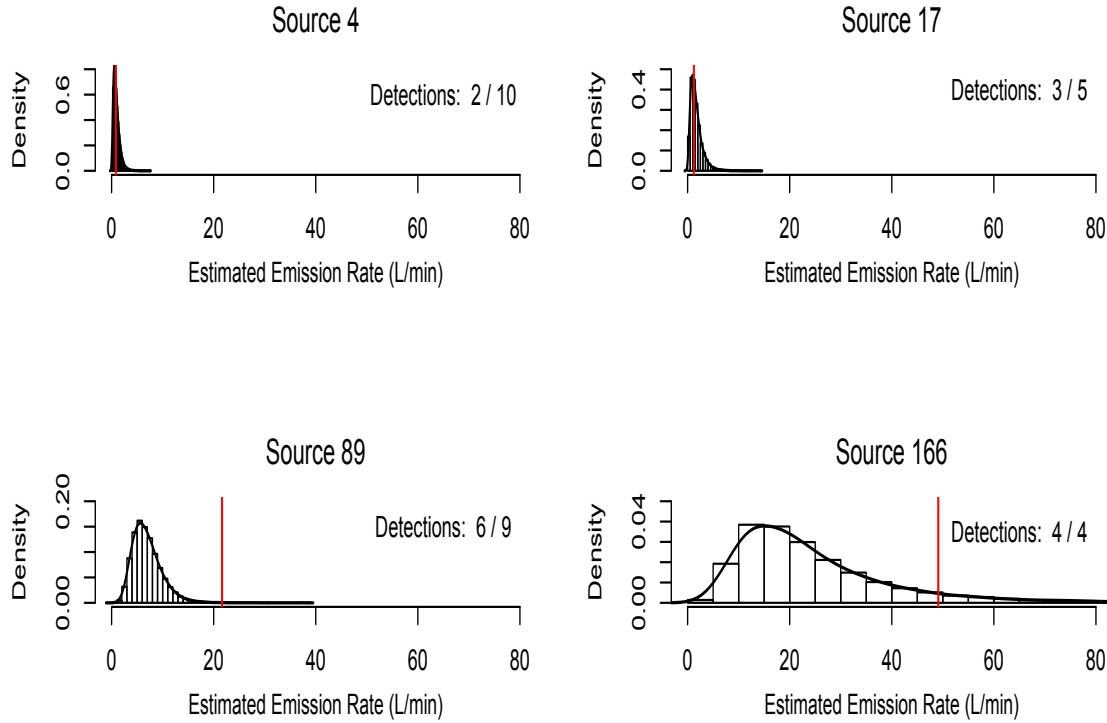


Figure 5.8: Posterior distributions of individual emission rates. The posterior distributions for four detected sources in City B. The distributions reflect uncertainty in individual emission rates based on the detection probability and the observed excess  $CH_4$  values for that source. The vertical red lines are the naive estimates of the emission rate using a least squares fit of the controlled release data and the observed excess  $CH_4$ .

due to the presence of a larger number of sources with a large emission rate in City B. The posterior distributions of total emissions per 100 miles sampled in the two cities is displayed in Figure 5.7.

Our model output can also be used to prioritize methane source investigation and repair and to understand the impacts of those repairs. Figure 5.8 displays the posterior distribution of emission rate for 4 sources detected in City B. Using the posterior distributions of the emission rate for our  $n$  detected sources, we can readily answer the question, “What is the probability that this source’s emission rate is greater than  $r$  L/min?”. After computing this

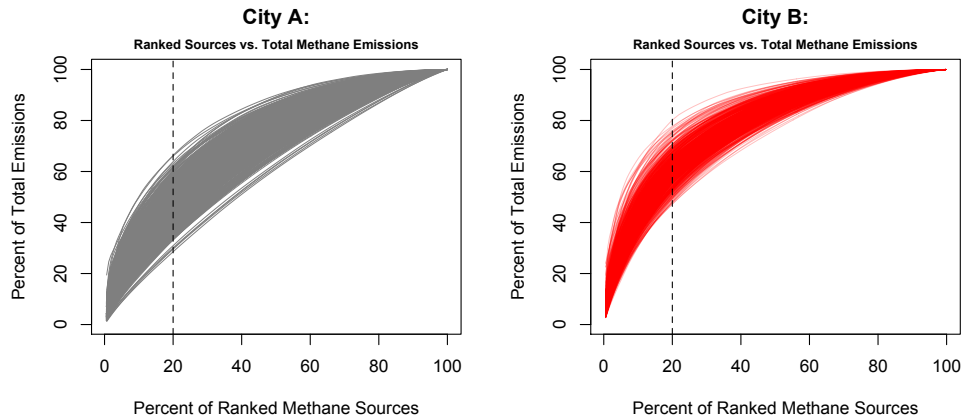


Figure 5.9: Percentage of total methane emissions versus percentage of ranked methane sources. These plots display the percentage of total methane emissions as a function the percentage of ranked methane sources and can be used to quantify the impact of removing the largest methane sources. The sources are ranked by estimated methane emission rate. For example, in City A, removing the largest 20% of sources would remove an estimated 47% (35.5%, 59.5%) of the total methane emissions. In City B, removing the largest 20% of sources would remove an estimated 60% (50.8%, 72.8%) of the total methane emissions.

value for each of the sources, we can compile a ranked list of the sources' potential to be large. This list would be informative for individual source investigation/repair. Aggregating estimated source emissions spatially may also be informative. For example, we could estimate total emissions over several contiguous blocks and compare this estimate to the estimates over other aggregated blocks to prioritize regional repairs. We can also quantify the impact of repairing or remedying the largest methane sources. The plots in Figure 5.9 can be used to estimate the impact of removing the largest methane sources in the sampling region for each city. These plots display the percent of estimated total emissions versus the percent of ranked methane sources. For example, in City B it is estimated that removing the largest 20% of sources would remove approximately 47% (35.5%, 59.5%) of the total emissions. In City A removing the largest 20% of sources results in removing an estimated 60% (50.8%, 72.8%) of total emissions.

## 5.7 Discussion

With the development of mobile sensing technology to identify methane sources, such as natural gas leaks, there is an increasing need for methods that translate the raw  $CH_4$  concentration data into meaningful quantities. By coupling a calibration and capture-recapture model, we developed a calibration-capture-recapture (CCR) model to estimate the total number of methane sources within a designated sampling region. Our model estimates population characteristics by leveraging data from controlled release experiments and data collected by sampling in urban areas and is therefore an integrated population model (Schaub and Abadi, 2011). In addition to estimating the total number of sources, we are also able to estimate individual emission rates and the total methane emissions with our CCR model. By coupling the calibration and capture-recapture model, we are able to appropriately account for uncertainty in the detection process and individual emission rate estimation process. In the spirit of Royle (2009) and Royle and Dorazio (2012), we formulated our model as a Bayesian hierarchical model and used parameter-expanded data augmentation to enable estimation.

Our CCR model has features similar to other population models. Ruiz-Gutierrez et al. (2016) also develop an integrated population model that uses calibration to improve inference. They use calibration to improve estimation of false-positive detection errors within an occupancy model (MacKenzie et al., 2002). In our model we use calibration to improve estimates of individual methane source emission rates within a capture-recapture model.

We use PX-DA to estimate our CCR model. Based on the work of J.D. Nichols (Nichols and Karanth, 2002), Royle and Dorazio (2012) provide a heuristic explanation of the duality and connection between occupancy models and closed population capture-recapture models developed using PX-DA. Using this duality, we can think of our model in the framework of occupancy models. Like an occupancy model, we can think of our  $K$  sources (observed plus augmented) as representing sites or patches that are sampled to assess the presence or absence of a methane emitting source. In fact, the detections of a source over different drive-



bys are aggregated to create source sites. The augmented sites, having all-zero encounters, represent both sites where a methane source is present, but was never detected, and sites where there is no methane source. Within this framework we can make inferences about the parameter of interest  $\psi$ , the probability that a site contains a methane source.

However, our model is different from traditional occupancy models and constitutes a capture-recapture model because our sites represent individual members of the population rather than a location which may hold several individuals. In an occupancy model, site abundance typically measures the number of animals at the site. In our model we can think of site abundance as the methane emission rate, a characteristic of the individual source at that site. Royle and Nichols (2003) develop an occupancy model linking unobserved site abundance with probability of detecting occupancy at the site. Similarly, our CCR model links unobserved abundance (source emission rate) with the probability of detection. However, the models differ in how abundance is modeled. Royle and Nichols (2003) use a discrete distribution to model abundance (counts) while in our CCR model the distribution for abundance (emission rates) is continuous. Additionally, during sampling we record information related to abundance ( $\max CH_4$ ), and we leverage controlled release experiments to estimate the abundance.

Unlike capture-recapture models with a fixed number of sampling attempts, our model accounts for the preferential sampling that arises due to the logistics of driving a sampling vehicle in an urban area. A variable number of sampling attempts may naturally arise in other capture-recapture settings. For example, ease of physical accessibility or weather may lead to a variable number of sampling attempts in a spatial capture-recapture setting. For camera trapping surveys, camera failure can lead to a variable number of sampling attempts at a given site. Our model successfully incorporates the information from the varying sampled regions to estimate a common detection function and provides insight into the effect of sampling effort on the detection of individuals. For detection of methane sources, our results

indicates that five or more sampling attempts greatly reduces the uncertainty in the number of unobserved sources.

## CONCLUSIONS AND FUTURE WORK

### 6.1 Future Work: Nonparametric Tests of Spatial Isotropy

There are several avenues of future work with nonparametric tests of spatial isotropy. Methods to more formally characterize the optimal block size and bandwidth parameters for the tests in the spatial domain would enhance the applicability of the tests. The performance of the tests for non-gridded data in Guan et al. (2004) and Maity and Sherman (2012) are sensitive to these choices and their optimality remains an open question. Zhang et al. (2014) developed a nonparametric method for estimating the asymptotic variance-covariance matrix of statistics derived from spatial data that avoids choosing tuning parameters which could simplify test implementation. Another area of future investigation is the effect of changing the form of the covariance function and its parameters on test size and power. Our simulations in Chapter 2 used exponential covariance functions (i.e., a Matérn covariance function with smoothness equal to 0.5). Results may differ for other covariance functions (e.g., spherical covariance function or Matérn covariance function with smoothness other than 0.5). A third area of further investigation is to compare nonparametric methods to parametric methods for testing isotropy (e.g., Scaccia and Martin (2011)). A final area of future work is development of a formal definition and more careful quantification of power of tests of spatial isotropy. For example, the degree of geometric anisotropy could be quantified using different characteristics of the covariance function, including the ratio of the major and minor axes of the ellipse, degree of rotation of the ellipse relative to one of the coordinate axes, and range of the process. Furthermore, it is important to consider the effects of density

and design of sampling locations, sample size, and the amount of noise (nugget and sill) in the observations on a test’s ability to detect anisotropy.

Improvements to our software package `spTest` is another area of future work. First, our new nonparametric test of spatial isotropy in Chapter 4 could readily be incorporated into the package. Additionally, as new hypothesis tests and graphical techniques are developed, adding them to the `spTest` package would provide a convenient suite of diagnostic software for practitioners. Next, it would be useful to extend the functionality of `spTest` to handle non-rectangular sampling domains. Finally, the computational efficiency of `spTest` functions can be improved by programming them in C++.

There are several areas of future work related to our new hypothesis test in Chapter 4, including extending our test to handle missing data and non-gridded sampling locations (e.g., Fuentes, 2007). The asymptotic results derived in Bandyopadhyay and Lahiri (2009) may be useful for this extension. Additionally, improvements to the variance approximations in Section 4.4.2 might amend the concerns of inflated test size when data are observed on square grids. Finally, the plots in Figure 4.4 suggest that when  $\theta = 0$ , the power of the test might be improved by considering the maximum difference between periodogram values at frequencies with the same norm. For example, the test statistic  $M_{D_k} = \max \{D_k\}$  may be informative. Developing a distribution for this test statistic will present a new set of challenges relative to those presented in Section 4.5.3. Finally, the use of data tapering or pre-whitening may also improve test performance, as these methods reduce the bias in the periodogram values.

Our investigation into the aliasing concern in Chapter 4 suggests that we should replace  $f(\boldsymbol{\omega})$  with  $f_{\Delta}(\boldsymbol{\omega})$  in the Whittle approximation (49). An issue with this approach is that using  $f_{\Delta}(\boldsymbol{\omega})$  greatly increases the computation time, as the aliased spectral density (25) needs to be computed at each step of the optimization. Additionally, the double infinite sum in (25) needs to be truncated. Therefore, an area of future research is developing “efficient and accurate approximations to  $f_{\Delta}(\boldsymbol{\omega})$ ” (Stein, 2016). Additionally, we only explored the effect of aliasing on point estimation using the Whittle approximation. An area of future

research would be to explore the effect of aliasing on the uncertainty (i.e., standard error) estimates given by the Whittle approximation. Finally, like other graphical techniques, our graphical assessment of aliasing may be subjective and open to interpretation. Thus, a final area of further work would be developing a formal hypothesis test for aliasing.

## 6.2 Future Work: Calibration-Capture-Recapture

There are many areas of future work in analyzing methane concentration data from mobile detection platforms. Our capture-recapture model does not discriminate between natural gas and non-natural gas methane sources. While we attempted to remove non-natural gas sources during data processing, discriminating types of methane sources is important for identifying areas of potential infrastructure repair. For example, the locations of natural gas pipelines, wetlands, and storm drains could be combined with the estimated source sites to provide a probabilistic classification of the source type.

Factors other than source emission rate, such as wind, traffic-generated air movement, and changing source expression, will also affect the probability of detecting a source. For example, if wind is blowing emissions away from the roadway, the source is less likely to be detected by the GSV car’s instrumentation. Likewise, rain saturated soil can act as a barrier to methane gas’ movement to the atmosphere. In our CCR model, the probability of detection is a function only of source emission rate. As a result, non-detection of a source is attributed only to source emission rate and is evidence that the source’s emission rate is small. For this reason, we suspect that our individual emission rates are conservative. Thus, an area of future work is fully understanding the effects of these variables on source detection and incorporating them into a model for source detection.

We did not include these other influential covariates here because they are difficult to quantify at the small, localized scales at which methane sources occur. Air movement due to wind, traffic, and urban geography poses a particularly onerous challenge. We plan to perform more controlled methane releases and incorporate more covariates in the future to

gain a better understanding of the source generating, detection, and emission processes. For example, age, number of pipeline connections, pipeline gas pressure, material, and soil types are known to affect the prevalence of natural gas leaks in pipeline infrastructure. Wind, soil moisture content, underground gas migration, distance from the GSV vehicle, and surrounding traffic all influence the probability of detecting elevated methane levels. Given that elevated readings occur, these same variables also influence the levels of  $CH_4$  measured by the GSV instrumentation.

Other capture-recapture models might also be useful for analyzing data collected by GSV vehicles. Over time, known natural gas leaks are repaired and new gas leaks arise, so it would be natural to extend our closed-population capture-recapture model to an open population of methane sources. An  $M_{ht}$  capture-recapture model might also be useful for analyzing the data. Because wind speed and direction can change from sampling attempt to sampling attempt, the detection probability of an individual source is likely to change with sampling attempt. In addition to the challenge of quantifying winds at small scales, we do not obtain an estimate of a source’s emission rate when it is not detected which poses another challenge in estimating model  $M_{ht}$ .

Finally, incorporating the spatial nature of the methane concentration data is another area of tangential and future work. Maps play a key role in informing discussion on infrastructure repair. Our capture-recapture models provide estimates for the number of undetected sources, classified by sampling effort, but they do not provide inference for the specific locations of those undetected sources. We may consider methane sources as arising due to a non-homogeneous point process, and inference on the underlying source generating process may be a surrogate for the quality of infrastructure in a given region. A modified spatial capture-recapture model may also be appropriate for these data and could be used to infer the location (i.e., “home site”) of methane sources. In this work, we did not use a spatially explicit capture-recapture model because we do not know the exact locations of methane sources. Additionally, an assumption that methane sources occupy a two-dimensional region

may not be appropriate for our analysis. The locations of natural gas leaks, the methane source of highest importance and interest for our analysis, tend to be restricted to locations along roadways where pipeline infrastructure is buried.

### **6.3 Conclusions**

In this dissertation we presented applied, theoretical, and methodological advances in the statistical analysis of spatially-referenced and capture-recapture data. Our review of nonparametric tests of isotropy provides insight and guidelines for using the tests in practice. The software package we developed enables practitioners to easily implement the various methods. Our new nonparametric tests of isotropy provides a new tool for testing spatial isotropy and overcomes shortfalls of the other methods. We also discuss many avenues of future work related to nonparametric tests of spatial isotropy. Finally, our new capture-recapture model provides a framework to draw inference from data collected by new mobile sensing technology. The methodology for our capture-recapture model is useful in other capture-recapture settings, and the results of our model are important for informing climate change and infrastructure discussions.

## REFERENCES

- Adler, R. J. (1981). *The Geometry of Random Fields*, volume 62. Siam.
- Amstrup, S. C., McDonald, T. L., and Manly, B. F. (2010). *Handbook of Capture-Recapture Analysis*. Princeton University Press.
- Baczkowski, A. (1990). A test of spatial isotropy. In *Compstat*, pages 277–282. Springer.
- Baczkowski, A. and Mardia, K. (1987). Approximate lognormality of the sample semi-variogram under a Gaussian process. *Communications in Statistics-Simulation and Computation*, 16(2):571–585.
- Baczkowski, A. and Mardia, K. (1990). A test of spatial symmetry with general application. *Communications in Statistics-Theory and Methods*, 19(2):555–572.
- Baddeley, A. J. and Turner, R. (2005). **spatstat**: An R package for analyzing spatial point patterns. *Journal of Statistical Software*, 12.
- Bandyopadhyay, S. and Lahiri, S. (2009). Asymptotic properties of discrete fourier transforms for spatial data. *Sankhyā: The Indian Journal of Statistics, Series A (2008-)*, pages 221–259.
- Bandyopadhyay, S., Lahiri, S. N., Nordman, D. J., et al. (2015). A frequency domain empirical likelihood method for irregularly spaced spatial data. *The Annals of Statistics*, 43(2):519–545.
- Bandyopadhyay, S. and Rao, S. S. (2015). A test for stationarity for irregularly spaced spatial data. *Journal of the Royal Statistical Society (B)*.
- Banerjee, S., Carlin, B. P., and Gelfand, A. E. (2014). *Hierarchical Modeling and Analysis for Spatial Data*. CRC Press.
- Bivand, R., Keitt, T., and Rowlingson, B. (2016). *rgdal: Bindings for the Geospatial Data Abstraction Library*. R package version 1.1-9.
- Borgman, L. and Chao, L. (1994). Estimation of a multidimensional covariance function in case of anisotropy. *Mathematical Geology*, 26(2):161–179.
- Boucher, O., Friedlingstein, P., Collins, B., and Shine, K. P. (2009). The indirect global warming potential and global temperature change potential due to methane oxidation. *Environmental Research Letters*, 4(4):044007.
- Bowman, A. W. and Azzalini, A. (2014). *R Package sm: Nonparametric Smoothing Methods (version 2.2-5.4)*. University of Glasgow, UK and Università di Padova, Italia.



- Bowman, A. W. and Crujeiras, R. M. (2013). Inference for variograms. *Computational Statistics & Data Analysis*, 66:19–31.
- Brillinger, D. R. (1970). The frequency analysis of relations between stationary spatial series. In *Proc. Twelfth Biennial Seminar Canadian Math. Congress (ed. R. Pyke). Canadian Math. Congress, Montreal*, pages 39–81.
- Brillinger, D. R. (1974). Fourier analysis of stationary processes. *Proceedings of the IEEE*, 62(12):1628–1643.
- Brillinger, D. R. (1981). *Time Series: Data Analysis and Theory*, volume 36. Siam.
- Brockwell, P. J. and Davis, R. A. (2009). *Introduction to Time Series and Forecasting*. Springer.
- Cabana, E. (1987). Affine processes: a test of isotropy based on level sets. *SIAM Journal on Applied Mathematics*, 47(4):886–891.
- Chorti, A. and Hristopulos, D. T. (2008). Nonparametric identification of anisotropic (elliptic) correlations in spatially distributed data sets. *Signal Processing, IEEE Transactions on*, 56(10):4738–4751.
- Corstanje, R., Grunwald, S., and Lark, R. (2008). Inferences from fluctuations in the local variogram about the assumption of stationarity in the variance. *Geoderma*, 143(1):123–132.
- Coull, B. A. and Agresti, A. (1999). The use of mixed logit models to reflect heterogeneity in capture-recapture studies. *Biometrics*, 55(1):294–301.
- Cressie, N. (1993). *Statistics for Spatial Data: Wiley Series in Probability and Statistics*. Wiley-Interscience New York.
- Crujeiras, R. M. (2006). *Contributions to spectral spatial statistics*. PhD thesis, University Santiago de Compostela. unpublished PhD thesis.
- Crujeiras, R. M. and Fernández-Casal, R. (2010). On the estimation of the spectral density for continuous spatial processes. *Statistics*, 44(6):587–600.
- Crujeiras, R. M., Fernández-Casal, R., and González-Manteiga, W. (2010). Goodness-of-fit tests for the spatial spectral density. *Stochastic Environmental Research and Risk Assessment*, 24(1):67–79.
- Dahlhaus, R. and Künsch, H. (1987). Edge effects and efficient parameter estimation for stationary random fields. *Biometrika*, 74(4):877–882.
- Ecker, M. D. and Gelfand, A. E. (1999). Bayesian modeling and inference for geometrically anisotropic spatial data. *Mathematical Geology*, 31(1):67–83.
- Ecker, M. D. and Gelfand, A. E. (2003). Spatial modeling and prediction under stationary non-geometric range anisotropy. *Environmental and Ecological Statistics*, 10(2):165–178.

- Fuentes, M. (2001). A high frequency kriging approach for non-stationary environmental processes. *Environmetrics*, 12(5):469–483.
- Fuentes, M. (2002). Spectral methods for nonstationary spatial processes. *Biometrika*, 89(1):197–210.
- Fuentes, M. (2005). A formal test for nonstationarity of spatial stochastic processes. *Journal of Multivariate Analysis*, 96(1):30–54.
- Fuentes, M. (2006). Testing for separability of spatial–temporal covariance functions. *Journal of Statistical Planning and Inference*, 136(2):447–466.
- Fuentes, M. (2007). Approximate likelihood for large irregularly spaced spatial data. *Journal of the American Statistical Association*, 102(477):321–331.
- Fuentes, M. (2013). Spectral methods. *Wiley StatsRef: Statistics Reference Online*.
- Fuentes, M. and Reich, B. (2010). Spectral domain. *Handbook of Spatial Statistics*, pages 57–77.
- Fuentes, M. and Smith, R. L. (2001). A new class of nonstationary spatial models. Technical report, Technical report, North Carolina State University, Raleigh, NC.
- García-Soidán, P. (2007). Asymptotic normality of the Nadaraya–Watson semivariogram estimators. *Test*, 16(3):479–503.
- García-Soidán, P. H., Febrero-Bande, M., and González-Manteiga, W. (2004). Nonparametric kernel estimation of an isotropic variogram. *Journal of Statistical Planning and Inference*, 121(1):65–92.
- Gelfand, A. E. and Schliep, E. M. (2016). Spatial statistics and Gaussian processes: A beautiful marriage. *Spatial Statistics*.
- Gneiting, T., Genton, M., and Guttorp, P. (2007). Geostatistical space-time models, stationarity, separability and full symmetry. *Statistical Methods for Spatio-Temporal Systems*, pages 151–175.
- Guan, Y., Sherman, M., and Calvin, J. A. (2004). A nonparametric test for spatial isotropy using subsampling. *Journal of the American Statistical Association*, 99(467):810–821.
- Guan, Y., Sherman, M., and Calvin, J. A. (2006). Assessing isotropy for spatial point processes. *Biometrics*, 62(1):119–125.
- Guan, Y., Sherman, M., and Calvin, J. A. (2007). On asymptotic properties of the mark variogram estimator of a marked point process. *Journal of Statistical Planning and Inference*, 137(1):148–161.
- Guan, Y. T. (2003). *Nonparametric methods of assessing spatial isotropy*. PhD thesis, Texas A&M University.

- Guyon, X. (1982). Parameter estimation for a stationary process on a d-dimensional lattice. *Biometrika*, 69(1):95–105.
- Hall, P. and Patil, P. (1994). Properties of nonparametric estimators of autocovariance for stationary random fields. *Probability Theory and Related Fields*, 99(3):399–424.
- Haskard, K. A. (2007). *An anisotropic Matérn spatial covariance model: REML estimation and properties*. PhD thesis, University of Adelaide.
- Irvine, K. M., Gitelman, A. I., and Hoeting, J. A. (2007). Spatial designs and properties of spatial correlation: effects on covariance estimation. *Journal of Agricultural, Biological, and Environmental Statistics*, 12(4):450–469.
- Isaaks, E. H. and Srivastava, R. M. (1989). *Applied Geostatistics*, volume 2. Oxford University Press New York.
- Jona-Lasinio, G. (2001). Modeling and exploring multivariate spatial variation: A test procedure for isotropy of multivariate spatial data. *Journal of Multivariate Analysis*, 77(2):295–317.
- Journel, A. G. and Huijbregts, C. J. (1978). *Mining geostatistics*. Academic Press.
- Jun, M. and Genton, M. G. (2012). A test for stationarity of spatio-temporal random fields on planar and spherical domains. *Statistica Sinica*, 22(4):1737.
- Kim, H.-J. and Fuentes, M. (2000). Spectral analysis with spatial periodogram and data tapers. In *Proceedings Joint Statistical Meeting*. Citeseer.
- Kim, T. Y. and Park, J.-S. (2012). On nonparametric variogram estimation. *Journal of the Korean Statistical Society*, 41(3):399–413.
- Lahiri, S. and Zhu, J. (2006). Resampling methods for spatial regression models under a class of stochastic designs. *The Annals of Statistics*, 34(4):1774–1813.
- Lahiri, S. N. (2003). *Resampling Methods for Dependent Data*. Springer Science & Business Media.
- Li, B., Genton, M. G., and Sherman, M. (2007). A nonparametric assessment of properties of space–time covariance functions. *Journal of the American Statistical Association*, 102(478):736–744.
- Li, B., Genton, M. G., and Sherman, M. (2008a). Testing the covariance structure of multivariate random fields. *Biometrika*, 95(4):813–829.
- Li, B., Genton, M. G., Sherman, M., et al. (2008b). On the asymptotic joint distribution of sample space–time covariance estimators. *Bernoulli*, 14(1):228–248.
- Lim, C. Y. and Stein, M. (2008). Properties of spatial cross-periodograms using fixed-domain asymptotics. *Journal of Multivariate Analysis*, 99(9):1962–1984.

- Link, W. A. (2013). A cautionary note on the discrete uniform prior for the binomial n. *Ecology*, 94(10):2173–2179.
- Liu, J. S. and Wu, Y. N. (1999). Parameter expansion for data augmentation. *Journal of the American Statistical Association*, 94(448):1264–1274.
- Lu, H. and Zimmerman, D. (2001). Testing for isotropy and other directional symmetry properties of spatial correlation. *preprint*.
- Lu, H.-C. (1994). *On the Distributions of the Sample Covariogram and Semivariogram and Their Use in Testing for Isotropy*. PhD thesis, University of Iowa.
- Lu, N. and Zimmerman, D. L. (2005). Testing for directional symmetry in spatial dependence using the periodogram. *Journal of Statistical Planning and Inference*, 129(1):369–385.
- MacKenzie, D. I., Nichols, J. D., Lachman, G. B., Droege, S., Andrew Royle, J., and Langtimm, C. A. (2002). Estimating site occupancy rates when detection probabilities are less than one. *Ecology*, 83(8):2248–2255.
- Maity, A. and Sherman, M. (2012). Testing for spatial isotropy under general designs. *Journal of Statistical Planning and Inference*, 142(5):1081–1091.
- Matheron, G. (1961). Precision of exploring a stratified formation by boreholes with rigid spacing-application to a bauxite deposit. In *International Symposium of Mining Research, University of Missouri*, volume 1, pages 407–22.
- Matheron, G. (1962). *Traité de Géostatistique Appliquée*. Editions Technip.
- Matsuda, Y. and Yajima, Y. (2009). Fourier analysis of irregularly spaced data on rd. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 71(1):191–217.
- Mearns, L. O., Gutowski, W., Jones, R., Leung, R., McGinnis, S., Nunes, A., and Qian, Y. (2009). A regional climate change assessment program for north america. *Eos, Transactions American Geophysical Union*, 90(36):311–311.
- Modjeska, J. S. and Rawlings, J. (1983). Spatial correlation analysis of uniformity data. *Biometrics*, pages 373–384.
- Molina, A. and Feito, F. R. (2002). A method for testing anisotropy and quantifying its direction in digital images. *Computers & Graphics*, 26(5):771–784.
- Nadaraya, E. A. (1964). On estimating regression. *Theory of Probability & Its Applications*, 9(1):141–142.
- Nichols, J. D. (1992). Capture-recapture models. *BioScience*, 42(2):94–102.
- Nichols, J. D. and Karanth, K. U. (2002). Statistical concepts: assessing spatial distributions. In *Monitoring tigers and their prey: A manual for wildlife researchers, managers, and conservationists in tropical Asia*, pages 29–38. Centre for Wildlife Studies.

- Nicolis, O., Mateu, J., and D'Ércole, R. (2010). Testing for anisotropy in spatial point processes. In *Proceedings of the Fifth International Workshop on Spatio-Temporal Modelling*, pages 1990–2010.
- Nychka, D., Furrer, R., Paige, J., and Sain, S. (2014). *fields: Tools for Spatial Data*. R package version 8.3-5.
- Osborne, C. (1991). Statistical calibration: a review. *International Statistical Review/Revue Internationale de Statistique*, pages 309–336.
- Otis, D. L., Burnham, K. P., White, G. C., and Anderson, D. R. (1978). Statistical inference from capture data on closed animal populations. *Wildlife Monographs*, (62):3–135.
- Pagano, M. (1971). Some asymptotic properties of a two-dimensional periodogram. *Journal of Applied Probability*, 8(4):841–847.
- Park, M. S. and Fuentes, M. (2008). Testing lack of symmetry in spatial–temporal processes. *Journal of Statistical Planning and Inference*, 138(10):2847–2866.
- Politis, D. N. and Sherman, M. (2001). Moment estimation for statistics from marked point processes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(2):261–275.
- Pollock, K. H. (2002). The use of auxiliary variables in capture-recapture modelling: an overview. *Journal of Applied Statistics*, 29(1-4):85–102.
- Porcu, E., Crujeiras, R., Mateu, J., and Gonzalez-Manteiga, W. (2009). On the second order properties of the multidimensional periodogram for regularly spaced data. *Theory of Probability & Its Applications*, 53(2):349–356.
- Possolo, A. (1991). Subsampling a random field. *Lecture Notes-Monograph Series*, pages 286–294.
- Priestley, M. and Rao, T. S. (1969). A test for non-stationarity of time-series. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 140–149.
- Priestley, M. B. (1981). *Spectral Analysis and Time Series*. Academic Press.
- R Core Team (2015). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Ribeiro Jr., P. and Diggle, P. (2001). *geoR: A package for geostatistical analysis*. *R-NEWS*, 1(2):15–18.
- Rosenblatt, M. (1985). *Stationary Sequences and Random Fields*. Birkhäuser Boston, Inc.
- Royle, J. A. (2009). Analysis of capture–recapture models with individual covariates using data augmentation. *Biometrics*, 65(1):267–274.

- Royle, J. A. and Dorazio, R. M. (2012). Parameter-expanded data augmentation for Bayesian analysis of capture–recapture models. *Journal of Ornithology*, 152(2):521–537.
- Royle, J. A., Dorazio, R. M., and Link, W. A. (2007). Analysis of multinomial models with unknown index using data augmentation. *Journal of Computational and Graphical Statistics*, 16(1):67–85.
- Royle, J. A. and Nichols, J. D. (2003). Estimating abundance from repeated presence–absence data or point counts. *Ecology*, 84(3):777–790.
- Ruiz-Gutierrez, V., Hooten, M. B., and Campbell Grant, E. H. (2016). Uncertainty in biological monitoring: a framework for data collection and analysis to account for multiple sources of sampling bias. *Methods in Ecology and Evolution*.
- Scaccia, L. and Martin, R. (2002). Testing for simplification in spatial models. In *Compstat*, pages 581–586. Springer.
- Scaccia, L. and Martin, R. (2005). Testing axial symmetry and separability of lattice processes. *Journal of Statistical Planning and Inference*, 131(1):19–39.
- Scaccia, L. and Martin, R. (2011). Model-based tests for simplification of lattice processes. *Journal of Statistical Computation and Simulation*, 81(1):89–107.
- Schabenberger, O. and Gotway, C. A. (2004). *Statistical Methods for Spatial Data Analysis*. CRC Press.
- Schaub, M. and Abadi, F. (2011). Integrated population models: a novel analysis framework for deeper insights into population dynamics. *Journal of Ornithology*, 152(1):227–237.
- Shao, X. and Li, B. (2009). A tuning parameter free test for properties of space–time covariance functions. *Journal of Statistical Planning and Inference*, 139(12):4031–4038.
- Sherman, M. (1996). Variance estimation for statistics computed from spatial lattice data. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 509–523.
- Sherman, M. (2011). *Spatial Statistics and Spatio-Temporal Data: Covariance Functions and Directional Properties*. John Wiley & Sons.
- Sherman, M. and Carlstein, E. (1994). Nonparametric estimation of the moments of a general statistic computed from spatial data. *Journal of the American Statistical Association*, 89(426):496–500.
- Spiliopoulos, I., Hristopulos, D. T., Petrakis, M., and Chorti, A. (2011). A multigrid method for the estimation of geometric anisotropy in environmental data from sensor networks. *Computers & Geosciences*, 37(3):320–330.
- Stein, M. (2016). Personal Communication.
- Stein, M. L. (1988). Asymptotically efficient prediction of a random field with a misspecified covariance function. *The Annals of Statistics*, 16(1):55–63.

- Stein, M. L. (1995). Fixed-domain asymptotics for spatial periodograms. *Journal of the American Statistical Association*, 90(432):1277–1288.
- Stein, M. L., Chi, Z., and Welty, L. J. (2004). Approximating likelihoods for large spatial data sets. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 66(2):275–296.
- Tanner, M. A. and Wong, W. H. (1987). The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association*, 82(398):528–540.
- Thon, K., Geilhufe, M., and Percival, D. B. (2015). A multiscale wavelet-based test for isotropy of random fields on a regular lattice. *Image Processing, IEEE Transactions on*, 24(2):694–708.
- Van Hala, M., Bandyopadhyay, S., Lahiri, S. N., and Nordman, D. J. (2014). A frequency domain empirical likelihood for estimation and testing of spatial covariance structure. *preprint*.
- Vecchia, A. V. (1988). Estimation and model identification for continuous spatial processes. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 297–312.
- Venables, W. N. and Ripley, B. D. (2002). *Modern Applied Statistics with S*. Springer, New York, fourth edition. ISBN 0-387-95457-0.
- von Fischer, J. C., Cooley, D., Gaylord, A., Griebenow, C. J., Hamburg, S. P., Salo, J., Schumacher, R., Theobald, D., and Ham, J. (2017). Rapid identification of location and magnitude of urban natural gas pipeline leaks. *to appear, Environmental Science and Technology*.
- Watson, G. S. (1964). Smooth regression analysis. *Sankhyā: The Indian Journal of Statistics, Series A*, pages 359–372.
- Weller, Z. D. (2016a). read-WRFG-data. <https://github.com/zdweller/Read-NARCCAP/blob/master/read-WRFG-data>.
- Weller, Z. D. (2016b). spTest-JSS-Code. <https://github.com/zdweller/spTest-JSS-Code/blob/master/spTest-JSS-Code-script>.
- Weller, Z. D. (2016c). **spTest**: an R package implementing nonparametric tests of isotropy. *to appear, Journal of Statistical Software*.
- Weller, Z. D. (2016d). **spTest**: *Nonparametric Hypothesis Tests of Isotropy and Symmetry*. R package version 0.2.4.
- Weller, Z. D. and Hoeting, J. A. (2016). A review of nonparametric hypothesis tests of isotropy properties in spatial data. *Statistical Science*, 31(3):305–324.
- Whittle, P. (1954). On stationary processes in the plane. *Biometrika*, pages 434–449.

- Yaglom, A. (1987). *Correlation Theory of Stationary and Related Random Functions*. Springer-Verlag.
- Zhang, H. (2004). Inconsistent estimation and asymptotically equal interpolations in model-based geostatistics. *Journal of the American Statistical Association*, 99(465):250–261.
- Zhang, X., Li, B., and Shao, X. (2014). Self-normalization for spatial data. *Scandinavian Journal of Statistics*, 41(2):311–324.
- Zimmerman, D. L. (1993). Another look at anisotropy in geostatistics. *Mathematical Geology*, 25(4):453–470.