

NONSTATIONARY GAUSSIAN RANDOM FIELDS WITH APPLICATION TO SPACE
AND SPACE–TIME MODELING

By

Abolfazl Safikhani

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Statistics – Doctor of Philosophy

2015

ABSTRACT

NONSTATIONARY GAUSSIAN RANDOM FIELDS WITH APPLICATION TO SPACE AND SPACE–TIME MODELING

By

Abolfazl Safikhani

Space-time models have become increasingly popular in scientific studies. Gaussian random fields (GRFs) are an important tool of this field. Developing covariance structures for GRFs, parameter estimation, and prediction in big spatial datasets are central challenges in spatio-temporal modeling. My thesis consists of three main chapters dealing with these main challenges in this area. In chapter 2, we investigate the problem of equivalence of GRFs with stationary increments, and obtain sufficient conditions for equivalence in terms of the behavior of the spectral measures at infinity. Further, the main results are applied to a rich family of nonstationary space-time models with possible anisotropy behavior. In chapter 3, we propose a nonstationary parametric model, in which the underlying Gaussian random field may have different regularities in different directions, thus can be applied to model anisotropy. Using the theory of equivalence of Gaussian measures under nonstationary assumption, strong consistency of the tapered likelihood based estimation of the variance component under fixed domain asymptotics are derived by putting mild conditions on the spectral behavior of the tapering covariance function. In chapter 4, we merge two ideas of factor modeling and low-rank spatial processes approximation to develop and propose a class of hierarchical low-rank spatial factor models which offer a rich and flexible modeling option for dealing with large vector of outcomes observed at large number of locations. A Markov chain Monte Carlo algorithm is developed for estimation, and further, the full posterior distributions are recovered in a Bayesian predictive framework.

Keywords: Space–time models, nonstationary Gaussian random fields, Equivalence of Gaussian measures, Covariance tapering, Fixed–domain asymptotics, Spatial factor modeling, Gaussian predictive processes, Hierarchical Bayesian

I dedicate this dissertation to my wonderful parents.

ACKNOWLEDGMENTS

I wish to express my deepest gratitude to my advisors Dr. Yimin Xiao and Dr. Alla Sikorskii. Their guidance, patience and thoughtfulness made this work possible. It was a privilege to work with them.

I would like to thank my thesis committee members Drs. Andrew Finley, V. S. Mandrekar, and Chae Young Lim, for their time and interest. Also, I appreciate the help and support of the Department of Statistics and Probability in last five years.

TABLE OF CONTENTS

LIST OF TABLES	vii
LIST OF FIGURES	viii
Chapter 1 Introduction	1
Chapter 2 Spectral Conditions for Equivalence of Gaussian Random Fields with stationary increments	4
2.1 Introduction	4
2.2 Preliminary	7
2.3 Main Results	13
2.4 Application	28
Chapter 3 Covariance Tapering for Anisotropic Nonstationary Gaussian Random Fields with Application to Large Scale Spatial Data Sets	35
3.1 Introduction	36
3.2 Preliminary	37
3.3 Main Result	38
3.4 Simulation Results	42
Chapter 4 Spatial factor models and dimension reducing Gaussian pre- dictive processes for high dimensional remotely sensed data and forest variables	44
4.1 Model	45
4.2 Identifiability Issue	46
4.3 Prior Specification	47
4.4 Approximation using Gaussian predictive processes	48
4.5 Data equation	49
4.6 Computations	50
4.6.1 Updating the mean part b	51
4.6.2 Updating ω^*	51
4.6.3 Updating the loading factors Λ	52
4.6.4 Updating the nuggets τ	53
4.7 Simulations	53
4.8 Real Data Analysis	54
BIBLIOGRAPHY	59

LIST OF TABLES

Table 3.1	Results for estimation of σ^2	43
Table 4.1	Parameter estimates with their 95% credible intervals	54
Table 4.2	Parameter estimates with their 95% credible intervals	58

LIST OF FIGURES

Figure 3.1	3D Simulated Surface on $[0, 1]^2$ with increments 0.02	39
Figure 3.2	2D Projection of the Simulated Surface on $[0, 1]^2$ with increments 0.02	40
Figure 3.3	Simulated Surface on $[0, 1]^2$ with increments 0.07	42
Figure 4.1	AGB measurements from the Penobscot Experimental Forest in Bradley, Maine	55
Figure 4.2	Fitted plot of AGB measurements together with the observed values	56
Figure 4.3	LiDAR data with fitted values and credible bands	57

Chapter 1

Introduction

Space-time models have become increasingly popular in scientific studies such as geology, climatology, geophysics, environmental and atmospheric sciences ([9] and [41]). Gaussian random fields (GRFs) are an important tool of this field. Developing covariance structures for GRFs, parameter estimation, and prediction in big spatial datasets are central challenges in spatio-temporal modeling. My thesis consists of three main chapters dealing with these main challenges in this area.

Finding conditions for equivalence of Gaussian measures is necessary for studying the asymptotics in the framework of covariance tapering. Equivalence of Gaussian measures is a classical problem in probability theory that has been studied since early 60s and 70s. Necessary and sufficient conditions for the equivalence of general Gaussian measures are given in [22], [15], [19]. However, the classical conditions are rather abstract, and may not be directly applicable to concrete cases. More verifiable conditions are then stated by restricting to a particular family of Gaussian measures. There are criteria for equivalence of stationary Gaussian processes in terms of their spectral densities in [25], and more general for stationary GRFs in [37] and [48]. However, investigation on the equivalence of nonstationary GRFs is limited to some special cases which are discussed in recent works (see, e.g., [7], [42], and [44]). In chapter 2, we study the equivalence of GRFs with stationary increments, and give a spectral condition under which the two GRFs will induce equivalent measures. This condition is easy to verify, and the results are applicable to different families of GRFs such

as the ones with the power law as their spectral densities with different power exponents in different dimensions, thus allowing anisotropy.

Researchers across scientific fields are asking increasingly complex questions that require spatio-temporal models able to accommodate rich association structures. In such settings, traditional assumptions such as stationarity and isotropy on the underlying GRFs may not be appropriate. These assumptions become increasingly tenuous as the size of the dataset and domain increase, which is increasingly common in many applied settings. There have been considerable attempts in recent literature to construct nonstationary anisotropic spatial models (See [24], [17], [30], [31], [26], [1] and references therein for the construction of such nonstationary models). In chapter 3, we study covariance estimation for a family of anisotropic GRFs with stationary increments using covariance tapering (See [18] for definition). Covariance tapering introduces sparsity to allow for more efficient matrix operations which is essential in dealing with large scale data sets. Consistency of MLE under the tapered covariance structure is established by putting a mild condition on the decay rate of the spectral density of the tapering function at infinity. This result also deals with the problem of equivalence of nonstationary GRFs.

Researchers are increasingly seeking inference about multiple outcome variables across spatial and temporal domains. This interest in applied community is due in large part to widely available geographic information systems (GIS), geographic position systems (GPS), and open source long-term monitoring networks. In such settings, models should capture both within location (i.e., among outcome variable) and between location covariance structures. Delivering this inference is often difficult due to the massive size of the given dataset. In chapter 4, we merge two ideas of factor modeling and low-rank spatial processes approximation to develop and propose a class of hierarchical low-rank spatial factor models

which offer a rich and flexible modeling option for dealing with large vector of outcomes observed at large number of locations. A Markov chain Monte Carlo algorithm is developed for estimation, and further, the full posterior distributions are recovered in a Bayesian predictive framework. We illustrate our methodology with simulation experiments and an environmental data set involving Light Detection and Ranging (LiDAR) signals together with above-ground biomass (AGB) from the Penobscot Experimental Forest in Bradley, Maine.

Chapter 2

Spectral Conditions for Equivalence of Gaussian Random Fields with stationary increments

Space-time models have become increasingly popular in scientific studies. Recent works in this field consist of construction of nonstationary anisotropic Gaussian random fields (GRFs). However, the problem of equivalence of measures corresponding to GRFs, which has direct consequences on the study of such models, is discussed mostly for the stationary case, and results for the nonstationary case are limited. In this chapter, we investigate the problem of equivalence of GRFs with stationary increments, and obtain sufficient conditions for equivalence in terms of the behavior of the spectral measures at infinity. Further, the main results are applied to a rich family of nonstationary space-time models with possible anisotropy behavior.

2.1 Introduction

Space-time models have become increasingly popular in scientific studies such as geology, climatology, geophysics, environmental and atmospheric sciences, etc ([9] and [41]). Gaussian random fields (GRFs) are basically the main tools in space-time modeling. Most of the para-

metric models proposed are GRFs with specific parametric covariance structure (See [10], [20], [43] and [21] for rich families of space-time covariance functions). The main job then is to find consistent estimates for the parameters, and finally use them for prediction of the underlying random field at unobserved locations, and also in performing statistical inference. Given a parametric family of Gaussian distributions, an important question is whether all the parameters are consistently estimable. First step to answer this question demands an investigation on the equivalence of the corresponding Gaussian measures involved in the model, since if two parameters in the parameter space give equivalent Gaussian distributions, then it is impossible to find consistent estimators for all the parameters involved regardless of the method chosen for estimation (See for example [50] for discussion on inconsistent estimation in Matern covariance functions). Another application of equivalence of Gaussian measures comes from covariance structure misspecification, and its effect on spatial interpolation ([41], chapter 4). In spatial statistics, an important goal is to predict the underlying GRF at unobserved locations. The common method is the best linear unbiased predictor (BLUP), which is known as kriging (See [41] for definition and properties of kriging). This, however needs information about the covariance function of the underlying GRF, which is in fact unknown. Using misspecified covariance functions in prediction may increase the prediction error. [41] proved that covariance functions which yield to equivalent Gaussian measures, have asymptotically the same prediction error rate (See also [39] and [40]). Therefore, finding sufficient conditions under which two covariance functions result in equivalent Gaussian distributions has direct impact in evaluating the prediction error in interpolation of spatial data, and thus proving asymptotically optimal prediction under misspecified covariance structure. There are other applications of equivalence and perpendicularity of GRFs in spatial modeling. For example, we refer to [18] and [28] for the application in covariance tapering.

Equivalence of Gaussian measures is a classical problem in probability theory that has been studied since early 60's and 70's. Necessary and sufficient conditions for the equivalence of general Gaussian measures are given in [22], [15], [19], and references therein (See also [32], [27] for conditions in terms of time domain reproducing kernel Hilbert spaces). These conditions are rather abstract, and may not be directly applicable to concrete cases. More specific conditions are then stated by restricting to a particular family of Gaussian measures. There are criteria for equivalence of stationary Gaussian processes in terms of their spectral densities in [25], and more general for stationary GRFs in [37] and [48]. However, investigation on the equivalence of nonstationary GRFs is limited to some special cases which are discussed in recent works. For instance, we refer to [7] on mixed fractional Brownian motion, [5] on Volterra processes, [38] on multiparameter Gaussian processes, and [42] on a family of power law covariance functions. Recently, [44] found sufficient conditions for equivalence of Gaussian processes with stationary increments in terms of their spectral densities similar to one in [25] for the stationary case. However, this result is only for one dimensional Gaussian processes, and the case for multidimensional GRFs with stationary increments is not discussed.

Space-time models have become more complex, and traditional assumptions such as stationarity and isotropy on the underlying GRFs might not be appropriate in data analysis in different fields such as geophysics, environmental and atmospheric sciences, etc. There have been considerable attempts in recent works for constructing nonstationary anisotropic spatial models (We refer to [24], [17], [30], [31],[26], [1] and references therein for the construction of such nonstationary models). Therefore, studying the equivalence of nonstationary GRFs is of interest in the study of such nontrivial families of space-time models. In this paper, we will investigate the equivalence of GRFs with stationary increments which might have

different regularities in each direction, and thus allows anisotropy as well, and find sufficient conditions for their equivalence in terms of the behavior of their spectral measures at infinity.

The rest of the chapter is organized as follows. We start section 2.2 by introducing some useful Hilbert spaces connected to the frequency domain, and study their structure. In section 2.3, we state the main result of the paper, which is sufficient conditions for equivalence of GRFs with stationary increments using the tail behavior of their spectral densities. In the last section, we apply the main results to a rich family of anisotropic nonstationary spatio-temporal Gaussian models.

2.2 Preliminary

Let $X = \{X_t : t \in \mathbb{R}^d\}$ be a centered GRF with stationary increments with $X(0) = 0$. According to [49], we have the following spectral representations for X , and for the covariance structure of X , respectively:

$$X(t) \stackrel{d}{=} \int_{\mathbb{R}^d} \left(e^{i\langle t, \lambda \rangle} - 1 \right) W(d\lambda), \quad (2.2.1)$$

$$C(t, s) = \mathbf{E}(X_t X_s) = \int_{\mathbb{R}^d} \left(e^{i\langle t, \lambda \rangle} - 1 \right) \left(e^{-i\langle s, \lambda \rangle} - 1 \right) F(d\lambda), \quad (2.2.2)$$

where W is a complex-valued Gaussian random measure with control measure F , which is a non-negative symmetric measure on $\mathbb{R}^d \setminus \{0\}$, called the spectral measure of X satisfying

$$\int_{\mathbb{R}^d} \frac{|\lambda|^2}{1 + |\lambda|^2} F(d\lambda) < \infty. \quad (2.2.3)$$

If the spectral measure is absolutely continuous with respect to the Lebesgue measure

on \mathbb{R}^d , we will call its Radon-Nikodym derivative spectral density, denoted by $f(\lambda)$. We will give conditions in terms of the spectral densities for the equivalence of GRFs with stationary increments, but first, we need to define precisely equivalent GRFs.

Definition 2.2.1. *For a fixed $D \subseteq \mathbb{R}^d$, we call two GRFs $X = \{X_t : t \in \mathbb{R}^d\}$, $Y = \{Y_t : t \in \mathbb{R}^d\}$ equivalent if they induce equivalent measures on the measurable space $(\mathbb{R}^D, \mathcal{B}(\mathbb{R}^D))$, in which $\mathcal{B}(\mathbb{R}^D)$ is the σ -field generated by the Borel subsets of \mathbb{R}^D . Moreover, we call them locally equivalent if they are equivalent on any bounded subset of \mathbb{R}^d .*

Such spectral representations for the covariance structure of GRFs of the type given in 2.2.2 makes an important bridge between the problem of equivalence of GRFs and the description of the space generated by the linear combinations of the kernel functions in the spectral representation. For that purpose, in this section we need to study such linear spaces, and explore their properties which will be valuable in the next section. For that, let's define for a fixed $D \subseteq \mathbb{R}^d$ bounded, and a measure F on \mathbb{R}^d satisfying 2.2.3, an incomplete Hilbert space $\mathcal{L}_D^e = \text{span} \left\{ e_t(\lambda) := e^{i\langle t, \lambda \rangle} - 1 : t \in D \right\}$ with the inner product $\langle e_t, e_s \rangle_F = \int_{\mathbb{R}^d} e_t(\lambda) \overline{e_s(\lambda)} F(d\lambda)$. We denote the closure in $L^2(F)$ of \mathcal{L}_D^e by $\mathcal{L}_D(F)$. Also, for $T > 0$ define

$$\Pi_T := \left\{ \lambda = (\lambda_1, \dots, \lambda_d) \in \mathbb{R}^d : -T \leq \lambda_j \leq T, \forall j = 1, \dots, d \right\}.$$

By looking at the basis functions generating \mathcal{L}_D^e , we realize that they are entire functions defined on \mathbb{C}^d , and they are of finite exponential type (See [36] for definition and properties). However, the elements in the completed Hilbert space, $\mathcal{L}_D(F)$, may not have the same properties. These properties enable us to apply Paley-Wiener type theorems, and get nice description of the elements in theses Hilbert spaces. This problem is discussed in details in

[34] and [33], and there are conditions mentioned in [33] on the measure F for this purpose. Therefore, we need to put further assumption on the spectral measure F in order to get the same nice properties. Here is the crucial assumption on the tail behavior of the spectral density $f(\lambda)$, assuming it exists:

$$(C1) \quad \text{There exist } c, k, \eta > 0 \text{ such that for all } |\lambda| > k, \text{ we have } f(\lambda) \geq \frac{c}{|\lambda|^\eta}.$$

This assumption on the spectral density will force the elements in $\mathcal{L}_D(F)$ to be entire functions of finite exponential type. The next two lemmas will prove this statement. The following lemma is taken from [46], Lemma 2.2, and we state it here again for completeness.

Lemma 2.2.1. *Suppose that the spectral density f satisfies (C1). Then for fixed $T > 0$, there exists positive constants C and M such that for all functions ϕ of the form*

$$\phi(\lambda) = \sum_{k=1}^n a_k \left(e^{i\langle t^k, \lambda \rangle} - 1 \right), \quad (2.2.4)$$

where $a_k \in \mathbb{R}$ and $t^k \in \Pi_T$, we have for all $z \in \mathbb{C}^d$

$$|\phi(z)| \leq C \|\phi\|_F \exp\{M|z|\}. \quad (2.2.5)$$

Moreover, for fixed $C_1 > 0$, there exists a positive constant C_2 such that for all functions of the form 2.2.4, we have

$$|\phi(z)| \leq C_2 |z| \|\phi\|_F \quad (2.2.6)$$

for all $|z| \leq C_1$.

One can use 2.2.5 to define the limiting functions in $\mathcal{L}_{\Pi_T}(F)$ in such a way that they also satisfy both 2.2.5 and 2.2.6. We will prove this in the next Lemma.

Lemma 2.2.2. *Let F be a spectral measure on \mathbb{R}^d which satisfies (C1). Then, for each $T > 0$, the space $\mathcal{L}_{\Pi_T}(F)$ consists of the restriction to \mathbb{R}^d of entire functions on \mathbb{C}^d of finite exponential type. Moreover, 2.2.6 is also true for all functions in $\mathcal{L}_{\Pi_T}(F)$.*

Proof. The idea of the proof is similar to [33], p. 304. Take a sequence $\phi_n \in \mathcal{L}_{\Pi_T}^e$, such that $\|\phi_n - \phi\|_F \rightarrow 0$ for some $\phi \in \mathcal{L}_{\Pi_T}(F)$. Then, it is a Cauchy sequence in $L^2(F)$, and using 2.2.5, we get

$$|\phi_n(z) - \phi_m(z)| \leq C \|\phi_n - \phi_m\|_F \exp\{M|z|\} \quad (2.2.7)$$

This means for each fixed $z \in \mathbb{C}^d$, the sequence $\phi_n(z)$ is a Cauchy sequence in \mathbb{C} . So, it is convergent. Let's denote the limit by $\tilde{\phi}(z)$. Now, since limit in $L^2(F)$ sense implies the almost sure convergence for a subsequence, $\phi = \tilde{\phi}$ a.e. with respect to F . From now on, we will take $\tilde{\phi}$ as our favorite version of the limits of functions in $\mathcal{L}_{\Pi_T}^e$. Therefore, the elements in the space $\mathcal{L}_{\Pi_T}(F)$, are not only the $L^2(F)$ limits of functions in $\mathcal{L}_{\Pi_T}^e$, but also the pointwise limits as well. Thus, both 2.2.5 and 2.2.6 are true for all the elements in $\mathcal{L}_{\Pi_T}(F)$. The only thing left to prove is that these functions are entire functions on \mathbb{C}^d . But this is true since any element of the space $\mathcal{L}_{\Pi_T}(F)$ is the locally uniform limit of functions of the form 2.2.4 which are obviously entire functions, and thus, they are also entire functions as well (This is called Weierstrass Theorem. See for example [14], Proposition 2.8, p. 52).

□

This lemma shows that if the spectral density satisfies the assumption (C1), we can complete the space $\mathcal{L}_{\Pi_T}^e$ in such a way that the resulting functions are locally uniform limits of entire functions, and hence, they are entire functions as well, and they are also of finite exponential type. Further, since 2.2.5 is true for all the elements in the Hilbert space $\mathcal{L}_{\Pi_T}(F)$, we can see that the point evaluators, i.e. the functionals on $\mathcal{L}_{\Pi_T}(F)$ of the

form $\phi \mapsto \phi(z)$ for each fixed $z \in \mathbb{C}^d$ are bounded operators. Now, we can apply Riesz representation Theorem (See [23], Theorem 3. p. 31) to prove that the space $\mathcal{L}_{\Pi_T}(F)$ is a Reproducing Kernel Hilbert Space (RKHS) in the sense of [2]. This means there exists a family of functions $K_T(.,.)$ from $\mathbb{R}^d \times \mathbb{R}^d$ to \mathbb{C} such that first of all $K_T(\omega,.) \in \mathcal{L}_{\Pi_T}(F)$ for all $\omega \in \mathbb{R}^d$, and secondly, for every $\phi \in \mathcal{L}_{\Pi_T}(F)$ and $\omega \in \mathbb{R}^d$, we have the following kernel property

$$\langle \phi, K_T(\omega, .) \rangle_F = \int_{\mathbb{R}^d} \phi(\lambda) \overline{K_T(\omega, \lambda)} F(d\lambda) = \phi(\omega). \quad (2.2.8)$$

Also, it is worthwhile to mention that the set of all functions $\{K_T(\omega, .) : \omega \in \mathbb{R}^d\}$ is dense in $\mathcal{L}_{\Pi_T}(F)$ (To see this, note that if $\phi \in \mathcal{L}_{\Pi_T}(F)$ is orthogonal to $K_T(\omega, .)$ for all $\omega \in \mathbb{R}^d$, then $\phi(\omega) = \langle \phi, K_T(\omega, .) \rangle_F = 0$, which implies $\phi = 0$.), and further for all $\psi \in L^2(F)$, the function

$$\omega \mapsto \langle \psi, K_T(\omega, .) \rangle_F \quad (2.2.9)$$

is the orthogonal projection of ψ on $\mathcal{L}_{\Pi_T}(F)$ (See the proof in [2], p. 345). We denote this projection by $\pi_{\mathcal{L}_{\Pi_T}(F)}\psi$.

Finding explicit forms of the reproducing kernels is not an easy job. However, in order to prove the results in section 2.3, we need upper bounds for the growth rate of the diagonal elements of the reproducing kernels at origin and also at infinity. The following proposition proves an important growth rate for the diagonal elements in the reproducing kernels.

Proposition 2.2.1. *Suppose that the spectral density of F , $f(\lambda)$ satisfies (C1). Then, for fixed $T > 0$ and $C_1 > 0$, there exists a positive constant C_2 such that*

$$K_T(\omega, \omega) \leq C_2 |\omega|^2 \quad (2.2.10)$$

for all $|\omega| < C_1$.

Proof. Since for fixed $\omega \in \mathbb{R}^d$, $K_T(\omega, \cdot) \in \mathcal{L}_{\Pi_T}(F)$, we can apply Lemma 2.2.2 to these functions. This means that the inequality 2.2.6 can be applied to them. After doing so, we get

$$\begin{aligned} |K_T(\omega, \lambda)| &\leq C_2 |\lambda| \|K_T(\omega, \cdot)\|_F \\ &= C_2 |\lambda| (K_T(\omega, \omega))^{1/2} \end{aligned}$$

for all $\omega \in \mathbb{R}^d$ and $\lambda \in \mathbb{C}^d$ with $|\lambda| < C_1$. Now, if we put $\lambda = \omega$, since the constant C_2 doesn't depend on ω , we will get the desired result. \square

We also need to define another Hilbert space based on the tensor product of the elements in $\mathcal{L}_{\Pi_T}(F)$. For that purpose, first we define $\mathcal{L}_{\Pi_T}^e \otimes \mathcal{L}_{\Pi_T}^e$ to be the span of functions $(e_t \otimes e_s)(\omega, \lambda) := e_t(\omega) \overline{e_s(\lambda)}$ with $t, s \in \Pi_T$. Now, denote by $\mathcal{L}_{\Pi_T}(F) \otimes \mathcal{L}_{\Pi_T}(F)$ the closure in $L^2(F \otimes F)$ of the space $\mathcal{L}_{\Pi_T}^e \otimes \mathcal{L}_{\Pi_T}^e$. According to [2], Theorem 1, p. 361, the new Hilbert space $\mathcal{L}_{\Pi_T}(F) \otimes \mathcal{L}_{\Pi_T}(F)$ is also a RKHS with reproducing kernel

$$((\omega_1, \lambda_1), (\omega_2, \lambda_2)) \mapsto K_T(\omega_1, \omega_2) \overline{K_T(\lambda_1, \lambda_2)}. \quad (2.2.11)$$

This means in fact that for $\psi \in \mathcal{L}_{\Pi_T}(F) \otimes \mathcal{L}_{\Pi_T}(F)$,

$$\langle \psi, K_T(\omega, \cdot) \otimes K_T(\lambda, \cdot) \rangle_{F \otimes F} = \psi(\omega, \lambda). \quad (2.2.12)$$

We finish this section by a lemma stating that the norm of the elements in spaces $\mathcal{L}_{\Pi_T}(F)$ only depends on the tail behavior of the spectral measure F .

Lemma 2.2.3. *Suppose f_0 and f_1 are two spectral densities satisfying the condition (C1), and further $f_0(\lambda) \asymp f_1(\lambda)$ as $|\lambda| \rightarrow \infty$. Then, $\mathcal{L}_{\Pi_T}(f_0) = \mathcal{L}_{\Pi_T}(f_1)$, and further there exist positive constants C_3 and C_4 such that for all $\phi \in \mathcal{L}_{\Pi_T}(f_0)$*

$$C_3 \|\phi\|_{f_1} \leq \|\phi\|_{f_0} \leq C_4 \|\phi\|_{f_1}.$$

Proof. Suppose $g \in \mathcal{L}_{\Pi_T}(f_0)$. There exists functions g_n of the form 2.2.4 such that

$\|g - g_n\|_{f_0} \rightarrow 0$ as $n \rightarrow \infty$. Now, using lemma 2.2.2, we get

$$\begin{aligned} \|g - g_n\|_{f_1}^2 &= \int_{\mathbb{R}^d} |g(\lambda) - g_n(\lambda)|^2 f_1(\lambda) d\lambda \\ &\leq C_2 \|g - g_n\|_{f_0}^2 \int_{|\lambda| < C_1} |\lambda|^2 f_1(\lambda) d\lambda \\ &\quad + C \|g - g_n\|_{f_0}^2. \end{aligned}$$

Therefore, $g \in \mathcal{L}_{\Pi_T}(f_1)$ as well. Similarly by simply putting $g - g_n = \phi$ in the above calculations, we can verify the second part of the lemma. \square

2.3 Main Results

In this section, we study the equivalence of GRFs with stationary increments, and clarify its connection to the Hilbert spaces constructed in section 2.2. In particular, the role of the reproducing kernels of the RKHS $\mathcal{L}_{\Pi_T}(F)$ will be emphasized. We start this section by rephrasing Theorem 5, p. 84. in [25] into our context (See also [48], Theorem 1, p. 149).

Theorem 2.3.1. *Two centered GRFs with stationary increments, and the spectral measures F_0 and F_1 are equivalent on D if and only if:*

$$(i) \|\phi\|_{F_0} \asymp \|\phi\|_{F_1}, \forall \phi \in \mathcal{L}_D^e,$$

(ii) There exists a function $\psi \in \mathcal{L}_D(F_0) \otimes \mathcal{L}_D(F_0)$ such that for all $t, s \in D$

$$\langle e_t, e_s \rangle_{F_0} - \langle e_t, e_s \rangle_{F_1} = \langle \psi, e_t \otimes e_s \rangle_{F_0 \otimes F_0}. \quad (2.3.1)$$

Proof. Proof is actually reconstruction of the proof of the Theorem 5, p. 84 of [25]. The starting point is however, Theorem 4, p. 80. This Theorem can be adapted to our context since its proof only involves the “entropy distance” between the Gaussian measures, and thus is true for general GRFs (See also [6], Theorem 4.1, 4.4 pp. 180-185). After doing so, we get that two GRFs with stationary increments and spectral measures F_0 and F_1 are equivalent on $D \subseteq \mathbb{R}^d$, if and only if, $\|\phi\|_{F_0} \asymp \|\phi\|_{F_1}, \forall \phi \in \mathcal{L}_D^e$, and $\Delta = I - A^*A$ is a Hilbert-Schmidt operator in $\mathcal{L}_D(F_0)$, in which I is the identity operator on $\mathcal{L}_D(F_0)$, and $A : \mathcal{L}_D(F_0) \mapsto \mathcal{L}_D(F_1)$ with $A\phi = \phi$ for all $\phi \in \mathcal{L}_D(F_0)$. Now, since Δ is a self-adjoint operator, if it is also a Hilbert-Schmidt operator, by the Spectral Theorem (See [13], corollary 5 p. 905), we can conclude that there exists an orthonormal basis for $\mathcal{L}_D(F_0)$ consisting of the eigenvectors of Δ , denoting them by $\phi_j, j = 1, 2, \dots$, with corresponding eigenvalues $\lambda_j, j = 1, 2, \dots$ with $\sum_j \lambda_j^2 < \infty$. Note that we can write $\sum_j \lambda_j^2 = \sum_{j,k} \langle \Delta \phi_j, \phi_k \rangle_{F_0}^2$. The square root of this quantity is called the Hilbert-Schmidt norm. This norm doesn’t depend on the choice of the orthonormal basis (See [13], Lemma 2, p. 1010). Therefore, we can rephrase Theorem 4 in the following form: two GRFs with stationary increments and spectral measures F_0 and F_1 are equivalent on D , if and only if, $\|\phi\|_{F_0} \asymp \|\phi\|_{F_1}, \forall \phi \in \mathcal{L}_D^e$, and $\sum_{j,k} \langle \Delta \phi_j, \phi_k \rangle_{F_0}^2 < \infty$ for any orthonormal basis for $\mathcal{L}_D(F_0)$.

Now, take an arbitrary orthonormal basis for $\mathcal{L}_D(F_0)$, ϕ_1, ϕ_2, \dots , and suppose $\sum_{j,k} \langle \Delta \phi_j, \phi_k \rangle_{F_0}^2 < \infty$. Define $\psi_0(\omega, \lambda) = \sum_{j,k} \langle \Delta \phi_j, \phi_k \rangle_{F_0} \phi_j(\omega) \overline{\phi_k(\lambda)}$. We can see that

$\|\psi_0\|_{F_0 \otimes F_0}^2 = \sum_{j,k} \langle \Delta \phi_j, \phi_k \rangle_{F_0}^2 < \infty$, and thus by the form of ψ_0 , it's clear that it belongs to $\mathcal{L}_D(F_0) \otimes \mathcal{L}_D(F_0)$. Also, observe that

$$\begin{aligned} \langle \psi_0, \phi_j \otimes \phi_k \rangle_{F_0 \otimes F_0} &= \langle \Delta \phi_j, \phi_k \rangle_{F_0} \\ &= \langle (I - A^* A) \phi_j, \phi_k \rangle_{F_0} \\ &= \langle \phi_j, \phi_k \rangle_{F_0} - \langle \phi_j, \phi_k \rangle_{F_1}. \end{aligned}$$

This shows that 2.3.1 holds for orthonormal basis of the space $\mathcal{L}_D(F_0)$. Therefore, by continuity of inner product 2.3.1 will be true for all the elements in $\mathcal{L}_D(F_0)$, especially for e_t and e_s when $t, s \in D$.

Conversely, suppose there exists a function $\psi_0 \in \mathcal{L}_D(F_0) \otimes \mathcal{L}_D(F_0)$, such that $\langle \phi_j, \phi_k \rangle_{F_0} - \langle \phi_j, \phi_k \rangle_{F_1} = \langle \psi_0, \phi_j \otimes \phi_k \rangle_{F_0 \otimes F_0}$ for an orthonormal basis ϕ_j 's for $\mathcal{L}_D(F_0)$. Then, we have

$$\begin{aligned} \sum_{j,k} \langle \Delta \phi_j, \phi_k \rangle_{F_0}^2 &= \sum_{j,k} \left(\langle \phi_j, \phi_k \rangle_{F_0} - \langle \phi_j, \phi_k \rangle_{F_1} \right)^2 \\ &= \sum_{j,k} \langle \psi_0, \phi_j \otimes \phi_k \rangle_{F_0 \otimes F_0}^2 \\ &\leq \|\psi_0\|_{F_0 \otimes F_0}^2 \\ &< \infty. \end{aligned}$$

This completes the proof. □

Previous theorem is stated in general form where equivalence of GRFs is guaranteed with no restriction on the spectral measure. However, verifying the second condition in this theorem which involves finding a function $\psi \in \mathcal{L}_D(F_0) \otimes \mathcal{L}_D(F_0)$ with the property 2.3.1, seems to be hard. If we put the condition (C1) on the tail behavior of the spectral density

of one of the measures involved, we will get the following theorem using the reproducing kernels of $\mathcal{L}_{\Pi_T}(F_0)$. In fact, this theorem clarifies what must be the function ψ in theorem 2.3.1.

Theorem 2.3.2. *Two centered GRFs with stationary increments, and spectral measures F_0 and F_1 , with F_0 satisfying assumption (C1), are equivalent on Π_T for some $T > 0$ if and only if:*

- (i) $\|\phi\|_{F_0} \asymp \|\phi\|_{F_1}, \forall \phi \in \mathcal{L}_{\Pi_T}^e$,
- (ii) $\psi(\omega, \lambda) = K_T^0(\omega, \lambda) - \int_{\mathbb{R}^d} K_T^0(\omega, \gamma) \overline{K_T^0(\lambda, \gamma)} F_1(d\gamma) \in \mathcal{L}_{\Pi_T}(F_0) \otimes \mathcal{L}_{\Pi_T}(F_0)$ where $K_T^0(., .)$ are the reproducing kernels of the space $\mathcal{L}_{\Pi_T}(F_0)$.

Proof. First, assume that the measures induced by them are equivalent. Then, by theorem 2.3.1, there exists a function $\psi \in \mathcal{L}_{\Pi_T}(F_0) \otimes \mathcal{L}_{\Pi_T}(F_0)$ such that the equation 2.3.1 will hold. Now, because of bilinearity and continuity of inner product together with the fact that $\mathcal{L}_{\Pi_T}(F_0) = \mathcal{L}_{\Pi_T}(F_1)$, we get

$$\langle \phi_1, \phi_2 \rangle_{F_0} - \langle \phi_1, \phi_2 \rangle_{F_1} = \langle \psi, \phi_1 \otimes \phi_2 \rangle_{F_0 \otimes F_0} \quad (2.3.2)$$

for all $\phi_1, \phi_2 \in \mathcal{L}_{\Pi_T}(F_0) = \mathcal{L}_{\Pi_T}(F_1)$. Now, simply choose for fixed $\omega, \lambda \in \mathbb{R}^d$, $\phi_1(\gamma) = K_T^0(\omega, \gamma)$ and $\phi_2(\gamma) = K_T^0(\lambda, \gamma)$, and replace them in equation 2.3.2 to get

$$\psi(\omega, \lambda) = K_T^0(\omega, \lambda) - \int_{\mathbb{R}^d} K_T^0(\omega, \gamma) \overline{K_T^0(\lambda, \gamma)} F_1(d\gamma).$$

Conversely, since $\psi \in \mathcal{L}_{\Pi_T}(F_0) \otimes \mathcal{L}_{\Pi_T}(F_0)$, by reproducing kernel property we get $\psi(\omega, \lambda) = \langle \psi, K_T^0(\omega, .) \otimes K_T^0(\lambda, .) \rangle_{F_0 \otimes F_0}$. Also, note that by the form of ψ , we have

$\psi(\omega, \lambda) = \langle K_T^0(\omega, \cdot), K_T^0(\lambda, \cdot) \rangle_{F_0} - \langle K_T^0(\omega, \cdot), K_T^0(\lambda, \cdot) \rangle_{F_1}$. Combining them together, we get

$$\langle K_T^0(\omega, \cdot), K_T^0(\lambda, \cdot) \rangle_{F_0} - \langle K_T^0(\omega, \cdot), K_T^0(\lambda, \cdot) \rangle_{F_1} = \langle \psi, K_T^0(\omega, \cdot) \otimes K_T^0(\lambda, \cdot) \rangle_{F_0 \otimes F_0}. \quad (2.3.3)$$

Now, since the $\text{span}\{K_T^0(\omega, \cdot); \omega \in \mathbb{R}^d\}$ is dense in $\mathcal{L}_{\Pi_T}(F_0)(= \mathcal{L}_{\Pi_T}(F_1))$, the equality 2.3.3 will be true for all the elements in $\mathcal{L}_{\Pi_T}(F_0)$.

□

Remark 2.3.1. *In order to prove that the function ψ appeared in Theorem 2.3.2 is in the space $\mathcal{L}_{\Pi_T}(F_0) \otimes \mathcal{L}_{\Pi_T}(F_0)$, we only need to show that it belongs to $L^2(F_0 \otimes F_0)$. To prove this fact, we will show that the orthogonal projection of ψ on $\mathcal{L}_{\Pi_T}(F_0) \otimes \mathcal{L}_{\Pi_T}(F_0)$ is in fact itself. For that, observe*

$$\begin{aligned} \left(\pi_{\mathcal{L}_{\Pi_T}(F_0) \otimes \mathcal{L}_{\Pi_T}(F_0)}(\psi) \right) (\omega, \lambda) &= \left\langle \psi, K_T^0(\omega, \cdot) \otimes K_T^0(\lambda, \cdot) \right\rangle_{F_0 \otimes F_0} \\ &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \psi(x, y) \overline{K_T^0(\omega, x)} \overline{K_T^0(\lambda, y)} F_0(dx) F_0(dy) \\ &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K_T^0(x, \gamma) \overline{K_T^0(y, \gamma)} \overline{K_T^0(\omega, x)} \overline{K_T^0(\lambda, y)} \\ &\quad \times F_0(dx) F_0(dy) (F_0(d\gamma) - F_1(d\gamma)) \\ &= \int_{\mathbb{R}^d} K_T^0(\omega, \gamma) \overline{K_T^0(\lambda, \gamma)} (F_0(d\gamma) - F_1(d\gamma)) \\ &= \psi(\omega, \lambda). \end{aligned}$$

This is a huge reduction which strengthen the applicability of Theorem 2.3.2.

By Remark 2.3.1, checking the second assumption in Theorem 2.3.2 is reduced to only checking $\psi \in L^2(F_0 \otimes F_0)$. However, checking the first assumption seems to be not an easy job since we need to compare the norms of all the elements in the space $\mathcal{L}_{\Pi_T}^e$ under

two different measures. For that purpose, in the following, we will try to find equivalent conditions which may be easier to verify in application.

It is well known (See [13], p. 1009) that for $\psi \in L^2(F \otimes F)$, one can define a Hilbert-Schmidt operator on $L^2(F)$ in the following way

$$(\psi\phi)(\omega) = \int_{\mathbb{R}^d} \psi(\omega, \lambda) \phi(\lambda) F(d\lambda) \quad (2.3.4)$$

for every $\phi \in L^2(F)$. If we use specifically the ψ defined in Theorem 2.3.2, and restrict the domain to $\mathcal{L}_{\Pi_T}(F)$, we will have again a Hilbert-Schmidt operator on $\mathcal{L}_{\Pi_T}(F)$. Note that the image of the operator ψ is in fact inside the $\mathcal{L}_{\Pi_T}(F)$. To prove this, observe that for $\phi \in \mathcal{L}_{\Pi_T}(F)$,

$$\begin{aligned} \left(\pi_{\mathcal{L}_{\Pi_T}}(\psi\phi) \right)(\omega) &= \int_{\mathbb{R}^d} (\psi\phi)(x) \overline{K_T(\omega, x)} F(dx) \\ &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \phi(y) \psi(x, y) \overline{K_T(\omega, x)} F(dx) F(dy) \\ &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \phi(y) K_T(x, \gamma) \overline{K_T(y, \gamma) K_T(\omega, x)} F(dx) F(dy) \\ &\quad \times (F(d\gamma) - F_1(d\gamma)) \\ &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \phi(y) K_T(\omega, \gamma) \overline{K_T(y, \gamma)} (F(d\gamma) - F_1(d\gamma)) F(dy) \\ &= \int_{\mathbb{R}^d} \phi(y) \psi(\omega, y) F(dy) \\ &= (\psi\phi)(\omega). \end{aligned}$$

This argument shows that $\psi\phi \in \mathcal{L}_{\Pi_T}(F)$ for any $\phi \in \mathcal{L}_{\Pi_T}(F)$. Also, observe that since $\psi(\omega, \lambda) = \overline{\psi(\lambda, \omega)}$, the operator ψ is self-adjoint. This fact together with compactness of this operator (Since ψ is a Hilbert-Schmidt operator, it is already compact, see [13], p. 1009) enable us to use the spectral Theorem for compact normal operators (See [13], Corollary 5,

p. 905), which we will use in the proof of the next Theorem. In fact, next theorem shows that the first condition in Theorem 2.3.2 can be replaced by $1 \notin \sigma(\psi)$, where $\sigma(\psi)$ is the spectrum of the operator ψ (See [13], p. 902 for definition).

Theorem 2.3.3. *Two GRFs with stationary increments and spectral measures F_0 and F_1 with F_0 satisfying the condition (C1), are equivalent on Π_T if and only if the function defined by*

$$\psi(\omega, \lambda) = K_T^0(\omega, \lambda) - \int_{\mathbb{R}^d} K_T^0(\omega, \gamma) \overline{K_T^0(\lambda, \gamma)} F_1(d\gamma) \quad (2.3.5)$$

belongs to $\mathcal{L}_{\Pi_T}(F_0) \otimes \mathcal{L}_{\Pi_T}(F_0)$, and $1 \notin \sigma(\psi)$.

Proof. From 2.3.2, by putting $\phi_1 = \phi_2 = \phi$, and the definition of the operator ψ in 2.3.4, we get

$$\frac{\|\phi\|_{F_1}^2}{\|\phi\|_{F_0}^2} = 1 - \frac{\langle \psi\phi, \phi \rangle_{F_0}}{\|\phi\|_{F_0}^2} \quad (2.3.6)$$

for all $\phi \in \mathcal{L}_{\Pi_T}(F_0)$. This simply implies that first $\sigma(\psi) \subseteq (-\infty, 1]$, and second there exists a finite positive constant C such that $\|\phi\|_{F_1} \leq C\|\phi\|_{F_0}$ for all $\phi \in \mathcal{L}_{\Pi_T}^e$ since ψ is a bounded operator. This fact shows that proving $\psi \in L^2(F_0 \otimes F_0)$ is helping us to verify half of what we need in the first condition of theorem 2.3.2 as well. What remains is to show that $1 \notin \sigma(\psi)$ if and only if there exists a positive constant c such that $\|\phi\|_{F_0} \leq c\|\phi\|_{F_1}$ for all $\phi \in \mathcal{L}_{\Pi_T}^e$.

First, suppose that $\|\phi\|_{F_0} \leq c\|\phi\|_{F_1}$ for some $c > 0$. If $1 \in \sigma(\psi)$, it means that there exists $\phi \in \mathcal{L}_{\Pi_T}(F_0)$ with $\|\phi\|_{F_0} = 1$ such that $\psi\phi = \phi$. Putting it in 2.3.6, we get $\|\phi\|_{F_1} = 0$ which is contradiction. Conversely, suppose $1 \notin \sigma(\psi)$, and also there exists a sequence $\phi_n \in \mathcal{L}_{\Pi_T}^e$ such that $\|\phi_n\|_{F_0} = 1$ for all $n \geq 1$, and $\|\phi_n\|_{F_1} \rightarrow 0$ as $n \rightarrow \infty$. Since ψ is a self-adjoint compact operator, by Corollary 5 p. 905 in [13], there exists a

countable orthonormal basis for $\mathcal{L}_{\Pi_T}(F_0)$ consisting of eigenvectors of ψ , denoting them by $g_j, j = 1, 2, \dots$ with corresponding eigenvalues λ_j . Now, each ϕ_n has the representation $\phi_n = \sum_{j=1}^{\infty} a_{nj} g_j$ for $a_{nj} \in \mathbb{R}$. Putting this sequence back to the equation 2.3.6, we get that $\langle \psi \phi_n, \phi_n \rangle_{F_0} \rightarrow 1$ which means $\sum_{j=1}^{\infty} a_{nj}^2 \lambda_j \rightarrow 1$ as $n \rightarrow \infty$. Now, since $1 = \|\phi_n\|_{F_0}^2 = \sum_{j=1}^{\infty} a_{nj}^2$, we can rewrite the above equation as $0 \leq \sum_{j=1}^{\infty} a_{nj}^2 (1 - \lambda_j) \rightarrow 0$ (This quantity is non-negative since all the eigenvalues are bounded above by 1). Since $1 \notin \sigma(\psi)$, and $\{\lambda_j, j = 1, 2, \dots\}$ has no accumulation points in \mathbb{C} except possibly 0 (See [13], Corollary 5 p. 905), there exists $\epsilon > 0$ such that $\sup \{\lambda_j, j \geq 1\} = 1 - \epsilon$. However, this implies that $\sum_{j=1}^{\infty} a_{nj}^2 (1 - \lambda_j) \geq \epsilon \sum_{j=1}^{\infty} a_{nj}^2 = \epsilon$ for all $n \in \mathbb{N}$, which is contradiction by the fact that this sequence must go to 0 when n goes to ∞ . This completes the proof. \square

Remark 2.3.2. *Note that based on the proof of theorem 2.3.3, we can change the first condition in theorem 2.3.2 to*

(i)' *There exists a positive constant c such that $\|\phi\|_{F_0} \leq c \|\phi\|_{F_1}$ for all $\phi \in \mathcal{L}_{\Pi_T}^e$.*

As lemma 2.2.3 emphasizes that the behavior of the spectral measure at origin does not affect the structure of the space $\mathcal{L}_{\Pi_T}(F)$, one might expect the same formation in terms of the equivalence of Gaussian measures. The following Theorem shows that changing the spectral measure on bounded subsets of \mathbb{R}^d will not affect the equivalence of the corresponding GRFs. In other words, for checking the equivalence of GRFs, only the behavior of their spectral measures at infinity is important.

Theorem 2.3.4. *Suppose two GRFs with stationary increments have spectral measures F_0 and F_1 such that F_0 satisfies the condition (C1), and $F_0 = F_1$ on I^c , where I is a bounded subset of \mathbb{R}^d . Then, these two GRFs are locally equivalent.*

Proof. Define $\tilde{F}_1(d\lambda) = 1_{I^c}(\lambda)F_0(d\lambda)$. First, we show that F_0 and \tilde{F}_1 will produce locally equivalent GRFs with stationary increments. For that, fix $T > 0$. We will investigate the equivalence of measures on Π_T . The function ψ appearing in Theorem 2.3.3 in this case is given by

$$\psi(\omega, \lambda) = \int_I K_T^0(\omega, \gamma) \overline{K_T^0(\lambda, \gamma)} F_0(d\gamma).$$

Now, by Fubini's Theorem, and the reproducing kernel property we get

$$\begin{aligned} \|\psi\|_{F_0 \otimes F_0}^2 &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \left| \int_I K_T^0(\omega, \gamma) \overline{K_T^0(\lambda, \gamma)} F_0(d\gamma) \right|^2 F_0(d\omega) F_0(d\lambda) \\ &= \int_I \int_I \left(\int_{\mathbb{R}^d} K_T^0(\omega, \gamma) \overline{K_T^0(\omega, \gamma')} F_0(d\omega) \int_{\mathbb{R}^d} K_T^0(\lambda, \gamma') \overline{K_T^0(\lambda, \gamma)} F_0(d\lambda) \right) \\ &\quad \times F_0(d\gamma) F_0(d\gamma') \\ &= \int_I \int_I \left| K_T^0(\gamma, \gamma') \right|^2 F_0(d\gamma) F_0(d\gamma'). \end{aligned}$$

Now, using Proposition 2.2.1, and the fact that $\left| K_T^0(\gamma, \gamma') \right|^2 \leq K_T^0(\gamma, \gamma) K_T^0(\gamma', \gamma')$, we get

$$\begin{aligned} \|\psi\|_{F_0 \otimes F_0}^2 &\leq \left(\int_I K_T^0(\gamma, \gamma) F_0(d\gamma) \right)^2 \\ &\leq \text{const.} \left(\int_I |\gamma|^2 F_0(d\gamma) \right)^2 \\ &< \infty \end{aligned}$$

by 2.2.3. This implies $\psi \in \mathcal{L}_{\Pi_T}(F_0) \otimes \mathcal{L}_{\Pi_T}(F_0)$.

It remains to show that $1 \notin \sigma(\psi)$. For that purpose, take an arbitrary $\phi \in \mathcal{L}_{\Pi_T}(F_0)$, and observe that (We use the fact that $K_T^0(\omega, \lambda) = \overline{K_T^0(\lambda, \omega)}$. See [2], p. 344)

$$\begin{aligned}
(\psi\phi)(\lambda) &= \int_{\mathbb{R}^d} \phi(\omega) \psi(\lambda, \omega) F_0(d\omega) \\
&= \int_{\mathbb{R}^d} \phi(\omega) \left(\int_I K_T^0(\lambda, \gamma) \overline{K_T^0(\omega, \gamma)} F_0(d\gamma) \right) F_0(d\omega) \\
&= \frac{\int_I \int_{\mathbb{R}^d} \overline{\phi(\omega)} K_T^0(\lambda, \gamma) \overline{K_T^0(\gamma, \omega)} F_0(d\omega) F_0(d\gamma)}{\int_I \overline{\phi(\gamma)} K_T^0(\lambda, \gamma) F_0(d\gamma)} \\
&= \frac{\int_I \overline{\phi(\gamma)} K_T^0(\lambda, \gamma) F_0(d\gamma)}{\int_I \overline{\phi(\gamma)} K_T^0(\lambda, \gamma) F_0(d\gamma)} \\
&= \pi_{\mathcal{L}_{\Pi_T}(F_0)}(\overline{\phi} 1_I)(\lambda).
\end{aligned}$$

Therefore, if $\psi\phi = \phi$, it implies in particular that $\|\phi\|_{F_0} \leq \|\overline{\phi} 1_I\|_{F_0}$. This means $\phi = 0$ almost everywhere with respect to F_0 in I^c . Hence, since ϕ is an entire function, this implies that $\phi = 0$. Thus, 1 cannot be in the spectrum of ψ .

So far, we proved GRFs with spectral measures F_0 and \tilde{F}_1 are locally equivalent, but since $F_0 = F_1$ on I , similarly, we can say that F_1 and \tilde{F}_1 produce locally equivalent GRFs. Putting these two together, we get the desired result. \square

Theorems 2.3.2, and 2.3.3 give necessary and sufficient conditions for equivalence of GRFs with stationary increments, but still verifying the conditions mentioned in these two theorems for some concrete cases might be difficult. In the literature, there are sufficient conditions for equivalence of certain GRFs in terms of only their spectral densities. These conditions are easily verifiable once the two spectral densities are given. For example, we refer to [25], theorem 17, p. 104, [37], theorem 4, and [48], theorem 4, p. 156 for stationary GRFs, and also [44], theorem 5.3, [45], theorem 1, and [42] theorem 1 for some nonstationary cases. Therefore, in the rest of this section, we attempt to find similar sufficient conditions for equivalence, only in terms of the spectral measures.

Theorem 2.3.5. *Suppose that the spectral measure F_0 and F_1 have positive densities f_0 and f_1 with respect to the Lebesgue measure, and F_0 satisfies the condition (C1). If there exists a finite constant $C > 0$ such that $\|\phi\|_{F_0} \leq C\|\phi\|_{F_1}$ for all $\phi \in \mathcal{L}_{\Pi_T}^e$, and*

$$\int_{|\lambda|>k} \left(\frac{f_1(\lambda) - f_0(\lambda)}{f_0(\lambda)} \right)^2 K_T^0(\lambda, \lambda) f_0(\lambda) d\lambda < \infty \quad (2.3.7)$$

for some $k > 0$, then GRFs with stationary increments and spectral measures F_0 and F_1 are equivalent on Π_T .

Proof. Applying theorem 2.3.4, we can change the value of f_1 on any bounded set, without having any consequences on the equivalence. So, we assume here that $f_0 = f_1$ on $|\lambda| \leq k$.

The function ψ in Theorem 2.3.3 here will be of the form

$$\begin{aligned} \psi(\omega, \lambda) &= \int_{\mathbb{R}^d} K_T^0(\omega, \gamma) \overline{K_T^0(\lambda, \gamma)} (f_0(\lambda) - f_1(\lambda)) d\gamma \\ &= \pi_{\mathcal{L}_{\Pi_T}(F_0)} \left(K_T^0(\omega, \cdot) \frac{f_0 - f_1}{f_0} \right) (\lambda). \end{aligned}$$

(observe that since $|K_T^0(\omega, \lambda)|^2 \leq K_T^0(\omega, \omega) K_T^0(\lambda, \lambda)$, 2.3.7 implies that $K_T^0(\omega, \cdot) \frac{f_0 - f_1}{f_0} \in L^2(F_0)$ for all $\omega \in \mathbb{R}^d$, and hence using the orthogonal projection makes sense). Now, it

follows that

$$\begin{aligned}
\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |\psi(\omega, \lambda)|^2 F_0(d\lambda) F_0(d\omega) &= \int_{\mathbb{R}^d} \left\| \pi_{\mathcal{L}_D(F_0)} \left(K_T^0(\omega, \cdot) \frac{f_0 - f_1}{f_0} \right) \right\|_{F_0}^2 F_0(d\omega) \\
&\leq \int_{\mathbb{R}^d} \left\| K_T^0(\omega, \cdot) \frac{f_0 - f_1}{f_0} \right\|_{F_0}^2 F_0(d\omega) \\
&= \int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} |K_T^0(\omega, \gamma)|^2 \left(\frac{f_0(\gamma) - f_1(\gamma)}{f_0(\gamma)} \right)^2 F_0(d\gamma) \right) \\
&\quad \times F_0(d\omega) \\
&= \int_{\mathbb{R}^d} \left(\frac{f_0(\gamma) - f_1(\gamma)}{f_0(\gamma)} \right)^2 K_T^0(\gamma, \gamma) f_0(\gamma) d\gamma \\
&= \int_{|\gamma| > k} \left(\frac{f_1(\gamma) - f_0(\gamma)}{f_0(\gamma)} \right)^2 K_T^0(\gamma, \gamma) f_0(\gamma) d\gamma \\
&< \infty
\end{aligned}$$

by the integrability assumption 2.3.7. Hence $\psi \in L^2(F_0 \otimes F_0)$, and this completes the proof. \square

In the condition 2.3.7, in addition to the behavior of the spectral densities at infinity, the growth rate of the diagonal elements of the reproducing kernels of the space $\mathcal{L}_{\Pi_T}(F_0)$ at infinity is also playing an important role. Since finding explicit forms of reproducing kernels are difficult, we need at least to find upper bounds for the growth rate of the diagonal terms. The following condition on spectral density helps us to accomplish this task:

(C2) For spectral density f , there exist an entire function ϕ_0 on \mathbb{C}^d of finite exponential type such that $f(\lambda) \asymp |\phi_0(\lambda)|^2$ as $|\lambda| \rightarrow \infty$ on \mathbb{R}^d .

As it was mentioned, condition (C2) will force an upper bound for the behavior of the reproducing kernels on the diagonal at infinity, and the following lemma explores this fact.

Lemma 2.3.1. *Suppose f_0 is a spectral density such that it satisfies (C1) and (C2) for some entire function ϕ_0 . Then, for $T > 0$, there exists a finite constant $C > 0$ such that the reproducing kernel K_T^0 of $\mathcal{L}_{\Pi_T}(f_0)$ satisfies*

$$\left| K_T^0(\omega, \lambda) \right|^2 \leq C \frac{K_T^0(\omega, \omega)}{f_0(\lambda)}$$

for all real ω , and all λ large enough. In particular,

$$\left| K_T^0(\lambda, \lambda) \right| \leq \frac{C}{f_0(\lambda)} \quad (2.3.8)$$

for all λ large enough.

Proof. The idea of the proof is similar to the one in Lemma 3 p. 57 in [45]. Put $f(\lambda) = |\phi_0(\lambda)|^2$. Since f and f_0 are comparable at ∞ , and f is bounded around 0, it is clear that f is satisfying both conditions 2.2.3 and (C1). This means we can define $\mathcal{L}_{\Pi_T}(f)$ the same way, and this space is also a RKHS. Consider an arbitrary orthonormal basis for this space, and denote them by $\psi_k, k = 1, 2, \dots$. Now, by Lemma 2.2.2, they are entire functions on \mathbb{C}^d with finite exponential type which doesn't depend on k . Further, we know $\psi_k \phi_0 \in L^2(\mathbb{R}^d)$ since

$$\int_{\mathbb{R}^d} |\psi_k(\lambda) \phi_0(\lambda)|^2 d\lambda = \int_{\mathbb{R}^d} |\psi_k(\lambda)|^2 f(\lambda) d\lambda = 1 < \infty.$$

Therefore, we can apply the Paley-Wiener Theorem ([36], Theorem 3.4.2. p. 171) to get $\psi_k \phi_0 = \widehat{g_k}$ for certain functions $g_k \in L^2(B)$ where B is a bounded subset of \mathbb{R}^d (Here \widehat{h} stands for the Fourier transform of h). By Parseval relation, we can deduce that g_k 's are

orthonormal in $L^2(B)$. It follows from Bessel's inequality that

$$\begin{aligned} \sum_k |\psi_k(\lambda)|^2 f(\lambda) &= \sum_k \left| \int_B e^{-i\langle t, \lambda \rangle} g_k(t) dt \right|^2 \\ &\leq \int_B |e^{i\langle t, \lambda \rangle}|^2 dt \\ &= m(B), \end{aligned}$$

where $m(B)$ is the Lebesgue measure of B . Therefore,

$$\sum_k |\psi_k(\lambda)|^2 \leq m(B)/f(\lambda) \leq C/f_0(\lambda)$$

for large enough λ . Now, consider the reproducing kernels of $\mathcal{L}_{\Pi_T}(f_0)$, denoting them by $K_T^0(\omega, \cdot)$. Since $f(\lambda) \asymp f_0(\lambda)$ as $|\lambda| \rightarrow \infty$, by lemma (2.2.3), $K_T^0(\omega, \cdot)$ belong to $\mathcal{L}_{\Pi_T}(f)$ as well for all $\omega \in \mathbb{R}^d$. Thus, we can expand it using the basis ψ_k , and get $K_T^0(\omega, \lambda) = \sum_k \langle K_T^0(\omega, \cdot), \psi_k \rangle_f \psi_k(\lambda)$, and then by Cauchy-Schwarz and lemma 2.2.3, we get

$$\begin{aligned} |K_T^0(\omega, \lambda)|^2 &\leq \|K_T^0(\omega, \cdot)\|_f^2 \sum_k |\psi_k(\lambda)|^2 \\ &\leq c \|K_T^0(\omega, \cdot)\|_{f_0}^2 \sum_k |\psi_k(\lambda)|^2 \\ &= c K_T^0(\omega, \omega) \sum_k |\psi_k(\lambda)|^2, \end{aligned}$$

which makes the proof complete. □

Theorem 2.3.5 in combination with the lemma 2.3.1 will lead to an appealing result. If the relative difference between two spectral densities is square integrable at infinity, then the corresponding GRFs with stationary increments will be locally equivalent. We finish this

section by proving this fact.

Theorem 2.3.6. *Suppose that the spectral measures F_0 and F_1 have positive densities f_0 and f_1 with respect to the Lebesgue measure, with f_0 satisfying (C1) and (C2) for some entire function ϕ_0 on \mathbb{C}^d of finite exponential type. If there exists a finite constant $k > 0$ such that*

$$\int_{|\lambda|>k} \left(\frac{f_1(\lambda) - f_0(\lambda)}{f_0(\lambda)} \right)^2 d\lambda < \infty, \quad (2.3.9)$$

then GRFs with stationary increments having spectral measures F_0 and F_1 are locally equivalent.

Proof. Combine Lemma 2.3.1, and Theorem 2.3.5. The only thing we need to additionally prove is that $1 \notin \sigma(\psi)$. In spirit of Theorem 2.3.4, we can assume that $f_0 = f_1$ on $|\lambda| \leq k$. Now, take an arbitrary element $\phi \in \mathcal{L}_D(f_0)$, and observe that using the multidimensional Paley-Wiener Theorem ([36], Theorem 3.4.2. p. 171), we get $\phi\phi_0$ is the inverse Fourier transform of a squared integrable function g with bounded support, B in \mathbb{R}^d . This implies that

$$\begin{aligned} |\phi(\lambda)\phi_0(\lambda)|^2 &= \left| \int_B e^{-i\langle \lambda, \gamma \rangle} g(\gamma) d\gamma \right|^2 \\ &\leq \int_B \left| e^{-i\langle \lambda, \gamma \rangle} \right|^2 g^2(\lambda) d\lambda \\ &= \int_B g^2(\lambda) d\lambda \\ &< \infty \end{aligned}$$

for all $\lambda \in \mathbb{R}^d$. This means $\phi\phi_0$ is bounded on \mathbb{R}^d . This fact together with 2.3.9 imply that

$\phi \frac{f_0 - f_1}{f_0} \in L^2(f_0)$. Now, observe that

$$\begin{aligned}
(\psi\phi)(\lambda) &= \int_{\mathbb{R}^d} \phi(\omega) \psi(\lambda, \omega) F_0(d\omega) \\
&= \int_{\mathbb{R}^d} \phi(\omega) \left(\int_{|\gamma|>k} K_T^0(\lambda, \gamma) \overline{K_T^0(\omega, \gamma)} \left(\frac{f_0(\gamma) - f_1(\gamma)}{f_0(\gamma)} \right) F_0(d\gamma) \right) F_0(d\omega) \\
&= \frac{\int_{|\gamma|>k} \int_{\mathbb{R}^d} \overline{\phi(\omega)} \overline{K_T^0(\lambda, \gamma)} K_T^0(\gamma, \omega) \left(\frac{f_0(\gamma) - f_1(\gamma)}{f_0(\gamma)} \right) F_0(d\omega) F_0(d\gamma)}{\int_{|\gamma|>k} \overline{\phi(\gamma)} \left(\frac{f_0(\gamma) - f_1(\gamma)}{f_0(\gamma)} \right) \overline{K_T^0(\lambda, \gamma)} F_0(d\gamma)} \\
&= \frac{\int_{|\gamma|>k} \overline{\phi(\gamma)} \left(\frac{f_0(\gamma) - f_1(\gamma)}{f_0(\gamma)} \right) \overline{K_T^0(\lambda, \gamma)} F_0(d\gamma)}{\int_{|\gamma|>k} \overline{\phi(\gamma)} \left(\frac{f_0(\gamma) - f_1(\gamma)}{f_0(\gamma)} \right) \overline{K_T^0(\lambda, \gamma)} F_0(d\gamma)} \\
&= \pi_{\mathcal{L}_{\Pi_T}(F_0)} \left(\frac{\overline{\phi} f_0 - f_1}{f_0} 1_{\{|\gamma|>k\}} \right) (\lambda).
\end{aligned}$$

Now, similar to the proof of the Theorem 2.3.4, if $\psi\phi = \phi$, we get that

$\|\phi\|_{F_0} \leq \left\| \frac{\overline{\phi} f_0 - f_1}{f_0} 1_{\{|\gamma|>k\}} \right\|_{F_0}$. Letting $k \rightarrow \infty$, by the Dominated Convergence Theorem, we get $\|\phi\|_{F_0} = 0$. This implies that $\phi = 0$. Thus, 1 cannot be in the spectrum of ψ . The proof is complete. \square

2.4 Application

In this section, we apply the results in section 2.3 to some anisotropic GRFs with stationary increments. In particular, we consider GRFs with stationary increments and spectral density of the form

$$f(\lambda) = \frac{1}{\left(\sum_{j=1}^d |\lambda_j|^{\beta_j} \right)^\gamma}, \quad (2.4.1)$$

where $\lambda = (\lambda_1, \dots, \lambda_d) \in \mathbb{R}^d \setminus \{0\}$, $\beta_j > 0$ for all $j = 1, \dots, d$, and $\gamma > \sum_{j=1}^d \frac{1}{\beta_j}$. The latter condition guaranties the integrability condition in 2.2.3 for spectral measures (See Proposition 2.1 in [47]). Fractal and smoothness properties of this family of spatio-temporal

models are discussed in [47]. In order to apply the main results to this model, we need to verify the assumptions needed. First of all, condition (C1) is obviously satisfied for spectral densities of the form 2.4.1. Next lemma provides the proof that these spectral densities also satisfy (C2).

Lemma 2.4.1. *Spectral density functions of the form 2.4.1 satisfy condition (C2).*

Proof. First of all, it is obvious that

$$\frac{1}{\left(\sum_{j=1}^d |\lambda_j|^{\beta_j}\right)^\gamma} \asymp \frac{1}{\left(1 + \sum_{j=1}^d |\lambda_j|^{\beta_j}\right)^\gamma},$$

as $|\lambda| \rightarrow \infty$. Therefore, it suffices to prove the lemma for functions of the form on the right hand side. Now, similar to the construction made in the proof of lemma 2.3 in [29], we can find a function $\phi \in L^2(B)$ for some bounded subset $B \subseteq \mathbb{R}^d$ such that

$$\frac{1}{\left(1 + \sum_{j=1}^d |\lambda_j|^{\beta_j}\right)^\gamma} \asymp \left|\widehat{\phi}(\lambda)\right|^2,$$

as $|\lambda| \rightarrow \infty$, which $\widehat{\phi}$ is the Fourier transform of ϕ . By Paley-Wiener Theorem ([36], Theorem 3.4.2. p. 171), $\widehat{\phi}$ is actually the restriction on \mathbb{R}^d of an entire function on \mathbb{C}^d with finite exponential type. Therefore, the proof is complete. \square

In order to apply the main results, first we consider a similar family of spectral densities of the form

$$f(\lambda) = \frac{1}{\left(\sum_{j=1}^d |\lambda_j|^{H_j}\right)^{Q+2}}, \tag{2.4.2}$$

where $\lambda = (\lambda_1, \dots, \lambda_d) \in \mathbb{R}^d \setminus \{0\}$, $0 < H_j < 1$ for all $j = 1, \dots, d$, and $Q = \sum_{j=1}^d \frac{1}{H_j}$ (See the connection between this family and the ones of the form 2.4.1 in remark 2.2 in [47]).

Theorem 2.4.1. *Suppose f_0 and f_1 are spectral densities of the form 2.4.2 with parameters H_j^0 and H_j^1 , respectively. Then, GRFs with stationary increments and spectral densities f_0 and f_1 are locally equivalent if and only if $H_j^0 = H_j^1$ for all $j = 1, \dots, d$.*

Proof. Suppose for some $k \in \{1, \dots, d\}$, $H_k^0 < H_k^1$. By Lemma 3.2 in [47], there exist $c_1, c_2 > 0$, such that for all $t \in \mathbb{R}^d$

$$c_1 \sum_{j=1}^d |t|^{2H_j^i} \leq \|e_t\|_{f_i}^2 \leq c_2 \sum_{j=1}^d |t|^{2H_j^i}, \quad (2.4.3)$$

for $i = 0, 1$. If we simply choose $t \in \mathbb{R}^d$ with $t_k = l$, and $t_j = 0$ for $j \neq k$, we get

$$\frac{\|e_t\|_{f_1}^2}{\|e_t\|_{f_0}^2} \leq \frac{c_2}{c_1} l^{2(H_k^1 - H_k^0)} \rightarrow 0 \quad \text{as } l \rightarrow 0.$$

This violates the necessary condition for equivalence of Gaussian measures in Theorem 2.3.1. Therefore, the proof is complete. □

Now, we go back to the more flexible spectral densities for space-time models with stationary increments given in 2.4.1. Next theorem proves that under certain conditions, the mixture of spectral densities of the form 2.4.1 will be equivalent to the one with the lowest decay rate at infinity. Similar results of this type for Gaussian processes can be seen in [44] and [7].

Theorem 2.4.2. *Suppose X and Y are two independent centered GRFs with stationary increments with spectral densities of the form 2.4.1 with parameters (β_j, γ) and (β'_j, γ') ,*

respectively. Then, if

$$\gamma' > \frac{1}{2} \sum_{j=1}^d \left(\frac{1}{\beta'_j} \right) + \max_{1 \leq j \leq d} \left\{ \frac{\beta_j}{\beta'_j} \right\} \gamma, \quad (2.4.4)$$

then X and $X + Y$ will be locally equivalent.

Proof. Using theorem 2.3.6, we only need to show

$$\int_{|\lambda| > 1} \frac{\left(\sum_{j=1}^d |\lambda_j|^{\beta_j} \right)^{2\gamma}}{\left(\sum_{j=1}^d |\lambda_j|^{\beta'_j} \right)^{2\gamma'}} d\lambda < \infty. \quad (2.4.5)$$

By using the inequality $(a + b)^p \leq 2^p(a^p + b^p)$, we can break the integral in 2.4.5 into d integrals. Thus, it's enough to show for each fixed $j = 1, \dots, d$

$$I_j := \int_{|\lambda| > 1} \frac{|\lambda_j|^{2\beta_j\gamma}}{\left(\sum_{k=1}^d |\lambda_k|^{\beta'_k} \right)^{2\gamma'}} d\lambda < \infty. \quad (2.4.6)$$

Since $|\lambda| > 1$, this implies that $|\lambda_k| > 1/\sqrt{d}$ for some $k \in \{1, \dots, d\}$. We need to make two cases. Case I is when $k = j$, and case II is when $k \neq j$. In both cases, we use the following fact that given positive constants β and γ , and nonnegative constant b , there exists a finite positive constant c such that for all $a > 0$

$$\begin{aligned} \int_0^\infty \frac{x^b}{(a + x^\beta)^\gamma} dx &= a^{-(\gamma - \frac{1}{\beta} - \frac{b}{\beta})} \int_0^\infty \frac{y^b}{(1 + y^\beta)^\gamma} dy \\ &= \begin{cases} c a^{-(\gamma - \frac{1}{\beta} - \frac{b}{\beta})} & \text{if } \beta\gamma - b > 1, \\ \infty & \text{if } \beta\gamma - b \leq 1. \end{cases} \end{aligned} \quad (2.4.7)$$

First, let's consider case I. By, applying d times the formula 2.4.7,

$$\begin{aligned}
I_j &\leq \int_{\frac{1}{\sqrt{d}}}^{\infty} |\lambda_j|^{2\beta_j\gamma} \underbrace{\int_0^{\infty} \dots \int_0^{\infty}}_{d-1} \frac{d\lambda}{\left(\sum_{r=1}^d |\lambda_r|^{\beta'_r}\right)^{2\gamma'}} \\
&\leq c' \int_{\frac{1}{\sqrt{d}}}^{\infty} \frac{|\lambda_j|^{2\beta_j\gamma}}{\left(|\lambda_j|^{\beta'_j}\right)^{2\gamma' - \sum_{r \neq j} \frac{1}{\beta'_r}}} d\lambda_j \\
&< \infty
\end{aligned}$$

since $\beta'_j \left(2\gamma' - \sum_{r \neq j} \frac{1}{\beta'_r}\right) > 2\beta_j\gamma + 1$ due to condition 2.4.4.

Next, we consider case II, where $k \neq j$. Similar to case I, we use formula 2.4.7 iteratively, but we do the integration in different order. Denote integration in λ_i for $i \neq j, k$ by $d\lambda_{\setminus j, k}$, and observe

$$\begin{aligned}
I_j &\leq \int_{\frac{1}{\sqrt{d}}}^{\infty} d\lambda_k \underbrace{\int_0^{\infty} \dots \int_0^{\infty}}_{d-1} \frac{|\lambda_j|^{2\beta_j\gamma}}{\left(\sum_{r=1}^d |\lambda_r|^{\beta'_r}\right)^{2\gamma'}} d\lambda_j d\lambda_{\setminus j, k} \\
&\leq c \int_{\frac{1}{\sqrt{d}}}^{\infty} d\lambda_k \underbrace{\int_0^{\infty} \dots \int_0^{\infty}}_{d-2} \frac{1}{\left(\sum_{r=1}^d |\lambda_r|^{\beta'_r}\right)^{2\gamma' - \frac{1}{\beta'_j} - \frac{2\beta_j\gamma}{\beta'_j}}} d\lambda_{\setminus j, k} \\
&\leq c' \int_{\frac{1}{\sqrt{d}}}^{\infty} \frac{1}{\left(|\lambda_k|^{\beta'_k}\right)^{2\gamma' - \frac{2\beta_j\gamma}{\beta'_j} - \sum_{r \neq k} \frac{1}{\beta'_r}}} d\lambda_k \\
&< \infty,
\end{aligned}$$

where the second and the third inequality follow since $2\gamma'\beta'_j > 2\beta_j\gamma + 1$ and $\beta'_k \left(2\gamma' - \frac{2\beta_j\gamma}{\beta'_j} - \sum_{r \neq k} \frac{1}{\beta'_r}\right) > 1$, respectively using the assumption 2.4.4. Therefore, the proof is complete. \square

Next, we will consider similar situation as in theorem 2.4.2, but this time we put discrete

spectral measure mixed with the ones of the form 2.4.1. For that purpose, consider discrete spectral measure of the form

$$F(\{-\gamma^n\}) = F(\{\gamma^n\}) = \alpha_n, \quad (2.4.8)$$

where $\gamma^n \in \mathbb{R}^d$, $\alpha_n \geq 0$, for $n \geq 1$, and $\sum_{n=1}^{\infty} \frac{|\gamma^n|^2}{1+|\gamma^n|^2} \alpha_n < \infty$. If $\{\gamma^n, n = 1, 2, \dots\}$ is a bounded subset of \mathbb{R}^d , then in view of theorem 2.3.4, this spectral measure won't affect the equivalence of Gaussian measures. Therefore, we consider here only the case where $|\gamma^n| \rightarrow \infty$ as $n \rightarrow \infty$.

Theorem 2.4.3. *Let X and Y be two independent centered GRFs with stationary increments with spectral measures F_X and F_Y . Suppose F_X has density with respect to Lebesgue measure on \mathbb{R}^d , denoted by f , which satisfies both conditions (C1) and (C2), and F_Y is a discrete measure of the form 2.4.8. Then, if*

$$\sum_{n>N} \frac{\alpha_n}{f(\gamma^n)} < \infty, \quad (2.4.9)$$

for some $N \geq 1$, then X and $X + Y$ are locally equivalent.

Proof. First of all, using theorem 2.3.4, we can assume $\alpha_n = 0$ for $n = 1, \dots, N$, without having any consequences on the equivalence of Gaussian measures. Second, observe that for all $\phi \in \mathcal{L}_{\Pi_T}^e$, $\|\phi\|_{F_X} \leq \|\phi\|_{F_X+F_Y}$, which by remark 2.3.2, is equivalent to condition (i) in theorem 2.3.2. All we need to prove is then to show that the function ψ in theorem 2.3.2 is in $L^2(f \otimes f)$. For that, using Foubini's theorem and the reproducing kernel property we

have

$$\begin{aligned}
\|\psi\|_{f \otimes f}^2 &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \left| \sum_{n>N} \alpha_n K_T(\omega, \gamma^n) \overline{K_T(\lambda, \gamma^n)} \right|^2 f(\omega) f(\lambda) d\omega d\lambda \\
&= \sum_{n,m>N} \alpha_n \alpha_m \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K_T(\omega, \gamma^n) \overline{K_T(\lambda, \gamma^n)} \overline{K_T(\omega, \gamma^m)} K_T(\lambda, \gamma^m) \\
&\quad \times f(\omega) f(\lambda) d\omega d\lambda \\
&= \sum_{n,m>N} \alpha_n \alpha_m |K_T(\gamma^m, \gamma^n)|^2 \\
&\leq \left(\sum_{n>N} \frac{\alpha_n}{f(\gamma^n)} \right)^2 \\
&< \infty
\end{aligned}$$

by the assumption 2.4.9. The proof is complete. \square

Chapter 3

Covariance Tapering for Anisotropic Nonstationary Gaussian Random Fields with Application to Large Scale Spatial Data Sets

Estimating the covariance structure of spatial random processes is an important step in spatial data analysis. Maximum likelihood estimation is a popular method in spatial models based on Gaussian random fields. But calculating the likelihood in large scale data sets is computationally infeasible due to the heavy computation of the precision matrix. One way to mitigate this issue, which is due to Furrer et al. (2006), is to “taper” the covariance matrix. While most of the results in the current literature focus on isotropic tapering for stationary Gaussian processes, there are many cases in application that require modeling of anisotropy and/or nonstationarity. In this chapter, we propose a nonstationary parametric model, in which the underlying Gaussian random field may have different regularities in different directions, thus can be applied to model anisotropy. Using the theory of equivalence of Gaussian measures under nonstationary assumption, strong consistency of the tapered likelihood based estimation of the variance component under fixed domain asymptotics are

derived by putting mild conditions on the spectral behavior of the tapering covariance function. The procedure is illustrated with numerical simulation.

3.1 Introduction

Spatial Statistics is nowadays a very active research field in Statistics, and has many applications in Geology, Agricultural Science, Environmental Science, Climate data, etc. [9, 41]. A common problem in this field is the estimation of the covariance structure in spatial models based on Gaussian random fields. Likelihood-based estimators are the most popular method for estimating the covariance parameters. However, in large scale spatial data sets, calculating the likelihood is computationally infeasible due to the heavy calculation of the precision matrix. There are different ways to overcome this issue. The first idea is to set the off-diagonal entries of the covariance matrix to zero, or to keep the first k subdiagonal entries for some integer k and put the rest of them to be zero, which is called banding. In this way, the resulting matrix becomes sparse, and one can use the existing algorithms dealing with sparse matrices to handle the new covariance matrix efficiently. But, the problem with banding is that the final covariance matrix may not be positive definite, which is a huge drawback since all the theoretical covariance matrices must be positive definite. An alternative approach is to multiply the covariance function by a positive definite compactly supported correlation function. By this way, the resulting covariance matrix is again sparse, but still positive definite. This is called covariance tapering [18]. A natural question is that how we can use tapering to construct consistent estimates for the covariance parameters in spatial regression models. Kaufman et al. (2008) proposed two different likelihood-based estimations of the parameters in the Matern covariance function, and proved the strong con-

sistency of the estimates using the results in the equivalence of stationary Gaussian measures [37]. See also [12] for more results in the asymptotics of tapered maximum likelihood estimators. There are two strong assumptions under which the consistency is proved, isotropy, and stationarity of the underlying Gaussian random field. However, there are many cases in application in which the data sets resemble anisotropy and nonstationarity. There have been interests in general (not necessarily in the context of covariance tapering) for dropping these conditions in spatial modellings (See [11], and [1]). In this chapter, we relax both of these conditions simultaneously. For that purpose, in section 3.2, we introduce a class of parametric models for spatial modeling which are anisotropic and non-stationary. In section 3.3, we state the main result of the paper which is deriving strong consistency of the tapered likelihood-based estimates of the variance parameter. In the last section, we demonstrate the procedure proposed here via numerical simulation.

3.2 Preliminary

In this section, we introduce a class of Gaussian random fields which can be used as spatial models. Any Gaussian field model needs a mean structure and a covariance structure to be uniquely defined. However, finding new models for covariance structure might be hard since one can only choose covariance structures from the family of positive definite functions and it is complicated to verify the positive definiteness. An alternative is to use the spectral representations of positive definite functions or variograms. We consider here specifically Gaussian field models with stationary increments. By applying the results in [49], the covariance functions of Gaussian random fields with stationary increments on \mathbb{R}^d can be determined by a symmetric non-negative measures on $\mathbb{R}^d \setminus \{0\}$ satisfying certain integrability

condition (See e.g., [47] for more details). This measure is called spectral measure. If this measure is absolutely continuous with respect to Lebesgue measure on \mathbb{R}^d , we will call its Radon-Nikodym derivative the “spectral density”.

Now, we propose the following spectral density:

$$f(\lambda) = \frac{\sigma^2}{\left(1 + \sum_j |\lambda_j|^{H_j}\right)^{Q+\nu}} \quad (3.2.1)$$

where $\lambda = (\lambda_1, \dots, \lambda_d) \in \mathbb{R}^d$, $Q = \sum_j^d 1/H_j$, and σ^2 , ν and H_j ’s are all positive parameters. σ^2 is the variance component, and H_j ’s are related to the smoothness of the model in different directions. Figures 3.1 and 3.2 are realizations of such a Gaussian field over the two-dimensional grid $[0, 1]^2$ with increments of 0.02 with parameters $H_1 = 1.5$, $H_2 = 3$, and $\nu = 0.4$.

Having different parameters, H_j , $j = 1, \dots, d$, for different components of $\lambda \in \mathbb{R}^d$, allows the underlying random field to have different regularities in different directions, and therefore, this model is able to capture anisotropy appearing in spatial data sets. (See more about fractal and smoothness properties of these models in [47]. See also [43]).

3.3 Main Result

In this section, we state the main result of the article, which is the consistency of the tapered likelihood estimation of the variance parameter σ^2 in the model introduced in the last section under the fixed domain asymptotics. The following theorem, shows how to get equivalent Gaussian measures in tapering set up, which is necessary in proving strong consistency, by putting mild conditions on the spectral behavior of the tapering covariance function.

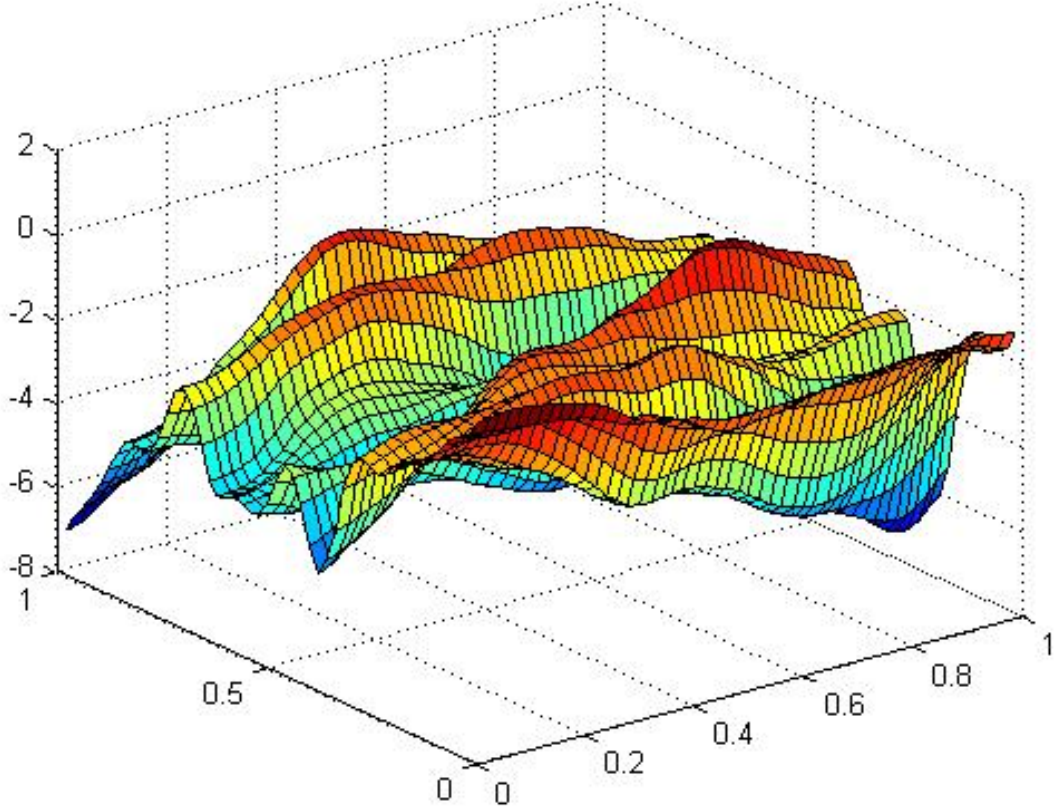


Figure 3.1 3D Simulated Surface on $[0, 1]^2$ with increments 0.02

Theorem 1. Suppose $\{Z(t), t \in T\}$, where $T \subseteq \mathbb{R}^d$ is a mean zero Gaussian random field with stationary increments, and the spectral density of the form 3.2.1 for some positive constants σ^2 , ν , and H_j , $j = 1, \dots, d$. Let's denote the covariance function of the process by $K_0(x)$. If $K_1(x) = K_0(x)K_t(x)$ where $K_t(x)$ is a correlation function, and its Fourier transform exists and satisfies the following condition:

$$f_t(\lambda) \leq \frac{M}{\left(1 + \sum_j |\lambda_j|^{H_j}\right)^{Q+\nu+\epsilon}} \quad (3.3.1)$$

for some $\epsilon > \max\{Q/2, 1/\min\{H_j/2\}\}$, and $M > 0$. Now, if $H_j > 1$, $j = 1, \dots, d$, then the Gaussian measures induced by K_0 and K_1 will be equivalent on the paths of $\{Z(t), t \in T\}$

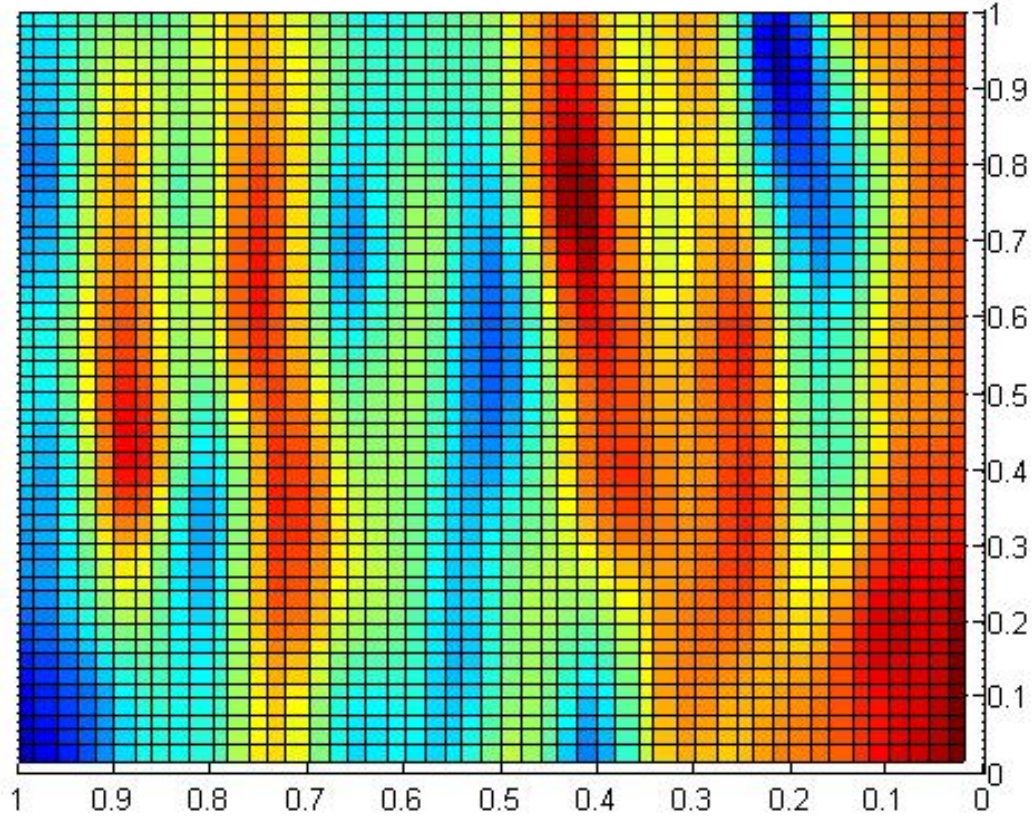


Figure 3.2 2D Projection of the Simulated Surface on $[0, 1]^2$ with increments 0.02

for any bounded subset $T \subset \mathbb{R}^d$.

Remark 1. *Proof of this theorem is similar to that of Theorem 1 in [28]. However, due to the nonstationarity assumption, it relies on the study of the equivalence of Gaussian random fields with stationary increments, which is discussed in details in the second chapter.*

Suppose one observes the random process at locations s_j 's $j = 1, \dots, n$ on some bounded domain in \mathbb{R}^d , and $K_t(x)$ is the tapering covariance function. Fix a tapering parameter $\gamma > 0$. Define the tapering matrix $T(\gamma)$ to be $T(\gamma)_{ij} = K_t(\|s_i - s_j\|; \gamma)$. This means that if $\|s_i - s_j\| > \gamma$, then $T(\gamma)_{ij} = 0$. Denote the true covariance matrix of the process by $\Sigma(\theta)$. Now, the idea of tapering is to use $\Sigma(\theta) \circ T(\gamma)$ as the new covariance matrix in the

likelihood of the process. Here, \circ means the element wise matrix product. Therefore, the Tapered Maximum likelihood estimator (Tapered MLE) of σ^2 will be:

$$\begin{aligned}\widehat{\sigma_t^2} &= \operatorname{argmax}_{\sigma^2} \text{ Tapered likelihood} \\ &= \operatorname{argmax}_{\sigma^2} \left(-\frac{1}{2} \log |\Sigma(\theta) \circ T(\gamma)| - \frac{1}{2} Z' (\Sigma(\theta) \circ T(\gamma))^{-1} Z \right),\end{aligned}$$

where Z is the vector of observed values of the random field at the specified locations. Evaluating the inverse of $\Sigma(\theta) \circ T(\gamma)$ is much faster than $\Sigma(\theta)$ in many cases, especially with large n . This is the benefit of using Tapered MLE. However, this estimator might be biased in general. In [28], another estimator based on the same idea of tapering is presented which is unbiased. We call it here the Adjusted Tapered MLE, and it is defined as follows:

$$\begin{aligned}\widehat{\sigma_{adj}^2} &= \operatorname{argmax}_{\sigma^2} \text{ Adjusted tapered likelihood} \\ &= \operatorname{argmax}_{\sigma^2} \left(-\frac{1}{2} \log |\Sigma(\theta) \circ T(\gamma)| - \frac{1}{2} Z' \left((\Sigma(\theta) \circ T(\gamma))^{-1} \circ T(\gamma) \right) Z \right).\end{aligned}$$

We will see in section 4 that the Adjusted Tapered MLE performs better comparing to the Tapered MLE due to its unbiasedness. Theorem 1 is the key part in proving the consistency of the tapered maximum likelihood estimators for the variance component (σ^2). The outline of the proof is similar to Theorem 2 in [28], and also it relies on the above theorem. We omit the proof here.

3.4 Simulation Results

In this section, we illustrate the methods discussed in previous parts by applying them into some simulated spatial data sets. For this purpose, we simulated 1000 data sets, each consisting of a multivariate Gaussian vector of length 14×14 . The locations are two-dimensional grid over $[0, 1]^2$ with increments 0.07. We added a random noise in each coordinate, uniformly distributed on $[-0.01, 0.01]$. We used the spectral density 3.2.1 to generate the normal vector with $H_1 = 1.5$, $H_2 = 2$, $\nu = 0.5$, and $\sigma^2 = 1$. Figure 3.3 shows a simulated surface using the above parameters over the specified grid.

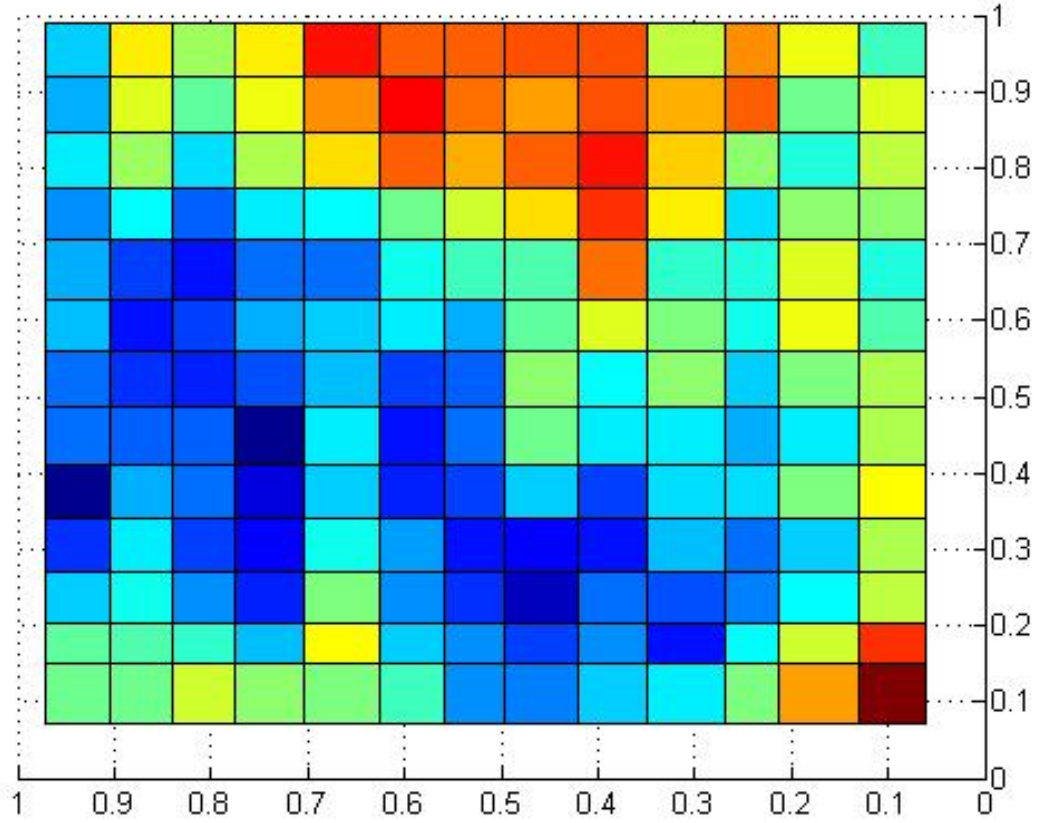


Figure 3.3 Simulated Surface on $[0, 1]^2$ with increments 0.07

We used the same tapering function used in [28] with different tapering parameter $\gamma =$

0.7, 0.4, 0.2, respectively. We applied three different estimators for estimating σ^2 : MLE, Tapered MLE and Adjusted Tapered MLE (See [28] for details.). Table 3.1 shows the results for this procedure for different tapering parameters. The values in the table are the averages of the estimated value over the 1000 repetitions with their standard errors in the bracket. As it was expected, by decreasing the tapering parameter, the Tapered MLE becomes more and more biased. However, the Adjusted Tapered MLE is almost unbiased regardless of the changes in the tapering parameter.

Table 3.1 Results for estimation of σ^2

$\widehat{\sigma^2}$	$\gamma = 0.7$	$\gamma = 0.4$	$\gamma = 0.2$
MLE	0.999(0.10)	0.999(0.10)	0.999(0.10)
Tapered MLE	0.437(0.06)	0.409(0.07)	0.400(0.17)
Adjusted Tapered MLE	1.025(0.53)	1.019(0.47)	1.031(0.61)

Simulations show that comparing to MLE, Adjusted Tapered MLE is a good alternative for estimating the covariance parameters in spatial data analysis, and it is computationally feasible. Further, this method has the potential to be generalized to more complicated models which have nonstationarity and anisotropy.

There are other methods dealing with covariance estimation in large spatial data sets. For example, in [11], nonstationary spatial models are defined through some fixed basis functions, and then weighted least squares method (rather than the MLE approach) is chosen for covariance parameter estimation with the emphasis of finding the best linear unbiased prediction (BLUP).

Chapter 4

Spatial factor models and dimension reducing Gaussian predictive processes for high dimensional remotely sensed data and forest variables

Combining spatially-explicit long-term forest inventory and remotely sensed information from Light Detection and Ranging (LiDAR) datasets through statistical models can be a powerful tool for predicting and mapping above-ground biomass (AGB) at a range of geographic scales. We present and examine a novel modeling approach to improve prediction of AGB using LiDAR data. More specifically, we introduce a multivariate spatial regression model in which both LiDAR data and AGB play as target variables. There are two sources of computational infeasibility in the proposed model which need to be addressed. First, the number of target variables are too high since we put both LiDAR and AGB simultaneously (In our real dataset, we have more than 100 LiDAR data points at different heights for each location). Second, the number of locations are also too high (more than 451 locations in our

data). In other words, dimension reduction is needed in two aspects: (i) the length of the vector of outcomes, and (ii) the very large number of spatial locations. We use spatial factor modeling for the first aspect, and low-rank spatial processes for the second one. Latent variable (factor) models are usually used to address the large vector of outcomes, and low-rank spatial processes offer a rich and flexible modeling option for dealing with a large number of locations.

We introduce the model in the first section, and it is followed by a brief discussion on the identifiability issue of the model in section 4.2. Preparation for the Bayesian hierarchical framework for the estimation and prediction is done in sections 4.3, 4.5 and 4.6. In section 4.4, we explain in detail the low-rank spatial processes chosen here to reduce the computational time. We finish this chapter by a simulation study and real data analysis.

4.1 Model

The following hierarchical factor model is considered

$$\begin{pmatrix} z(s) \\ y(s) \end{pmatrix} = \begin{pmatrix} x(s)' \beta \\ x_y(s)' \eta \end{pmatrix} + \begin{pmatrix} A & \mathbf{0} \\ \alpha'_f & \alpha_\nu \end{pmatrix} \begin{pmatrix} f(s) \\ \nu(s) \end{pmatrix} + \begin{pmatrix} \epsilon_z(s) \\ \epsilon_y(s) \end{pmatrix} \quad (4.1.1)$$

where $z(s)$ is $m \times 1$ vector, $y(s) \in \mathbb{R}$, $x(s)$ is $t \times m$ set of covariates for the high-dimensional signal $z(s)$, $x_y(s)$ is $r \times 1$ vector of covariates for the target variable $y(s)$, β and η are $t \times 1$ and $r \times 1$ vectors of parameters, respectively. A is $m \times q$ matrix of factor loadings with elements satisfying $a_{1j} > 0 \ \forall j = 1, \dots, q$. $\alpha'_f = (\alpha_{f,1}, \alpha_{f,2}, \dots, \alpha_{f,q})$ is a $1 \times q$ vector of coefficients, and $\alpha_\nu > 0$. Also, $f(s) = (f_1(s), f_2(s), \dots, f_q(s))'$ where each $f_j(s)$ follows an independent

spatial $GP(0, C_j(s_1, s_2; \theta_j))$ with unit variance, and $\nu(s)$ follows a $GP(0, C_j(., \theta_\nu))$ with unit variance and independent from all $f'_j s$. $\epsilon_z(s)$ and $\epsilon_y(s)$ are $N(0, \Sigma_{\epsilon_z} := \text{diag}(\tau_1^2, \dots, \tau_m^2))$ and $N(0, \tau_y^2)$ where τ^2 and τ_y^2 are representing the usual geostatistical nuggets for predicting $z(s)$ and $y(s)$, respectively.

In this model, $z(s)$ represents the LiDAR data, $y(s)$ is the biomass, and $f_j(s)$'s are the factors with the total number of q factors.

4.2 Identifiability Issue

Define $\Lambda := \begin{pmatrix} A & \mathbf{0} \\ \alpha'_f & \alpha_\nu \end{pmatrix}$, and $\omega(s) := \begin{pmatrix} f(s) \\ \nu(s) \end{pmatrix}$. The indentifiability issue in model 4.1.1 relates to the existence of an $(q+1) \times (q+1)$ orthogonal matrix P (i.e. $P'P = PP' = I_q$) such that $P'w(s)$ has the same covariance structure as $w(s)$. Following the argument in [35], the only possible matrices in this case are the ones whose entries are in $\{-1, 0, 1\}$ with exactly one nonzero element in each row and each column. In other words, changing the sign of some of the components in $w(s)$, and permuting some of the elements in $w(s)$ are the only two groups of transformations which can lead to another Gaussian random field with the same distribution as $w(s)$. Both cases need to be handled in order to have an identifiable model. The first group can be prevented by putting constrain on the sign of the parameters involved in Λ . To be precise, at each column in Λ , the sign of one of the elements needs to be fixed. Here, we have positive elements in the first row of A with $\alpha_\nu > 0$, which together will solve the first part. Switching the orders of the elements in $w(s)$ can be prevented by putting constrain on the priors related to the covarince parameters in $f(s)$ and $\nu(s)$, which will be addressed in the next section.

4.3 Prior Specification

We put a multivariate normal prior on the mean parameters β and η , with mean μ_β and μ_η , and variance Σ_β and Σ_η , respectively (we will consider a flat prior, i.e. $\Sigma_\beta^{-1} = \Sigma_\eta^{-1} = 0$, and also $\mu_\beta = \mu_\eta = 0$). Both nugget effects τ^2 and τ_y^2 are assigned Inverse Gamma distribution $IG(a, b)$ with choices of $a = 2$ for having an infinite variance and $b = 5$. All the parameters in the matrix Λ are considered to be independent normal random variables or truncated normal random variables depending on whether their sign are determined beforehand. More precisely, $a_{ij} \sim N(0, c_a^2)$ for $i \neq 1$, $\alpha_{f,i} \sim N(0, c_\alpha^2)$ for $i = 1, \dots, q$, and $a_{1i} \sim N(0, c_a^2)I(a_{1i} > 0)$ for $i = 1, \dots, q$, and $\alpha_\nu \sim N(0, c_\alpha^2)I(\alpha_\nu > 0)$. Correlation functions related to $f'_j s$ and ν are considered to be exponential correlation function, i.e. $\text{cov}(f_j(s), f_j(t)) := \rho_j(s, t; \phi_j) = \exp(-\phi_j \|s - t\|)$ for $j = 1, \dots, q$, and also $\text{cov}(\nu(s), \nu(t)) := \rho_\nu(s, t; \phi_\nu) = \exp(-\phi_\nu \|s - t\|)$. Because of identifiability issue discussed in section 4.2, we construct a joint distribution for the $\phi'_j s$ to ensure ordering. In particular, we set

$$\pi(\phi) = \pi(\phi_1)\pi(\phi_2 | \phi_1)\dots\pi(\phi_q | \phi_{q-1}, \dots, \phi_1) \quad (4.3.1)$$

where $\pi(\phi_1)$ is a uniform density with support (ϕ_l, ϕ_u) with $\phi_l = -\log(0.05)/d_{max}$ and $\phi_u = -\log(0.01)/d_{min}$, where d_{min} and d_{max} are the minimum and maximum distance across all the locations, and for $j = 2, 3, \dots, q$

$$\pi(\phi_j | \phi_{j-1}, \dots, \phi_1) \propto \exp\left(-\frac{c_j}{\phi_j - \phi_{j-1}}\right) I(\phi_{j-1} < \phi_j < \phi_u). \quad (4.3.2)$$

Based on the discussion in [35], we fix $c_j = 2j$ as a reasonable choice. Also, $\phi_\nu \sim \text{uniform}(\phi_l, \phi_u)$.

4.4 Approximation using Gaussian predictive processes

Due to computational burden, we need to replace the random effects $f(s)$ and $\nu(s)$ by a low-rank approximation of them. One way to address this issue is to project the mentioned Gaussian random fields into the space generated by them at certain locations, called “knots” (We refer to [4] and [16] for the complete discussion on Gaussian predictive processes). Suppose $S^* = \{s_1^*, \dots, s_{n^*}^*\}$ is the set of knots considered for approximating both fields $f(s)$ and $\nu(s)$. We replace $f(s)$ and $\nu(s)$ with $f^*(s)$ and $\nu^*(s)$ with the following details. Define the Gaussian field

$$\tilde{f}_k^*(s) \sim N\left(f_k^*(s), \sigma_k^{*2}(s)\right) \quad (4.4.1)$$

where

$$f_k^*(s) := d_k(s; \phi_k)' D_k^*(\phi_k)^{-1} f_k(S^*), \quad \sigma_k^{*2}(s) := 1 - d_k(s; \phi_k)' D_k^*(\phi_k)^{-1} d_k(s; \phi_k), \quad (4.4.2)$$

and

$$\tilde{\nu}^*(s) \sim N\left(\nu^*(s), \sigma_\nu^{*2}(s)\right) \quad (4.4.3)$$

where

$$\nu^*(s) := d_\nu(s; \phi_\nu)' D_\nu^*(\phi_\nu)^{-1} \nu(S^*), \quad \sigma_\nu^{*2}(s) := 1 - d_\nu(s; \phi_\nu)' D_\nu^*(\phi_\nu)^{-1} d_\nu(s; \phi_\nu), \quad (4.4.4)$$

and $d_k(s; \phi_k)$ is an $n^* \times 1$ vector whose i th element is $\rho_k(s, s_i^*; \phi_k)$, $D_k^*(\phi_k)$ is an $n^* \times n^*$ covariance matrix whose (i, j) th entry is $\rho_k(s_i^*, s_j^*; \phi_k)$, $f_k(S^*)$ is an $n^* \times 1$ vector with elements as the evaluation of the field f_k at the knot points. The case for $\nu(s)$ is exactly similar to the fields f_k^* s by just simply replacing the index k by ν . After this replacement,

model 4.1.1 can be written as

$$\begin{pmatrix} z(s) \\ y(s) \end{pmatrix} = \begin{pmatrix} x(s)' \beta \\ x_y(s)' \eta \end{pmatrix} + \begin{pmatrix} A & \mathbf{0} \\ \alpha'_f & \alpha_\nu \end{pmatrix} \begin{pmatrix} f^*(s) \\ \nu^*(s) \end{pmatrix} + \begin{pmatrix} \epsilon_z^*(s) \\ \epsilon_y^*(s) \end{pmatrix} \quad (4.4.5)$$

where $f^*(s) = (f_1^*(s), f_2^*(s), \dots, f_q^*(s))'$, $\epsilon_z^*(s) \sim N(0, \Sigma_{\epsilon_z^*(s)} := A \Sigma_{\tilde{f}^*} A' + \Sigma_{\epsilon_z})$, and $\epsilon_y^*(s) \sim N(0, \sigma_{\epsilon_y^*(s)}^2 := \alpha'_f \Sigma_{\tilde{f}^*} \alpha_f + \alpha_\nu^2 \sigma_\nu^{*2}(s) + \tau_y^2)$ with $\Sigma_{\tilde{f}^*}$ to be a $q \times q$ diagonal matrix with diagonal elements $\sigma_k^{*2}(s)$, $k = 1, \dots, q$.

4.5 Data equation

Let

$$\begin{aligned} Y(s) &= \begin{pmatrix} z(s) \\ y(s) \end{pmatrix}, \quad X'(s) = \begin{pmatrix} x(s)' & \mathbf{0} \\ \mathbf{0} & x_y(s)' \end{pmatrix}, \quad b = \begin{pmatrix} \beta \\ \eta \end{pmatrix}, \\ \omega^*(s) &= \begin{pmatrix} f^*(s) \\ \nu^*(s) \end{pmatrix}, \quad \epsilon(s) = \begin{pmatrix} \epsilon_z^*(s) \\ \epsilon_y^*(s) \end{pmatrix}. \end{aligned} \quad (4.5.1)$$

Now, the model can be written as

$$Y(s_i) = X(s_i)b + \Lambda \omega^*(s_i) + \epsilon(s_i) \quad (4.5.2)$$

for $i = 1, \dots, n$. Also, the matrix form of the model is

$$Y = Xb + \bar{\Lambda} \omega^* + \epsilon \quad (4.5.3)$$

where

$$\begin{aligned}
Y &= \begin{pmatrix} Y(s_1) \\ \vdots \\ Y(s_n) \end{pmatrix}, X = \begin{pmatrix} X(s_1) \\ \vdots \\ X(s_n) \end{pmatrix}, \bar{\Lambda} = \begin{pmatrix} \Lambda & \mathbf{0} \\ & \ddots \\ \mathbf{0} & \Lambda \end{pmatrix}, \\
\epsilon &= \begin{pmatrix} \epsilon(s_1) \\ \vdots \\ \epsilon(s_n) \end{pmatrix}, \omega^* = \begin{pmatrix} \omega^*(s_1) \\ \vdots \\ \omega^*(s_n) \end{pmatrix}.
\end{aligned} \tag{4.5.4}$$

Here, $\omega^* \sim N(0, \Sigma_{\omega^*})$, and $\epsilon \sim N(0, \Sigma_{\epsilon})$,

$$\Sigma_{\epsilon} = \begin{pmatrix} \Sigma_{\epsilon(s_1)} & \mathbf{0} \\ & \ddots \\ \mathbf{0} & \Sigma_{\epsilon(s_n)} \end{pmatrix}, \Sigma_{\omega^*} = \begin{pmatrix} \Sigma_{\omega^*(s_1)} & \mathbf{0} \\ & \ddots \\ \mathbf{0} & \Sigma_{\omega^*(s_n)} \end{pmatrix} \tag{4.5.5}$$

with

$$\Sigma_{\epsilon(s_i)} = \begin{pmatrix} \Sigma_{\epsilon_z^*(s_i)} & \mathbf{0} \\ \mathbf{0} & \sigma_{\epsilon_y^*(s_i)}^2 \end{pmatrix}, \Sigma_{\omega^*(s_i)} = \begin{pmatrix} \Sigma_{f^*(s_i)} & \mathbf{0} \\ \mathbf{0} & \sigma_{\nu^*(s_i)}^2 \end{pmatrix} \tag{4.5.6}$$

where $\Sigma_{f^*(s)}$ is a diagonal matrix with diagonal elements $d_k(s; \phi_k)' D_k^*(\phi_k)^{-1} d_k(s; \phi_k)$, and $\sigma_{\nu^*(s)}^2 = d_{\nu}(s; \phi_{\nu})' D_{\nu}^*(\phi_{\nu})^{-1} d_{\nu}(s; \phi_{\nu})$.

4.6 Computations

The first step for computation purposes is to find the full conditionals of the parameters involved in the model.

4.6.1 Updating the mean part b

The posterior of β follows $N_t(\tilde{\mu}_\beta, \tilde{\Sigma}_\beta)$ with:

$$\tilde{\Sigma}_\beta = \left(\Sigma_\beta^{-1} + \sum_{i=1}^n x(s_i) \Sigma^\star(s_i) x(s_i)' \right)^{-1} \quad (4.6.1)$$

and

$$\tilde{\mu}_\beta = \tilde{\Sigma}_\beta \left(\Sigma_\beta^{-1} \mu_\beta + \sum_{i=1}^n x(s_i) \Sigma^\star(s_i) (z(s_i) - Af(s_i)) \right), \quad (4.6.2)$$

where $\Sigma^\star(s)$ is an $(m) \times (m)$ diagonal matrix with j -th diagonal element τ_j^{-2} for $j = 1, \dots, m$.

The posterior distribution of η follows a $N_r(\tilde{\mu}_\eta, \tilde{\Sigma}_\eta)$ with:

$$\tilde{\Sigma}_\eta = \left(\Sigma_\eta^{-1} + \sum_{i=1}^n x_y(s_i) x_y(s_i)' / \tau_\nu^2 \right)^{-1} \quad (4.6.3)$$

and

$$\tilde{\mu}_\eta = \tilde{\Sigma}_\eta \left(\Sigma_\eta^{-1} \mu_\eta + \sum_{i=1}^n x_y(s_i) \tau_\nu^{-2} (y(s_i) - \lambda_\nu \omega(s_i)) \right), \quad (4.6.4)$$

where λ_ν is the last row of the matrix Λ .

4.6.2 Updating ω^\star

ω^\star consists of two main components: f^\star and ν^\star , and they will be updated similarly. For

$k = 1, \dots, q$, f_k^\star will follow a $N_{n^\star}(\tilde{\mu}_{f_k^\star}, \tilde{\Sigma}_{f_k^\star})$ with:

$$\tilde{\Sigma}_{f_k^\star} = D_k^\star(\phi_k) \left(D_k^\star(\phi_k) + D_k(\phi_k)' \Sigma_{f_k}(\phi_k)^{-1} D_k(\phi_k) \right)^{-1} D_k^\star(\phi_k) \quad (4.6.5)$$

and

$$\tilde{\mu}_{f_k^\star} = \tilde{\Sigma}_{f_k^\star} D_k^\star(\phi_k)^{-1} D_k(\phi_k)' \Sigma_{f_k}(\phi_k)^{-1} f_k, \quad (4.6.6)$$

where $\Sigma_{f_k}(\phi_k)$ is an $n^\star \times n^\star$ diagonal matrix whose j -th diagonal element is $\sigma_k^{\star 2}(s_j^\star)$, and $D_k(\phi_k)$ is the covariance matrix between the observation locations and the knots for the k -th factor f_k .

4.6.3 Updating the loading factors Λ

For $j = 2, \dots, m$, the full conditionals for the j -th row of Λ follows $N_q(m_j, K_j)$, where

$$K_j^{-1} = c_a^{-2} I_q + \tau_j^{-2} (F' F) \quad (4.6.7)$$

and

$$m_j = \tau_j^{-2} K_j F' (z_j - \mu^j), \quad (4.6.8)$$

where $F = (f(s_1), \dots, f(s_n))'$, and μ^j is the mean vector of z_j . The first row of Λ is normal distribution truncated at 0 with mean m_1 and variance K_1 . The last row of Λ follows $N_{q+1}(m_\nu, K_\nu)$ with:

$$K_\nu^{-1} = c_a^{-2} I_{q+1} + \tau_\nu^{-2} (W' W) \quad (4.6.9)$$

and

$$m_\nu = \tau_\nu^{-2} K_\nu W' (y - \mu^y), \quad (4.6.10)$$

where $W = (\omega(s_1), \dots, \omega(s_n))'$, and μ^y is the mean vector of y_j . The last element of this row is truncated at 0.

4.6.4 Updating the nuggets τ

For $j = 1, \dots, m$, the full conditional distribution of τ_j^2 will be Inverse Gamma with shape parameter $a + n/2$, and scale parameter $b + \frac{1}{2}(z_j - \mu^j - F\lambda_j)'(z_j - \mu^j - F\lambda_j)$ where λ_j' is the j -th row of the matrix Λ . τ_ν^2 will be also Inverse Gamma with the same shape parameter, but with scale parameter $b + \frac{1}{2}(y - \mu^y - W\lambda_\nu)'(y - \mu^y - W\lambda_\nu)$ where λ_ν' is the last row of the matrix Λ .

4.7 Simulations

We illustrate the performance of the model using a simulation study. Here, we generated data over 256 locations across a $[0, 50] \times [0, 50]$ square using the following parameters:

$$b = \begin{pmatrix} 5 \\ 10 \\ 12 \end{pmatrix}, \Lambda = \begin{pmatrix} 2 & 1 & 0 \\ -1 & 2 & 0 \\ 1 & -1 & 2 \end{pmatrix}, \phi = \begin{pmatrix} 0.1 \\ 0.4 \\ 0.12 \end{pmatrix}, \Sigma_\epsilon = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

In this simulation, $t = r = 1$ and $m = q = 2$. We used Gaussian predictive processes to reduce the computation time by selecting 81 knots. Table 4.1 shows the parameter estimates together with their credible interval. It can be observed from this table that all the true parameter values are within their 95% credible interval calculated by the posterior distributions.

Table 4.1 Parameter estimates with their 95% credible intervals

parameter	true value	estimate	credible interval
β_1	5	6.278622	(4.213607,8.506537)
β_2	10	9.152451	(8.141619,10.197632)
η	12	12.13515	(10.67103,13.93106)
τ_1^2	2	1.7069400	(0.9133798,2.6578697)
τ_2^2	5	2.6038261	(0.8570427,6.4833734)
τ_ν^2	1	1.4963317	(0.7562288,2.4505370)
$\Lambda_{1,1}$	2	2.235054	(1.418790,2.978264)
$\Lambda_{1,2}$	1	0.6389909	(0.2377660,1.4958831)
$\Lambda_{2,1}$	-1	-1.0007031	(-1.8632592,-0.3081513)
$\Lambda_{2,2}$	2	2.371756	(1.055427,2.886742)
$\Lambda_{3,1}$	1	0.6106929	(-0.1468869,2.2537335)
$\Lambda_{3,2}$	-1	-0.8785957	(-1.3010258,-0.5326576)
$\Lambda_{3,3}$	2	2.430128	(1.571113,3.469093)
ϕ_1	0.1	0.06601322	(0.04838031,0.12743472)
ϕ_2	0.4	0.6352006	(0.3916554,0.8975183)
ϕ_ν	0.12	0.08413867	(0.04986396,0.19860435)

4.8 Real Data Analysis

In this section, the proposed model is assessed using LiDAR data acquired from NASA Goddards LiDAR, Hyper-spectral & Thermal imager and field inventory data from the Penobscot Experimental Forest in Bradley, Maine (NASA Goddards LiDAR, Hyper-spectral, and Thermal (G-LiHT) imager 28 is an air-borne platform developed, in part, to examine how future space-originating LiDAR, e.g., ICESat-2, 29 GEDI, or other platforms, may be combined with field-based validation measurements to build predictive models for AGB and other forest variables; See [3] and [8] for more information). The dataset consist of two parts: (i) biomass measurements for over 451 locations, and (ii) LiDAR signal data observed at more than 100 different heights in more than 5000 locations including those we have their biomass (AGB) measurements. Figure 4.1 plots the Penobscot Experimental Forest together with the AGB measurements.

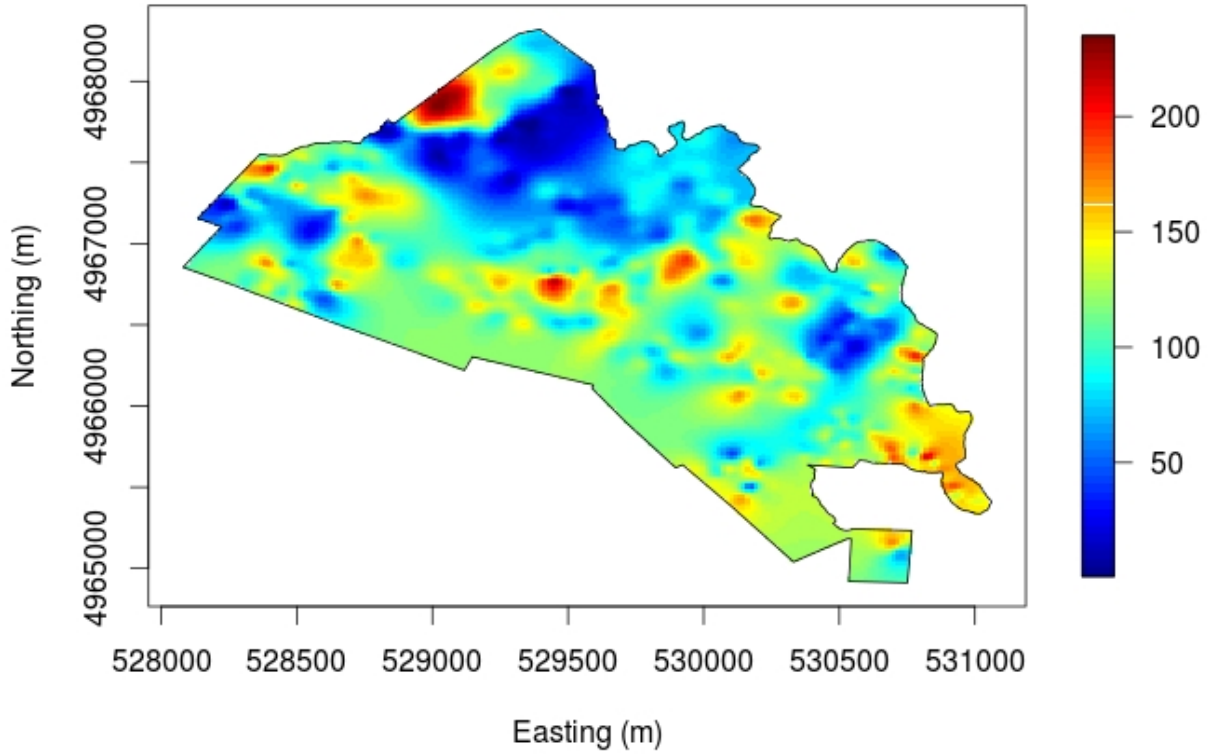


Figure 4.1 AGB measurements from the Penobscot Experimental Forest in Bradley, Maine

We only used the portion of the data in which we have both AGB measurement and LiDAR data for the analysis so that we would be able to use the multivariate model proposed in this chapter for data fitting. This portion consists of 451 locations with 64 outcome variables (63 from the LiDAR data, and one from the AGB). The maximum number of factors is set to be 10. The stochastic factor selection is done similarly as in [35] by putting Bernoulli priors on keeping each factor in the model independently (See section 2.3 in [35] for more explanation on the adaptive Bayesian factor selection). Further, 100 knots are chosen for the low-rank Gaussian predictive processes approximation. We used 20,000 MCMC iterations, including 2,000 samples for burn-in. Table 4.2 summarizes the fitted model which includes

some of the parameter estimates with their credible intervals. Out of all the 10 factors, factor # 5 was not selected in the stochastic selection procedure we have chosen in our model. One could observe the wide credible interval for the coefficient $\alpha_{f,5}$ as compared to all the other factor loadings as a sign of its corresponding factor not being selected. Moreover, figure 4.2 shows both the real data and the fitted values coming from the fitted model. Also, in figure 4.3 one can see the fitted values for the LiDAR data at several different locations.

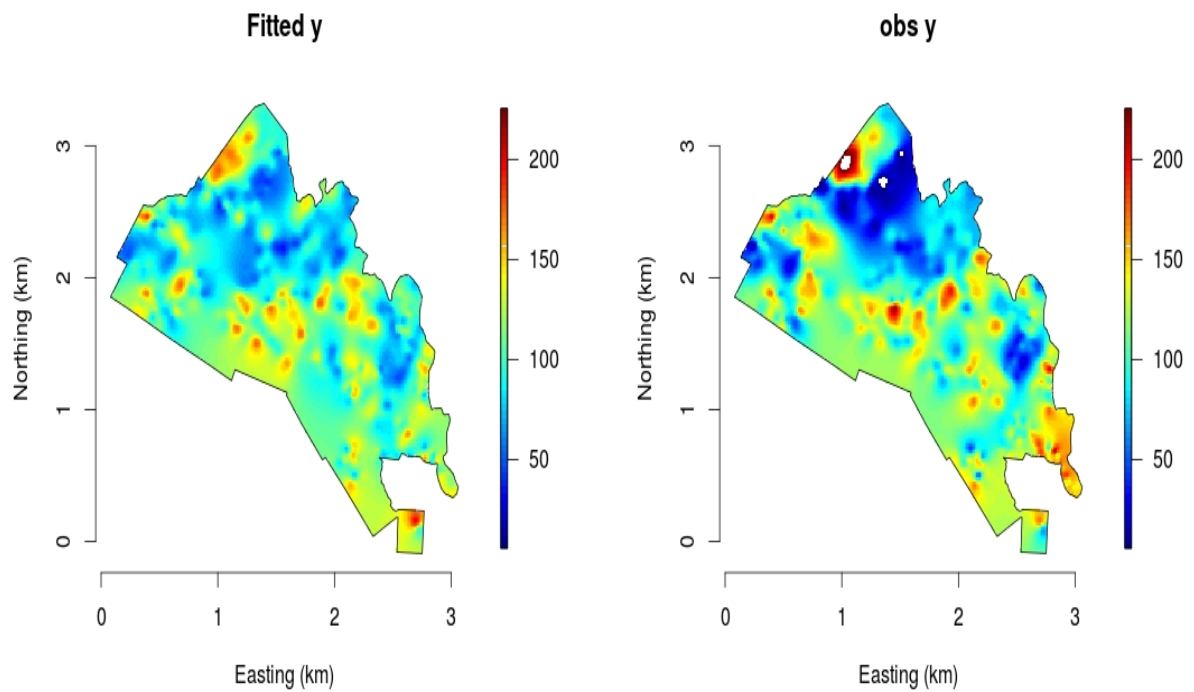


Figure 4.2 Fitted plot of AGB measurements together with the observed values

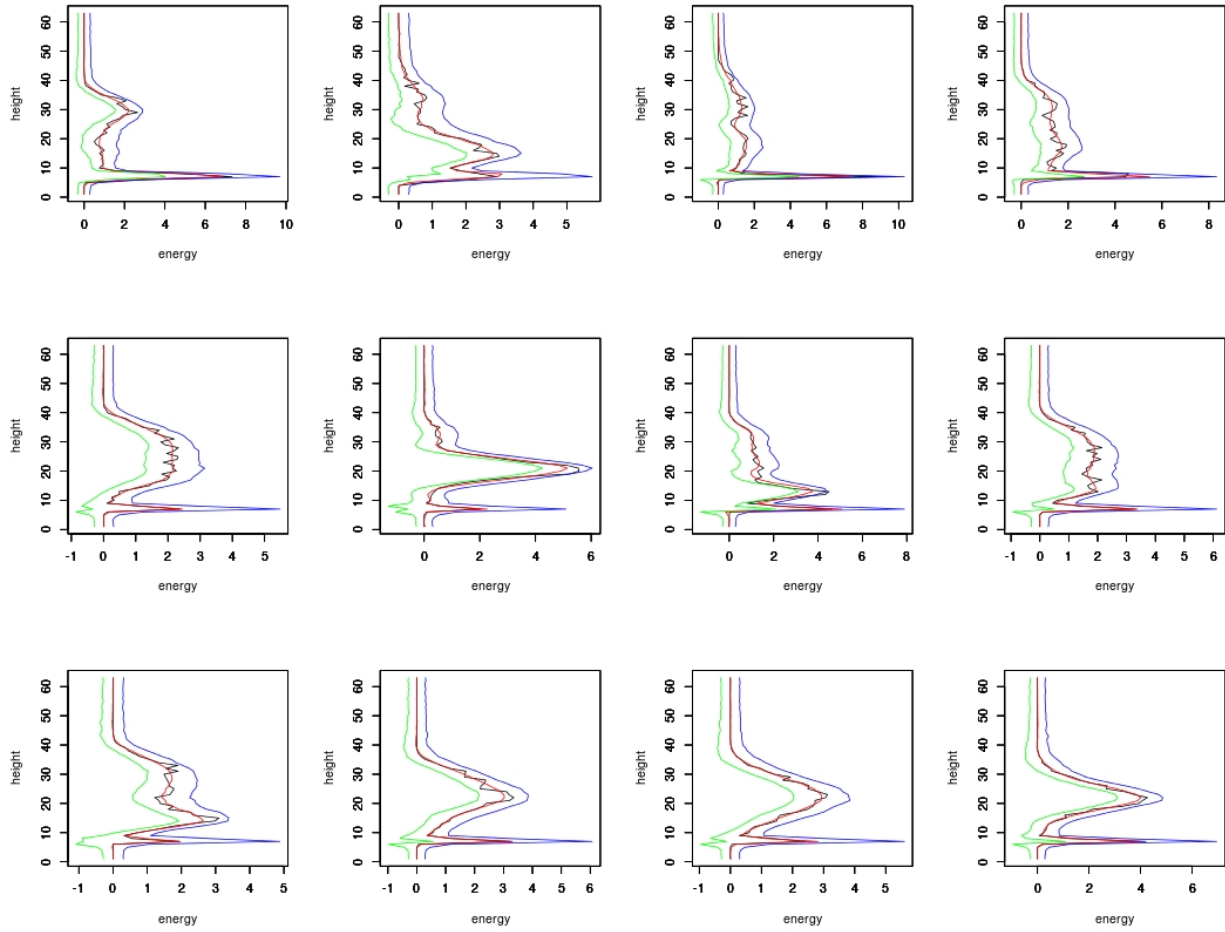


Figure 4.3 LiDAR data with fitted values and credible bands

Table 4.2 Parameter estimates with their 95% credible intervals

parameter	estimate	credible intervals
η_0	0.9912472	(0.9559148,1.0267694)
τ_y^2	0.1186832	(0.1039515,0.1366400)
$\alpha_{f,1}$	0.1887652	(-0.2071523,0.5462206)
$\alpha_{f,2}$	-0.1370774	(-0.4929575,0.1614393)
$\alpha_{f,3}$	0.0883730	(-0.16048225,0.32833808)
$\alpha_{f,4}$	0.0604460	(-0.16171807,0.27297671)
$\alpha_{f,5}$	-0.016114	(-6.11216314,6.18284331)
$\alpha_{f,6}$	-0.021537	(-0.24357843,0.08823335)
$\alpha_{f,7}$	0.0242004	(-0.13238865,0.18407468)
$\alpha_{f,8}$	-0.115645	(-0.30288243,0.01502838)
$\alpha_{f,9}$	0.1121742	(-0.2331279,0.2656526)
$\alpha_{f,10}$	-0.165493	(-0.2868342,0.1519692)
$\alpha_{f,\nu}$	0.112052	(0.005656521,0.357829027)
ϕ_1	0.7551075	(0.7501629,0.7775895)
ϕ_2	0.9855927	(0.9161703,1.0884193)
ϕ_3	1.375376	(1.211139,1.602903)
ϕ_4	2.140675	(1.734234,2.813785)
ϕ_5	14.867140	(3.371117,35.734567)
ϕ_6	31.397823	(5.248678,53.407113)
ϕ_7	46.783236	(8.355846,67.885794)
ϕ_8	61.83730	(18.15725,81.25511)
ϕ_9	77.90531	(35.90097,91.89012)
ϕ_{10}	95.05987	(55.68514,99.91256)
ϕ_ν	0.7620842	(0.7504561,0.8102750)

BIBLIOGRAPHY

BIBLIOGRAPHY

- [1] Anderes, E. B. and Stein, M. L.: “Estimating deformations of isotropic Gaussian random fields on the plane”, *The Annals of Statistics*, 719–741(2008)
- [2] Aronszajn, N.: “Theory of reproducing kernels”, *Transactions of the American mathematical society*, 68.3, 337–404(1950)
- [3] Awadallah, A., Abbott, V., Wynne, R., and Nelson, R.: “Estimating forest canopy height and biophysical parameters 454 using photon-counting laser altimetry”, *Proc. 13th International Conference on LiDAR Applications for Assessing Forest 455 Ecosystems*, 1291-136 (2013)
- [4] Banerjee, S., Gelfand, A. E., Finley, A. O., and Sang, H.: “Gaussian predictive process models for large spatial data sets”, *Journal of the Royal Statistical Society, Series B*, 70(4), 825–848 (2008)
- [5] Baudoin, F. and Nualart, D.: “Equivalence of Volterra processes”, *Stochastic processes and their applications*, 107(2), 327–350 (2003)
- [6] Chatterji, S. D. and Mandrekar, V.: “Equivalence and singularity of Gaussian measures and applications”, *Probabilistic analysis and related topics*, 1, 169–197 (1978)
- [7] Cheridito, P.: “Mixed fractional Brownian motion”, *Bernoulli*, 7(6), 913–934 (2001)
- [8] Cook, B., Corp, L., Nelson, R., Middleton, E., Morton, D., McCorkel, J., Masek, J., Ranson, K., Ly, V., and Montesano, P.: “Nasa goddards lidar, hyperspectral and thermal (g-liht) airborne imager”, *Remote Sensing*, 5, 4045-4066 (2013)
- [9] Cressie, N.A.C.: “Statistics for Spatial Data”, 2nd ed. New York: Jone Wiley & Sons (1993)
- [10] Cressie, N. and Huang, H. C.: “Classes of nonseparable, spatio-temporal stationary covariance functions”, *Journal of the American Statistical Association*, 94(448), 1330–1339(1999)
- [11] Cressie, N., Johannesson, G.: “Fixed rank kriging for very large spatial data sets”, *Journal of the Royal Statistical Society: Series B*, Vol. 70(1), 209–226 (2008)

- [12] Du, J., Zhang, H., Mandrekar, V. S.: “Fixed-domain asymptotic properties of tapered maximum likelihood estimators”, *The Annals of Statistics*, Vol. 37(6A), 3330–3361 (2009)
- [13] Dunford, N. and Schwartz, J. T.: “Linear operators. Part 2: Spectral theory. Self adjoint operators in Hilbert space”, New York (1963)
- [14] Ebeling, W.: “Functions of several complex variables and their singularities”, Vol. 83. American Mathematical Soc. (2007)
- [15] Feldman, J.: “Equivalence and perpendicularity of Gaussian processes”, *Pacific J. Math.*, 8(4), 699–708 (1958)
- [16] Finley, A. O., Sang, H., Banerjee, S., and Gelfand, A. E.: “Improving the performance of predictive process modeling for large datasets”, *Computational statistics and data analysis*, 53(8), 2873–2884 (2009)
- [17] Fuentes, M. and Smith, R. L.: “A new class of nonstationary spatial models”, Technical report, North Carolina State University, Raleigh, NC, (2001)
- [18] Furrer, R. and Genton, M. G. and Nychka, D.: “Covariance tapering for interpolation of large spatial datasets”, *Journal of Computational and Graphical Statistics*, 15(3), (2006)
- [19] Gikhman, I. I. and Skorokhod, A. V.: “The Theory of Stochastic Processes: I”, Springer (1965)
- [20] Gneiting, T.: “Nonseparable, stationary covariance functions for space–time data”, *Journal of the American Statistical Association*, 97(458), 590–600 (2002)
- [21] Gneiting, T. and Genton, M. and Guttorp, P.: “Geostatistical space–time models, stationarity, separability and full symmetry”, *Statistical Methods for Spatio–Temporal Systems*, 151–175 (2007)
- [22] Hajek, J.: “On a Property of Normal Distribution of Any Stochastic Processes”, *Math. Statist. Prob.*, 245–252 (1958)
- [23] Halmos, P. R.: “Introduction to Hilbert space and the theory of spectral multiplicity”, New York: Chelsea (1957)
- [24] Higdon, D. and Swall, J. and Kern, J.: “Non-stationary spatial modeling”, *Bayesian statistics*, 6(1), 761–768 (1999)

- [25] Ibragimov, I. A. and Rozanov, I. A.: “Gaussian random processes”, New York: Springer-Verlag. (1978)
- [26] Jun, M. and Stein, M. L.: “Nonstationary covariance models for global data”, The Annals of Applied Statistics, 1271–1289 (2008)
- [27] Kallianpur, G. and Oodaira, H.: “Non-anticipative representations of equivalent Gaussian processes”, The Annals of Probability, 1(1), 104–122 (1973)
- [28] Kaufman, C. G. and Schervish, M. J. and Nychka, D. W.: “Covariance tapering for likelihood-based estimation in large spatial data sets”, Journal of the American Statistical Association, 103(484), 1545–1555 (2008)
- [29] Luan, N. and Xiao, Y.: “Spectral conditions for strong local nondeterminism and exact Hausdorff measure of ranges of Gaussian random fields”, Journal of Fourier Analysis and Applications, 18(1), 118–145 (2012)
- [30] Nychka, D. and Wikle, C. and Royle, J. A.: “Multiresolution models for nonstationary spatial covariance functions”, Statistical Modeling, 2(4), 315–331 (2002)
- [31] Paciorek, C. J. and Schervish, M. J.: “Spatial modeling using a new class of nonstationary covariance functions”, Environmetrics, 17(5), 483–506 (2006)
- [32] Parzen, E.: “Probability density functionals and reproducing kernel Hilbert spaces”, Proceedings of the Symposium on Time Series Analysis, 196, 155–169 (1963)
- [33] Pitt, L. D.: “Stationary Gaussian Markov fields on \mathbb{R}^d with a deterministic component”, Journal of Multivariate Analysis, 5.3, 300–311 (1975)
- [34] Pitt, L. D.: “Some problems in the spectral theory of stationary processes on \mathbb{R}^d ”, Indiana Univ. Math. J., 23, 343–365 (1973)
- [35] Ren, Q., and Banerjee, S.: “Hierarchical Factor Models for Large Spatially Misaligned Data: A LowRank Predictive Process Approach”, Biometrics, 69(1), 19–30 (2013)
- [36] Ronkin, L. I.: “Introduction to the theory of entire functions of several variables”, American Mathematical Soc (1974)
- [37] Skorokhod, A. V. and Yadrenko, M. I.: “On absolute continuity of measures corresponding to homogeneous Gaussian fields”, Theory of Probability and Its Applications, 18(1), 27–40 (1973)

- [38] Sottinen, T. and Tudor, C. A.: “On the equivalence of multiparameter Gaussian processes”, *Journal of Theoretical Probability*, 19(2), 461–485 (2006)
- [39] Stein, M. L.: “Uniform asymptotic optimality of linear predictions of a random field using an incorrect second-order structure”, *The Annals of Statistics*, 850–872 (1990)
- [40] Stein, M. L.: “Predicting random fields with increasing dense observations”, *The Annals of Applied Probability*, 9(1), 242–273 (1999)
- [41] Stein, M. L.: “Interpolation of spatial data: some theory for kriging”, Springer (1999)
- [42] Stein, M. L.: “Equivalence of Gaussian measures for some nonstationary random fields”, *Journal of Statistical Planning and Inference*, 123(1), 1–11 (2004)
- [43] Stein, M. L.: “Space–time covariance functions”, *Journal of the American Statistical Association*, 100(469), 310–321 (2005)
- [44] Van Zanten, H.: “When is a linear combination of independent fBms equivalent to a single fBm?”, *Stochastic processes and their applications*, 117(1), 57–70 (2007)
- [45] Van Zanten, H.: “A remark on the equivalence of Gaussian processes”, *Electronic Communications in Probability*, 13, 54–59 (2008)
- [46] Xiao, Y.: “Strong local nondeterminism and sample path properties of Gaussian random Fields”, *Asymptotic Theory in Probability and Statistics with Applications* (T.-L. Lai, Q.-M. Shao and L. Qian, eds), Higher Education Press, Beijing., 136–176 (2007)
- [47] Xue, Y. and Xiao, Y.: “Fractal and smoothness properties of space-time Gaussian models”, *Frontiers of Mathematics in China*, 6(6), 1217–1248 (2011)
- [48] Yadrenko, M. I.: “Spectral theory of random fields”, A. V. Balakrisn’an (Ed.). *Optimization Software. Publications Division* (1983)
- [49] Yaglom, A. M.: “Some classes of random fields in n-dimensional space, related to stationary random processes”, *Theory of Probability and Its Applications*, 2(3), 273–320 (1957)
- [50] Zhang, H.: “Inconsistent estimation and asymptotically equal interpolations in model-based geostatistics”, *Journal of the American Statistical Association*, 99(465), 250–261 (2004)