

Pontificia Universidad Católica de Chile Faculty of Mathematics Department of Statistics

Spatiotemporal modeling of count data

Diego Fabián Morales Navarrete

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR IN STATISTICS

December 2021

© Copyright by Diego Morales-Navarrete, 2021.

All rights reserved. No part of the publication may be reproduced in any form by print, photoprint, microfilm, electronic or any other means without the prior written permission of one of the copyright holders.

PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE DEPARTMENT OF STATISTICS

The undersigned hereby certify that they have read and recommend to the Faculty of Mathematics for acceptance a thesis entitled **Spatiotemporal modeling of count data** by **Diego Fabián Morales Navarrete** in partial fulfillment of the requirements for the degree of **Doctor in Statistics**.

Dated: December, 2021

Research Supervisor :	
-	Luis Mauricio Castro Cepero
	Pontificia Universidad Católica de Chile
Research Co-Supervisor :	
	Moreno Bevilacqua
	Universidad Adolfo Ibañez
Examing Committee :	
-	Marcos Oliveira Prates
	Universidade Federal de Minas Gerais
	Poppy Vallaios
	Ronny vanejos
	Universidad Técnica Federico Santa María

Jonathan Daniel Acosta Pontificia Universidad Católica de Chile

PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE

Date: December 2021

Author	:	Diego Fabián Morales Navarrete
Title	:	Spatiotemporal modeling of count data
Department	:	Statistics
Degree	:	Doctor in Statistics
Convocation	:	December
Year	:	2021

Permission is herewith granted to Pontificia Universidad Católica de Chile to circulate and to have copied for non-commercial purposes, at its discretion, the above title upon the request of individuals or institutions.

Signature of Author

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

The author attests that permission has been obtained for the use of any copyrighted material appearing in this thesis (other than brief excerpts requiring only proper acknowledgement in scholarly writing) and that all such use is clearly acknowledged.

IN MEMORY OF MY GRANDMA MARÍA PERGRINA AND MY GODFATHER PEDRO FABIÁN

Acknowledgements

First, I would like to thank to my supervisors Mauricio Castro and Moreno Bevilacqua. I deeply appreciate their guidance and motivation throughout my doctoral training. Thank you for all the fruitful discussions that made my project to constantly improve and helped me to develop as a statistician and scientist. I am especially grateful that you believed in me and my skills as researcher. Thanks for your them mentoring and constant support.

Thanks to my family, which has been essential for me to be able to achieve my goals. Their unconditional support has made difficult times more bearable. I would like to thank my grandparents, my parents, my brothers Andres and Daniel and may aunts -especially Susana-. They have been always with me, in good and bad times.

I also would like to thank my friends and doctoral classmates, especially Gaby for the long talks about statistics and life. Thanks to Christian Caamaño for his help, suggestions, and comments on the thesis.

Finally, I acknowledge the founding of this thesis to VRI-UC scholarship and the partial support of ANID - Millennium Nucleus Center for the Discovery of Structures in Complex Data - MIDAS - NCN17_059, FONDECYT grant 1200068, and FONDECYT grant 1170258, from the Chilean government.

Contents

Li	st of t	ables		iii
Li	st of f	igures		vii
Al	bstrac	:t		ix
In	trodu	ction		xi
1	The	oretical	background	1
	1.1	Spatia	l and Spatio-temporal modeling	1
		1.1.1	Spatial and Spatio-temporal processes	1
		1.1.2	Stationarity and Isotropy	4
		1.1.3	Separability, full symmetry and compactly supported	9
		1.1.4	Properties of covariance functions	11
		1.1.5	Isotropic parametric models	18
		1.1.6	Continuity of the sample paths	19
		1.1.7	Modeling spatial and spatio-temporal data	20
		1.1.8	Kriging	23
	1.2	Renew	al processes	25
		1.2.1	Poisson process	27

CONTENTS

2	A ne	w class of counting random fields	31
	2.1	A random field with exponential marginal distributions	31
	2.2	Spatial Poisson random fields	33
		2.2.1 Correlation function	35
		2.2.2 Bivariate distribution	39
	2.3	Estimation and prediction	42
		2.3.1 Weighted pairwise likelihood estimation	42
		2.3.2 Optimal linear prediction	44
	2.4	Simulation studies	45
		2.4.1 Performance of the weighted pairwise likelihood estimation	45
		2.4.2 Performance of the optimal linear prediction	50
3	Mod	eling zero inflated Poisson spatial data	53
	3.1	A random field with Bernoulli marginal distribution	53
	3.2	Spatial zero inflated Poisson random fields	54
		3.2.1 Correlation function	57
		3.2.2 Bivariate distribution	59
	3.3	Simulation studies	61
		3.3.1 Performance of the weighted pairwise likelihood estimation	61
		3.3.2 Performance of the optimal linear prediction	66
4	Mod	eling over-dispersed spatial data	71
	4.1	A random field with Erlang marginal distribution	71
	4.2	Spatial Poisson-Erlang mixture random fields	72
		4.2.1 Correlation function	75
		4.2.2 Bivariate distribution	77

CONTENTS

4.3	Simulation studies		
	4.3.1	Performance of the weighted pairwise likelihood estimation	81
	4.3.2	Performance of the optimal linear prediction	85
Applications to real data			89
5.1	Applica	ation to the reindeer pellet-group survey in Sweden	89
5.2	Applica	ation to weed from the Bjertorp farm in Sweden	95
Conclusions and discussion		99	
Proofs of Theorems and Corollaries		eorems and Corollaries	103
A.1	Proofs	of Chapter 2	103
A.2	Proofs	of Chapter 3	116
A.3	Proofs	of Chapter 4	120
	4.3 Appl 5.1 5.2 Cond A.1 A.2 A.3	 4.3 Simula 4.3.1 4.3.2 Applications 5.1 Applications 5.2 Applications Conclusions Proofs of The A.1 Proofs A.2 Proofs A.3 Proofs	 4.3 Simulation studies

CONTENTS

List of Tables

2.1	Bias and MSE associated with Poisson wpl, misspecified Gaussian wpl	
	and misspecified Gaussian ML when the true random field is Poisson with	
	$\lambda(\mathbf{s}) = e^{\beta} \text{ and } \rho(\mathbf{h}) = (1 - \mathbf{h} /\alpha)_+^4. \dots \dots \dots \dots \dots \dots$	46
2.2	Bias and MSE associated with the Poisson wpl, misspecified Gaussian wpl	
	and misspecified Gaussian ML when estimating a non-stationary Poisson	
	random field with $\lambda(\mathbf{s}) = \exp\{\beta + \beta_1 u_1(\mathbf{s}) + \beta_2 u_2(\mathbf{s})\}$ and $\rho(\mathbf{h}) = (1 - \beta_1 u_1(\mathbf{s}) + \beta_2 u_2(\mathbf{s}))$	
	$ \mathbf{h} /\alpha)_+^4$	47
2.3	Bias and MSE associated with the Poisson wpl, misspecified Gaussian	
	wpl and misspecified Gaussian ML when estimating a non-stationary	
	spatio-temporal Poisson random field with $\lambda(\mathbf{s},t^*) = \exp\{eta + $	
	$\beta_1 u_1(\mathbf{s}, t^*) + \beta_2 u_2(\mathbf{s}, t^*) \}$ and $\rho(\mathbf{h}, t^*) = (1 - \mathbf{h} / \alpha_{\mathbf{s}})^4_+ (1 - t^* / \alpha_{t^*})^4_+.$	49
2.4	Empirical mean of RMSE and MAE associated with the optimal linear	
	predictor of the proposed Poisson random field (P), the optimal predictor	
	based on the Poisson GC (PGC) and Poisson LG (PLG) approaches	
	when the datasets are simulated from the P and PGC models under three	
	scenarios, namely, Scenario 1 (low counts), Scenario 2 (medium counts)	
	and Scenario 3 (large counts).	52
3.1	Bias and MSE associated with ZIP wpl and misspecified ZIP (MZIP) wpl	
	when the true random field is ZIP with $\lambda(\mathbf{s}) = e^{\beta}$, $p = \Phi(\theta)$ and $\rho(\mathbf{h}) =$	
	$(1 - \mathbf{h} /\alpha)_+^4$.	63

3.2	Bias and MSE associated with ZIP wpl and misspecified ZIP (MZIP) wpl	
	when estimating a non-stationary ZIP random field with $\lambda(\mathbf{s}) = \exp\{eta +$	
	$\beta_1 u_1(\mathbf{s}) + \beta_2 u_2(\mathbf{s})\}, p = \Phi(\theta) \text{ and } \rho(\mathbf{h}) = (1 - \mathbf{h} /\alpha)_+^4. \dots \dots$	64
3.3	Bias and MSE associated with ZIP wpl and misspecified ZIP (MZIP)	
	wpl when estimating a non-stationary ZIP random field with $\lambda(\mathbf{s},t^*)$ =	
	$\exp\{\beta + \beta_1 u_1(\mathbf{s}, t^*) + \beta_2 u_2(\mathbf{s}, t^*)\}, \ p(\mathbf{s}, t^*) = \Phi(\theta) \text{ and } \rho(\mathbf{h}) = (1 - 1)$	
	$ \mathbf{h} /\alpha)_{+}^{4}$	65
3.4	Empirical mean of RMSE and MAE associated with the optimal linear	
	predictor of the proposed ZIP random field (ZIP), the optimal predictor	
	based on the ZIP GC (ZIPGC) and ZIP LG (ZIPLG) approaches when	
	the datasets are simulated from the ZIP and ZIPGC models under three	
	scenarios, namely, Scenario 1 (low zeros), Scenario 2 (medium zeros) and	
	Scenario 3 (large zeros).	69
4.1	Bias and MSE associated with Poisson-Erlang wpl and misspecified PEM	
	(MPEM) wpl when the true random field is Poisson-Erlang mixture with	
	$\kappa = 1, 5, \mu(\mathbf{s}) = e^{\beta}$, and $\rho(\mathbf{h}) = (1 - \mathbf{h} /\alpha)_+^4 \dots \dots \dots \dots$	82
4.2	Bias and MSE associated with PEM wpl and misspecified PEM (MPEM)	
	wpl when estimating a non-stationary PEM random field with $\mu(\mathbf{s})$ =	
	$\exp\{\beta + \beta_1 u_1(\mathbf{s}) + \beta_2 u_2(\mathbf{s})\}, \kappa = 1 \text{ and } \rho(\mathbf{h}) = (1 - \mathbf{h} /\alpha)_+^4. \ldots$	83
4.3	Bias and MSE associated with PEM wpl and misspecified PEM (MPEM)	
	<i>wpl</i> when estimating a non-stationary PEM random field with $\mu(\mathbf{s},t^*)=$	
	$\exp\{\beta + \beta_1 u_1(\mathbf{s}, t^*)\}, \kappa = 1 \text{ and } \rho(\mathbf{h}) = (1 - \mathbf{h} /\alpha)_+^4. \dots \dots \dots$	84
4.4	Empirical mean of RMSE and MAE associated with the optimal linear	
	predictor of the proposed PEM random field (PEM), the optimal predictor	
	based on the NB GC (NBGC) and NB LG (NBLG) approaches when the	
	datasets are simulated from the PEM and Negative Binomial (NB) GC	
	models under three scenarios, namely, Scenario 1 ($\kappa = 1$) and Scenario 2	
	$(\kappa = 5)$.	88

LIST OF TABLES

5.1	Parameter estimates for the reindeer pellet-group survey data obtained	
	under the ZIP, ZIP GC and ZIP LG random fields. The associated standard	
	errors are in parenthesis. The last row shows the associated empirical	
	mean of the RMSE for each model	94
5.2	Parameter estimates for the weed data obtained under the PEM, Negative	
	Binomial GC and Negative Binomial LG random fields. The associated	
	standard errors are in parenthesis. The last row shows the associated	
	empirical mean of the RMSE for each model.	98

iv

List of Figures

A realization of a Poisson LG random field where the LG random field is given by e^{μ+√σ²G(s)} with G a standard Gaussian random field with parameters μ = 0.5 and σ² = 0.05 (a) and its associated histogram (c). A realization of our proposed Poisson random field with λ = e^{0.5+0.05/2} (b) and its associated histogram (d). In both cases the underlying isotropic correlation is ρ(r) = (1 - r/0.5)⁴₊.

xii

38

- 2.1 From left to right: (a) correlation functions $\rho_Y(\mathbf{h}, \mu, \sigma^2)$ of the Poisson LG random field with $\mu = 0.5, \sigma^2 = 0.05$ and $\mu = 2.5, \sigma^2 = 0.1$, and $\mu = 4.5, \sigma^2 = 0.2$; (b) correlation function $\rho_N(\mathbf{h}, \lambda)$ of our proposed Poisson random field for $\lambda = 1.69, 12.81$, and, 99.48; (c) correlation function $\rho_C(\mathbf{h}, \lambda)$ of the Poisson GC random field for $\lambda = 1.69, 12.81$, and, 99.48. The black line in the Figures depicts the underlying correlation model given by $\rho(\mathbf{h}) = (1 - ||\mathbf{h}||/0.5)_+^4$.

- 2.3 Centred box-plots of estimates under the Poisson *wpl* (P), misspecified Gaussian *wpl* (PG), and misspecified Gaussian ML (MG) when estimating a non-stationary Poisson random field with $\lambda(\mathbf{s}) = \exp\{\beta + \beta_1 u_1(\mathbf{s}) + \beta_2 u_2(\mathbf{s})\}, \beta = 0.5, \beta_1 = -0.2, \beta_2 = 0.3 \text{ and } \rho(\mathbf{h}) = (1 - ||\mathbf{h}||/\alpha)_+^4$ with $\alpha = 0.2$. 48
- 2.4 Centred box-plots of estimates under the Poisson *wpl* (P), misspecified Gaussian *wpl* (PG), and misspecified Gaussian ML (MG) when estimating a non-stationary space-time Poisson random field with $\lambda(\mathbf{s}, t^*) = \exp\{\beta + \beta_1 u_1(\mathbf{s}, t^*) + \beta_2 u_2(\mathbf{s}, t^*)\}$ with $\beta = 0.5$, $\beta_1 = -0.2$ and $\beta_2 = 0.3$ and $\rho(\mathbf{h}, t^*) = (1 - ||\mathbf{h}|| / \alpha_{\mathbf{s}})^4_+ (1 - |t^*| / \alpha_{t^*})^4_+$ with $\alpha_{\mathbf{s}} = 0.2$, $\alpha_{t^*} = 1$ 49
- 3.1 From the top to the bottom. A realization (column a) of a ZIP random field with p(s) = Φ(-2), Φ(-1), Φ(0), Φ(1) and its associated histogram (column b), respectively. In all cases λ(s) = 5 and ρ(h) = (1-||h||/0.2)⁴)₊.
- 3.3 For each row (from left to right): bivariate ZIP GC distribution, our proposed bivariate ZIP distribution and the difference between them. The first, second and third row are obtained setting $\rho(\mathbf{h}) = 0.1, 0.5, 0.9$ for the underlying correlation.
- 3.4 Centred box-plots of estimates under ZIP *wpl* (ZIP) and misspecified ZIP (MZIP) *wpl* when estimating a non-stationary Poisson random field with $\lambda(\mathbf{s}) = \exp\{\beta + \beta_1 u_1(\mathbf{s}) + \beta_2 u_2(\mathbf{s})\}, \beta = 2, \beta_1 = 0.8, \beta_2 = -1.2,$ $p = \Phi(\theta), \theta = 1 \text{ and } \rho(\mathbf{h}) = (1 - ||\mathbf{h}||/\alpha)_+^4 \text{ with } \alpha = 0.2. \dots 64$

4.1	From the top to the bottom. A realization (column a) of a Poisson-Erlang	
	mixture random field with $\kappa = 1, 5, 100$ and its associated histogram	
	(column b), respectively. In all cases $\mu(\mathbf{s}) = 5$ and $\rho(\mathbf{h}) = (1 - \mathbf{h} /0.2)_+^4$.	74
4.2	Correlation function $ ho_M({f h},\mu,\kappa)$ of the Poisson-Erlang mixture random	
	field with $\kappa = 1, 5, 100, \mu = 5, \rho_N(\mathbf{h}, 5)$ and $\rho(\mathbf{h}) = (1 - \mathbf{h} /\alpha)_+^4$ with	
	$\alpha = 0.5.$	76
4.3	For each row (from left to right): bivariate Negative Binomial GC	
	distribution, our proposed bivariate Poisson-Erlang mixture distribution	
	and the difference between them. The first, second and third row are	
	obtained setting $ ho({f h})=0.1, 0.5, 0.9$ for the underlying correlation	80
4.4	Centred box-plots of estimates under PEM wpl (PEM) and misspecified	
	PEM (MPEM) wpl when estimating a non-stationary Poisson random field	
	with $\mu(\mathbf{s}) = \exp\{\beta + \beta_1 u_1(\mathbf{s})\}, \beta = 1.5, \beta_1 = -0.5, \kappa = 1 \text{ and } \rho(\mathbf{h}) =$	
	$(1 - \mathbf{h} /\alpha)_+^4$ with $\alpha = 0.2$.	83
4.5	Centred box-plots of estimates under PEM wpl (PEM) and misspecified	
	PEM wpl (MPEM) when estimating a non-stationary Poisson random field	
	with $\mu(\mathbf{s}, t^*) = \exp\{\beta + \beta_1 u_1(\mathbf{s}, t^*)\}, \beta = 1.5, \beta_1 = -0.5, \kappa = 1$ and	
	$ \rho(\mathbf{h}, t^{\star}) = (1 - \mathbf{h} / \alpha_{\mathbf{s}})_{+}^{4} (1 - t^{\star} / \alpha_{t^{\star}})_{+}^{4} \text{ with } \alpha_{s} = 0.2, \alpha_{t^{\star}} = 1. \dots $	85
5.1	Examples of point transect survey design	90
5.2	Spatial location of reindeer pellet-group survey data.	91
5.3	Empirical semi-variogram of reindeer pellet-group survey data	91
5.4	Spatial location of weed.	96
5.5	Empirical semi-variogram of weed data	96

Abstract

Modeling spatial and spatio-temporal data is a challenging task in statistics. In many applications, the observed data can be modeled using Gaussian, skew-Gaussian or even restricted random field models. However, in several fields, such as population genetics, epidemiology, aquaculture, among others, the data of interest are often count data, and therefore the mentioned models are not suitable for the analysis of this type of data. Consequently, there is a need for spatial and spatio-temporal models that are able to properly describe data coming from counting processes. Commonly two approaches are used to model this type of data: generalized linear mixed models (GLMMs) with Gaussian random field (GRF) effects, and copula models. Unfortunately, these approaches do not give an explicit characterization of the count random field such us their q-dimensional distribution or correlation function. It is important to stress that GLMMs models induces a discontinuity in the path. Therefore, the correlation function is not continuous at the origin and samples located nearby are more dissimilar than in the continuous case. Moreover, there are cases in which the copula representation for discrete distributions is not unique, so it is unidentifiable. Hence, to deal with the latter mentioned issues, we propose a novel approach to model spatial and spatio-temporal count data in an efficient and accurate manner. Briefly, starting from independent copies of a "parent" GRF, a set of transformations can be applied, and the result is a non-Gaussian random field. This approach is based on the characterization of count random fields that inherit some of the well-known geometric properties from GRFs. For instance, if one chooses an isotropic correlation function defined in the parent GFR, then the count random fields have an

isotropic correlation function. Firstly, we define a general class of count random fields. Then, three particular count random fields are studied. The first one is a Poisson random field, the second one is a count random field that considers excess zeros and the last one is a count random field that considers over-dispersion. Additionally, a simulation study will be developed to assess the performance of the proposed models. In that way, we are going to evaluate them through several simulation scenarios, making variations in the parameters. The results show accurate estimations of the parameters for different scenarios. Additionally, we assess the performance of the optimal linear prediction of the proposed models and it is compared with GLMMs and copula models. The results show that the proposed models have a better performance than GLMMs models and a quite similar performance with copula models. Finally, we analyze two real data applications. The first one considers a zero inflated version of the proposed Poisson random field to deal with excess zeros and the second one considers an over-dispersed count random field.

Introduction

Spatial or spatio-temporal count data are routinely collected in many earth and social sciences, such as ecology, epidemiology, demography, agriculture and geography. For instance, in ecological sciences, an important goal is to estimate and predict the temporal evolution of species distribution (in terms of abundance) in a region (Wang et al., 2015; Quiroz and Prates, 2018).

The analysis of spatial and spatio-temporal count data requires the development of statistical models for geo-referenced count data that take into account the spatial and spatio-temporal dependence. Random fields or stochastic processes are useful models when dealing with geo-referenced spatial or spatio-temporal data (Stein, 1999; Cressie and Wikle, 2011; Banerjee et al., 2014). In particular, the Gaussian random field is widely used due to its attractive properties and mathematical tractability (Gelfand and Schliep, 2016). Gaussianity is clearly a restrictive assumption when dealing with counting data. However, many models of current use for spatial count data employ Gaussian random fields as building blocks.

The first example is the hierarchical model approach proposed by Diggle et al. (1998), which can be viewed as a generalized linear mixed model (Diggle and Ribeiro, 2007; Diggle and Giorgi, 2019). Under this framework, non-Gaussian models for spatial data can be specified using a link function and a latent Gaussian random field through a conditional independence assumption. In particular, the Poisson Log-Gaussian random field (Poisson LG hereafter) has been widely applied for modeling count spatial data (see for instance Christensen and Waagepetersen, 2002; Guillot et al., 2009; De Oliveira, 2013,

for interesting applications and in-depth study of its properties). Similar models, that can be defined hierarchically in terms of the specification of the first two moments and a correlation function have been proposed in Monestiez et al. (2006) and De Oliveira (2014).

It is important to stress that the conditional independence assumption underlying these kind of models leads to (a) random fields with marginal distributions that are not Poisson and (b) random fields with a "forced" nugget effects that implies no mean square continuity.



Figure (1) A realization of a Poisson LG random field where the LG random field is given by $e^{\mu+\sqrt{\sigma^2}G(\mathbf{s})}$ with G a standard Gaussian random field with parameters $\mu = 0.5$ and $\sigma^2 = 0.05$ (a) and its associated histogram (c). A realization of our proposed Poisson random field with $\lambda = e^{0.5+0.05/2}$ (b) and its associated histogram (d). In both cases the underlying isotropic correlation is $\rho(r) = (1 - r/0.5)_{+}^4$.

xii

To illustrate this situation, Figure 1 (a) shows a realization on the unit square of a Poisson LG random field, assuming $e^{\mu+\sqrt{\sigma^2}G(\mathbf{s})}$ as an LG random field, where G is the standard Gaussian random field with isotropic correlation $\rho(r) = (1 - r/0.5)^4_+$ belonging to the Generalized Wendland family (Bevilacqua et al., 2019), $\mu = 0.5$, $\sigma^2 = 0.05$, r is the spatial distance and $(\cdot)_+$ denotes the positive part. In this case, the mean of the Poisson LG field is given by $\lambda = e^{0.5+0.05/2}$. The associated histogram is shown in 1 (d). Additionally, Figure 1 (b) shows a realization and the associated histogram of our proposed random field (see Equation (2.2)), with the same mean and underlying correlation function of the Poisson LG model. A quick analysis of both figures reveal a "whitening" effect on the Poisson LG random field's paths because of the "forced" discontinuity at the origin of the correlation function of the Poisson LG (see Section 2.2.1). This potential problem, which has also been highlighted by De Oliveira (2013), indicates that the Poisson LG random field may impose severe restrictions on the correlation structure and may be inadequate to model spatial count data mostly consisting of small counts.

The second example is the Poisson spatial model obtained using Gaussian copula (Kazianka and Pilz, 2010; Masarotto and Varin, 2012; Joe, 2014), which is referred to as the Poisson GC random field hereafter. This approach has some potential benefits with respect to the hierarchical models (see Han and De Oliveira, 2016, for a comparison between these two approaches). For instance, the resulting random field has Poisson marginals and can be mean square continuous or not depending on if the latent Gaussian random field is mean square continuous or not. In addition to some criticisms concerning the lack of uniqueness of the copula when applied to discrete data (Genest and Neslehova, 2007; Trivedi and Zimmer, 2017), in this approach it is not clear what the underlying physical mechanism generating the data is, making it less interesting from an interpretability perspective.

Our proposal tries to solve the drawbacks of the Poisson LG and of the Poisson GC approaches by specifying a new class of spatial counting random fields based on the Poisson counting process (Cox, 1970; Mainardi et al., 2007; Ross, 2008) applied to the spatial setting. Specifically, we first consider a random field with exponential

marginal distributions obtained as a rescaled sum of two independent copies of an underlying standard Gaussian random field. Then, by considering independent copies of the exponential random field as 'inter-arrival times' in the counting renewal processes framework, we obtain a (non-)stationary random field with Poisson marginal distributions. As a consequence, the proposed model can be viewed as a spatial generalization of the Poisson process.

For the novel Poisson random field, we provide the covariance function and analytic expressions for the bivariate distribution in terms of the regularized incomplete Gamma and confluent hypergeometric functions (Gradshteyn and Ryzhik, 2007). It turns out that the dependence of the proposed Poisson random field is indexed by the correlation function of the underlying Gaussian random field and by the mean parameter. It is important to stress that our theoretical results are inspired by the two-dimensional renewal theory described in Hunter (1974). We propose two additional random fields that can deal with the excessive number of zeros and over-dispersion in the data. Specifically, using a Bernoulli random field we develop a zero inflated version of the propose a random field version of the Poisson-Erlang mixture model, *i.e.*, a mixture of Poisson random fields with Erlang random field mixing weights.

The Poisson random field estimation is conducted using the weighted pairwise likelihood (*wpl*) method (Lindsay, 1988; Varin et al., 2011; Bevilacqua and Gaetan, 2015) to exploit the results obtained about the bivariate distribution. In particular, in an extensive simulation study, we explore the efficiency of the *wpl* method when estimating the parameters of the proposed Poisson random field. We also explore the statistical efficiency of a Gaussian misspecified version of the *wpl* method (Gouriéroux et al., 2017; Bevilacqua et al., 2020), which is also called Gaussian quasi-likelihood in some literature (Masuda, 2013). The findings show that the misspecified *wpl* leads to a less efficient estimator, in particular for low counts. However, the method has some computational benefits. This method is also used to the zero inflated Poisson random fields and the Poisson-Erlang mixture random field and the results are similar.

Introduction

In addition, we compare the performance of the optimal linear predictor under the Poisson and zero inflated Poisson proposed models with the optimal predictors obtained using the GC and LG models. The proposed over-dispersed model was compared with the Negative Binomal GC and Negative Binomal LG models. Finally, in the real data application, we consider a zero inflated Poisson random field to deal with excess zeros in the reindeer pellet-group counts data using the LG and GC models as benchmarks; and to deal with the over-dispersion in the weed counts from the Bjertorp farm, we consider a Poisson-Erlang mixture random field using the Negative Binomial LG and Negative Binomial GC models as benchmarks . The proposed methods in this thesis are implemented in the R (R Core Team, 2020) package GeoModels (Bevilacqua et al., 2019).

The manuscript is organized as follows: Chapter 1 provides background material of spatial modeling, spatio-temporal modeling and renewal theory. Chapter 2,3 and 4 provides the details of Poisson, zero inflated Poisson, and Poisson-Erlang mixture random fields, respectively. Chapter 5 contains applications with excess of zeros and over dispersed spatial count data. Chapter 2,3 and Section 1 of Chapter 5 form part of a submitted paper, while Chapter 4 and Section 2 of Chapter 5 are part of a working paper.

Introduction

xvi

Chapter 1

Theoretical background

In this chapter we review few basic concepts of spatial and spatio-temporal processes such as stationarity, covariance functions, isotropy, geometric properties and kriging, among others. Additionally, we review concepts of renewal processes, particularly, the Poisson process.

1.1 Spatial and Spatio-temporal modeling

1.1.1 Spatial and Spatio-temporal processes

Spatial processes

A spatial process or spatial stochastic process Q, which can be written as $Q = \{Q(\mathbf{s}), \mathbf{s} \in A\}$, is a collection of random variables $Q(\mathbf{s})$ at location s, and A is a set that indexes all possible spatial locations of interest. Furthermore, A is a subset of the d-dimensional Euclidean space \mathbb{R}^d , *i.e.*, $A \subseteq \mathbb{R}^d$. If d = 1, Q is called stochastic (random) process. Otherwise, if $d \ge 2$, Q is called random field (RF). A formal definition for a spatial stochastic process is given bellow.

Definition 1.1.1. Let (Ω, \mathcal{F}, P) be a probability space and $A \subseteq \mathbb{R}^d$ an arbitrary set.

For each $\mathbf{s} \in A$ the real valued function $Q(\mathbf{s}, \cdot) : \Omega \to \mathbb{R}, \omega \mapsto (\mathbf{s}, \omega)$ is a random variable. Thus, any collection of random variables $Q = \{Q(\mathbf{s}, \cdot), \mathbf{s} \in A \subseteq \mathbb{R}^d\}$ defined on (Ω, \mathcal{F}, P) is a stochastic process with index set A.

Hereafter we call spatial process, stochastic process or random field interchangeably.

If we fix any event $\omega \in \Omega$, then $\{Q(\mathbf{s}, \omega), \mathbf{s} \in A \subseteq \mathbb{R}^d\}$ is a realization or sample path of the random field Q. Formally,

Definition 1.1.2. A sample path of a stochastic process Q is a mapping $Q : A \to \mathbb{R}$, s $\mapsto Q(\mathbf{s}, \omega)$ which to every event $\omega \in \Omega$ corresponds a sample path or realization of stochastic process Q.

Let $\{\mathbf{s}_1, \ldots, \mathbf{s}_n\} \subset A$ be a finite set of spatial locations with $n \in \mathbb{N}$, then the *finite* dimensional joint distribution of the stochastic process $\{Q(\mathbf{s}), \mathbf{s} \in A\}$ is defined as

$$F_{\mathbf{s}_1,\mathbf{s}_2,\ldots,\mathbf{s}_n}(q_1,\ldots,q_n) = \Pr(Q(\mathbf{s}_1) \le q_1,\ldots,Q(\mathbf{s}_n) \le q_n).$$

A random field can be defined through its finite dimensional joint distribution if it has the properties of permutation invariance and projection invariance. In fact, the following theorem of Kolmogorov supports this statement.

Theorem 1.1.1 (Kolmogorov Consistency Theorem). Assume that for each $n \in \mathbb{N}$ and for each set of indexing points $\{s_1, \ldots, s_n\}$, we define a finite dimensional joint distribution F_{s_1,s_2,\ldots,s_n} . If the collection of all such F_{s_1,s_2,\ldots,s_n} satisfy

1. Projection invariance. For all $n \in \mathbb{N}$ *and all* $\mathbf{s}_1, \ldots, \mathbf{s}_n \in A$ *it holds that*

$$F_{\mathbf{s}_1,\mathbf{s}_2,\dots,\mathbf{s}_n,\mathbf{s}_{n+1}}(q_1,q_2,\dots,q_n,q_{n+1}) \to F_{\mathbf{s}_1,\mathbf{s}_2,\dots,\mathbf{s}_n}(q_1,q_2,\dots,q_n) \text{ as } q_{n+1} \to \infty.$$

2. Permutation invariance. Let π : $\{1, \ldots, n\} \rightarrow \{1, \ldots, n\}$ be a permutation of a subset of $\{1, \ldots, n\}$. Then, for all $n \in \mathbb{N}$ and all $\mathbf{s}_1, \ldots, \mathbf{s}_n \in A$,

$$F_{\mathbf{s}_1,\mathbf{s}_2,\ldots,\mathbf{s}_n}(q_1,q_2,\ldots,q_n) = F_{\mathbf{s}_{\pi(1)},\mathbf{s}_{\pi(2)},\ldots,\mathbf{s}_{\pi(n)}}(q_{\pi(1)},q_{\pi(2)},\ldots,q_{\pi(n)})$$

Then, there exists a probability space (Ω, \mathcal{F}, P) and a stochastic process $\{Q(\mathbf{s}), \mathbf{s} \in A\}$ whose finite dimensional joint distributions are given by the collection of all such $F_{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_n}$.

Thus, using the multivariate normal distribution we can define an important random field as follows.

Definition 1.1.3. $G = \{G(\mathbf{s}), \mathbf{s} \in A\}$ is called Gaussian process or Gaussian random field if for all n and admissible $\mathbf{s}_1, \ldots, \mathbf{s}_n$, the joint distribution of $G(\mathbf{s}_1), \ldots, G(\mathbf{s}_n)$ is multivariate normal.

Finally, we define a second order random field as shown below.

Definition 1.1.4. A stochastic process $\{Q(\mathbf{s}), \mathbf{s} \in A\}$ is said to be a second order process (random field) if for all $\mathbf{s} \in A$, $\mathbb{E}(Q(\mathbf{s})^2) < \infty$.

Therefore, a Gaussian random field is a second order random field and it is characterized in terms of its mean vector and covariance matrix, *i.e.*, its first and second moments.

Spatio-temporal processes

A spatio-temporal process is a generalization of a spatial process, where every spatio-temporal location (index) can be seen as a point on $\mathbb{R}^d \times \mathbb{R}$, *i.e.*, the collection of random variables varies in both the spatial and temporal domains. Formally, $\mathbb{R}^d \times \mathbb{R} = \mathbb{R}^{d+1}$, *i.e.*, time can be considered as an additional coordinate. Moreover, some results on spatial covariance functions, kriging, among others apply to space–time process as well. But from a physical viewpoint, this is insufficient because time differs intrinsically from space. Hereafter we set the space-time domain as $\mathbb{R}^d \times \mathbb{R}$ and a spatio-temporal process will be called spatio-temporal random field. A formal definition is given as follows.

Definition 1.1.5. Let (Ω, \mathcal{F}, P) be a probability space, $A \subseteq \mathbb{R}^d$ a finite domain in space and $T \subseteq \mathbb{R}$ a finite domain in time. For each $\mathbf{s} \in A$ and $t \in T$ the real valued function $Q(\mathbf{s}, t, \cdot) : \Omega \to \mathbb{R}, \omega \mapsto (\mathbf{s}, t, \omega)$ is a random variable. Thus, any collection of random variables $Q = \{Q(\mathbf{s}, t, \cdot), (\mathbf{s}, t) \in A \times T \subseteq \mathbb{R}^d \times \mathbb{R}\}$ is a spatio-temporal random field with spatio-temporal domain $A \times T$.

As in the case of spatial random fields, the spatio-temporal Gaussian random field can be defined as follows.

Definition 1.1.6. A spatio-temporal random field $\{G(\mathbf{s},t), (\mathbf{s},t) \in A \times T \subseteq \mathbb{R}^d \times \mathbb{R}\}$ is said to be Gaussian if the random vector $(G(\mathbf{s}_1, t_1), \dots, G(\mathbf{s}_n, t_n))^{\top}$, for any set of spatio-temporal locations $\{(\mathbf{s}_1, t_1), \dots, (\mathbf{s}_n, t_n)\}$, follows a multivariate normal distribution.

Moreover, a second order spatio-temporal random field is defined as follows.

Definition 1.1.7. A spatio-temporal random field $\{Q(\mathbf{s},t), (\mathbf{s},t) \in A \times T\}$ is said to be a second order random field if for all $(\mathbf{s},t) \in A \times T$, $\mathbb{E}(Q(\mathbf{s},t)^2) < \infty$.

1.1.2 Stationarity and Isotropy

On the one hand, stationarity in spatial (spatio-temporal) process intuitively means shift-invariance on space (space-time) of the random field's probabilistic properties. On the other hand, isotropy means invariance under rotations of the random field. The most common definitions of stationarity are strict stationarity and weak stationarity, but Matheron (1965) introduced a weaker concept called intrinsic stationarity. The different kinds of stationarity are defined as follow.

Definition 1.1.8. A spatial process $\{Q(\mathbf{s}), \mathbf{s} \in A\}$ is said strictly stationary if for all $\mathbf{s}_1, \ldots, \mathbf{s}_n$ and any $\mathbf{h} \in \mathbb{R}^d$, the joint distribution of $Q(\mathbf{s}_1), \ldots, Q(\mathbf{s}_n)$ is identical with the joint distribution of $Q(\mathbf{s}_1 + \mathbf{h}), \ldots, Q(\mathbf{s}_n + \mathbf{h}), i.e.$,

 $\Pr(Q(\mathbf{s}_1) \le q_1, \dots, Q(\mathbf{s}_n) \le q_n) = \Pr(Q(\mathbf{s}_1 + \mathbf{h}) \le q_1, \dots, Q(\mathbf{s}_n + \mathbf{h}) \le q_n),$

where $q_1, \ldots, q_n \in \mathbb{R}$.

Definition 1.1.9. A spatio-temporal random field $\{Q(\mathbf{s},t), (\mathbf{s},t) \in A \times T\}$ is strictly stationary if its probability distribution is translation invariant. In other words, if, in the case of any two given vectors, $\mathbf{h} \in A$ and $u \in T$:

$$(Q(\mathbf{s}_1, t_1), Q(\mathbf{s}_2, t_2), \dots, Q(\mathbf{s}_n, t_n))^{\top}$$

and

$$(Q(\mathbf{s}_1 + \mathbf{h}, t_1 + u), Q(\mathbf{s}_2 + \mathbf{h}, t_2 + u), \dots, Q(\mathbf{s}_n + \mathbf{h}, t_n + u))^{\top}$$

have the same multivariate distribution function.

Definition 1.1.10. A second order process $\{Q(\mathbf{s}), \mathbf{s} \in A\}$ is weakly stationary if:

$$\mathbb{E}(Q(\mathbf{s})) = \mu \ \forall \, \mathbf{s},$$

and

$$\forall i, j \in \mathbb{N}, \ \operatorname{Cov}(Q(\mathbf{s}_i), Q(\mathbf{s}_j)) = \operatorname{C}(\mathbf{s}_j - \mathbf{s}_i) = \operatorname{C}(\mathbf{h}).$$

If $Q(\mathbf{s}, t)$ is a second order spatio-temporal random field, then we define weak stationarity as follows.

Definition 1.1.11. A spatio-temporal random field $Q(\mathbf{s}, t)$ is weakly stationary if:

$$\mathbb{E}(Q(\mathbf{s},t)) = \mu(\mathbf{s},t) = \mu, \quad \forall (\mathbf{s},t) \in \mathbb{R}^2 \times \mathbb{R},$$

and

$$\operatorname{Cov}(Q(\mathbf{s}_i, t_i), Q(\mathbf{s}_i, t_i)) = \operatorname{C}(\mathbf{s}_j - \mathbf{s}_i, t_j - t_i), \quad \forall (\mathbf{s}, t) \in \mathbb{R}^2 \times \mathbb{R}$$

 $C(\mathbf{h})$ and $C(\mathbf{h}, u)$ are called the covariance function of $\{Q(\mathbf{s}), \mathbf{s} \in A\}$ and $\{Q(\mathbf{s}, t), (\mathbf{s}, t) \in \mathbb{R}^d \times \mathbb{R}\}$, respectively. Moreover, for a weakly stationary process, the correlation in the spatial and spatio-temporal case is defined as:

$$\operatorname{Corr}(Q(\mathbf{s}_i), Q(\mathbf{s}_j)) = \frac{\operatorname{C}(\mathbf{h})}{\operatorname{C}(\mathbf{0})} = \rho(\mathbf{h}),$$
$$\operatorname{Corr}(Q(\mathbf{s}_i, t_i), Q(\mathbf{s}_j, t_j)) = \frac{\operatorname{C}(\mathbf{h}, t)}{\operatorname{C}(\mathbf{0}, 0)} = \rho(\mathbf{h}, t),$$

respectively, where $C(\mathbf{0}) = Var(Q(\mathbf{s}))$ and $C(\mathbf{0}, 0) = Var(Q(\mathbf{s}, t))$. The next section shows the definition of a valid covariance function and some of its properties.

Definition 1.1.12. The spatio-temporal random field $Q(\mathbf{s}, t)$ is said to have a spatially stationary covariance function if, for any two pairs (\mathbf{s}_i, t_i) and (\mathbf{s}_j, t_j) on $\mathbb{R}^d \times \mathbb{R}$, the covariance function $C((\mathbf{s}_i, t_i), (\mathbf{s}_j, t_j))$ only depends on the distance between the locations $(\mathbf{s}_i - \mathbf{s}_j)$ and the times t_i and t_j .

Definition 1.1.13. The spatio-temporal random field $Q(\mathbf{s}, t)$ is said to have a temporally stationary covariance function if, for any two pairs (\mathbf{s}_i, t_i) and (\mathbf{s}_j, t_j) on $\mathbb{R}^d \times \mathbb{R}$, the covariance function $C((\mathbf{s}_i, t_i), (\mathbf{s}_j, t_j))$ only depends on the distance between the times $(t_i - t_j)$ and the spatial locations s_i and s_j .

Definition 1.1.14. If the spatio-temporal random field $Q(\mathbf{s}, t)$ has a stationary covariance function in both spatial and temporal terms, then it is said to have a stationary covariance function. In this case, the covariance function can be expressed as

$$C((\mathbf{s}_i, t_i), (\mathbf{s}_j, t_j)) = C(\mathbf{h}, u),$$

 $\mathbf{h} = \mathbf{s}_i - \mathbf{s}_j$ and $u = t_i - t_j$ being the distances in space and time, respectively.

Strict stationarity implies weak stationarity when the random field is a second order random field. In general the reverse is not true, but for Gaussian random fields, weak stationarity implies strict stationarity.

In the cases that the assumptions of stationarity may not be satisfied or the covariance function does not exist, Matheron (1965) introduced a weaker concept of stationarity. The hypothesis is that the increments of the process are weakly stationary.

Definition 1.1.15. A random field $\{Q(\mathbf{s}), \mathbf{s} \in A\}$ is intrinsic stationary if the increment process $Q(\mathbf{s} + \mathbf{h}) - Q(\mathbf{s})$ is weakly stationary for any fixed $\mathbf{h} \in \mathbb{R}^d$, *i.e.*,:

$$\mathbb{E}(Q(\mathbf{s} + \mathbf{h}) - Q(\mathbf{s})) = 0$$

and

$$\operatorname{Var}(Q(\mathbf{s} + \mathbf{h}) - Q(\mathbf{s})) = 2\gamma(\mathbf{h}).$$

If the random field is a second-order spatio-temporal random field, then it is intrinsic stationary if the increment process $Q(\mathbf{s} + \mathbf{h}, t + u) - Q(\mathbf{s}, t)$ is weakly stationary in space-time for any fixed $(\mathbf{h}, t) \in \mathbb{R}^d \times \mathbb{R}$, *i.e.*,

$$\mathbb{E}(Q(\mathbf{s} + \mathbf{h}, t + u) - Q(\mathbf{s}, t)) = 0$$

and

$$\operatorname{Var}(Q(\mathbf{s} + \mathbf{h}, t + u) - Q(\mathbf{s}, t)) = 2\gamma(\mathbf{h}, u).$$

The function $2\gamma(\mathbf{h})$ $(2\gamma(\mathbf{h}, u))$ is called variogram and $\gamma(\mathbf{h})$ $(\gamma(\mathbf{h}, u))$ is called semi-variogram or semi-variance. Moreover, if the random field is weakly stationary, then $\gamma(\mathbf{h}) = C(\mathbf{0}) - C(\mathbf{h})$ for a spatial random field and $\gamma(\mathbf{h}, u) = C(\mathbf{0}, 0) - C(\mathbf{h}, u)$ for a spatio-temporal random field. Thus, a weakly stationary random field with covariance function $C(\cdot)$ is intrinsic stationary with variogram $2\gamma(\mathbf{h}) = 2(C(\mathbf{0}) - C(\mathbf{h}))$ or $2\gamma(\mathbf{h}, u) =$ $2(C(\mathbf{0}, 0) - C(\mathbf{h}, u))$ in the spatio-temporal case. However, the converse is not true, but it holds in the case when the semi-variogram is bounden.

Definition 1.1.16. A spatio-temporal random field $Q(\mathbf{s}, t)$ is said to have an intrinsically stationary semi-variogram in space if, for any pair of spatio-temporal locations (\mathbf{s}_i, t_i) and (\mathbf{s}_i, t_i) on $\mathbb{R}^d \times \mathbb{R}$, the semi-variogram $\gamma((\mathbf{s}_i, t_i), (\mathbf{s}_j, t_j))$ only depends on $\mathbf{h} = \mathbf{s}_i - \mathbf{s}_j$ and the times t_i and t_j .

Definition 1.1.17. A spatio-temporal random field $Q(\mathbf{s}, t)$ is said to have an intrinsically stationary semi-variogram in time if, for any pair of spatio-temporal locations (\mathbf{s}_i, t_i) and (\mathbf{s}_i, t_i) on $\mathbb{R}^d \times \mathbb{R}$, the semi-variogram $\gamma((\mathbf{s}_i, t_i), (\mathbf{s}_j, t_j))$ only depends on $u = t_i - t_j$ and the spatial locations \mathbf{s}_i and \mathbf{s}_j .

Definition 1.1.18. If the spatio-temporal random field $Q(\mathbf{s}, t)$ has an intrinsically stationary semi-variogram in both space and time, then it is said to have an intrinsically stationary semi-variogram. In this case, the semivariogram can be expressed as

$$\gamma((\mathbf{s}_i, t_i), (\mathbf{s}_j, t_j)) = \gamma(\mathbf{h}, u),$$

 $\mathbf{h} = \mathbf{s}_i - \mathbf{s}_j$ and $u = t_i - t_j$ representing the distance in space and time, respectively. The constraints $\gamma(\cdot, u)$ and $\gamma(\mathbf{h}, \cdot)$ are called purely spatial and purely temporal semi-variograms, respectively.

As we outlined at the beginning of the section, an isotropic random field is invariant under rotations. Thus, we can define it as follows.

Definition 1.1.19. The random field $\{Q(\mathbf{s}), \mathbf{s} \in \mathbb{R}^d\}$ is isotropic if, for all R rotation of \mathbb{R}^d , the random field $\{Q(R\mathbf{s}), R\mathbf{s} \in \mathbb{R}^d\}$ has the same distribution than $Q(\mathbf{s})$, *i.e.*,

$$\Pr(Q(R\mathbf{s}_1) \le q_1, \dots, Q(R\mathbf{s}_n) \le q_n) = \Pr(Q(\mathbf{s}_1) \le q_1, \dots, Q(\mathbf{s}_n) \le q_n),$$

where $q_1, \ldots, q_n \in \mathbb{R}$.

Nevertheless, if $\{Q(\mathbf{s}), \mathbf{s} \in \mathbb{R}^d\}$ is a weakly stationary random field, then isotropy implies that the covariance function is invariant under rotations and depends only on the Euclidean distance $|| \cdot ||$. The covariance is called isotropic covariance and it is defined as follows.

Definition 1.1.20. A weakly stationary random field $\{Q(\mathbf{s}), \mathbf{s} \in \mathbb{R}^d\}$ has isotropic covariance if:

$$\forall i, j \in \mathbb{N}, \ \operatorname{Cov}(Q(\mathbf{s}_i), Q(\mathbf{s}_j)) = \operatorname{C}(||\mathbf{s}_j - \mathbf{s}_i||) = \operatorname{C}(||\mathbf{h}||).$$

Isotropy can be defined for spatio-temporal covariance functions as follows.

Definition 1.1.21. A stationary spatio-temporal random field $Q(\mathbf{s}, t)$ has a spatially isotropic covariance function if

$$C(\mathbf{h}, u) = C(\|\mathbf{h}\|, u), \quad \forall (\mathbf{s}, t) \in \mathbb{R}^d \times \mathbb{R}.$$

Definition 1.1.22. A stationary spatio-temporal random field $Q(\mathbf{s}, t)$ has a temporally isotropic (or symmetric) covariance function if

$$C(\mathbf{h}, u) = C(\mathbf{h}, |u|), \quad \forall (\mathbf{s}, t) \in \mathbb{R}^d \times \mathbb{R}.$$
Definition 1.1.23. A stationary spatio-temporal random field $Q(\mathbf{s}, t)$ has a isotropic covariance function in space and time if

$$C(\mathbf{h}, u) = C(\|\mathbf{h}\|, |u|), \quad \forall (\mathbf{s}, t) \in \mathbb{R}^d \times \mathbb{R}.$$

Even though the isotropy assumption could be violated in real data applications, the isotropic random fields form the basic building blocks for more sophisticated, anisotropic and nonstationary random field models. For instance, Zimmerman (1993) proposed geometrically anisotropic models which are built up from isotropic random fields, and Sampson and Guttorp (1992) developed nonstationary covariance models using deformation of the geographic coordinate space in order to obtain a new one where the covariance is isotropic.

1.1.3 Separability, full symmetry and compactly supported

Separability and full symmetry are interesting concepts in spatio-temporal modeling. The first one, permits more efficient inferences in terms of computing and the other one is motivated by atmospheric, environmental and geophysical processes (Gneiting, 2002c; Huang and Hsu, 2004; Stein, 2005).

Definition 1.1.24. A spatio-temporal random field $Q(\mathbf{s}, t)$ has a separable covariance function if there is a purely spatial covariance function $C_s(\mathbf{s}_i, \mathbf{s}_j)$ and a purely temporal covariance function $C_t(t_i, t_j)$ such that

$$C((\mathbf{s}_i, t_i), (\mathbf{s}_j, t_j)) = C_s(\mathbf{s}_i, \mathbf{s}_j)C_t(t_i, t_j)$$

for any pair of spatio-temporal locations (\mathbf{s}_i, t_i) and $(\mathbf{s}_j, t_j) \in \mathbb{R}^d \times \mathbb{R}$. If this breakdown is not possible, the covariance function will be called non-separable.

Definition 1.1.25. A spatio-temporal random field $Q(\mathbf{s}, t)$ has fully symmetric covariance function if

$$C((\mathbf{s}_i, t_i), (\mathbf{s}_j, t_j)) = C((\mathbf{s}_i, t_j), (\mathbf{s}_j, t_i))$$

for any pair of spatio-temporal locations (\mathbf{s}_i, t_i) and $(\mathbf{s}_j, t_j) \in \mathbb{R}^d \times \mathbb{R}$. If $Q(\mathbf{s}, t)$ is stationary, then fully symmetry condition reduces to

$$C(\mathbf{h}, u) = C(\mathbf{h}, -u) = C(-\mathbf{h}, u) = C(-\mathbf{h}, -u), \quad \forall (\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R}.$$

If the covariance function of a stationary spatio-temportal random field is isotropic in space and time, then it is fully symmetric. Note that separability is a particular case of full symmetry and, as such, any test to verify full symmetry can be used to reject separability. Moreover, under stationarity, verifying full symmetry comes down to confirming that $C(\mathbf{h}, u) = C(\mathbf{h}, -u)$ or $C(-\mathbf{h}, u) = C(-\mathbf{h}, -u)$.

Unlike spatio-temporal covariance function, separability does not make sense for semi-variograms because the product of semiv-ariograms does not assure a valid semi-variogram. On the other hand, the property of fully symmetry does not have this problem and it can be defined as follows.

Definition 1.1.26. A spatio-temporal random field $Q(\mathbf{s}, t)$ has fully symmetric variogram structure if

$$\operatorname{Var}(Q(\mathbf{s}_i, t_i) - Q(\mathbf{s}_j, t_2)) = \operatorname{Var}(Q(\mathbf{s}_i, t_j) - Q(\mathbf{s}_j, t_i))$$

for all locations (\mathbf{s}_i, t_i) and (\mathbf{s}_i, t_i) on $\mathbb{R}^d \times \mathbb{R}$.

Finally, we define the compactly supported covariance functions for spatio-temporal random fields.

Definition 1.1.27. A spatio-temporal rf has a compactly supported covariance function if, for any pair of spatio-temporal locations (\mathbf{s}_i, t_i) and $(\mathbf{s}_j, t_j) \in \mathbb{R}^d \times \mathbb{R}$, the covariance function $C((\mathbf{s}_i, t_i), (\mathbf{s}_j, t_j))$ tends towards zero when the spatial and/or temporal distance is sufficiently large.

Spatio-temporal random fields with a compactly supported covariance functions are attractive from a computing viewpoint, because they allow a computationally efficient estimation and prediction (Gneiting, 2002a; Bevilacqua et al., 2019).

1.1.4 Properties of covariance functions

Spatial covarince functions

The covariance functions and variograms are important for the analysis of spatial data. They allow us to study how spatial dependence varies with distance and play an important role in spatial prediction. Both covariance function and variogram may not be defined arbitrary, *i.e.*, not just any function can be a valid covariance function or variogram. Therefore, to be a valid covariance function or variogram, they need to be positive semi-definite and conditionally negative definite, respectively.

Definition 1.1.28. Let $\{Q(\mathbf{s}), \mathbf{s} \in A\}$ be a weakly stationary random field with covariance matrix $C = \{C(s_i, s_j)\}_{i,j}^n$. Then, C is positive semi-definite if:

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j \mathcal{C}(\mathbf{s}_i, \mathbf{s}_j) \ge 0$$
(1.1)

for any set of $\mathbf{s}_1, \ldots, \mathbf{s}_n$ and for all $a_1, \ldots, a_n \in \mathbb{R}$.

Since a second order random field is fully characterized by its moments, a necessary condition for their existence is that C is positive semi-definite. Therefore, positive semi-definiteness is a necessary and sufficient condition for the existence of a random field with finite second moments. In particular, this result holds for Gaussian random fields.

On the other hand, if a spatial process is weakly stationary, necessary and sufficient conditions for a valid covariance function is provided by Bochner's theorem (1933).

Theorem 1.1.2. (Bochner's Theorem). Let $\{Q(\mathbf{s}), \mathbf{s} \in A\}$ be a weakly stationary random field with $A = \mathbb{R}^d$. Then $C(\mathbf{h})$ is positive semi-definite if and only if it can be represented as:

$$C(\boldsymbol{h}) = \int e^{i\boldsymbol{\omega}^T \boldsymbol{h}} dF(\boldsymbol{\omega}), \qquad (1.2)$$

where F is a positive, symmetric, and finite measure and is called the spectral measure of C(h). If F is absolutely continuous with respect to Lebesgue measure, i.e., $dF(\boldsymbol{\omega}) = f(\boldsymbol{\omega})d\boldsymbol{\omega}$, then $f(\boldsymbol{\omega})$ is called the spectral density. Analogously, Schoenberg (1938) and later Yaglom (1987) showed that a valid isotropic covariance function can be characterized by the following theorem.

Theorem 1.1.3. For any $d \ge 2$, a function C(h) is a continuous isotropic covariance function of a weakly stationary random field on \mathbb{R}^d if and only if it can be represented as:

$$C(\mathbf{h}) = 2^{\frac{d-2}{2}} \Gamma\left(\frac{d}{2}\right) \int_0^\infty (\boldsymbol{\omega}||\mathbf{h}||)^{-\frac{d-2}{2}} J_{\frac{d-2}{2}}(\boldsymbol{\omega}||\mathbf{h}||) dG(\boldsymbol{\omega}),$$
(1.3)

where J_k is the Bessel function of the first kind of order k and the measure $G(\cdot)$ is nondecreasing bounded in \mathbb{R}^+ and $G(\mathbf{0}) = 0$.

A valid variogram must be conditionally negative definite. It is defined as follows.

Definition 1.1.29. Let $\{Q(\mathbf{s}), \mathbf{s} \in A\}$ be a intrinsic stationary random field with variogram 2γ . Let $\mathbf{s}_1, \ldots, \mathbf{s}_n$ be a set of locations and a_1, \ldots, a_n a set of real numbers such that $\sum_{i=1}^{n} a_i = 0$. Then, a variogram is conditionally negative definite if:

$$\sum_{i}^{n} \sum_{j}^{n} a_{i} a_{j} \gamma(\mathbf{s}_{i} - \mathbf{s}_{j}) \leq 0.$$
(1.4)

Note that if the random field is weakly stationary then C(h) or $\gamma(h)$ can be used interchangeably for inference or prediction. Additionally, Proposition 1.1.1 shows some properties of covariance function and semi-variogram.

Proposition 1.1.1. Let $\{Q(\mathbf{s}), \mathbf{s} \in A\}$ be a weakly stationary random field with covariance function $C(\mathbf{h})$ and semi-variogram $\gamma(\mathbf{h})$. Then, for $C(\mathbf{h})$ the following properties hold:

- (i) $\forall \mathbf{h} \in A, \mathbf{C}(\mathbf{0}) \geq 0.$
- (ii) $\forall \mathbf{h} \in A$, $|C(\mathbf{h})| \leq C(\mathbf{0}) = Var(Q(\mathbf{s}))$.
- (iii) C(h) = C(-h).
- (iv) If $C_i(\mathbf{h})$ is a valid covariance functions for i = 1, ..., k, then, if $b_i \ge 0$, $\forall j$, $\sum_{j=1}^k b_i C_j(\mathbf{h})$ is a valid covariance function.

- (v) If $C_i(\mathbf{h})$ is a valid covariance function for i = 1, ..., k, then $\prod_{i=1}^{k} C_i(\mathbf{h})$ is a valid covariance function.
- (vi) A valid covariance function in \mathbb{R}^d , it is also a valid covariance function in \mathbb{R}^p , with $p \leq d$.
- (vii) If C is continuous at the origin, then C is everywhere uniformly continuous.

Analogous for $\gamma(\mathbf{h})$ the following properties hold:

- (i) $\gamma(0) = 0$.
- (ii) $\forall \mathbf{h} \in A, \gamma(\mathbf{h}) = \gamma(-\mathbf{h}).$
- (iii) $\forall \mathbf{h} \in A, \gamma(\mathbf{h}) \geq 0.$
- (iv) If $\gamma_i(\mathbf{h})$ is a valid semi-variograms for i = 1, ..., k, then, if $b_i \ge 0$, $\forall j, \sum_{i=1}^{k} b_i \gamma_j(\mathbf{h})$ is a valid variogram.
- (v) If γ is continuous at 0, then γ is continuous at every site s where γ is locally bounded.
- (vi) If γ is bounded in a neighborhood of 0, $\exists a \text{ and } b \ge 0$ such that for any $\mathbf{h} \in A$, $\gamma(\mathbf{h}) \le a ||\mathbf{h}||^2 + b.$
- (vii) If T is a linear transformation in \mathbb{R}^d and γ a valid semi-variogram, then $\gamma(T\mathbf{h})$ is too.

Note that for a weakly stationary random field, the semi-variogram has a *sill* at height C(0) as $\|\mathbf{h}\| \to \infty$. It holds because if $C(\mathbf{h}) \to 0$ as $\|\mathbf{h}\| \to \infty$, then $\gamma(\mathbf{h}) \to C(\mathbf{0})$ as $\|\mathbf{h}\| \to \infty$. Moreover, the distance at which the semi-variogram reaches its sill is called the *range* and it is called *practical range* when 95% of the value of the sill is reached.

It is important to point out that the semi-variogram could be discontinuous at the origin, *i.e.*, for any $\mathbf{h} \neq \mathbf{0} \in A$, $\gamma(\mathbf{h}) \geq C(\mathbf{0}) > 0$. This phenomenon is called *nugget effect*, and represents the effect of measurement error and micro-scale variability.

Spatio-temporal covariance functions

The spatio-temporal covariance functions play an important roll for the analysis of spatio-temporal data. Therefore, we are interested in valid covariance functions on $\mathbb{R}^d \times \mathbb{R}$.

Definition 1.1.30. Let $\{Q(\mathbf{s},t), (\mathbf{s},t) \in \mathbb{R}^d \times \mathbb{R}\}$ be a second order spatio-temporal random field with covariance function $C((\mathbf{s}_i, t_i), (\mathbf{s}_j, t_j))$. A necessary and sufficient condition for a real-valued function $C((\mathbf{s}_i, t_i), (\mathbf{s}_j, t_j))$ defined on $\mathbb{R}^d \times \mathbb{R}$ to be a valid covariance function is for it to be symmetric, $C((\mathbf{s}_i, t_i), (\mathbf{s}_j, t_j)) = C((\mathbf{s}_j, t_j), (\mathbf{s}_i, t_i))$, and positive semi-definite, that is,

$$\sum_{i}^{n} \sum_{j}^{n} a_{i} a_{j} \mathcal{C}((\mathbf{s}_{i}, t_{i}), (\mathbf{s}_{j}, t_{j})) \ge 0$$

$$(1.5)$$

for any $n \in \mathbb{N}$, and for any $(\mathbf{s}_i, t_i) \in \mathbb{R}^d \times \mathbb{R}$ and $a_1, \ldots, a_n \in \mathbb{R}$.

If $\{Q(\mathbf{s},t) \text{ is stationary then the following definition is obtained.}$

Definition 1.1.31. Let $\{Q(\mathbf{s},t), (\mathbf{s},t) \in \mathbb{R}^d \times \mathbb{R}\}$ be a stationary spatio-temporal random field with covariance function $C(\mathbf{h}, u)$). A necessary and sufficient condition for a real-valued function $C(\mathbf{h}, u)$ defined on $\mathbb{R}^d \times \mathbb{R}$ to be a stationary covariance function is it to be an even function $(C(\mathbf{h}, u) = C(-\mathbf{h}, -u))$ and positive semi-definite, that is,

$$\sum_{i}^{n} \sum_{j}^{n} a_{i} a_{j} \mathcal{C}(\mathbf{h}_{i}, u_{i}) \ge 0$$
(1.6)

for any $n \in \mathbb{N}$, and for any $(\mathbf{h}_i, u_i) \in \mathbb{R}^d \times \mathbb{R}$ and $a_1, \ldots, a_n \in \mathbb{R}$

Moreover, Bochner's theorem allows us to characterize the stationary space-time covariance functions as follow.

Theorem 1.1.4 (Bochner's Theorem). A function $C(\mathbf{h}, u)$ remains defined on $\mathbb{R}^d \times \mathbb{R}$ and is a stationary covariance function if, and only if, it has the following form:

$$C(\mathbf{h}, u) = \int \int e^{i(\mathbf{w}^{\top}\mathbf{h} + vu)} dF(\mathbf{w}, v), \quad (\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R},$$
(1.7)

where the function F is a non-negative finite measure defined on $\mathbb{R}^d \times \mathbb{R}$, which is known as a spectral measure.

If C is integrable, then the spectral measure F is absolutely continuous and the representation (1.7) can be written as follows

$$C(\mathbf{h}, u) = \int \int e^{i(\mathbf{w}^{\top}\mathbf{h} + vu)} f(\mathbf{w}, v) d\mathbf{w} dv, \quad (\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R},$$

where f is a non-negative, continuous and integrable function that is known as a spectral density function. Note that the covariance function and the spectral density function form a pair of Fourier transforms, and

$$f(\mathbf{w}, v) = (2\pi)^{-(d+1)} \int \int e^{-i(\mathbf{w}^{\top}\mathbf{h} + vu)} \mathcal{C}(\mathbf{h}, u) d\mathbf{h} du.$$

If the spatio-temporal covariance function is fully symmetric, then Theorem 1.1.4 can be can be specialized as follows

Theorem 1.1.5. Let $C(\mathbf{h}, u)$ be a continuous function defined on $\mathbb{R}^d \times \mathbb{R}$, then $C(\mathbf{h}, u)$ is a fully symmetric stationary covariance function if, and only if, it is of the form:

$$C(\mathbf{h}, u) = \int \int \cos(\mathbf{w}^{\top} \mathbf{h}) \cos(vu) dF(\mathbf{w}, v), \quad (\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R},$$
(1.8)

where the function F is a non-negative finite measure defined on $\mathbb{R}^d \times \mathbb{R}$, which is known as a spectral distribution function. If the spectral density function f also exists, then f is fully symmetric, that is:

$$f(\mathbf{w}, v) = f(\mathbf{w}, -v) = f(-\mathbf{w}, v) = f(-\mathbf{w}, -v), \quad \forall (\mathbf{w}, v) \in \mathbb{R}^d \times \mathbb{R}.$$

Following the definition 1.1.24 we can state that a stationary spatio-temporal covariance function $C(\mathbf{h}, u)$ was separable if there were two stationary covariance functions $C_s(\mathbf{h})$ and $C_t(u)$ that are purely spatial and purely temporal, respectively, such that

$$C(\mathbf{h}, u) = C_s(\mathbf{h})C_t(u), \quad \forall (\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R}.$$

Moreover, applying Theorem 1.1.4 we can obtain that the spectral measure of a stationary separable covariance function can be written as the product of a spectral measure on the spatial domain and a spectral measure on the temporal domain. If the spectral density

function exists, then it also can be expressed as the product of spectral density functions on the respective domains.

Although separable spatio-temporal covariance functions have some advantages, there are some cases that they may not represent a physical phenomena. This is why non-separable spatio-temporal covariance functions have been sought. Cressie and Huang (1999) provide a characterization of the class of stationary spatio-temporal covariance functions under the additional hypothesis of integrability. Then, they use this characterization to construct non-separable stationary space-time covariance. The main result of their research is given by the following theorem.

Theorem 1.1.6. Let $C(\mathbf{h}, u)$ be a stationary spatio-temporal covariance function. Suppose that $C(\mathbf{h}, u)$ is continuous and integrable. Then, $C(\mathbf{h}, u)$ is a positive semi-definite continuous spatio-temporal stationary covariance function on $\mathbb{R}^d \times \mathbb{R}$ if, and only if, it has the following form:

$$C(\mathbf{h}, u) = \int e^{i\mathbf{w}^{\top}\mathbf{h}} \rho(\mathbf{w}, u) k(\mathbf{w}) d\mathbf{w}, \quad (\mathbf{h}, u) \in \mathbb{R}^{d} \times \mathbb{R}$$

where for each $\mathbf{w} \in \mathbb{R}^d$, $\rho(\cdot, u)$ is a continuous correlation function, $\int \rho(\mathbf{w}, u) du < \infty$, $k(\mathbf{w}) > 0$ and $\int k(\mathbf{w}) d\mathbf{w} < \infty$.

For instance, if we set $\rho(\mathbf{w}, u) = \exp(\|\mathbf{w}\|^2 u^2/4)$ and $k(\mathbf{w}) = \exp(-\sigma^2 \|\mathbf{w}\|^2/4)$, $\sigma^2 > 0$. Then, a three-paremeter non-separable spatio-temporal stationary covariance function on $\mathbb{R}^d \times \mathbb{R}$ is given as,

$$C(\mathbf{h}, u) = \frac{\sigma^2}{(a^2 u^2 + 1)^{d/2}} \exp\left(-\frac{b^2 \|\mathbf{h}\|}{a^2 u^2 + 1}\right),$$

where a > 0 is the scaling parameter of time, b > 0 is the scaling parameter of space, and $\sigma^2 = C(\mathbf{0}, 0)$.

Later on, Gneiting (2002b) proposed a method that is based on this construction, but does not depend on Fourier inversion, and De Iaco et al. (2002) and Ma (2005) constructed fully symmetric stationary space-time covariance functions by mixtures of separable covariances.

Finally, we generalize some properties of the spatial covariance function to the spatio-temporal case.

Proposition 1.1.2. Let $\{Q(\mathbf{h}, u), (\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R}\}$ be a stationary random field with covariance function $C(\mathbf{h}, u)$. Then, for $C(\mathbf{h}, u)$ the following properties hold:

(i) If $C_1(\mathbf{h}, u)$ is a covariance function defined on $\mathbb{R}^d \times \mathbb{R}$ and b > 0, then

 $C(\mathbf{h}, u) = bC_1(\mathbf{h}, u)$

is also a covariance function defined on $\mathbb{R}^d \times \mathbb{R}$.

(ii) Let $C_1(\mathbf{h}, u)$ and $C_2(\mathbf{h}, u)$ be two covariance functions defined on $\mathbb{R}^d \times \mathbb{R}$, then

 $C(\mathbf{h}, u) = C_1(\mathbf{h}, u) + C_2(\mathbf{h}, u)$

is a covariance function defined on $\mathbb{R}^d \times \mathbb{R}$.

(iii) Let $C_1(\mathbf{h}, u)$ and $C_2(\mathbf{h}, u)$ be two covariance functions defined on $\mathbb{R}^d \times \mathbb{R}$, then

 $C(\mathbf{h}, u) = C_1(\mathbf{h}, u)C_2(\mathbf{h}, u)$

is a covariance function defined on $\mathbb{R}^d \times \mathbb{R}$ *.*

(iv) In the case of a stationary spatial covariance function $C_s(\mathbf{h})$ defined on \mathbb{R}^d , and a stationary temporal covariance function $C_t(u)$ defined on \mathbb{R} , the functions

$$C(\mathbf{h}, u) = C_s(\mathbf{h} + \boldsymbol{\theta} u),$$

$$C(\mathbf{h}, u) = C_t(u + \boldsymbol{\theta}^\top \mathbf{h}),$$

with $\boldsymbol{\theta} \in \mathbb{R}^d$, are stationary covariance functions on $\mathbb{R}^d \times \mathbb{R}$.

(v) In the spatio-temporal case, the nugget effect can be exclusively spatial, exclusively temporal, or spatio-temporal and will therefore be given by

$$C(\mathbf{h}, u) = a \mathbb{1}_{(\mathbf{0}, 0)}(\mathbf{h}, u) + b \mathbb{1}_{(\mathbf{0})}(\mathbf{h}) + c \mathbb{1}_{(0)}(u),$$
(1.9)

where a, b, c are non-negative constants and $\mathbb{1}_A(\cdot)$ is an indicator function.

(vi) The product and sum of continuous spatio-temporal covariance functions with a nugget effect of the type (1.9), yield valid covariance models.

1.1.5 Isotropic parametric models

We now introduce the family of isotropic parametric covariance models and describe some members of this family. We can define a general form of an isotropic parametric covariance function by:

$$C(\mathbf{h}, \boldsymbol{\psi}) = \begin{cases} C(0)\rho(\|\mathbf{h}\|, \boldsymbol{\psi}), & \|\mathbf{h}\| > 0\\ C(0) + \tau^2, & \|\mathbf{h}\| = 0 \end{cases},$$
(1.10)

where τ^2 represents the nugget effect and $\rho(\|\mathbf{h}\|, \boldsymbol{\psi})$ is a parametric correlation function which depends on the parameter vector $\boldsymbol{\psi} \in \Psi \subseteq \mathbb{R}^p$. From the extensive list of correlation models in the literature, some of the most popular are the following:

• The powered exponential family (Diggle et al., 1998, among others) is defined by:

$$\rho(\|\mathbf{h}\|, \alpha, p) = \exp\left(-\left(\frac{\|\mathbf{h}\|}{\alpha}\right)^p\right), \quad 0 (1.11)$$

where $\alpha > 0$ is a spatial scale parameter. If p = 1, the exponential correlation model is obtained, while the Gaussian correlation model arises when p = 2. Moreover, the sample paths of a Gaussian random field are infinitely differentiable when p = 2and not differentiable at all when p < 2.

• The Matérn correlation model (Matèrn, 1986) is defined by:

$$\mathcal{M}_{\nu,\alpha}(\|\mathbf{h}\|) = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\|\mathbf{h}\|}{\alpha}\right)^{\nu} \mathcal{K}_{\nu}\left(\frac{\|\mathbf{h}\|}{\alpha}\right), \quad \|\mathbf{h}\| \ge 0, \quad (1.12)$$

where, \mathcal{K}_v is a modified Bessel function of the second kind of order v > 0 and $\alpha > 0$ (range) a spatial scale parameter. If v = 1/2, the exponential correlation model is obtained, while the Gaussian correlation model arises when $v \to \infty$. Moreover, for a positive integer k, the sample paths of a Gaussian random field are k times differentiable if and only if v > k. • The compactly supported Generalized Wendland correlation model (Gneiting, 2002a) is defined by:

$$\mathcal{GW}_{\zeta,\delta,\alpha}(\|\mathbf{h}\|) = \begin{cases} \frac{1}{B(2\zeta,\delta+1)} \int_{\|\mathbf{h}\|/\alpha}^{1} u \left(u^2 - \frac{\|\mathbf{h}\|^2}{\alpha^2}\right)^{\zeta-1} (1-u)^{\delta} du & \|\mathbf{h}\| < \alpha, \\ 0 & \text{otherwise} \end{cases}$$
(1.13)

where $B(\cdot, \cdot)$ is the beta function, $\zeta \ge 0$, $\delta > (d+1)/2 + \zeta$ and $\alpha > 0$ is the spatial compact support. The integral in Equation (1.13) has a closed form solution when ζ is a positive integer. For instance, if $\zeta = 0$ the correlation model is $\mathcal{GW}_{0,\delta,\alpha}(h) =$ $(1 - ||h||/\alpha)^{\delta}_{+}$ where $(\cdot)_{+}$ denotes the positive part, defined as Askey function. Moreover, for a positive integer k, the sample paths of a Gaussian random field are k times differentiable if and only if $\zeta > k - 1/2$.

Matérn and Generalized Wendland correlation models are more flexible than the powered exponential family. Additionally, Generalized Wendland correlation model is compactly supported. Thus, the covariance matrix is sparse and sparse matrix algorithms can then be used to evaluate efficiently an approximate likelihood (Bevilacqua et al., 2019). Hence, compactly supported is an interesting feature from a computational point of view because it reduces the computational burden.

1.1.6 Continuity of the sample paths

Continuity of the sample paths is a basic geometrical property of a random field and it has implications for the smoothness of random field realizations. We look at three types of continuity, continuity in probability, almost surely continuity and mean square continuity.

Definition 1.1.32. Let $\{Q(\mathbf{s}), \mathbf{s} \in A\}$ be a random field with $A = \mathbb{R}^d$.

1. A random field $Q(\mathbf{s})$ has continuous sample path with probability one in A if for all $\omega \in \Omega$ and every \mathbf{s}_0 for which $\|\mathbf{s}_0 - \mathbf{s}\| \to 0$ as $\mathbf{s} \to \mathbf{s}_0$, then

$$\Pr(\omega: |Q(\mathbf{s}_0, \omega) - Q(\mathbf{s}, \omega)| \to 0 \text{ as } \mathbf{s} \to \mathbf{s}_0) = 1 \text{ for all } \mathbf{s} \in A.$$

2. A random field $Q(\mathbf{s})$ is almost surely continuous in A if for almost all $\omega \in \Omega$ and every \mathbf{s}_0 for which $\|\mathbf{s}_0 - \mathbf{s}\| \to 0$ as $\mathbf{s} \to \mathbf{s}_0$, then

$$\Pr(\omega: |Q(\mathbf{s}_0, \omega) - Q(\mathbf{s}, \omega)| \to 0 \text{ as } \mathbf{s} \to \mathbf{s}_0) = 1 \text{ for all } \mathbf{s} \in A$$

3. A random field $Q(\mathbf{s})$ is mean square continuous in A if for every \mathbf{s}_0 for which $||\mathbf{s}_0 - \mathbf{s}|| \to 0$ as $\mathbf{s} \to \mathbf{s}_0$, then

$$E\{[Q(\mathbf{s}) - Q(\mathbf{s}_0)]^2\} \to 0 \text{ as } \mathbf{s} \to \mathbf{s}_0 \text{ for all } \mathbf{s} \in A$$

In general, one form of continuity does not imply the other. However, if a random field is a bounded process, then almost surely continuity implies mean square continuity.

Under some assumptions there is a relationship between mean square continuity and the convariance function of a random field. In fact, Stein (1999) shows that, for a weakly stationary random field, mean square continuity is equivalent to the covariance function C(h) being continuous at 0, *i.e.*,

$$\lim_{\mathbf{h}\to 0} C(\mathbf{h}) = C(\mathbf{0}).$$

Therefore, the nugget effect, which is noted by τ^2 , implies non mean square continuity because $C(\mathbf{h}) \rightarrow \tau^2$ as $\mathbf{h} \rightarrow \mathbf{0}$. Hence, the path of random field is not smooth.

1.1.7 Modeling spatial and spatio-temporal data

Modeling spatial data

The classical spatial random field model is specified by:

$$Y(\mathbf{s}) = \mu(\mathbf{s}) + Q(\mathbf{s}), \tag{1.14}$$

where $\mu(\mathbf{s})$ is a deterministic and continuous function and represents the mean of the response $Y(\mathbf{s})$. A common specification of $\mu(\mathbf{s})$ is the parametric form $g(X(\mathbf{s})^{\top}\boldsymbol{\beta})$, where $g(\cdot)$ is a continuous function, $X(\mathbf{s})$ is a *p*-dimensional vector of explanatory variables at

location s and β is a *p*-dimensional vector of parameters. Moreover, Q(s) is a zero-mean random field with a parametric covariance.

In addition, for capturing the micro-scale spatial variation and measurement error that may occur in the data collection process, *i.e.*, nugget effect, we can incorporate an error component e(s) to model (1.14). This error component typically has no spatial structure and can be assumed as a zero-mean pure error process with variance equal to τ^2 . Hence, the model (1.14) can be rewritten as follows:

$$Y(\mathbf{s}) = \mu(\mathbf{s}) + Q(\mathbf{s}) + e(\mathbf{s}). \tag{1.15}$$

A typical assumption for $Q(\mathbf{s})$ is Gaussianity, *i.e.*, considering Gaussian random fields. However, in some cases, the observed data cannot be modeled using a Gaussian random field, because it is discrete. In fact, in environmental analysis, population genetics, epidemiology and aquaculture, among other fields, the observed data are in general count data. To solve this issue, non-Gaussian data with spatial dependence are analysed using generalized linear mixed models (GLMM), where the spatial dependence is captured by a Gaussian randon field effect (see Breslow and Clayton, 1993). Later on, Diggle et al. (1998) proposed hierarchical models by assuming that, $Q(\mathbf{s})$ is a Gaussian random field and conditional on $Q(\mathbf{s})$, the $Y(\mathbf{s}_i)$ are mutually independent with a distribution belonging to the exponential family, conditional means $\mathbb{E}[Y(\mathbf{s}|Q(\mathbf{s}))] = g(Q(\mathbf{s}))$ and conditional variances τ^2 .

It is important to stress that this type of model induces a discontinuity in the path (De Oliveira, 2013). Consequently, samples located nearby are more dissimilar in value than in the case when the correlation function is continuous at the origin (Morgan, 2005).

In contrast to hierarchical models, Gaussian copula models allow us to model specific marginal distributions taking into account specific correlation structures. For example, Han and De Oliveira (2016) described a class of random field models for geo-statistical count data based on Gaussian copulas. Similarly, Kazianka and Pilz (2010) proposed geo-statistical copula-based models that are able to deal with random fields having discrete marginal distributions. Unfortunately, there are cases in which the copula representation

for discrete distributions is not unique, so it is unidentifiable (Genest and Neslehova, 2007).

Modeling spatio-temporal data

We can generalize the classical spatial random field model to spatio-temporal case by

$$Y(\mathbf{s},t) = \mu(\mathbf{s},t) + Q(\mathbf{s},t), \tag{1.16}$$

where $(\mathbf{s}, t) \in \mathbb{R}^d \times \mathbb{R}$, $\mu(\mathbf{s}, t)$ is a deterministic space-time trend function and $Q(\mathbf{s}, t)$ a zero-mean spatio-temporal random field.

Even if spatio-temporal random fields are defined on $\mathbb{R}^d \times \mathbb{R}$, there are others domains of interest in real data. For instance, the time domain can be considered discrete, *i.e.*, $(\mathbf{s},t) \in \mathbb{R}^d \times \mathbb{Z}$. Storvik et al. (2002) proposed time autoregressive Gaussian models to consider a discrete time domain, and Stein (2005) used a temporal Markov structure. On the other hand, the global curvature of the earth is a feature that must be taken into account in applications related with atmospheric and geophysical data. Therefore, the space domain has to considered on a sphere, *i.e.*, the domain could be $\mathbb{S}^d \times \mathbb{R}$ or $\mathbb{S}^d \times \mathbb{Z}$. Parametric covariance models on global spatial or spatio-temporal domains are proposed by Gneiting (1999), Stein (2005) and more recently by Jeong and Jun (2015), Porcu et al. (2016) and Berg and Porcu (2017).

The deterministic space-time trend function $\mu(\mathbf{s}, t)$, can be parametrized through a continuous function of $X(\mathbf{s}, t)^{\top}\beta$, where $X(\mathbf{s}, t)$ is a *p*-dimensional vector of explanatory variables at spatio-temporal location (\mathbf{s}, t) and β is a *p*-dimensional vector of parameters. The simplest case occurs when $\mu(\mathbf{s}, t)$ is decomposed as the sum of a purely spatial and a purely temporal trend component. Temporal trends are often periodic and can be modeled with trigonometric functions or non-parametric alternatives.

The spatio-temporal random fiel Q(s, t) is usually assumed weakly stationary and isotropic, but non-stationary and anisotropic models can be considered. (see Porcu et al., 2006; Schlather, 2010, for instance).

Similar to the spatial case, the nugget effect can be modeled by an error component $e(\mathbf{s}, t)$ with covariance function of the type (1.9), that is

$$C_{e}(\mathbf{h}, u) = \tau_{st}^{2} \mathbb{1}_{(\mathbf{0}, 0)}(\mathbf{h}, u) + \tau_{s}^{2} \mathbb{1}_{\mathbf{0}}(\mathbf{h}) + \tau_{t}^{2} \mathbb{1}_{0}(u), \qquad (1.17)$$

where $\tau_{st}^2, \tau_s^2, \tau_t^2$ are non-negative constants and represent the spatio-temporal, the purely spatial and a purely temporal nugget, respectively. Then, the spatio-temporal model can be written as follows

$$Y(\mathbf{s},t) = \mu(\mathbf{s},t) + Q(\mathbf{s},t) + e(\mathbf{s},t).$$
(1.18)

Finally, for the non-Gaussian space-random field, the models based on the hierarchical and copula approach, which are described in Section 1.1.7, can be used if we consider a space-time domain instead of a spatial domain.

1.1.8 Kriging

One of the primary goals of spatial and spatio-temporal statistics modeling is to make predictions at spatial locations without observations. Kriging aims to predict the value of a random field at one or more non-observed points from a observed data at n spatial locations, and provides the best linear unbiased predictor (BLUP).

Let $\{Y(\mathbf{s}), \mathbf{s} \in A \subset \mathbb{R}^d\}$ be a random field with the following model assumption :

$$Y(\mathbf{s}) = \mu(\mathbf{s}) + Q(\mathbf{s}),$$

where $\mu(\mathbf{s})$ is a deterministic function and $\{Q(\mathbf{s}), \mathbf{s} \in A \subset \mathbb{R}^d\}$ is a zero-mean weakly stationary random field. Let $\mathbf{s}_1, \ldots, \mathbf{s}_n$ be *n* known spatial locations and $\mathbf{Y} = (Y(\mathbf{s}_1), \ldots, Y(\mathbf{s}_n))^\top$ a vector of observed data at the spatial locations, and $\widehat{Y}(\mathbf{s}_0)$ a predictor at a spatial location \mathbf{s}_0 . Then, the kriging predictor minimizes the mean-squared prediction error, $\mathbb{E}[(Y(\mathbf{s}_0) - \widehat{Y}(\mathbf{s}_0))^2]$, and satisfies the following properties:

•
$$\widehat{Y}(\mathbf{s}_0) = \sum_{i=1}^n \lambda_i Y(\mathbf{s}_i).$$

• $E[\widehat{Y}(\mathbf{s}_0)] = E[Y(\mathbf{s}_0)] = \mu(\mathbf{s}_0)$

Hence, we must find $\lambda_1, \ldots, \lambda_1$ that minimizes $\mathbb{E}[(Y(\mathbf{s}_0) - \widehat{Y}(\mathbf{s}_0))^2]$ such as $\sum_{i=1}^n \lambda_i = 1$. If $Q(\mathbf{s})$ is a Gaussian random field, $\widehat{Y}(\mathbf{s}_0)$ is exactly the conditional expectation $\mathbb{E}[Y(\mathbf{s}_0)|\mathbf{Y}]$, *i.e.*, the optimal predictor (in the mean square sense). Therefore, the kriging predictor matches the optimal predictor.

Kriging also makes assumptions about the mean of the random field. Depending on this assumption, numerous variants of kriging are available. We describe the cases when μ , which is unknown, is constant (ordinary kriging) or not (universal kriging).

1. Ordinari Kriging (OK).

Let $\mu(\mathbf{s}) = \mu$ be a unknown constant and $C = [C(\mathbf{s}_i - \mathbf{s}_j)]_{i,j=1}^n$ the known covariance matrix of the random field $Q(\mathbf{s})$. Note that in this case the unbiased property implies that $\sum_{i=1}^n = 1$. Then, the kriging predictor $\widehat{Y}(\mathbf{s}_0)_{OK}$ is:

$$\widehat{Y}(\mathbf{s}_0)_{\mathrm{OK}} = \widehat{\mu} + \boldsymbol{c}^\top C^{-1} (\boldsymbol{Y} - \mathbf{1}\widehat{\mu}),$$

where $\boldsymbol{c} = [C(\mathbf{s}_0 - \mathbf{s}_i)]_{i=1}^n$, and the ordinary kriging variance is

$$\sigma^2(\mathbf{s}_0)_{\mathrm{OK}} = \mathrm{C}(\mathbf{0}) - \left(\boldsymbol{c} + \mathbf{1}\frac{(1-\mathbf{1}^{\top}C^{-1}\boldsymbol{c})}{\mathbf{1}^{\top}C^{-1}\mathbf{1}}\right)^{\top}C^{-1}\boldsymbol{c} + \frac{1-\mathbf{1}^{\top}C^{-1}\boldsymbol{c}}{\mathbf{1}^{\top}C^{-1}\mathbf{1}}.$$

2. Universal Kriging (UK).

Let $\mu(\mathbf{s}) = X(\mathbf{s})^{\top} \boldsymbol{\beta}$ be a unknown deterministic function where $X(\mathbf{s})$ and is $\boldsymbol{\beta}$ are *p*-dimensional vectors, $X = [X(\mathbf{s}_i)]$ a $n \times p$ matrix and $C = [C(\mathbf{s}_i - \mathbf{s}_j)]_{i,j=1}^n$ the known covariance matrix of the random field $Q(\mathbf{s})$. Note that, in this case, the unbiased property implies that $\boldsymbol{\lambda}^{\top} X = X(\mathbf{s}_0)$. Then, the kriging predictor $\widehat{Y}(\mathbf{s}_0)_{\text{UK}}$ is:

$$\widehat{Y}(\mathbf{s}_0)_{\mathrm{OK}} = X(\mathbf{s}_0)^\top \widehat{\boldsymbol{\beta}} + \boldsymbol{c}^\top C^{-1} (\boldsymbol{Y} - X(\mathbf{s}_0)^\top \widehat{\boldsymbol{\beta}}),$$

where $\boldsymbol{c} = [C(\mathbf{s}_0 - \mathbf{s}_i)]_{i=1}^n$, and the universal kriging variance is

$$\sigma^2(\mathbf{s}_0)_{\mathrm{UK}} = \mathrm{C}(\mathbf{0}) - \boldsymbol{c}^{\top} C^{-1} \boldsymbol{c} + (X(\mathbf{s}_0) - X^{\top} C^{-1} \boldsymbol{c})^{\top} (X^{\top} C^{-1} X)^{-1} (X(\mathbf{s}_0) - X^{\top} C^{-1} \boldsymbol{c})$$

The term, $C(\mathbf{0}) - \mathbf{c}^{\top} C^{-1} \mathbf{c}$, corresponds to the mean squared error of the BLUP, and the last term is the penalty for having to estimate β .

1.2 Renewal processes

Renewal processes play an important roll in many areas such as system analysis in queueing theory, physical modeling, seismology, among others. Firstly, we introduce the concept of counting process by the following definition.

Definition 1.2.1. A stochastic process $\{N(t), t \ge 0\}$ is called a counting process if N(t) represents the total number of events that have occurred up to time t.

Thus a formal definition of a renewal process is given as follows.

Definition 1.2.2. Let $\{X_n, n = 1, 2, ...\}$ be a sequence of independently and identically distributed non-negative random variables with distribution function $F(\cdot)$. We call these inter-arrival times or waiting times. Letting

$$S_0 = 0, \quad S_n = \sum_{i=1}^n X_i, \quad n \ge 1,$$

it follows that S_n is the time of the *n*th event and we call them renewal times. Thus, the counting process $\{N(t), t \ge 0\}$ defined by

$$N(t) := \begin{cases} 0 & \text{if } 0 \le t < S_1 \\ \max_{n \ge 1} \{S_n \le t\} & \text{if } S_1 \le t \end{cases},$$
(1.19)

is called a renewal process.

Note that $N(t) \ge n$ if only if $S_n \le t$, then the probability mass function of N(t) can be obtained as follows:

$$Pr(N(t) = n) = Pr(N(t) \ge n) - Pr(N(t) \ge n+1)$$
$$= Pr(S_n \le t) - Pr(S_{n+1} \le t)$$
$$= F_n(t) - F_{n+1}(t)$$

where $F_n(t)$ are the *n*-fold convolution of the cumulative distribution functions F of the interarrival times.

Definition 1.2.3 (Renewal function). For a renewal process $\{N(t); t \ge 0\}$,

$$m(t) = \mathbb{E}(N(t))$$

is called as a renewal function. Moreover, it can be shown that

$$m(t) = \sum_{n=1}^{\infty} F_n(t).$$

Now, we introduce the definition of bivariate renewal process given by Hunter (1974).

Definition 1.2.4. Let $\{(X_n, Y_n), n = 1, 2, ...\}$ be a sequence of independently and identically distributed non-negative random variables with common joint distribution function $F(x, y) = \Pr(X_n \le x, Y_n \le y)$. Let

$$S_n = (S_n^{(1)}, S_n^{(2)}) = \left(\sum_{i=1}^n X_i, \sum_{i=1}^n Y_i\right).$$

Define

$$N(t_1) := \begin{cases} 0 & \text{if } 0 \le t_1 < S_1^{(1)} \\ \max_{n \ge 1} \{S_n^{(1)} \le t_1\} & \text{if } S_1^{(1)} \le t_1 \end{cases}, \quad N(t_2) := \begin{cases} 0 & \text{if } 0 \le t_2 < S_1^{(2)} \\ \max_{n \ge 1} \{S_n^{(2)} \le t_2\} & \text{if } S_1^{(2)} \le t_2 \end{cases}$$

Then, the random pair $(N(t_1), N(t_2))$ is called the bivariate renewal process.

The joint distribution of $(N(t_1), N(t_2))$ is given by (Hunter, 1974):

$$\Pr(N(t_1) = n, N(t_2) = m) = \begin{cases} [F_0 - F^1 - F^2 + F] * *F_n(\lambda(\mathbf{s}_i), \lambda(\mathbf{s}_j)) & \text{if } n = m \\ [F_r^1 - F_{r+1}^1 - F_{r-1}^1 * *F + F_r^1 * *F] * *F_m(t_1, t_2) & \text{if } n > m, \\ n = m + r, \\ [F_r^2 - F_{r+1}^2 - F_{r-1}^2 * *F + F_r^2 * *F] * *F_n(t_1, t_2) & \text{if } n < m, \\ m = n + r \end{cases}$$

where ** is the double convolution, F^1 and F^2 are the probability mass function of $N(t_1)$ and $N(t_2)$, respectively, and F_n is the *n*-fold bivariate convolution of F. The moments of $(N(t_1), N(t_2))$ for low orders are given by:

$$\begin{split} \mathbb{E}[N(t_1)] &= \sum_{i=1}^{\infty} F_i^1(t_1), \\ \mathbb{E}[N(t_2)] &= \sum_{j=1}^{\infty} F_j^2(t_2), \\ \mathbb{E}[N(t_1)N(t_2)] &= \left(F_0 + \sum_{i=1}^{\infty} F_i^1(t_1) + \sum_{j=1}^{\infty} F_j^2(t_2)\right) * * \left(\sum_{r=1}^{\infty} F_r(t_1, t_2)\right) \\ \operatorname{Cov}(N(t_1), N(t_2)) &= \left(F_0 + \sum_{i=1}^{\infty} F_i^1(t_1) + \sum_{j=1}^{\infty} F_j^2(t_2)\right) * * \left(\sum_{r=1}^{\infty} F_r(t_1, t_2)\right) - \sum_{i=1}^{\infty} F_i^1(t_1) \sum_{j=1}^{\infty} F_j^2(t_2), \end{split}$$

where $F_0 = 1$. Finally, we show an interesting result concerning the independence of the renewal processes $N(t_1)$ and $N(t_2)$.

Theorem 1.2.1. The following conditions are equivalent.

(i) X_1 and Y_1 are independent.

(ii)
$$Cov(N(t_1), N(t_2)) = 0$$
 for all $t_1 \ge 0, t_2 \ge 0$.

(iii) $N(t_1)$ and $N(t_2)$ are independent for all $t_1 \ge 0, t_2 \ge 0$.

1.2.1 Poisson process

The Poisson process is a particular case of a renewal process, where inter-arrival times are exponential random variables with some parameter. Moreover, a Poisson process can be homogeneous or non-homogeneous.

Definition 1.2.5 (Homogeneous Poisson process). Let $\{N(t), t \ge 0\}$ be a renewal process. If inter-arrival times are exponential random variables with parameter λ , then $\{N(t), t \ge 0\}$ is a homogeneous Poisson process with rate (or intensity) $\lambda, \lambda > 0$.

Another definition can be given using an axiomatic way as follows:

Definition 1.2.6. A counting process $\{N(t), t \ge 0\}$ is said to be a homogeneous Poisson process with rate (or intensity) $\lambda, \lambda > 0$, if:

(i) N(0) = 0.

- (ii) The process has independent increments.
- (iii) The number of events in any interval is Poisson distributed such that the probability that k events will occur in the time interval (s, s + t] will be

$$\Pr(N(s+t) - N(s) = k) = \Pr(N(t) = k) = \frac{e^{-\lambda t} (\lambda t)^k}{k!},$$

and

$$\mathbb{E}[N(s+t) - N(s)] = \mathbb{E}[N(t)] = \lambda t,$$

$$\operatorname{Var}(N(s+t) - N(s)) = \operatorname{Var}(N(t)) = \lambda t.$$

Note that in this case the renewal times have Erlang distribution with cumulative distribution function given by:

$$F_n(t) = 1 - \sum_{k=0}^{n-1} \frac{e^{-\lambda t} (\lambda t)^k}{k!}, \quad t > 0$$

Moreover, the renewal function is linear in time:

$$m(t) = \lambda t.$$

Definition 1.2.7 (Non-Homogeneous Poisson process). A counting process $\{N(t), t \ge 0\}$ is said to be a non-homogeneous Poisson process with rate (or intensity) $\lambda(t), t > 0$, if:

(i) N(0) = 0.

(ii) The process has independent increments.

(iii) The number of events in any interval is Poisson distributed such that the probability that k events will occur in the time interval (s, s + t] will be

$$\Pr(N(s+t) - N(s) = k) = \Pr(N(t) = k) = \frac{e^{-\int_{s}^{s+t} \lambda(y)dy} \left(\int_{s}^{s+t} \lambda(y)dy\right)^{k}}{k!},$$

and

$$\mathbb{E}[N(s+t) - N(s)] = = \int_{s}^{s+t} \lambda(y) dy,$$
$$\operatorname{Var}(N(s+t) - N(s)) = \int_{s}^{s+t} \lambda(y) dy.$$

The probability mass function of N(t) is

$$\Pr(N(t) = k) = \frac{e^{-\int_{0}^{t} \lambda(y)dy} \left(\int_{0}^{t} \lambda(y)dy\right)^{k}}{k!},$$

and the renewal function is

$$m(t) = \int_{0}^{t} \lambda(y) dy.$$

Chapter 2

A new class of counting random fields

2.1 A random field with exponential marginal distributions

To make the manuscript self-contained, we start by introducing some notation in this section. For the rest of the thesis, given a second order real-valued random field $Q = \{Q(\mathbf{s}), \mathbf{s} \in A \subseteq \mathbb{R}^d\}$, we denote by $f_{Q(\mathbf{s})}$ and $F_{Q(\mathbf{s})}$ the marginal probability density function (pdf) and cumulative distribution function (cdf) of $Q(\mathbf{s})$, respectively. Moreover, for any set of distinct points $(\mathbf{s}_1, \ldots, \mathbf{s}_n)^{\top}$, $n \in \mathbb{N}$ and $\mathbf{s}_i \in A$, we denote the correlation function by $\rho_Q(\mathbf{s}_i, \mathbf{s}_j) = \operatorname{Corr}(Q(\mathbf{s}_i), Q(\mathbf{s}_j))$. In the stationary case, the notation adopted is $\rho_Q(\mathbf{h}) = \operatorname{Corr}(Q(\mathbf{s}_i), Q(\mathbf{s}_j))$, where $\mathbf{h} = \mathbf{s}_i - \mathbf{s}_j$ is the lag separation vector. Finally, $f_{Q_{ij}}$ denotes the *pdf* of the bivariate random vector $Q_{ij} = (Q(\mathbf{s}_i), Q(\mathbf{s}_j))^{\top}$, $i \neq j$. If the random field Q is a discrete-valued random field, then $\operatorname{Pr}(Q(\mathbf{s}) = l)$ and $\operatorname{Pr}(Q(\mathbf{s}_i) = n, Q(\mathbf{s}_j) = m)$, $l, m, n \in \mathbb{N}$ will denote the marginal and bivariate discrete probability functions, respectively.

Let $G = \{G(\mathbf{s}), \mathbf{s} \in A\}$ be a zero mean and unit variance weakly stationary Gaussian random field with correlation function $\rho_G(\mathbf{h})$. Henceforth, we call G the Gaussian underlying random field, and with some abuse of notation, we set $\rho(\mathbf{h}) := \rho_G(\mathbf{h})$, denoting this as the underlying correlation function. Let G_1, G_2 be two independent copies of G and let us define the random field $W = \{W(\mathbf{s}), \mathbf{s} \in A\}$ as follows:

$$W(\mathbf{s}) := \frac{1}{2\lambda(\mathbf{s})} \sum_{k=1}^{2} G_k^2(\mathbf{s}), \qquad (2.1)$$

where $\lambda(\mathbf{s}) > 0$ is a non-random function. W is a stationary random field with a marginal exponential distribution, with parameter $\lambda(\mathbf{s})$ denoted by $W(\mathbf{s}) \sim \text{Exp}(\lambda(\mathbf{s}))$ with $\mathbb{E}(W(\mathbf{s})) = 1/\lambda(\mathbf{s})$, $\text{Var}(W(\mathbf{s})) = 1/\lambda^2(\mathbf{s})$, and it can be easily observed that $\rho_W(\mathbf{h}) = \rho^2(\mathbf{h})$.

The associated multivariate exponential density was discussed earlier by Krishnamoorthy and Parthasarathy (1951), and its properties have been studied since then by several authors (Krishnaiah and Rao, 1961; Royen, 2004). However, likelihood-based methods for exponential random fields can be troublesome since the analytical expressions of the multivariate density can be derived only in some special cases. For example, when d = 1 and the underlying correlation function is exponential and the multivariate *pdf* is given by (Bevilacqua et al., 2020):

$$f_W(w_1, \dots, w_n) = \exp\left[-\frac{w_1\lambda_1}{(1-\rho_{1,2}^2)} - \frac{w_n\lambda_n}{(1-\rho_{n-1,n}^2)} - \sum_{i=2}^{n-1} \frac{(1-\rho_{i-1,i}^2\rho_{i,i+1}^2)\lambda_i w_i}{(1-\rho_{i-1,i}^2)(1-\rho_{i,i+1}^2)}\right] \times \prod_{i=1}^{n-1} I_0\left(\frac{2\rho_{i,i+1}\sqrt{w_i\lambda_i w_{i+1}\lambda_{i+1}}}{(1-\rho_{i,i+1}^2)}\right) \times \left(\prod_{i=1}^{n-1} (1-\rho_{i,i+1}^2)\right)^{-1},$$

with $\rho_{ij} := \exp\{-|s_i - s_j|/\phi\}, \lambda_i = \lambda(s_i), \phi > 0$ and $I_a(x)$ being the modified Bessel function of the first kind of order *a*. Regardless of the dimension of the space *A* and the type of correlation function, the bivariate exponential *pdf* is given by (Kibble, 1941; Vere-Jones, 1997):

$$f_{W_{ij}}(w_i, w_j) = \frac{e^{-\frac{(\lambda(\mathbf{s}_i)w_i + \lambda(\mathbf{s}_j)w_j)}{(1 - \rho^2(\mathbf{h}))}}}{(1 - \rho^2(\mathbf{h}))} I_0\left(\frac{2\sqrt{\rho^2(\mathbf{h})\lambda(\mathbf{s}_i)\lambda(\mathbf{s}_j)w_iw_j}}{(1 - \rho^2(\mathbf{h}))}\right)$$

The exponential random field W will be used next for defining a new random field with Poisson marginal distributions.

2.2 Spatial Poisson random fields

Our proposal relies on considering an infinite sequence of independent copies $Y_1, Y_2...$, of $Y = \{Y(\mathbf{s}), \mathbf{s} \in A\}$, a positive continuous random field. First, we define a new class of counting random fields, $N_t := \{N_t(\mathbf{s}), \mathbf{s} \in A\}$, $t \ge 0$, as follows:

$$N_t(\mathbf{s}) := \begin{cases} 0 & \text{if } 0 \le t < S_1(\mathbf{s}) \\ \max_{n \ge 1} \{ S_n(\mathbf{s}) \le t \} & \text{if } S_1(\mathbf{s}) \le t \end{cases},$$
(2.2)

where $S_n(\mathbf{s}) = \sum_{i=1}^n Y_i(\mathbf{s})$ is the *n*-fold convolution of *Y*. This model can be viewed as a spatial generalization of the renewal counting processes (Cox, 1970; Mainardi et al., 2007), where we consider independent copies of a positive random field as "inter-arrival times" instead of an independent and identically distributed sequence of positive random variables.

For each $s \in A$, and using the classical results from the renewal counting processes theory, the marginal discrete probability function of N is given by:

$$\Pr(N_t(\mathbf{s}) = n) = F_{S_n(\mathbf{s})}(t) - F_{S_{n+1}(\mathbf{s})}(t).$$
(2.3)

In addition, the marginal mean (the so-called renewal function) and the variance of N_t are given, respectively, by:

$$\mathbb{E}(N_t(\mathbf{s})) = \sum_{i=1}^{\infty} F_{S_i(\mathbf{s})}(t), \quad \operatorname{Var}(N_t(\mathbf{s})) = \left(2\sum_{i=1}^{\infty} iF_{S_i(\mathbf{s})}(t) - \mathbb{E}(N_t(\mathbf{s}))\right) - (\mathbb{E}(N_t(\mathbf{s})))^2.$$

Different elections of the positive random field Y lead to counting random fields with specific marginal distributions.

In this thesis, we assume that $Y \equiv W$, where W is the positive random field defined in (2.1), with $\text{Exp}(\lambda(\mathbf{s}))$ marginal distribution and *cdf* given by $F_{Y(\mathbf{s})}(x) = 1 - e^{-\lambda(\mathbf{s})x}$, x > 0. In this case, $S_n(\mathbf{s}) \sim \text{Gamma}(n, \lambda(\mathbf{s}))$ with $n \in \mathbb{N}$ is an Erlang distribution, with *cdf* given by:

$$F_{S_n(\mathbf{s})}(x) = 1 - \sum_{k=0}^{n-1} \frac{e^{-\lambda(\mathbf{s})x} (\lambda(\mathbf{s})x)^k}{k!}, \quad x > 0$$

and from (2.3), we can obtain the marginal distribution of N_t as:

$$\Pr(N_t(\mathbf{s}) = n) = e^{-t\lambda(\mathbf{s})} [t\lambda(\mathbf{s})]^n / n!, \quad n = 0, 1, 2, \dots$$
(2.4)

with $\mathbb{E}(N_t(\mathbf{s})) = \operatorname{Var}(N_t(\mathbf{s})) = t\lambda(\mathbf{s})$. Then, $N_t(\mathbf{s}) \sim \operatorname{Poisson}(t\lambda(\mathbf{s}))$ is the random number of renewals occurring in the temporal interval (0, t] and spatial location s. Additionally, $t\lambda(\mathbf{s})$ is the expected number of arrivals in an interval of length t for each location site s. Hereafter and without loss of generality, t is set to one, and N_t is denoted as N. We will call N a Poisson random field with underlying correlation $\rho(\mathbf{h})$ because N is marginally Poisson distributed and the dependence is indexed by a correlation function.

Note that, when the spatially varying mean (and variance) $\lambda(\mathbf{s})$ is not constant then N is not stationary. A typical parametric specification for the mean is given by $\lambda(\mathbf{s}) = e^{X(\mathbf{s})^{\top}\boldsymbol{\beta}}$, where $X(\mathbf{s}) \in \mathbb{R}^k$ is a vector of covariates and $\boldsymbol{\beta} \in \mathbb{R}^k$ even though other types of parametric and non-parametric specifications can be used.

It is important to note that although the proposed Poisson random field is defined on the *d*-dimensional Euclidean space A, the proposed method can be easily adapted to other spaces, such as the space-time space or the spherical spaces. The key for this extension is the specification of a suitable underlying correlation function $\rho(\mathbf{h})$. For instance, a correlation function defined on the space-time setting, *i.e.*, $A \subseteq \mathbb{R}^d \times \mathbb{R}$ (Gneiting, 2002c) or on the sphere of arbitrary radius *i.e*, $A \subseteq \mathbb{S}^2 = \{\mathbf{s} \in \mathbb{R}^3, ||\mathbf{s}|| = M\}, M > 0$ (Gneiting, 2013; Porcu et al., 2016). In the case of lattice or areal data, a suitable precision matrix with an appropriate neighbourhood structure should be specified for the underlying Gaussian Markov random field (Rue and Held, 2005).

To close this section we provide a closed expression for the joint probability generating function (jpgf) of a pair $(N(\mathbf{s}_i), N(\mathbf{s}_j))$ from a N Poison random. In fact, the jpgf is defined as follow:

$$\mathcal{P}(\lambda(\mathbf{s}_i), \lambda(\mathbf{s}_j); c_1, c_2) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \Pr(N(\mathbf{s}_i) = n, N(\mathbf{s}_j) = m) c_1^n c_2^m.$$

Moreover, Theorem 2.2.1 provides an expression for the *jpgf* which depends on a confluent hypergeometric function of two variables or Humbert series (Gradshteyn and

Ryzhik, 2014) defined as:

$$\Phi_3(\alpha,\beta;x,y) = \sum_{m,n}^{\infty} \frac{(\alpha)_m}{(\beta)_{m+n}m!n!} x^m y^n,$$

where $(\cdot)_{\ell}$ is the Pochhammer symbol (Abramowitz and Stegun, 1965).

Theorem 2.2.1. Let N be a Poisson random field with underlying correlation $\rho = \rho(\mathbf{h})$ and mean $\mathbb{E}(N(\mathbf{s}_k)) = \lambda(\mathbf{s}_k) = \lambda_k$. Then the jpgf, $\mathcal{P}(\lambda(\mathbf{s}_i), \lambda(\mathbf{s}_j); c_1, c_2)$, is given by:

$$\begin{aligned} \mathcal{P}(\lambda_i,\lambda_j;c_1,c_2) &= \exp\left\{-\frac{\lambda_i}{(1-\rho^2)} - \frac{\lambda_j}{(1-\rho^2)}\right\} \times \\ &\left[\Phi_3\left(1,1;\left(-(1-c_2) + \frac{1}{(1-\rho^2)}\right)\lambda_j, -\left(\frac{1-c_1c_2}{(1-\rho^2)} - \frac{1}{(1-\rho^2)^2}\right)\lambda_i\lambda_j\right) \right. \\ &\left. -\Phi_3\left(1,1;\left(-\frac{c_1c_2-c_1}{[(1-c_1)(1-\rho^2)-1]} + \frac{1}{(1-\rho^2)}\right)\lambda_j, -\left(\frac{1-c_1c_2}{(1-\rho^2)} - \frac{1}{(1-\rho^2)^2}\right)\lambda_i\lambda_j\right)\right] \\ &\left. + \exp\left\{-(1-c_1)\lambda_i - \left(\frac{c_1c_2-c_1}{[(1-c_1)(1-\rho^2)-1]}\right)\lambda_j\right\}, \end{aligned}$$

Proof. See the Appendix.

Hunter (1974) obtained the bivariate Laplace transform of the *jpgf*, but not an explicit expression for the *jpgf*, even though he mentioned that it can be obtained theoretically by inverting his proposed expression. Thus, Theorem 2.2.1 becomes an important contribution to Hunter (1974) by giving a closed expression for the *jpgf*.

2.2.1 Correlation function

The following result, which can be obtained from the pioneering work of Hunter (1974), provides the correlation function $\rho_N(\mathbf{s}_i, \mathbf{s}_j)$ of the non-stationary Poisson random field with underlying correlation $\rho(\mathbf{h})$ depending on the regularized lower incomplete gamma function:

$$\gamma^*(a,x) = \frac{\gamma(a,x)}{\Gamma(a)} = \frac{1}{\Gamma(a)} \int_0^\infty t^{a-1} e^{-t} dt$$
(2.5)

where $\gamma(\cdot, \cdot)$ is the lower incomplete gamma function and $\Gamma(\cdot)$ is the Gamma function. Additionally, we define the function $\gamma^*(a, x, x') = \gamma^*(a, x) \gamma^*(a, x')$, which considers the product of two regularized lower incomplete gamma function sharing a common parameter.

Theorem 2.2.2. Let N be a non-stationary Poisson random field with underlying correlation $\rho(\mathbf{h})$. Then,

$$\rho_N(\mathbf{s}_i, \mathbf{s}_j) = \frac{\rho^2(\mathbf{h})(1 - \rho^2(\mathbf{h}))}{\sqrt{\lambda(\mathbf{s}_i)\lambda(\mathbf{s}_j)}} \sum_{r=0}^{\infty} \gamma^* \left(r + 1, \frac{\lambda(\mathbf{s}_i)}{1 - \rho^2(\mathbf{h})}, \frac{\lambda(\mathbf{s}_j)}{1 - \rho^2(\mathbf{h})}\right),$$

with $\mathbf{h} = \mathbf{s}_i - \mathbf{s}_j$.

Proof. For details, refer to Hunter (1974), section 5.2 (pages 38-39).

Corollary 2.2.1. In Theorem 2.2.2, when $\lambda(\mathbf{s}) = \lambda$, the Poisson random field is weakly stationary with the correlation function given by:

$$\rho_N(\mathbf{h},\lambda) = \rho^2(\mathbf{h}) \left[1 - \exp\left(-z(\mathbf{h},\lambda)\right) \left(I_0\left(z(\mathbf{h},\lambda)\right) + I_1\left(z(\mathbf{h},\lambda)\right) \right) \right], \quad (2.6)$$

where $z(\mathbf{h}, \lambda) = 2\lambda(1 - \rho^2(\mathbf{h}))^{-1}$.

Proof. See the Appendix.

Note that $\rho_N(\mathbf{h})$ is well defined at the origin since $\rho_N(\mathbf{h}) = 1$, as $\mathbf{h} \to \mathbf{0}$, implying that the Poisson random field is weakly stationary and mean square continuous. Additionally, if $\rho(\mathbf{h}) = 0$, then $\rho_N(\mathbf{h}) = 0$ and if $\lambda \to \infty$ then $\rho_N(\mathbf{h}) = \rho^2(\mathbf{h})$, *i.e.*, it converges to the correlation function of an exponential random field.

Following the graphical example given in Figure 1, we now compare the correlation functions of the proposed Poisson random field with the correlation of the Poisson LG random field, which is defined hierarchically by first considering a LG random field $Z = \{Z(\mathbf{s}), \mathbf{s} \in A\}$ defined as $Z(\mathbf{s}) = e^{\mu + \sqrt{\sigma^2}G(\mathbf{s})}$, where G is a standard Gaussian random field with correlation $\rho(\mathbf{h})$, and then assuming $Y(\mathbf{s}) \mid Z(\mathbf{s}) \sim \text{Poisson}(Z(\mathbf{s}))$ with $Y(\mathbf{s}_i) \perp Y(\mathbf{s}_j) \mid Z$ for $i \neq j$. In this case, the first two moments of $Y(\mathbf{s})$ are given by $\mathbb{E}(Y(\mathbf{s})) =$ $e^{\mu+0.5\sigma^2}$ and $Var(Y(\mathbf{s})) = \mathbb{E}(Y(\mathbf{s}))(1 + \mathbb{E}(Y(\mathbf{s}))(e^{\sigma^2} - 1))$. Consequently, and following Aitchison and Ho (1989), the correlation function is given by:

$$\rho_Y(\mathbf{h}, \mu, \sigma^2) = \frac{e^{\sigma^2 \rho(\mathbf{h})} - 1}{e^{\sigma^2} - 1 + \mathbb{E}(Y(\mathbf{s}))^{-1}}.$$

This correlation is discontinuous at the origin and the nugget effect is given by:

$$\frac{\mathbb{E}(Y(\mathbf{s}))^{-1}}{\mathbb{E}(Y(\mathbf{s}))^{-1} + e^{\sigma^2} - 1} > 0.$$

It is apparent that the marginal mean $\mathbb{E}(Y(\mathbf{s}))$ has a strong impact on the nugget effect.

Figure 2.1 (a) depicts the correlation functions $\rho_Y(\mathbf{h}, 0.5, 0.05)$, $\rho_Y(\mathbf{h}, 2.5, 0.1)$, and $\rho_Y(\mathbf{h}, 4.5, 0.2)$, which correspond to Poisson LG random fields with mean $\mathbb{E}(Y(\mathbf{s})) = 1.69, 12.81$, and, 99.48, respectively. As underlying correlation model we assume $\rho(\mathbf{h}) = (1 - ||\mathbf{h}||/0.5)^4)_+$. It can be appreciated that for large mean values, the nugget effect is negligible. However, for small mean values, the nugget effect can be huge, and it is the cause of the "whitening" effect observed in Figure 1 (a). This fact has been also highlighted in De Oliveira (2013), which indicates that the Poisson LG random field may impose severe restrictions on the correlation structures, which is inadequate for spatial count data models consisting primarily of small counts.

Figure 2.1 (b) depicts the correlation function $\rho_N(\mathbf{h}, \lambda)$ of the proposed Poisson random field using the same means and underlying correlation function of the Poisson LG random field. It can be appreciated that the correlation is well defined at the origin and covers the entire range between 0 and 1, irrespective of the mean values.

Finally, Figure 2.1 (c) depicts the correlation function of the Poisson GC random field (Han and De Oliveira, 2016) $C = \{C(\mathbf{s}), \mathbf{s} \in A\}$ defined as $C(\mathbf{s}) = F_{\mathbf{s}}^{-1}(\Phi(G(\mathbf{s})), \lambda)$, where $\Phi(\cdot)$ is the *cdf* of the standard Gaussian distribution and $F_{\mathbf{s}}^{-1}(\cdot, \lambda)$ is the quantile function of the Poisson distribution and G is a standard Gaussian random field with correlation $\rho(\mathbf{h})$. The correlation function in this case is given by

$$\rho_C(\mathbf{h},\lambda) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \lambda^{-1} F_{\mathbf{s}_i}^{-1}(\Phi(z_i),\lambda) F_{\mathbf{s}_j}^{-1}(\Phi(z_j),\lambda) \phi_2(z_i, z_j, \rho(\mathbf{h})) dz_i dz_j - \lambda$$

where ϕ_2 is the *pdf* of the bivariate standard Gaussian distribution. It is apparent that the Poisson GC correlation $\rho_C(\mathbf{h}, \lambda)$ is much stronger than $\rho_N(\mathbf{h}, \lambda)$, and it does not seem to be affected by the different mean values.



Figure (2.1) From left to right: (a) correlation functions $\rho_Y(\mathbf{h}, \mu, \sigma^2)$ of the Poisson LG random field with $\mu = 0.5, \sigma^2 = 0.05$ and $\mu = 2.5, \sigma^2 = 0.1$, and $\mu = 4.5, \sigma^2 = 0.2$; (b) correlation function $\rho_N(\mathbf{h}, \lambda)$ of our proposed Poisson random field for $\lambda = 1.69, 12.81$, and, 99.48; (c) correlation function $\rho_C(\mathbf{h}, \lambda)$ of the Poisson GC random field for $\lambda = 1.69, 12.81$, 12.81, and, 99.48. The black line in the Figures depicts the underlying correlation model given by $\rho(\mathbf{h}) = (1 - ||\mathbf{h}||/0.5)_+^4$.

It is important to stress that Poisson and Poisson GC random fields that are

not mean-square continuous can be obtained by introducing a nugget effect, *i.e.*, a discontinuity at the origin of $\rho_N(\mathbf{h})$. This can be achieved by replacing the underlying correlation function $\rho(\mathbf{h})$ with $\rho^*(\mathbf{h}) = \rho(\mathbf{h})(1 - \tau^2) + \tau^2 \mathbb{1}_0(||\mathbf{h}||)$, where $0 \le \tau^2 < 1$ represents the underlying nugget effect.

2.2.2 Bivariate distribution

In this section, we provide the bivariate distribution of the Poisson random field. This distribution can be written in terms of an infinite series depending on the regularized lower incomplete Gamma function defined in (2.5) and the regularized hypergeometric confluent function (Gradshteyn and Ryzhik, 2014), defined as:

$${}_{1}\widetilde{\mathbf{F}}_{1}(a;b;x) = \frac{{}_{1}\mathbf{F}_{1}(a;b;x)}{\Gamma(b)} = \sum_{k=0}^{\infty} \frac{(a)_{k}x^{k}}{\Gamma(b+k)k!},$$

where $_{1}F_{1}$ is the standard hypergeometric confluent function.

For the sake of simplicity, we analyse the following cases separately: (a) n = m = 0, (b) $n = 0, m \ge 1$ and $m = 0, n \ge 1$, (c) n = m = 1, 2..., and (d) $n, m \ge 1, n \ne m$. Moreover, we set $p_{nm} = \Pr(N(\mathbf{s}_i) = n, N(\mathbf{s}_j) = m)$, $\lambda_i = \lambda(\mathbf{s}_i), \lambda_j = \lambda(\mathbf{s}_j)$ and $\rho = \rho(\mathbf{h})$ for notational convenience. We additionally define the function S as follows:

$$\mathcal{S}\left(\begin{smallmatrix}a&;b\\c&,x,x'\end{smallmatrix}\right) = {}_{1}\widetilde{\mathrm{F}}_{1}(a;b;x)\gamma^{*}\left(c,x'\right).$$

Theorem 2.2.3. Let N be a Poisson random field with underlying correlation ρ and mean $\mathbb{E}(N(\mathbf{s}_k)) = \lambda_k$. Then the bivariate distribution p_{nm} is given by: (a) Case n = m = 0:

$$p_{00} = -1 + e^{-\lambda_i} + e^{-\lambda_j} + (1 - \rho^2) \sum_{k=0}^{\infty} \rho^{2k} \gamma^* \left(k + 1, \frac{\lambda_i}{1 - \rho^2}, \frac{\lambda_j}{1 - \rho^2}\right).$$

(b) Cases $n \ge 1, m = 0$ and $m \ge 1, n = 0$, $p_{n0} = g(n, \lambda_i, \lambda_j, \rho)$ and $p_{0m} = g(m, \lambda_j, \lambda_i, \rho)$, respectively, where

$$g(b, x, y, \rho) = \frac{x^b}{b!} e^{-x} - x^b e^{-\frac{x}{1-\rho^2}} \sum_{\ell=0}^{\infty} \left(\frac{\rho^2 x}{1-\rho^2}\right)^{\ell} \mathcal{S}\left(\begin{smallmatrix}b & b+\ell+1\\ \ell+1 & \frac{\rho^2 x}{1-\rho^2}, \frac{y}{1-\rho^2}\end{smallmatrix}\right).$$

(c) Case $n = m \ge 1$:

$$p_{nn} = -(1-\rho^2)^n \sum_{k=0}^{\infty} \frac{\rho^{2k}(n)_k}{k!} \gamma^{\star} \left(n+k, \frac{\lambda_i}{1-\rho^2}, \frac{\lambda_j}{1-\rho^2} \right) \\ + \left(\frac{1-\rho^2}{\rho^2} \right)^n \sum_{k=0}^{\infty} \sum_{\ell=0}^{1} \frac{(n)_k}{k!} e^{-\lambda(\mathbf{s}_i)(1-\ell)-\lambda(\mathbf{s}_j)\ell} \gamma^{\star} \left(n+k, \frac{\rho^{2(1-\ell)}\lambda_i}{1-\rho^2}, \frac{\rho^{2\ell}\lambda_j}{1-\rho^2} \right) \\ + (1-\rho^2)^{n+1} \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \frac{\rho^{2k+2\ell}(n)_\ell}{\ell!} \gamma^{\star} \left(n+\ell+k+1, \frac{\lambda_i}{1-\rho^2}, \frac{\lambda_j}{1-\rho^2} \right).$$

(d) Cases $n \ge 2, m \ge 1$ with n > m, and $m \ge 2, n \ge 1$ with m > n, $p_{nm} = h(n, m, \lambda_i, \lambda_j, \rho)$ and $p_{nm} = h(m, n, \lambda_j, \lambda_i, \rho)$, respectively, where

$$h(a, b, x, y, \rho) = x^{m} e^{-\frac{x}{1-\rho^{2}}} \left[\sum_{\ell=0}^{\infty} \frac{(b)_{\ell}}{\ell!} \left(\frac{\rho^{2} x}{1-\rho^{2}} \right)^{\ell} \mathcal{S} \left(a^{-b+1}; a^{+\ell+1}; \frac{\rho^{2} x}{1-\rho^{2}}, \frac{y}{1-\rho^{2}} \right) - \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \frac{(b)_{\ell}}{\ell!} \left(\frac{\rho^{2} x}{1-\rho^{2}} \right)^{k+\ell} \mathcal{S} \left(a^{-b}; a^{+k+\ell+1}; \frac{\rho^{2} x}{1-\rho^{2}}, \frac{y}{1-\rho^{2}} \right) \right].$$

Proof. See the Appendix.

The evaluation of the bivariate distribution can be troublesome at first sight. However, it can be performed by truncating the series and taking into account that efficient numerical computation of the regularized lower incomplete Gamma and hypergeometric confluent functions can be found in different libraries such as the GNU scientific library (Gough, 2009) and the most important statistical softwares including R, MATLAB and Python. In particular, the R package Geomodels (Bevilacqua et al., 2019) uses the Python implementations in the SciPy library (Virtanen et al., 2020).

The bivariate distribution can be written as the product of two independent Poisson distributions when $\rho_N(\mathbf{h}) = 0$. This result, provided by Hunter (1974) in Theorem 3.6, establishes that the independence of two renewal counting processes is equivalent to a zero correlation between them. As outlined in Section 2.2.1, $\rho(\mathbf{h}) = 0$ implies $\rho_N(\mathbf{h}) = 0$.

Consequently, pairwise independence at the level of the underlying Gaussian random field implies pairwise independence for the Poisson random field.

We now compare the type of bivariate dependence induced by the proposed model and the GC one when $\lambda = 5$.



Figure (2.2) For each row (from left to right): bivariate Poisson GC distribution, our proposed bivariate Poisson distribution and the difference between them. The first, second and third row are obtained setting $\rho(\mathbf{h}) = 0.1, 0.5, 0.9$ for the underlying correlation.

Figure (2.2) (from left to right) presents the bivariate GC distribution, the bivariate Poisson distribution in Theorem 2.2.3 and a coloured image representing the differences between them. Note that a positive value of that difference implies that the probabilities associated with the bivariate distribution in Theorem 2.2.3 are greater than the probabilities of the bivariate GC one. Only the probabilities $Pr(N(\mathbf{s}_i) = n, N(\mathbf{s}_j) = m)$ for n, m = 0, 1, ..., 12 are considered in the plots. The first, second and third rows consider increasing levels of underlying correlations $\rho(\mathbf{h}) = 0.1, 0.5, 0.9$.

The higher the correlation, the more significant the difference between the bivariate distributions. In addition, it can be observed that the probabilities of the GC distribution are the largest along the diagonal. This is not surprising since the Poisson GC model inherits the type of dependence of the underlying Gaussian random field. On the contrary, the proposed bivariate distribution tends to assign more probabilities outside the main diagonal with respect to the GC case.

2.3 Estimation and prediction

In this section, we start by describing the weighted pairwise likelihood (*wpl*) estimation method; then, we focus on the optimal linear prediction.

2.3.1 Weighted pairwise likelihood estimation

Composite likelihood is a general class of objective functions that combine low-dimensional terms based on the likelihood of marginal or conditional events to construct a pseudo likelihood (Lindsay, 1988; Varin et al., 2011). A particular case of the composite likelihood class is the pairwise likelihood (see for example Heagerty and Lele, 1998; Bevilacqua and Gaetan, 2015; Alegría et al., 2017; Bevilacqua et al., 2020, for application of pairwise likelihood in the spatial setting) that combines the bivariate distributions of all possible distinct pairs of observations. Let $N = (n_1, n_2, ..., n_l)^{\top}$ be a realization of the Poisson random field N observed at distinct spatial locations $s_1, s_2, \ldots, s_l, s_i \in A$ and let $\theta = (\beta^{\top}, \alpha^{\top})$ be the vector of unknown parameters where α is the vector parameter associated with the underlying correlation model and β the regression parameters. The pairwise likelihood function is defined as follows:

$$pl(\boldsymbol{\theta}) := \sum_{i=1}^{l-1} \sum_{j=i+1}^{l} \log(\Pr(N(\mathbf{s}_i) = n_i, N(\mathbf{s}_j) = n_j))\zeta_{ij},$$

where $Pr(N(\mathbf{s}_i) = n_i, N(\mathbf{s}_j) = n_j)$ is the bivariate density given in Theorem 2.2.3 and ζ_{ij} is a non-negative suitable weight. The choice of cut-off weights, namely,

$$\zeta_{ij} = \begin{cases} 1 & \| \mathbf{s}_i - \mathbf{s}_j \| \le \xi \\ 0 & \text{otherwise} \end{cases},$$
(2.7)

for a positive value of ξ , can be motivated by its simplicity and by observing that the dependence between observations that are distant is weak (Joe and Lee, 2009; Bevilacqua and Gaetan, 2015).

The maximum weighted pairwise likelihood (wpl) estimator is given by:

$$\widehat{\boldsymbol{\theta}} := \operatorname{argmax}_{\boldsymbol{\theta}} \operatorname{pl}(\boldsymbol{\theta}).$$

Under some mixing conditions of the Poisson random field (Bevilacqua et al., 2012; Bevilacqua and Gaetan, 2015), it can be shown that, in the case of increasing domain asymptotic, $\hat{\theta}$ is consistent and asymptotically Gaussian distributed, with the covariance matrix given by $\mathcal{G}_n^{-1}(\theta)$, *i.e.*, the inverse of the Godambe information $\mathcal{G}_n(\theta) :=$ $\mathcal{H}_n(\theta)\mathcal{J}_n(\theta)^{-1}\mathcal{H}_n(\theta)$, where $\mathcal{H}_n(\theta) := \mathbb{E}[-\nabla^2 \operatorname{pl}(\theta)]$ and $\mathcal{J}_n(\theta) := \operatorname{Var}[\nabla \operatorname{pl}(\theta)]$. The standard error estimation can be obtained from the square root diagonal elements of $\mathcal{G}_n^{-1}(\hat{\theta})$.

It is important to stress that the computation of the standard errors requires the evaluation of the matrices $\mathcal{H}_n(\hat{\theta})$ and $\mathcal{J}_n(\hat{\theta})$. However, the evaluation of $\mathcal{J}_n(\hat{\theta})$ is computationally unfeasible for large datasets, and in this case, subsampling techniques can be used, as in Heagerty and Lele (1998) and Bevilacqua et al. (2012). A straightforward and more robust alternative is the parametric bootstrap estimation of $\mathcal{G}_n^{-1}(\theta)$ (Bai et al., 2014). Another critical issue related to large datasets is that the computation of the *wpl* estimator can be computationally demanding due to the computational complexity associated with the bivariate Poisson distribution given in Theorem 2.2.3. An estimator that requires a smaller computational burden can be obtained by considering a misspecified *wpl* (Masuda, 2013; Gouriéroux et al., 2017; Bevilacqua et al., 2020). Specifically, suppose that in the estimation procedure, we assume a non-stationary Gaussian random field with mean and variance equal to $\lambda(s)$ and correlation $\rho_N(s_i, s_j)$ given in Theorem 2.2.2. In that case, a misspecified Gaussian *wpl* only requires the computation of the Gaussian bivariate distribution.

Gaussian misspecification is a useful inferential tool when the likelihood computation cannot be calculated for some reason, but the first two moments and the correlation are known. Note that the misspecified Gaussian random field matches the mean, variance, and correlation function of the Poisson random field. Additionally, standard maximum likelihood estimation can be performed under the Gaussian misspecification setting.

2.3.2 Optimal linear prediction

The random field's optimal predictor concerning the mean squared error criterion requires the knowledge of the finite-dimensional distribution, which is not available for the Poisson random field. As in the estimation step, once again, the Gaussian misspecification allows to build an optimal linear predictor based on the correlation of the Poisson random field given in Theorem 2.2.2. Specifically, if the goal is the prediction of N at s_0 given the vector of spatial observations N observed at s_1, s_2, \ldots, s_l , then the optimal linear Gaussian prediction is given by:

$$\widehat{N(\mathbf{s}_0)} = \lambda(\mathbf{s}_0) + \boldsymbol{c}^{\top} \Sigma^{-1} (\boldsymbol{N} - \boldsymbol{\lambda})$$
 (2.8)

where $\lambda = (\lambda(\mathbf{s}_1), \dots, \lambda(\mathbf{s}_l))^{\top}$, $\mathbf{c} = [\sqrt{\lambda(\mathbf{s}_0)\lambda(\mathbf{s}_i)}\rho_N(\mathbf{s}_0, \mathbf{s}_i)]_{i=1}^l$ and $\Sigma = \sqrt{\lambda\lambda^{\top}} \odot [\rho_N(\mathbf{s}_i, \mathbf{s}_j)]_{i,j=1}^l$ is the variance-covariance matrix (\odot the matrix Schur product). In practice, the mean and covariance matrix are not known and must be estimated. The
associated mean squared error is:

$$MSE(\widehat{N(\mathbf{s}_0)}) = \lambda(\mathbf{s}_0) - \boldsymbol{c}^{\top} \Sigma^{-1} \boldsymbol{c}_0$$

Note that this kind of prediction does not guarantee the positivity and discreteness of the prediction. However, in general, optimal linear prediction can be a useful approximation of the optimal predictor, as was shown, for example, in De Oliveira (2006) and recently in Bevilacqua et al. (2020).

2.4 Simulation studies

In this section, we focus on two simulation studies. The first one analyses the performance of the *wpl* method when estimating the Poisson random field under the spatial and spatio-temporal settings. The second one analyses the Poisson optimal linear predictor's performance, comparing our approach with the Poisson GC and Poisson LG models.

2.4.1 Performance of the weighted pairwise likelihood estimation

In this study, we consider 1000 realizations from a stationary spatial Poisson random field observed at $\mathbf{s}_i \in [0, 1]^2$, i = 1, ..., l, l = 441. Specifically, we considered a regular grid with increments of size 0.05 over the unit square $[0, 1]^2$. The grid points were perturbed, adding a uniform random value over [-0.015, 0.015] to each coordinate. A perturbed grid allows us to obtain more stable estimates since different sets of small distances are available and very close location points are avoided.

For the Poisson random field we, consider $\lambda(\mathbf{s}) = e^{\beta}$ with $\beta = \log(2), \log(5), \log(10), \log(20)$, and an underlying isotropic correlation model $\rho(\mathbf{h}) = (1 - ||\mathbf{h}||/\alpha)_+^4$ with $\alpha = 0.2$. As outlined in Section 1.1.5, the use of a compactly supported correlation function simplifies the computation of the bivariate Poisson distribution proposed in Theorem 2.2.3.

We study the performance of the Poisson *wpl*, the misspecified Gaussian *wpl* and the misspecified Gaussian maximum likelihood (ML) estimation methods. In the

(misspecified) wpl estimation, we consider a cut-off weight function, as in (2.7), with $\xi = 0.1$.

	Poisson wpl		Gaussian wpl		Gaussian ML	
	Bias	MSE	Bias	MSE	Bias	MSE
$\beta = \log(2)$	-0.00251	0.00151	-0.00348	0.00161	-0.00366	0.00165
$\alpha = 0.2$	-0.00663	0.00113	-0.00828	0.00208	-0.00748	0.00203
$\beta = \log(5)$	-0.00113	0.00065	-0.00147	0.00068	-0.00161	0.00068
$\alpha = 0.2$	-0.00435	0.00098	-0.00422	0.00149	-0.00344	0.00145
$\beta = \log(10)$	0.00052	0.00033	0.00031	0.00033	0.00014	0.00033
$\alpha = 0.2$	-0.00261	0.00096	-0.00336	0.00120	-0.00296	0.00115
$\beta = \log(20)$	-0.00026	0.00018	-0.00039	0.00019	-0.00037	0.00018
$\alpha = 0.2$	-0.00449	0.00094	-0.00499	0.00099	-0.00402	0.00095

Table (2.1) Bias and MSE associated with Poisson *wpl*, misspecified Gaussian *wpl* and misspecified Gaussian ML when the true random field is Poisson with $\lambda(\mathbf{s}) = e^{\beta}$ and $\rho(\mathbf{h}) = (1 - ||\mathbf{h}||/\alpha)_{+}^{4}$.

Table 2.1 shows the bias and mean squared error associated with β and α through the four scenarios and three estimation methods. As expected, the misspecified Gaussian ML performs slightly better than the misspecified Gaussian *wpl*. More importantly, it can be recognized that the Poisson *wpl* shows the best performance, particularly when estimating the spatial dependence parameter. This fact is more evident for low counts *i.e.*, when β is decreasing. However, when increasing the mean, the performances of the three methods of estimation tend to be considerably similar, in particular when the mean of the Poisson random field is 20.

To summarize, the Poisson *wpl* is the best method for estimating the Poisson random field when the mean is small (lower than 20 as a rule of thumb in our experiments). For large counts, the misspecified Gaussian *wpl* or ML methods show approximately the same performance as the Poisson *wpl* method.

We also study the proposed methods' performance when estimating a non-stationary version of the Poisson random field. Under the previous simulation setting we changed the constant mean by considering a regression model, that is, $\lambda(\mathbf{s}) = \exp\{\beta + \beta_1 u_1(\mathbf{s}) + \beta_2 u_2(\mathbf{s})\}\$ with $\beta = 1.5$, $\beta_1 = -0.2$ and $\beta_2 = 0.3$, where $u_1(\mathbf{s})$ and $u_2(\mathbf{s})$ are independent realizations from a (0, 1) uniform random variable. Table 2.2 shows the bias and MSE associated with β , β_1 , β_2 and α for the three methods of estimation, and Figure 2.3 plots the associated centred box-plots.

Additionally, in this case, the Poisson *wpl* method shows the best MSE for each estimation. Notice that all three methods of estimations show approximately the same performance as in the stationary case.

	Poisson wpl		Gaussian wpl		Gaussian ML	
	Bias	MSE	Bias	MSE	Bias	MSE
$\beta = 1.5$	-0.00263	0.00419	-0.00359	0.00445	-0.00320	0.00428
$\beta_1 = -0.2$	0.00189	0.00618	0.00148	0.00665	0.00046	0.00627
$\beta_2 = 0.3$	0.00185	0.00608	0.00202	0.00661	0.00182	0.00621
$\alpha = 0.2$	-0.00148	0.00091	-0.00030	0.00124	0.00096	0.00122

Table (2.2) Bias and MSE associated with the Poisson *wpl*, misspecified Gaussian *wpl* and misspecified Gaussian ML when estimating a non-stationary Poisson random field with $\lambda(\mathbf{s}) = \exp\{\beta + \beta_1 u_1(\mathbf{s}) + \beta_2 u_2(\mathbf{s})\}$ and $\rho(\mathbf{h}) = (1 - ||\mathbf{h}||/\alpha)_+^4$.

Finally, we consider a simulation scheme under a spatio-temporal setting. Specifically, we consider 1000 simulations from a non-stationary space-time Poisson random field observed at $\mathbf{s}_i \in [0, 1]^2$, i = 1, ..., l, l = 40 spatial location sites, uniformly distributed within the unit square and $t_1^* = 0, t_2^* = 0.25, ..., t_{25}^* = 6$, 25 time points. We consider a regression model for the spatio-temporal mean $\lambda(\mathbf{s}, t^*) = \exp\{\beta + \beta_1 u_1(\mathbf{s}, t^*) + \beta_2 u_2(\mathbf{s}, t^*)\}$, where $u_k(\mathbf{s}, t^*)$, k = 1, 2 are independent realizations from a (0, 1) uniform random variable. We set $\beta = 1.5$, $\beta_1 = -0.2$ and $\beta_2 = 0.3$ as in the previous simulation scheme.

Additionally, as the underlying space-time correlation, we use a simple isotropic and temporal symmetric space-time Wendland separable model $\rho(\mathbf{h}, t^*) = (1 - ||\mathbf{h}|| / \alpha_s)_+^4 (1 - |t^*| / \alpha_{t^*})_+^4$ with $\alpha_s = 0.2$ and $\alpha_{t^*} = 1$, where $t^* = t_i^* - t_j^*$ with $i, j \in \{1, 2, \dots, 25\}$.



Figure (2.3) Centred box-plots of estimates under the Poisson *wpl* (P), misspecified Gaussian *wpl* (PG), and misspecified Gaussian ML (MG) when estimating a non-stationary Poisson random field with $\lambda(\mathbf{s}) = \exp\{\beta + \beta_1 u_1(\mathbf{s}) + \beta_2 u_2(\mathbf{s})\}, \beta = 0.5, \beta_1 = -0.2, \beta_2 = 0.3$ and $\rho(\mathbf{h}) = (1 - ||\mathbf{h}||/\alpha)_+^4$ with $\alpha = 0.2$.

Finally, for the misspecified *wpl* estimation, we consider a cut-off weight function as in (2.7) extended to the space time case, with $\xi_s = 0.2$ and $\xi_{t^*} = 0.5$.

The results concerning this simulation study are shown in Table 2.3, including the bias and MSE associated with β , β_1 , β_2 and α_s , α_{t^*} for the three estimation methods. In addition, Figure 2.4 shows the associated box plots. As it can be observed, the Poisson *wpl* approach outperforms the misspecified Gaussian *wpl* and ML as expected for each parameter.

We want to highlight that all of the estimation methods showed similar behaviours as in the purely spatial case.

	Poisson wpl		Gaussian wpl		Gaussian ML	
	Bias	MSE	Bias	MSE	Bias	MSE
$\beta = 1.5$	-0.00058	0.00166	-0.00120	0.00180	-0.00110	0.00167
$\beta_1 = -0.2$	-0.00079	0.00257	-0.00056	0.00274	-0.00102	0.00249
$\beta_2 = 0.3$	0.00036	0.00284	0.00070	0.00302	0.00062	0.00267
$\alpha_{\rm s}=0.2$	-0.01057	0.00464	-0.01323	0.00630	-0.01343	0.00629
$\alpha_{t^*} = 1$	-0.00124	0.01846	0.00165	0.02534	0.00032	0.02415

Table (2.3) Bias and MSE associated with the Poisson *wpl*, misspecified Gaussian *wpl* and misspecified Gaussian ML when estimating a non-stationary spatio-temporal Poisson random field with $\lambda(\mathbf{s}, t^*) = \exp\{\beta + \beta_1 u_1(\mathbf{s}, t^*) + \beta_2 u_2(\mathbf{s}, t^*)\}$ and $\rho(\mathbf{h}, t^*) = (1 - ||\mathbf{h}||/\alpha_{\mathbf{s}})^4_+(1 - |t^*|/\alpha_{t^*})^4_+$.



Figure (2.4) Centred box-plots of estimates under the Poisson *wpl* (P), misspecified Gaussian *wpl* (PG), and misspecified Gaussian ML (MG) when estimating a non-stationary space-time Poisson random field with $\lambda(\mathbf{s}, t^*) = \exp\{\beta + \beta_1 u_1(\mathbf{s}, t^*) + \beta_2 u_2(\mathbf{s}, t^*)\}$ with $\beta = 0.5$, $\beta_1 = -0.2$ and $\beta_2 = 0.3$ and $\rho(\mathbf{h}, t^*) = (1 - ||\mathbf{h}||/\alpha_{\mathbf{s}})^4_+(1 - |t^*|/\alpha_{t^*})^4_+$ with $\alpha_{\mathbf{s}} = 0.2$, $\alpha_{t^*} = 1$.

2.4.2 Performance of the optimal linear prediction

In this section, we compare the performance of the optimal linear predictor of the proposed Poisson random field with the optimal predictors based on the Poisson GC and Poisson LG approaches. To compare the prediction performance of the three approaches, we consider the following steps:

- 1. Set j = 1. Repeat until j = 100.
- 2. Simulate the j-th spatial dataset from the proposed Poisson random field by considering 300 location sites uniformly distributed on the unit square.
- 3. Set k = 1. Repeat until k = 50.
- 4. Randomly split the *j*-th dataset by using 80% of the data for estimation and 20% as the validation dataset.
- 5. Estimate using *wpl* under our model and using ML for the Poisson GC and Poisson LG models.
- Compute the optimal linear predictor (2.8) and the optimal predictor for the Poisson GC and Poisson LG models at the coordinates associated with the validation dataset, given the estimates obtained at the previous step.
- 7. Compute, for each model, $RMSE_k$ and MAE_k .
- 8. k = k + 1.

9. Compute, for each model $\overline{\text{RMSE}}_j = \sum_{k=1}^{50} \text{RMSE}_k / 50$ and $\overline{\text{MAE}}_j = \sum_{k=1}^{50} \text{MAE}_k / 50$.

- 10. j = j + 1.
- 11. Compute, for each model RMSE = $\sum_{j=1}^{100} \overline{\text{RMSE}}_j / 100$ and MAE = $\sum_{j=1}^{100} \overline{\text{MAE}}_j / 100$.

This numerical experiment has been replicated by simulating (at step 2) from a Poisson GC random field. The simulation settings have been chosen such that the means (and variances) are the same for both models. Specifically, we consider three scenarios with increasing mean, that is 1.69 (scenario 1), 12.81 (scenario 2) and 99.48 (scenario 3). This election generates a Poisson marginal distribution with low, medium and large counts, respectively. Additionally, as the underlying correlation model, we consider $\rho(\mathbf{h}) = e^{-3||\mathbf{h}||/\alpha}$, with $\alpha = 0.15$ for the Poisson GC model and $\alpha = 0.35$ for the Poisson model. This specific setting allows us to obtain similar correlations in both cases.

For the *wpl* estimation of our Poisson model, we set $\xi = 0.05$ in (2.7). For the Poisson GC and Poisson LG, we use the maximum likelihood estimation method implemented in the R (R Core Team, 2020) packages gcKrig (Han and Oliveira, 2018) and spaMM (Rousset and Ferdy, 2014). The gcKrig package implements a variant of the sequential importance sampling algorithm and it is used to approximate the Poisson GC likelihood (see Masarotto and Varin, 2017; Han and Oliveira, 2018, for the computational details). The Poisson LG maximum likelihood estimation is computed using Laplace approximations through the spaMM package (Rousset and Ferdy, 2014). Finally, the Poisson GC model's optimal prediction, which involves the evaluation of an *n*-dimensional integral, is approximated using a variant of the sequential importance sampling algorithm, as in the estimation step.

Table 2.4 summarizes the results of our experiment, showing the RMSE and MAE for each model under the different scenarios and types of model generation. As expected, the prediction using the Poisson LG is the worst in all scenarios. The performance predictions in terms of RMSE and MAE of our Poisson and Poisson GC models are quite similar in all scenarios, with a slight preference for the Poisson GC prediction in some of them. This is not surprising since, under the GC model, the optimal predictor is computed. However, the proposed optimal linear predictor is general very competitive.

In addition, it is important to note that the computations of the ML estimates and the optimal prediction for the Poisson GC model are computationally very intensive, even for a relatively small dataset. To explain, in the previous example (300 location sites) the

	Scenario 1		Scenario 2		Scenario 3	
	Р	PGC	Р	PGC	Р	PGC
RMSE _P	1.101410	1.060940	2.792362	2.806396	7.479191	7.751566
RMSE _{PGC}	1.102478	1.048888	2.792176	2.786705	7.474412	7.739166
RMSE _{PLG}	1.182258	1.155893	3.087286	3.135605	8.437809	8.720756
MAE _P	0.839773	0.839212	2.155351	2.217256	5.852510	6.127105
MAE _{PGC}	0.837037	0.825414	2.152248	2.200265	5.847146	6.117363
MAE _{PLG}	0.930753	0.924698	2.453938	2.510324	6.739108	6.981031

Table (2.4) Empirical mean of RMSE and MAE associated with the optimal linear predictor of the proposed Poisson random field (P), the optimal predictor based on the Poisson GC (PGC) and Poisson LG (PLG) approaches when the datasets are simulated from the P and PGC models under three scenarios, namely, Scenario 1 (low counts), Scenario 2 (medium counts) and Scenario 3 (large counts).

time in seconds (measured using the R function system.time in a computer laptop with a 2.4 GHz processor and 8 GB of memory) to estimate, using 80% of the data, and to predict, using 20% of the data, require approximately 2.453 seconds and 10.125 seconds, respectively, for the GC model using the package gcKrig. On the other hand, the computations of the *wpl* estimates and of the Poisson optimal linear prediction require 0.685 seconds and 0.125 seconds, respectively, using the package GeoModels. As a consequence the proposed methods of estimation and prediction are clearly more scalable.

Chapter 3

Modeling zero inflated Poisson spatial data

3.1 A random field with Bernoulli marginal distribution

A Bernoulli random field $B = \{B(\mathbf{s}), \mathbf{s} \in A\}$ is defined by thresholding a weakly stationary Gaussian random field $G = \{G(\mathbf{s}), \mathbf{s} \in A\}$ (see for example Heagerty and Lele, 1998) with $\mathbb{E}(G(\mathbf{s})) = \mu(\mathbf{s})$ and covariance function $Cov(G(\mathbf{s}_1), G(\mathbf{s}_2)) = \alpha^2 \rho(\mathbf{h})$ with $\mathbf{h} = \mathbf{s}_1 - \mathbf{s}_2 \in A$, that is:

$$B(\mathbf{s}) = \mathbb{1}_{(c,+\infty)}(G(\mathbf{s}) + \epsilon(\mathbf{s})).$$
(3.1)

Here $\rho(\mathbf{h})$ is a correlation function, $\mu(\mathbf{s})$ is a deterministic function, for instance $\mu(\mathbf{s}) = X^{\top}(\mathbf{s})\beta$, and $\epsilon(\mathbf{s}) \sim N(0, \tau^2)$ is an optional white noise independent of G. With this definition the marginal distribution of $B(\mathbf{s})$ is given by the probit model

$$\Pr(B(\mathbf{s}) = 1) = 1 - \Pr(B(\mathbf{s}) = 0) = \Phi\left(\frac{\mu(\mathbf{s}) - c}{\sigma}\right), \tag{3.2}$$

where Φ is the univariate standard Gaussian cdf and $\sigma^2 = Var(G(s) + \epsilon(s)) = \tau^2 + \alpha^2$. In the geostatistical terminology such random field is called Gaussian excursion set or truncated Gaussian random function (see for example Lantuèjoul, 2002). Moreover, given two different locations s_i and s_j , we can set $p_k = \Pr(B(s_k) = 1)$, k = i, j and we denote the joint probability of success as

$$p_{11} = \Pr(B(\mathbf{s}_i) = 1, B(\mathbf{s}_j) = 1) = \Phi_2\left(\frac{\mu(\mathbf{s}_i) - c}{\sigma}, \frac{\mu(\mathbf{s}_j) - c}{\sigma}; \rho(\mathbf{h})\right),$$

where $\Phi_2(\cdot, \cdot, \rho(\mathbf{h}))$ is the bivariate standard Gaussian *cdf* with correlation $\rho(\mathbf{h})$. Then the bivariate distribution of *B* can be written as:

$$\Pr(B(\mathbf{s}_i) = u, B(\mathbf{s}_j) = v) = (1 + p_{11} - p_i - p_j)^{(1-u)(1-v)} p_{11}^{uv} (p_i - p_{11})^{u(1-v)} (p_j - p_{11})^{v(1-u)},$$
(3.3)

for $u, v \in \{0, 1\}$. It is easy to see that the covariance function for the Bernoulli random field is:

$$Cov(B(\mathbf{s}_i), B(\mathbf{s}_j)) = Pr(B(\mathbf{s}_1) = 1, B(\mathbf{s}_2) = 1) - Pr(B(\mathbf{s}_i) = 1) Pr(B(\mathbf{s}_j) = 1)$$
$$= p_{11} - p_i p_j.$$

Then, the correlation function $\rho_B(\mathbf{s}_i, \mathbf{s}_j)$ of the non-stationary Bernoulli random field with underlying correlation $\rho(\mathbf{h})$ is given by:

$$\rho_B(\mathbf{s}_i, \mathbf{s}_j) = \frac{p_{11} - p_i p_j}{\sqrt{p_i p_j (1 - p_i)(1 - p_j)}}.$$

Finally, if $\mu(\mathbf{s}_i) = \mu(\mathbf{s}_j) = \mu$ and therefore $p_i = p_j = p$ with $p = \Phi(\mu)$, then the Bernoulli random field is weakly stationary with correlation function given by:

$$\rho_B(\mathbf{h}) = \frac{p_{11} - p^2}{p(1-p)}$$

3.2 Spatial zero inflated Poisson random fields

We extend the proposed Poisson random field to a zero inflated Poisson (ZIP) random field to deal with excess zeros in data. Specifically, let $B = \{B(\mathbf{s}), \mathbf{s} \in A\}$, be a Bernoulli random field such that $B(\mathbf{s}) = \mathbb{1}_{(-\infty,0)}(G(\mathbf{s}))$ where G is a Gaussian random field with $\mathbb{E}(G(\mathbf{s})) = \theta(\mathbf{s})$, unit variance and correlation function $\rho^*(\mathbf{h})$. The marginal probability of having an excess zero is then given by:

$$p(\mathbf{s}) := \Pr(B(\mathbf{s}) = 0) = \Phi(\theta(\mathbf{s})),$$

where Φ is the univariate standard Gaussian *cdf*.

Let N be a Poisson random field with $\mathbb{E}(N(\mathbf{s})) = \lambda(\mathbf{s})$ and underlying correlation $\rho(\mathbf{h})$. If B and N are independent, our proposed ZIP random field is given by the random field $N^* := \{N^*(\mathbf{s}), \mathbf{s} \in A\}$ defined as:

$$N^*(\mathbf{s}) = B(\mathbf{s})N(\mathbf{s}). \tag{3.4}$$

The marginal distribution is given by

$$\Pr(N^*(\mathbf{s}) = n^*(\mathbf{s})) = \begin{cases} p(\mathbf{s}) + (1 - p(\mathbf{s}))e^{-\lambda(\mathbf{s})} & \text{if } n^*(\mathbf{s}) = 0\\ (1 - p(\mathbf{s}))\frac{\lambda(\mathbf{s})^{n^*(\mathbf{s})}e^{-\lambda(\mathbf{s})}}{n^*(\mathbf{s})!} & \text{if } n^*(\mathbf{s}) = 1, 2, \dots \end{cases},$$
(3.5)

with $\mathbb{E}(N^*(\mathbf{s})) = (1 - p(\mathbf{s}))\lambda(\mathbf{s})$ and $\operatorname{Var}(N^*(\mathbf{s})) = \mathbb{E}(N^*(\mathbf{s}))[1 + \frac{p(\mathbf{s})}{1 - p(\mathbf{s})}\mathbb{E}(N^*(\mathbf{s}))]$. Note that the ZIP random field is over-dispersed and when $p(\mathbf{s}) \to 0$ then the Poisson random field is obtained as special case.

We exemplify this feature by generating a realization of a stationary ZIP random field for an increasing probability of excess zeros. Specifically, Figure 3.1 shows the realization and the associated histogram for a ZIP random field with probability of excess zeros $p(\mathbf{s}) = \Phi(-2), \Phi(-1), \Phi(0), \Phi(1)$, respectively. All the cases consider $\lambda(\mathbf{s}) = 5$ and an underlying correlation model given by $\rho(\mathbf{h}) = (1 - ||\mathbf{h}||/0.2)_+^4$. Note that the histogram of the ZIP random field with the smallest probability of excess zeros, as we expected, is quite similar to a Poisson one. Hence, the case when $p(\mathbf{s}) = \Phi(-2)$ will not be considered in the numerical examples.



Figure (3.1) From the top to the bottom. A realization (column a) of a ZIP random field with $p(\mathbf{s}) = \Phi(-2), \Phi(-1), \Phi(0), \Phi(1)$ and its associated histogram (column b), respectively. In all cases $\lambda(\mathbf{s}) = 5$ and $\rho(\mathbf{h}) = (1 - ||\mathbf{h}||/0.2)^4)_+$.

3.2.1 Correlation function

Let N be a Poisson random field with underlying correlation $\rho(\mathbf{h})$ and $\mathbb{E}(N(\mathbf{s}_i)) = \lambda_i$. Let B be a Bernoulli random field with underlying correlation $\rho^*(\mathbf{h})$ and $\Pr(B(\mathbf{s}_i) = 0) = \Phi(\theta(\mathbf{s}_i))$. Setting $p_i = \Pr(B(\mathbf{s}_i) = 0)$, $p_j = \Pr(B(\mathbf{s}_j) = 0)$ and $p_{kl}^* = \Pr(B(s_i) = k, B(s_j) = l)$ for $k, l \in \{0, 1\}$. If B and N are independent, then the covariance of the ZIP random field N* is given by (see proof of Theorem 3.2.1):

$$\operatorname{Cov}(N^*(s_i), N^*(s_j)) = p_{11}^* \operatorname{Cov}(N(s_i), N(s_j)) + \lambda_i \lambda_j \operatorname{Cov}(B(s_i), B(s_j)),$$

where $p_{11}^* = 1 - (p_{01}^* + p_{10}^* + p_{00}^*)$ with $p_{01}^* = p_i - p_{00}^*$, $p_{10}^* = p_j - p_{00}^*$, $p_{00}^* = \Phi_2(\theta(\mathbf{s}_i), \theta(\mathbf{s}_j)); \rho^*(\mathbf{h}))$ and $\Phi_2(\cdot, \cdot, \rho^*(\mathbf{h}))$ is a bivariate standard Gaussian *cdf* with correlation $\rho^*(\mathbf{h})$.

Theorem 3.2.1. Let N^* be a non-stationary zero inflated Poisson random field with underlying correlations $\rho^*(\mathbf{h})$ and $\rho(\mathbf{h})$. Then:

$$\rho_{N^*}(\mathbf{s}_i, \mathbf{s}_j) = \frac{p_{11}^* \rho_N(\mathbf{s}_i, \mathbf{s}_j)}{\sqrt{(1 - p_i)(1 - p_j)(1 + p_i\lambda_i)(1 + p_j\lambda_j)}} + \frac{\sqrt{\lambda_i \lambda_j p_i p_j} \rho_B(\mathbf{s}_i, \mathbf{s}_j)}{\sqrt{(1 + p_i\lambda_i)(1 + p_j\lambda_j)}},$$

with $\mathbf{h} = \mathbf{s}_i - \mathbf{s}_j$.

Proof. See the Appendix.

Corollary 3.2.1. In Theorem 3.2.1, when $\lambda(\mathbf{s}) = \lambda$ and $p_i = p_j = p = \Phi(\theta)$, the zero inflated Poisson random field is weakly stationary with correlation function given by:

$$\rho_{N^*}(\mathbf{h}, \lambda, \theta) = \frac{p_{11}^* \rho_N(\mathbf{h})}{(1-p)(1+p\lambda)} + \frac{p\lambda\rho_B(\mathbf{h})}{1+p\lambda}$$

Proof. See the Appendix.

The latent correlation functions involved in B and N ($\rho^*(\mathbf{h})$ and $\rho(\mathbf{h})$ respectively) can be assumed equal in order to simplify the inference. Note that $\rho_{N^*}(\mathbf{h})$ is well defined at the origin since $\rho_N(\mathbf{h}) = 1$, $\rho_B(\mathbf{h}) = 1$ and $p_{11}^* = 1 - p$, as $\mathbf{h} \to \mathbf{0}$, implying that the zero inflated Poisson random field is weakly stationary and mean square continuous. Additionally, if $\rho(\mathbf{h}) = 0$, then $\rho_{N^*}(\mathbf{h}) = 0$ and if $p \to 0$ then $\rho_{N^*}(\mathbf{h}) = \rho_N(\mathbf{h})$, *i.e.*, it converges to the correlation function of a Poisson random field.

Figure 3.2 illustrates the effect of excess of zeros in the correlation function of the zero inflated Poisson random field. We set $\theta = -2, -1, 0, 1, \lambda = 5$ and an underlying correlation model given by $\rho(\mathbf{h}) = (1 - ||\mathbf{h}||/0.5)^4)_+$.



Figure (3.2) Correlation functions $\rho_{N^*}(\mathbf{h}, \lambda, \theta)$ of the ZIP random field with $\theta = -2, -1, 0, 1, \lambda = 5$ and $\rho(\mathbf{h}) = (1 - ||\mathbf{h}||/\alpha)_+^4$ with $\alpha = 0.5$.

Furthermore, a not mean-square continuous version of the ZIP random field can be obtained by introducing a nugget effect for B and other for N. Specifically, by replacing $\rho^*(\mathbf{h})$ and $\rho(\mathbf{h})$ with $\rho^*(\mathbf{h})$ and $\rho^{**}(\mathbf{h})$ respectively:

$$\rho^{\star}(\mathbf{h}) = (1 - \tau_2^2)\rho^{\star}(\mathbf{h}) + \tau_2^2 \mathbb{1}_0(||\mathbf{h}||),$$

$$\rho^{\star\star}(\mathbf{h}) = (1 - \tau_1^2)\rho(\mathbf{h}) + \tau_1^2 \mathbb{1}_0(||\mathbf{h}||),$$

where $0 \leq \tau_1^2, \tau_2^2 < 1$, represents the underlying nugget effects.

3.2.2 Bivariate distribution

The bivariate distribution of the zero inflated Poisson random field can be written in terms of the marginal and bivariate distributions of the underlying Bernoulli and Poisson random fields. Therefore, similar to the case of Poisson random fields, we analyze the following cases separately: (a) n = m = 0, (b) n = 0, m > 0, (c) m = 0, n > 0 and (d) n > 0, m > 0.

Moreover, we set $p_{nm} = \Pr(N(\mathbf{s}_i) = n, N(\mathbf{s}_j) = m)$, $\lambda_i = \lambda(\mathbf{s}_i)$, $\lambda_j = \lambda(\mathbf{s}_j)$, $p(\mathbf{s}_i) = p_i$ and $p(\mathbf{s}_j) = p_j$. Thus, the bivariate distribution is given by the following result.

Theorem 3.2.2. Let N^* be a zero inflated Poisson random field with underlying correlation $\rho(\mathbf{h})$ and mean $\mathbb{E}(N^*(\mathbf{s}_k)) = (1 - p_k)\lambda_k$. Then the bivariate distribution $\Pr(N^*(\mathbf{s}_i) = n, N^*(\mathbf{s}_j) = m)$ is given by:

(a) Case n = m = 0:

$$\Pr(N^*(s_i) = 0, N^*(s_j) = 0) = p_{00}^* + p_{01}^* \Pr(N(s_j) = 0) + p_{10}^* \Pr(N(s_i) = 0) + p_{11}^* p_{00}.$$

(b) Case n = 0, m > 0:

$$\Pr(N^*(s_i) = 0, N^*(s_j) = m) = p_{01}^* \Pr(N(s_j) = m) + p_{11}^* p_{0m}$$

(c) Case n > 0, m = 0:

$$\Pr(N^*(s_i) = n, N^*(s_j) = 0) = p_{10}^* \Pr(N(s_i) = n) + p_{11}^* p_{n0}$$

(d) Case n > 0, m > 0:

$$\Pr(N^*(s_i) = n, N^*(s_j) = m) = p_{11}^* p_{nm}.$$

Proof. See the Appendix.

It is important to point out that if $\rho_{N^*}(\mathbf{h}) = 0$ then the bivariate distribution can be written as the product of two independent zero inflated Poisson distributions. Therefore,

since $\rho(\mathbf{h}) = 0$ implies $\rho_{N^*}(\mathbf{h}) = 0$, pairwise independence of the underlying Gaussian random field implies pairwise independence of the zero inflated Poisson random field.

We now compare the type of bivariate dependence induced by the proposed model and the GC one when $\lambda = 5$ and $\theta = -1$.



Figure (3.3) For each row (from left to right): bivariate ZIP GC distribution, our proposed bivariate ZIP distribution and the difference between them. The first, second and third row are obtained setting $\rho(\mathbf{h}) = 0.1, 0.5, 0.9$ for the underlying correlation.

Figure (3.3) (from left to right) presents the bivariate GC distribution, the bivariate zero inflated Poisson distribution in Theorem 3.2.2 and a coloured image representing the differences between them. As in the previous chapter, a positive value of the difference implies that the probabilities associated with the bivariate distribution in Theorem 3.2.2 are greater than the probabilities of the bivariate GC one. Only the probabilities $Pr(N(\mathbf{s}_i) = n, N(\mathbf{s}_j) = m)$ for n, m = 0, 1, ..., 10 are considered in the plots. The first, second and third rows consider increasing levels of underlying correlations $\rho(\mathbf{h}) = 0.1, 0.5, 0.9$.

Similar to the Poisson case, the underlying correlation increases as the difference between the bivariate distributions becomes larger. Note that the probability of the bivariate GC distribution at the pair (0,0) is larger than the bivariate distribution in Theorem 3.3. Moreover, the bivariate probabilities of the bivariate GC distribution at the pairs (n,0) and (0,m) are bigger as well as n and m are closest to zero. On the other hand, the probabilities of the proposed bivariate distribution tend to be larger along the main diagonal and, when n and m are furthest from zero.

3.3 Simulation studies

3.3.1 Performance of the weighted pairwise likelihood estimation

We set up the simulations as in the previous chapter, *i.e.*, considering 1000 realizations from a stationary spatial zero inflated Poisson random field observed at $s_i \in [0, 1]^2$, i = 1, ..., l, l = 441. The grid for the spatial locations is the same perturbed regular grid that we presented in Section 2.4.1.

For the zero inflated Poisson random field we first assume a constant increasing mean and success probability that is $\lambda(\mathbf{s}) = e^{\beta}$ with $\beta = \log(5), \log(20), p(\mathbf{s}) = \Phi(\theta)$ with $\theta = -1, 0, 1$ and an underlying isotropic correlation model $\rho(\mathbf{h}) = (1 - ||\mathbf{h}||/\alpha)^4_+$ with $\alpha = 0.2$. As outlined in Section 1.1.5, the use of a compactly supported correlation function clearly simplify the computation of the bivariate zero inflated Poisson distribution in Theorem 3.2.2. The misspecified *wpl* is obtained when the Poisson random field in 3.4 is replaced with a Gaussian random field with mean and variance equal to $\lambda(\mathbf{s})$ and $\rho_N(\mathbf{h})$ as correlation function, *i.e.*, $N^*(\mathbf{s}) = B(\mathbf{s})G(\mathbf{s})$ with $\mathbb{E}(G(\mathbf{s})) = \operatorname{Var}(G(\mathbf{s})) = \lambda(\mathbf{s})$ and $\rho_G(\mathbf{h}) = \rho_N(\mathbf{h})$. Thus, the bivariate distribution for the misspecified zero inflated Poisson is given by the following theorem.

Theorem 3.3.1. Let N^* be a misspecified zero inflated Poisson random field with underlying correlation $\rho(\mathbf{h})$ and mean $\mathbb{E}(N^*(\mathbf{s}_k)) = (1 - p_k)\lambda_k$. Then the bivariate distribution $f_{N^*(\mathbf{s}_i),N^*(\mathbf{s}_i)}(n,m)$ is given by:

(a) Case n = m = 0:

$$f_{N^*(\mathbf{s}_i),N^*(\mathbf{s}_j)}(0,0) = p_{00}^* + p_{01}^* f_{G(s_i)}(0) + p_{10}^* f_{G(s_j)}(0) + p_{11}^* f_{G(s_i),G(s_j)}(0,0).$$

(b) Case n = 0, m > 0:

$$f_{N^*(\mathbf{s}_i),N^*(\mathbf{s}_j)}(0,m) = p_{01}^* f_{G(s_j)}(m) + p_{11}^* f_{G(s_i),G(s_j)}(0,m)$$

(c) Case n > 0, m = 0:

$$f_{N^*(\mathbf{s}_i),N^*(\mathbf{s}_i)}(n,0) = p_{10}^* f_{G(s_i)}(n) + p_{11}^* f_{G(s_i),G(s_i)}(n,0).$$

(d) Case n > 0, m > 0:

$$f_{N^*(\mathbf{s}_i),N^*(\mathbf{s}_j)}(n,m) = p_{11}^* f_{G(s_i),G(s_j)}(n,m).$$

Proof. See the Appendix.

We study the performance of the zero inflated Poisson wpl and the misspecified wpl. In the (misspecified) wpl estimation we consider a cut-off weight function as in (2.7) with $\xi = 0.1$.

As we can observe in Table 3.1, the ZIP *wlp* shows the best performance throughout the six proposed scenarios, specially at the estimation of the spatial dependence parameter. Note that the performance of the two methods is quite similar for large number of zeros, *i.e.*, $\theta = 1$. Moreover, the performance similarity gap decreases for a large β .

	ZIP	wpl	MZIF	• wpl
	Bias	MSE	Bias	MSE
$\beta = \log(5)$	-0.001589	0.001305	-0.006626	0.001398
$\theta = 0$	-0.001195	0.009364	-0.011589	0.009651
$\alpha = 0.2$	-0.003197	0.000675	-0.001500	0.000746
$\beta = \log(5)$	-0.003955	0.003612	-0.009775	0.003984
$\theta = 1$	0.002750	0.011497	-0.002673	0.011572
$\alpha = 0.2$	-0.004372	0.001347	-0.003432	0.001380
$\beta = \log(5)$	-0.000839	0.000763	-0.005567	0.000813
$\theta = -1$	-0.007386	0.011877	-0.037222	0.014200
$\alpha = 0.2$	-0.005195	0.000710	-0.000850	0.000816
$\beta = \log(20)$	-0.000201	0.000311	-0.000309	0.000312
$\theta = 0$	-0.001863	0.009799	-0.001869	0.009800
$\alpha = 0.2$	-0.002982	0.000644	-0.003021	0.000653
$\beta = \log(20)$	0.000840	0.000885	0.000616	0.000896
$\theta = 1$	0.005714	0.011995	0.005707	0.011996
$\alpha = 0.2$	-0.004833	0.001338	-0.004760	0.001350
$\beta = \log(20)$	-0.000653	0.000205	-0.000763	0.000206
$\theta = -1$	-0.003308	0.012204	-0.003329	0.012207
$\alpha = 0.2$	-0.004732	0.000618	-0.004894	0.000649

Table (3.1) Bias and MSE associated with ZIP *wpl* and misspecified ZIP (MZIP) *wpl* when the true random field is ZIP with $\lambda(\mathbf{s}) = e^{\beta}$, $p = \Phi(\theta)$ and $\rho(\mathbf{h}) = (1 - ||\mathbf{h}||/\alpha)_{+}^{4}$.

We also study the performance of the proposed methods when estimating a non stationary version of the ZIP random field. Under the previous simulation setting we change the constant mean of the ZIP random field by considering a regression model that is $\lambda(\mathbf{s}) = \exp\{\beta + \beta_1 u_1(\mathbf{s}) + \beta_2 u_2(\mathbf{s})\}\$ with $\beta = 2.0$, $\beta_1 = 0.8$ and $\beta_2 = -1.2$ where $u_1(\mathbf{s})$ and $u_2(\mathbf{s})$ are independent realizations from a standard uniform random variable, and we set the success probability $p(\mathbf{s}) = \Phi(\theta)$ with $\theta = 1$. Table 3.2 shows the bias and mean

squared error associated with β , β_1 , β_2 , θ and α for the two estimation methods and Figure 3.4 shows the associated centred box-plots. Overall, the ZIP *wpl* method shows the highest statistical efficiency. Moreover, the smallest MSE belongs to it for each estimation.

	ZIP	wpl	MZIP wpl		
	Bias	MSE	Bias	MSE	
$\beta = 2$	0.004371	0.019139	-0.000514	0.020694	
$\beta_1 = 0.8$	-0.007326	0.034080	-0.004460	0.036361	
$\beta_2 = -1.2$	-0.009287	0.031906	-0.014940	0.034470	
$\theta = 1$	0.003800	0.012842	0.000748	0.012874	
$\alpha = 0.2$	-0.005598	0.001483	-0.005108	0.001485	

Table (3.2) Bias and MSE associated with ZIP *wpl* and misspecified ZIP (MZIP) *wpl* when estimating a non-stationary ZIP random field with $\lambda(\mathbf{s}) = \exp\{\beta + \beta_1 u_1(\mathbf{s}) + \beta_2 u_2(\mathbf{s})\}, p = \Phi(\theta)$ and $\rho(\mathbf{h}) = (1 - ||\mathbf{h}||/\alpha)_+^4$.



Figure (3.4) Centred box-plots of estimates under ZIP *wpl* (ZIP) and misspecified ZIP (MZIP) *wpl* when estimating a non-stationary Poisson random field with $\lambda(\mathbf{s}) = \exp\{\beta + \beta_1 u_1(\mathbf{s}) + \beta_2 u_2(\mathbf{s})\}, \beta = 2, \beta_1 = 0.8, \beta_2 = -1.2, p = \Phi(\theta), \theta = 1 \text{ and } \rho(\mathbf{h}) = (1 - ||\mathbf{h}||/\alpha)_+^4$ with $\alpha = 0.2$.

Finally, we consider a simulation scheme under the spatio-temporal setting. Specifically,

we consider 1000 realizations from a non stationary space-time ZIP random field observed at $\mathbf{s}_i \in [0,1]^2$, $i = 1, \ldots, l$, l = 40 spatial location sites uniformly distributed in the unit square and $t_1^* = 0, t_2^* = 0.25, \ldots, t_{25}^* = 6$, 25 temporal instants. We consider a regression model for the spatio-temporal mean $\lambda(\mathbf{s}, t^*) = \exp\{\beta + \beta_1 u_1(\mathbf{s}, t^*) + \beta_2 u_2(\mathbf{s}, t^*)\}$, where $u_k(\mathbf{s}, t^*)$, k = 1, 2 are independent realizations from a standard uniform random variable, and a spatio-temporal success probability $p(\mathbf{s}, t^*) = \Phi(\theta)$. As in the previous simulation, we set $\beta = 2.0, \beta_1 = 0.8, \beta_2 = -1.2$ and $\theta = 1$. Additionally, as underlying space-time correlation we use a simple isotropic and temporal symmetric space-time Wendland separable model $\rho(\mathbf{h}, t^*) = (1 - ||\mathbf{h}||/\alpha_{\mathbf{s}})_+^4(1 - |t^*|/\alpha_{t^*})_+^4$ with $\alpha_{\mathbf{s}} = 0.2$ and $\alpha_{t^*} = 1$, where $t^* = t_i^* - t_j^*$ with $i, j \in \{1, 2, \ldots, 25\}$. Finally, in the (misspecified) *wpl* estimation we consider a cut-off weight function as in (2.7) extended to the space time case with $\xi_{\mathbf{s}} = 0.2$ and $\xi_{t^*} = 0.25$.

The results are depicted in Table 3.3 where bias and mean squared error associated with β , β_1 , β_2 , θ and α_s , α_{t^*} for the two methods of estimation are shown. In addition, Figure 3.5 shows the associated box-plots.

	ZIP	wpl	MZIP wpl		
	Bias	MSE	Bias	MSE	
$\beta = 2$	0.004617	0.007276	0.001995	0.007567	
$\beta_1 = 0.8$	-0.009988	0.012907	-0.008197	0.013298	
$\beta_2 = -1.2$	-0.001096	0.013481	-0.006318	0.014104	
$\theta = 1$	0.001253	0.003983	-0.001686	0.003994	
$\alpha_s = 0.2$	-0.006759	0.003840	-0.006509	0.003865	
$\alpha_{t^*} = 1$	0.002967	0.024062	0.006996	0.024988	

Table (3.3) Bias and MSE associated with ZIP *wpl* and misspecified ZIP (MZIP) *wpl* when estimating a non-stationary ZIP random field with $\lambda(\mathbf{s}, t^*) = \exp\{\beta + \beta_1 u_1(\mathbf{s}, t^*) + \beta_2 u_2(\mathbf{s}, t^*)\}, p(\mathbf{s}, t^*) = \Phi(\theta) \text{ and } \rho(\mathbf{h}) = (1 - ||\mathbf{h}||/\alpha)_+^4.$



Figure (3.5) Centred box-plots of estimates under ZIP wpl (ZIP) and misspecified ZIP wpl (MZIP) when estimating a non-stationary Poisson random field with $\lambda(\mathbf{s}, t^*) = \exp\{\beta + \beta_1 u_1(\mathbf{s}, t^*) + \beta_2 u_2(\mathbf{s}, t^*)\}, \beta = 2, \beta_1 = 0.8, \beta_2 = -1.2, p(\mathbf{s}, t^*) = \Phi(\theta), \theta = 1 \text{ and } \rho(\mathbf{h}, t^*) = (1 - ||\mathbf{h}||/\alpha_{\mathbf{s}})^4_+ (1 - |t^*|/\alpha_{t^*})^4_+ \text{ with } \alpha_s = 0.2, \alpha_{t^*} = 1.$

3.3.2 Performance of the optimal linear prediction

In this section, we compare the performance of the optimal linear predictor of the proposed ZIP random field with the optimal predictors based on the ZIP GC and ZIP LG approaches. We propose the following optimal linear predictor for the ZIP random field:

$$\widehat{N^*(\mathbf{s}_0)} = (1-p)\lambda(\mathbf{s}_0) + \boldsymbol{c}^{\top} \Sigma^{-1} (\boldsymbol{N^*} - (1-p)\boldsymbol{\lambda}), \qquad (3.6)$$

where

$$\begin{split} \boldsymbol{\lambda} &= (\lambda(\mathbf{s}_1), \dots, \lambda(\mathbf{s}_l))^{\top}, \\ \boldsymbol{c} &= [(1-p)\sqrt{\lambda(\mathbf{s}_0)\lambda(\mathbf{s}_i)(1+p\lambda(\mathbf{s}_0))(1+p\lambda(\mathbf{s}_i))}\rho_{N^*}(\mathbf{s}_0, \mathbf{s}_i)]_{i=1}^l \quad \text{and} \\ \boldsymbol{\Sigma} &= [(1-p)\sqrt{\lambda(\mathbf{s}_i)\lambda(\mathbf{s}_j)(1+p\lambda(\mathbf{s}_i))(1+p\lambda(\mathbf{s}_j))}\rho_{N^*}(\mathbf{s}_i, \mathbf{s}_j)]_{i,j=1}^l \end{split}$$

is the variance-covariance matrix. The associated mean squared error is:

$$MSE(\widehat{N^*(\mathbf{s}_0)}) = (1-p)\lambda(\mathbf{s}_0)(1+p\lambda(\mathbf{s}_0)) - \boldsymbol{c}^{\top}\Sigma^{-1}\boldsymbol{c}.$$

The simulation configuration is given by the following steps:

- 1. Set j = 1. Repeat until j = 100.
- 2. Simulate the j-th spatial dataset from the proposed ZIP random field by considering 300 location sites uniformly distributed on the unit square.
- 3. Set k = 1. Repeat until k = 50.
- 4. Randomly split the *j*-th dataset by using 80% of the data for estimation and 20% as the validation dataset.
- 5. Estimate using *wpl* under our model and using ML for the ZIP GC and ZIP LG models.
- 6. Compute the optimal linear predictor (3.6) and the optimal predictor for the ZIP GC and ZIP LG models at the coordinates associated with the validation dataset, given the estimates obtained at the previous step.
- 7. Compute, for each model, $RMSE_k$ and MAE_k .
- 8. k = k + 1.

9. Compute, for each model
$$\overline{\text{RMSE}}_j = \sum_{k=1}^{50} \text{RMSE}_k/50$$
 and $\overline{\text{MAE}}_j = \sum_{k=1}^{50} \text{MAE}_k/50$.

10. j = j + 1.

11. Compute, for each model RMSE =
$$\sum_{j=1}^{100} \overline{\text{RMSE}}_j / 100$$
 and MAE = $\sum_{j=1}^{100} \overline{\text{MAE}}_j / 100$.

This numerical experiment has been replicated by simulating (at step 2) from a ZIP GC random field. The simulation settings have been chosen such that the means (and variances) are the same for both models. Specifically, we consider three scenarios with $\lambda = 5$ and increasing probability of excess zeros, that is $\theta = -1$ (scenario 1), $\theta = 0$ (scenario 2) and $\theta = 1$ (scenario 3). This election generates a ZIP marginal distribution with low, medium and large zeros, respectively. Additionally, as the underlying correlation model, we consider $\rho(\mathbf{h}) = e^{-3||\mathbf{h}||/\alpha}$, with $\alpha = 0.2$ for the ZIP GC model and $\alpha = 0.4$ for the ZIP model. This specific setting allows us to obtain similar correlations in both cases.

For the *wpl* estimation of our ZIP model, we set $\xi = 0.05$ in (2.7). For the ZIP GC and ZIP LG, we use the ML estimation method implemented in the R (R Core Team, 2020) packages gcKrig (Han and Oliveira, 2018) and INLA (Rue et al., 2009; Lindgren et al., 2011; Martins et al., 2013). The gcKrig package implements a variant of the sequential importance sampling algorithm and it is used to approximate the Poisson GC likelihood (see Masarotto and Varin, 2017; Han and Oliveira, 2018, for the computational details). The ZIP LG estimation exploits the integrated nested Laplace approximation, under a Bayesian framework, considering the default settings for the prior probability distribution of the parameters. Finally, the ZIP GC model's optimal prediction, which involves the evaluation of an *n*-dimensional integral, is approximated using a variant of the sequential importance sampling algorithm, as in the estimation step.

The results of the experiment are presented in Table 3.4. We can observe that the prediction using the ZIP LG is the worst in all scenarios. On the other hand, the performance predictions of our ZIP and ZIP GC models are quite similar in all scenarios. However, as in the Poisson case, the proposed optimal linear predictor is very competitive.

Finally, we compare the computations of the estimates and prediction between the proposed ZIP and the ZIP GC model. We use the same system configuration presented in Section 2.4.2. On the one hand, for the GC model, the time to estimate and predict

	Scenario 1		Scenario 2		Scenario 3	
	ZIP	ZIP GC	ZIP	ZIP GC	ZIP	ZIP GC
RMSEZIP	2.152476	2.026175	2.268950	2.292984	1.594206	1.730472
RMSEZIPGC	2.166845	1.968744	2.295804	2.234183	1.607570	1.707594
RMSEZIPLG	3.257610	2.426053	3.656097	3.629683	3.338754	4.236931
MAEZIP	1.637398	1.620287	1.662866	1.878494	0.855923	1.063258
MAEZIPGC	1.637634	1.564363	1.660078	1.751010	0.815557	0.983061
MAEZIPLG	2.237918	1.949775	3.059709	3.169689	2.234141	3.793925

Table (3.4) Empirical mean of RMSE and MAE associated with the optimal linear predictor of the proposed ZIP random field (ZIP), the optimal predictor based on the ZIP GC (ZIPGC) and ZIP LG (ZIPLG) approaches when the datasets are simulated from the ZIP and ZIPGC models under three scenarios, namely, Scenario 1 (low zeros), Scenario 2 (medium zeros) and Scenario 3 (large zeros).

was 8.512 seconds and 52.489 seconds, respectively. On the other hand, for the proposed ZIP model, the computations for estimation and prediction required 3.524 seconds and 0.215 seconds, respectively. Therefore, the proposed estimation and prediction methods are more scalable than the ones proposed by GC approach.

Chapter 4

Modeling over-dispersed spatial data

4.1 A random field with Erlang marginal distribution

Let G_1, G_2, \ldots be a infinite sequence of a zero mean and unit variance weakly stationary Gaussian random field with correlation function $\rho(\mathbf{h})$. A random field $\Lambda = \{\Lambda(\mathbf{s}), \mathbf{s} \in A\}$ defined as:

$$\Lambda(\mathbf{s}) := \frac{\mu(\mathbf{s})}{2\kappa} \sum_{k=1}^{2\kappa} G_k^2(\mathbf{s}), \tag{4.1}$$

where $\mu(\mathbf{s}) > 0$ is a non-random function, is a random field with marginal Erlang distribution denoted by $\Lambda(\mathbf{s}) \sim \operatorname{Erlang}\left(\kappa, \frac{\kappa}{\mu(\mathbf{s})}\right)$, where κ is the shape parameter and $\kappa/\mu(\mathbf{s})$ is the rate parameter. Thus, the mean and variance of the random field are $\mathbb{E}(\Lambda(\mathbf{s})) = \mu(\mathbf{s})$ and $\operatorname{Var}(\Lambda(\mathbf{s})) = \frac{\mu(\mathbf{s})^2}{\kappa}$, respectively. The covariance function is:

$$\operatorname{Cov}(\Lambda(\mathbf{s}_i), \Lambda(\mathbf{s}_j)) = \frac{\mu(\mathbf{s}_i)\mu(\mathbf{s}_j)}{\kappa} \rho^2(\mathbf{h}),$$

and the correlation function is $\rho_{\Lambda}(\mathbf{h}) = \rho^2(\mathbf{h})$.

Moreover, setting $\lambda(\mathbf{s}_i) = \lambda_i$ and $\lambda(\mathbf{s}_j) = \lambda_j$, the bivariate Erlang *pdf* is given by

(Kibble, 1941; Vere-Jones, 1997; Bevilacqua et al., 2020):

$$f_{\Lambda_{ij}}(\lambda_i,\lambda_j) = \frac{\kappa^{\kappa+1}(\mu(\mathbf{s}_i)\mu(\mathbf{s}_j))^{-\frac{\kappa+1}{2}}}{(1-\rho(\mathbf{h})^2)\Gamma(\kappa)} \left(\frac{\lambda_i\lambda_j}{\rho(\mathbf{h})^2}\right)^{\frac{\kappa-1}{2}} \times e^{-\frac{\kappa}{(1-\rho(\mathbf{h})^2)}\left(\frac{\lambda_i}{\mu(\mathbf{s}_i)} + \frac{\lambda_j}{\mu(\mathbf{s}_j)}\right)} I_{\kappa-1}\left(\frac{2\kappa\rho(\mathbf{h})}{1-\rho(\mathbf{h})^2}\sqrt{\frac{\lambda_i\lambda_j}{\mu(\mathbf{s}_i)\mu(\mathbf{s}_j)}}\right)$$
(4.2)

An interesting result highlighted by Bevilacqua et al. (2020) shows that pairwise independence of the underling Gaussian random field implies pairwise independence of the Erlang random field. Moreover, when $\rho_{\Lambda}(\mathbf{h}) = 0$ the bivariate distribution in 4.2 can be written as the product of two independent Erlang distributions. This result will be useful for the next section.

4.2 Spatial Poisson-Erlang mixture random fields

In this section, we propose to extent the Mixed Poisson models into Mixed Poisson random fields. We use the hierarchical specification of the Mixed Poisson models as starting point, and we extent it for random fields which have Mixed Poisson distribution marginals.

Let $Y = \{Y(s), s \in A\}$ be a positive continuous random field. A Mixed Poisson random field is defined as

$$M := \{ M(\mathbf{s}), \mathbf{s} \in A \},\tag{4.3}$$

such that, hierarchically

- $M(\mathbf{s})|Y(\mathbf{s}) \sim \text{Poisson}(Y(\mathbf{s})),$
- $\operatorname{Cov}(M(\mathbf{s}_i), M(\mathbf{s}_j)|Y(\mathbf{s}_i), Y(\mathbf{s}_j)) = \rho_N(\mathbf{s}_i, \mathbf{s}_j).$

By abuse of notation, we can denote M|Y as a Poisson random field defined in 2.2 with $\mathbb{E}(M(\mathbf{s})|Y(\mathbf{s})) = \operatorname{Var}(M(\mathbf{s})|Y(\mathbf{s})) = Y(\mathbf{s})$ and correlation function given by Theorem 2.2.2. Moreover, we set $\rho(\mathbf{h})$ as the underlying correlation function of Y and M|Y, *i.e.*, the random fields Y and M|Y have the same "parent" Gaussian random field.

Note that the marginal probability distribution of the proposed Mixed Poisson random field is a Mixed Poisson distribution given by (Titterington et al., 1985; Lindsay, 1995; McLachlan and Peel, 2004):

$$\Pr(M(\mathbf{s}) = m) = \int_{\mathbb{R}^+} \Pr(N(\mathbf{s}) = m | y(\mathbf{s})) f_{Y(\mathbf{s})}(y(\mathbf{s})) dy(\mathbf{s}).$$
(4.4)

with $\mathbb{E}(M(\mathbf{s})) = \mathbb{E}(Y(\mathbf{s}))$, $\operatorname{Var}(M(\mathbf{s})) = \operatorname{Var}(Y(\mathbf{s})) + \mathbb{E}(Y(\mathbf{s}))$. Note that this type of model can handle over-dispersion.

Under this framework, a wide range of over-dispersed random fields can be obtained. In this work, we consider $Y \equiv \Lambda$, *i.e.*, the positive random field is set as an Erlang random field defined in 4.1. Therefore, the marginal probability distribution in 4.4 is given by:

$$\Pr(M(\mathbf{s}) = m) = \frac{\Gamma(\kappa + m)}{m!\Gamma(\kappa)} \left(\frac{\mu(\mathbf{s})}{\kappa + \mu(\mathbf{s})}\right)^m \left(\frac{\kappa}{\kappa + \mu(\mathbf{s})}\right)^{\kappa},$$

which belongs to a marginal probability distribution of a Negative Binomial random variable with

$$\mathbb{E}(M(\mathbf{s})) = \mu(\mathbf{s}),$$

$$\operatorname{Var}(M(\mathbf{s})) = \mu(\mathbf{s}) \left(1 + \frac{\mu(\mathbf{s})}{\kappa}\right).$$

We call κ as a shape parameter and it will control the over-dispersion of the random field.

Hereafter M will be called as a Poisson-Erlang mixture random field with underlying correlation function $\rho(\mathbf{h})$. Note that the Poisson random field is a particular case of this class of random fields when $\kappa \to \infty$, and the highest over-dispersion is achieved when $\kappa = 1$.

We exemplify this feature by generating a realization of a stationary Poisson-Erlang mixture random field for an increasing over-dispersion. Specifically, Figure 4.1 shows the realization and the associated histogram for a Poisson-Erlang mixture random field with shape parameter $\kappa = 1, 5, 100$, respectively. All the cases consider $\mu(\mathbf{s}) = 5$ and an underlying correlation model given by $\rho(\mathbf{h}) = (1 - ||\mathbf{h}||/0.2)_+^4$. Note that the histogram of the Poisson-Erlang mixture random field with the highest shape parameter, as we expected,

is quite similar to a Poisson one. Hence, the case when $\kappa(\mathbf{s}) = 100$ will not be considered in the numerical examples.



Figure (4.1) From the top to the bottom. A realization (column a) of a Poisson-Erlang mixture random field with $\kappa = 1, 5, 100$ and its associated histogram (column b), respectively. In all cases $\mu(\mathbf{s}) = 5$ and $\rho(\mathbf{h}) = (1 - ||\mathbf{h}||/0.2)_+^4$.

Note that if $\mu(\mathbf{s})$ depend on \mathbf{s} then M is not stationary. A parametric specification for the mean is given by $\mu(\mathbf{s}) = e^{X(\mathbf{s})^\top \boldsymbol{\beta}}$, where $X(\mathbf{s}) \in \mathbb{R}^k$ is a vector of covariates and $\boldsymbol{\beta} \in \mathbb{R}^k$.

4.2.1 Correlation function

The following result provides the correlation function $\rho_M(\mathbf{s}_i, \mathbf{s}_j)$ of the non-stationary Poisson-Erlang mixture random field with underlying correlation $\rho(\mathbf{h})$ depending on the regularized Gauss hypergeometric function (Olver et al., 2010) defined by:

$${}_{2}\widetilde{\mathrm{F}}_{1}(a,b;c;x) = \sum_{k=0}^{\infty} \frac{(a)_{k}(b)_{k}x^{k}}{\Gamma(c+k)k!}, \quad |x| < 1,$$
(4.5)

and by analytic continuation elsewhere. Moreover, we set $\nu(\mathbf{s}_i) = \kappa/\mu(\mathbf{s}_i)$ and $\nu(\mathbf{s}_j) = \kappa/\mu(\mathbf{s}_j)$. Furthermore, the function \mathcal{H} is defined as follows:

$$\mathcal{H}\left(\begin{smallmatrix}a&,b\\c&,x,x'\end{smallmatrix}\right) = {}_{2}\widetilde{\mathrm{F}}_{1}(a,b;c;x){}_{2}\widetilde{\mathrm{F}}_{1}(a,b;c;x').$$

Theorem 4.2.1. Let M be a non-stationary Poisson-Erlang mixture random field with underlying correlation $\rho(\mathbf{h})$. Then:

$$\rho_{M}(\mathbf{s}_{i}, \mathbf{s}_{j}) = \rho^{2}(\mathbf{h}) \sqrt{\frac{\nu(\mathbf{s}_{i})\nu(\mathbf{s}_{j})}{(1 + \nu(\mathbf{s}_{i}))(1 + \nu(\mathbf{s}_{j}))}} \\
+ \frac{\rho(\mathbf{h})^{2}(1 - \rho(\mathbf{h})^{2})^{\kappa+1}(\nu(\mathbf{s}_{i})\nu(\mathbf{s}_{j}))^{\kappa-1/2}}{\Gamma(\kappa)((\nu(\mathbf{s}_{i}) + 1)(\nu(\mathbf{s}_{j}) + 1))^{\kappa+1/2}\sqrt{\mu(\mathbf{s}_{i})\mu(\mathbf{s}_{j})}} \times \\
\sum_{r=0}^{\infty} \sum_{\ell=0}^{\infty} \frac{\rho(\mathbf{h})^{2\ell}(\nu(\mathbf{s}_{i})\nu(\mathbf{s}_{j}))^{\ell}\Gamma(r + \kappa + \ell + 1)^{2}}{\ell!\Gamma(\ell + \kappa)((\nu(\mathbf{s}_{i}) + 1)(\nu(\mathbf{s}_{j}) + 1))^{r+\ell}} \mathcal{H}\left(\begin{smallmatrix} 1 & , 1 - \kappa - \ell \\ r + 2 \end{smallmatrix}, -\frac{1}{\nu(\mathbf{s}_{i})}, -\frac{1}{\nu(\mathbf{s}_{j})} \right) \\$$

with $\mathbf{h} = \mathbf{s}_i - \mathbf{s}_j$.

Proof. See the Appendix.

Corollary 4.2.1. In Theorem 4.2.1, when $\mu(\mathbf{s}) = \mu$ and $\nu(\mathbf{s}) = \kappa/\mu$, the Poisson-Erlang mixture random field is weakly stationary with correlation function given by:

$$\rho_{M}(\mathbf{h}) = \rho(\mathbf{h})^{2} \left\{ 1 - \frac{\nu(\nu(1-\rho(\mathbf{h})^{2}))^{\frac{1}{2}} \left(2 + \nu(1-\rho(\mathbf{h})^{2})\right)^{\kappa}}{(1+\nu) \left(4 + \nu(1-\rho(\mathbf{h})^{2})\right)^{\kappa+\frac{1}{2}}} \times \left[{}_{2}\widetilde{F}_{1} \left(\frac{1-\kappa}{2}, -\frac{\kappa}{2}; 1; \frac{4}{(2+\nu(1-\rho(\mathbf{h})^{2}))^{2}} \right) + \frac{\kappa+1}{2+\nu(1-\rho(\mathbf{h})^{2})} {}_{2}\widetilde{F}_{1} \left(\frac{2-\kappa}{2}, \frac{1-\kappa}{2}; 2; \frac{4}{(2+\nu(1-\rho(\mathbf{h})^{2}))^{2}} \right) \right] \right\}$$

Proof. See the Appendix.

Notice that if $\mathbf{h} = 0$ then $\rho_M(\mathbf{h}) = 1$, *i.e.*, $\rho_M(\mathbf{h})$ is well defined at the origin. Consequently, the Poisson-Erlang mixture random field is weakly stationary and mean square continuous. Moreover, if $\rho(\mathbf{h}) = 0$, then $\rho_M(\mathbf{h}) = 0$ and if $\kappa \to \infty$ then $\rho_M(\mathbf{h}) = \rho_N(\mathbf{h})$, *i.e.*, it converges to the correlation function of a Poisson random field. The last implication can be observed in Figure 4.2, since the red and black dot-dashed line overlap.

The effect of over-dispersion in the correlation function of Poisson-Erlang mixture random field is illustrated by the Figure 4.2. We set $\kappa = 1, 5, 100, \mu = 5$ and an underlying correlation model given by $\rho(\mathbf{h}) = (1 - ||\mathbf{h}||/0.5)^4)_+$.



Figure (4.2) Correlation function $\rho_M(\mathbf{h}, \mu, \kappa)$ of the Poisson-Erlang mixture random field with $\kappa = 1, 5, 100, \mu = 5, \rho_N(\mathbf{h}, 5)$ and $\rho(\mathbf{h}) = (1 - ||\mathbf{h}||/\alpha)_+^4$ with $\alpha = 0.5$.

Finally, a Poisson-Erlang mixture random field that is not mean-square continuous can be obtained by introducing a nugget effect, *i.e.*, a discontinuity at the origin of $\rho_M(\mathbf{h})$. This can be achieved by replacing the underlying correlation function $\rho(\mathbf{h})$ with $\rho^*(\mathbf{h}) = \rho(\mathbf{h})(1 - \tau^2) + \tau^2 \mathbb{1}_0(||\mathbf{h}||)$, where $0 \le \tau^2 < 1$ represents the underlying nugget effect.

4.2.2 Bivariate distribution

The following result provides the bivariate distribution of the Poisson-Erlang mixture random field. This distribution can be written in terms of an infinite series depending on the regularized Gauss hypergeometric function defined in (4.5).

We analyse the following cases separately: (a) n = m = 0, (b) $n = 0, m \ge 1$ and $m = 0, n \ge 1$, (c) n = m = 1, 2..., and (d) $n, m \ge 1, n \ne m$. Moreover, we set $\tilde{p}_{nm} = \Pr(M(\mathbf{s}_i) = n, M(\mathbf{s}_j) = m)$, $\mu_i = \mu(\mathbf{s}_i)$, $\mu_i = \mu(\mathbf{s}_i)$, $\nu_i = \nu(\mathbf{s}_i)$, $\nu_j = \nu(\mathbf{s}_j)$ and $\rho = \rho(\mathbf{h})$. We additionally define the function $\tilde{\mathcal{H}}$ as follows:

$$\widetilde{\mathcal{H}}\left(\begin{smallmatrix}a&,a'\\b\\c&,c'\end{smallmatrix},x,x'\right) = {}_{2}\widetilde{\mathrm{F}}_{1}(a,b;c;x){}_{2}\widetilde{\mathrm{F}}_{1}(a',b;c';x').$$

Theorem 4.2.2. Let M be a Poisson-Erlang mixture random field with underlying correlation ρ , mean $\mathbb{E}(M(\mathbf{s}_k)) = \mu_k$ and variance $\operatorname{Var}(M(\mathbf{s}_k)) = \mu_k(1+1/\nu_k)$. Then the bivariate distribution \tilde{p}_{nm} is given by:

(a) Case n = m = 0:

$$\begin{split} \widetilde{p}_{00} &= -1 + \left(\frac{\nu_i}{1+\nu_i}\right)^{\kappa} + \left(\frac{\nu_j}{1+\nu_j}\right)^{\kappa} \\ &+ \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \frac{(\nu_i \nu_j)^{\ell+\kappa-1} \rho^{2k+2\ell} (1-\rho^2)^{\kappa+1} \Gamma(k+\ell+\kappa+1)^2}{\ell! \Gamma(\kappa) \Gamma(\ell+\kappa) ((1+\nu_i)(1+\nu_j))^{k+\ell+\kappa}} \mathcal{H}\left(\begin{smallmatrix} 1 & , 1-\ell-\kappa \\ & k+2 \end{smallmatrix}, -\frac{1}{\nu_i}, -\frac{1}{\nu_j} \right). \end{split}$$

(b) Cases $n \ge 1, m = 0$ and $m \ge 1, n = 0$:

$$\widetilde{p}_{n0} = \widetilde{g}(n, \nu_i, \nu_j, \kappa, \rho), \quad \widetilde{p}_{0m} = \widetilde{g}(m, \nu_i, \nu_j, \kappa, \rho),$$

respectively, where

$$\begin{split} \widetilde{g}(b,x,y,a,\rho) &= \frac{(a)_b x^a}{b!(1+x)^{b+a}} \\ &\quad - \frac{(1-\rho^2)^{a+b}}{(1+x-\rho^2)^b y \Gamma(a)} \left(\frac{xy}{(1+x)(1+y)}\right)^a \\ &\quad \times \sum_{\ell=0}^{\infty} \sum_{\ell_1=0}^{\infty} \frac{(xy)^{\ell_1} \rho^{2\ell+2\ell_1} (\ell_1+a)_{\ell+b} (\ell_1+1)_{\ell+a}}{((1+x)(1+y))^{\ell+\ell_1}} \widetilde{\mathcal{H}} \left(\begin{smallmatrix} b, & 1\\ 1-\ell_1-a & \\ \ell+b+1, & \ell+2 \end{smallmatrix}, -\frac{\rho^2}{1+x-\rho^2}, -\frac{1}{y} \right). \end{split}$$

$$\begin{aligned} (c) \ Case \ n &= m \ge 1; \\ \widetilde{p}_{nn} = (1 - \rho^2)^{n+\kappa} \sum_{k=0}^{\infty} \sum_{\ell_1=0}^{\infty} \frac{(n)_k (\nu_i \nu_j)^{\ell_1 + \kappa} \rho^{2k + 2\ell_1} (\ell_1 + \kappa)_{n+k} (\kappa)_{n+k+\ell_1}}{k! \ell_1! ((\nu_i + 1)(\nu_j + 1))^{n+k+\ell_1 + \kappa}} \\ & \times \left[-\frac{(1 + \nu_i)(1 + \nu_j)}{\nu_i \nu_j} \mathcal{H} \left(\frac{1 \ , 1 - \ell_1 - \kappa}{r+k+1} \ , -\frac{1}{\nu_i} \ , -\frac{1}{\nu_j} \right) \right. \\ & + \frac{(1 + \nu_i)(1 + \nu_j)}{(1 + \nu_i - \rho^2)\nu_j} \mathcal{H} \left(\frac{1 \ , 1 - \ell_1 - \kappa}{r+k+1} \ , -\frac{1}{\nu_i} \ , -\frac{\rho^2}{1 + \nu_j - \rho^2} \right) \right] \\ & + (1 - \rho^2)^{n+\kappa+1} \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \sum_{\ell_1=0}^{\infty} \frac{(n)_\ell (\nu_i \nu_j)^{\ell_1 + \kappa - 1} \rho^{2k+2\ell+2\ell_1} (\ell_1 + \kappa)_{n+\ell+k+1} (\kappa)_{n+\ell+k+\ell_1 + 1}}{\ell! \ell_1! ((\nu_i + 1)(\nu_j + 1))^{n+\ell+k+\ell_1 + \kappa}} \\ & \times \mathcal{H} \left(\frac{1 \ , 1 - \ell_1 - \kappa}{r+k+\ell+2} \ , -\frac{1}{\nu_i} \ , -\frac{1}{\nu_j} \right). \end{aligned}$$

(d) Cases $n \ge 2, m \ge 1$ with n > m, and $m \ge 2, n \ge 1$ with m > n,

$$\widetilde{p}_{nm} = \widetilde{h}(n, m, \nu_i, \nu_j, \kappa, \rho), \quad \widetilde{p}_{nm} = \widetilde{h}(m, n, \nu_j, \nu_i, \kappa, \rho),$$

respectively, where

Proof. See the appendix.

Similarly to the Poisson case, the bivariate distribution can be written as the product of two independent Negative Binomial distributions when $\rho(\mathbf{h})_M = 0$. If $\rho(\mathbf{h})_M = 0$ then

 $\rho(\mathbf{h})_N = 0$ and $\rho(\mathbf{h})_{\Lambda} = 0$. Therefore, the Erlang and Poisson bivariate distributions can be written as the product of two independent Erlang and Poisson distributions, respectively. Using the law of total probability on the Poisson-Erlang mixture random field, it can be shown that its bivariate distribution can be written as the product of two independent Negative Binomial distributions.

Moreover, one of the results from the section 4.2.1 establishes that $\rho(\mathbf{h}) = 0$ implies $\rho(\mathbf{h})_M = 0$. Then, pairwise independence of the underlying Gaussian random field implies pairwise independence of the Poisson-Erlang mixture random field.

We also compare the type of bivariate dependence induced by the proposed model and the GC one when $\mu = 5$ and $\kappa = 5$. Figure (4.3) (from left to right) presents the bivariate GC distribution, the bivariate Poisson-Erlang mixture distribution in Theorem 4.2.2 and a coloured image representing the differences between them. As in the previous chapter, a positive value of the difference implies that the probabilities associated with the bivariate distribution in Theorem 4.2.2 are greater than the probabilities of the bivariate GC one. Only the probabilities $Pr(M(\mathbf{s}_i) = n, M(\mathbf{s}_j) = m)$ for n, m = 0, 1, ..., 12 are considered in the plots. The first, second and third rows consider increasing levels of underlying correlations $\rho(\mathbf{h}) = 0.1, 0.5, 0.9$.

Similar to the Poisson and zero inflated Poisson cases, the larger of the difference between the bivariate distributions, the greater of the underlying correlation. Moreover, the bivariate probabilities of the bivariate GC distribution at the pairs (n, 0) and (0, m)are bigger as well as n and m are closest to zero. Furthermore, the probabilities of the proposed bivariate distribution tend to be smaller along the diagonal and larger outside it.



Figure (4.3) For each row (from left to right): bivariate Negative Binomial GC distribution, our proposed bivariate Poisson-Erlang mixture distribution and the difference between them. The first, second and third row are obtained setting $\rho(\mathbf{h}) = 0.1, 0.5, 0.9$ for the underlying correlation.
4.3 Simulation studies

4.3.1 Performance of the weighted pairwise likelihood estimation

In this case the estimation procedure may require more effort than in previous models. Note that, since the shape parameter κ is an integer, the optimization problem belongs to a general class of optimization problems called mixed-integer nonlinear programming. This type of problem cannot be solved in a simple easy way. However, if the κ parameter is fixed, the estimation problem can be performed using nonlinear optimization as in the previous models.

Therefore, we set κ as a know parameter and assess the performance of the weighted pairwise likelihood estimation. Nevertheless, the shape parameter κ can be estimated using a grid of admissible values or by using the floor and ceiling of the moment estimator, and choosing the integer that reaches the highest value of the logarithm of the weighted pairwise likelihood.

We first consider 1000 realizations from a stationary spatial Poisson-Erlang mixture random field observed at $s_i \in [0, 1]^2$, i = 1, ..., l, l = 225. Specifically, we have considered a regular grid with increments 0.07 over $[0, 1]^2$. Then the grid points have been perturbed, adding a uniform random value on [-0.02, 0.02] to each coordinate. The use of perturbed grid helps to get more stable estimates because different sets of small distances are available and too close location points are avoided.

For the Poisson-Erlang mixture random field we first assume a constant mean and increasing shape that is $\mu(\mathbf{s}) = e^{\beta}$ with $\beta = \log(5)$, $\kappa = 1, 5$ and an underlying isotropic correlation model $\rho(\mathbf{h}) = (1 - ||\mathbf{h}||/\alpha)^4_+$ with $\alpha = 0.2$. As discussed in previous chapters, the use of a compactly supported correlation function clearly simplify the computation of the bivariate Poisson-Erlang mixture distribution in Theorem 4.2.2.

The misspecified wpl is obtained when the Poisson-Erlang mixture random field in 4.3 is replaced with a Gaussian random field with mean equal to $\mu(\mathbf{s})$ and variance equal to $\mu(\mathbf{s})(1 + \mu(\mathbf{s})/\kappa)$ and $\rho_M(\mathbf{h})$ as correlation function.

We study the performance of the Poisson-Erlang wpl and the misspecified wpl. In the (misspecified) wpl estimation we consider a cut-off weight function as in (2.7) with $\xi = 0.1$.

			PEM wpl		MPEM wpl	
			Bias	MSE	Bias	MSE
	$\kappa = 1$	$\beta = \log(5)$	-0.006072	0.006029	-0.012227	0.008489
		$\alpha = 0.2$	-0.026533	0.007898	-0.033060	0.008901
	$\kappa = 5$	$\beta = \log(5)$	-0.003236	0.001955	-0.003694	0.002563
		$\alpha = 0.2$	-0.015549	0.006535	-0.017601	0.006875

Table (4.1) Bias and MSE associated with Poisson-Erlang *wpl* and misspecified PEM (MPEM) *wpl* when the true random field is Poisson-Erlang mixture with $\kappa = 1, 5, \mu(\mathbf{s}) = e^{\beta}$, and $\rho(\mathbf{h}) = (1 - ||\mathbf{h}||/\alpha)_{+}^{4}$.

Table 4.1 shows that the PEM *wlp* performs better throughout all the proposed scenarios. Moreover, for the maximum over-dispersion ($\kappa = 1$) the estimation of the spatial dependence parameter clearly shows the best performance.

We also study the performance of the proposed methods when estimating a non stationary version of the Poisson-Erlang mixture random field. Under the previous simulation settings we change the constant mean of the Poison-Erlang mixture random field by considering a regression model that is $\mu(\mathbf{s}) = \exp\{\beta + \beta_1 u_1(\mathbf{s})\}$ with $\beta = 1$ and $\beta_1 = -0.5$ where $u_1(\mathbf{s})$ are independent realizations from a standard uniform random variable, and we set $\kappa = 1$. Table 4.2 shows the bias and mean squared error associated with β , β_1 , and α when κ is set to 1 for the two methods of estimation and Figure 4.4 show the associated centred box-plots. Once again, the Poisson-Erlang mixture wpl method shows the best statistical efficiency.

	PEM wpl		MPEM wpl	
	Bias MSE		Bias	MSE
$\beta = 1.5$	-0.013716	0.022232	-0.025370	0.029153
$\beta_1 = -0.5$	0.010838	0.072491	0.011591	0.094542
$\alpha = 0.2$	-0.031100	0.008749	-0.032777	0.009544

Table (4.2) Bias and MSE associated with PEM *wpl* and misspecified PEM (MPEM) *wpl* when estimating a non-stationary PEM random field with $\mu(\mathbf{s}) = \exp\{\beta + \beta_1 u_1(\mathbf{s}) + \beta_2 u_2(\mathbf{s})\}, \kappa = 1 \text{ and } \rho(\mathbf{h}) = (1 - ||\mathbf{h}||/\alpha)_+^4.$



Figure (4.4) Centred box-plots of estimates under PEM *wpl* (PEM) and misspecified PEM (MPEM) *wpl* when estimating a non-stationary Poisson random field with $\mu(\mathbf{s}) = \exp\{\beta + \beta_1 u_1(\mathbf{s})\}, \beta = 1.5, \beta_1 = -0.5, \kappa = 1 \text{ and } \rho(\mathbf{h}) = (1 - ||\mathbf{h}||/\alpha)^4_+ \text{ with } \alpha = 0.2.$

Finally, we consider a simulation under the spatio-temporal setting. Specifically, we consider 1000 simulations from a non stationary space-time Poisson-Erlang mixture

random field observed at $\mathbf{s}_i \in [0, 1]^2$, i = 1, ..., l, l = 35 spatial location sites uniformly distributed in the unit square and $t_1^* = 0, t_2^* = 0.25, ..., t_{21}^* = 5$, 21 temporal instants. We consider a regression model for the spatio-temporal mean $\mu(\mathbf{s}, t^*) = \exp\{\beta + \beta_1 u_1(\mathbf{s}, t^*)\}$, where $u_1(\mathbf{s}, t^*)$ are independent realizations from a standard uniform random variable, and a shape κ . As in the previous simulation, we set $\beta = 1.5$, $\beta_1 = -0.5$, and $\kappa = 1$. Additionally, as underlying space-time correlation we use a simple isotropic and temporal symmetric space-time Wendland separable model $\rho(\mathbf{h}, t^*) = (1 - ||\mathbf{h}||/\alpha_{\mathbf{s}})_{+}^4(1 - |t^*|/\alpha_{t^*})_{+}^4$ with $\alpha_{\mathbf{s}} = 0.2$ and $\alpha_{t^*} = 1$, where $t^* = t_i^* - t_j^*$ with $i, j \in \{1, 2, ..., 25\}$. Finally, in the (misspecified) wpl estimation we consider a cut-off weight function as in (2.7) extended to the space-time case with $\xi_{\mathbf{s}} = 0.2$ and $\xi_{t^*} = 0.25$.

The results are depicted in Table 4.3 where bias and mean squared error associated with β , β_1 , β_2 , κ and α_s , α_{t^*} for the two methods of estimation are shown. In addition, Figure 4.5 shows the associated boxplots.

	PEM wpl		MPEM wpl	
	Bias MSE		Bias	MSE
$\beta = 1.5$	-0.001869	0.001869 0.008601 -0.0047		0.012292
$\beta_1 = -0.5$	-0.001181	0.025096	-0.001109	0.036370
$\alpha_s = 0.2$	-0.009451	0.005781	-0.010931	0.007256
$\alpha_t = 1$	-0.018836	0.039714	-0.018692	0.044585

Table (4.3) Bias and MSE associated with PEM *wpl* and misspecified PEM (MPEM) *wpl* when estimating a non-stationary PEM random field with $\mu(\mathbf{s}, t^*) = \exp\{\beta + \beta_1 u_1(\mathbf{s}, t^*)\},\ \kappa = 1$ and $\rho(\mathbf{h}) = (1 - ||\mathbf{h}||/\alpha)_+^4$.



Figure (4.5) Centred box-plots of estimates under PEM *wpl* (PEM) and misspecified PEM *wpl* (MPEM) when estimating a non-stationary Poisson random field with $\mu(\mathbf{s}, t^*) = \exp\{\beta + \beta_1 u_1(\mathbf{s}, t^*)\}, \beta = 1.5, \beta_1 = -0.5, \kappa = 1 \text{ and } \rho(\mathbf{h}, t^*) = (1 - ||\mathbf{h}||/\alpha_{\mathbf{s}})^4_+ (1 - |t^*|/\alpha_{t^*})^4_+$ with $\alpha_s = 0.2, \alpha_{t^*} = 1$.

4.3.2 Performance of the optimal linear prediction

In this section, we compare the performance of the optimal linear predictor of the proposed Poisson-Erlang mixture random field with the optimal predictors based on the Negative Binomial GC and Negative Binomial LG approaches. We propose the following optimal linear predictor for the Poisson-Erlang mixture random field:

$$\widehat{M(\mathbf{s}_0)} = \mu(\mathbf{s}_0) + \boldsymbol{c}^{\top} \Sigma^{-1} (\boldsymbol{M} - \boldsymbol{\mu}), \qquad (4.6)$$

where
$$\boldsymbol{\mu} = (\mu(\mathbf{s}_1), \dots, \mu(\mathbf{s}_l))^{\top}, \boldsymbol{c} = [\sqrt{\mu(\mathbf{s}_0)\mu(\mathbf{s}_i)(1+\mu(\mathbf{s}_0)/\kappa)(1+\mu(\mathbf{s}_i)/\kappa)}\rho_M(\mathbf{s}_0, \mathbf{s}_i)]_{i=1}^l$$

and $\boldsymbol{\Sigma} = [\sqrt{\mu(\mathbf{s}_i)\mu(\mathbf{s}_j)(1+\lambda(\mathbf{s}_i)/\kappa)(1+\mu(\mathbf{s}_j)/\kappa)}\rho_M(\mathbf{s}_i, \mathbf{s}_j)]_{i,j=1}^l$ is the

variance-covariance matrix. The associated mean squared error is:

$$MSE(\widehat{M(\mathbf{s}_0)}) = \mu(\mathbf{s}_0)(1 + \lambda(\mathbf{s}_0)/\kappa) - \boldsymbol{c}^{\top} \boldsymbol{\Sigma}^{-1} \boldsymbol{c}.$$

We consider a simulation setup similar to the one presented in Section 2.4.2. The steps are the following :

- 1. Set j = 1. Repeat until j = 100.
- 2. Simulate the j-th spatial dataset from the proposed PEM random field by considering 300 location sites uniformly distributed on the unit square.
- 3. Set k = 1. Repeat until k = 20.
- 4. Randomly split the *j*-th dataset by using 80% of the data for estimation and 20% as the validation dataset.
- 5. Estimate using *wpl* under our model and using ML for the Negative Binomial GC and Negative Binomial LG models.
- 6. Compute the optimal linear predictor (4.6) and the optimal predictor for the Negative Binomial GC and Negative Binomial LG models at the coordinates associated with the validation dataset, given the estimates obtained at the previous step.
- 7. Compute, for each model, $RMSE_k$ and MAE_k .
- 8. k = k + 1.

9. Compute, for each model $\overline{\text{RMSE}}_j = \sum_{k=1}^{20} \text{RMSE}_k/20$ and $\overline{\text{MAE}}_j = \sum_{k=1}^{20} \text{MAE}_k/20$.

- 10. j = j + 1.
- 11. Compute, for each model RMSE = $\sum_{j=1}^{100} \overline{\text{RMSE}}_j / 100$ and MAE = $\sum_{j=1}^{100} \overline{\text{MAE}}_j / 100$.

This numerical experiment has been replicated by simulating random draws (at step 2) from a Negative Binomial GC random field. The simulation settings have been chosen such that the means (and variances) were the same for both models. Specifically, we consider two scenarios with $\mu = 5$ and increasing shape, that is $\kappa = 1$ (scenario 1) and $\kappa = 5$ (scenario 2). This election generates a Negative Binomial marginal distribution with large and medium over-dispersion, respectively. Additionally, as the underlying correlation model, we consider $\rho(\mathbf{h}) = e^{-3||\mathbf{h}||/\alpha}$, with $\alpha = 0.25$ for the Negative Binomial GC model and $\alpha = 0.45$ for the Poisson-Erlang mixture model. This specific setting allows us to obtain similar correlations in both cases.

For the *wpl* estimation of our Poisson-Erlang mixture model, we set $\xi = 0.1$ in (2.7). For the Negative Binomial GC and Negative Binomial LG, we use the ML estimation method implemented in the R (R Core Team, 2020) packages gcKrig (Han and Oliveira, 2018) and spaMM (Rousset and Ferdy, 2014). The gcKrig package implements a variant of the sequential importance sampling algorithm and it is used to approximate the Negative Binomial GC likelihood (see Masarotto and Varin, 2017; Han and Oliveira, 2018, for the computational details). The Negative Binomial LG estimation exploits Laplace approximations for the likelihood. Finally, Negative Binomial GC model's optimal prediction, which involves the evaluation of an *n*-dimensional integral, is approximated using a variant of the sequential importance sampling algorithm, as in the estimation step.

Table 4.4 summarizes the results of our experiment, showing the RMSE and MAE for each model under the different scenarios and types of model generation. Note that the prediction using the Negative Binomial LG is the worst in all scenarios. The performance predictions in terms of RMSE and MAE of our models and Negative Binomial GC models are quite similar in all scenarios, with a slight preference for the Negative Binomial GC prediction in some of them. However, as in the Poisson case, the proposed optimal linear predictor is very competitive.

Finally, as in the previous chapters, the computations of the ML estimates and the optimal prediction for the Negative Binomial GC model are computationally very intensive. Although the computations of the *wpl* estimates are also computationally very

	Scenario 1		Scenario 2	
	PEM NB GC		PEM	NB GC
RMSEPEM	3.825707	3.760064	2.304298	2.090512
RMSE _{NBGC}	3.826071	3.719257	2.303032	2.083224
RMSE _{NBLG}	4.538499	4.464333	2.610666	2.480698
MAEPEM	2.721068	2.621559	1.735116	1.595893
MAE _{NBGC}	2.728842	2.585248	1.735787	1.589779
MAE _{NBLG}	3.227349	3.085764	2.020598	1.916848

Table (4.4) Empirical mean of RMSE and MAE associated with the optimal linear predictor of the proposed PEM random field (PEM), the optimal predictor based on the NB GC (NBGC) and NB LG (NBLG) approaches when the datasets are simulated from the PEM and Negative Binomial (NB) GC models under three scenarios, namely, Scenario 1 ($\kappa = 1$) and Scenario 2 ($\kappa = 5$).

intensive, but using the misspecified *wpl* estimates we can deal with this issue. As in the previous chapter, we use the system configuration presented in Section 2.4.2 to measure the time in seconds for the estimation and prediction procedures. On the one hand, the GC model require approximately 4.022 seconds for the estimation and 38.668 seconds for the prediction. On the other hand, the computations of the misspecified *wpl* estimates and of the Poisson-Erlang mixture optimal linear prediction require 0.104 seconds and 0.012 seconds, respectively. Even if the computations of the proposed *wpl* estimation are intensives, the proposed method of estimation with the misspecified *wlp* and prediction are scalable.

Chapter 5

Applications to real data

5.1 Application to the reindeer pellet-group survey in Sweden

The faecal pellet count technique is one of the most popular tools for estimating an animal species' abundance. Specifically, this technique uses the number of observed droppings combined with their decay time and the target animal species' defecation rate. With these ingredients, it is possible to obtain an accurate density estimation of an animal population. This method was proposed by Bennett et al. (1940) and has been improved by several authors since then (see for instance Etten and Bennett, 1965; Mayle et al., 1999; Krebs et al., 2001, among others).

The study motivating our research is a reindeer pellet-group survey conducted in the northern forest area of Sweden and previously analysed by Lee et al. (2016). The goal of this survey was to assess the impact of newly established wind farms on reindeer habitat selection. This choice is crucial for the reindeer since it involves trade-offs between fulfilling necessities for feeding, mating, parental care, and risk mitigation of predation (Sivertsen et al., 2016).

Survey data were collected over the years 2009-2010 and present a large number of

zero counts. This situation frequently occurs when spatial species count data are collected because the survey is conducted using a point transect design (Buckland et al., 2001). This design considers a set of K points as systematically spaced points along lines located throughout the survey region, where K should be at least 20 for obtaining robust estimates of the abundance. In our case, the study area was 250 km², which is the distance between each transect of 300 m. On each transect, the distance between each plot 100 m. Figure 5.1, taken from Buckland et al. (2001), presents two examples of point transect surveys in which the points are systematically spaced along lines.



Figure (5.1) Examples of point transect survey design

As mentioned before, the pellet-group survey is a technique that provides a general idea of species distribution over a specific geographic area. To assess the impact of newly established wind farms on reindeer habitat, a reindeer pellet-group survey was conducted on Storliden Mountain in the northern forest area of Sweden (Lee et al., 2016). The dataset, which corresponds to the year 2009, has 357 geo-referenced data $y(s_i)$, i = 1, 2, ..., 357and it consists of pellet-group counts where, a pellet-group is defined as a cluster of 20 or more pellets.

The dataset possesses two challenging features. The first one is that 73.67% of the counts are zeros because the animal might move as it defecates, and some plots present zero pellet-group counts (see Figure 5.2). The second one is that the empirical semi-variogram (see Figure 5.3) exhibits both spatial correlation and a nugget effect.



Figure (5.2) Spatial location of reindeer pellet-group survey data.



Figure (5.3) Empirical semi-variogram of reindeer pellet-group survey data

For analysing the reindeer pellet-group survey data, we consider the proposed ZIP random field and we compare it with the ZIP Gaussian copula (ZIP GC) using the R package gcKrig (Han and Oliveira, 2018). In addition, we consider the ZIP

Log-Gaussian (ZIP LG) random field as implemented in the R package INLA (Rue et al., 2009; Lindgren et al., 2011; Martins et al., 2013) which exploits the integrated nested Laplace approximation, under a Bayesian framework, in the estimation step.

Following the results of Lee et al. (2016), we include the following three covariates: Northwest slopes (NS), Elevation (Eln) and Distance to power lines (DPL). In particular we specify $\lambda(s)$ as:

$$\lambda(\mathbf{s}) = \exp(\beta_0 + \beta_{NS} \mathrm{NS}(\mathbf{s}) + \beta_{Eln} \mathrm{Eln}(\mathbf{s}) + \beta_{DPL} \mathrm{DPL}(\mathbf{s}))$$

The parametrization for the marginal mean and variance are slightly different for the three models. Specifically, assuming a constant probability of excess of zero counts p, the marginal mean and variance specifications are given by $\mathbb{E}(Y(\mathbf{s})) = \lambda(\mathbf{s})(1-p)$ and $\operatorname{Var}(Y(\mathbf{s})) = \mathbb{E}(Y(\mathbf{s})) \left[1 + \frac{p}{1-p}\mathbb{E}(Y(\mathbf{s}))\right]$ for the proposed model, $\mathbb{E}(Y(\mathbf{s})) = \lambda(\mathbf{s})$ and $\operatorname{Var}(Y(\mathbf{s})) = \mathbb{E}(Y(\mathbf{s})) \left[1 + \theta_{GC}\mathbb{E}(Y(\mathbf{s}))\right]$ for the GC model, $\mathbb{E}(Y(\mathbf{s})) = \lambda(\mathbf{s}) \exp(0.5\sigma^2)(1-p)$ and $\operatorname{Var}(Y(\mathbf{s})) = \mathbb{E}(Y(\mathbf{s})) \left[1 + \frac{p}{1-p}\mathbb{E}(Y(\mathbf{s}))\right] + \mathbb{E}(Y(\mathbf{s}))^2 \left[\exp(\sigma^2) - 1\right]$ for the LG model.

Here, p is specified as $\Phi(\theta)$, with $\theta \in \mathbb{R}$, as $\frac{\theta_{GC}}{1+\theta_{GC}}$, with $\theta_{GC} > 0$, and as $\frac{\exp(\theta_{LG})}{1+\exp(\theta_{LG})}$, with $\theta_{LG} \in \mathbb{R}$, respectively, so θ , θ_{GC} , θ_{LG} can be interpreted as over-dispersion parameters. It is important to remark that β_0 can not be compared between the different approaches, but β_{NS} , β_{Eln} and β_{DPL} can be compared.

We assume an underlying exponential correlation model with nugget effect $\rho(\mathbf{h}) = (1 - \tau^2)e^{-||\mathbf{h}||/\alpha} + \tau^2 \mathbb{1}_0(||\mathbf{h}||)$ for the ZIP GC and ZIP LG random fields. On the other hand for the proposed ZIP model we specify $\rho_1(\mathbf{h}) = (1 - \tau_2^2)e^{-||\mathbf{h}||/\alpha} + \tau_2^2 \mathbb{1}_0(||\mathbf{h}||)$ and $\rho_2(\mathbf{h}) = (1 - \tau_1^2)e^{-||\mathbf{h}||/\alpha} + \tau_1^2 \mathbb{1}_0(||\mathbf{h}||)$ that is two different underlying correlation models for *B* and *N* sharing a common exponential correlation model and different nugget effects τ_1^2, τ_2^2 .

We use maximum *wpl* estimation with $\xi = 150$ in (2.7) for our ZIP random field. For the ZIP GC model we perform maximum likelihood estimation as explained in Section 2.4.2, and for ZIP LG model, we perform approximate Bayesian inference using the INLA approach (Rue et al., 2009).

Table 5.1 summarizes the results of the estimates, including their standard error for the

three models. In the case of our ZIP random field, standard errors were computed by using parametric bootstrap (Efron and Tibshirani, 1986). For the Poisson LG model the reported estimates are the means of the posterior distributions with associated standard error.

Note that if β_{DPL} is a positive value, then the counts of pellet-groups increase at larger distances from the power lines, *i.e.*, there is a greater reindeer population far from the wind farms. The estimation of the regression parameters are quite similar for our ZIP and the ZIP GC models with lower standard error estimations for the GC model. On the other hand our ZIP model shows the smallest standard error estimation of the spatial scale parameter α . Finally, the estimates of p, the excess of zero counts, which depend on $\theta_{,,,GC}$ and θ_{LG} for the ZIP, ZIP GC and ZIP LG random fields, are given by 0.481, 0.477, 0.573, respectively.

Finally, we want to assess the predictive performances of the three models. To do so, we randomly choose 80% of the spatial locations (*i.e.*, 286 location sites) for the parameter estimation and use the remaining 20% (*i.e.*, 71 location sites) for the predictions. We repeat this procedure a 100 times, recording the RMSE each time. Specifically, for each *j*-th left-out sample $(y_j(\mathbf{s}_1), y_j(\mathbf{s}_2), \ldots, y_j(\mathbf{s}_{71}))$, we compute

RMSE_j =
$$\left(\frac{1}{71}\sum_{i=1}^{71} (y_j(\mathbf{s}_i) - \widehat{Y}_j(\mathbf{s}_i))^2\right)^{1/2}$$
,

where $\hat{Y}_j(\mathbf{s}_i)$ is the optimal linear predictor for our ZIP random field (computed using the correlation given in Corollary 2.2.1), the optimal predictor for the ZIP GC random field and the mean of the posterior predictive distribution for the ZIP LG random field. Finally, we report in Table 5.1, the empirical mean of the RMSE obtained for each left-out sample, *i.e.*, $\overline{\text{RMSE}} = \sum_{j=1}^{100} \text{RMSE}_j/100$. The ZIP and ZIP GC random fields' clearly outperform the ZIP LG random field in terms of prediction performance. In particular the proposed ZIP random field provides the smallest $\overline{\text{RMSE}}$.

CHAPTER 5. APPLICATIONS TO REAL DATA

	ZIP	ZIP GC	ZIP LG
β_0	-23.423890	-19.096181	-17.905000
	(6.473000)	(2.308969)	(9.514000)
β_{NS}	-0.534781	-0.464979	-0.826000
	(0.469695)	(0.364517)	(0.388000)
β_{Eln}	0.005203	0.003836	0.010000
	(0.004136)	(0.002988)	(0.005000)
β_{DPL}	2.594934	2.060382	1.622000
	(0.742912)	(0.265003)	(1.177000)
θ	-0.048763		
	(1.048744)		
θ_{GC}		0.912081	
		(0.261074)	
θ_{LG}			0.293000
			(0.098000)
α	339.132404	298.926862	685.922925
	(92.981918)	(186.990640)	(354.614010)
$ au_1^2$	0.868296		
	(0.263721)		
$ au_2^2$	0.623987		
	(0.308045)		
τ^2		0.714227	0.084655
		(0.128536)	(0.022749)
σ^2			0.735290
			(0.396332)
RMSE	0.797449	0.801876	0.835611

Table (5.1) Parameter estimates for the reindeer pellet-group survey data obtained under the ZIP, ZIP GC and ZIP LG random fields. The associated standard errors are in parenthesis. The last row shows the associated empirical mean of the RMSE for each model.

5.2 Application to weed from the Bjertorp farm in Sweden

One of the problems in agricultural fields is the infestations of weeds. Farmers are interested in killing them by varying the amount or type of herbicide depending on the intensity of the weed (Webster, 2010). Therefore, accurately characterizing and mapping the spatial distribution of weeds is a useful tool for farmers. Thus, spatial or spatio-temporal models are used to analyse counts of weed (see Donald, 1994; Cardina et al., 1995; Johnson et al., 1996).

The study, which motivates our research, belongs to the Bjertorp farm in the South-West of Sweden and previously analysed by Guillot et al. (2009) and De Oliveira (2013). The dataset consists of weed counts, $y(s_i)$, in frames of 0.5 by 0.75 m at 89 sampling locations, s_1, \ldots, s_{89} on a regular grid (see Figure 5.4). There are not covariates and no spatial trend, thus the mean and variance can be assumed constant.

The dataset possess two challenging features. The first one is that the sample variance (3211.562) is larger than the sample mean (76.40449), *i.e.*, there is evidence that the dataset is over-dispersed. The second one is that the empirical semi-variogram (see Figure 5.5) exhibits spatial correlation.

For analysing the weed data, we consider the proposed PEM random field and we compare it with the Negative Binomial Gaussian copula (NB GC) using the R package gcKrig (Han and Oliveira, 2018). In addition, we consider the Negative Binomial Log-Gaussian (NB LG) random field as implemented in the R package spaMM (Rousset and Ferdy, 2014).

The parametrization for the marginal mean and variance are slightly different for the three models. Specifically, setting $\mu = \exp\{\beta_0\}$, the marginal mean and variance specifications are given by $\mathbb{E}(Y(\mathbf{s})) = \mu$ and $\operatorname{Var}(Y(\mathbf{s})) = \mathbb{E}(Y(\mathbf{s})) \left[1 + \frac{\mathbb{E}(Y(\mathbf{s}))}{\kappa}\right]$ for the proposed model, $\mathbb{E}(Y(\mathbf{s})) = \mu$ and $\operatorname{Var}(Y(\mathbf{s})) = \mathbb{E}(Y(\mathbf{s})) \left[1 + \theta \mathbb{E}(Y(\mathbf{s}))\right]$ for the GC model where $\theta \in \mathbb{R}^+/\{0\}$, $\mathbb{E}(Y(\mathbf{s})) = \mu \exp(0.5\sigma^2)$ and $\operatorname{Var}(Y(\mathbf{s})) =$



Figure (5.4) Spatial location of weed.



Figure (5.5) Empirical semi-variogram of weed data

$$\mathbb{E}(Y(\mathbf{s}))\left[1 + \frac{\mathbb{E}(Y(\mathbf{s}))}{\kappa}\right] + \mathbb{E}(Y(\mathbf{s}))^2 \left[\exp(\sigma^2) - 1\right] \text{ for the LG model.}$$

Now we have to set the shape parameter κ . Following the procedure outlined in Section 4.3.1, two possible candidates are the floor $(\hat{\kappa}_f)$ and ceiling $(\hat{\kappa}_c)$ of the moment estimator

of κ . We compute them as follow:

$$\widehat{\kappa}_{f} = \left[\frac{\left(\frac{1}{89} \sum_{i=1}^{89} y(\mathbf{s}_{i})\right)^{2}}{\left(\frac{1}{89} \sum_{i=1}^{89} \left(y(\mathbf{s}_{i}) - \frac{1}{89} \sum_{i=1}^{89} y(\mathbf{s}_{i})\right)^{2} - \frac{1}{89} \sum_{i=1}^{89} y(\mathbf{s}_{i})}\right] = \lfloor 1.86 \rfloor = 1,$$

$$\widehat{\kappa}_{c} = \left[\frac{\left(\frac{1}{89} \sum_{i=1}^{89} y(\mathbf{s}_{i})\right)^{2}}{\left(\frac{1}{89} \sum_{i=1}^{89} y(\mathbf{s}_{i}) - \frac{1}{89} \sum_{i=1}^{89} y(\mathbf{s}_{i})\right)^{2} - \frac{1}{89} \sum_{i=1}^{89} y(\mathbf{s}_{i})}\right] = \lceil 1.86 \rceil = 2.$$

We assume an underlying exponential correlation model without nugget effect $\rho(\mathbf{h}) = e^{-||\mathbf{h}||/\alpha}$ for the three models. However, the parametrization used by the R package spaMM is $\rho(\mathbf{h}) = e^{-||\mathbf{h}||\alpha^*}$, *i.e.*, the range considered in the other formulation is the reciprocal of α^* .

We use maximum *wpl* estimation with $\xi = 60$ in (2.7) for our PEM random field with the two different settings of κ . The maximum log-composite-likelihood value is -1630.17when $\kappa = 1$ and -1617.75 when $\kappa = 2$. Then, we prefer the model with $\kappa = 2$.

Similarly, we perform maximum likelihood estimation for Negative Binomial LG model with $\kappa = 1, 2$. The maximum log-likelihood value is -475.4877 when $\kappa = 1$ and -467.7218 when $\kappa = 2$. Then, we prefer the model with $\kappa = 2$. For the Negative Binomial GC and Negative Binomial LG models we perform ML estimation as explained in Section 4.3.2.

Table 5.2 summarizes the final results of the estimates, including their standard error for the three models. In the case of our PEM random field, standard errors were computed by using parametric bootstrap as in the previous section.

Finally, we want to assess the predictive performances of the three models. To do so, we randomly choose 80% of the spatial locations (*i.e.*, 71 location sites) for the parameter estimation and use the remaining 20% (*i.e.*, 18 location sites) for the predictions. We repeat this procedure a 100 times, recording the RMSE each time. Specifically, for each

	PEM	PEM GC	PEM LG
β_0	4.343	4.3939	4.207
	(0.1253)	(0.2004)	(0.2521)
θ		0.6488	
		(0.1432)	
α	81.141	57.5346	110.769
	(11.5231)	(17.4152)	
σ^2			0.3604
RMSE	50.40589	50.69107	52.93922

Table (5.2) Parameter estimates for the weed data obtained under the PEM, Negative Binomial GC and Negative Binomial LG random fields. The associated standard errors are in parenthesis. The last row shows the associated empirical mean of the RMSE for each model.

j-th left-out sample $(y_j(\mathbf{s}_1), y_j(\mathbf{s}_2), \dots, y_j(\mathbf{s}_{18}))$, we compute

RMSE_j =
$$\left(\frac{1}{18}\sum_{i=1}^{18} (y_j(\mathbf{s}_i) - \widehat{Y}_j(\mathbf{s}_i))^2\right)^{1/2}$$
,

where $\hat{Y}_j(\mathbf{s}_i)$ is the optimal linear predictor for our PEM random field (computed using the correlation given in Corollary 4.2.1), the optimal predictor for the Negative Binomial GC random field and the Negative Binomial LG random field. Finally, we report in Table 5.2, the empirical mean of the RMSE obtained for each left-out sample, *i.e.*, $\overline{\text{RMSE}} = \sum_{j=1}^{100} \text{RMSE}_j/100$. The PEM and Negative Binomial GC random fields' clearly outperform the Negative Binomial LG random field in terms of prediction performance. In particular the proposed PEM random field provides the smallest $\overline{\text{RMSE}}$.

Chapter 6

Conclusions and discussion

This project has introduced a model based on a Poisson random field, *i.e.*, a random field with Poisson marginal distributions, for regression and dependence analysis when addressing point-referenced count data defined on a spatial Euclidean space. However, the proposed methodology can be easily adapted to other types of data, such as space-time (Gneiting, 2013), areal (Rue and Held, 2005) or spherical data (Gneiting, 2013).

Our model can be viewed as a spatial generalization of the standard Poisson process since it is obtained by considering sequences of independent copies of a random field with an exponential marginal distribution of inter-arrival times in the counting renewal processes framework. The resulting (non-)stationary random field is marginally Poisson distributed and the dependence is indexed by a correlation function as in the Gaussian case.

They key features of the proposed Poisson random field with respect to the well-known hierarchical Poisson Log-Gaussian random field are that its marginal distribution is Poisson distributed and it can be mean square continuous or not. The Poisson Gaussian copula approach shares these good features with our model. However, the generating mechanisms (*i.e* the Poisson process), underlying our model makes it more appealing from interpretability viewpoint.

In our proposal, a possible limitation is that inference based on full likelihood cannot be performed due to the lack of amenable expressions of the associated multivariate distributions. Nevertheless, the simulations studies we conducted showed that our approach based on a weighted pairwise likelihood estimation seems to be an effective solution for estimating the unknown parameters involved in the Poisson, zero inflated Poisson and Poisson-Erlang mixture random field. Another potential limitation is that the optimal predictor that minimizes the mean square prediction error is not available. However, our numerical experiments show that our solution based on optimal linear predictor performs very well if compared with the optimal predictors of the Poisson Gaussian copula and Poisson Log Gaussian models.

A well-known restriction of the Poisson distribution is equidispersion. Unfortunately, this situation is not always observed in real spatial data. The class of random fields proposed in (2.2) can be used to obtain random fields with flexible marginal models that take into account over or under dispersion. In this case, a possible solution is to consider random fields with a more flexible marginal distribution than the exponential marginal distribution, such as the Gamma or Weibull random fields (Bevilacqua et al., 2020). Another alternative for obtaining over-dispersed random fields is considering scale mixtures of Poisson random fields. We use the last alternative to obtain a Poisson-Erlang mixture random field which has Negative Binomial marginals. Even if the computation of the bivariate has a high comtutational cost, it is not the case of the prediction. There are alternatives that allows us to deal with this problem. For instance, we can use the misspecified *wlp* or compactly supported correlation functions. In literature, another way to model over-dispersion is using compound random variables. In a future work, we could generalize it to space (space-time) with the results of this thesis.

We can enrich the proposed new class of counting random fields (see equation 2.2) if we consider positive random fields other than exponential. For instance, if the positive random field is an Erlang random field then the resulting random field is under-dispersed.

In many fields, the data counts may display a feature of excess zeros and be spatially correlated. To deal with this kind of situations, we propose a random field that can handle spatial (spatio-temporal) data with excess of zeros. Our simulation studies reveal that the proposed zero inflated Poisson random field is suitable for modeling excess of zeros in

spatial counts. Moreover, the proposed model performs very well if compared with zero inflated Poisson gaussian copula and zero inflated Poisson Log Gaussian models.

Even if Poisson models do not appear commonly in real data, they can be used as building blocks. In fact, we use the proposed Poisson random field to develop random fields with over dispersion and excess of zeros.

Finally, the application of our model to the reindeer pellet-group survey data in Sweden shows that our approach can be easily adapted to handle spatial count data with an excessive number of zeros. Moreover, the application to the weed counts from the Bjertorp farm shows a good performance of the Poisson-Erlang mixture random field. The results show that the proposed models are good candidates to model real data and an alternative to Gaussian copula models, without forgetting that are they are less computationally expensive.

Appendix A

Proofs of Theorems and Corollaries

A.1 Proofs of Chapter 2

Proof of Corollary 2.2.1. Using Theorem 2.2.2, the correlation function of N when $\lambda(\mathbf{s}_i) = \lambda(\mathbf{s}_j) = \lambda$ has the following expression

$$\rho_{\scriptscriptstyle N}(\mathbf{h}) = \frac{\rho^2(\mathbf{h})(1-\rho^2(\mathbf{h}))}{\lambda} \sum_{r=1}^{\infty} \left(\frac{\gamma\left(r, \frac{\lambda}{1-\rho^2(\mathbf{h})}\right)}{\Gamma(r)} \right)^2,$$

where $\rho(\mathbf{h})$ is the underlying correlation model. The last expression is equivalent to

$$\begin{split} \rho_{N}(\mathbf{h}) &= \frac{\rho^{2}(\mathbf{h})(1-\rho^{2}(\mathbf{h}))}{\lambda} \sum_{r=0}^{\infty} \left(\frac{\gamma \left(r+1, \frac{\lambda}{1-\rho^{2}(\mathbf{h})}\right)}{\Gamma(r+1)} \right)^{2} \\ &= \frac{\rho^{2}(\mathbf{h})(1-\rho^{2}(\mathbf{h}))}{\lambda} \sum_{r=0}^{\infty} \left(\frac{\gamma \left(r+1, \frac{\lambda}{1-\rho^{2}(\mathbf{h})}\right)}{r!} \right)^{2} \\ &= \frac{\rho^{2}(\mathbf{h})\lambda}{1-\rho^{2}(\mathbf{h})^{2}} F_{2}\left((1, 1.5), (2, 3); -\frac{4\lambda}{1-\rho^{2}(\mathbf{h})}\right). \end{split}$$

The last equality follows from the identity (Brychkov, 2008, p. 460):

$$\sum_{k=0}^{\infty} \frac{1}{k(\nu)_k} \gamma \nu + k, z = \frac{z^{2\nu}}{\nu^2} {}_2F_2\left(\left(\nu, \nu + \frac{1}{2}\right), (\nu + 1, 2\nu + 1); -4z\right).$$

Here $_2F_2$ is special case of the generalized hypergeometric function. Now, let $z = \frac{4\lambda}{1-\rho^2(\mathbf{h})}$. Then,

$$\rho_{N}(\mathbf{h}) = \frac{\rho^{2}(\mathbf{h})\lambda}{1-\rho^{2}(\mathbf{h})^{2}}F_{2}\left((1,1.5),(2,3);-z\right).$$

Using the identity (Ng and Geller, 1970, A10b):

$$2z_2F_2\left(\left(1,\nu+\frac{1}{2}\right), (2,2\nu+1); 2z\right) = \frac{1}{\left(\nu-\frac{1}{2}\right)} \left[\left(\frac{2}{z}\right)^{\nu} z \exp(z)\Gamma(\nu+1)\{I_{\nu-1}(z) - I_{\nu}(z)\} - 2\nu\right],$$

 $\rho_{N}(\mathbf{h})$ can be written as follow:

$$\rho_{N}(\mathbf{h}) = \frac{\rho^{2}(\mathbf{h})\lambda}{1-\rho^{2}(\mathbf{h})} \frac{4\left[1-\exp\left(-\frac{z}{2}\right)\left(I_{0}\left(\frac{z}{2}\right)+I_{1}\left(\frac{z}{2}\right)\right)\right]}{z}$$
$$= \rho^{2}(\mathbf{h})\left[1-\exp\left(-\frac{z}{2}\right)\left(I_{0}\left(\frac{z}{2}\right)+I_{1}\left(\frac{z}{2}\right)\right)\right]$$

Proof of Theorem 2.2.1. Using results of Hunter (1974), the bivariate Laplace transformation of the join probability generating function (jpgf) is

$$\mathcal{P}^{\circ}(p,q;c_1,c_2) = \frac{\rho^2(1-c_1)(1-c_2)}{[1-c_1c_2+p+q+(1-\rho^2)pq][1-c_1+p][1-c_2+q]} + \frac{1}{[1-c_1+p][1-c_2+q]}$$

The inverse Laplace transform is given by

$$\mathcal{P}(\lambda(\mathbf{s}_i), \lambda(\mathbf{s}_j); c_1, c_2) = \mathcal{L}^{-1} \left(\frac{\rho^2 (1 - c_1)(1 - c_2)}{[1 - c_1 c_2 + p + q + (1 - \rho^2)pq][1 - c_1 + p][1 - c_2 + q]} \right) \\ + \frac{1}{[1 - c_1 + p][1 - c_2 + q]} \right) \\ = \mathcal{L}^{-1} \left(\frac{\rho^2 (1 - c_1)(1 - c_2)}{[1 - c_1 - c_2 + p + q + (1 - \rho^2)pq][1 - c_1 + p][1 - c_2 + q]} \right) \\ + \mathcal{L}^{-1} \left(\frac{1}{[1 - c_1 + p][1 - c_2 + q]} \right).$$

Note that

$$\mathcal{L}^{-1}\left(\frac{1}{[1-c_1+p][1-c_2+q]}\right) = \mathcal{L}^{-1}\left(\frac{1}{[(1-c_1)+p][(1-c_2)+q]}\right)$$
$$= \exp\left\{-(1-c_1)\lambda(\mathbf{s}_i) - (1-c_2)\lambda(\mathbf{s}_j)\right\}.$$

$$\begin{split} \mathcal{L}_{p,q}^{-1} & \left(\frac{\rho^2 (1-c_1)(1-c_2)}{[1-c_1c_2+p+q+(1-\rho^2)pq][1-c_1+p][1-c_2+q]} \right) \\ &= \rho^2 (1-c_1)(1-c_2) \times \mathcal{L}_q^{-1} \left(\frac{1}{[1-c_2+q]} \mathcal{L}_p^{-1} \left(\frac{1}{[1-c_1c_2+p+q+(1-\rho^2)pq][1-c_1+p]} \right) \right) \\ &= \frac{\rho^2 (1-c_1)(1-c_2)}{(1-\rho^2)} \times \\ & \mathcal{L}_q^{-1} \left(\frac{1}{[(1-c_2)+q]} \frac{1}{[\frac{1}{(1-\rho^2)}+q]} \mathcal{L}_p^{-1} \left(\frac{1}{[\frac{1-c_1c_2+q}{(1-\rho^2)}]} + p \right] [(1-c_1)+p] \right) \right) \\ &= \frac{\rho^2 (1-c_1)(1-c_2)}{(1-\rho^2)} \times \\ & \mathcal{L}_q^{-1} \left(\frac{1}{[(1-c_2)+q]} \frac{1}{[\frac{1}{(1-\rho^2)}+q]} \left(\frac{\exp\left\{ -\left(\frac{1-c_1c_2+q}{(1-\rho^2)}\right) \lambda(\mathbf{s}_i) \right\} - \exp\left\{ -(1-c_1)\lambda(\mathbf{s}_i) \right\} \right) \\ &= \rho^2 (1-c_1)(1-c_2) \times \\ &= \rho^2 (1-c_1)(1-c_2) \times \end{split}$$

$$\begin{split} \mathcal{L}_{q}^{-1} \left(\frac{1}{[(1-s_{2})+q]} \left(\frac{\exp\left\{ - \left(\frac{\frac{1-c_{1}c_{2}+q}{(1-\rho^{2})}}{[\frac{1}{(1-\rho^{2})}+q]} \right) \lambda(\mathbf{s}_{i}) \right\} - \exp\left\{ - (1-c_{1}) \lambda(\mathbf{s}_{i}) \right\} \right) \\ = \frac{\rho^{2}(1-c_{1})(1-c_{2})}{[(1-c_{1})(1-\rho^{2})-1]} \\ \times \mathcal{L}_{q}^{-1} \left(\frac{1}{[(1-c_{2})+q]} \frac{1}{\frac{c_{1}c_{2}-c_{1}}{[(1-c_{1})(1-\rho^{2})-1]}+q} \exp\left\{ - \left(\frac{\frac{1-c_{1}c_{2}+q}{(1-\rho^{2})}}{[\frac{1}{(1-\rho^{2})}+q]} \right) \lambda(\mathbf{s}_{i}) \right\} \right) \\ - \frac{\rho^{2}(1-c_{1})(1-c_{2})}{[(1-c_{1})(1-\rho^{2})-1]} \\ \times \mathcal{L}_{q}^{-1} \left(\frac{1}{[(1-c_{2})+q]} \frac{1}{\frac{c_{1}c_{2}-c_{1}}{[(1-c_{1})(1-\rho^{2})-1]}+q} \exp\left\{ - (1-c_{1}) \lambda(\mathbf{s}_{i}) \right\} \right) \\ = \exp\left\{ - \frac{\lambda(\mathbf{s}_{i})}{(1-\rho^{2})} - \frac{\lambda(\mathbf{s}_{j})}{(1-\rho^{2})} \right\} \\ \times \mathcal{L}_{q}^{-1} \left(\frac{1}{1-c_{2}} - \frac{1}{(1-\rho^{2})}+q} \exp\left\{ - \left(\frac{\frac{1-c_{1}c_{2}}{(1-\rho^{2})ab} - \frac{1}{(1-\rho^{2})^{2}}}{q} \right) \lambda(\mathbf{s}_{i}) \right\} \right) \\ - \exp\left\{ - \frac{\lambda(\mathbf{s}_{i})}{(1-\rho^{2})} - \frac{\lambda(\mathbf{s}_{j})}{(1-\rho^{2})} \right\} \\ \times \mathcal{L}_{q}^{-1} \left(\frac{1}{\frac{c_{1}c_{2}-c_{1}}{(1-\rho^{2})} - \frac{1}{(1-\rho^{2})}+q} \exp\left\{ - \left(\frac{\frac{1-c_{1}c_{2}}{(1-\rho^{2})ab} - \frac{1}{(1-\rho^{2})^{2}}}{q} \right) \lambda(\mathbf{s}_{i}) \right\} \right) \\ - \exp\left\{ - \frac{\lambda(\mathbf{s}_{i})}{(1-\rho^{2})} - \frac{\lambda(\mathbf{s}_{j})}{(1-\rho^{2})} \right\} \\ \times \mathcal{L}_{q}^{-1} \left(\frac{1}{\frac{c_{1}c_{2}-c_{1}}{(1-\rho^{2})-1} - \frac{1}{(1-\rho^{2})}+q} \exp\left\{ - \left(\frac{\frac{1-c_{1}c_{2}}{(1-\rho^{2})} - \frac{1}{(1-\rho^{2})^{2}}}{q} \right) \lambda(\mathbf{s}_{i}) \right\} \right) \\ \end{array}$$

$$-\frac{\rho(1-c_{1})(1-c_{2})}{[(1-c_{1})(1-\rho^{2})-1]} \times \exp\left\{-(1-c_{1})\lambda(\mathbf{s}_{i})\right\} \times \mathcal{L}_{q}^{-1}\left(\frac{1}{[1-c_{2}+q]}\frac{1}{\frac{c_{1}c_{2}-c_{1}}{[(1-c_{1})(1-\rho^{2})-1]}+q}\right)$$

$$= \exp\left\{-\frac{\lambda(\mathbf{s}_{i})}{(1-\rho^{2})} - \frac{\lambda(\mathbf{s}_{j})}{(1-\rho^{2})}\right\}$$

$$\times \Phi_{3}\left(1,1;\left(-(1-c_{2})+\frac{1}{(1-\rho^{2})}\right)\lambda(\mathbf{s}_{j}), -\left(\frac{1-c_{1}c_{2}}{(1-\rho^{2})}-\frac{1}{(1-\rho^{2})^{2}}\right)\lambda(\mathbf{s}_{i})\lambda(\mathbf{s}_{j})\right)$$

$$-\exp\left\{-\frac{\lambda(\mathbf{s}_{i})}{(1-\rho^{2})} - \frac{\lambda(\mathbf{s}_{j})}{(1-\rho^{2})}\right\}$$

$$\times \Phi_{3}\left(1,1;\left(-\frac{c_{1}c_{2}-c_{1}}{[(1-c_{1})(1-\rho^{2})-1]}+\frac{1}{(1-\rho^{2})}\right)\lambda(\mathbf{s}_{j}), -\left(\frac{1-c_{1}c_{2}}{(1-\rho^{2})}-\frac{1}{(1-\rho^{2})^{2}}\right)\lambda(\mathbf{s}_{i})\lambda(\mathbf{s}_{j})\right)$$

$$-\exp\left\{-(1-c_{1})\lambda(\mathbf{s}_{i})-(1-c_{2})\lambda(\mathbf{s}_{j})\right\} + \exp\left\{-(1-c_{1})\lambda(\mathbf{s}_{i})-\left(\frac{c_{1}c_{2}-c_{1}}{[(1-c_{1})(1-\rho^{2})-1]}\right)\lambda(\mathbf{s}_{j})\right\}.$$

Therefore, $\mathcal{P}(\lambda(\mathbf{s}_i), \lambda(\mathbf{s}_j); c_1, c_2)$, denoted by \mathcal{P}_{ij} , is as follows:

$$\begin{aligned} \mathcal{P}_{ij} &= \exp\left\{-\frac{\lambda(\mathbf{s}_i)}{(1-\rho^2)} - \frac{\lambda(\mathbf{s}_j)}{(1-\rho^2)}\right\} \\ &\times \Phi_3\left(1, 1; \left(-(1-c_2) + \frac{1}{(1-\rho^2)}\right) \lambda(\mathbf{s}_j), -\left(\frac{1-c_1c_2}{(1-\rho^2)} - \frac{1}{(1-\rho^2)^2}\right) \lambda(\mathbf{s}_i)\lambda(\mathbf{s}_j)\right) \\ &- \exp\left\{-\frac{\lambda(\mathbf{s}_i)}{(1-\rho^2)} - \frac{\lambda(\mathbf{s}_j)}{(1-\rho^2)}\right\} \\ &\times \Phi_3\left(1, 1; \left(-\frac{c_1c_2-c_1}{[(1-c_1)(1-\rho^2)-1]} + \frac{1}{(1-\rho^2)}\right) \lambda(\mathbf{s}_j), -\left(\frac{1-c_1c_2}{(1-\rho^2)} - \frac{1}{(1-\rho^2)^2}\right) \lambda(\mathbf{s}_i)\lambda(\mathbf{s}_j)\right) \\ &- \exp\left\{-(1-c_1)\lambda(\mathbf{s}_i) - (1-c_2)\lambda(\mathbf{s}_j)\right\} \\ &+ \exp\left\{-(1-c_1)\lambda(\mathbf{s}_i) - \left(\frac{c_1c_2-c_1}{[(1-c_1)(1-\rho^2)-1]}\right)\lambda(\mathbf{s}_j)\right\} \\ &+ \exp\left\{-(1-c_1)\lambda(\mathbf{s}_i) - (1-c_2)\lambda(\mathbf{s}_j)\right\} \\ &= \exp\left\{-\frac{\lambda(\mathbf{s}_i)}{(1-\rho^2)} - \frac{\lambda(\mathbf{s}_j)}{(1-\rho^2)}\right\} \\ &\times \Phi_3\left(1, 1; \left(-(1-c_2) + \frac{1}{(1-\rho^2)}\right)\lambda(\mathbf{s}_j), -\left(\frac{1-c_1c_2}{(1-\rho^2)} - \frac{1}{(1-\rho^2)^2}\right)\lambda(\mathbf{s}_i)\lambda(\mathbf{s}_j)\right) \\ &- \exp\left\{-\frac{\lambda(\mathbf{s}_i)}{(1-\rho^2)} - \frac{\lambda(\mathbf{s}_j)}{(1-\rho^2)}\right\} \end{aligned}$$

$$\times \Phi_{3} \left(1, 1; \left(-\frac{c_{1}c_{2}-c_{1}}{\left[(1-c_{1})(1-\rho^{2})-1 \right]} + \frac{1}{(1-\rho^{2})} \right) \lambda(\mathbf{s}_{j}), - \left(\frac{1-c_{1}c_{2}}{(1-\rho^{2})} - \frac{1}{(1-\rho^{2})^{2}} \right) \lambda(\mathbf{s}_{i}) \lambda(\mathbf{s}_{j}) \right)$$

$$+ \exp \left\{ - \left(1-c_{1} \right) \lambda(\mathbf{s}_{i}) - \left(\frac{c_{1}c_{2}-c_{1}}{\left[(1-c_{1})(1-\rho^{2})-1 \right]} \right) \lambda(\mathbf{s}_{j}) \right\}$$

Proof of Theorem 2.2.3.

Consider two bivariate independent non-negative random variables random variables $(A_1, A_2)^{\top}$ and $(B_1, B_2)^{\top}$ with cdf $F(\cdot, \cdot)$ and $G(\cdot, \cdot)$, and pdf $f(\cdot, \cdot)$ and $g(\cdot, \cdot)$ respectively. Following Hunter (1974), we define the cdf of $(Z_1, Z_2)^{\top} = (A_1, A_2)^{\top} + (B_1, B_2)^{\top}$ as their double convolution as :

$$[F * *G](a,b) = \int_0^a \int_0^b F(a-u,b-v)dG(u,v) = \Pr(Z_1 \le a, Z_2 \le b)$$
(A.1)

and we define $F_0(a,b) = 1$, $F_1(a,b) = F(a,b)$ and $F_{n+1}(a,b) = [F_1 * *F_n](a,b)$. Moreover, $F^1(a,b) = \lim_{b\to\infty} F(a,b) = F(a,\infty) = F(a) = F^1(a)$, $F^2(a,b) = \lim_{a\to\infty} F(a,b) = F(\infty,b) = F(b) = F^2(b)$, $F_n^1(a,b) = \lim_{b\to\infty} F_n(a,b) = F_n(a,\infty) = F_n(a) = F_n(a, \infty) = F_n(a) = F_n(a,b) = \lim_{a\to\infty} F_n(a,b) = F_n(\infty,b) = F_n(b) = F_n^2(b)$.

The general expression for the bivariate pdf of N is given by Theorem. 3.1 of Hunter (1974) as follow:

$$\Pr(N(\mathbf{s}_{i}) = n, N(\mathbf{s}_{j}) = m) = \begin{cases} [F_{0} - F^{1} - F^{2} + F] * *F_{n}(\lambda(\mathbf{s}_{i}), \lambda(\mathbf{s}_{j})) & if \quad n = m \\ [F_{r}^{1} - F_{r+1}^{1} - F_{r-1}^{1} * *F + F_{r}^{1} * *F] * *F_{m}(\lambda(\mathbf{s}_{i}), \lambda(\mathbf{s}_{j})) & if \quad n > m, \\ n = m + r \\ [F_{r}^{2} - F_{r+1}^{2} - F_{r-1}^{2} * *F + F_{r}^{2} * *F] * *F_{n}(\lambda(\mathbf{s}_{i}), \lambda(\mathbf{s}_{j})) & if \quad n < m, \\ m = n + r \\ (A.2) \end{cases}$$

where, following our notation, $F_n(a, b)$ is the *n*-fold bivariate convolution of $(Y(\mathbf{s}_i), Y(\mathbf{s}_j))^{\top}$, that is,

$$F_n(a,b) = \Pr\left(\sum_{k=1}^n Y_k(\mathbf{s}_i) \le a, \sum_{k=1}^n Y_k(\mathbf{s}_j) \le b\right) = \Pr(S_n(\mathbf{s}_i) \le a, S_n(\mathbf{s}_j) \le b) = F_{S_{n;ij}}(a,b)$$

Let $Y(\mathbf{s}_i)$ and $Y(\mathbf{s}_j)$ a pair observations from Y. Then the following results hold

$$F^{1}(\lambda(\mathbf{s}_{i})) = 1 - \exp\{-\lambda(\mathbf{s}_{i})\}$$

$$F^{2}(\lambda(\mathbf{s}_{j})) = 1 - \exp\{-\lambda(\mathbf{s}_{j})\}$$

$$F(\lambda(\mathbf{s}_{i}), \lambda(\mathbf{s}_{j})) = \int_{0}^{\lambda(\mathbf{s}_{i})} \int_{0}^{\lambda(\mathbf{s}_{j})} \frac{1}{1 - \rho^{2}} \exp\left\{-\frac{1}{1 - \rho^{2}}(u + v)\right\} I_{0}\left(\frac{2\rho}{1 - \rho^{2}}\sqrt{uv}\right) dv du$$

$$F_{n}^{1}(\lambda(\mathbf{s}_{i})) = \gamma^{*}(n, \lambda(\mathbf{s}_{i}))$$

$$F_{n}^{2}(\lambda(\mathbf{s}_{i})) = \gamma^{*}(n, \lambda(\mathbf{s}_{j}))$$

$$F_{n}(\lambda(\mathbf{s}_{i}), \lambda(\mathbf{s}_{j})) = \int_{0}^{\lambda(\mathbf{s}_{i})} \int_{0}^{\lambda(\mathbf{s}_{j})} \frac{1}{(1 - \rho^{2})\Gamma(n)} \left(\frac{t_{1}t_{2}}{\rho^{2}}\right)^{\frac{n-1}{2}} \exp\left\{-\frac{(t_{1} + t_{2})}{1 - \rho^{2}}\right\} I_{n-1}\left(\frac{2\rho}{1 - \rho^{2}}\sqrt{t_{1}t_{2}}\right) dt_{2}dt_{1}$$

$$(A.3)$$

Let $p_{n,m}$ denote $Pr(N(\mathbf{s}_i) = n, N(\mathbf{s}_j) = m)$ and, plug in A.1 and A.3 on A.2 the following results are obtained:

If n = m = 0

$$p_{0,0} = [F_0 - F^1 - F^2 + F](\lambda(\mathbf{s}_i), \lambda(\mathbf{s}_j))$$

$$= -1 + \exp\{-\lambda(\mathbf{s}_i)\} + \exp\{-\lambda(\mathbf{s}_j)\}$$

$$+ \int_{0}^{\lambda(\mathbf{s}_i)} \int_{0}^{\lambda(\mathbf{s}_j)} \frac{1}{1 - \rho^2} \exp\{-\frac{u + v}{1 - \rho^2}\} I_0\left(\frac{2\rho}{1 - \rho^2}\sqrt{uv}\right) dv du$$

$$= 1 + \exp\{-\lambda(\mathbf{s}_i)\} + \exp\{-\lambda(\mathbf{s}_j)\}$$

$$+ (1 - \rho^2) \sum_{k=0}^{\infty} \frac{\rho^{2k}}{(k!)^2} \gamma\left(k + 1, \frac{\lambda(\mathbf{s}_i)}{1 - \rho^2}\right) \gamma\left(k + 1, \frac{\lambda(\mathbf{s}_j)}{1 - \rho^2}\right)$$

$$= -1 + \exp\{-\lambda(\mathbf{s}_i)\} + \exp\{-\lambda(\mathbf{s}_j)\}$$

$$+ (1 - \rho^2) \sum_{k=0}^{\infty} \rho^{2k} \gamma^*\left(k + 1, \frac{\lambda(\mathbf{s}_i)}{1 - \rho^2}\right) \gamma^*\left(k + 1, \frac{\lambda(\mathbf{s}_j)}{1 - \rho^2}\right)$$

$$= -1 + \exp\{-\lambda(\mathbf{s}_{i})\} + \exp\{-\lambda(\mathbf{s}_{j})\} + (1-\rho^{2})\sum_{k=0}^{\infty}\rho^{2k}\gamma^{\star}\left(k+1, \frac{\lambda(\mathbf{s}_{i})}{1-\rho^{2}}, \frac{\lambda(\mathbf{s}_{j})}{1-\rho^{2}}\right)$$

If $n = m \ge 1$

$$\begin{split} p_{n,n} &= [\mathbf{F}_0 - \mathbf{F}^1 - \mathbf{F}^2 + \mathbf{F}] * * \mathbf{F}_n(\lambda(\mathbf{s}_i), \lambda(\mathbf{s}_j)) \\ &= \int_0^{\lambda(\mathbf{s}_i) \lambda(\mathbf{s}_j)} \left[-1 + \exp\{-(\lambda(\mathbf{s}_i) - t_1)\} + \exp\{-(\lambda(\mathbf{s}_j) - t_2)\} \\ &+ \int_0^{\lambda(\mathbf{s}_i) - t_1 \lambda(\mathbf{s}_j) - t_2} \int_0^{1} \frac{1}{1 - \rho^2} \exp\left\{-\frac{u + v}{1 - \rho^2}\right\} I_0\left(\frac{2\rho}{1 - \rho^2} \sqrt{uv}\right) dv du \right] \times \\ &\left[\frac{1}{(1 - \rho^2)(n - 1)!} \left(\frac{t_1 t_2}{\rho^2}\right)^{\frac{n - 1}{2}} \exp\left\{-\frac{t_1 + t_2}{1 - \rho^2}\right\} I_{n-1}\left(\frac{2\rho}{1 - \rho^2} \sqrt{t_1 t_2}\right) \right] dt_2 dt_1 \\ &= -\frac{(1 - \rho^2)^n}{\Gamma(n)} \sum_{k=0}^{\infty} \frac{\rho^{2k} \Gamma(n + k)}{\Gamma(k + 1)} \frac{\gamma\left(n + k, \frac{\lambda(\mathbf{s}_i)}{1 - \rho^2}\right)}{\Gamma(n + k)} \frac{\gamma\left(n + k, \frac{\lambda(\mathbf{s}_j)}{1 - \rho^2}\right)}{\Gamma(n + k)} \\ &+ \frac{(1 - \rho^2)^n \exp\{-\lambda(\mathbf{s}_i)\}}{\Gamma(n)\rho^{2n}} \sum_{k=0}^{\infty} \frac{\Gamma(n + k)}{\Gamma(k + 1)} \frac{\gamma\left(n + k, \frac{\lambda(\mathbf{s}_i)}{1 - \rho^2}\right)}{\Gamma(n + k)} \frac{\gamma\left(n + k, \frac{\lambda(\mathbf{s}_j)}{1 - \rho^2}\right)}{\Gamma(k + 1)} \frac{\gamma\left(n + k, \frac{\lambda(\mathbf{s}_j)}{1 - \rho^2}\right)}{\Gamma(k + 1)} \\ &+ \frac{(1 - \rho^2)^n \exp\{-\lambda(\mathbf{s}_j)\}}{0} \sum_{k=0}^{\infty} \frac{\Gamma(n + k)}{\Gamma(k + 1)} \frac{\gamma\left(n + k, \frac{\lambda(\mathbf{s}_i)}{1 - \rho^2}\right)}{\Gamma(k + 1)} \frac{\gamma\left(n + k, \frac{\lambda(\mathbf{s}_j)}{1 - \rho^2}\right)}{\Gamma(k + 1)} \frac{\gamma\left(n + k, \frac{\lambda(\mathbf{s}_j)}{1 - \rho^2}\right)}{\Gamma(k + 1)} \\ &+ \int_0^{\lambda(\mathbf{s}_i) \lambda(\mathbf{s}_j)} \left[(1 - \rho^2) \sum_{k=0}^{\infty} \frac{\rho^{2k}}{(k!)^2} \gamma\left(k + 1, \frac{\lambda(\mathbf{s}_i) - t_1}{1 - \rho^2}\right) \gamma\left(k + 1, \frac{\lambda(\mathbf{s}_j) - t_2}{1 - \rho^2}\right) \right] \times \\ &\left[\frac{1}{(1 - \rho^2)(n - 1)!} \left(\frac{t_1 t_2}{\rho^2}\right)^{\frac{n - 1}{2}} \exp\left\{-\frac{t_1 + t_2}{1 - \rho^2}\right\} I_{n-1} \left(\frac{2\rho}{1 - \rho^2} \sqrt{t_1 t_2}\right) \right] dt_2 dt_1 \\ &= -\frac{(1 - \rho^2)^n}{\Gamma(n)} \sum_{k=0}^{\infty} \frac{\rho^{2k} \Gamma(n + k)}{\Gamma(k + 1)} \gamma^* \left(n + k, \frac{\lambda(\mathbf{s}_i)}{1 - \rho^2}\right) \gamma^* \left(n + k, \frac{\lambda(\mathbf{s}_j)}{1 - \rho^2}\right) \end{split}$$

$$\begin{split} &+ \frac{(1-\rho^2)^n \exp\{-\lambda(\mathbf{s}_i)\}}{\Gamma(n)\rho^{2n}} \sum_{k=0}^{\infty} \frac{\Gamma(n+k)}{\Gamma(k+1)} \gamma^* \left(n+k, \frac{\rho^2\lambda(\mathbf{s}_i)}{1-\rho^2}\right) \gamma^* \left(n+k, \frac{\lambda(\mathbf{s}_j)}{1-\rho^2}\right) \\ &+ \frac{(1-\rho^2)^n \exp\{-\lambda(\mathbf{s}_j)\}}{\Gamma(n)\rho^{2n}} \sum_{k=0}^{\infty} \frac{\Gamma(n+k)}{\Gamma(k+1)} \gamma^* \left(n+k, \frac{\lambda(\mathbf{s}_i)}{1-\rho^2}\right) \gamma^* \left(n+k, \frac{\rho^2\lambda(\mathbf{s}_j)}{1-\rho^2}\right) \\ &+ \int_{0}^{\lambda(\mathbf{s}_i)\lambda(\mathbf{s}_j)} \left[(1-\rho^2) \sum_{k=0}^{\infty} \frac{\rho^{2k}}{(k!)^2} \gamma \left(k+1, \frac{\lambda(\mathbf{s}_i)-t_1}{1-\rho^2}\right) \gamma \left(k+1, \frac{\lambda(\mathbf{s}_j)-t_2}{1-\rho^2}\right) \right] \times \\ &\left[\frac{1}{(1-\rho^2)(n-1)!} \left(\frac{t_1t_2}{\rho^2} \right)^{\frac{n-1}{2}} \exp\left\{ -\frac{t_1+t_2}{1-\rho^2} \right\} I_{n-1} \left(\frac{2\rho}{1-\rho^2} \sqrt{t_1t_2} \right) \right] dt_2 dt_1 \\ &= -\frac{(1-\rho^2)^n}{\Gamma(n)} \sum_{k=0}^{\infty} \frac{\rho^{2k}\Gamma(n+k)}{\Gamma(k+1)} \gamma^* \left(n+k, \frac{\lambda(\mathbf{s}_i)}{1-\rho^2}\right) \gamma^* \left(n+k, \frac{\lambda(\mathbf{s}_j)}{1-\rho^2}\right) \\ &+ \frac{(1-\rho^2)^n \exp\{-\lambda(\mathbf{s}_i)\}}{\Gamma(n)\rho^{2n}} \sum_{k=0}^{\infty} \frac{\Gamma(n+k)}{\Gamma(k+1)} \gamma^* \left(n+k, \frac{\rho^2\lambda(\mathbf{s}_i)}{1-\rho^2}\right) \gamma^* \left(n+k, \frac{\rho^2\lambda(\mathbf{s}_j)}{1-\rho^2}\right) \\ &+ \frac{\sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \frac{\rho^{2k+2\ell}}{(1-\rho^2)^{2\ell+n-1}} \frac{1}{(k!)^2(n-1)!\ell!\Gamma(n+\ell)} \times \\ &\int_{0}^{\lambda(\mathbf{s}_j)} t_1^{\ell+n-1} \exp\left\{-\frac{t_2}{1-\rho^2}\right\} \gamma \left(k+1, \frac{\lambda(\mathbf{s}_i)-t_1}{1-\rho^2}\right) \times \\ &= -\frac{(1-\rho^2)^n}{\Gamma(n)} \sum_{k=0}^{\infty} \frac{\rho^{2k}\Gamma(n+k)}{\Gamma(k+1)} \gamma^* \left(n+k, \frac{\lambda(\mathbf{s}_i)}{1-\rho^2}\right) \times \\ &= -\frac{(1-\rho^2)^n \exp\{-\lambda(\mathbf{s}_i)\}}{\sum_{k=0}^{\infty} \frac{\Gamma(n+k)}{\Gamma(k+1)}} \gamma^* \left(n+k, \frac{\lambda(\mathbf{s}_i)}{1-\rho^2}\right) \\ &+ \frac{(1-\rho^2)^n \exp\{-\lambda(\mathbf{s}_i)\}}{\sum_{k=0}^{\infty} \frac{\Gamma(n+k)}{\Gamma(k+1)}} \gamma^* \left(n+k, \frac{\lambda(\mathbf{s}_i)}{1-\rho^2}\right) \\ &+ \frac{(1-\rho^2)^n \exp\{-\lambda(\mathbf{s}_i)\}}{\sum_{k=0}^{\infty} \frac{\Gamma(n+k)}{\Gamma(k+1)}} \gamma^* \left(n+k, \frac{\lambda(\mathbf{s}_i)}{1-\rho^2}\right) \gamma^* \left(n+k, \frac{\lambda(\mathbf{s}_j)}{1-\rho^2}\right) \\ &+ \frac{(1-\rho^2)^n \exp\{-\lambda(\mathbf{s}_i)\}}{\Gamma(n)\rho^{2n}} \sum_{k=0}^{\infty} \frac{\Gamma(n+k)}{\Gamma(k+1)} \gamma^* \left(n+k, \frac{\lambda(\mathbf{s}_i)}{1-\rho^2}\right) \gamma^* \left(n+k, \frac{\lambda(\mathbf{s}_j)}{1-\rho^2}\right) \\ &+ \frac{(1-\rho^2)^n \exp\{-\lambda(\mathbf{s}_i)\}}{\Gamma(n)\rho^{2n}} \sum_{k=0}^{\infty} \frac{\Gamma(n+k)}{\Gamma(k+1)} \gamma^* \left(n+k, \frac{\lambda(\mathbf{s}_i)}{1-\rho^2}\right) \gamma^* \left(n+k, \frac{\lambda(\mathbf{s}_j)}{1-\rho^2}\right) \\ &+ \frac{(1-\rho^2)^n \exp\{-\lambda(\mathbf{s}_j)\}}{\Gamma(n)\rho^{2n}} \sum_{k=0}^{\infty} \frac{\Gamma(n+k)}{\Gamma(k+1)} \gamma^* \left(n+k, \frac{\lambda(\mathbf{s}_i)}{1-\rho^2}\right) \gamma^* \left(n+k, \frac{\lambda(\mathbf{s}_j)}{1-\rho^2}\right) \\ &+ \frac{(1-\rho^2)^n \exp\{-\lambda(\mathbf{s}_j)\}}{\Gamma(n)\rho^{2n}}} \sum_{k=0}^{\infty} \frac{\Gamma(n+k)}{\Gamma(k+1)} \gamma^* \left(n+k, \frac{\lambda(\mathbf{s}_i)}{1-\rho^2}\right) \gamma^* \left(n+k, \frac{\lambda(\mathbf{s}_j)}{1-\rho^2}\right) \\ &+ \frac{(1-\rho^2)^n \exp\{-\lambda(\mathbf{s}_j)\}}{\Gamma(n)\rho^{2n}$$

$$\begin{split} &+ \frac{(1-\rho^2)^{n+1}}{\Gamma(n)} \times \\ &\sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \frac{\Gamma(n+\ell)\rho^{2k+2\ell}}{\Gamma(\ell+1)} \gamma^* \left(n+\ell+k+1, \frac{\lambda(\mathbf{s}_i)}{1-\rho^2}\right) \gamma^* \left(n+\ell+k+1, \frac{\lambda(\mathbf{s}_j)}{1-\rho^2}\right) \\ &= -\frac{(1-\rho^2)^n}{\Gamma(n)} \sum_{k=0}^{\infty} \frac{\rho^{2k}\Gamma(n+k)}{\Gamma(k+1)} \gamma^* \left(n+k, \frac{\lambda(\mathbf{s}_i)}{1-\rho^2}, \frac{\lambda(\mathbf{s}_j)}{1-\rho^2}\right) \\ &+ \frac{(1-\rho^2)^n \exp\{-\lambda(\mathbf{s}_j)\}}{\Gamma(n)\rho^{2n}} \sum_{k=0}^{\infty} \frac{\Gamma(n+k)}{\Gamma(k+1)} \gamma^* \left(n+k, \frac{\rho^2\lambda(\mathbf{s}_i)}{1-\rho^2}, \frac{\lambda(\mathbf{s}_j)}{1-\rho^2}\right) \\ &+ \frac{(1-\rho^2)^n \exp\{-\lambda(\mathbf{s}_j)\}}{\Gamma(n)\rho^{2n}} \sum_{k=0}^{\infty} \frac{\Gamma(n+\ell)\rho^{2k+2\ell}}{\Gamma(\ell+1)} \gamma^* \left(n+k, \frac{\lambda(\mathbf{s}_i)}{1-\rho^2}, \frac{\rho^2\lambda(\mathbf{s}_j)}{1-\rho^2}\right) \\ &+ \frac{(1-\rho^2)^{n+1}}{\Gamma(n)} \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \frac{\Gamma(n+\ell)\rho^{2k+2\ell}}{\Gamma(\ell+1)} \gamma^* \left(n+k, \frac{\lambda(\mathbf{s}_i)}{1-\rho^2}, \frac{\lambda(\mathbf{s}_j)}{1-\rho^2}\right) \\ &= -\frac{(1-\rho^2)^n}{\Gamma(n)} \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \frac{\Gamma(n+\ell)\rho^{2k+2\ell}}{\Gamma(\ell+1)} \gamma^* \left(n+k, \frac{\lambda(\mathbf{s}_i)}{1-\rho^2}, \frac{\lambda(\mathbf{s}_j)}{1-\rho^2}\right) \\ &+ \frac{(1-\rho^2)^{n+1}}{\Gamma(n)} \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \frac{\Gamma(n+\ell)\rho^{2k+2\ell}}{\Gamma(\ell+1)} \gamma^* \left(n+\ell+k+1, \frac{\lambda(\mathbf{s}_i)}{1-\rho^2}, \frac{\lambda(\mathbf{s}_j)}{1-\rho^2}\right) \\ &= -(1-\rho^2)^n \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \frac{\Gamma(n+\ell)\rho^{2k+2\ell}}{\Gamma(\ell+1)} \gamma^* \left(n+k, \frac{\lambda(\mathbf{s}_j)}{1-\rho^2}\right) \\ &+ \left(\frac{1-\rho^2}{\rho^2}\right)^n \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \frac{\Gamma(n+\ell)\rho^{2k+2\ell}}{k!} \gamma^* \left(n+k, \frac{\lambda(\mathbf{s}_i)}{1-\rho^2}\right) \\ &+ \left(1-\rho^2)^{n+1} \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \frac{\Gamma(n+\ell)\rho^{2k+2\ell}}{k!} \gamma^* \left(n+k, \frac{\lambda(\mathbf{s}_j)}{1-\rho^2}\right) \\ &+ \left(1-\rho^2)^{n+1} \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \frac{\Gamma(n+\ell)\rho^{2k+2\ell}}{k!} \gamma^* \left(n+k, \frac{\lambda(\mathbf{s}_j)}{1-\rho^2}\right) \\ &+ \left(1-\rho^2)^{n+1} \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \frac{\rho^{2k}(n)_k}{k!} \gamma^* \left(n+k, \frac{\lambda(\mathbf{s}_j)}{1-\rho^2}\right) \\ &+ \left(1-\rho^2)^{n+1} \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \frac{\rho^{2k+2\ell}(n)_\ell}{k!} \gamma^* \left(n+\ell+k+1, \frac{\lambda(\mathbf{s}_j)}{1-\rho^2}\right) \\ &+ \left(1-\rho^2)^{n+1} \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \frac{\rho^{2k+2\ell}(n)_\ell}{\ell!} \gamma^* \left(n+\ell+k+1, \frac{\lambda(\mathbf{s}_j)}{1-\rho^2}\right) \\ &+ \left(1-\rho^2\right)^{n+1} \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \frac{\rho^{2k+2\ell}(n)_\ell}{\ell!} \gamma^* \left(n+\ell+k+1, \frac{\lambda(\mathbf{s}_j)}{1-\rho^2}\right) \\ &+ \left(1-\rho^2\right)^{n+1} \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \frac{\rho^{2k+2\ell}(n)_\ell}{\ell!} \gamma^* \left(n+\ell+k+1, \frac{\lambda(\mathbf{s}_j)}{1-\rho^2}\right) \\ \\ &+ \left(1-\rho^2\right)^{n+1} \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \frac{\rho^{2k+2\ell}(n)_\ell}{\ell!} \gamma^* \left(n+\ell+k+1, \frac{\lambda(\mathbf{s}_j)}{1-\rho^2}\right) \\ \\ &+ \left(1-\rho^2\right)^{n+1} \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \frac{\rho^{2k+2\ell}(n)_\ell}{\ell!} \gamma^* \left(n+\ell+$$

If $n \ge 1$ and m = 0

$$p_{n,0} = [\mathbf{F}_n^1 - \mathbf{F}_{n+1}^1 - \mathbf{F}_{n-1}^1 * *\mathbf{F} + \mathbf{F}_n^1 * *\mathbf{F}](\lambda(\mathbf{s}_i), \lambda(\mathbf{s}_j))$$
$$= \frac{1}{n!} \lambda(\mathbf{s}_i)^n \exp\{-\lambda(\mathbf{s}_i)\}$$

$$\begin{split} &- \int_{0}^{\lambda(\mathbf{s}_{i})} \frac{\gamma(n-1,\lambda(\mathbf{s}_{i})-u)}{(n-2)!} \frac{1}{1-\rho^{2}} \exp\left\{-\frac{u+v}{1-\rho^{2}}\right\} I_{0}\left(\frac{2\rho}{1-\rho^{2}}\sqrt{uv}\right) dvdu \\ &+ \int_{0}^{\lambda(\mathbf{s}_{i})} \int_{0}^{\lambda(\mathbf{s}_{i})} \frac{\gamma(n,\lambda(\mathbf{s}_{i})-u)}{(n-1)!} \frac{1}{1-\rho^{2}} \exp\left\{-\frac{u+v}{1-\rho^{2}}\right\} I_{0}\left(\frac{2\rho}{1-\rho^{2}}\sqrt{uv}\right) dvdu \\ &= \frac{1}{n!} \lambda(\mathbf{s}_{i})^{n} \exp\{-\lambda(\mathbf{s}_{i})\} + \frac{1}{1-\rho^{2}} \times \\ &\int_{0}^{\lambda(\mathbf{s}_{i})} \int_{0}^{\lambda(\mathbf{s}_{i})} \frac{\gamma(n-\lambda(\mathbf{s}_{i})-u)}{(n-1)!} - \frac{\gamma(n-1,\lambda(\mathbf{s}_{i})-u)}{(n-2)!}\right) \exp\left\{-\frac{u+v}{1-\rho^{2}}\right\} I_{0}\left(\frac{2\rho\sqrt{uv}}{1-\rho^{2}}\right) dvdu \\ &= \frac{1}{n!} \lambda(\mathbf{s}_{i})^{n} \exp\{-\lambda(\mathbf{s}_{i})\} - \frac{1}{(n-1)!1-\rho^{2}} \times \\ &\int_{0}^{\lambda(\mathbf{s}_{i})} \int_{0}^{\lambda(\mathbf{s}_{i})} \left(\lambda(\mathbf{s}_{i})-u\right)^{n-1} \exp\{-\lambda(\mathbf{s}_{i})-u\} \exp\left\{-\frac{u+v}{1-\rho^{2}}\right\} I_{0}\left(\frac{2\rho}{1-\rho^{2}}\sqrt{uv}\right) dvdu \\ &= \frac{1}{n!} \lambda(\mathbf{s}_{i})^{n} \exp\{-\lambda(\mathbf{s}_{i})\} - \frac{\exp\{-\lambda(\mathbf{s}_{i})-u\}}{(n-1)!1-\rho^{2}} \times \\ &\sum_{0}^{\lambda(\mathbf{s}_{i})} \left(\lambda(\mathbf{s}_{i})-u\right)^{n-1} \exp\{-\lambda(\mathbf{s}_{i})-u\} \exp\left\{-\frac{\mu+v}{1-\rho^{2}}\right\} I_{0}\left(\frac{2\rho}{1-\rho^{2}}\sqrt{uv}\right) dvdu \\ &= \frac{1}{n!} \lambda(\mathbf{s}_{i})^{n} \exp\{-\lambda(\mathbf{s}_{i})\} - \frac{\exp\{-\lambda(\mathbf{s}_{i})\}}{(n-1)!1-\rho^{2}} \times \\ &\sum_{0}^{\infty} \left(\frac{\rho}{1-\rho^{2}}\right)^{2\ell} \frac{1}{\ell!^{2}} \int_{0}^{\lambda(\mathbf{s}_{i})} \left(\lambda(\mathbf{s}_{i})-u\right)^{n-1} u^{\ell} \exp\left\{-\frac{\rho^{2}u}{1-\rho^{2}}\right\} du \int_{0}^{\lambda(\mathbf{s}_{i})} v^{\ell} \exp\left\{-\frac{v}{1-\rho^{2}}\right\} dv \\ &= \frac{1}{n!} \lambda(\mathbf{s}_{i})^{n} \exp\{-\lambda(\mathbf{s}_{i})\} - \frac{\exp\{-\lambda(\mathbf{s}_{i})\}}{(n-1)!} \exp\left\{-\frac{\rho^{2}\lambda(\mathbf{s}_{i})}{1-\rho^{2}}\right\} \\ &\quad I^{1}_{i} \left(n; n+\ell+1; \frac{\rho^{2}\lambda(\mathbf{s}_{i})}{1-\rho^{2}}\right) \gamma\left(\ell+1; \frac{\lambda(\mathbf{s}_{i})}{1-\rho^{2}}\right) \\ &= \frac{1}{n!} \lambda(\mathbf{s}_{i})^{n} \exp\{-\lambda(\mathbf{s}_{i})\} - \lambda(\mathbf{s}_{i})^{n} \exp\left\{-\frac{\lambda(\mathbf{s}_{i})}{1-\rho^{2}}\right\} \times \\ &\sum_{\ell=0}^{\infty} \left(\frac{\rho^{2}\lambda(\mathbf{s}_{i})}{1-\rho^{2}}\right)^{\ell} \frac{1}{\Gamma} \left(n; n+\ell+1; \frac{\rho^{2}\lambda(\mathbf{s}_{i})}{1-\rho^{2}}\right) \gamma^{*} \left(\ell+1; \frac{\lambda(\mathbf{s}_{i})}{1-\rho^{2}}\right) \\ &= \frac{1}{n!} \lambda(\mathbf{s}_{i})^{n} \exp\{-\lambda(\mathbf{s}_{i})\} \end{aligned}$$

$$-\lambda(\mathbf{s}_i)^n \exp\left\{-\frac{\lambda(\mathbf{s}_i)}{1-\rho^2}\right\} \sum_{\ell=0}^{\infty} \left(\frac{\rho^2 \lambda(\mathbf{s}_i)}{1-\rho^2}\right)^{\ell} \mathcal{S}\left(\begin{smallmatrix}n \ ; n+\ell+1\\ \ell+1\end{smallmatrix}, \frac{\rho^2 \lambda(\mathbf{s}_i)}{1-\rho^2}, \frac{\lambda(\mathbf{s}_j)}{1-\rho^2}\right)$$

If $n > m, n \ge 2$ and $m \ge 1$

$$\begin{split} p_{n,m} &= [\mathbf{F}_{r}^{1} - \mathbf{F}_{r+1}^{1} - \mathbf{F}_{r-1}^{1} * *\mathbf{F} + \mathbf{F}_{r}^{1} * *\mathbf{F}] * *\mathbf{F}_{m}(\lambda(\mathbf{s}_{i}), \lambda(\mathbf{s}_{j})) \\ &= \int_{0}^{\lambda(\mathbf{s}_{i}) \lambda(\mathbf{s}_{j})} \int_{0}^{\lambda(\mathbf{s}_{i}) - t_{1}} \left[\frac{1}{r!} (\lambda(\mathbf{s}_{i}) - t_{1})^{r} \exp\{-\lambda(\mathbf{s}_{i}) - t_{1}\} \\ &- \int_{0}^{\lambda(\mathbf{s}_{i}) - t_{1} \lambda(\mathbf{s}_{j}) - t_{2}} \frac{\gamma(r - 1, \lambda(\mathbf{s}_{i}) - t_{1} - u)}{(r - 2)!(1 - \rho^{2})} \exp\left\{-\frac{u + v}{1 - \rho^{2}}\right\} I_{0}\left(\frac{2\rho}{1 - \rho^{2}}\sqrt{uv}\right) dvdu \\ &+ \int_{0}^{\lambda(\mathbf{s}_{i}) - t_{1} \lambda(\mathbf{s}_{j}) - t_{2}} \frac{\gamma(r, \lambda(\mathbf{s}_{i}) - t_{1} - u)}{(r - 1)!(1 - \rho^{2})} \exp\left\{-\frac{u + v}{1 - \rho^{2}}\right\} I_{0}\left(\frac{2\rho}{1 - \rho^{2}}\sqrt{uv}\right) dvdu \\ &\left[\frac{1}{(1 - \rho^{2})(m - 1)!}\left(\frac{t_{1}t_{2}}{\rho^{2}}\right)^{\frac{m-1}{2}} \exp\left\{-\frac{t_{1} + t_{2}}{1 - \rho^{2}}\right\} I_{m-1}\left(\frac{2\rho}{1 - \rho^{2}}\sqrt{t_{1}t_{2}}\right)\right] dt_{2}dt_{1} \\ &= \int_{0}^{\lambda(\mathbf{s}_{i}) \lambda(\mathbf{s}_{i})} \int_{0}^{1} \left[\frac{1}{r!} (\lambda(\mathbf{s}_{i}) - t_{1})^{r} \exp\{-\lambda(\mathbf{s}_{i}) - t_{1}\} \\ &+ \int_{0}^{\lambda(\mathbf{s}_{i}) - t_{1} \lambda(\mathbf{s}_{j}) - t_{2}} \left(\frac{\gamma(r, \lambda(\mathbf{s}_{i}) - t_{1} - u)}{(r - 1)!} - \frac{\gamma(r - 1, \lambda(\mathbf{s}_{i}) - t_{1} - u)}{(r - 2)!}\right) \\ &= \frac{1}{1 - \rho^{2}} \exp\left\{-\frac{u + v}{1 - \rho^{2}}\right\} I_{0}\left(\frac{2\rho}{1 - \rho^{2}}\sqrt{uv}\right) dvdu\right] \times \\ &\left[\frac{1}{(1 - \rho^{2})(m - 1)!}\left(\frac{t_{1}t_{2}}{\rho^{2}}\right)^{\frac{m-1}{2}} \exp\left\{-\frac{t_{1} + t_{2}}{1 - \rho^{2}}\right\} I_{m-1}\left(\frac{2\rho}{1 - \rho^{2}}\sqrt{t_{1}t_{2}}\right)\right] dt_{2}dt_{1} \\ &= \int_{0}^{\lambda(\mathbf{s}_{i}) \lambda(\mathbf{s}_{i})} \left[\frac{1}{r!}(\lambda(\mathbf{s}_{i}) - t_{1})^{r} \exp\{-\lambda(\mathbf{s}_{i}) - t_{1}\} \end{aligned}$$

$$\begin{split} &- \int_{0}^{\lambda(\mathbf{s}_{i})-t_{1}} \int_{0}^{\lambda(\mathbf{s}_{i})-t_{1}} \frac{(\lambda(\mathbf{s}_{i})-t_{1}-u))^{r-1}}{(r-1)!} \exp\{-\lambda(\mathbf{s}_{i})-t_{1}-u\} \\ &\quad \frac{1}{1-\rho^{2}} \exp\{-\frac{u+v}{1-\rho^{2}}\} I_{0}\left(\frac{2\rho}{1-\rho^{2}}\sqrt{uv}\right) dv du\right] \times \\ &\left[\frac{1}{(1-\rho^{2})(m-1)!} \left(\frac{t_{1}t_{2}}{\rho^{2}}\right)^{\frac{m-1}{2}} \exp\{-\frac{t_{1}+t_{2}}{1-\rho^{2}}\} I_{m-1}\left(\frac{2\rho}{1-\rho^{2}}\sqrt{t_{1}t_{2}}\right)\right] dt_{2} dt_{1} \\ &= \int_{0}^{\lambda(\mathbf{s}_{i})} \int_{0}^{\lambda(\mathbf{s}_{i})} \frac{1}{r!} (\lambda(\mathbf{s}_{i})-t_{1})^{r} \exp\{-\lambda(\mathbf{s}_{i})-t_{1}\} \\ &\left[\frac{1}{(1-\rho^{2})(m-1)!} \left(\frac{t_{1}t_{2}}{\rho^{2}}\right)^{\frac{m-1}{2}} \exp\{-\frac{t_{1}+t_{2}}{1-\rho^{2}}\} I_{m-1}\left(\frac{2\rho}{1-\rho^{2}}\sqrt{t_{1}t_{2}}\right)\right] dt_{2} dt_{1} \\ &- \int_{0}^{\lambda(\mathbf{s}_{i})} \int_{0}^{\lambda(\mathbf{s}_{i})-t_{1}} \int_{0}^{\lambda(\mathbf{s}_{i})-t_{1}} \frac{(\lambda(\mathbf{s}_{i})-t_{1}-u)}{(r-1)!} \exp\{-\lambda(\mathbf{s}_{i})-t_{1}-u\} \\ &\quad \frac{1}{1-\rho^{2}} \exp\{-\frac{u+v}{1-\rho^{2}}\} I_{0}\left(\frac{2\rho}{1-\rho^{2}}\sqrt{uv}\right) dv du\right] \times \\ &\left[\frac{1}{(1-\rho^{2})(m-1)!} \left(\frac{t_{1}t_{2}}{\rho^{2}}\right)^{\frac{m-1}{2}} \exp\{-\frac{t_{1}+t_{2}}{1-\rho^{2}}\} I_{m-1}\left(\frac{2\rho}{1-\rho^{2}}\sqrt{t_{1}t_{2}}\right)\right] dt_{2} dt_{1} \\ &= \sum_{\ell=0}^{\infty} \frac{\exp\{-\lambda(\mathbf{s}_{i})\}\rho^{2\ell}}{\ell!r!(m-1)!\Gamma(r+m)(1-\rho^{2})^{2\ell+m}} \times \\ &\int_{0}^{\lambda(\mathbf{s}_{i})} t_{1}^{i+m-1}(\lambda(\mathbf{s}_{i})-t_{1})^{r} \exp\{-\frac{\rho^{2}t_{1}}{1-\rho^{2}}dt_{1}} \int_{0}^{\lambda(\mathbf{s}_{i})} t_{2}^{i+m-1} \exp\{-\frac{t_{2}}{1-\rho^{2}}\} dt_{2} \\ &- \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \frac{\exp\{-\frac{\lambda(\mathbf{s}_{i})}{k!\ell!(r+k)!(m-1)!\Gamma(m+\ell)(1-\rho^{2})^{k+m+2\ell}} \times \\ &\int_{0}^{\lambda(\mathbf{s}_{i})} (\lambda(\mathbf{s}_{i})-t_{1})^{r+k} t_{1}^{m+\ell-1} {}_{1}F_{1}\left(r;r+k+1;\frac{\rho^{2}(\lambda(\mathbf{s}_{i})-t_{1})}{1-\rho^{2}}\right) dt_{1} \end{split}$$

$$\begin{split} &\int_{0}^{\lambda(\mathbf{s}_{j})} t_{2}^{m+\ell-1} \exp\left\{-\frac{t_{2}}{1-\rho^{2}}\right\} \gamma\left(k+1, \frac{\lambda(\mathbf{s}_{j})-t_{2}}{1-\rho^{2}}\right) dt_{2} \\ &= \lambda(\mathbf{s}_{i})^{m+r} \exp\left\{-\frac{\lambda(\mathbf{s}_{i})}{1-\rho^{2}}\right\} \times \\ &\left[\sum_{\ell=0}^{\infty} \frac{\Gamma(m+\ell)}{\ell!(m-1)!} \left(\frac{\rho^{2}\lambda(\mathbf{s}_{i})}{1-\rho^{2}}\right)^{\ell} \times \right. \\ &\left. {}_{1}\widetilde{F}_{1}\left(r+1;m+\ell+r+1;\frac{\rho^{2}\lambda(\mathbf{s}_{i})}{1-\rho^{2}}\right) \gamma^{*}\left(m+\ell,\frac{\lambda(\mathbf{s}_{j})}{1-\rho^{2}}\right) - \right. \\ &\left. {}_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \frac{\Gamma(m+\ell)}{\ell!(m-1)!} \left(\frac{\rho^{2}\lambda(\mathbf{s}_{i})}{1-\rho^{2}}\right)^{k+\ell} \times \right. \\ &\left. {}_{1}\widetilde{F}_{1}\left(r;1+k+\ell+r+m;\frac{\rho^{2}\lambda(\mathbf{s}_{i})}{1-\rho^{2}}\right) \gamma^{*}\left(m+\ell+k+1,\frac{\lambda(\mathbf{s}_{j})}{1-\rho^{2}}\right) \right] \\ &= \lambda(\mathbf{s}_{i})^{n} \exp\left\{-\frac{\lambda(\mathbf{s}_{i})}{1-\rho^{2}}\right\} \times \\ &\left[\sum_{\ell=0}^{\infty} \frac{(m)_{\ell}}{\ell!} \left(\frac{\rho^{2}\lambda(\mathbf{s}_{i})}{1-\rho^{2}}\right)^{\ell} \mathcal{S}\left(\frac{n-m+1}{m+\ell},\frac{\rho^{2}\lambda(\mathbf{s}_{i})}{1-\rho^{2}},\frac{\lambda(\mathbf{s}_{j})}{1-\rho^{2}}\right) - \right. \\ &\left. \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \frac{(m)_{\ell}}{\ell!} \left(\frac{\rho^{2}\lambda(\mathbf{s}_{i})}{1-\rho^{2}}\right)^{k+\ell} \mathcal{S}\left(\frac{n-m+1}{m+k+\ell+1},\frac{\rho^{2}\lambda(\mathbf{s}_{i})}{1-\rho^{2}},\frac{\lambda(\mathbf{s}_{j})}{1-\rho^{2}}\right) \right] \end{split}$$

The cases $m \ge 1, n = 0$ and $m > n, m \ge 2, n \ge 1$ are analogous to cases $n \ge 1, m = 0$ and $n > m, n \ge 2, m \ge 1$ respectively. Then the result holds.

A.2 Proofs of Chapter 3

Proof of Theorem 3.2.1. Note that the covariance of the ZIP random field N^* can be written as follows:

$$Cov(N^*(\mathbf{s}_i), N^*(\mathbf{s}_j)) = E[N^*(\mathbf{s}_i)N^*(\mathbf{s}_j)] - E[N^*(\mathbf{s}_i)]E[N^*(\mathbf{s}_j)]$$
$$= E[N(\mathbf{s}_i)B(\mathbf{s}_i)N(\mathbf{s}_j)B(\mathbf{s}_j)] - E[N(\mathbf{s}_i)B(\mathbf{s}_i)]E[N(\mathbf{s}_j)B(\mathbf{s}_j)]$$
$$= E[B(\mathbf{s}_i)B(\mathbf{s}_j)]E[N(\mathbf{s}_i)N(\mathbf{s}_j)] - E[B(\mathbf{s}_i)]E[N(\mathbf{s}_i)]E[N(\mathbf{s}_j)]E[N(\mathbf{s$$
$$= p_{11}^* E[N(\mathbf{s}_i)N(\mathbf{s}_j)] - \Pr(B(\mathbf{s}_i) = 1)E[N(\mathbf{s}_i)]\Pr(B(\mathbf{s}_j) = 1)E[N(\mathbf{s}_j)]$$

$$= p_{11}^* E[N(\mathbf{s}_i)N(\mathbf{s}_j)] - \Pr(B(\mathbf{s}_i) = 1)\lambda_i\Pr(B(\mathbf{s}_j) = 1)\lambda_j$$

$$= p_{11}^*[\operatorname{Cov}(N(\mathbf{s}_i), N(\mathbf{s}_j)) + \lambda_i\lambda_j] - \Pr(B(\mathbf{s}_i) = 1)\Pr(B(\mathbf{s}_j) = 1)\lambda_i\lambda_j$$

$$= p_{11}^*\operatorname{Cov}(N(\mathbf{s}_i), N(\mathbf{s}_j)) + \lambda_i\lambda_j[p_{11}^* - \Pr(B(\mathbf{s}_i) = 1)\Pr(B(\mathbf{s}_j) = 1)]$$

$$= p_{11}^*\operatorname{Cov}(N(\mathbf{s}_i), N(\mathbf{s}_j)) + \lambda_i\lambda_j\operatorname{Cov}(B(\mathbf{s}_i), B(\mathbf{s}_j))$$

Using this result and $Var(N^*(\mathbf{s}_k)) = (1 - p_k)\lambda_k[1 + p_k\lambda_k]$, k = i, j, the correlation of the ZIP random field can be written as follows:

$$\begin{split} \rho_{N^*}(\mathbf{s}_i, \mathbf{s}_i) = & \frac{p_{11}^* \operatorname{Cov}(N(\mathbf{s}_i), N(\mathbf{s}_j)) + \lambda_i \lambda_j \operatorname{Cov}(B(\mathbf{s}_i), B(\mathbf{s}_j))}{\sqrt{(1 - p_i)\lambda_i [1 + p_i\lambda_i](1 - p_j)\lambda_j [1 + p_j\lambda_j]}} \\ = & \frac{p_{11}^* \operatorname{Cov}(N(\mathbf{s}_i), N(\mathbf{s}_j))}{\sqrt{(1 - p_i)\lambda_i [1 + p_i\lambda_i](1 - p_j)\lambda_j [1 + p_j\lambda_j]}} + \\ & \frac{\lambda_i \lambda_j \operatorname{Cov}(B(\mathbf{s}_i), B(\mathbf{s}_j))}{\sqrt{(1 - p_i)\lambda_i [1 + p_i\lambda_i](1 - p_j)\lambda_j [1 + p_j\lambda_j]}} \\ = & \frac{p_{11}^*}{\sqrt{(1 - p_i)[1 + p_i\lambda_i](1 - p_j)[1 + p_j\lambda_j]}} \frac{\operatorname{Cov}(N(\mathbf{s}_i), N(\mathbf{s}_j))}{\sqrt{\lambda_i\lambda_j}} + \\ & \frac{\lambda_i \lambda_j \sqrt{p_i p_j}}{\sqrt{\lambda_i [1 + p_i\lambda_i]\lambda_j [1 + p_j\lambda_j]}} \frac{\operatorname{Cov}(B(\mathbf{s}_i), B(\mathbf{s}_j))}{\sqrt{p_i p_j (1 - p_i)(1 - p_j)(1 + p_i\lambda_i)(1 + p_j\lambda_j)}} + \frac{\sqrt{\lambda_i \lambda_j p_i p_j} \rho_B(\mathbf{s}_i, \mathbf{s}_j)}{\sqrt{(1 + p_i\lambda_i)(1 + p_j\lambda_j)}} \end{split}$$

Proof of Corollary 3.2.1. If the ZIP random field is s weakly stationary then the underlying correlations functions are stationary. Then $\rho_N(\mathbf{s}_i, \mathbf{s}_j) = \rho_N(\mathbf{h})$ and $\rho_B(\mathbf{s}_i, \mathbf{s}_j) = \rho_B(\mathbf{h})$. Moreover, $\lambda(\mathbf{s}) = \lambda$ and $p_i = p_j = p = \Phi(\theta)$. Thus, using these in the theorem 3.2.1 and making algebra the result holds, that is,

$$\rho_{N^*}(\mathbf{h},\lambda,\theta) = \frac{p_{11}^* \rho_N(\mathbf{h})}{(1-p)(1+p\lambda)} + \frac{p\lambda\rho_B(\mathbf{h})}{1+p\lambda}.$$

Proof of Theorem 3.2.2. Using the law of total probability ,the substitution theorem of conditional probabilities and the independence of N and B in $Pr(N^*(\mathbf{s}_i) = n, N^*(\mathbf{s}_j) = n)$

m), the bivariate distribution of de ZIP random field N^* can be obtained as follows:

$$\begin{aligned} \Pr(N^*(\mathbf{s}_i) = n, N^*(\mathbf{s}_j) = m) &= \Pr(B(\mathbf{s}_i)N(\mathbf{s}_i) = n, B(\mathbf{s}_j)N(\mathbf{s}_j) = m) \\ &= \Pr(B(\mathbf{s}_i) = 0, B(\mathbf{s}_j) = 0) \Pr(B(\mathbf{s}_i)N(\mathbf{s}_i) = n, B(\mathbf{s}_j)N(\mathbf{s}_j) = m|B(\mathbf{s}_i) = 0, B(\mathbf{s}_j) = 0) \\ &+ \Pr(B(\mathbf{s}_i) = 0, B(\mathbf{s}_j) = 1) \Pr(B(\mathbf{s}_i)N(\mathbf{s}_i) = n, B(\mathbf{s}_j)N(\mathbf{s}_j) = m|B(\mathbf{s}_i) = 0, B(\mathbf{s}_j) = 1) \\ &+ \Pr(B(\mathbf{s}_i) = 1, B(\mathbf{s}_j) = 0) \Pr(B(\mathbf{s}_i)N(\mathbf{s}_i) = n, B(\mathbf{s}_j)N(\mathbf{s}_j) = m|B(\mathbf{s}_i) = 1, B(\mathbf{s}_j) = 0) \\ &+ \Pr(B(\mathbf{s}_i) = 1, B(\mathbf{s}_j) = 1) \Pr(B(\mathbf{s}_i)N(\mathbf{s}_i) = n, B(\mathbf{s}_j)N(\mathbf{s}_j) = m|B(\mathbf{s}_i) = 1, B(\mathbf{s}_j) = 1) \\ &= p_{00}^* \Pr(B(\mathbf{s}_i)N(\mathbf{s}_i) = n, B(\mathbf{s}_j)N(\mathbf{s}_j) = m|B(\mathbf{s}_i) = 0, B(\mathbf{s}_j) = 0) \\ &+ p_{01}^* \Pr(B(\mathbf{s}_i)N(\mathbf{s}_i) = n, B(\mathbf{s}_j)N(\mathbf{s}_j) = m|B(\mathbf{s}_i) = 1, B(\mathbf{s}_j) = 0) \\ &+ p_{10}^* \Pr(B(\mathbf{s}_i)N(\mathbf{s}_i) = n, B(\mathbf{s}_j)N(\mathbf{s}_j) = m|B(\mathbf{s}_i) = 1, B(\mathbf{s}_j) = 0) \\ &+ p_{11}^* \Pr(B(\mathbf{s}_i)N(\mathbf{s}_i) = n, B(\mathbf{s}_j)N(\mathbf{s}_j) = m|B(\mathbf{s}_i) = 1, B(\mathbf{s}_j) = 1). \end{aligned}$$

If n = m = 0,

$$Pr(N^{*}(\mathbf{s}_{i}) = 0, N^{*}(\mathbf{s}_{j}) = 0) = p_{00}^{*} + p_{01}^{*} Pr(N(\mathbf{s}_{j}) = 0) + p_{10}^{*} Pr(N(\mathbf{s}_{i}) = 0) + p_{11}^{*} Pr(N(\mathbf{s}_{i}) = 0, N(\mathbf{s}_{j}) = 0) = p_{00}^{*} + p_{01}^{*} Pr(N(\mathbf{s}_{j}) = 0) + p_{10}^{*} Pr(N(\mathbf{s}_{i}) = 0) + p_{11}^{*} p_{00}$$

If n = 0 and m > 0,

$$\Pr(N^*(\mathbf{s}_i) = 0, N^*(\mathbf{s}_j) = m) = p_{01}^* \Pr(N(\mathbf{s}_j) = m) + p_{11}^* \Pr(N(\mathbf{s}_i) = 0, N(\mathbf{s}_j) = m)$$
$$= p_{01}^* \Pr(N(\mathbf{s}_j) = m) + p_{11}^* p_{0m}$$

If n > 0 and m = 0,

$$Pr(N^*(\mathbf{s}_i) = n, N^*(\mathbf{s}_j) = 0) = p_{10}^* Pr(N(\mathbf{s}_i) = n) + p_{11}^* Pr(N(\mathbf{s}_i) = n, N(\mathbf{s}_j) = 0)$$
$$= p_{10}^* Pr(N(\mathbf{s}_i) = n) + p_{11}^* p_{n0}$$

If n > 0 and m > 0,

$$\Pr(N^*(\mathbf{s}_i) = n, N^*(\mathbf{s}_j) = m) = p_{11}^* \Pr(N(\mathbf{s}_i) = n, N(\mathbf{s}_j) = m)$$
$$= p_{11}^* p_{nm}$$

Proof of Theorem 3.3.1. Let $\psi_{N^*(\mathbf{s}_i),N^*(\mathbf{s}_j)}(t_i,t_j)$ be the characteristic function for $(N^*(\mathbf{s}_i),N^*(\mathbf{s}_j))^{\top}$, then the bivariate density function of $N^*(s)$ is defined as follows:

$$f_{N^{*}(\mathbf{s}_{i}),N^{*}(\mathbf{s}_{j})}(n,m) = \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{1}{(2\pi)^{2}} \exp\{-i(t_{i}n+t_{j}m)\}\psi_{N^{*}(\mathbf{s}_{i}),N^{*}(\mathbf{s}_{j})}(t_{i},t_{j})dt_{i}dt_{j}.$$

Now, note that $\psi_{{}^{N^*(\mathbf{s}_i),N^*(\mathbf{s}_j)}}(t_i,t_j)$ can be written as:

$$\begin{split} \psi_{N^*(\mathbf{s}_i),N^*(\mathbf{s}_j)}(t_i,t_j) =& E[\exp\{i(t_iN^*(\mathbf{s}_i) + t_jN^*(s_j))\}] \\ =& E[\exp\{i(t_iB(\mathbf{s}_i)G(\mathbf{s}_i) + t_jB(\mathbf{s}_j)G(\mathbf{s}_j))\}|B(\mathbf{s}_i) = 1, B(\mathbf{s}_j) = 1]P(B(\mathbf{s}_i) = 1, B(\mathbf{s}_j) = 1) \\ &+ E[\exp\{i(t_iB(\mathbf{s}_i)G(\mathbf{s}_i) + t_jB(\mathbf{s}_j)G(\mathbf{s}_j))\}|B(\mathbf{s}_i) = 1, B(\mathbf{s}_j) = 0]P(B(\mathbf{s}_i) = 1, B(\mathbf{s}_j) = 0) \\ &+ E[\exp\{i(t_iB(\mathbf{s}_i)G(\mathbf{s}_i) + t_jB(\mathbf{s}_j)G(\mathbf{s}_j))\}|B(\mathbf{s}_i) = 0, B(\mathbf{s}_j) = 1]P(B(\mathbf{s}_i) = 0, B(\mathbf{s}_j) = 1) \\ &+ E[\exp\{i(t_iB(\mathbf{s}_i)G(\mathbf{s}_i) + t_jB(\mathbf{s}_j)G(\mathbf{s}_j))\}|B(\mathbf{s}_i) = 0, B(\mathbf{s}_j) = 1]P(B(\mathbf{s}_i) = 0, B(\mathbf{s}_j) = 1) \\ &+ E[\exp\{i(t_iB(\mathbf{s}_i)G(\mathbf{s}_i) + t_jB(\mathbf{s}_j)G(\mathbf{s}_j))\}|B(\mathbf{s}_i) = 0, B(\mathbf{s}_j) = 0]P(B(\mathbf{s}_i) = 0, B(\mathbf{s}_j) = 0) \\ &= E[\exp\{i(t_iG(\mathbf{s}_i) + t_jG(\mathbf{s}_j))\}]p_{11}^* \\ &+ E[\exp\{i(t_iG(\mathbf{s}_i) + t_jG(\mathbf{s}_j))\}]p_{10}^* + p_{00}^* \\ &= p_{11}^*\psi_{G(\mathbf{s}_i),G(\mathbf{s}_j)}(t_i,t_j) + p_{10}^*\psi_{G(\mathbf{s}_i)}(t_i) + p_{01}^*\psi_{G(\mathbf{s}_j)}(t_j) + p_{00}^* \end{split}$$

Then,

$$\begin{split} f_{N^*(\mathbf{s}_i),N^*(\mathbf{s}_j)}(n,m) &= \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{1}{(2\pi)^2} \exp\{-i(t_in+t_jm)\} [p_{11}^*\psi_{G(\mathbf{s}_i)G(\mathbf{s}_j)}(t_i,t_j) + p_{10}^*\psi_{G(\mathbf{s}_i)}(t_i)] dt_i dt_j \\ &+ \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{1}{(2\pi)^2} \exp\{-i(t_in+t_jm)\} [p_{01}^*\psi_{G(\mathbf{s}_j)}(t_j) + p_{00}^*] dt_i dt_j \\ &= p_{11}^* f_{G(\mathbf{s}_i),G(\mathbf{s}_j)}(n,m) + p_{10}^* f_{G(\mathbf{s}_i)}(n) \int_{\mathbb{R}} \frac{1}{2\pi} \exp\{-i(t_jm)\} dt_j \\ &+ p_{01}^* f_{G(\mathbf{s}_j)}(m) \int_{\mathbb{R}} \frac{1}{2\pi} \exp\{-i(t_in)\} dt_i \\ &+ p_{00}^* \int_{\mathbb{R}} \frac{1}{2\pi} \exp\{-i(t_in)\} dt_i \int_{\mathbb{R}} \frac{1}{2\pi} \exp\{-i(t_jm)\} dt_j \end{split}$$

Note that $\psi_{\delta_0}(t) = 1$ is the characteristic function of the degenerate distribution at 0. Moreover, $f_{\delta_0}(y) = \int_{\mathbb{R}} \frac{1}{2\pi} \exp\{-i(ty)\} dt_j = 1$ if y = 0 and $f_{\delta_0}(y) = 0$ if $y \neq 0$. Thus, we obtain that:

$$f_{N^{*}(\mathbf{s}_{i}),N^{*}(\mathbf{s}_{j})}(n,m) = p_{11}^{*}f_{G(\mathbf{s}_{i}),G(\mathbf{s}_{j})}(n,m) + p_{10}^{*}f_{G(\mathbf{s}_{i})}(n)f_{\delta_{0}}(m) + p_{01}^{*}f_{G(\mathbf{s}_{j})}(m)f_{\delta_{0}}(m) + p_{00}^{*}f_{\delta_{0}}(n)f_{\delta_{0}}(m)$$

Finally,

If n = 0 and m = 0

$$f_{N^*(\mathbf{s}_i),N^*(\mathbf{s}_j)}(0,0) = p_{11}^* f_{G(\mathbf{s}_i),G(\mathbf{s}_j)}(0,0) + p_{10}^* f_{G(\mathbf{s}_i)}(0) + p_{01}^* f_{G(\mathbf{s}_j)}(0) + p_{00}^*$$

If n = 0 and m > 0

$$f_{N^*(\mathbf{s}_i),N^*(\mathbf{s}_j)}(0,m) = p_{11}^* f_{G(\mathbf{s}_i),G(\mathbf{s}_j)}(0,m) + p_{01}^* f_{G(\mathbf{s}_j)}(m)$$

If n > 0 and m = 0

$$f_{N^*(\mathbf{s}_i),N^*(\mathbf{s}_j)}(n,0) = p_{11}^* f_{G(\mathbf{s}_i),G(\mathbf{s}_j)}(n,0) + p_{10}^* f_{G(\mathbf{s}_i)}(n)$$

If
$$n > 0$$
 and $m > 0$

 $f_{N^*(\mathbf{s}_i),N^*(\mathbf{s}_j)}(y_i,y_j) = p_{11}^* f_{G(\mathbf{s}_i),G(\mathbf{s}_j)}(n,m)$

A.3 Proofs of Chapter 4

Proof of Theorem 4.2.1 For the simplicity of notation, let $\nu(\mathbf{s}_i) = \nu_i$, $\nu(\mathbf{s}_j) = \nu_j$, $\lambda(\mathbf{s}_i) = \lambda_i$, $\lambda(\mathbf{s}_j) = \lambda_j$ and $\rho(\mathbf{h}) = \rho$. Note that the covariance can be written as follows:

 $Cov(M(\mathbf{s}_i), M(\mathbf{s}_j)) = \mathbb{E}[Cov(N(\mathbf{s}_i), N(\mathbf{s}_j) | \Lambda(\mathbf{s}_i), \Lambda(\mathbf{s}_j))]$

$$+ \operatorname{Cov}(\mathbb{E}[N(\mathbf{s}_{i})|\Lambda(\mathbf{s}_{i})], \mathbb{E}[N(\mathbf{s}_{j})|\Lambda(\mathbf{s}_{j})])$$

$$= \mathbb{E}[\operatorname{Cov}(N(\mathbf{s}_{i}), N(\mathbf{s}_{j})|\Lambda(\mathbf{s}_{i}), \Lambda(\mathbf{s}_{j}))] + \operatorname{Cov}(\Lambda(\mathbf{s}_{i}), \Lambda(\mathbf{s}_{j}))$$

$$= \mathbb{E}\left[\rho^{2}(1-\rho^{2})\sum_{r=0}^{\infty}\frac{1}{\Gamma(r+1)^{2}}\gamma\left(r+1, \frac{\Lambda(\mathbf{s}_{i})}{1-\rho^{2}}\right)\gamma\left(r+1, \frac{\Lambda(\mathbf{s}_{j})}{1-\rho^{2}}\right)\right]$$

$$+ \rho^{2}\frac{\kappa}{\nu_{i}\nu_{j}}$$

$$= \sum_{r=0}^{\infty}\frac{\rho^{2}(1-\rho^{2})}{\Gamma(r+1)^{2}}\mathbb{E}\left[\gamma\left(r+1, \frac{\Lambda(\mathbf{s}_{i})}{1-\rho^{2}}\right)\gamma\left(r+1, \frac{\Lambda(\mathbf{s}_{j})}{1-\rho^{2}}\right)\right]$$

$$+ \rho^{2}\frac{\kappa}{\nu_{i}\nu_{j}}$$
(A.4)

Now, $\mathbb{E}\left[\gamma\left(r+1,\frac{\Lambda(\mathbf{s}_i)}{1-\rho^2}\right)\gamma\left(r+1,\frac{\Lambda(\mathbf{s}_j)}{1-\rho^2}\right)\right]$, which is denoted by \mathcal{I} , is calculated as follows.

$$\begin{split} \mathcal{I} &= \int_{0}^{\infty} \int_{0}^{\infty} \gamma \left(r+1, \frac{\lambda_{i}}{1-\rho^{2}} \right) \gamma \left(r+1, \frac{\lambda_{j}}{1-\rho^{2}} \right) \frac{\nu_{i}\nu_{j}}{(1-\rho^{2})\Gamma(\kappa)} \left(\frac{\nu_{i}\nu_{j}\lambda_{i}\lambda_{j}}{\rho^{2}} \right)^{\frac{\kappa-1}{2}} \\ &\qquad \times \exp \left\{ -\frac{\nu_{i}\lambda_{i}+\nu_{j}\lambda_{j}}{1-\rho^{2}} \right\} I_{\kappa-1} \left(\frac{2\rho}{1-\rho^{2}} \sqrt{\nu_{i}\nu_{j}\lambda_{i}\lambda_{j}} \right) d\lambda_{i}d\lambda_{j} \\ &= \int_{0}^{\infty} \int_{0}^{\infty} \exp \left\{ -\frac{\lambda_{i}}{1-\rho^{2}} \right\} \sum_{k_{1}=0}^{\infty} \frac{\Gamma(r+1)}{\Gamma(r+k_{1}+2)} \left(\frac{\lambda_{i}}{1-\rho^{2}} \right)^{k_{1}+r+1} \\ &\qquad \times \exp \left\{ -\frac{\lambda_{j}}{1-\rho^{2}} \right\} \sum_{k_{2}=0}^{\infty} \frac{\Gamma(r+1)}{\Gamma(r+k_{2}+2)} \left(\frac{\lambda_{j}}{1-\rho^{2}} \right)^{k_{2}+r+1} \\ &\qquad \times \frac{(\nu_{i}\nu_{j})^{\frac{\kappa+1}{2}}}{(1-\rho^{2})\Gamma(\kappa)} \left(\frac{\lambda_{i}\lambda_{j}}{\rho^{2}} \right)^{\frac{\kappa-1}{2}} \exp \left\{ -\frac{\nu_{i}\lambda_{i}+\nu_{j}\lambda_{j}}{1-\rho^{2}} \right\} \\ &\qquad \times \sum_{\ell=0}^{\infty} \frac{1}{\ell!\Gamma(\ell+\kappa)} \left(\frac{\rho}{1-\rho^{2}} \sqrt{\nu_{i}\nu_{j}\lambda_{i}\lambda_{j}} \right)^{2\ell+\kappa-1} d\lambda_{i}d\lambda_{j} \\ &= \sum_{\ell=0}^{\infty} \sum_{k_{1}=0}^{\infty} \sum_{k_{2}=0}^{\infty} \frac{\Gamma(r+1)2(\nu_{i}\nu_{j})^{\ell+\kappa}\rho^{2\ell}}{\Gamma(r+k_{1}+2)\Gamma(r+k_{2}+2)\Gamma(\kappa)\ell!\Gamma(\ell+\kappa)(1-\rho^{2})^{k_{1}+k_{2}+2r+2\ell+2+\kappa}} \\ &\qquad \times \int_{0}^{\infty} \lambda_{i}^{k_{1}+r+\ell+\kappa} \exp \left\{ -\frac{(1+\nu_{i})\lambda_{i}}{1-\rho^{2}} \right\} d\lambda_{i} \int_{0}^{\infty} \lambda_{j}^{k_{2}+r+\ell+\kappa} \exp \left\{ -\frac{(1+\nu_{j})\lambda_{j}}{1-\rho^{2}} \right\} d\lambda_{j} \end{split}$$

$$\begin{split} &= \sum_{\ell=0}^{\infty} \frac{\Gamma(r+1)^2 (\nu_i \nu_j)^{\ell+\kappa} \rho^{2\ell} ((1+\nu_i)(1+\nu_j))^{-(r+\ell+\kappa+1)}}{\Gamma(\kappa)\ell!\Gamma(\ell+\kappa)(1-\rho^2)^{-\kappa}} \\ &\qquad \times \sum_{k_1=0}^{\infty} \frac{\Gamma(k_1+r+\ell+\kappa+1)}{\Gamma(r+k_1+2)} (1+\nu_i)^{-k_1} \sum_{k_2=0}^{\infty} \frac{\Gamma(k_2+r+\ell+\kappa+1)}{\Gamma(r+k_2+2)} (1+\nu_j)^{-k_2} \\ &= \sum_{\ell=0}^{\infty} \frac{\Gamma(r+1)^2 (\nu_i \nu_j)^{\ell+\kappa} \rho^{2\ell}}{\Gamma(\kappa)\ell!\Gamma(\ell+\kappa)(1-\rho^2)^{-\kappa} ((1+\nu_i)(1+\nu_j))^{r+\ell+\kappa+1}} \\ &\qquad \times \left(\frac{\Gamma(r+\ell+\kappa+1)}{\Gamma(r+2)}\right)^2 {}_2F_1 \left(1,r+\ell+\kappa+1;r+2;\frac{1}{\nu_i+1}\right) \\ &\qquad \times {}_2F_1 \left(1,r+\ell+\kappa+1;r+2;\frac{1}{\nu_j+1}\right) \\ &= \sum_{\ell=0}^{\infty} \frac{\Gamma(r+1)^2 (\nu_i \nu_j)^{\ell+\kappa-1} \rho^{2\ell}}{\Gamma(\kappa)\ell!\Gamma(\ell+\kappa)(1-\rho^2)^{-\kappa} ((1+\nu_i)(1+\nu_j))^{r+\ell+\kappa}} \\ &\qquad \times \left(\frac{\Gamma(r+\ell+\kappa+1)}{\Gamma(r+2)}\right)^2 {}_2F_1 \left(1,1-\ell-\kappa;r+2;-\frac{1}{\nu_i}\right) {}_2F_1 \left(1,1-\ell-\kappa;r+2;-\frac{1}{\nu_j}\right) \\ &= \sum_{\ell=0}^{\infty} \frac{\Gamma(r+1)^2 (\nu_i \nu_j)^{\ell+\kappa-1} \rho^{2\ell} \Gamma(r+\ell+\kappa+1)^2}{\Gamma(\kappa)\ell!\Gamma(\ell+\kappa)(1-\rho^2)^{-\kappa} ((1+\nu_i)(1+\nu_j))^{r+\ell+\kappa}} \mathcal{H} \left(\frac{1,1-\kappa-\ell}{r+2},-\frac{1}{\nu_i},-\frac{1}{\nu_j}\right) \end{split}$$

Plugging in this on A.4, dividing the result by $\sqrt{\operatorname{Var}(M(\mathbf{s}_i))\operatorname{Var}(M(\mathbf{s}_j))}$ and doing some algebra, the result holds.

Proof of Corollary 4.2.1. Let $\rho(\mathbf{h}) = \rho$. Note that the stationary covariance function can be written as follows:

$$\begin{aligned} \operatorname{Cov}(M(\mathbf{s}_{i}), M(\mathbf{s}_{j})) &= \mathbb{E}[\operatorname{Cov}(N(\mathbf{s}_{i}), N(\mathbf{s}_{j}) | \Lambda(\mathbf{s}))] + \operatorname{Cov}(\mathbb{E}[N(\mathbf{s}_{i}) | \Lambda(\mathbf{s}_{i})], \mathbb{E}[N(\mathbf{s}_{j}) | \Lambda(\mathbf{s}_{j})]) \\ &= \mathbb{E}[\operatorname{Cov}(N(\mathbf{s}_{i}), N(\mathbf{s}_{j}) | \Lambda(\mathbf{s}))] + \operatorname{Cov}(\Lambda(\mathbf{s}_{i}), \Lambda(\mathbf{s}_{j})) \\ &= \mathbb{E}\left[\Lambda(\mathbf{s})\rho^{2}\left(1 - \exp\left\{-\frac{2\Lambda(\mathbf{s})}{1 - \rho^{2}}\right\}\left(I_{0}\left(\frac{2\Lambda(\mathbf{s})}{1 - \rho^{2}}\right) + I_{1}\left(\frac{2\Lambda(\mathbf{s})}{1 - \rho^{2}}\right)\right)\right)\right] \\ &+ \rho^{2}\frac{\kappa}{\nu^{2}} \\ &= \rho^{2}\frac{\kappa}{\nu} - \rho^{2}\mathbb{E}\left[\Lambda(\mathbf{s})\exp\left\{-\frac{2\Lambda(\mathbf{s})}{1 - \rho^{2}}\right\}I_{0}\left(\frac{2\Lambda(\mathbf{s})}{1 - \rho^{2}}\right)\right] \end{aligned}$$

$$-\rho^{2}\mathbb{E}\left[\Lambda(\mathbf{s})\exp\left\{-\frac{2\Lambda(\mathbf{s})}{1-\rho^{2}}\right\}I_{1}\left(\frac{2\Lambda(\mathbf{s})}{1-\rho^{2}}\right)\right]+\rho^{2}\frac{\kappa}{\nu^{2}}$$
(A.5)

Then,

and

$$\mathcal{I}_1 = \mathbb{E}\left[\Lambda(\mathbf{s}) \exp\left\{-\frac{2\Lambda(\mathbf{s})}{1-\rho^2}\right\} I_0\left(\frac{2\Lambda(\mathbf{s})}{1-\rho^2}\right)\right]$$
$$\mathcal{I}_2 = \mathbb{E}\left[\Lambda(\mathbf{s}) \exp\left\{-\frac{2\Lambda(\mathbf{s})}{1-\rho^2}\right\} I_1\left(\frac{2\Lambda(\mathbf{s})}{1-\rho^2}\right)\right]$$

are calculated as follow:

$$\begin{split} \mathcal{I}_{1} &= \int_{0}^{\infty} \lambda \exp\left\{-\frac{2\lambda}{1-\rho^{2}}\right\} \sum_{m=0}^{\infty} \frac{1}{m!\Gamma(m+1)} \left(\frac{\lambda}{1-\rho^{2}}\right)^{2m} \frac{\nu^{\kappa}\lambda^{\kappa-1}}{\Gamma(\kappa)} \exp\left\{-\nu\lambda\right\} d\lambda \\ &= \sum_{m=0}^{\infty} \frac{\nu^{\kappa}}{m!^{2}(1-\rho^{2})^{2m}\Gamma(\kappa)} \int_{0}^{\infty} \lambda^{2m+\kappa} \exp\left\{-\left(\frac{2}{1-\rho^{2}}+\nu\right)\lambda\right\} d\lambda \\ &= \sum_{m=0}^{\infty} \frac{\nu^{\kappa}}{m!^{2}(1-\rho^{2})^{2m}\Gamma(\kappa)} \left(\frac{2}{1-\rho^{2}}+\nu\right)^{-(2m+\kappa+1)} \Gamma(2m+\kappa+1) \\ &= \frac{\nu^{\kappa}}{\Gamma(\kappa)} \left(\frac{2}{1-\rho^{2}}+\nu\right)^{-(\kappa+1)} \sum_{m=0}^{\infty} \left(\frac{1}{2+\nu(1-\rho^{2})}\right)^{2m} \frac{\Gamma(2m+\kappa+1)}{m!^{2}} \\ &= \nu^{\kappa}\kappa \left(\frac{1-\rho^{2}}{2+\nu(1-\rho^{2})}\right)^{\kappa+1} {}_{2}F_{1}\left(\frac{\kappa+1}{2},\frac{\kappa+2}{2};1;\frac{4}{(2+\nu(1-\rho^{2}))^{2}}\right) \\ &= \kappa \left(\frac{1-\rho^{2}}{\nu}\right)^{\frac{1}{2}} \frac{(2+\nu(1-\rho^{2}))^{\kappa}}{(4+\nu(1-\rho^{2}))^{\kappa+\frac{1}{2}}} {}_{2}F_{1}\left(\frac{1-\kappa}{2},-\frac{\kappa}{2};1;\frac{4}{(2+\nu(1-\rho^{2}))^{2}}\right) \end{split}$$

$$\begin{aligned} \mathcal{I}_2 &= \int_0^\infty \lambda \exp\left\{-\frac{2\lambda}{1-\rho^2}\right\} \sum_{m=0}^\infty \frac{1}{m!\Gamma(m+2)} \left(\frac{\lambda}{1-\rho^2}\right)^{2m+1} \frac{\nu^{\kappa}\lambda^{\kappa-1}}{\Gamma(\kappa)} \exp\left\{-\nu\lambda\right\} d\lambda \\ &= \sum_{m=0}^\infty \frac{\nu^{\kappa}}{m!\Gamma(m+2)(1-\rho^2)^{2m+1}\Gamma(\kappa)} \int_0^\infty \lambda^{2m+\kappa+1} \exp\left\{-\left(\frac{2}{1-\rho^2}+\nu\right)\lambda\right\} d\lambda \\ &= \sum_{m=0}^\infty \frac{\nu^{\kappa}}{m!\Gamma(m+2)(1-\rho^2)^{2m+1}\Gamma(\kappa)} \left(\frac{2+\nu(1-\rho^2)}{1-\rho^2}\right)^{-(2m+\kappa+2)} \Gamma(2m+\kappa+2) \end{aligned}$$

$$= \frac{\nu^{\kappa}(2+\nu(1-\rho^2))^{-(\kappa+2)}}{\Gamma(\kappa)(1-\rho^2)^{-(\kappa+1)}} \sum_{m=0}^{\infty} \frac{\Gamma(2m+\kappa+2)}{m!\Gamma(m+2)} \left(2+\nu(1-\rho^2)\right)^{-2m}$$

$$= \frac{\nu^{\kappa}(2+\nu(1-\rho^2))^{-(\kappa+2)}}{\Gamma(\kappa)(1-\rho^2)^{-(\kappa+1)}} \Gamma(\kappa+2)_2 F_1\left(\frac{\kappa+2}{2},\frac{\kappa+3}{2};2;\frac{4}{(2+\nu(1-\rho^2))^2}\right)$$

$$= \kappa(\kappa+1) \left(\frac{1-\rho^2}{\nu}\right)^{\frac{1}{2}} \frac{(2+\nu(1-\rho^2))^{\kappa-1}}{(4+\nu(1-\rho^2))^{\kappa+\frac{1}{2}}} F_1\left(\frac{2-\kappa}{2},\frac{1-\kappa}{2};2;\frac{4}{(2+\nu(1-\rho^2))^2}\right)$$

Plugging in \mathcal{I}_1 and \mathcal{I}_2 on A.5, dividing the result by Var(M(s)) and doing some algebra, the result holds.

Proof of Theorem 4.2.2. Note that bivariate distribution can be obtained as follows

$$\widetilde{p}_{nm} = \int_{0}^{\infty} \int_{0}^{\infty} \Pr(N(\mathbf{s}_i) = n, N(\mathbf{s}_j) = m | \lambda_i \lambda_j) f_{\Lambda_{ij}}(\lambda_i, \lambda_j) d\lambda_i \lambda_j$$

Then, using theorem 2.2.3 and 4.2 the following results are obtained:

If n = m = 0:

$$\begin{split} \widetilde{p}_{00} &= \int_{0}^{\infty} \int_{0}^{\infty} \left[-1 + \exp\{-\lambda_{i}\} + \exp\{-\lambda_{j}\} \right] f_{\Lambda_{ij}}(\lambda_{i},\lambda_{j}) d\lambda_{i}\lambda_{j} \\ &+ \int_{0}^{\infty} \int_{0}^{\infty} \sum_{k=0}^{\infty} \frac{(1-\rho^{2})\rho^{2k}}{\Gamma(k+1)^{2}} \gamma \left(k+1,\frac{\lambda_{i}}{1-\rho^{2}}\right) \gamma \left(k+1,\frac{\lambda_{i}}{1-\rho^{2}}\right) f_{\Lambda_{ij}}(\lambda_{i},\lambda_{j}) d\lambda_{i}\lambda_{j} \\ &= -\int_{0}^{\infty} \int_{0}^{\infty} f_{\Lambda_{ij}}(\lambda_{i},\lambda_{j}) d\lambda_{i}\lambda_{j} \\ &+ \int_{0}^{\infty} \int_{0}^{\infty} \exp\{-\lambda_{i}\} f_{\Lambda_{ij}}(\lambda_{i},\lambda_{j}) d\lambda_{i}\lambda_{j} \\ &+ \int_{0}^{\infty} \int_{0}^{\infty} \exp\{-\lambda_{j}\} f_{\Lambda_{ij}}(\lambda_{i},\lambda_{j}) d\lambda_{i}\lambda_{j} \end{split}$$

$$+ \int_{0}^{\infty} \int_{0}^{\infty} \sum_{k=0}^{\infty} \frac{(1-\rho^2)\rho^{2k}}{\Gamma(k+1)^2} \gamma\left(k+1, \frac{\lambda_i}{1-\rho^2}\right) \gamma\left(k+1, \frac{\lambda_i}{1-\rho^2}\right) f_{\Lambda_{ij}}(\lambda_i, \lambda_j) d\lambda_i \lambda_j$$

Note that the double integral of the first term is equal to 1. The double integral of the second term, noted by \mathcal{I}_1 , is obtained as follow:

The double integral of the third term, noted by \mathcal{I}_2 , is calculated as the previous one. The result is:

$$\mathcal{I}_2 = \left(\frac{\nu_j}{1+\nu_j}\right)^{\kappa}$$

The double integral of the last term, noted by \mathcal{I}_3 , is obtained as follows:

$$\begin{split} \mathcal{I}_{3} &= \int_{0}^{\infty} \int_{0}^{\infty} \gamma \left(k+1, \frac{\lambda_{i}}{1-\rho^{2}}\right) \gamma \left(k+1, \frac{\lambda_{j}}{1-\rho^{2}}\right) \frac{\nu_{i}\nu_{j}}{(1-\rho^{2})\Gamma(\kappa)} \left(\frac{\nu_{i}\nu_{j}\lambda_{i}\lambda_{j}}{\rho^{2}}\right)^{\frac{\kappa-1}{2}} \\ &\quad \times \exp\left\{-\frac{\nu_{i}\lambda_{i}+\nu_{j}\lambda_{j}}{1-\rho^{2}}\right\} I_{\kappa-1} \left(\frac{2\rho}{1-\rho^{2}}\sqrt{\nu_{i}\nu_{j}\lambda_{i}\lambda_{j}}\right) d\lambda_{i}d\lambda_{j} \\ &= \sum_{\ell=0}^{\infty} \sum_{k_{1}=0}^{\infty} \sum_{k_{2}=0}^{\infty} \frac{\Gamma(k+1)^{2}(\nu_{i}\nu_{j})^{\ell+\kappa}\rho^{2\ell}}{\Gamma(k+k_{1}+2)\Gamma(k+k_{2}+2)\Gamma(\kappa)\ell!\Gamma(\ell+\kappa)(1-\rho^{2})^{k_{1}+k_{2}+2k+2\ell+2+\kappa}} \\ &\quad \times \int_{0}^{\infty} \lambda_{i}^{k_{1}+k+\ell+\kappa} \exp\left\{-\frac{(1+\nu_{i})\lambda_{i}}{1-\rho^{2}}\right\} d\lambda_{i} \int_{0}^{\infty} \lambda_{j}^{k_{2}+k+\ell+\kappa} \exp\left\{-\frac{(1+\nu_{j})\lambda_{j}}{1-\rho^{2}}\right\} d\lambda_{j} \\ &= \sum_{\ell=0}^{\infty} \frac{\Gamma(k+1)^{2}(\nu_{i}\nu_{j})^{\ell+\kappa}\rho^{2\ell}((1+\nu_{i})(1+\nu_{j}))^{-(k+\ell+\kappa+1)}}{\Gamma(\kappa)\ell!\Gamma(\ell+\kappa)(1-\rho^{2})^{-\kappa}} \\ &\quad \times \sum_{k_{1}=0}^{\infty} \frac{\Gamma(k_{1}+k+\ell+\kappa+1)}{\Gamma(k+k_{1}+2)} (1+\nu_{i})^{-k_{1}} \sum_{k_{2}=0}^{\infty} \frac{\Gamma(k_{2}+k+\ell+\kappa+1)}{\Gamma(k+k_{2}+2)} (1+\nu_{j})^{-k_{2}} \\ &= \sum_{\ell=0}^{\infty} \frac{\Gamma(k+1)^{2}(\nu_{i}\nu_{j})^{\ell+\kappa-1}\rho^{2\ell}}{\Gamma(\kappa)\ell!\Gamma(\ell+\kappa)(1-\rho^{2})^{-\kappa}((1+\nu_{i})(1+\nu_{j}))^{k+\ell+\kappa}} \\ &\quad \times \left(\frac{\Gamma(k+\ell+\kappa+1)}{\Gamma(k+2)}\right)^{2} {}_{2}F_{1} \left(1,1-\ell-\kappa;k+2;-\frac{1}{\nu_{i}}\right) {}_{2}F_{1} \left(1,1-\ell-\kappa;k+2;-\frac{1}{\nu_{j}}\right) \\ &= \sum_{\ell=0}^{\infty} \frac{\Gamma(r+1)^{2}(\nu_{i}\nu_{j})^{\ell+\kappa-1}\rho^{2\ell}\Gamma(k+\ell+\kappa+1)^{2}}{\Gamma(\kappa)\ell!\Gamma(\ell+\kappa)(1-\rho^{2})^{-\kappa}((1+\nu_{i})(1+\nu_{j}))^{r+\ell+\kappa}} \times \mathcal{H} \left(\frac{1,1-\kappa-\ell}{k+2},-\frac{1}{\nu_{j}},-\frac{1}{\nu_{j}}\right) \end{split}$$

Thus, plugging in \mathcal{I}_1 , \mathcal{I}_2 and \mathcal{I}_3 on \widetilde{p}_{00} and doing some algebra, the result holds. If $n \ge 1, m = 0$:

$$\widetilde{p}_{n0} = \int_{0}^{\infty} \int_{0}^{\infty} \frac{\lambda_{i}^{n}}{n!} \exp\{-\lambda_{i}\} f_{\Lambda_{ij}}(\lambda_{i},\lambda_{j}) d\lambda_{i} d\lambda_{j} - \int_{0}^{\infty} \int_{0}^{\infty} \exp\{-\frac{\lambda_{i}}{1-\rho^{2}}\} \sum_{\ell=0}^{\infty} \left(\frac{\rho^{2}}{1-\rho^{2}}\right)^{\ell} \frac{\lambda_{i}^{\ell+n}}{\Gamma(\ell+n+1)\Gamma(\ell+1)}$$

$$\times {}_{1}F_{1}\left(n,\ell+n+1;\frac{\rho^{2}\lambda_{i}}{1-\rho^{2}}\right)\gamma\left(\ell+1,\frac{\lambda_{i}}{1-\rho^{2}}\right)f_{\Lambda_{ij}}(\lambda_{i},\lambda_{j})d\lambda_{i}d\lambda_{j}$$

$$= \frac{1}{n!}\int_{0}^{\infty}\int_{0}^{\infty}\lambda_{i}^{n}\exp\{-\lambda_{i}\}f_{\Lambda_{ij}}(\lambda_{i},\lambda_{j})d\lambda_{i}d\lambda_{j}$$

$$-\sum_{\ell=0}^{\infty}\left(\frac{\rho^{2}}{1-\rho^{2}}\right)^{\ell}\frac{1}{\Gamma(\ell+n+1)\Gamma(\ell+1)}$$

$$\times\int_{0}^{\infty}\int_{0}^{\infty}\exp\left\{-\frac{\lambda_{i}}{1-\rho^{2}}\right\}\lambda_{i}^{\ell+n}{}_{1}F_{1}\left(n,\ell+n+1;\frac{\rho^{2}\lambda_{i}}{1-\rho^{2}}\right)\gamma\left(\ell+1,\frac{\lambda_{i}}{1-\rho^{2}}\right)f_{\Lambda_{ij}}(\lambda_{i},\lambda_{j})d\lambda_{i}d\lambda_{j}$$

The double integral of the first term, noted by \mathcal{I}_1 , is obtained as follows:

$$\begin{split} \mathcal{I}_{1} &= \int_{0}^{\infty} \int_{0}^{\infty} \frac{\lambda_{i}^{n} \exp\{-\lambda_{i}\}\nu_{i}\nu_{j}}{(1-\rho^{2})\Gamma(\kappa)} \left(\frac{\nu_{i}\nu_{j}\lambda_{i}\lambda_{j}}{\rho^{2}}\right)^{\frac{\kappa-1}{2}} \exp\left\{-\frac{\nu_{i}\lambda_{i}+\nu_{j}\lambda_{j}}{1-\rho^{2}}\right\} I_{\kappa-1} \left(\frac{2\rho}{1-\rho^{2}}\sqrt{\nu_{i}\nu_{j}\lambda_{i}\lambda_{j}}\right) d\lambda_{i}d\lambda_{j} \\ &= \sum_{k=0}^{\infty} \frac{(\nu_{i}\nu_{j})^{k+\kappa}\rho^{2k}}{\Gamma(\kappa)k!\Gamma(k+\kappa)(1-\rho^{2})^{2k+\kappa}} \\ &\qquad \times \int_{0}^{\infty} \lambda_{i}^{n+k+\kappa-1} \exp\left\{-\frac{(1-\rho^{2}+\nu_{i})\lambda_{i}}{1-\rho^{2}}\right\} d\lambda_{i} \int_{0}^{\infty} \lambda_{j}^{k+\kappa-1} \exp\left\{-\frac{\nu_{i}\lambda_{j}}{1-\rho^{2}}\right\} d\lambda_{j} \\ &= \sum_{k=0}^{\infty} \frac{\nu_{i}^{k+\kappa}\rho^{2k}(1-\rho^{2}+\nu_{i})^{-(n+\kappa+\kappa)}\Gamma(n+k+\kappa)}{\Gamma(\kappa)k!(1-\rho^{2})^{-(n+\kappa)}} \\ &= \frac{\nu_{i}^{\kappa}(1-\rho^{2}+\nu_{i})^{-(n+\kappa)}}{\Gamma(\kappa)(1-\rho^{2})^{-(n+\kappa)}} \sum_{k=0}^{\infty} \frac{\nu_{i}^{k}\rho^{2k}(1-\rho^{2}+\nu_{i})^{-k}\Gamma(n+k+\kappa)}{k!} \\ &= \frac{\nu_{i}^{\kappa}(1-\rho^{2}+\nu_{i})^{-(n+\kappa)}}{\Gamma(\kappa)(1-\rho^{2})^{-(n+\kappa)}} \left(1-\frac{\rho^{2}\nu_{i}}{1-\rho^{2}+\nu_{i}}\right)^{-(n+\kappa)} \Gamma(n+\kappa) \\ &= \frac{(\kappa)n\nu_{i}^{\kappa}}{(1+\nu_{i})^{n+\kappa}} \end{split}$$

The double integral of the second term, noted by \mathcal{I}_2 , is obtained as follows:

$$\begin{split} \mathcal{I}_{2} &= \int_{0}^{\infty} \int_{0}^{\infty} \exp\left\{-\frac{\lambda_{i}}{1-\rho^{2}}\right\} \lambda_{i}^{\ell+n} {}_{1}F_{1}\left(n,\ell+n+1;\frac{\rho^{2}\lambda_{i}}{1-\rho^{2}}\right) \gamma\left(\ell+1,\frac{\lambda_{i}}{1-\rho^{2}}\right) \\ &\quad \times \frac{\nu_{i}\nu_{j}}{(1-\rho^{2})\Gamma(\kappa)} \left(\frac{\nu_{i}\nu_{j}\lambda_{i}\lambda_{j}}{\rho^{2}}\right)^{\frac{\kappa-1}{2}} \exp\left\{-\frac{\nu_{i}\lambda_{i}+\nu_{j}\lambda_{j}}{1-\rho^{2}}\right\} I_{\kappa-1}\left(\frac{2\rho}{1-\rho^{2}}\sqrt{\nu_{i}\nu_{j}\lambda_{i}\lambda_{j}}\right) d\lambda_{i}d\lambda_{j} \\ &= \sum_{\ell_{1}=0}^{\infty} \sum_{k_{2}=0}^{\infty} \sum_{k_{2}=0}^{\infty} \frac{\Gamma(n+k_{1})\Gamma(\ell+n+1)\Gamma(\ell+1)(\nu_{i}\nu_{j})^{\ell_{1}+\kappa}\rho^{2k_{1}+2\ell_{1}}}{\Gamma(\kappa)\Gamma(\ell+n+k_{1}+1)\Gamma(n)k_{1}!\Gamma(\ell+k_{2}+2)\ell_{1}!\Gamma(\ell_{1}+\kappa)(1-\rho^{2})^{k_{1}+\ell+k_{2}+2\ell_{1}+\kappa+1}} \\ &\int_{0}^{\infty} \lambda_{i}^{\ell+n+k_{1}+\ell_{1}+\kappa-1} \exp\left\{-\frac{(\nu_{i}+1)\lambda_{i}}{1-\rho^{2}}\right\} d\lambda_{i} \int_{0}^{\infty} \lambda_{j}^{\ell+k_{2}+\ell_{1}+\kappa} \exp\left\{-\frac{(\nu_{j}+1)\lambda_{j}}{1-\rho^{2}}\right\} d\lambda_{j} \\ &= \sum_{\ell_{1}=0}^{\infty} \frac{\Gamma(\ell+n+1)\Gamma(\ell+1)(\nu_{i}\nu_{j})^{\ell_{1}+\kappa}\rho^{2\ell_{1}}(1-\rho^{2})^{\ell+\kappa+n}}{\Gamma(\kappa)\ell_{1}!\Gamma(\ell_{1}+\kappa)(\nu_{i}+1)^{\ell+n+\ell_{1}+\kappa}(\nu_{j}+1)^{\ell+\ell_{1}+\kappa+1}} \frac{\Gamma(\ell+n+\ell_{1}+\kappa)\Gamma(\ell+\ell_{1}+\kappa+1)}{\Gamma(\ell+n+1)\Gamma(\ell+2)} \\ &\quad \times {}_{2}F_{1}\left(n,\ell+n+\ell_{1}+\kappa;\ell+n+1;\frac{\rho^{2}}{\nu_{i}+1}\right) {}_{2}F_{1}\left(1,\ell+\ell_{1}+\kappa+1;\ell+2;\frac{1}{\nu_{i}+1}\right) \\ &= \sum_{\ell_{1}=0}^{\infty} \frac{\Gamma(\ell+n+1)\Gamma(\ell+1)(\nu_{i}\nu_{j})^{\ell_{1}+\kappa}\rho^{2\ell_{1}}(\ell_{1}+\kappa)_{\ell+n}(\ell_{1}+1)_{\ell+\kappa}}{\Gamma(\kappa)\ell_{1}!\Gamma(\ell_{1}+\kappa)(\nu_{i}+1)^{\ell+n+\ell_{1}+\kappa}(\nu_{j}+1)^{\ell+\ell_{1}+\kappa+1}(1-\rho^{2})^{-(\ell+\kappa+n)}} \\ &\quad \times \left(\frac{1-\nu_{i}-\rho^{2}}{\nu_{i}+1}\right)^{-n} \left(\frac{\nu_{j}}{\nu_{j}+1}\right)^{-1} \widetilde{\mathcal{H}}\left(\frac{n,-1}{\ell+n+\ell_{1}},\frac{\rho^{2}}{\ell+2},-\frac{\rho^{2}}{1+\nu_{i}-\rho^{2}},-\frac{1}{\nu_{j}}\right) \end{split}$$

Thus, plugging in \mathcal{I}_1 and \mathcal{I}_2 on \tilde{p}_{n0} and doing some algebra, the result holds. Moreover, the case $m \ge 1, n = 0$ is analogous to $n \ge 1, m = 0$.

If
$$n = m \ge 1$$
:

$$\begin{split} \widetilde{p}_{nn} &= \int_{0}^{\infty} \int_{0}^{\infty} -\frac{1-\rho^2}{\Gamma(n)} \sum_{k=0}^{\infty} \frac{\rho^{2k}}{k! \Gamma(n+k)} \gamma\left(n+k, \frac{\lambda_i}{1-\rho^2}\right) \gamma\left(n+k, \frac{\lambda_j}{1-\rho^2}\right) f_{\Lambda_{ij}}(\lambda_i, \lambda_j) d\lambda_i d\lambda_j \\ &+ \int_{0}^{\infty} \int_{0}^{\infty} \frac{(1-\rho^2)^n \exp\{-\lambda_i\}}{\Gamma(n)\rho^{2n}} \\ &\times \sum_{k=0}^{\infty} \frac{1}{k! \Gamma(n+k)} \gamma\left(n+k, \frac{\rho^2 \lambda_i}{1-\rho^2}\right) \gamma\left(n+k, \frac{\lambda_j}{1-\rho^2}\right) f_{\Lambda_{ij}}(\lambda_i, \lambda_j) d\lambda_i d\lambda_j \end{split}$$

The double integral of the first term, noted by \mathcal{I}_1 , is obtained as follows:

$$\mathcal{I}_1 = \int_{0}^{\infty} \int_{0}^{\infty} \gamma \left(n+k, \frac{\lambda_i}{1-\rho^2} \right) \gamma \left(n+k, \frac{\lambda_j}{1-\rho^2} \right) \frac{\nu_i \nu_j}{(1-\rho^2)\Gamma(\kappa)} \left(\frac{\nu_i \nu_j \lambda_i \lambda_j}{\rho^2} \right)^{\frac{\kappa-1}{2}}$$

$$\begin{split} & \times \exp\left\{-\frac{\nu_{i}\lambda_{i}+\nu_{j}\lambda_{j}}{1-\rho^{2}}\right\}I_{\kappa-1}\left(\frac{2\rho}{1-\rho^{2}}\sqrt{\nu_{i}\nu_{j}\lambda_{i}\lambda_{j}}\right)d\lambda_{i}d\lambda_{j} \\ &=\sum_{\ell_{1}=0}^{\infty}\sum_{k_{2}=0}^{\infty}\sum_{k_{2}=0}^{\infty}\frac{\Gamma(n+k+1)\Gamma(n+k+k_{2}+1)\Gamma(\ell_{1}+\kappa)\Gamma(\kappa)\ell_{1}!(1-\rho^{2})^{k_{1}+k_{2}+2n+2k+2\ell_{1}+\kappa}}{\Gamma(n+k+1+1)\Gamma(n+k+k_{2}+1)\Gamma(\ell_{1}+\kappa)\Gamma(\kappa)\ell_{1}!(1-\rho^{2})^{k_{1}+k_{2}+2n+2k+2\ell_{1}+\kappa}} \\ & \times\int_{0}^{\infty}\lambda_{i}^{n+k+k_{1}+\ell_{1}+\kappa-1}\exp\left\{-\frac{(\nu_{i}+1)\lambda_{i}}{1-\rho^{2}}\right\}d\lambda_{i}\int_{0}^{\infty}\lambda_{j}^{n+k+k_{2}+\ell_{1}+\kappa-1}\exp\left\{-\frac{(\nu_{j}+1)\lambda_{j}}{1-\rho^{2}}\right\}d\lambda_{j} \\ &=\sum_{\ell_{1}=0}^{\infty}\frac{\Gamma(n+k)^{2}(\nu_{i}\nu_{j})^{\ell_{1}+\kappa}\rho^{2\ell_{1}}}{\Gamma(\ell_{1}+\kappa)\Gamma(\kappa)\ell_{1}!(1-\rho^{2})^{-\kappa}((\nu_{i}+1)(\nu_{j}+1))^{n+k+\ell_{1}+\kappa}}\left(\frac{\Gamma(n+k+\ell_{1}+\kappa)}{\Gamma(n+k+1)}\right)^{2} \\ & \times {}_{2}F_{1}\left(1,n+k+\ell_{1}+\kappa;n+k+1;\frac{1}{\nu_{i}+1}\right){}_{2}F_{1}\left(1,n+k+\ell_{1}+\kappa;n+k+1;\frac{1}{\nu_{j}+1}\right) \\ &=\sum_{\ell_{1}=0}^{\infty}\frac{\Gamma(n+k)^{2}(\nu_{i}\nu_{j})^{\ell_{1}+\kappa-1}\rho^{2\ell_{1}}(\ell_{1}+\kappa)_{n+k}(\kappa)_{n+\ell_{1}+\kappa}}{\ell_{1}!(1-\rho^{2})^{\kappa}((\nu_{i}+1)(\nu_{j}+1))^{n+k+\ell_{1}+\kappa}}\mathcal{H}\left(\frac{1,1-\ell_{1}-\kappa}{r+k+1},-\frac{1}{\nu_{i}},-\frac{1}{\nu_{j}}\right) \end{split}$$

The double integral of the second term, noted by \mathcal{I}_2 , is obtained as follows:

$$\begin{split} \mathcal{I}_{2} &= \int_{0}^{\infty} \int_{0}^{\infty} \exp\{-\lambda_{i}\}\gamma \left(n+k, \frac{\lambda_{i}}{1-\rho^{2}}\right)\gamma \left(n+k, \frac{\lambda_{j}}{1-\rho^{2}}\right) \frac{\nu_{i}\nu_{j}}{(1-\rho^{2})\Gamma(\kappa)} \left(\frac{\nu_{i}\nu_{j}\lambda_{i}\lambda_{j}}{\rho^{2}}\right)^{\frac{\kappa-1}{2}} \\ &\quad \times \exp\left\{-\frac{\nu_{i}\lambda_{i}+\nu_{j}\lambda_{j}}{1-\rho^{2}}\right\} I_{\kappa-1} \left(\frac{2\rho}{1-\rho^{2}}\sqrt{\nu_{i}\nu_{j}\lambda_{i}\lambda_{j}}\right) d\lambda_{i}d\lambda_{j} \\ &= \sum_{\ell_{1}=0}^{\infty} \sum_{k_{1}=0}^{\infty} \sum_{k_{2}=0}^{\infty} \frac{\Gamma(n+k)^{2}(\nu_{i}\nu_{j})^{\ell_{1}+\kappa}\rho^{2n+2k+2k_{1}+2\ell_{1}}}{\Gamma(n+k+k_{1}+1)\Gamma(n+k+k_{2}+1)\Gamma(\ell_{1}+\kappa)\Gamma(\kappa)\ell_{1}!(1-\rho^{2})^{k_{1}+k_{2}+2n+2k+2\ell_{1}+\kappa}} \\ &\quad \times \int_{0}^{\infty} \lambda_{i}^{n+k+k_{1}+\ell_{1}+\kappa-1} \exp\left\{-\frac{(\nu_{i}+1)\lambda_{i}}{1-\rho^{2}}\right\} d\lambda_{i} \int_{0}^{\infty} \lambda_{j}^{n+k+k_{2}+\ell_{1}+\kappa-1} \exp\left\{-\frac{(\nu_{j}+1)\lambda_{j}}{1-\rho^{2}}\right\} d\lambda_{j} \\ &= \sum_{\ell_{1}=0}^{\infty} \frac{\Gamma(n+k)^{2}(\nu_{i}\nu_{j})^{\ell_{1}+\kappa}\rho^{2n+2k+2\ell_{1}}}{\Gamma(\ell_{1}+\kappa)\Gamma(\kappa)\ell_{1}!(1-\rho^{2})^{-\kappa}((\nu_{i}+1)(\nu_{j}+1))^{n+k+\ell_{1}+\kappa}} \left(\frac{\Gamma(n+k+\ell_{1}+\kappa)}{\Gamma(n+k+1)}\right)^{2} \\ &\quad \times {}_{2}F_{1}\left(1,n+k+\ell_{1}+\kappa;n+k+1;\frac{\rho^{2}}{\nu_{i}+1}\right) {}_{2}F_{1}\left(1,n+k+\ell_{1}+\kappa;n+k+1;\frac{1}{\nu_{j}+1}\right) \\ &= \sum_{\ell_{1}=0}^{\infty} \frac{\Gamma(n+k)^{2}(\nu_{i}\nu_{j})^{\ell_{1}+\kappa}\rho^{2\ell_{1}}(\ell_{1}+\kappa)_{n+k}(\kappa)_{n+\ell_{1}+\kappa}}{\ell_{1}!(1-\rho^{2})^{\kappa}((\nu_{i}+1)(\nu_{j}+1))^{n+k+\ell_{1}+\kappa-1}} \end{split}$$

$$\times \frac{1}{(1-\nu_{i}-\rho^{2})\nu_{j}}\mathcal{H}\left(\begin{smallmatrix} 1 & , 1-\ell_{1}-\kappa \\ r+k+1 \end{smallmatrix}, -\frac{\rho^{2}}{1+\nu_{i}-\rho^{2}}, -\frac{1}{\nu_{j}} \right)$$

The double integral of the third term, noted by \mathcal{I}_3 , is calculated as the previous one. The result is:

$$\mathcal{I}_{3} = \sum_{\ell_{1}=0}^{\infty} \frac{\Gamma(n+k)^{2} (\nu_{i}\nu_{j})^{\ell_{1}+\kappa} \rho^{2\ell_{1}} (\ell_{1}+\kappa)_{n+k} (\kappa)_{n+\ell_{1}+\kappa}}{\ell_{1}! (1-\rho^{2})^{\kappa} ((\nu_{i}+1)(\nu_{j}+1))^{n+k+\ell_{1}+\kappa-1}} \\ \times \frac{1}{(1-\nu_{j}-\rho^{2})\nu_{i}} \mathcal{H} \left({}^{1}, {}^{1-\ell_{1}-\kappa}_{r+k+1}, -\frac{1}{\nu_{i}}, -\frac{\rho^{2}}{1+\nu_{j}-\rho^{2}} \right)$$

The double integral of the last term, noted by \mathcal{I}_4 , is obtained as follows:

$$\begin{split} \mathcal{I}_{4} &= \int_{0}^{\infty} \int_{0}^{\infty} \gamma \left(n + \ell + k + 1, \frac{\lambda_{i}}{1 - \rho^{2}} \right) \gamma \left(n + \ell + k + 1, \frac{\lambda_{j}}{1 - \rho^{2}} \right) \frac{\nu_{i}\nu_{j}}{(1 - \rho^{2})\Gamma(\kappa)} \\ & \times \left(\frac{\nu_{i}\nu_{j}\lambda_{i}\lambda_{j}}{\rho^{2}} \right)^{\frac{\kappa-1}{2}} \exp \left\{ -\frac{\nu_{i}\lambda_{i} + \nu_{j}\lambda_{j}}{1 - \rho^{2}} \right\} I_{\kappa-1} \left(\frac{2\rho}{1 - \rho^{2}} \sqrt{\nu_{i}\nu_{j}\lambda_{i}\lambda_{j}} \right) d\lambda_{i}d\lambda_{j} \\ &= \sum_{\ell_{1}=0}^{\infty} \sum_{k_{2}=0}^{\infty} \sum_{k_{2}=0}^{\infty} \frac{\Gamma(n + \ell + k + 1)^{2}(\nu_{i}\nu_{j})^{\ell_{1}+\kappa}\rho^{2\ell_{1}}(1 - \rho^{2})^{-(k_{1}+k_{2}+2n+2k+2\ell+2\ell_{1}+\kappa+2)}}{\Gamma(n + \ell + k + k_{1}+2)\Gamma(n + \ell + k + k_{1}+2)\Gamma(\ell_{1} + \kappa)\Gamma(\kappa)\ell_{1}!} \\ & \times \int_{0}^{\infty} \lambda_{i}^{n+\ell+k+k_{1}+\ell_{1}+\kappa} \exp \left\{ -\frac{(\nu_{i}+1)\lambda_{i}}{1 - \rho^{2}} \right\} d\lambda_{i} \int_{0}^{\infty} \lambda_{j}^{n+\ell+k+k_{2}+\ell_{1}+\kappa} \exp \left\{ -\frac{(\nu_{j}+1)\lambda_{j}}{1 - \rho^{2}} \right\} d\lambda_{j} \\ &= \sum_{\ell_{1}=0}^{\infty} \frac{\Gamma(n + \ell + k + 1)^{2}(\nu_{i}\nu_{j})^{\ell_{1}+\kappa}\rho^{2\ell_{1}}}{\Gamma(\ell_{1} + \kappa)\Gamma(\kappa)\ell_{1}!(1 - \rho^{2})^{-\kappa}((\nu_{i}+1)(\nu_{j}+1))^{n+\ell+k+\ell_{1}+\kappa+1}} \left(\frac{\Gamma(n + \ell + k + \ell_{1} + \kappa + 1)}{\Gamma(n + \ell + k + 2)} \right)^{2} \\ & \times {}_{2}F_{1} \left(1, n + \ell + k + \ell_{1} + \kappa + 1; n + \ell + k + 2; \frac{1}{\nu_{i}+1} \right) \\ & \times {}_{2}F_{1} \left(1, n + \ell + k + \ell_{1} + \kappa + 1; n + \ell + k + 2; \frac{1}{\nu_{j}+1} \right) \\ &= \sum_{\ell_{1}=0}^{\infty} \frac{\Gamma(n + \ell + k + 1)^{2}(\nu_{i}\nu_{j})^{\ell_{1}+\kappa-1}\rho^{2\ell_{1}}(\ell_{1} + \kappa)_{n+\ell+k+\ell_{1}+\kappa}}{\ell_{1}!(1 - \rho^{2})^{\kappa}((\nu_{i}+1)(\nu_{j}+1))^{n+\ell+k+\ell_{1}+\kappa}}} \mathcal{H} \left(\frac{1}{1, \frac{1-\ell_{1}-\kappa}{r+k+\ell_{2}}, -\frac{1}{\nu_{i}}, -\frac{1}{\nu_{j}}} \right) \end{split}$$

Thus, plugging in $\mathcal{I}_1, \mathcal{I}_2, \mathcal{I}_3$ and \mathcal{I}_4 on \tilde{p}_{nn} and doing some algebra, the result holds.

If $n \ge 2$, $m \ge 1$ with n > m:

Let n = m + r with r > 0, then the bivariate distribution is

$$\begin{split} \widetilde{p}_{nm} &= \int_{0}^{\infty} \int_{0}^{\infty} \lambda_{i}^{m+r} \exp\left\{-\frac{\lambda_{i}}{1-\rho^{2}}\right\} \sum_{\ell=0}^{\infty} \frac{\Gamma(m+\ell)}{\ell!\Gamma(m)} \left(\frac{\rho^{2}\lambda_{i}}{1-\rho^{2}}\right)^{\ell} \frac{1}{\Gamma(m+\ell+r+1)\Gamma(m+\ell)} \\ &\quad \times {}_{1}F_{1}\left(r+1;m+\ell+r+1,\frac{\rho^{2}\lambda_{i}}{1-\rho^{2}}\right)\gamma\left(m+\ell,\frac{\lambda_{j}}{1-\rho^{2}}\right)f_{\Lambda_{ij}}(\lambda_{i},\lambda_{j})d\lambda_{i}d\lambda_{j} \\ &\quad -\int_{0}^{\infty} \int_{0}^{\infty} \lambda_{i}^{m+r} \exp\left\{-\frac{\lambda_{i}}{1-\rho^{2}}\right\}\sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \frac{\Gamma(m+\ell)}{\ell!\Gamma(m)} \left(\frac{\rho^{2}\lambda_{i}}{1-\rho^{2}}\right)^{k+\ell} \\ &\quad \times \frac{1}{\Gamma(m+\ell+r+k+1)\Gamma(m+\ell+k+1)} \\ &\quad \times {}_{1}F_{1}\left(r;m+\ell+r+k+1,\frac{\rho^{2}\lambda_{i}}{1-\rho^{2}}\right)\gamma\left(m+\ell+k+1,\frac{\lambda_{j}}{1-\rho^{2}}\right)f_{\Lambda_{ij}}(\lambda_{i},\lambda_{j})d\lambda_{i}d\lambda_{j} \\ &= \sum_{\ell=0}^{\infty} \frac{\Gamma(m+\ell)}{\ell!\Gamma(m)} \left(\frac{\rho^{2}}{1-\rho^{2}}\right)^{\ell} \frac{1}{\Gamma(m+\ell+r+1)\Gamma(m+\ell)} \\ &\quad \times \int_{0}^{\infty} \int_{0}^{\infty} \lambda_{i}^{m+r+\ell} \exp\left\{-\frac{\lambda_{i}}{1-\rho^{2}}\right)I_{1}F_{1}\left(r+1;m+\ell+r+1,\frac{\rho^{2}\lambda_{i}}{1-\rho^{2}}\right) \\ &\quad \times \gamma\left(m+\ell,\frac{\lambda_{j}}{1-\rho^{2}}\right)f_{\Lambda_{ij}}(\lambda_{i},\lambda_{j})d\lambda_{i}d\lambda_{j} \\ &- \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \frac{\Gamma(m+\ell)}{\ell!\Gamma(m)} \left(\frac{\rho^{2}}{1-\rho^{2}}\right)^{k+\ell} \frac{1}{\Gamma(m+\ell+r+k+1)\Gamma(m+\ell+k+1)} \\ &\quad \times \int_{0}^{\infty} \int_{0}^{\infty} \lambda_{i}^{m+rk+\ell} \exp\left\{-\frac{\lambda_{i}}{1-\rho^{2}}\right\} {}_{1}F_{1}\left(r;m+\ell+r+k+1,\frac{\rho^{2}\lambda_{i}}{1-\rho^{2}}\right) \\ &\quad \times \gamma\left(m+\ell+k+1,\frac{\lambda_{j}}{1-\rho^{2}}\right)f_{\Lambda_{ij}}(\lambda_{i},\lambda_{j})d\lambda_{i}d\lambda_{j} \end{aligned}$$

The double integral of the first term, noted by \mathcal{I}_1 , is obtained as follows:

$$\mathcal{I}_1 = \int_0^\infty \int_0^\infty \lambda_i^{m+r+\ell} \exp\left\{-\frac{\lambda_i}{1-\rho^2}\right\} {}_1F_1\left(r+1;m+\ell+r+1,\frac{\rho^2\lambda_i}{1-\rho^2}\right)\gamma\left(m+\ell,\frac{\lambda_j}{1-\rho^2}\right)$$

$$\begin{split} & \times \frac{\nu_{i}\nu_{j}}{(1-\rho^{2})\Gamma(\kappa)} \left(\frac{\nu_{i}\nu_{j}\lambda_{i}\lambda_{j}}{\rho^{2}}\right)^{\frac{n-1}{2}} \exp\left\{-\frac{\nu_{i}\lambda_{i}+\nu_{j}\lambda_{j}}{1-\rho^{2}}\right\} I_{\kappa-1}\left(\frac{2\rho}{1-\rho^{2}}\sqrt{\nu_{i}\nu_{j}\lambda_{i}\lambda_{j}}\right) d\lambda_{i}d\lambda_{j} \\ & = \sum_{\ell_{1}=0}^{\infty}\sum_{k_{2}=0}^{\infty}\sum_{k_{2}=0}^{\infty}\frac{\Gamma(r+k_{1}+1)\Gamma(m+\ell+r+1)\Gamma(m+\ell)(\nu_{i}\nu_{j})^{\ell_{1}+\kappa}}{\Gamma(\kappa)\Gamma(m+\ell+r+k_{1}+1)\Gamma(r+1)k_{1}!} \\ & \times \frac{\rho^{2k_{1}+2\ell_{1}}(1-\rho^{2})^{-(k_{1}+k_{2}+m+\ell+2\ell_{1}+\kappa)}}{\Gamma(m+\ell+k_{2}+1)\ell_{1}!\Gamma(\ell_{1}+\kappa)} \\ & \int_{0}^{\infty}\lambda_{i}^{m+r+\ell+k_{1}+\ell_{1}+\kappa-1} \exp\left\{-\frac{(\nu_{i}+1)\lambda_{i}}{1-\rho^{2}}\right\} d\lambda_{i}\int_{0}^{\infty}\lambda_{j}^{m+\ell+k_{2}+\ell_{1}+\kappa} \exp\left\{-\frac{(\nu_{j}+1)\lambda_{j}}{1-\rho^{2}}\right\} d\lambda_{j} \\ & = \sum_{\ell_{1}=0}^{\infty}\frac{\Gamma(m+\ell+r+1)\Gamma(m+\ell)(\nu_{i}\nu_{j})^{\ell_{1}+\kappa}\rho^{2\ell_{1}}(1-\rho^{2})^{m+\ell+\kappa+r}}{\Gamma(\kappa)\ell_{1}!\Gamma(\ell_{1}+\kappa)(\nu_{i}+1)^{m+r+\ell+\ell_{1}+\kappa}(\nu_{j}+1)^{m+\ell+\ell_{1}+\kappa}} \\ & \times \frac{\Gamma(m+r+\ell+\ell_{1}+\kappa)\Gamma(m+\ell+\ell_{1}+\kappa)}{\Gamma(m+\ell+r+1)\Gamma(m+\ell+1)} \\ & \times 2F_{1}\left(r+1,m+r+\ell+\ell_{1}+\kappa;m+\ell+r+1;\frac{1}{\nu_{i}+1}\right) \\ & \times 2F_{1}\left(1,m+\ell+\ell_{1}+\kappa;m+\ell+1;\frac{1}{\nu_{i}+1}\right) \\ & = \sum_{\ell_{1}=0}^{\infty}\frac{\Gamma(m+\ell+r+1)\Gamma(m+\ell)(\nu_{i}\nu_{j})^{\ell_{1}+\kappa}\rho^{2\ell_{1}}(\ell+\kappa)_{m+r+\ell}(\kappa)_{m+\ell+\ell_{1}}}{\Gamma(\kappa)\ell_{1}!\Gamma(\ell_{1}+\kappa)((\nu_{i}+1)(\nu_{j}+1))^{m+\ell+\ell_{1}+\kappa-1}(1-\rho^{2})^{-(m+\ell+\kappa+r)}} \\ & \times \left(\frac{1}{1-\nu_{i}-\rho^{2}}\right)^{-r-1}\left(\frac{1}{\nu_{j}}\right)^{-1}\widetilde{\mathcal{H}}\left(\frac{r+1, \ 1}{1-\ell_{i}-\kappa}, \ 1}{1-\ell_{i}-\kappa}, \ 1}{1-\nu_{i}-\rho^{2}}, -\frac{1}{\nu_{j}}\right) \\ & \times \left(\frac{1}{1-\nu_{i}-\rho^{2}}\right)^{-n+m-1}\left(\frac{1}{\nu_{j}}\right)^{-1}\widetilde{\mathcal{H}}\left(\frac{n-m+1, \ 1}{1-\ell_{i}-\kappa}, \ 1}{1-\ell_{i}-\kappa}, \ 1}{1-\nu_{i}-\rho^{2}}, -\frac{1}{\nu_{j}}\right) \end{split}$$

The double integral of the second term, noted by \mathcal{I}_2 , is obtained as follows:

$$\mathcal{I}_2 = \int_0^\infty \int_0^\infty \lambda_i^{m+rk+\ell} \exp\left\{-\frac{\lambda_i}{1-\rho^2}\right\} {}_1F_1\left(r;m+\ell+r+k+1,\frac{\rho^2\lambda_i}{1-\rho^2}\right)\gamma\left(m+\ell+k+1,\frac{\lambda_j}{1-\rho^2}\right)$$

$$\begin{split} & \times \frac{\nu_{i}\nu_{j}}{(1-\rho^{2})\Gamma(\kappa)} \left(\frac{\nu_{i}\nu_{j}\lambda_{i}\lambda_{j}}{\rho^{2}}\right)^{\frac{\kappa-1}{2}} \exp\left\{-\frac{\nu_{i}\lambda_{i}+\nu_{j}\lambda_{j}}{1-\rho^{2}}\right\} I_{\kappa-1}\left(\frac{2\rho}{1-\rho^{2}}\sqrt{\nu_{i}\nu_{j}\lambda_{i}\lambda_{j}}\right) d\lambda_{i}d\lambda_{j} \\ &= \sum_{\ell_{1}=0}^{\infty}\sum_{k_{1}=0}^{\infty}\sum_{k_{2}=0}^{\infty} \frac{\Gamma(r+k_{1})\Gamma(m+r+\ell+k+1)\Gamma(m+\ell+k+1)(\nu_{i}\nu_{j})^{\ell_{1}+\kappa}}{\Gamma(\kappa)\Gamma(m+r+\ell+k+k_{1}+1)\Gamma(r)k_{1}!} \\ & \times \frac{\rho^{2k_{1}+2\ell_{1}}(1-\rho^{2})^{-(k_{1}+k_{2}+m+\ell+2\ell_{1}+\kappa+1+k)}}{\Gamma(m+\ell+k+k_{2}+2\ell_{1}+\kappa)} \\ & \int_{0}^{\infty}\lambda_{i}^{m+r+k+\ell+k_{1}+\ell_{1}+\kappa-1} \exp\left\{-\frac{(\nu_{i}+1)\lambda_{i}}{1-\rho^{2}}\right\} d\lambda_{i}\int_{0}^{\infty}\lambda_{j}^{m+\ell+k+k_{2}+\ell_{1}+\kappa} \exp\left\{-\frac{(\nu_{j}+1)\lambda_{j}}{1-\rho^{2}}\right\} d\lambda_{j} \\ &= \sum_{\ell_{1}=0}^{\infty}\frac{\Gamma(m+r+\ell+k+1)\Gamma(m+\ell+k+1)(\nu_{i}+\ell_{j})^{\ell_{1}+\kappa}\rho^{2\ell_{1}}(1-\rho^{2})^{m+\ell+k+r+\kappa}}{\Gamma(\kappa)\ell_{1}!\Gamma(\ell_{1}+\kappa)(\nu_{i}+1)^{m+r+k+\ell+\ell_{1}+\kappa}(\nu_{j}+1)^{m+\ell+k+\ell_{1}+\kappa+1}} \\ & \times \frac{\Gamma(m+r+k+\ell+\ell_{1}+\kappa)\Gamma(m+\ell+k+\ell_{1}+\kappa+1)}{\Gamma(m+\ell+k+1)\Gamma(m+\ell+k+2)} \\ & \times 2F_{1}\left(r,m+r+k+\ell+\ell_{1}+\kappa,m+r+\ell+k+1;\frac{\rho^{2}}{\nu_{i}+1}\right) \\ & \times 2F_{1}\left(1,m+\ell+k+\ell_{1}+\kappa+1;m+\ell+k+2;\frac{1}{\nu_{i}+1}\right) \\ &= \sum_{\ell_{1}=0}^{\infty}\frac{\Gamma(m+r+\ell+k+1)\Gamma(m+\ell+k+1)(\nu_{i}\nu_{j})^{\ell_{1}+\kappa}\rho^{2\ell_{1}}(\ell+\kappa)_{m+r+k+\ell}(\kappa)_{m+k+\ell+\ell_{1}+1}}{\Gamma(\kappa)\ell_{1}!\Gamma(\ell_{1}+\kappa)((\nu_{i}+1)(\nu_{j}+1))^{m+\ell+k+\ell_{1}+\kappa}(1-\rho^{2})^{-(m+\ell+k+\ell_{1}+\kappa)}} \\ & \times \left(\frac{1}{1-\nu_{i}-\rho^{2}}\right)^{-r}\left(\frac{1}{\nu_{j}}\right)^{-1}\widetilde{\mathcal{H}}\left(\frac{n-m}{\ell+\kappa-1},\frac{1}{\ell+\kappa+1},\frac{1}{\ell+m+k+2},-\frac{\rho^{2}}{1+\nu_{i}-\rho^{2}},-\frac{1}{\nu_{j}}\right) \\ & \times \left(\frac{1}{1-\nu_{i}-\rho^{2}}\right)^{-r}\left(\frac{1}{\nu_{j}}\right)^{-1}\widetilde{\mathcal{H}}\left(\frac{n-m}{\ell+\kappa+1},\frac{1}{\ell+m+k+2},-\frac{\rho^{2}}{1+\nu_{i}-\rho^{2}},-\frac{1}{\nu_{j}}\right) \end{split}$$

Thus, plugging in \mathcal{I}_1 and \mathcal{I}_2 on \tilde{p}_{nm} and doing some algebra, the result holds. Moreover, the case $m \ge 2, n \ge 1$ with m > n is analogous to $n \ge 2, m \ge 1$ with n > m.

Bibliography

- Abramowitz, M. and I. Stegun (1965). *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables.* Applied mathematics series. Dover Publications.
- Aitchison, J. and C. H. Ho (1989). The multivariate poisson-log normal distribution. *Biometrika* 76(4), 643–653.
- Alegría, A., S. Caro, M. Bevilacqua, E. Porcu, and J. Clarke (2017). Estimating covariance functions of multivariate skew-gaussian random fields on the sphere. *Spatial Statistics* 22, 388 – 402.
- Bai, Y., J. Kang, and P. Song (2014). Efficient pairwise composite likelihood estimation for spatial-clustered data. *Biometrics* 7(3), 661–670.
- Banerjee, S., B. P. Carlin, and A. E. Gelfand (2014). *Hierarchical Modeling and Analysis for Spatial Data*. Boca Raton: FL: Chapman & Hall/CRC Press.
- Bennett, L., P. English, and R. McCain (1940). A study of deer populations by use of pellet-group counts. *The Journal of Wildlife Management 4*, 398–403.
- Berg, C. and E. Porcu (2017). From schoenberg coefficients to schoenberg functions. *Constructive Approximation* 45(2), 217–241.
- Bevilacqua, M., C. Caamaño, and C. Gaetan (2020). On modelling positive continuous data with spatio-temporal dependence. *Environmetrics*. To appear.

- Bevilacqua, M., C. Caamaño, R. Arellano-Valle, and V. Morales-Oñate (2020). Non-gaussian geostatistical modeling using (skew) t processes. *Scandinavian Journal* of Statistics, 1–34.
- Bevilacqua, M., T. Faouzi, R. Furrer, and E. Porcu (2019). Estimation and prediction using generalized wendland functions under fixed domain asymptotics. *The Annals of Statistics* 47, 828–856.
- Bevilacqua, M. and C. Gaetan (2015). Comparing composite likelihood methods based on pairs for spatial gaussian random fields. *Statistics and Computing* 25, 877–892.
- Bevilacqua, M., C. Gaetan, J. Mateu, and E. Porcu (2012). Estimating space and space-time covariance functions for large data sets: a weighted composite likelihood approach. *Journal of the American Statistical Association 107*, 268–280.
- Bevilacqua, M., V. Morales-Oñate, and C. Caamaño-Carrillo (2019). Geomodels: A package for geostatistical gaussian and non gaussian data analysis. https://vmoprojs.github.io/GeoModels-page/. R package version 1.0.3-4.
- Breslow, N. E. and D. G. Clayton (1993). Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association* 88(421), 9–25.
- Brychkov, Y. (2008, 01). *Handbook of Special Functions: Derivatives, Integrals, Series and Other Formulas.* A Chapman & Hall book. Taylor & Francis.
- Buckland, S., D. R. Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers, and L. Thomas (2001). *Introduction to Distance Sampling-Estimating Abundance of Biological Populations*. New York, NY., Oxford University Press.
- Cardina, J., D. H. Sparrow, and E. L. McCoy (1995). Analysis of spatial distribution of common lambsquarters (chenopodium album) in no-till soybean (glycine max). Weed Science 43(2), 258–268.

- Christensen, O. and Waagepetersen (2002). Bayesian prediction of spatial count data using generalized linear mixed models. *Biometrics* 58, 280–286.
- Cox, D. (1970). Renewal Theory. London: Methuen & Co.
- Cressie, N. and H. Huang (1999). Classes of nonseparable, spatiotemporal stationary covariance functions. *Journal of the American Statistical Association* 94, 1330–1340.
- Cressie, N. and C. Wikle (2011). *Statistics for Spatio-Temporal Data*. New Jersey: Wiley Series in Probability and Statistics. Wiley.
- De Iaco, S., D. E. Myers, and T. Posa (2002). Nonseparable space-time covariance models: some parametric families. *Mathematical Geology 34*, 23–42.
- De Oliveira, V. (2006). On optimal point and block prediction in log-gaussian random fields. *Scandinavian Journal of Statistics 33*, 523–540.
- De Oliveira, V. (2013). Hierarchical poisson models for spatial count data. *Journal of Multivariate Analysis 122*, 93–408.
- De Oliveira, V. (2014). Poisson kriging: A closer investigation. Spatial Statistics 7, 1-20.
- Diggle, P. and E. Giorgi (2019, 03). *Model-based Geostatistics for Global Public Health: Methods and Applications*. Chapman & Hall/CRC Interdisciplinary Statistics. Chapman and Hall/CRC Press.
- Diggle, P., J. Tawn, and R. Moyeed (1998). Model-based geostatistics. *Journal of the Royal Statistical Society, Series C* 47, 299–350.
- Diggle, P. J. and P. J. Ribeiro (2007). Model-based Geostatistics. New York: Springer.
- Donald, W. W. (1994). Geostatistics for mapping weeds, with a canada thistle (cirsium arvense) patch as a case study. *Weed Science* 42(4), 648–657.
- Efron, B. and R. Tibshirani (1986). Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science 1*(1), 54–75.

- Etten, R. V. and C. Bennett (1965). Some sources of error in using pellet-group counts for censusing deer. *The Journal of Wildlife Management 29*, 723–729.
- Gelfand, A. E. and E. M. Schliep (2016). Spatial statistics and Gaussian processes: A beautiful marriage. *Spatial Statistics* 18, 86–104.
- Genest, C. and J. Neslehova (2007). A primer of copulas for count data. *ASTIN Bulletin 37*, 475–515.
- Gneiting, T. (1999). Correlation functions for atmospheric data analysis. *Quarterly Journal of the Royal Meteorological Society* 125, 2449–2464.
- Gneiting, T. (2002a). Compactly supported correlation functions. *Journal of Multivariate Analysis 83*, 493–508.
- Gneiting, T. (2002b). Nonseparable, stationary covariance functions for space-time data. *Journal of the American Statistical Association* 97(458), 590–600.
- Gneiting, T. (2002c). Stationary covariance functions for space-time data. *Journal of the American Statistical Association* 97, 590–600.
- Gneiting, T. (2013, 09). Strictly and non-strictly positive definite functions on spheres. *Bernoulli 19*(4), 1327–1349.
- Gough, B. (2009). GNU Scientific Library Reference Manual. Network Theory Ltd.
- Gouriéroux, C., A. Monfort, and E. Renault (2017). Consistent pseudo-maximum likelihood estimators. *Annals of Economics and Statistics*, 187–218.
- Gradshteyn, I. and I. Ryzhik (2007). Table of integrals, series, and products. In A. Jeffrey and D. Zwillinger (Eds.), *Table of Integrals, Series, and Products* (7 ed.). New York: Academic Press.
- Gradshteyn, I. and I. Ryzhik (2014). Special functions. In D. Zwillinger, V. Moll, I. Gradshteyn, and I. Ryzhik (Eds.), *Table of Integrals, Series, and Products* (Eighth ed.)., pp. 1014–1059. Boston: Academic Press.

- Guillot, G., N. Loren, and M. Rudemo (2009). Spatial prediction of weed intensities from exact count data and image-based estimates. *Journal of the Royal Statistical Society: Series C* 58, 525–542.
- Han, Z. and V. De Oliveira (2016). On the correlation structure of gaussian copula models for geostatistical count data. *Australian & New Zealand Journal of Statistics* 58, 47–69.
- Han, Z. and V. D. Oliveira (2018). gckrig: An R Package for the Analysis of Geostatistical Count Data Using Gaussian Copulas. *Journal of Statistical Software* 87(13), 1–32.
- Heagerty, P. and S. Lele (1998). A composite likelihood approach to binary spatial data. *Journal of the American Statistical Association 93*, 1099–1111.
- Huang, H.-C. and N.-J. Hsu (2004). Modeling transport effects on ground-level ozone using a non-stationary space–time model. *Environmetrics* 15(3), 251–268.
- Hunter, J. (1974). Renewal theory in two dimensions: basic results. *Advances in Applied Probability* 6, 376–391.
- Jeong, J. and M. Jun (2015). A class of Matérn-like covariance functions for smooth processes on a sphere. *Spatial Statistics 11*, 1–18.
- Joe, H. (2014). *Dependence Modeling with Copulas*. Boca Raton, FL: Chapman and Hall/CRC.
- Joe, H. and Y. Lee (2009). On weighting of bivariate margins in pairwise likelihood. *Journal of Multivariate Analysis 100*, 670–685.
- Johnson, G. A., D. A. Mortensen, and C. A. Gotway (1996). Spatial and temporal analysis of weed seedling populations using geostatistics. *Weed Science* 44(3), 704–710.
- Kazianka, H. and J. Pilz (2010). Copula-based geostatistical modeling of continuous and discrete data including covariates. *Stochastic Environmental Research and Risk Assessment* 24, 661–673.

- Kibble, W. F. (1941). A two-variate gamma type distribution. *Sankhyā: The Indian Journal of Statistics 5*, 137–150.
- Krebs, C., R. Boonstra, V. Nams, M. O. M, K. Hodges, and S. Boutin (2001). Estimating snowshoe hare population density from pellet plots: A further evaluation. *Canadian Journal of Zoology* 79, 1–4.
- Krishnaiah, P. and M. Rao (1961). Remarks on a multivariate Gamma distribution. *The American Mathematical Monthly* 68(4), 342–346.
- Krishnamoorthy, A. S. and M. Parthasarathy (1951). A Multivariate Gamma-Type Distribution. *The Annals of Mathematical Statistics* 22(4), 549 557.
- Lantuèjoul, C. (2002). Geostatistical Simulation. Berlin: Springer.
- Lee, Y., A. Moudud, M. Noh, L. Rönnegård, and A. Skarin (2016). Spatial modeling of data with excessive zeros applied to reindeer pellet-group counts. *Ecology and Evolution* 6, 7047 7056.
- Lindgren, F., H. Rue, , and J. Lindstrom (2011). An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. *Journal of the Royal Statistical Society, Series B* 73(4), 423–498.
- Lindsay, B. (1988). Composite likelihood methods. *Contemporary Mathematics* 80, 221–239.
- Lindsay, B. G. (1995). Mixture models: Theory, geometry and applications. *NSF-CBMS Regional Conference Series in Probability and Statistics 5*, i–163.
- Ma, C. (2005). Spatio-temporal variograms and covariance models. *Advances in Applied Probability 37*, 706–725.
- Mainardi, F., R. Gorenflo, and A. Vivoli (2007). Beyond the Poisson renewal process: A tutorial survey. *Journal of Computational and Applied Mathematics* 205(2), 725 735.

- Martins, T. G., D. Simpson, F. Lindgren, and H. Rue (2013). Bayesian computing with INLA: new features. *Computational Statistics and Data Analysis* 67, 68–83.
- Masarotto, G. and C. Varin (2012). Gaussian copula marginal regression. *Electronic Journal of Statistics* 6, 1517–1549.
- Masarotto, G. and C. Varin (2017). Gaussian copula regression in R. *Journal of Statistical Software* 77(8), 1–26.
- Masuda, H. (2013). Convergence of Gaussian quasi-likelihood random fields for ergodic Lévy driven SDE observed at high frequency. *The Annals of Statistics* 41(3), 1593–1641.
- Matèrn, B. (1986). Spatial Variation: Stochastic Models and their Applications to Some Problems in Forest Surveys and Other Sampling Investigations (2nd ed.). Heidelberg: Springer.
- Matheron, G. (1965). Les Variables Régionalisées et Leur Estimation. Paris: Masson.
- Mayle, B., A. Peace, and R. Gill (1999). How many deer? a field guide to estimating deer population size. Technical Report Forestry Commission Field book 18, Forestry Commission, Edinburgh.
- McLachlan, G. and D. Peel (2004). *Finite Mixture Models*. Wiley Series in Probability and Statistics. Hoboken: Wiley.
- Monestiez, P., L. Dubroca, E. Bonnin, and J.-P. D. C. Guinet (2006). Geostatistical modelling of spatial distribution of balaenoptera physalus in the northwestern mediterranean sea from sparse count data and heterogeneous. *Ecological Modelling 193*, 615–628.
- Morgan, C. J. (2005). Analysing spatial data via geostatistical methods. Master's thesis, University of the Witwatersrand, Johannesburg.

- Ng, E. and M. Geller (1970). On some indefinite integrals of confluent hypergeometric functions. *Journal of Research of the Notional Bureau of Standards-B. Mathematica Sciences* 74B(2), 85–98.
- Olver, F., N. I. of Standards, T. (U.S.), D. Lozier, R. Boisvert, and C. Clark (2010). *NIST Handbook of Mathematical Functions Hardback and CD-ROM*. Cambridge University Press.
- Porcu, E., M. Bevilacqua, and M. G. Genton (2016). Spatio-temporal covariance and cross-covariance functions of the great circle distance on a sphere. *Journal of the American Statistical Association 111*(514), 888–898.
- Porcu, E., P. Gregori, and J. Mateu (2006). Nonseparable stationary anisotropic space time covariance functions. *Stochastic Environmental Research and Risk Assessment 21*, 113–122.
- Quiroz, Z. C. and M. O. Prates (2018). Bayesian spatio-temporal modelling of anchovy abundance through the spde approach. *Spatial Statistics* 28, 236–256. One world, one health.
- R Core Team (2020). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Ross, S. (2008). Stochastic Processes (2 ed.). Wiley. Wiley India Pvt. Limited.
- Rousset, F. and J.-B. Ferdy (2014). Testing environmental and genetic effects in the presence of spatial autocorrelation. *Ecography* 37(8), 781–790.
- Royen, T. (2004). Multivariate Gamma distributions II. In *Encyclopedia of Statistical Sciences*, pp. 419–425. New York: John Wiley & Sons.
- Rue, H. and L. Held (2005). Gaussian Markov Random Fields: Theory and Applications. London: Chapman & Hall.

- Rue, H., S. Martino, and N. Chopin (2009). Approximate Bayesian inference for latent Gaussian models using integrated nested Laplace approximations (with discussion). *Journal of the Royal Statistical Society, Series B* 71, 319–392.
- Sampson, P. D. and P. Guttorp (1992). Nonparametric estimation of nonstationary spatial covariance structure. *Journal of the American Statistical Association* 87(417), 108–119.
- Schlather, M. (2010). Some covariance models based on normal scale mixtures. *Bernoulli* 16(3), 780 – 797.
- Schoenberg, I. J. (1938). Metric spaces and completely monotone functions. Annals of Mathematics 39(4), 811–841.
- Sivertsen, T. R., B. Åhman, S. M. J. G. Steyaert, L. Rönnegård, J. Frank, P. Segerström, O.-G. Støen, and A. Skarin (2016). Reindeer habitat selection under the risk of brown bear predation during calving season. *Ecosphere* 7(11), e01583.
- Stein, M. (1999). *Interpolation of Spatial Data. Some Theory of Kriging*. New York: Springer-Verlag.
- Stein, M. (2005). Space-time covariance functions. Journal of the American Statistical Association 100, 310–321.
- Storvik, G., A. Frigessi, and D. Hirst (2002). Stationary space-time Gaussian fields and their time autoregressive representation. *Stochastic Modelling* 2, 139–161.
- Titterington, D., A. Smith, and U. Makov (1985). Statistical Analysis of Finite Mixture Distributions. Applied section. John Wiley & Sons.
- Trivedi, P. and D. Zimmer (2017). A note on identification of bivariate copulas for discrete count data. *Econometrics* 5(1), 1–11.
- Varin, C., N. Reid, and D. Firth (2011). An overview of composite likelihood methods. *Statistica Sinica* 21, 5–42.

- Vere-Jones, D. (1997). Alpha-permanents and their applications to multivariate gamma, negative binomial and ordinary binomial distributions. New Zealand Journal of Mathematics 26, 125–149.
- Virtanen, P., R. Gommers, T. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, B. Jonathan, van der Walt, S. J., M. Brett, J. Wilson, K. Jarrod Millman, N. Mayorov, and et al. (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods* 17, 261–272.
- Wang, X., M. Chen, R. Kou, and D. Dey (2015). Bayesian spatial-temporal modeling of ecological zero-inflated count data. *Statistica Sinica* 25, 189–204.
- Webster, R. (2010, 07). *Weeds, Worms and Geostatistics*, pp. 221–241. Springer Netherlands.
- Yaglom, A. M. (1987). Correlation Theory of Stationary and Related Random Functions. Volume I: Basic Results. New York: Springer.
- Zimmerman, D. L. (1993). Another look at anisotropy in geostatistics. *Mathematical Geology* 25(4), 453–470.