

**NON-FULLY SYMMETRIC SPACE-TIME MATERN-CAUCHY  
CORRELATION FUNCTIONS**

by

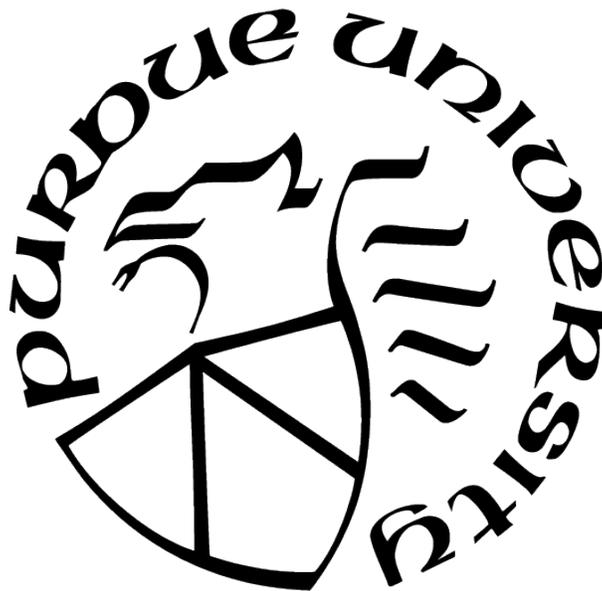
**Zizhuang Wu**

**A Dissertation**

*Submitted to the Faculty of Purdue University*

*In Partial Fulfillment of the Requirements for the degree of*

**Doctor of Philosophy**



Department of Statistics

West Lafayette, Indiana

May 2021

**THE PURDUE UNIVERSITY GRADUATE SCHOOL  
STATEMENT OF COMMITTEE APPROVAL**

**Dr. Tonglin Zhang, Co-chair**

Department of Statistics

**Dr. Hao Zhang, Co-chair**

Department of Statistics

**Dr. Jun Xie**

Department of Statistics

**Dr. Qifan Song**

Department of Statistics

**Approved by:**

Dr. Jun Xie

To my family

## ACKNOWLEDGMENTS

I would like to thank my advisors Dr. Tonglin Zhang and Dr. Hao Zhang for their invaluable guidance and support to my research, and for their passion and dedication of research which inspire me to achieve more in my Ph.D. career. I would also like to thank my thesis committee members, Dr. Jun Xie and Dr. Qifan Song for their precious advice to my research work and defense, and Dr. Kiseop Lee for his kindly advice and care throughout my Ph.D. life. And I really appreciate the faculty and staff in the Department of Statistics of Purdue University for their great support and help to me. I want to thank my fellow colleagues and friends in Statistics Department, including Jincheng Bai, Donglai Chen, Jinyuan Chen, Yao Chen, Eric Gerber, Botao Hao, Cheng Li, Jiapeng Liu, Ryan Murphy, Yixuan Qiu, Min Ren, Yuying Song, Hui Sun, Qi Wang, Zhanyu Wang, Wutao Wei, Jiasen Yang, Boqian Zhang, Rongrong Zhang, Yumin Zhang, Zhou Shen and so many that I can't list them all here. I enjoy and appreciate all kinds of cooperation with you and the happy time we had together. Those are all my important memories of years at Purdue. Last but not least, I would like to thank my parents for their greatest love and support.

# TABLE OF CONTENTS

LIST OF FIGURES . . . . .	7
ABBREVIATIONS . . . . .	10
ABSTRACT . . . . .	11
1 Introduction . . . . .	12
1.1 Geostatistics and correlation functions . . . . .	12
1.2 Spatio-temporal processes . . . . .	16
1.3 Our contribution . . . . .	19
2 Spectral methods . . . . .	21
2.1 Bochner’s Representation . . . . .	21
2.2 Our methodology . . . . .	23
3 Correlation Models . . . . .	26
3.1 Matérn Models . . . . .	26
3.2 Cauchy Models . . . . .	27
3.3 Cressie and Huang’s Model . . . . .	30
3.4 Gneiting’s Model . . . . .	31
3.5 Stein’s Model . . . . .	32
3.6 Space-Time Matérn Models . . . . .	34
4 Matérn-Cauchy Models . . . . .	53
4.1 Construction of Non-Fully Symmetric Space-Time Matérn-Cauchy Models . . . . .	53
4.2 Computation . . . . .	56
4.3 The Case When $d = 2$ . . . . .	61
4.4 Discussion . . . . .	66
5 Numerical Results . . . . .	82
5.1 Shandong Temperature Data . . . . .	82

5.2	Irish Wind Data . . . . .	84
5.3	North Atlantic Ocean Temperature Data . . . . .	85
6	Conclusion and Future Work . . . . .	88
6.1	Conclusion . . . . .	88
6.2	Future Work . . . . .	88
	REFERENCES . . . . .	90
	VITA . . . . .	94

## LIST OF FIGURES

3.1	Contour plots for one example of Cressie and Huang’s model . . . . .	32
3.2	Contour plots for one example of Gneiting’s model . . . . .	33
3.3	Contour plots for $\mathcal{M}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$ as functions of $\mathbf{h} = (h_1, h_2)$ , with $u = 0$ , $\mathbf{r} = (0.5, 0)$ , $a_1 = a_2 = 1$ , and $\nu_1 = \nu_2 = 0.35$ . . . . .	36
3.4	Contour plots for $\mathcal{M}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$ as functions of $\mathbf{h} = (h_1, h_2)$ , with $u = 0.1$ , $\mathbf{r} = (0.5, 0)$ , $a_1 = a_2 = 1$ , and $\nu_1 = \nu_2 = 0.35$ . . . . .	37
3.5	Contour plots for $\mathcal{M}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$ as functions of $\mathbf{h} = (h_1, h_2)$ , with $u = 0.2$ , $\mathbf{r} = (0.5, 0)$ , $a_1 = a_2 = 1$ , and $\nu_1 = \nu_2 = 0.35$ . . . . .	38
3.6	Contour plots for $\mathcal{M}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$ as functions of $\mathbf{h} = (h_1, h_2)$ , with $u = 0.5$ , $\mathbf{r} = (0.5, 0)$ , $a_1 = a_2 = 1$ , and $\nu_1 = \nu_2 = 0.35$ . . . . .	39
3.7	3d contour plot for $\mathcal{M}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$ as functions of $\mathbf{h} = (h_1, h_2)$ , with $u = 0$ , $\mathbf{r} = (0.5, 0)$ , $a_1 = a_2 = 1$ , and $\nu_1 = \nu_2 = 0.35$ . $h_1$ varies evenly from -0.1 to 0.1, and $h_2$ varies evenly from -0.1 to 0.1. . . . .	39
3.8	3d contour plot for $\mathcal{M}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$ as functions of $\mathbf{h} = (h_1, h_2)$ , with $u = 0.1$ , $\mathbf{r} = (0.5, 0)$ , $a_1 = a_2 = 1$ , and $\nu_1 = \nu_2 = 0.35$ . $h_1$ varies evenly from -0.1 to 0.1, and $h_2$ varies evenly from -0.1 to 0.1. . . . .	40
3.9	3d contour plot for $\mathcal{M}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$ as functions of $\mathbf{h} = (h_1, h_2)$ , with $u = 0.2$ , $\mathbf{r} = (0.5, 0)$ , $a_1 = a_2 = 1$ , and $\nu_1 = \nu_2 = 0.35$ . $h_1$ varies evenly from -0.1 to 0.1, and $h_2$ varies evenly from -0.1 to 0.1. . . . .	40
3.10	3d contour plot for $\mathcal{M}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$ as functions of $\mathbf{h} = (h_1, h_2)$ , with $u = 0.5$ , $\mathbf{r} = (0.5, 0)$ , $a_1 = a_2 = 1$ , and $\nu_1 = \nu_2 = 0.35$ . $h_1$ varies evenly from -0.1 to 0.1, and $h_2$ varies evenly from -0.1 to 0.1. . . . .	41
3.11	Contour plots for $\mathcal{M}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$ as functions of $\mathbf{h} = (h_1, h_2)$ , with $\mathbf{r} = (0, 0)$ , $u = 0.5$ , $a_1 = a_2 = 1$ , and $\nu_1 = \nu_2 = 0.35$ . . . . .	41
3.12	Contour plots for $\mathcal{M}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$ as functions of $\mathbf{h} = (h_1, h_2)$ , with $\mathbf{r} = (0.1, 0)$ , $u = 0.5$ , $a_1 = a_2 = 1$ , and $\nu_1 = \nu_2 = 0.35$ . . . . .	42
3.13	Contour plots for $\mathcal{M}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$ as functions of $\mathbf{h} = (h_1, h_2)$ , with $\mathbf{r} = (0.2, 0)$ , $u = 0.5$ , $a_1 = a_2 = 1$ , and $\nu_1 = \nu_2 = 0.35$ . . . . .	43
3.14	Contour plots for $\mathcal{M}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$ as functions of $\mathbf{h} = (h_1, h_2)$ , with $\mathbf{r} = (0.5, 0)$ , $u = 0.5$ , $a_1 = a_2 = 1$ , and $\nu_1 = \nu_2 = 0.35$ . . . . .	44
3.15	Contour plots for $\mathcal{M}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$ as functions of $\ \mathbf{h}\ $ and $u$ with $\eta = 0$ , $\mathbf{r} = (0.5, 0)$ , $a_1 = a_2 = 1$ , and $\nu_1 = \nu_2 = 0.35$ , where $\eta$ is the angle between $\mathbf{h}$ and the horizontal axis. . . . .	45

3.16	Contour plots for $\mathcal{M}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$ as functions of $\ \mathbf{h}\ $ and $u$ with $\eta = \pi/8$ , $\mathbf{r} = (0.5, 0)$ , $a_1 = a_2 = 1$ , and $\nu_1 = \nu_2 = 0.35$ , where $\eta$ is the angle between $\mathbf{h}$ and the horizontal axis. . . . .	46
3.17	Contour plots for $\mathcal{M}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$ as functions of $\ \mathbf{h}\ $ and $u$ with $\eta = \pi/4$ , $\mathbf{r} = (0.5, 0)$ , $a_1 = a_2 = 1$ , and $\nu_1 = \nu_2 = 0.35$ , where $\eta$ is the angle between $\mathbf{h}$ and the horizontal axis. . . . .	47
3.18	Contour plots for $\mathcal{M}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$ as functions of $\ \mathbf{h}\ $ and $u$ with $\eta = \pi/2$ , $\mathbf{r} = (0.5, 0)$ , $a_1 = a_2 = 1$ , and $\nu_1 = \nu_2 = 0.35$ , where $\eta$ is the angle between $\mathbf{h}$ and the horizontal axis. . . . .	48
3.19	Contour plots for $\mathcal{M}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$ as functions of $\ \mathbf{r}\ $ and $u$ , with $\eta = 0$ , $\mathbf{h} = (0.5, 0)$ , $a_1 = a_2 = 1$ , and $\nu_1 = \nu_2 = 0.35$ , where $\eta$ is the angle between $\mathbf{r}$ and the horizontal axis. . . . .	49
3.20	Contour plots for $\mathcal{M}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$ as functions of $\ \mathbf{r}\ $ and $u$ , with $\eta = \pi/8$ , $\mathbf{h} = (0.5, 0)$ , $a_1 = a_2 = 1$ , and $\nu_1 = \nu_2 = 0.35$ , where $\eta$ is the angle between $\mathbf{r}$ and the horizontal axis. . . . .	50
3.21	Contour plots for $\mathcal{M}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$ as functions of $\ \mathbf{r}\ $ and $u$ , with $\eta = \pi/4$ , $\mathbf{h} = (0.5, 0)$ , $a_1 = a_2 = 1$ , and $\nu_1 = \nu_2 = 0.35$ , where $\eta$ is the angle between $\mathbf{r}$ and the horizontal axis. . . . .	51
3.22	Contour plots for $\mathcal{M}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$ as functions of $\ \mathbf{r}\ $ and $u$ , with $\eta = \pi/2$ , $\mathbf{h} = (0.5, 0)$ , $a_1 = a_2 = 1$ , and $\nu_1 = \nu_2 = 0.35$ , where $\eta$ is the angle between $\mathbf{r}$ and the horizontal axis. . . . .	52
4.1	Contour plots for $\mathcal{N}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$ as functions of $\mathbf{h} = (h_1, h_2)$ , with $u = 0$ , $\mathbf{r} = (0.5, 0)$ , $a_1 = a_2 = 1$ , and $\nu_1 = \nu_2 = 0.35$ . . . . .	62
4.2	Contour plots for $\mathcal{N}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$ as functions of $\mathbf{h} = (h_1, h_2)$ , with $u = 0.1$ , $\mathbf{r} = (0.5, 0)$ , $a_1 = a_2 = 1$ , and $\nu_1 = \nu_2 = 0.35$ . . . . .	63
4.3	Contour plots for $\mathcal{N}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$ as functions of $\mathbf{h} = (h_1, h_2)$ , with $u = 0.2$ , $\mathbf{r} = (0.5, 0)$ , $a_1 = a_2 = 1$ , and $\nu_1 = \nu_2 = 0.35$ . . . . .	64
4.4	Contour plots for $\mathcal{N}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$ as functions of $\mathbf{h} = (h_1, h_2)$ , with $u = 0.5$ , $\mathbf{r} = (0.5, 0)$ , $a_1 = a_2 = 1$ , and $\nu_1 = \nu_2 = 0.35$ . . . . .	65
4.5	Contour plots for $\mathcal{N}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$ as functions of $\mathbf{h} = (h_1, h_2)$ , with $u = 0$ , $\mathbf{r} = (0.5, 0)$ , $a_1 = a_2 = 1$ , and $\nu_1 = 0.35, \nu_2 = 3$ . . . . .	66
4.6	Contour plots for $\mathcal{N}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$ as functions of $\mathbf{h} = (h_1, h_2)$ , with $u = 0.1$ , $\mathbf{r} = (0.5, 0)$ , $a_1 = a_2 = 1$ , and $\nu_1 = 0.35, \nu_2 = 3$ . . . . .	67
4.7	Contour plots for $\mathcal{N}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$ as functions of $\mathbf{h} = (h_1, h_2)$ , with $u = 0.2$ , $\mathbf{r} = (0.5, 0)$ , $a_1 = a_2 = 1$ , and $\nu_1 = 0.35, \nu_2 = 3$ . . . . .	68
4.8	Contour plots for $\mathcal{N}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$ as functions of $\mathbf{h} = (h_1, h_2)$ , with $u = 0.5$ , $\mathbf{r} = (0.5, 0)$ , $a_1 = a_2 = 1$ , and $\nu_1 = 0.35, \nu_2 = 3$ . . . . .	69

4.9	Contour plots for $\mathcal{N}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$ as functions of $\mathbf{h} = (h_1, h_2)$ , with $\mathbf{r} = (0, 0)$ , $u = 0.5$ , $a_1 = a_2 = 1$ , and $\nu_1 = \nu_2 = 0.35$ . . . . .	70
4.10	Contour plots for $\mathcal{N}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$ as functions of $\mathbf{h} = (h_1, h_2)$ , with $\mathbf{r} = (0.1, 0)$ , $u = 0.5$ , $a_1 = a_2 = 1$ , and $\nu_1 = \nu_2 = 0.35$ . . . . .	71
4.11	Contour plots for $\mathcal{N}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$ as functions of $\mathbf{h} = (h_1, h_2)$ , with $\mathbf{r} = (0.2, 0)$ , $u = 0.5$ , $a_1 = a_2 = 1$ , and $\nu_1 = \nu_2 = 0.35$ . . . . .	72
4.12	Contour plots for $\mathcal{N}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$ as functions of $\mathbf{h} = (h_1, h_2)$ , with $\mathbf{r} = (0.5, 0)$ , $u = 0.5$ , $a_1 = a_2 = 1$ , and $\nu_1 = \nu_2 = 0.35$ . . . . .	73
4.13	Contour plots for $\mathcal{N}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$ as functions of $\ \mathbf{h}\ $ and $u$ , with $\eta = 0$ , $\mathbf{r} = (0.5, 0)$ , $a_1 = a_2 = 1$ , and $\nu_1 = \nu_2 = 0.35$ , where $\eta$ is the angle between $\mathbf{h}$ and the horizontal axis. . . . .	74
4.14	Contour plots for $\mathcal{N}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$ as functions of $\ \mathbf{h}\ $ and $u$ , with $\eta = \pi/8$ , $\mathbf{r} = (0.5, 0)$ , $a_1 = a_2 = 1$ , and $\nu_1 = \nu_2 = 0.35$ , where $\eta$ is the angle between $\mathbf{h}$ and the horizontal axis. . . . .	75
4.15	Contour plots for $\mathcal{N}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$ as functions of $\ \mathbf{h}\ $ and $u$ , with $\eta = \pi/4$ , $\mathbf{r} = (0.5, 0)$ , $a_1 = a_2 = 1$ , and $\nu_1 = \nu_2 = 0.35$ , where $\eta$ is the angle between $\mathbf{h}$ and the horizontal axis. . . . .	76
4.16	Contour plots for $\mathcal{N}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$ as functions of $\ \mathbf{h}\ $ and $u$ , with $\eta = \pi/2$ , $\mathbf{r} = (0.5, 0)$ , $a_1 = a_2 = 1$ , and $\nu_1 = \nu_2 = 0.35$ , where $\eta$ is the angle between $\mathbf{h}$ and the horizontal axis. . . . .	77
4.17	Contour plots for $\mathcal{N}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$ as functions of $\ \mathbf{r}\ $ and $u$ , with $\eta = 0$ , $\mathbf{h} = (0.5, 0)$ , $a_1 = a_2 = 1$ , and $\nu_1 = \nu_2 = 0.35$ , where $\eta$ is the angle between $\mathbf{r}$ and the horizontal axis. . . . .	78
4.18	Contour plots for $\mathcal{N}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$ as functions of $\ \mathbf{r}\ $ and $u$ , with $\eta = \pi/8$ , $\mathbf{h} = (0.5, 0)$ , $a_1 = a_2 = 1$ , and $\nu_1 = \nu_2 = 0.35$ , where $\eta$ is the angle between $\mathbf{r}$ and the horizontal axis. . . . .	79
4.19	Contour plots for $\mathcal{N}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$ as functions of $\ \mathbf{r}\ $ and $u$ , with $\eta = \pi/4$ , $\mathbf{h} = (0.5, 0)$ , $a_1 = a_2 = 1$ , and $\nu_1 = \nu_2 = 0.35$ , where $\eta$ is the angle between $\mathbf{r}$ and the horizontal axis. . . . .	80
4.20	Contour plots for $\mathcal{N}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$ as functions of $\ \mathbf{r}\ $ and $u$ , with $\eta = \pi/2$ , $\mathbf{h} = (0.5, 0)$ , $a_1 = a_2 = 1$ , and $\nu_1 = \nu_2 = 0.35$ , where $\eta$ is the angle between $\mathbf{r}$ and the horizontal axis. . . . .	81
5.1	Locations of Weather Stations in Shandong Province, China in July 2007. . . . .	83
5.2	Values of $\hat{C}(\mathbf{s}_i - \mathbf{s}_j, 1)$ when they are greater than 0.4 corresponding to the direction parallel to $\hat{\mathbf{r}}$ (left) and the counterclockwise direction vertical to $\hat{\mathbf{r}}$ (right). . . . .	84
5.3	Locations of Weather Stations in North Atlantic Ocean. . . . .	86
5.4	Yearly average temperature process at Angmassalik, source <a href="http://rimfrost.no/">http://rimfrost.no/</a> . . . . .	87

## ABBREVIATIONS

NFSST	non-fully symmetric space-time
CDF	cumulative distribution function
PDF	probability density function
GCD	generalized Cauchy distribution
LRD	long-range dependence
iff	if and only if
MLE	maximum likelihood estimation

## ABSTRACT

In spatio-temporal data analysis, the problem of non-separable space-time covariance functions is important and hard to deal with. Most of the famous constructions of these covariance functions are fully symmetric, which is inappropriate in many spatiotemporal processes. The Non-Fully Symmetric Space-Time (NFSST) Matérn model by Zhang, T. and Zhang, H. (2015) [1] provides a way to construct a non-fully symmetric non-separable space-time correlation function from marginal spatial and temporal Matérn correlation functions.

In this work we use the relationship between the spatial Matérn and temporal Cauchy correlation functions and their spectral densities, and provide a modification to their Bochner's representation by including a space-time interaction term. Thus we can construct a non-fully symmetric space-time Matérn-Cauchy model, from any given marginal spatial Matérn and marginal temporal Cauchy correlation functions. We are able to perform computation and parameter estimate on this family, using the Taylor expansion of the correlation functions. This model has attractive properties: it has much faster estimation compared with NFSST Matérn model when the spatio-temporal data is large; it enables the existence of temporal long-range dependence (LRD), adding substantially to the flexibility of marginal correlation function in the time domain. Several spatio-temporal meteorological data sets are studied using our model, including one case with temporal LRD.

# 1. Introduction

The objective of this dissertation is to provide a new approach to constructing non-fully symmetric, non-separable correlation models with known marginal correlation functions in space-time geostatistics studies. In this chapter, we give a literature review of the history, research problems and traditional methodology of geostatistics, and introduce our motivation to propose the non-fully symmetric space-time (NFSST) correlation models.

## 1.1 Geostatistics and correlation functions

Spatial statistics, as a research field within the discipline of statistics, was originally developed from different areas of real world applications, including mining industry, spatially concerned agriculture, and forestry. Generally speaking, spatial statistics is consisted of three major branches: (1) geostatistics, which is focused on processes with continuous spatial variation; (2) lattice data study, concerns those processes with discrete spatial variation; (3) spatial point processes. During the past 30 years, due to the development of high-speed computing and the worldwide usage of geographical information systems (GIS), researchers are endowed with the ability to collect larger spatial and spatio-temporal datasets, and they are able to investigate more complicated and challenging models. As a result, there has been an explosive growing of interest in the studies of space and space-time processes.

Geostatistics originates from mining industry, and develops its methodology for prediction (kriging) by Mathéron in 1950s [2]. Its original real-world application was to predict the expected yield of a mine over a geographical region  $D$ , knowing the samples of ores collected from a finite set of locations. Using a stochastic process  $\{S(x), x \in D\}$  to represent the yield of ores, we observe data  $\{Y_i : i = 1, \dots, n\}$  as the realized value (along with random error) of the process  $S(x)$  at locations  $\{x_i : i = 1, \dots, n\}$ . The model is  $Y_i = S(x_i) + Z_i$ , where the  $Z_i$  are independent noise term with mean zero and variance  $\tau^2$ . For any unobserved location  $x$ , the point predictor of  $S(x)$  is the least mean square error linear estimator. In the kriging method, the process  $S(\cdot)$  is assumed to have mean  $\mu(x) = \mathbf{d}(x)^\top \boldsymbol{\beta}$ , where  $\mathbf{d}(x)$  is a vector of explanatory variables. The kriging method also takes account of the covariance structure of  $S(\cdot)$ .

In general, the aim of geostatistics is to model and predict a spatially distributed variable of interest in a geographical region. The model often includes a response variable and a set of explanatory variables. The response variable is defined in theory at every point over the geographical region, and the region is a bounded  $D \subset \mathbb{R}^d$ , where  $d = 2$  or  $3$ . In the circumstance of spatio-temporal study, the region is assumed to be a bounded  $D \subset \mathbb{R}^d \times \mathbb{R}$ , where  $d = 2$  or  $3$ ,  $\mathbb{R}^d$  is the spatial domain and  $\mathbb{R}$  is the temporal domain. Values of the response variable are observed at fixed stations (and time, in the spatio-temporal case) in the region. Values of explanatory variables can be derived from other source. Therefore, we often assume that the response is only observed at stations (and certain time) and the independent variables are available everywhere in the region. The aim of a geostatistical method is to predict the value of response at unobserved locations (and unobserved time).

Assume a response and a  $p$ -dimensional vector of explanatory variables are observed over space and time domain, and the observed values are denoted by  $y_i$  and the vector of explanatory variables are denoted by  $\mathbf{x}_i$  at location  $(\mathbf{s}_i, t_i)$ , for  $i = 1, \dots, n$ . Here  $y_i = Y(\mathbf{s}_i, t_i)$ , and  $Y(\mathbf{s}, t)$  is a Gaussian random field where  $(\mathbf{s}, t) \in \mathbb{R}^d \times \mathbb{R}$ . Let  $\mathbf{y} = (y_1, \dots, y_n)^\top$  be the column vector of the response and  $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_n)^\top$  be the matrix of the explanatory variables of the observations. Then, a geostatistical model is

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\delta}, \tag{1.1}$$

where  $\boldsymbol{\delta}$  is a spatial correlation error term usually assumed distributed of  $\mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{R})$ , where the spatial correlation matrix  $\mathbf{R} = \text{cor}(\boldsymbol{\delta}, \boldsymbol{\delta})$  is the correlation matrix which may depend on parameter vector  $\boldsymbol{\theta}$ .

Note that  $\{y_i : i = 1, \dots, n\}$  is the realized values of  $Y(s, t)$  at different space-time locations  $(\mathbf{s}_i, t_i)$ , for  $i = 1, \dots, n$ , which are usually distinct locations in real applications. This makes the estimate of the correlation matrix  $\mathbf{R}$  unrealistic, unless some more structures are imposed on it. Therefore we often assume the covariance have certain stationarity, and further we may assume some parametric models for the covariance structure.

It is often assumed the spatial covariance function is either stationary or isotropic. In the pure spatial case,  $t$  is absent. We say the covariance function is *stationary* [3] if

$\text{cov}(Y(\mathbf{s} + \mathbf{h}), Y(\mathbf{s})) = \sigma^2 r(\mathbf{h})$  only depends on  $\mathbf{h}$  for any  $\mathbf{s}$ . It means that the covariance of  $Y(\cdot)$  at two different locations only depends on their relative locations, or in other words, on their spatial lag. We say the covariance function is *isotropic* [3] if  $\text{cov}(Y(\mathbf{s} + \mathbf{h}), Y(\mathbf{h})) = \sigma^2 r(\mathbf{h})$  only depends on  $\|\mathbf{h}\|$ . It means that the covariance of  $Y(\cdot)$  at two different locations only depends on their (Euclidean) distance. It is clearly that isotropic implies stationary and if the covariance function of  $Y$  is stationary, then  $\text{var}(Y(\mathbf{s}))$  does not depend on  $\mathbf{s}$ . Therefore, we can write covariance of  $Y$  as  $\sigma^2 \mathbf{R}$ .

It is often model the spatial correlation matrix of  $\boldsymbol{\delta}$  by a parametric correlation function. Suppose the correlation is expressed as  $c_\theta(\mathbf{h})$ , where  $\mathbf{h}$  is the spatial distance. The correlation matrix is

$$\mathbf{R} = \begin{pmatrix} c_\theta(\mathbf{d}_{11}) & c_\theta(\mathbf{d}_{12}) & \cdots & c_\theta(\mathbf{d}_{1n}) \\ c_\theta(\mathbf{d}_{21}) & c_\theta(\mathbf{d}_{22}) & \cdots & c_\theta(\mathbf{d}_{2n}) \\ \vdots & \vdots & \ddots & \vdots \\ c_\theta(\mathbf{d}_{n1}) & c_\theta(\mathbf{d}_{n2}) & \cdots & c_\theta(\mathbf{d}_{nn}) \end{pmatrix}, \quad (1.2)$$

where  $\mathbf{d}_{ij}$  is the distance between site  $i$  and  $j$ . Clearly, we have  $c_\theta(\mathbf{0}) = 1$  and  $c_\theta(\mathbf{h}) < 1$  if  $\|\mathbf{h}\| > 0$ . If  $\boldsymbol{\theta}$  is derived, then  $c_\theta$  is derived which means  $\mathbf{R}$  is known. Then  $\boldsymbol{\beta}$  can be estimated by the generalized least square method as

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^\top \mathbf{R}^{-1} \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{R}^{-1} \mathbf{y}, \quad (1.3)$$

and  $\sigma^2$  is estimated by

$$\hat{\sigma}^2 = \frac{1}{n-p} \mathbf{y}^\top [\mathbf{R}^{-1} - \mathbf{R}^{-1} \mathbf{X} (\mathbf{X}^\top \mathbf{R}^{-1} \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{R}^{-1}] \mathbf{y}. \quad (1.4)$$

Based on the above, we can estimate  $\boldsymbol{\theta}$  by using the profile maximum likelihood method, which aims to maximize the function

$$\begin{aligned} \ell_p(\boldsymbol{\theta}) &= -\frac{n}{2} \left[ 1 + \log\left(\frac{2\pi}{n}\right) \right] - \frac{1}{2} \log |\det(\mathbf{R}_\theta)| \\ &\quad - \frac{n}{2} \log [\mathbf{y}^\top \mathbf{R}_\theta^{-1} \mathbf{y} - \mathbf{y}^\top \mathbf{R}_\theta^{-1} \mathbf{X} (\mathbf{X}^\top \mathbf{R}_\theta^{-1} \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{R}_\theta^{-1} \mathbf{y}] \\ &= -\frac{n}{2} \left[ 1 + \log\left(\frac{2\pi}{n}\right) \right] - \frac{1}{2} \log |\det(\mathbf{R}_\theta)| - \frac{n}{2} \log(\mathbf{y}^\top \mathbf{M}_\theta \mathbf{y}). \end{aligned} \quad (1.5)$$

where

$$\mathbf{M}_\theta = \mathbf{R}_\theta^{-1} - \mathbf{R}_\theta^{-1} \mathbf{X} (\mathbf{X}^\top \mathbf{R}_\theta^{-1} \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{R}_\theta^{-1}. \quad (1.6)$$

Therefore,  $\theta$  can be estimated using maximum likelihood estimate. Let  $\dot{\ell} = (\frac{\partial \ell_p}{\partial \theta_1}, \frac{\partial \ell_p}{\partial \theta_2}, \dots, \frac{\partial \ell_p}{\partial \theta_n})$  be the vector of gradient and  $\ddot{\ell}_p = (\frac{\partial^2 \ell_p}{\partial \theta_i \partial \theta_j})$  be the Hessian matrix. We have the first order partial derivatives with respect to each element of  $\theta$  is

$$\frac{\partial \ell_p}{\partial \theta_j} = -\frac{1}{2} \text{tr}(\mathbf{R}_\theta^{-1} \frac{\partial \mathbf{R}_\theta}{\partial \theta_j}) + \frac{n \mathbf{y}^\top \mathbf{M}_\theta \frac{\partial \mathbf{R}_\theta}{\partial \theta_j} \mathbf{M}_\theta \mathbf{y}}{2 \mathbf{y}^\top \mathbf{M}_\theta \mathbf{y}}. \quad (1.7)$$

and the second order partial derivative function is

$$\begin{aligned} \frac{\partial^2 \ell_p}{\partial \theta_{j_1} \partial \theta_{j_2}} = & -\frac{1}{2} \text{tr}(\mathbf{R}_\theta^{-1} \frac{\partial^2 \mathbf{R}_\theta}{\partial \theta_{j_2} \partial \theta_{j_1}}) + \frac{1}{2} \text{tr}(\mathbf{R}_\theta^{-1} \frac{\partial \mathbf{R}_\theta}{\partial \theta_{j_1}} \mathbf{R}_\theta^{-1} \frac{\partial \mathbf{R}_\theta}{\partial \theta_{j_2}}) \\ & + \frac{n \mathbf{y}^\top \mathbf{M}_\theta \frac{\partial^2 \mathbf{R}_\theta}{\partial \theta_{j_1} \partial \theta_{j_2}} \mathbf{M}_\theta \mathbf{y}}{2 \mathbf{y}^\top \mathbf{M}_\theta \mathbf{y}} + \frac{n (\mathbf{y}^\top \mathbf{M}_\theta \frac{\partial \mathbf{R}_\theta}{\partial \theta_{j_1}} \mathbf{M}_\theta \mathbf{y}) (\mathbf{y}^\top \mathbf{M}_\theta \frac{\partial \mathbf{R}_\theta}{\partial \theta_{j_2}} \mathbf{M}_\theta \mathbf{y})}{2 (\mathbf{y}^\top \mathbf{M}_\theta \mathbf{y})^2} \\ & - \frac{n \mathbf{y}^\top \mathbf{M}_\theta (\frac{\partial \mathbf{R}_\theta}{\partial \theta_{j_1}} \mathbf{M}_\theta \frac{\partial \mathbf{R}_\theta}{\partial \theta_{j_2}} + \frac{\partial \mathbf{R}_\theta}{\partial \theta_{j_2}} \mathbf{M}_\theta \frac{\partial \mathbf{R}_\theta}{\partial \theta_{j_1}}) \mathbf{M}_\theta \mathbf{y}}{\mathbf{y}^\top \mathbf{M}_\theta \mathbf{y}}. \end{aligned} \quad (1.8)$$

where  $j, j_1, j_2 = 1, \dots, k$ .

The maximum likelihood estimate of  $\theta$  can be solved by Newton-Raphson iterative method by

$$\theta^{(l+1)} = \theta^{(l)} - \ddot{\ell}^{-1}(\theta^{(l)}) \dot{\ell}(\theta^{(l)}). \quad (1.9)$$

The predicted value of  $y$  at an unobserved site  $s_0$ , denoted by  $y_0$ , is computed by the kriging (or co-kriging) method as

$$\begin{aligned} y_0^* = \text{E}(y_0 | \mathbf{y}) &= \text{E}(y_0 - \mathbf{c}_0^\top \mathbf{R}^{-1} \mathbf{y} | \mathbf{y}) + \mathbf{c}_0^\top \mathbf{R}^{-1} \mathbf{y} \\ &= \mathbf{x}_0^\top \boldsymbol{\beta} + \mathbf{c}_0^\top \mathbf{R}^{-1} (\mathbf{y} - \mathbf{X} \boldsymbol{\beta}), \end{aligned} \quad (1.10)$$

where  $\mathbf{c}_0 = (\text{cor}(y_0, y_1), \dots, \text{cor}(y_0, y_n))^\top$  is the correlation vector between  $y_0$  and  $\mathbf{y}$ .

From the procedure above, we learn that once the parametric family of the correlation function  $c_\theta(\mathbf{h})$  is determined, we are able to estimate the parameters in the model and

make predictions based on the likelihood approach. Therefore, in a geostatistical model the critical point is to figure out a valid spatial correlation function. Currently, a well-known parametric family of correlation function is the popular Matérn correlation function, which was firstly introduced by Bertil Matérn's doctoral dissertation [4]. An one-dimensional Matérn correlation function is given by

$$c_{\theta}(h) = \begin{cases} 1, & \text{when } h = 0, \\ \frac{\theta_1}{2^{\alpha-1}\Gamma(\alpha)} \left(\frac{h}{\theta_2}\right)^{\alpha} K_{\alpha}\left(\frac{h}{\theta_2}\right), & \text{when } h \neq 0. \end{cases} \quad (1.11)$$

where  $0 < \theta_1 < 1$ ,  $\theta_2 > 0$  and  $K_{\alpha}$  is the modified Bessel function of second kind.  $\theta_2$  is called the scale parameter and  $\alpha$  is called the smoothness parameter. This parametric family is related with the student-t family, and it has been the choice of correlation family in many geostatistics studies.

## 1.2 Spatio-temporal processes

In spatio-temporal geostatistics studies, we consider processes

$$\{Y(\mathbf{s}, t) : (\mathbf{s}, t) \in D \subseteq \mathbb{R}^d \times \mathbb{R}\}$$

defined on both spatial domain  $\mathbf{s} \in \mathbb{R}^d$  and temporal domain  $t \in \mathbb{R}$ . From a purely mathematical perspective, there is no much difference between the spatio-temporal processes and pure spatial processes. Since the temporal domain can be considered as an extra dimension of space. And a spatio-temporal process can be viewed as a process defined on  $\mathbb{R}^{d+1} = \mathbb{R}^d \times \mathbb{R}$ . The modelling, estimation and kriging procedure can also be applied using the results on domain  $\mathbb{R}^{d+1}$ . However, from a physical perspective, such viewing may not be appropriate. Since time can only move forward not backward while there is no such restriction on space. Space and time differ in their scale and units as well. Therefore we want to build realistic statistical models taking care of these issues, more specifically, we want to construct space-time correlation functions revealing both distinction and interaction between space and time domains.

For a second-order stationary spatiotemporal process  $Y(\mathbf{s}, t)$ ,  $\mathbf{s} \in \mathbb{R}^d$  and  $t \in \mathbb{R}$ , its covariance function,  $C(\mathbf{h}, u) = \text{Cov}(Y(\mathbf{s}, t), Y(\mathbf{s} + \mathbf{h}, t + u))$ ,  $\mathbf{h} \in \mathbb{R}^d$ ,  $u \in \mathbb{R}$ , is essential to estimation and prediction. A space-time covariance function  $C$  is called *fully symmetric* [5] if

$$C(\mathbf{h}, u) = C(-\mathbf{h}, u) = C(\mathbf{h}, -u) = C(-\mathbf{h}, -u), \forall \mathbf{h} \in \mathbb{R}^d, u \in \mathbb{R}. \quad (1.12)$$

In the purely spatial context, this property is also known as axial symmetry [6]. With fixed temporal lag, a fully symmetric model can not distinguish the different effects as spatial lag towards one direction or its opposite direction. In real applications, there are climate and geophysical processes under some predominant forces following certain directions. And these factors make the fully symmetric model unrealistic. For instance, in mainland China, during summer monsoon seasons there are prevailing air movement towards northeast. Then today's high pollutant concentration at a southwestern location will likely result in tomorrow's high pollutant concentration at a northeastern location, but not vice versa. There are similar effects in ocean currents as well, for example the north Atlantic ocean currents flow towards certain directions in every summer and winter. These effects are well known in the geophysical literature, and it has been pointed out by Gneiting (2002) [5] that the fully symmetry is often violated when environmental processes are considered dynamically, it is more appropriate to use non-fully symmetric space-time covariance functions in applications.

A space-time covariance function  $C$  is called *separable* [7] if

$$C(\mathbf{h}, u) = K_1(\mathbf{h})K_2(u), \forall \mathbf{h} \in \mathbb{R}^d, u \in \mathbb{R}, \quad (1.13)$$

where  $K_1(\cdot)$  is a covariance function on  $\mathbb{R}^d$  and  $K_2(\cdot)$  is a covariance function on  $\mathbb{R}$ . One clear drawback of the separability models is the lack of space-time interaction. Stein in 2005 [7] pointed out the ridge issue caused by separable models. Thus non-separable models is more appropriate in real-world studies.

One basic, non-fully symmetric and non-separable construction of space-time covariance functions from known spatial covariance function is the frozen field model

$$C(\mathbf{h}, u) = C_s(\mathbf{h} - \mathbf{v}u), \quad (1.14)$$

where  $\mathbf{h} \in \mathbb{R}^d, u \in \mathbb{R}$ ,  $C_s$  is a spatial covariance function on  $\mathbb{R}^d$ ,  $\mathbf{v}$  is a velocity vector showing the time-forward movement speed of  $C_s$ . This model was proposed by Briggs (1968) [8] and proved useful when there are transportation effects in spatio-temporal processes [3]. The frozen field model cooperates spatial lag and temporal lag via  $\mathbf{v}$ . For a fixed temporal lag, the spatial marginal covariance function follows  $C_s$ , however for fixed spatial lag  $\mathbf{h}$  the temporal marginal covariance function is not clear.

A space-time covariance function is a spatial covariance function if time is ignored or a temporal covariance function if space is ignored. As mentioned in last section, if time is not involved, a great interest is to consider the Matérn family of spatial correlation functions as

$$\begin{aligned} M_d(\mathbf{h}|\nu, a) &= \frac{(a\|\mathbf{h}\|)^\nu}{2^{\nu-1}\Gamma(\nu)} K_\nu(a\|\mathbf{h}\|) \\ &= M_\nu(a\|\mathbf{h}\|), \mathbf{h} \in \mathbb{R}^d, a, \nu > 0, \end{aligned} \tag{1.15}$$

where  $K_\nu$  is a modified Bessel function of the second kind,  $a$  and  $\nu$  are scale and smoothness parameters respectively, and  $M_\nu(z) = |z|^\nu K_\nu(z)/[2^{\nu-1}\Gamma(\nu)], z \in \mathbb{R}$ . The Matérn family is isotropic in space. It was proposed by Matérn (1986) [9] and has received more attention since some theoretical work by Handcock (1993) [10] and Stein (1999) [11]. A nice review and discussion on Matérn family is given by Guttorp and Gneiting (2006) [12]. The Matérn spatial correlation family has been used in many applications (Lee and Shaddick, 2010 [13]; North, Wang, and Genton, 2011 [14]).

As discussed, an isotropic correlation function is inappropriate in modeling a spatiotemporal process since the temporal axis and the spatial axis are of different scales and the Euclidean distance in the product space is not suitable. The construction of space-time correlation functions is an interesting but difficult problem. There have been good developments in recent years on the construction of space-time correlation functions. The simplest case is the separable one provided by the product of a spatial correlation function and a temporal correlation function, but it does not model the space-time interaction [15] [7] and also too restrictive for space-time data analysis. A more detailed discussion of the shortcomings of separable models can be found in Kyriakidis and Journel (1999) [16]. Recently, much effort has been put on the construction of nonseparable space-time covariance functions. Many

models have been proposed. Examples include the product-sum model (De Iaco, Myers, and Posa, 2001 [17]), the mixture models (Ma, 2002, 2003b [18] [19]), and anisotropic spatial component models (Mateu, Porcu, and Gregori, 2008 [20]; Porcu, Gregori, and Mateu, 2006 [21]; Porcu, Mateu, and Saura, 2008 [22]). In the construction of these models, many methods have been proposed. Examples include the spectral representation method (Cressie and Huang, 1999 [15]; Stein, 2005 [7]), the completely monotone function method (Gneiting, 2002 [5]), the linear combination method (Ma, 2005 [23]), the convolution-based method (Rodrigues and Diggle, 2010 [24]), the stochastic differential equation method (Kolovos, et. al, 2004 [25]), and the mixture representation method (Fonseca and Steel, 2011 [26]). There have been a number of studies on the properties of space-time correlation functions, including, for example, the test for separability (Brown, Karesen, and Tonellato, 2000 [27]; Fuentes, 2006 [28]; Li, Genton, Sherman, 2007 [29]; Mitchell, Genton, and Gumpertz, 2005 [30]), the evaluation of spatial or temporal margins (De Iaco, Posa, and Myers, 2013 [31]), and types of nonseparability (De Iaco and Posa, 2013 [32]).

An important issue is the assessment of full symmetry. Although fully symmetry is a desirable assumption from a computational point of view, it may not be appropriate in applications (Shao and Li, 2009 [33]). Atmospheric, environmental, and geophysical processes are often under the influence of prevailing air or water flows, resulting in a lack of full symmetry (Gneiting, Genton, and Guttorp, 2007 [34]). Non-separable and fully-symmetric space-time covariance functions can be constructed by mixtures of separable covariance functions (De Iaco, Myers, and Posa, 2002 [35]; Ma, 2003a [36]). A geometric non-fully symmetric space-time covariance function can be formulated using a geometric transformation on a fully symmetric space-time covariance function (Stein, 2005 [7]). There is a great need for covariance functions which are non-fully symmetric in the spatiotemporal domain (Mateu, Porcu, and Gregori, 2008 [20]).

### 1.3 Our contribution

In this work, we focus on the construction of a non-separable, non-fully symmetric space-time (NFSST) correlation model that satisfies any given spatial Matérn or temporal

Cauchy margins. Specifically, given any spatial Matérn correlation function  $M_d(\mathbf{h}|\nu_1, a_1)$  (called the spatial Matérn margin) and a temporal Cauchy correlation function  $D(u|\nu_2, a_2)$  (called the temporal Cauchy margin), we construct an NFSST correlation function  $C(\mathbf{h}, u)$  satisfying

$$C(\mathbf{h}, 0) = M_d(\mathbf{h}|\nu_1, a_1), C(\mathbf{0}, u) = D(u|\nu_2, a_2). \quad (1.16)$$

Although the model is non-fully symmetric, its spatial and temporal margins are both isotropic. We allow the two smoothness parameters and the two scale parameters to be arbitrary. In some settings, the temporal Cauchy margin can have long-memory property.

The remainder of this dissertation is organized as follows. In Chapter 2, we provide an approach of constructing valid space-time correlation functions from using of Bochner's representation. In Chapter 3, we review the construction of some space-time correlation models and an non-fully symmetric space-time (NFSST) Matérn model. In Chapter 4, we employ the spectral approach to construct an NFSST Matérn-Cauchy model. In Chapter 5, we apply our models to some meteorological data sets. In Chapter 6, we provide a discussion on our models and our future work.

## 2. Spectral methods

This chapter introduces the spectral methods we use in the construction of space-time correlation functions. Bochner's representation, or spectral representation builds a bridge between the space-time domain and frequency domain. We can construct valid covariance functions and introduce new covariance models using their spectral densities.

### 2.1 Bochner's Representation

The study of space-time correlation functions is essential in spatio-temporal geostatistics research. A correlation function with appropriate properties leads to better fitting and prediction. However, the construction of valid correlation functions is not trivial, and Bochner's representation works as a powerful tool in such construction. One important characteristic of a valid (stationary) correlation function is that it must be positive definite, which means that the variance-covariance matrix

$$\mathbf{R} = \begin{pmatrix} c(\mathbf{s}_1 - \mathbf{s}_1) & c(\mathbf{s}_1 - \mathbf{s}_2) & \cdots & c(\mathbf{s}_1 - \mathbf{s}_n) \\ c(\mathbf{s}_2 - \mathbf{s}_1) & c(\mathbf{s}_2 - \mathbf{s}_2) & \cdots & c(\mathbf{s}_2 - \mathbf{s}_n) \\ \vdots & \vdots & \ddots & \vdots \\ c(\mathbf{s}_n - \mathbf{s}_1) & c(\mathbf{s}_n - \mathbf{s}_2) & \cdots & c(\mathbf{s}_n - \mathbf{s}_n) \end{pmatrix} \quad (2.1)$$

must be non-negative definite: for real numbers  $a_1, \dots, a_n$ , there is

$$\sum_{i=1}^n \sum_{j=1}^n a_i a_j c(\mathbf{s}_i - \mathbf{s}_j) \geq 0. \quad (2.2)$$

Suppose a space-time correlation function follows a parametric family  $C_{\boldsymbol{\theta}}(\mathbf{h}, u)$ , it is positive definite thus has its Bochner's representation, or spectral representation

$$C(\mathbf{h}, u) = \int_{\mathbb{R}^d \times \mathbb{R}} e^{i\mathbf{h}^\top \mathbf{s} + iut} F(d\mathbf{s}, dt), \quad (2.3)$$

for any  $\mathbf{s}, \mathbf{h} \in \mathbb{R}^d$  and  $t, u \in \mathbb{R}$ , where  $F$  is the spectral measure satisfying  $F(A, B) = F(-A, -B)$  for any Borel  $A \subseteq \mathbb{R}^d$  and  $B \subseteq \mathbb{R}$ .

If  $F$  is a finite measure and  $C(\mathbf{h}, u)$  is integrable, then  $F$  is absolutely continuous with respect to the Lebesgue measure on  $\mathbb{R}^d \times \mathbb{R}$  and Bochner's representation becomes

$$C(\mathbf{h}, u) = \int_{\mathbb{R}} \int_{\mathbb{R}^d} e^{i\mathbf{h}^\top \mathbf{s} + iut} f(\mathbf{s}, t) d\mathbf{s} dt, \quad (2.4)$$

where the non-negative integrable function  $f(\mathbf{s}, t)$  is called the spectral density of the process satisfies  $f(\mathbf{s}, t) = f(-\mathbf{s}, -t)$ .

It can be shown that the spectral density representation is positive-definite, since

$$\begin{aligned} & \sum_{i=1}^n \sum_{j=1}^n a_i a_j \int e^{i\mathbf{z}^\top (\mathbf{s}_i - \mathbf{s}_j)} f(\mathbf{z}) d\mathbf{z} \\ &= \int \sum_{i=1}^n \sum_{j=1}^n a_i a_j e^{i\mathbf{z}^\top \mathbf{s}_i} e^{-i\mathbf{z}^\top \mathbf{s}_j} f(\mathbf{z}) d\mathbf{z} \\ &= \int \left( \sum_{i=1}^n a_i e^{i\mathbf{z}^\top \mathbf{s}_i} \right) \left( \sum_{j=1}^n a_j e^{i\mathbf{z}^\top \mathbf{s}_j} \right) f(\mathbf{z}) d\mathbf{z} \\ &= \int \left\| \sum_{i=1}^n a_i e^{i\mathbf{z}^\top \mathbf{s}_i} \right\|^2 f(\mathbf{z}) d\mathbf{z} \geq 0. \end{aligned}$$

According to Bochner's theorem [37], a real continuous space-time function  $C$  on  $\mathbb{R}^d \times \mathbb{R}$  is a space-time covariance function if and only if it can be represented in the form of 2.4. Thus, the structure of a space-time process can be analyzed using its spectral density, or equivalently by estimating its autocorrelation function.

It is useful to think of correlation functions and spectral densities as occupying two domains. The correlation functions are in the space-time domain, while the spectral densities are in the frequency domain. Because of the properties of Bochner's representation, if we perform operations in one domain, there are corresponding operations in the other [38]. For instance, one multiplication operation in the frequency domain becomes convolution operation in the space-time domain, and vice versa. It enables us to jump between the space-time domain and frequency domain, and perform operations when they are most convenient.

Thus Bochner's representation is very useful in the construction of space-time correlation functions. Those functions can be specified as spatial correlation functions if time is ignored or as temporal correlation functions if space is ignored. In the spatio-temporal context,

a few nonseparable space-time covariance models have been constructed using Bochner's representation, including Cressie and Huang's model, proposed in 199 and Stein's model proposed in 2005.

## 2.2 Our methodology

We use spectral methods in the construction of our NFSST model. And the methodology is based on the following two theorems.

**Theorem 2.2.1.** *Let  $A(\mathbf{h}, u)$  be a space-time covariance function. Let  $Z_1$  and  $Z_2$  be positive random variables with joint CDF  $G$ . Then*

$$C(\mathbf{h}, u) = \int_0^\infty \int_0^\infty A(\mathbf{h}z_1, uz_2)G(dz_1, dz_2) \quad (2.5)$$

*is a valid space-time covariance function.*

*Proof.* Let  $F$  be the spectral distribution of  $A(\mathbf{h}, u)$ . Without the loss of generality, assume  $F$  is a probability measure, which is generated by a random vector  $\mathbf{x}_1 \in \mathbb{R}^d$  and  $X_2 \in \mathbb{R}$ . If  $(\mathbf{x}_1, X_2)$  is independent of  $(Z_1, Z_2)$ , then

$$A(\mathbf{h}Z_1, uZ_2) = \mathbb{E}(e^{i(Z_1\mathbf{h}^\top \mathbf{x}_1 + Z_2uX_2)} | Z_1, Z_2)$$

and

$$\begin{aligned} C(\mathbf{h}, u) &= \int_0^\infty \int_0^\infty \mathbb{E}[e^{i(Z_1\mathbf{h}^\top \mathbf{x}_1 + Z_2uX_2)}]G(dz_1, dz_2) \\ &= \mathbb{E}[\mathbb{E}(e^{i(Z_1\mathbf{h}^\top \mathbf{x}_1 + Z_2uX_2)} | Z_1, Z_2)] \\ &= \mathbb{E}(e^{i(Z_1\mathbf{h}^\top \mathbf{x}_1 + Z_2uX_2)}). \end{aligned}$$

Therefore,  $C(\mathbf{h}, u)$  is a characteristic function of  $\mathbf{x}_1Z_1$  and  $X_2Z_2$ , implying that  $C(\mathbf{h}, u)$  is a valid space-time correlation function. if  $F$  is not a probability measure, then  $C(\mathbf{h}, u)$  is a covariance function but not a correlation function.  $\square$

Theorem 2.2.1 shows how can we construct a new space-time covariance function  $C(\mathbf{h}, u)$  given a known space-time covariance function  $A(\mathbf{h}, u)$ . By introducing two positive random variables, compounding them with space and time variables, and integration out, we can get a new valid covariance function. By careful selection on  $Z_1$  and  $Z_2$ , the new  $C(\mathbf{h}, u)$  may have ideal properties and its computation might be efficiency.

The properties of the original  $A(\mathbf{h}, u)$  also affect  $C(\mathbf{h}, u)$ , states by the following theorem.

**Theorem 2.2.2.** *Let  $Z_1$  and  $Z_2$  be independent positive random variables on  $(0, \infty)$  in Theorem 2.2.1. The necessary and sufficient conditions for  $C(\mathbf{h}, u)$  to be separable in 2.5 is that  $A(\mathbf{h}, u)$  be separable.*

*Proof.* The sufficiency is directly implied by the expression of  $C(\mathbf{h}, u)$ . For the necessity, if  $C(\mathbf{h}, u)$  is separable then, using the uniqueness of the inverse Fourier transformation, we conclude that  $\mathbf{x}_1 Z_1$  and  $X_2 Z_2$  are independent. Since  $(\mathbf{x}_1, X_2)$ ,  $Z_1$  and  $Z_2$  are independent,  $(\mathbf{x}_1 Z_1, Z_1)$  is independent of  $Z_2$  and  $(X_2 Z_2, Z_2)$  is independent of  $Z_1$ . Let  $z_1$  and  $z_2$  be real values such that  $P(Z_1 \in dz_1)$  and  $P(Z_2 \in dz_2)$  are positive if  $|dz_1|$  and  $|dz_2|$ , the Lebesgue measures of  $dz_1$  and  $dz_2$  respectively, are positive. For any  $B_1 \in \mathcal{B}(\mathbb{R}^d)$  and  $B_2 \in \mathcal{B}(\mathbb{R})$ ,

$$\begin{aligned}
& P(\mathbf{x}_1 \in B_1, X_2 \in B_2) \\
&= \lim_{|dz_1| \rightarrow 0, |dz_2| \rightarrow 0} P(Z_1 \mathbf{x}_1 \in z_1 B_1, Z_2 X_2 \in z_2 B_2 | Z_1 \in dz_1, Z_2 \in dz_2) \\
&= \lim_{|dz_1| \rightarrow 0, |dz_2| \rightarrow 0} \frac{P(Z_1 \mathbf{x}_1 \in z_1 B_1, Z_2 X_2 \in z_2 B_2, Z_1 \in dz_1, Z_2 \in dz_2)}{P(Z_1 \in dz_1, Z_2 \in dz_2)} \\
&= \lim_{|dz_1| \rightarrow 0, |dz_2| \rightarrow 0} \frac{P(Z_1 \mathbf{x}_1 \in z_1 B_1, Z_1 \in dz_1) P(Z_2 X_2 \in z_2 B_2, Z_2 \in dz_2)}{P(Z_1 \in dz_1) P(Z_2 \in dz_2)} \\
&= \lim_{|dz_1| \rightarrow 0, |dz_2| \rightarrow 0} P(Z_1 \mathbf{x}_1 \in z_1 B_1 | Z_1 \in dz_1) P(Z_2 X_2 \in z_2 B_2 | Z_2 \in dz_2) \\
&= P(\mathbf{x}_1 \in B_1) P(X_2 \in B_2).
\end{aligned}$$

Therefore,  $\mathbf{x}_1$  and  $X_2$  are independent, implying that  $A(\mathbf{h}, u)$  is separable.  $\square$

The general idea in our approach is to use a trivial  $A(\mathbf{h}, u)$  to construct a non-trivial  $C(\mathbf{h}, u)$ . As the choice of  $Z_1$  and  $Z_2$  is generally flexible, many different families of space-time

correlation functions may be obtained if different types of  $Z_1$  and  $Z_2$  are utilized. In the following sections, we provide a way to construct a non-fully symmetric space-time Matérn-Cauchy model by using particular  $Z_1$  and  $Z_2$ , along with a Gaussian characteristic function  $A(\mathbf{h}, u)$ .

### 3. Correlation Models

In this chapter we review some widely used correlation models and some famous construction of space-time correlation functions. We can find how Bochner's representation plays a key role in many of these constructions, and some properties of these correlation models are studied as well.

#### 3.1 Matérn Models

A well-known family of spatial correlation functions is the Matérn family.

$$M_d(\mathbf{h}|\nu, a) = \frac{(a\|\mathbf{h}\|)^\nu}{2^{\nu-1}\Gamma(\nu)} K_\nu(a\|\mathbf{h}\|), \quad \mathbf{h} \in \mathbb{R}^d, \quad a, \nu > 0,$$

where  $K_\nu$  is a modified Bessel function of the second kind,  $a$  and  $\nu$  are the scale and smoothness parameters respectively. It was proposed by Matérn (1986) [9] and studied by Handcock and Stein (1993) [10], Guttorp and Gneiting (2006) [12]. And it has been used in many applications such as Lee and Shaddick (2010) [13], North, Wang and Genton (2011) [14].

The spectral density of the spatial Matérn correlation function is

$$m_d(\mathbf{x}|\nu, a) = \frac{\Gamma(\nu + d/2)}{\pi^{d/2} a^d \Gamma(\nu)} \frac{1}{(1 + \frac{\|\mathbf{x}\|^2}{a^2})^{\nu+d/2}}, \quad \mathbf{x} \in \mathbb{R}^d, \quad a, \nu > 0. \quad (3.1)$$

There is relationship between the spatial Matérn correlation function and the characteristic function of a multivariate t-distribution. It has can also be defined by the following theorem.

**Theorem 3.1.1.** *Let  $\mathbf{u}$  be a  $d$ -dimensional  $\mathcal{N}(0, \mathbf{I}_d)$  random variable and  $V$  be an univariate  $\Gamma(\nu, 1/2)$  random variable. If  $\mathbf{u}$  and  $V$  are independent, then  $m_d(\mathbf{x}|\nu, a)$  is the PDF and  $M_d(\mathbf{h}|\nu, a)$  is the characteristics function of  $\mathbf{x} = a\mathbf{u}/\sqrt{V}$ .*

*Proof.*  $\mathbf{u}$  and  $V$  are independent, so the joint PDF of  $(\mathbf{u}, V)$  is

$$f(\mathbf{u}, v) = \frac{v^{\nu-1}}{2^{\nu+\frac{d}{2}} \pi^{\frac{d}{2}} \Gamma(\nu)} e^{-\frac{1}{2}(\|\mathbf{u}\|^2 + \nu)}.$$

The inverse transformation of  $(\mathbf{x}, V) = (a\mathbf{u}/\sqrt{V}, V)$  is  $(\mathbf{u}, v) = (\sqrt{V}\mathbf{x}/a, V)$ . The determinant of the Jacobian matrix is  $(V/a)^{\frac{d}{2}}$ . So the joint PDF of  $(\mathbf{x}, V)$  is

$$g(\mathbf{x}, v) = \frac{v^{\nu+\frac{d}{2}-1}}{2^{\nu+\frac{d}{2}} a^{\frac{d}{2}} \pi^{\frac{d}{2}} \Gamma(\nu)} e^{-\frac{1}{2}(1+\frac{\|\mathbf{u}\|^2}{a^2})v}.$$

Integrating out the  $v$  in  $g(\mathbf{x}, v)$ , we will get the marginal PDF of  $\mathbf{x}$  is  $m_d(\mathbf{x}|\nu, a)$ . And thus the characteristic function of  $\mathbf{x}$  is  $M_d(\mathbf{h}|\nu, a)$ .  $\square$

Theorem 3.1.1 gives the integral representation of a Matérn correlation function.

**Corollary 3.1.1.** *If  $\nu, a > 0$ , then for any  $\mathbf{h} \in \mathbb{R}^d$  there is*

$$M_d(\mathbf{h}|\nu, a) = \frac{1}{2^\nu \Gamma(\nu)} \int_0^\infty v^{\nu-1} e^{-\frac{1}{2}(\frac{a^2\|\mathbf{h}\|^2}{v}+v)} dv. \quad (3.2)$$

*Proof.* There is

$$M_d(\mathbf{h}|\nu, a) = \mathbb{E}(e^{i\mathbf{h}^\top \mathbf{x}}) = \mathbb{E}[\mathbb{E}(e^{i\mathbf{h}^\top (a\mathbf{u}/\sqrt{V})} | V)] = \mathbb{E}(e^{-\frac{a^2\|\mathbf{h}\|^2}{2V}}),$$

which yields 3.2.  $\square$

By comparing 3.2 with 1.15, another expression of the modified Bessel function is derived

:

$$\int_0^\infty v^{\nu-1} e^{-\frac{1}{2}(\frac{z^2}{v}+v)} dv = 2|z|^\nu K_\nu(z) = 2^\nu \Gamma(\nu) M_\nu(z), z \in \mathbb{R}, \nu > 0. \quad (3.3)$$

This expression is useful in the numerical computation for our NFSST Matérn-Cauchy model in Section 4.3.

## 3.2 Cauchy Models

In this section we introduce the Cauchy spatial correlation function:

$$C(u) = \frac{1}{(1+a^2u^2)^{1/\beta}}, u \in \mathbb{R}, a, \beta > 0, \quad (3.4)$$

where  $a$  and  $\beta$  are the scale and smoothness parameters respectively. It is a special case of Generalized Cauchy distribution (GCD) (Devroye, 1990). The whole family of Generalized Cauchy distribution (GCD) has a closed form PDF, and heavy, algebraic tails that makes it suitable for modeling many real-life impulsive processes (Rider PR, 1957 [39]; Miller J and Thomas J, 1972 [40]; Carrillo, Aysal and Barner, 2010) [41]. The proposition below shows the relationship between the generalized Cauchy distribution and the characteristic function of an exponential distribution.

**Proposition 3.2.1.** *Devroye (1990) Let  $\alpha \in (0, 2]$  and  $\beta > 0$  be given constants. Let  $S_\alpha$  be a symmetric stable random variable with characteristic function  $e^{-|t|^\alpha}$ . And let  $V_\beta$  be an independent random variable with density  $\exp(-\nu^\beta)/\Gamma(1 + 1/\beta)$ ,  $\nu > 0$ . When  $\beta = 1$ ,  $V_\beta \sim \exp(1)$ .*

*Then  $X = S_\alpha V_\beta^{\beta/\alpha}$  has characteristic function  $\varphi(t) = 1/(1 + |t|^\alpha)^{1/\beta}$ .*

*Proof.* The characteristic function of  $X = S_\alpha V_\beta^{\beta/\alpha}$  is

$$\begin{aligned} \mathbb{E}(e^{itX}) &= \mathbb{E}(e^{itS_\alpha V_\beta^{\beta/\alpha}}) \\ &= \mathbb{E}(e^{-V_\beta^\beta |t|^\alpha}) \\ &= \int_0^\infty \frac{\exp(-v^\beta)}{\Gamma(1 + 1/\beta)} e^{-v^\beta |t|^\alpha} dv \\ &= \frac{1}{(1 + |t|^\alpha)^{1/\beta}}. \end{aligned}$$

□

Let  $\alpha = 2$  in the above proposition, the GCD family reduces to the Cauchy distribution family, which is the characteristic function of  $y = \sqrt{2}aUV^{\beta/2}$ ,

$$C(u|a, \beta) = \mathbb{E}(e^{-a^2 u^2 V^\beta}), \tag{3.5}$$

where  $U \sim N(0, 1)$ ,  $V_\beta$  with density  $\exp(-\nu^\beta)/\Gamma(1 + 1/\beta)$ ,  $\nu > 0$  and they are independent with each other. Thus Cauchy distribution function is a valid one-dimension spatial covariance function.

The heavy tails of Cauchy distributions may lead to the long-range dependence (LRD) property of Cauchy correlation models. Long-range dependence, also called long memory or long-range persistence, is a phenomenon that may arise in the analysis of spatial or time series data . It relates to the rate of decay of statistical dependence of two points with increasing time interval or spatial distance between the points. A phenomenon is usually considered to have long-range dependence if the dependence decays more slowly than an exponential decay, typically a power-like decay (Granger, Joyeux, 1980 [42]). In time series analysis, the Hurst exponent  $H$  is a measure of the extent of long-range dependence [43].  $H$  takes on values from 0 to 1. A value of 0.5 indicates the absence of long-range dependence. The closer  $H$  is to 1, the greater the degree of persistence or long-range dependence (Beran, 1994, p.34 [44]) .

Many past studies concluded that some geographical processes, including in weather, show long range dependence, excessive temporal dependence in widely separated events. For example, William Hurst showed in a classic 1951 [43] study that Nile river levels showed unexpected correlation over long time intervals, providing valuable practical information for the development of reservoir sizes. As for climate processes, a millennial-scale climate model (Franke et al., 2018 [45]) suggests that LRD rarely exists on the land regions of the earth. However there are rich LRD processes at the ocean regions such as annual average temperature processes in the north Atlantic ocean, elevated-CO2 processes in the southern oceans, etc.

**Definition 3.2.1.** (*Stationary long-range dependence process*) *A stationary process  $X$  has the long-range dependence property, if for its autocorrelation function holds:*

$$\int_{-\infty}^{\infty} \rho(u)du = \infty, u \in \mathbb{R}. \quad (3.6)$$

The Cauchy correlation model

$$C(u|a, \beta) = \frac{1}{(1 + a^2u^2)^{1/\beta}}, u \in \mathbb{R}, a, \beta > 0. \quad (3.7)$$

has long-range dependence property when  $\beta \geq 2$ . However, the Matérn correlation model

$$M_d(\mathbf{h}|\nu, a) = \frac{(a\|\mathbf{h}\|)^\nu}{2^{\nu-1}\Gamma(\nu)} K_\nu(a\|\mathbf{h}\|), \quad \mathbf{h} \in \mathbb{R}^d, \quad a, \nu > 0 \quad (3.8)$$

does not have long-range dependence property no matter what scale and smoothness parameters it has.

### 3.3 Cressie and Huang's Model

In spatial statistics, a space-time covariance function  $C(\mathbf{s}, t)$  is separable if

$$C(\mathbf{s}, t) = K_1(\mathbf{s})K_2(t), \quad (3.9)$$

where  $K_1$  and  $K_2$  are covariance functions in  $\mathbb{R}^2$  and  $\mathbb{R}$ , respectively. Separable models is not recommended due to the lack of space-time interactions.

In the spatiotemporal context, a few well-known nonseparable space-time covariance models were constructed. Cressie and Huang in 1999 [15] proposed an approach to construct of non-separable space-time covariance functions based on Bochner's representation. They showed that if  $h(\boldsymbol{\omega}; u) = \rho(\boldsymbol{\omega}; u)k(\boldsymbol{\omega})$ , where the following two conditions are satisfied:

**Condition 1.** For each  $\boldsymbol{\omega} \in \mathbb{R}^d$ ,  $\rho(\boldsymbol{\omega}; \cdot)$  is a continuous autocorrelation function,

$$\int_{\mathbb{R}} \rho(\boldsymbol{\omega}; u) du < \infty, \text{ and } k(\boldsymbol{\omega}) > 0.$$

**Condition 2.**  $\int_{\mathbb{R}^d} k(\boldsymbol{\omega}) d\boldsymbol{\omega} < \infty$ .

Then  $C(\mathbf{h}; u) = \int_{\mathbb{R}^d \times \mathbb{R}} e^{i\mathbf{h}\boldsymbol{\omega}} \rho(\boldsymbol{\omega}; u) k(\boldsymbol{\omega}) d\boldsymbol{\omega}$  is a valid continuous space-time stationary covariance function on  $\mathbb{R}^d \times \mathbb{R}$ .

**Example 3.3.1.** (Cressie and Huang, 1999) Let

$$\rho(\boldsymbol{\omega}; u) = \exp\{-\|\boldsymbol{\omega}\|^2|u|^2/4\} \exp\{-\delta u^2\}; \quad \delta > 0, \quad (3.10)$$

and

$$k(\boldsymbol{\omega}) = \exp\{-c_0\|\boldsymbol{\omega}\|^2/4\}; \quad c_0 > 0. \quad (3.11)$$

Then both condition 1 and condition 2 are satisfied. From

$$C(\mathbf{h}; u) = \int_{\mathbb{R}^d \times \mathbb{R}} e^{i\mathbf{h}\boldsymbol{\omega}} \rho(\boldsymbol{\omega}; u) k(\boldsymbol{\omega}) d\boldsymbol{\omega}, \quad (3.12)$$

and Matérn (1960, p.17),

$$C(\mathbf{h}, u) \propto \frac{1}{(u^2 + c_0)^{d/2}} \exp\left\{-\frac{\|\mathbf{h}\|^2}{u^2 + c_0}\right\} \exp\{-\delta u^2\}; \quad \delta > 0, \quad (3.13)$$

is a valid space-time covariance function in  $\mathbb{R}^d \times \mathbb{R}$ . Let  $\delta \rightarrow 0$ , because the limit of a sequence of space-time covariance functions is still valid if the limit exists (Matérn 1960, p.17),

$$C(\mathbf{h}; u|\boldsymbol{\theta}) = \frac{\sigma^2}{(a^2 u^2 + 1)^{d/2}} \exp\left\{-\frac{b^2 \|\mathbf{h}\|^2}{a^2 u^2 + 1}\right\}, \quad (3.14)$$

where  $\boldsymbol{\theta} = (a, b, \sigma^2)$  is a 3-parameter nonseparable space-time stationary covariance function family. When  $d = 2$  and  $a = b = \sigma^2 = 1$ , the contour plot of the space-time covariance function is shown in Figure 3.1.

### 3.4 Gneiting's Model

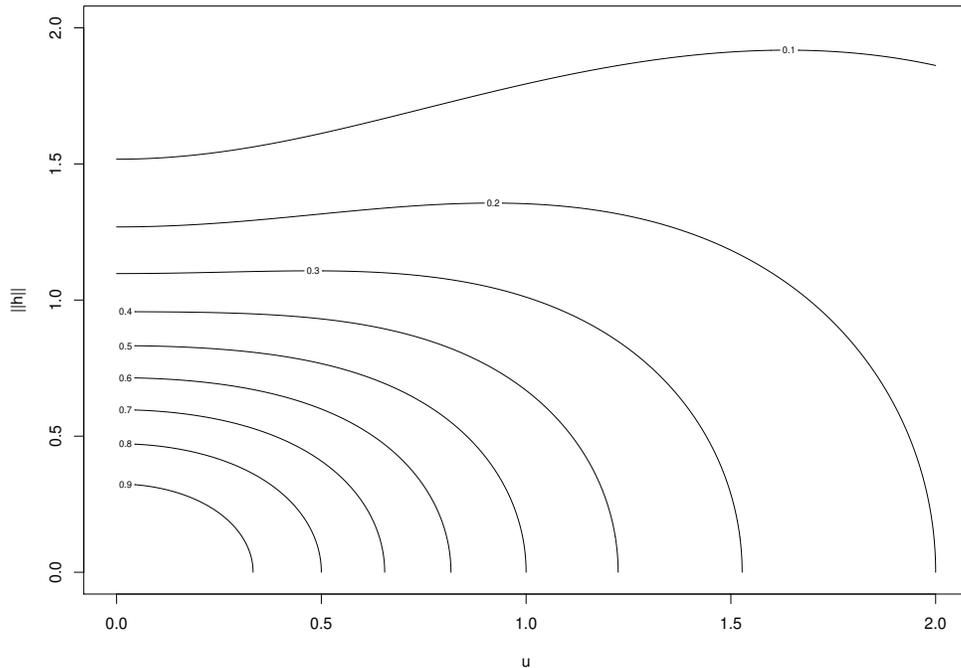
Gneiting's 2002 paper [5] made another construction of nonseparable space-time covariance functions via the spectral properties of completely monotone functions.

A continuous function  $\varphi(t)$ , defined for  $t > 0$  or  $t \geq 0$ , is said to be completely monotone if it possesses derivatives  $\varphi^{(n)}(t)$  of all orders and

$$(-1)^n \varphi^{(n)}(t) \geq 0 \quad (t > 0, n = 0, 1, 2, \dots).$$

It was shown by Bernstein's theorem (Feller 1966, p. 439) that if we let  $\varphi(t), t \geq 0$  be a completely monotone function and  $\psi(t), t \geq 0$  be any positive function with a completely monotone derivative, and let  $\sigma^2 > 0$ . Then

$$C(\mathbf{h}; u) = \frac{\sigma^2}{\psi(|u|^2)^{d/2}} \varphi\left(\frac{|\mathbf{h}|^2}{\psi(|u|^2)}\right), (\mathbf{h}; u) \in \mathbb{R}^d \times \mathbb{R} \quad (3.15)$$



**Figure 3.1.** Contour plot for  $C(\mathbf{h}; u|\boldsymbol{\theta}) = \frac{\sigma^2}{(a^2u^2+1)^{d/2}} \exp\{-\frac{b^2\|\mathbf{h}\|^2}{a^2u^2+1}\}$  as a function of  $u$  and  $\|\mathbf{h}\|$ , when  $d = 2$  and  $a = b = \sigma^2 = 1$ .

is a valid space-time covariance function.

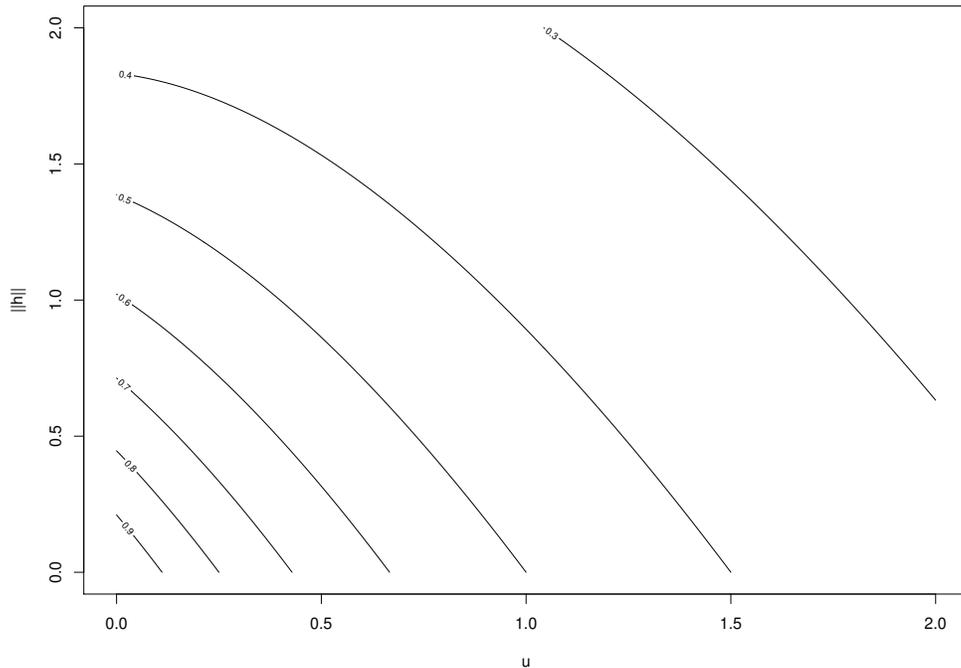
**Example 3.4.1.** (Gneiting, 2002) Let  $\varphi(t) = \exp(-ct^\gamma)$  and  $\psi(t) = (at^\alpha + 1)^\beta$  in 3.15, it leads to a parametric family of valid space-time covariance functions

$$C(\mathbf{h}; u) = \frac{\sigma^2}{(a|u|^{2\alpha} + 1)^{\beta d/2}} \exp\left(-\frac{c\|\mathbf{h}\|^{2\gamma}}{(a|u|^{2\alpha} + 1)^{\beta\gamma}}\right), (\mathbf{h}; u) \in \mathbb{R}^2 \times \mathbb{R}. \quad (3.16)$$

When  $d = 2$ ,  $\alpha = 1/2$  and  $a = \beta = c = \sigma^2 = 1$ , the contour plot of the space-time covariance function is shown in Figure 3.2.

### 3.5 Stein's Model

Stein in 2005 [7] had some discussion about the ridge issue among common space-time covariance functions. According to Proposition 1 of Stein (2005), suppose the space-time



**Figure 3.2.** Contour plot for  $C(\mathbf{h}; u) = \frac{\sigma^2}{(a|u|^{2\alpha}+1)^{\beta d/2}} \exp(-\frac{c\|\mathbf{h}\|^{2\gamma}}{(a|u|^{2\alpha}+1)^{\beta\gamma}})$  as a function of  $u$  and  $\|\mathbf{h}\|$ , when  $d = 2$ ,  $\alpha = 1/2$  and  $a = \beta = c = \sigma^2 = 1$ .

covariance function  $C(\mathbf{s}, t)$  is continuous in  $(\mathbf{s}, t) \in \mathbb{R}^3$ ,  $\alpha_1 < \dots < \alpha_p < 2$ ,  $c_1, \dots, c_p$  are even functions on  $\mathbb{R}$  with  $c_1(0) \neq 0$ , such that

$$C(\mathbf{s}, t) = C(0, 0) + \sum_{j=1}^p c_j(\mathbf{s})|t|^{\alpha_j} + R_{\mathbf{s}}(t), \quad (3.17)$$

where for any given  $\mathbf{s}$ ,  $R_{\mathbf{s}}(t)$  has a bounded second order derivative in  $t$  and  $R_{\mathbf{s}}(t) = O(t^2)$ . Let  $\rho_{\epsilon}(\mathbf{s}, t) = \text{Corr}(Y(\mathbf{0}, \epsilon) - Y(\mathbf{0}, 0), Y(\mathbf{s}, t + \epsilon) - Y(\mathbf{s}, t))$ . Then

$$\lim_{\epsilon \rightarrow 0} \rho_{\epsilon}(\mathbf{s}, t) = \begin{cases} c_1(\mathbf{s})/c_1(\mathbf{0}), & t = 0 \\ 0, & t \neq 0 \end{cases} \quad (3.18)$$

The lack of continuity of  $\rho_{\epsilon}(\mathbf{s}, t)$  along the  $\mathbf{s}$  axis leads to best linear unbiased predictors that depend on observations whose correlations are bounded away from 1 as  $\epsilon \rightarrow 0$ , implying

that correlation functions satisfying 3.17 have ridges along their axes. Stein points out that most separable correlation functions satisfy 3.17. All examples of Cressie and Huang (1999) satisfy 3.17. It would appear that nonseparable covariance functions proposed by Gneiting (2002) may not be smoother along their axes than at the origin. It was shown that one rich class of spectral densities of space-time covariance functions that may overcome the ridge issues is

$$f(\mathbf{w}) = \{c_1(a_1^2 + |\mathbf{w}_1|^2)^{\alpha_1} + c_2(a_2^2 + |\mathbf{w}_2|^2)^{\alpha_2}\}^{-\nu}. \quad (3.19)$$

for  $c_1$  and  $c_2$  positive,  $a_1^2 + a_2^2 > 0$ ,  $\alpha_1$  and  $\alpha_2$  positive integers and  $d_1/(\alpha_1\nu) + d_2/(\alpha_2\nu) < 2$ . However, this family of covariance functions may have no closed form expression.

### 3.6 Space-Time Matérn Models

The non-fully symmetric space-time (NFSST) Matérn Model was firstly proposed by T. Zhang and H. Zhang. Using the relationship between a spatial Matérn correlation function and the characteristic function of a multivariate t-distribution, they provided a way to construct a non-fully symmetric nonseparable space-time correlation function from any given marginal spatial Matérn and marginal temporal Matérn correlation functions.

**Definition 3.6.1.** *The NFSST Matérn Model. Let*

$$\mathcal{M}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2}(\mathbf{h}, u) = \mathbb{E}\left(e^{-\frac{1}{2}\left(\frac{a_1^2\|\mathbf{h}\|^2}{V_1} + \frac{2a_1a_2u\mathbf{r}^\top\mathbf{h}}{\sqrt{V_1V_2}} + \frac{a_2^2u^2}{V_2}\right)}\right), \quad (3.20)$$

where  $V_1$  and  $V_2$  are independent  $\Gamma(\nu_1, 1/2)$  and  $\Gamma(\nu_2, 1/2)$  random variables, respectively. If  $a_1, a_2, \nu_1, \nu_2 \in \mathbb{R}^+$  and  $\mathbf{r} \in \mathbb{R}^d$  with  $\|\mathbf{r}\| < 1$ , then  $\mathcal{M}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2}(\mathbf{h}, u)$  is a valid space-time correlation function. The class of space-time correlation functions

$$\mathcal{M} = \{\mathcal{M}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2} : a_1, a_2, \nu_1, \nu_2 \in \mathbb{R}^d, \|\mathbf{r}\| < 1\} \quad (3.21)$$

is called the NFSST Matérn model.

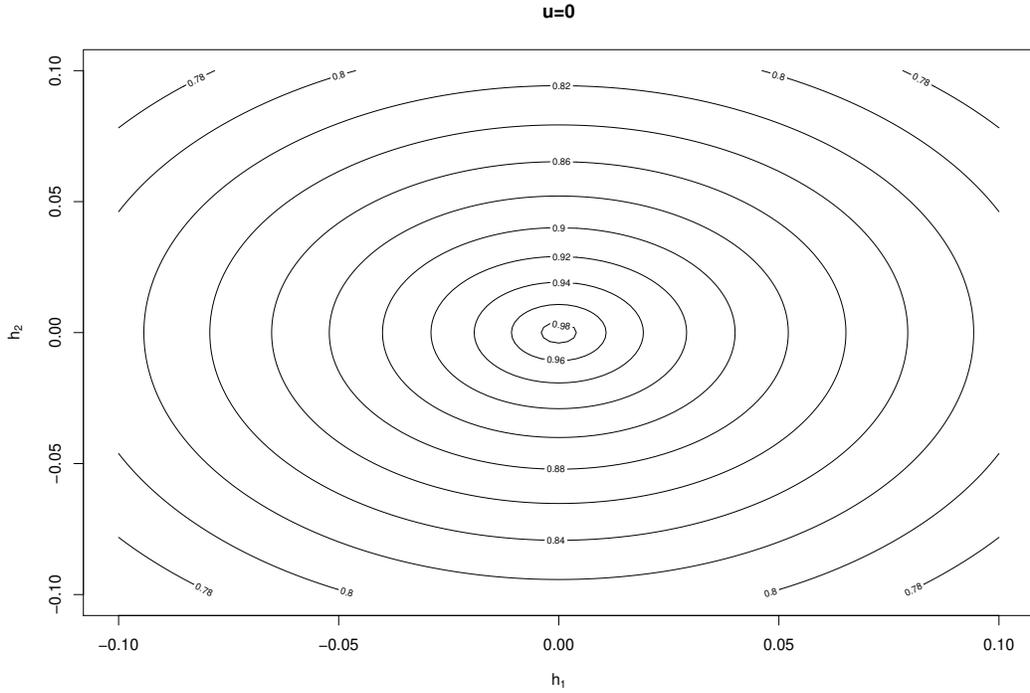
In such a space-time Matérn correlation function, the spatial margin is a spatial Matérn correlation function, and the temporal margin is a temporal Matérn correlation function.

There are 5 parameters in the NFSST Matérn model:  $(a_1, a_2, \nu_1, \nu_2, \mathbf{r})$ . They represent the scale parameter of the spatial marginal process, the scale parameter of the temporal marginal process, the smoothness parameter of the spatial marginal process, the smoothness parameter of the temporal marginal process, and the correlation coefficient between the spatial marginal process and temporal marginal process, respectively. The symmetry and separability NFSST Matérn correlation function is uniquely determined by  $\mathbf{r}$ . This correlation function can be numerically approximated using its Taylor expansion, and the MLE of parameters can be obtained using the profile likelihood functions. This model also overcomes the ridge problem mentioned in Stein's model. If  $\mathbf{r} \neq 0$  and the non-fully symmetry is satisfied,  $\mathcal{M}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2}$  does not belong to the covariance family with a covariance ridge problem considered by Proposition 1 of Stein (2005) [7]. The NFSST model has rich application in the analysis of space-time processes showing a linear space-time interaction.

**Example 3.6.1.** *Let  $d = 2$  in the NFSST Matérn model since it is important in our practical geostatitics research. Let  $\mathbf{r} = (r_1, r_2)$ , and  $\mathbf{h} = (h_1, h_2)$  when  $d = 2$ . Figure 3.3 to Figure 3.6 show the contour plots for  $\mathcal{M}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2}(\mathbf{h}, u)$  as functions of  $\mathbf{h}$  for different  $u$ , with fixed  $\mathbf{r} = (0.5, 0)$ ,  $a_1 = a_2 = 1$ , and  $\nu_1 = \nu_2 = 0.35$ . The contour plots are more symmetric for small positive  $u$  than those large positive  $u$  values. And the values of the correlation function strictly decreases in  $\|\mathbf{h}\|$  along a certain direction. The speed of the decrease depends on the direction of  $\mathbf{r}$ , which is maximized at the positive direction of  $\mathbf{r}$  and minimized at the negative direction of  $\mathbf{r}$ .*

*From Figure 3.7 to Figure 3.10, the 3d contour plots for  $\mathcal{M}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2}(\mathbf{h}, u)$  as functions of  $\mathbf{h} = (h_1, h_2)$  for different  $u$  are shown. It can be observed that as  $u$  increases, the asymmetry along  $x$  axis becomes stronger and stronger.*

*Figure 3.11 to Figure 3.14 show the contour plots for  $\mathcal{M}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2}(\mathbf{h}, u)$  as functions of  $\mathbf{h} = (h_1, h_2)$  for different  $\mathbf{r}$ , when  $u = 0.5$ ,  $a_1 = a_2 = 1$ , and  $\nu_1 = \nu_2 = 0.35$ . The contour plots in Figure 3.11 and in Figure 3.3 are actually paralleled and both of them are isotropic. In Figure 3.12 to Figure 3.14, the curves are not isotropic, and the magnitude of anisotropy increases as  $\|\mathbf{r}\|$  increases.*

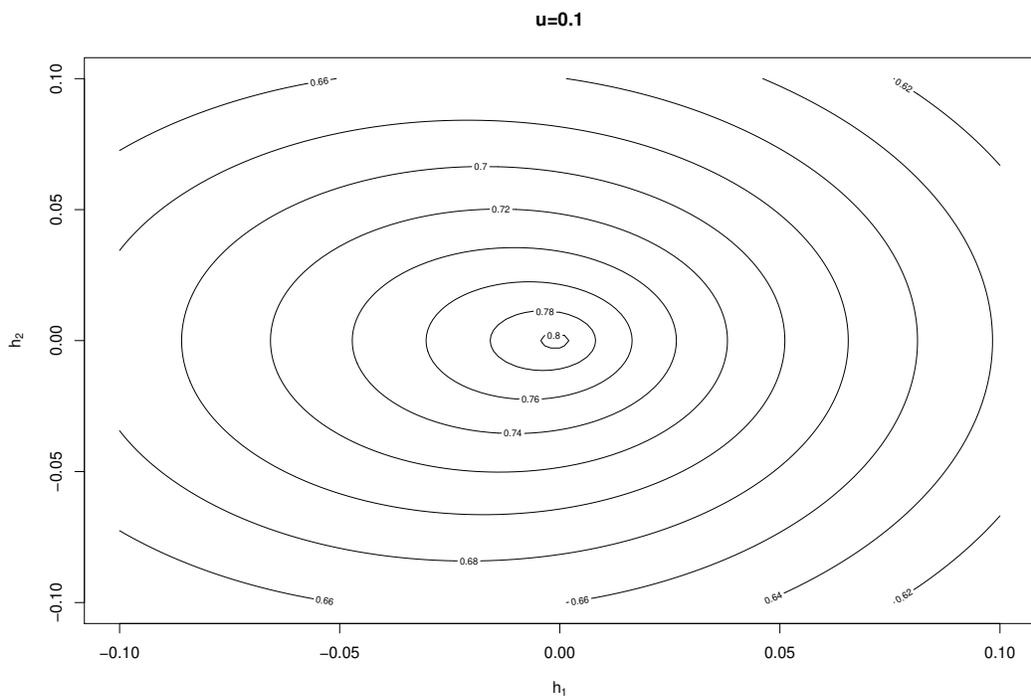


**Figure 3.3.** Contour plots for  $\mathcal{M}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$  as functions of  $\mathbf{h} = (h_1, h_2)$ , with  $u = 0$ ,  $\mathbf{r} = (0.5, 0)$ ,  $a_1 = a_2 = 1$ , and  $\nu_1 = \nu_2 = 0.35$ .

**Remark:** Both  $\mathbf{r}$  and  $\mathbf{h}$  are 2-dimensional vectors and can be expressed using their norms and angles. We can use polar coordinates to express the Taylor expansion of  $\mathcal{M}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$  as well. Let the angle of  $\mathbf{r}$  be  $\omega$  and the angle of  $\mathbf{h}$  be  $\eta$ ,  $\omega, \eta \in (-\pi, \pi]$ . Then we have  $\mathbf{r} = (r_1, r_2) = (\|\mathbf{r}\| \cos \omega, \|\mathbf{r}\| \sin \omega)$ ,  $\mathbf{h} = (h_1, h_2) = (\|\mathbf{h}\| \cos \eta, \|\mathbf{h}\| \sin \eta)$ , and  $\mathbf{r}^\top \mathbf{h} = r_1 h_1 + r_2 h_2 = \|\mathbf{r}\| \|\mathbf{h}\| \cos(\omega - \eta)$ , which yields  $\cos(-\eta) = (r_1 h_1 + r_2 h_2) / (\|\mathbf{r}\| \|\mathbf{h}\|)$ . In Figure 3.15 to Figure 3.18, we let  $\mathbf{r} = (0.5, 0)$ ,  $a_1 = a_2 = 1$ ,  $\nu_1 = \nu_2 = 0.35$  such that  $\omega = 0$  and  $\cos(\omega - \eta) = \cos \eta$ . It shows that the value of the correlation function is not symmetric about zero in  $u$  when  $\eta \neq \pi/2$ . If  $\eta = \pi/2$ , then  $\cos \eta = 0$ , implying that  $\mathcal{M}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$  is separable and fully symmetric.

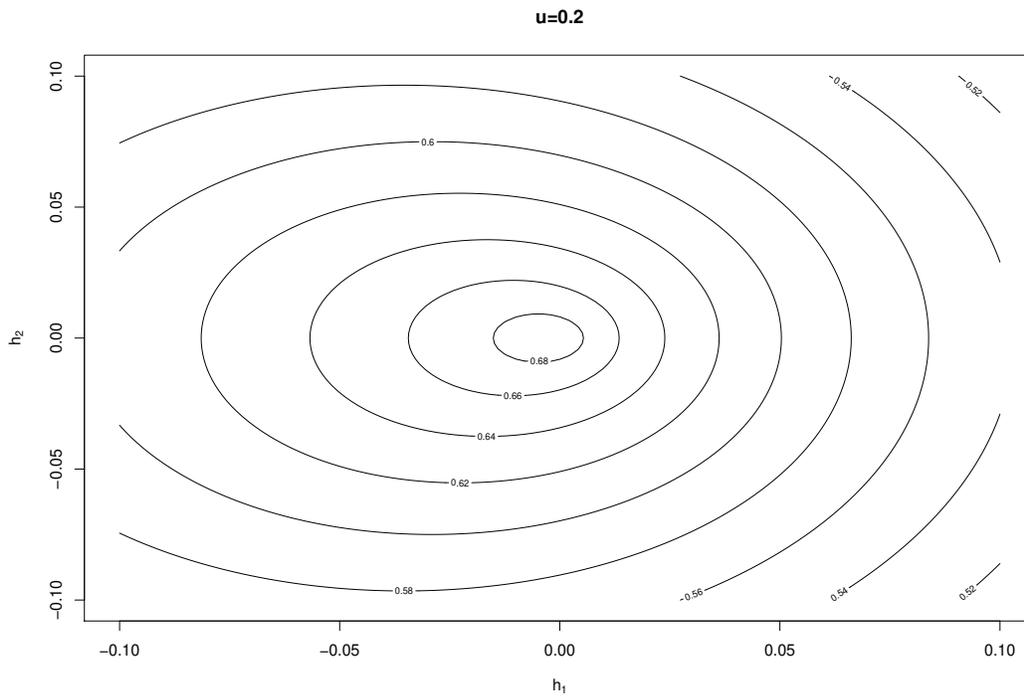
By the end, we evaluate the performance of  $\mathcal{M}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$  in the temporal domain. Let  $\mathbf{r}$  change and  $\mathbf{h} = (0.5, 0)$  be fixed, and  $a_1 = a_2 = 1$ ,  $\nu_1 = \nu_2 = 0.35$ . Figure 3.19 to Figure 3.22 show that values of  $\mathcal{M}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$  are symmetric about  $u$  if  $\eta = \pi/2$ .

$\mathcal{M}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$  has a nice property that it is symmetric about  $\mathbf{r}$  in the spatial domain. Particularly, for any  $\mathbf{h}_1, \mathbf{h}_2 \in \mathbb{R}^2$  satisfying  $\|\mathbf{h}_1\| = \|\mathbf{h}_2\|$ , if  $\mathbf{h}_1$  and  $\mathbf{h}_2$  are symmetric about

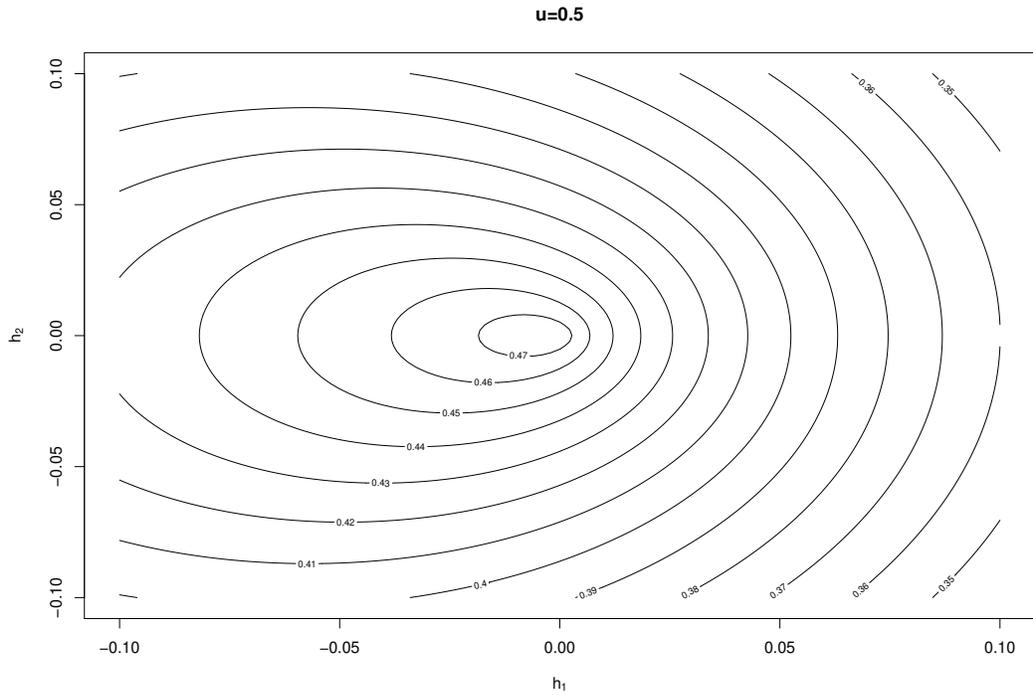


**Figure 3.4.** Contour plots for  $\mathcal{M}_{\mathbf{r},a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$  as functions of  $\mathbf{h} = (h_1, h_2)$ , with  $u = 0.1$ ,  $\mathbf{r} = (0.5, 0)$ ,  $a_1 = a_2 = 1$ , and  $\nu_1 = \nu_2 = 0.35$ .

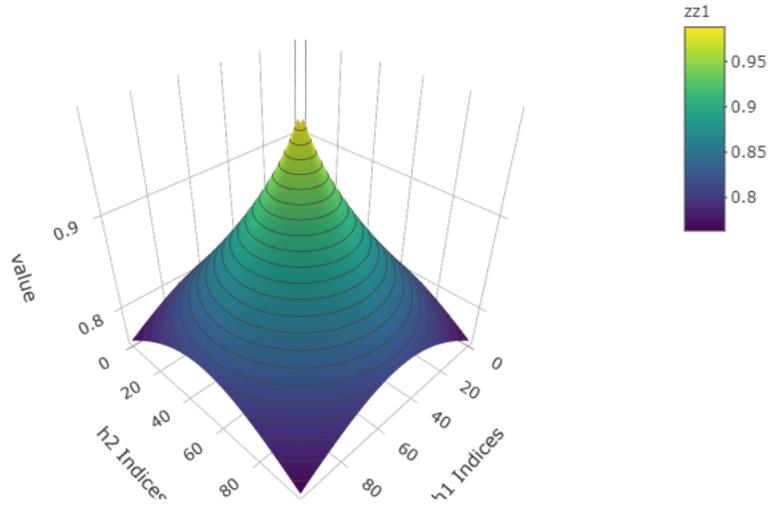
$\mathbf{r}$ , then  $\mathcal{M}_{\mathbf{r},a_1,a_2,\nu_1,\nu_2}(\mathbf{h}_1, u) = \mathcal{M}_{\mathbf{r},a_1,a_2,\nu_1,\nu_2}(\mathbf{h}_2, u)$ . This provides a nice way to interpret the NFSST Matérn model, especially in applications where data is dominated by some linear trends.



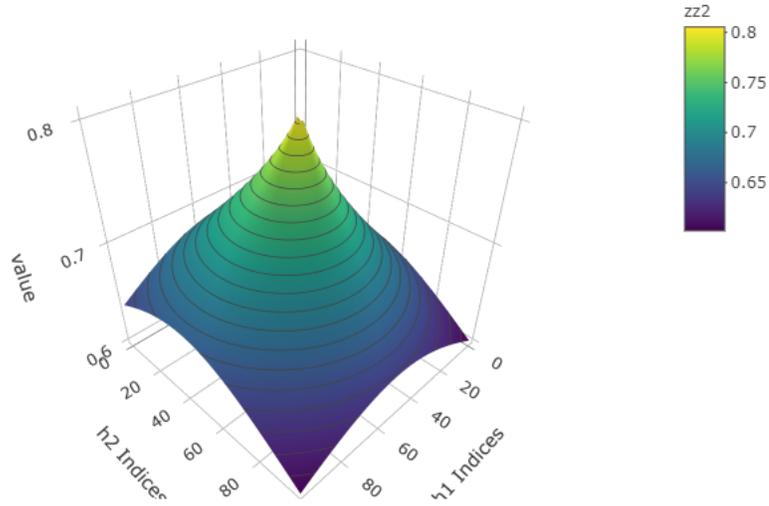
**Figure 3.5.** Contour plots for  $\mathcal{M}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2}(\mathbf{h}, u)$  as functions of  $\mathbf{h} = (h_1, h_2)$ , with  $u = 0.2$ ,  $\mathbf{r} = (0.5, 0)$ ,  $a_1 = a_2 = 1$ , and  $\nu_1 = \nu_2 = 0.35$ .



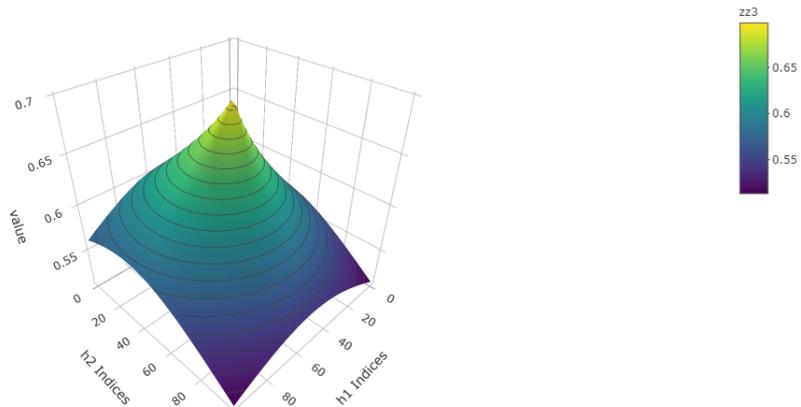
**Figure 3.6.** Contour plots for  $\mathcal{M}_{\mathbf{r},a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$  as functions of  $\mathbf{h} = (h_1, h_2)$ , with  $u = 0.5$ ,  $\mathbf{r} = (0.5, 0)$ ,  $a_1 = a_2 = 1$ , and  $\nu_1 = \nu_2 = 0.35$ .



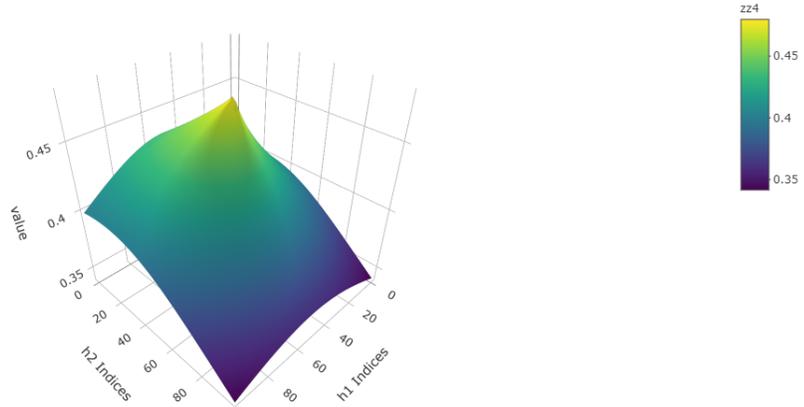
**Figure 3.7.** 3d contour plot for  $\mathcal{M}_{\mathbf{r},a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$  as functions of  $\mathbf{h} = (h_1, h_2)$ , with  $u = 0$ ,  $\mathbf{r} = (0.5, 0)$ ,  $a_1 = a_2 = 1$ , and  $\nu_1 = \nu_2 = 0.35$ .  $h_1$  varies evenly from -0.1 to 0.1, and  $h_2$  varies evenly from -0.1 to 0.1.



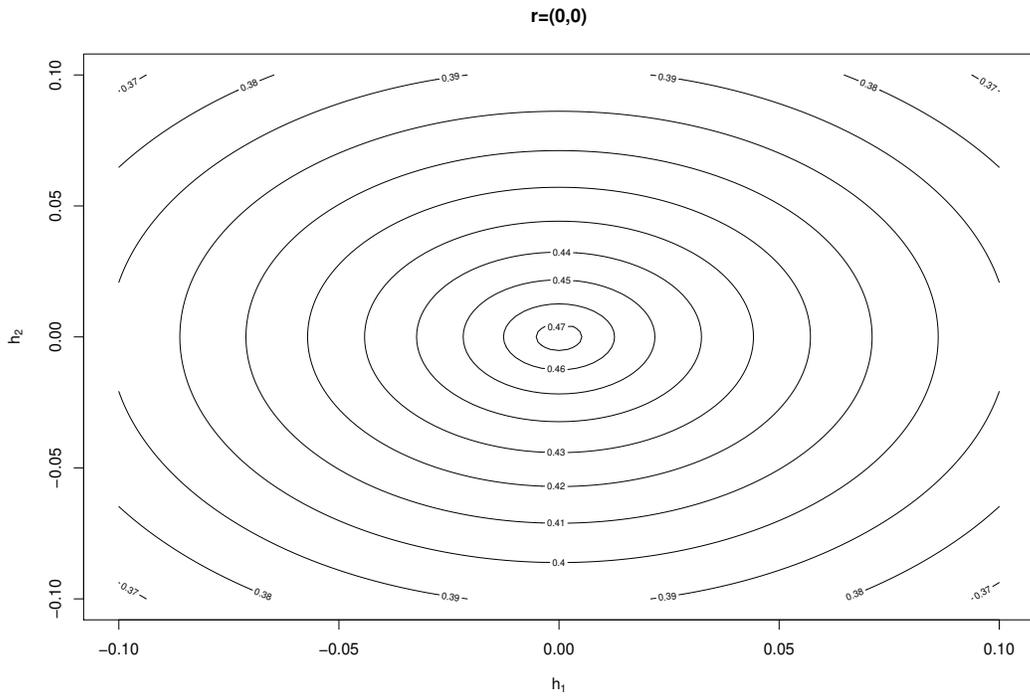
**Figure 3.8.** 3d contour plot for  $\mathcal{M}_{\mathbf{r},a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$  as functions of  $\mathbf{h} = (h_1, h_2)$ , with  $u = 0.1$ ,  $\mathbf{r} = (0.5, 0)$ ,  $a_1 = a_2 = 1$ , and  $\nu_1 = \nu_2 = 0.35$ .  $h_1$  varies evenly from -0.1 to 0.1, and  $h_2$  varies evenly from -0.1 to 0.1.



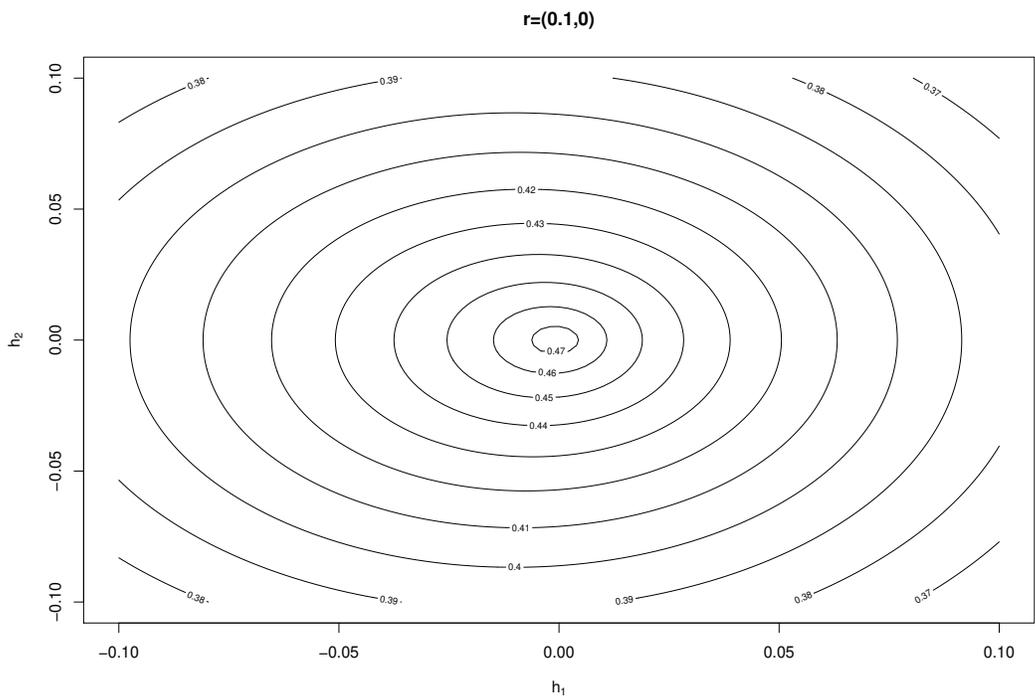
**Figure 3.9.** 3d contour plot for  $\mathcal{M}_{\mathbf{r},a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$  as functions of  $\mathbf{h} = (h_1, h_2)$ , with  $u = 0.2$ ,  $\mathbf{r} = (0.5, 0)$ ,  $a_1 = a_2 = 1$ , and  $\nu_1 = \nu_2 = 0.35$ .  $h_1$  varies evenly from -0.1 to 0.1, and  $h_2$  varies evenly from -0.1 to 0.1.



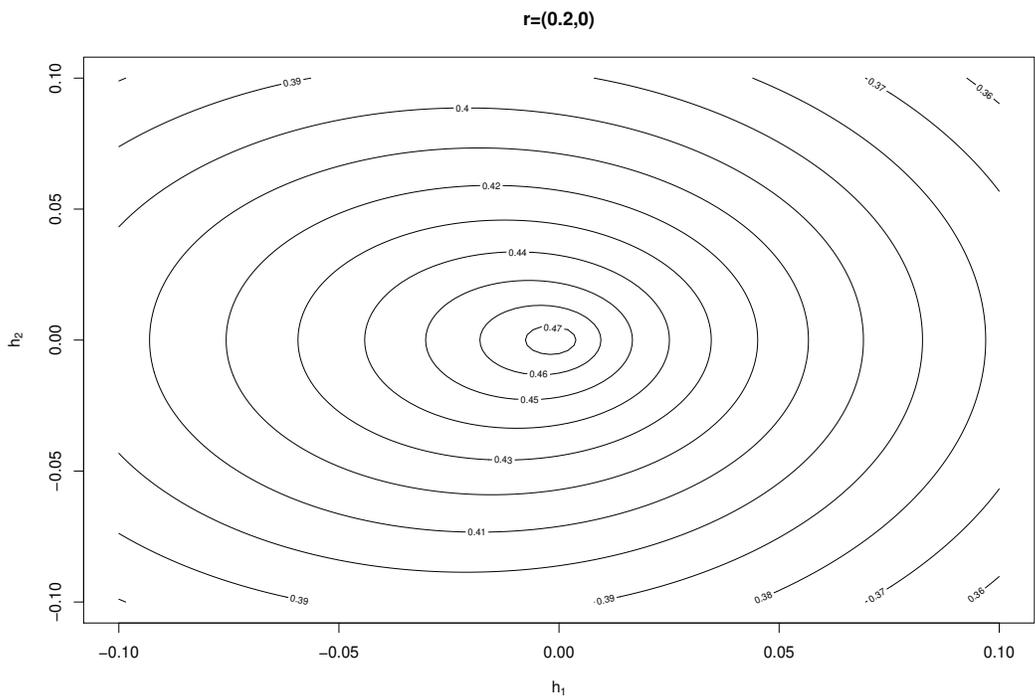
**Figure 3.10.** 3d contour plot for  $\mathcal{M}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$  as functions of  $\mathbf{h} = (h_1, h_2)$ , with  $u = 0.5$ ,  $\mathbf{r} = (0.5, 0)$ ,  $a_1 = a_2 = 1$ , and  $\nu_1 = \nu_2 = 0.35$ .  $h_1$  varies evenly from -0.1 to 0.1, and  $h_2$  varies evenly from -0.1 to 0.1.



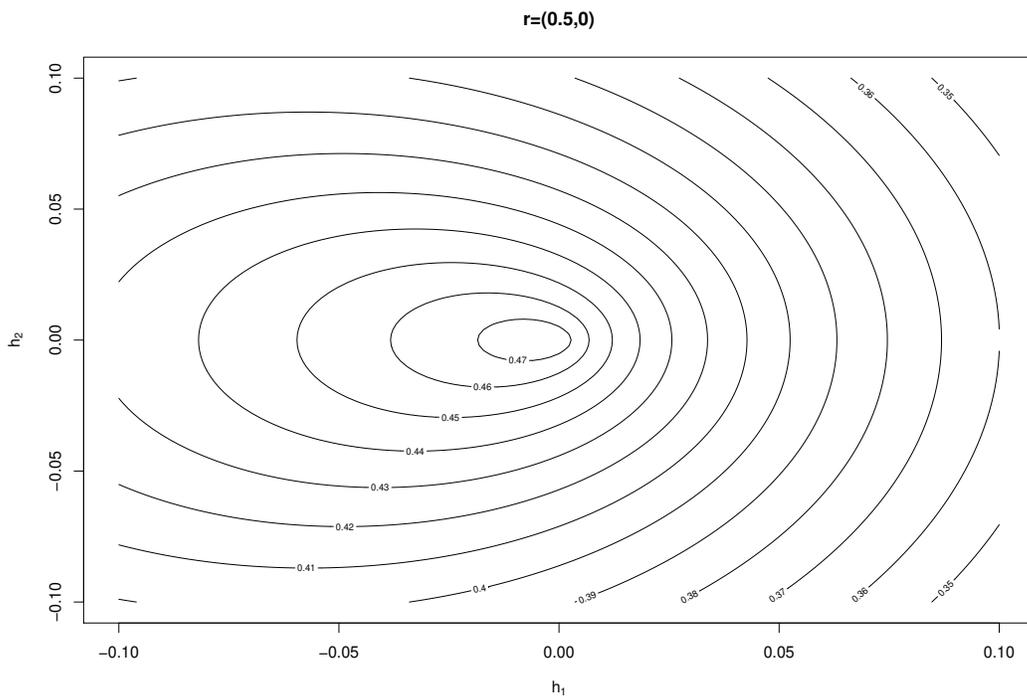
**Figure 3.11.** Contour plots for  $\mathcal{M}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$  as functions of  $\mathbf{h} = (h_1, h_2)$ , with  $\mathbf{r} = (0, 0)$ ,  $u = 0.5$ ,  $a_1 = a_2 = 1$ , and  $\nu_1 = \nu_2 = 0.35$



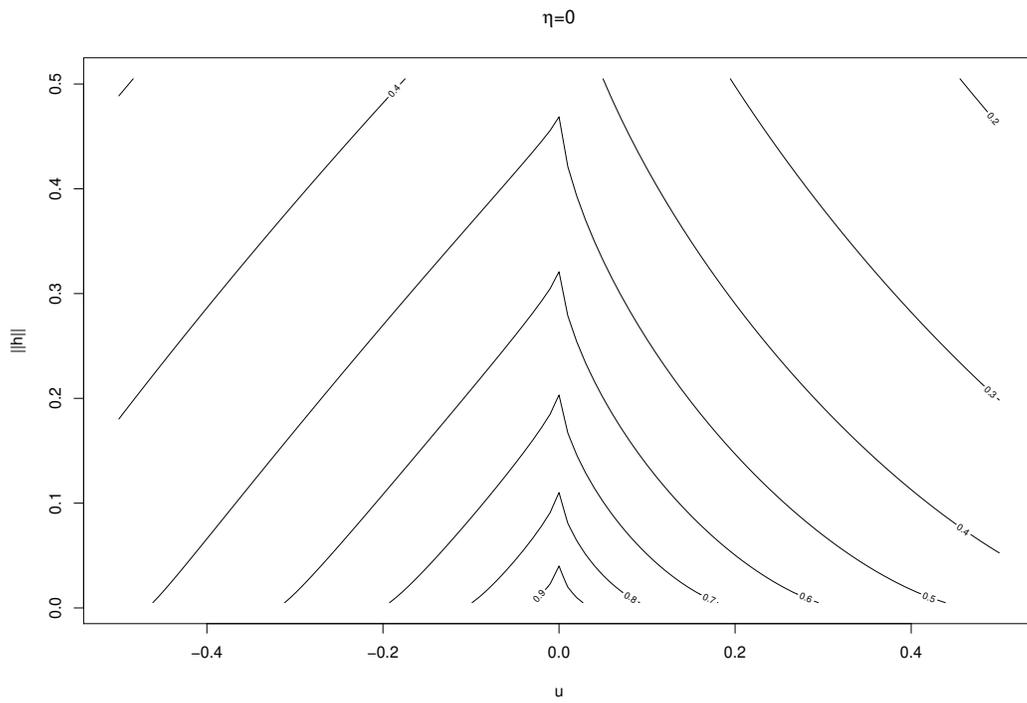
**Figure 3.12.** Contour plots for  $\mathcal{M}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2}(\mathbf{h}, u)$  as functions of  $\mathbf{h} = (h_1, h_2)$ , with  $\mathbf{r} = (0.1, 0)$ ,  $u = 0.5$ ,  $a_1 = a_2 = 1$ , and  $\nu_1 = \nu_2 = 0.35$



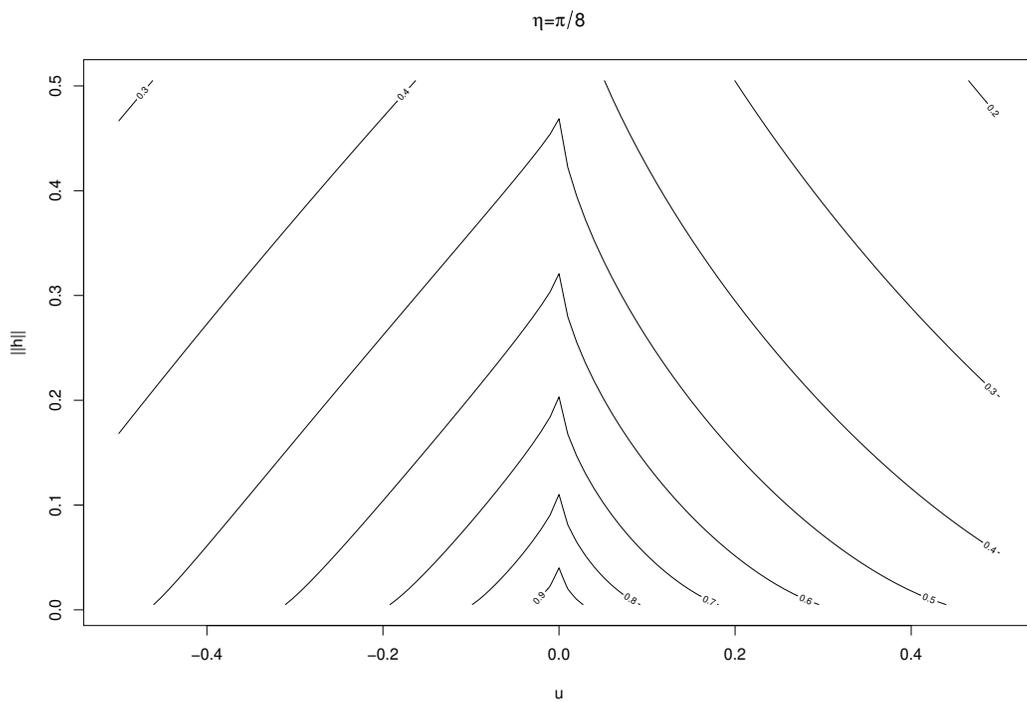
**Figure 3.13.** Contour plots for  $\mathcal{M}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2}(\mathbf{h}, u)$  as functions of  $\mathbf{h} = (h_1, h_2)$ , with  $\mathbf{r} = (0.2, 0)$ ,  $u = 0.5$ ,  $a_1 = a_2 = 1$ , and  $\nu_1 = \nu_2 = 0.35$



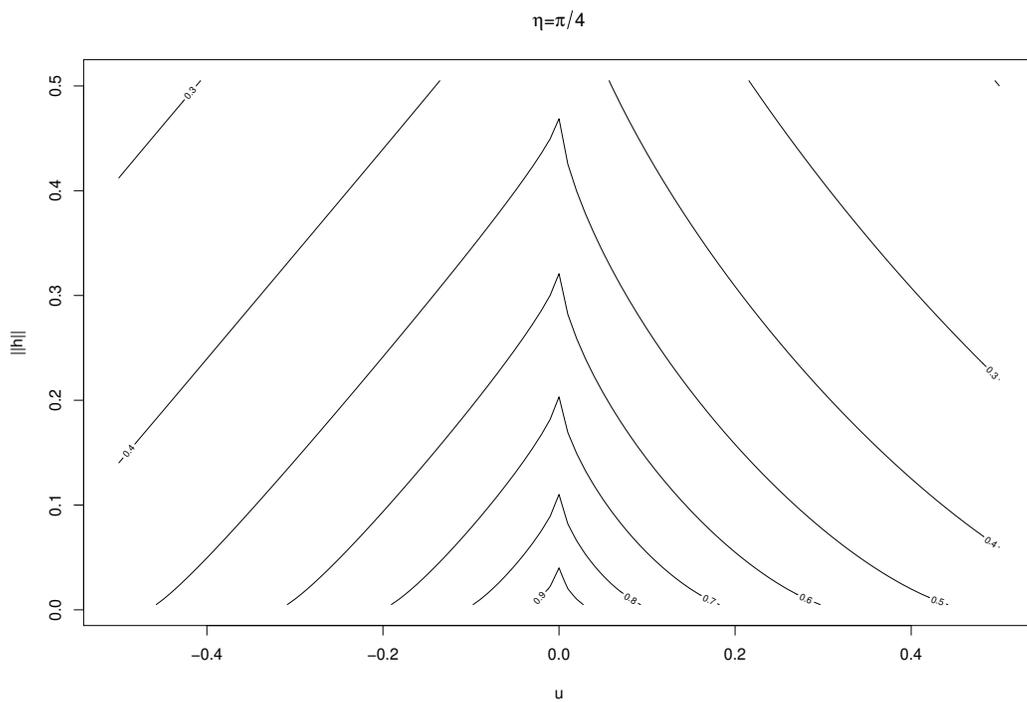
**Figure 3.14.** Contour plots for  $\mathcal{M}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$  as functions of  $\mathbf{h} = (h_1, h_2)$ , with  $\mathbf{r} = (0.5, 0)$ ,  $u = 0.5$ ,  $a_1 = a_2 = 1$ , and  $\nu_1 = \nu_2 = 0.35$



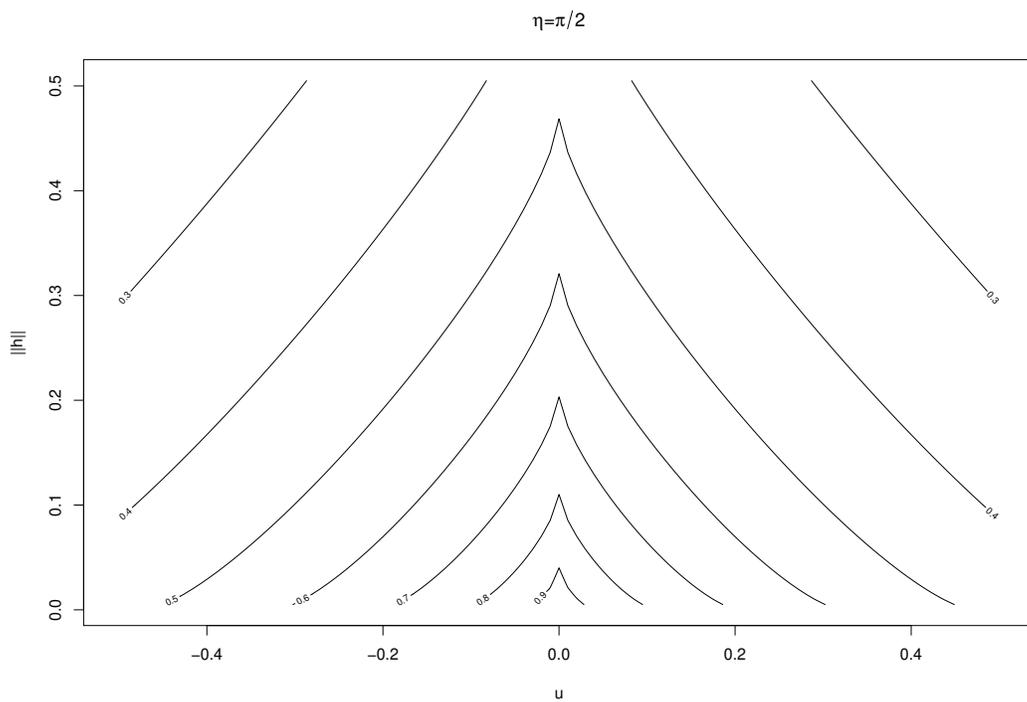
**Figure 3.15.** Contour plots for  $\mathcal{M}_{\mathbf{r},a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$  as functions of  $\|\mathbf{h}\|$  and  $u$  with  $\eta = 0$ ,  $\mathbf{r} = (0.5, 0)$ ,  $a_1 = a_2 = 1$ , and  $\nu_1 = \nu_2 = 0.35$ , where  $\eta$  is the angle between  $\mathbf{h}$  and the horizontal axis.



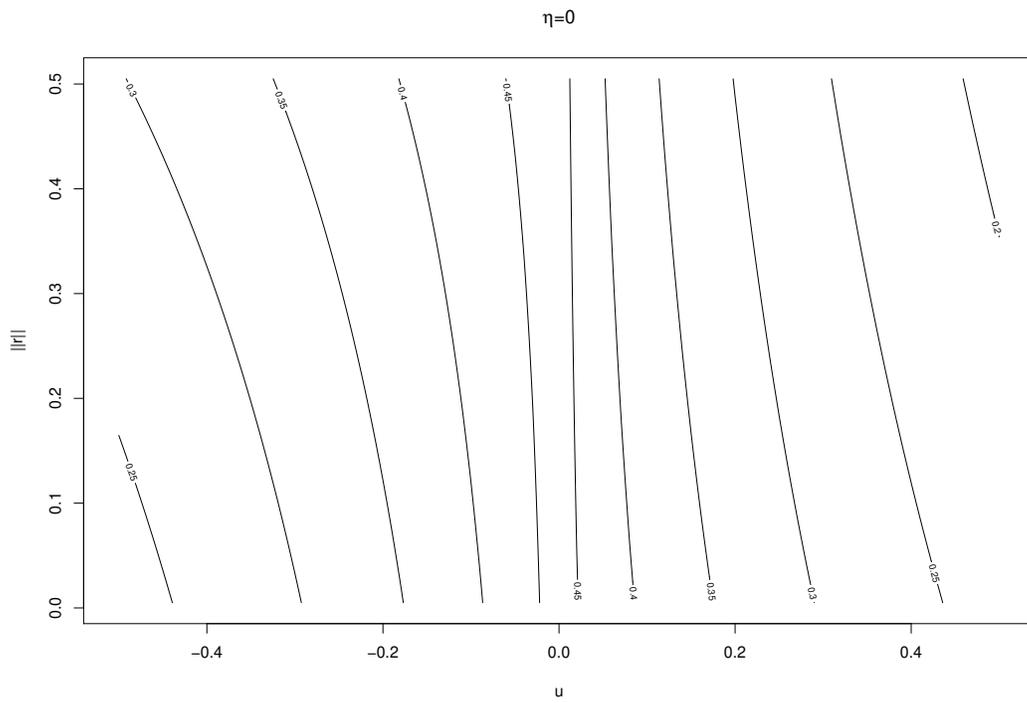
**Figure 3.16.** Contour plots for  $\mathcal{M}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2}(\mathbf{h}, u)$  as functions of  $\|\mathbf{h}\|$  and  $u$  with  $\eta = \pi/8$ ,  $\mathbf{r} = (0.5, 0)$ ,  $a_1 = a_2 = 1$ , and  $\nu_1 = \nu_2 = 0.35$ , where  $\eta$  is the angle between  $\mathbf{h}$  and the horizontal axis.



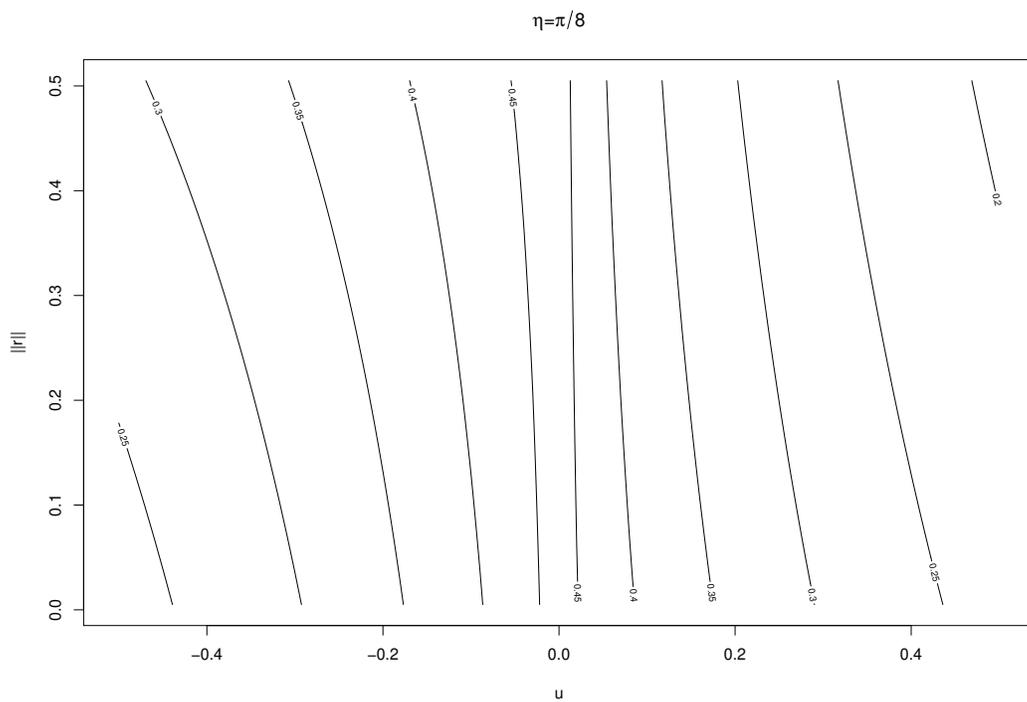
**Figure 3.17.** Contour plots for  $\mathcal{M}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2}(\mathbf{h}, u)$  as functions of  $\|\mathbf{h}\|$  and  $u$  with  $\eta = \pi/4$ ,  $\mathbf{r} = (0.5, 0)$ ,  $a_1 = a_2 = 1$ , and  $\nu_1 = \nu_2 = 0.35$ , where  $\eta$  is the angle between  $\mathbf{h}$  and the horizontal axis.



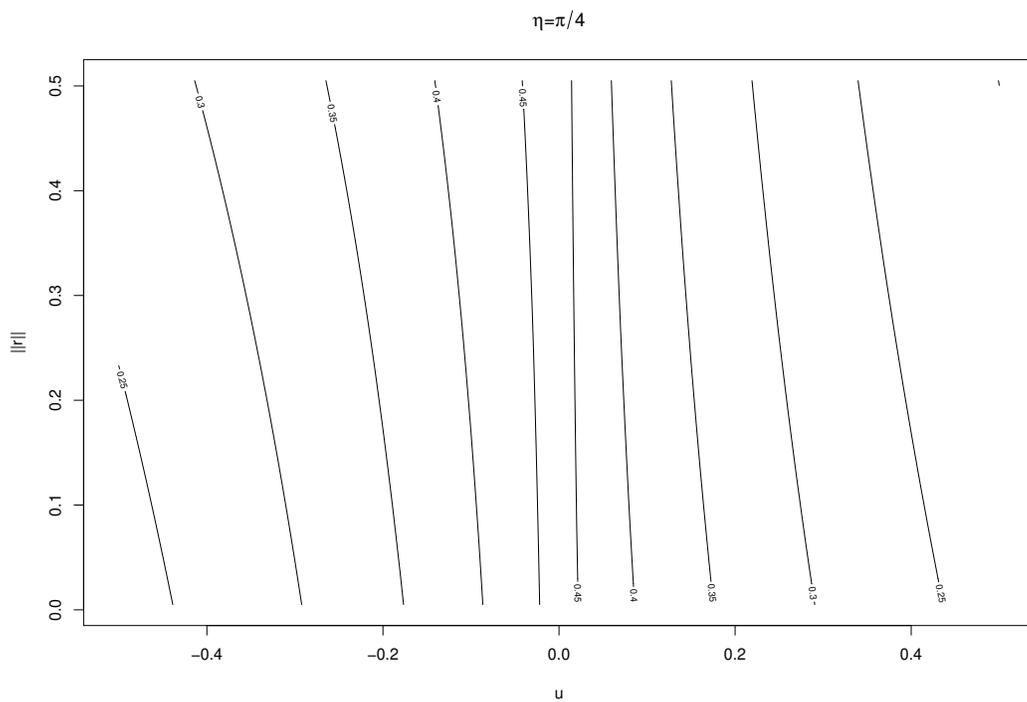
**Figure 3.18.** Contour plots for  $\mathcal{M}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2}(\mathbf{h}, u)$  as functions of  $\|\mathbf{h}\|$  and  $u$  with  $\eta = \pi/2$ ,  $\mathbf{r} = (0.5, 0)$ ,  $a_1 = a_2 = 1$ , and  $\nu_1 = \nu_2 = 0.35$ , where  $\eta$  is the angle between  $\mathbf{h}$  and the horizontal axis.



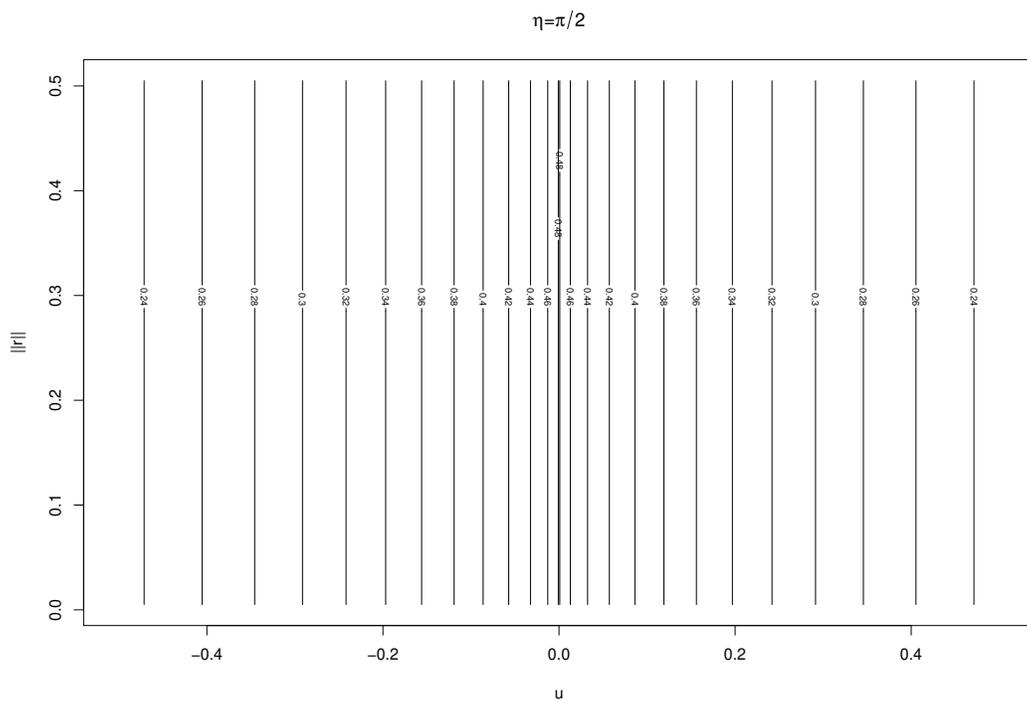
**Figure 3.19.** Contour plots for  $\mathcal{M}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$  as functions of  $\|\mathbf{r}\|$  and  $u$ , with  $\eta = 0$ ,  $\mathbf{h} = (0.5, 0)$ ,  $a_1 = a_2 = 1$ , and  $\nu_1 = \nu_2 = 0.35$ , where  $\eta$  is the angle between  $\mathbf{r}$  and the horizontal axis.



**Figure 3.20.** Contour plots for  $\mathcal{M}_{\mathbf{r},a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$  as functions of  $\|\mathbf{r}\|$  and  $u$ , with  $\eta = \pi/8$ ,  $\mathbf{h} = (0.5, 0)$ ,  $a_1 = a_2 = 1$ , and  $\nu_1 = \nu_2 = 0.35$ , where  $\eta$  is the angle between  $\mathbf{r}$  and the horizontal axis.



**Figure 3.21.** Contour plots for  $\mathcal{M}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$  as functions of  $\|\mathbf{r}\|$  and  $u$ , with  $\eta = \pi/4$ ,  $\mathbf{h} = (0.5, 0)$ ,  $a_1 = a_2 = 1$ , and  $\nu_1 = \nu_2 = 0.35$ , where  $\eta$  is the angle between  $\mathbf{r}$  and the horizontal axis.



**Figure 3.22.** Contour plots for  $\mathcal{M}_{\mathbf{r},a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$  as functions of  $\|\mathbf{r}\|$  and  $u$ , with  $\eta = \pi/2$ ,  $\mathbf{h} = (0.5, 0)$ ,  $a_1 = a_2 = 1$ , and  $\nu_1 = \nu_2 = 0.35$ , where  $\eta$  is the angle between  $\mathbf{r}$  and the horizontal axis.

## 4. Matérn-Cauchy Models

In this chapter we introduce the construction, properties, and numerical computation of our non-fully symmetric space-time Matérn-Cauchy models. Our main idea is to provide special  $A(\mathbf{h}, u)$ ,  $Z_1$ , and  $Z_2$  in Theorem 2.2.2 such that  $C(\mathbf{h}, u)$  is non-fully symmetric and satisfies 1.16. According to Theorem 2.2.1, if  $A(\mathbf{h}, u)$  is non-fully symmetric, then  $C(\mathbf{h}, u)$  is also non-fully symmetric. Therefore, we need to provide a non-fully symmetric  $A(\mathbf{h}, u)$  in the construction. As for the marginal properties of  $C(\mathbf{h}, u)$ , it is ideal that its spatial margin correlation function belongs to the Matérn family since the rich applications this family has been used in. However, we want to choose the Cauchy family for the temporal margin correlation function of  $C(\mathbf{h}, u)$ , since it may cover the circumstances of temporal long-range dependence. The way to choose  $A(\mathbf{h}, u)$ ,  $Z_1$ , and  $Z_2$  is motivated from the following.

### 4.1 Construction of Non-Fully Symmetric Space-Time Matérn-Cauchy Models

If  $\mathbf{u}_1 \sim \mathcal{N}(0, \mathbf{I}_d)$  and  $U_2 \sim \mathcal{N}(0, 1)$  be standard normal distributions and they are independent with each other, then the characteristic function of  $(\mathbf{u}_1, U_2)$  is  $E(e^{i(\mathbf{h}^\top \mathbf{u}_1 + u U_2)}) = e^{-\frac{1}{2}(\|\mathbf{h}\|^2 + u^2)}$ . Let  $V_1 \sim \Gamma(\nu_1, 1/2)$  and  $V_2$  has density  $\exp(-v^{\nu_2})/\Gamma(1+1/\nu_2)$ ,  $v > 0$  independently, which are also independent of  $(\mathbf{u}_1, U_2)$ . Then, according to Section 3.1 and Section 3.2,

$$C(\mathbf{h}, u) = E(e^{-\frac{1}{2}(\frac{a_1^2 \|\mathbf{h}\|^2}{V_1} + 2a_2^2 u^2 V_2^{\nu_2})}) \quad (4.1)$$

is a valid space-time correlation function, which is separable and satisfies 1.12 and 1.16. If the dependence between  $\mathbf{u}_1$  and  $U_2$  is imposed, then an NFSST Matérn-Cauchy model is obtained. Let

$$\begin{pmatrix} \mathbf{u}_1 \\ U_2 \end{pmatrix} \sim \mathcal{N} \left[ \begin{pmatrix} \mathbf{0} \\ 0 \end{pmatrix}, \begin{pmatrix} \mathbf{I}_d & \mathbf{r} \\ \mathbf{r}^\top & 1 \end{pmatrix} \right].$$

Since the eigenvalues of the variance-covariance matrix are either 1 or  $1 - \|\mathbf{r}\|^2$ , the distribution is valid if and only if  $\|\mathbf{r}\| < 1$ . The characteristic function of  $\mathbf{u}_1$  and  $U_2$  is

$$A_{\mathbf{r}}(\mathbf{h}, u) = e^{-\frac{1}{2}(\|\mathbf{h}\|^2 + 2u\mathbf{r}^\top \mathbf{h} + u^2)}. \quad (4.2)$$

According to 4.2,  $A_{\mathbf{r}}(\mathbf{h}, u) \neq A_{\mathbf{r}}(\mathbf{h}, -u)$  unless  $u\mathbf{r}^\top \mathbf{h} = 0$ . Generally, it is non-fully symmetric if  $\|\mathbf{r}\| \neq 0$ . Applying theorem 2.2.1, we derive a new family of non-fully symmetric space-time covariance functions.

**Definition 4.1.1.** (*The NFSST Matérn-Cauchy Model*). Let

$$\mathcal{N}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2}(\mathbf{h}, u) = \mathbb{E}\left(e^{-\frac{1}{2}\left(\frac{a_1^2 \|\mathbf{h}\|^2}{V_1} + \frac{2\sqrt{2}a_1 a_2 u \mathbf{r}^\top \mathbf{h} V_2^{\nu_2/2}}{\sqrt{V_1}} + 2a_2^2 u^2 V_2^{\nu_2}\right)}\right), \quad (4.3)$$

where  $V_1$  and  $V_2$  are independent random variables,  $V_1 \sim \Gamma(\nu_1, 1/2)$  and  $V_2$  has density  $\exp(-v^{\nu_2})/\Gamma(1 + 1/\nu_2)$ ,  $v > 0$ , respectively, If  $a_1, a_2, \nu_1, \nu_2 \in \mathbb{R}^+$  and  $\mathbf{r} \in \mathbb{R}^d$  with  $\|\mathbf{r}\| < 1$ , then  $\mathcal{N}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2}$  is called an NFSST Matérn-Cauchy correlation function. The class of space-time correlation functions

$$\mathcal{N} = \{\mathcal{N}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2} : a_1, a_2, \nu_1, \nu_2 \in \mathbb{R}^+, \mathbf{r} \in \mathbb{R}^d, \|\mathbf{r}\| < 1\} \quad (4.4)$$

is called the NFSST Matérn-Cauchy model.

**Lemma 4.1.1.** Let  $d \geq 2$ ,  $\|\mathbf{h}\|$ ,  $\|\mathbf{r}\|$ , and  $u$  be positive in 4.3. Let  $M_d(\mathbf{h}|\nu_1, a_1)$  be a Matérn correlation function with smoothness parameter  $\nu_1$  and scale parameter  $a_1$ ,  $C(u|\nu_2, a_2)$  be a Cauchy correlation function with smoothness parameter  $\nu_2$  and scale parameter  $a_2$ . If  $\mathbf{r}^\top \mathbf{h} = 0$ , then  $\mathcal{N}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2} = M_d(\mathbf{h}|\nu_1, a_1)C(u|\nu_2, a_2)$ ; if  $\mathbf{r}^\top \mathbf{h} > 0$ , then  $M_d(\mathbf{h}|\nu_1, a_1)C(u|\nu_2, a_2) < \mathcal{N}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2}$ ; if  $\mathbf{r}^\top \mathbf{h} < 0$ , then  $M_d(\mathbf{h}|\nu_1, a_1)C(u|\nu_2, a_2) > \mathcal{N}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2}$ .

*Proof.* The conclusion is implied by comparing

$$\mathcal{N}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2}(\mathbf{h}, u) = \mathbb{E}\left[e^{-\frac{\sqrt{2}a_1 a_2 u \mathbf{r}^\top \mathbf{h} V_2^{\nu_2/2}}{\sqrt{V_1}} e^{-\frac{1}{2}\left(\frac{a_1^2 \|\mathbf{h}\|^2}{V_1} + 2a_2^2 u^2 V_2^{\nu_2}\right)}}\right]$$

with

$$M_d(\mathbf{h}|\nu_1, a_1)C(u|\nu_2, a_2) = \mathbb{E}(e^{-\frac{1}{2}(\frac{a_1^2\|\mathbf{h}\|^2}{V_1} + 2a_2^2u^2V_2^{\nu_2})}),$$

since it is drawn by looking at the sign of  $\mathbf{r}^\top \mathbf{h}$ .  $\square$

Let

$$\mathcal{D}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2}(\mathbf{h}, u) = \mathcal{N}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2}(\mathbf{h}, u) - M_d(\mathbf{h}|\nu_1, a_1)C(u|\nu_2, a_2) \quad (4.5)$$

be the difference between the NFSST Matérn-Cauchy and the separable Matérn-Cauchy correlation functions. For any  $d \geq 2$ , there is

$$\mathcal{D}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2}(\mathbf{h}, u)\mathcal{D}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2}(-\mathbf{h}, u) \leq 0 \quad (4.6)$$

and the equality holds if and only if  $\mathbf{r}^\top \mathbf{h} = 0$  or  $u = 0$ . If  $d = 1$ , then

$$\mathcal{D}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2}(\mathbf{h}, u)\mathcal{D}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2}(-\mathbf{h}, u) \leq 0 \quad (4.7)$$

and the equality holds if and only if at least one of  $u, h \in \mathbb{R}$  and  $r \in (-1, 1)$  is zero.

**Theorem 4.1.1.**  $\mathcal{N}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2}$  is fully symmetric or separable if and only if  $\mathbf{r} = \mathbf{0}$ .

*Proof.* The necessary and sufficient conditions for  $\mathcal{N}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2}$  to be separable is implied by Theorem 2. The sufficiency of full symmetry is concluded using the definition of  $\mathcal{N}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2}$ . If  $d \geq 2$  then the necessity of full symmetry is implied by 4.6. If  $d = 1$ , then the necessity of fully symmetry is implied by 4.7.  $\square$

An NFSST Matérn-Cauchy correlation function can be either separable or nonseparable. If  $\mathbf{r} \neq \mathbf{0}$ , then  $\mathcal{N}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2}(\mathbf{h}, u)$  is nonseparable in the whole space but separable in the subspace  $\{\mathbf{h} \in \mathbb{R}^d : \mathbf{r}^\top \mathbf{h} = 0\}$ . Its spatial margin is a spatial Matérn correlation function and its temporal margin is a temporal Cauchy correlation function. The two margins can be arbitrary. Given the two margins,  $\mathcal{N}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2}$  can be constructed by introducing additional

parameter  $\mathbf{r}$ , which reflects the spatial-temporal interaction. Let  $\mathbf{r} = (r_1, r_2, \dots, r_d)$ . Then,  $\mathbf{r}$  can also be expressed via a polar transformation on  $\mathbb{R}^d$  as

$$\begin{aligned} r_1 &= \|\mathbf{r}\| \cos \theta_1, \\ r_j &= \|\mathbf{r}\| \left( \prod_{k=1}^{j-1} \sin \theta_k \right) \cos \theta_j, \quad j = 2, \dots, d-1, \\ r_d &= \|\mathbf{r}\| \left( \prod_{k=1}^{d-1} \sin \theta_k \right), \end{aligned}$$

where  $\theta_1, \dots, \theta_{d-2} \in [0, \pi]$  and  $\theta_{d-1} \in [0, 2\pi]$ . One can also use  $(\|\mathbf{r}\|, \theta_1, \dots, \theta_{d-1})$  to describe nonseparability. In practice, it gives geometric interpretation about the nonseparability.

## 4.2 Computation

As there is no closed-form expression for 4.3, we decided to provide a Taylor expansion for  $\mathcal{N}_{r, a_1, a_2, \nu_1, \nu_2}$  such that we can compute its numerical values. The basic way is to solve the Taylor expansion of the right hand side of 4.3 with positive  $\|\mathbf{h}\|, \|\mathbf{r}\|$ , and  $|u|$  as

$$\begin{aligned} \mathcal{N}_{r, a_1, a_2, \nu_1, \nu_2} &= \sum_{n=0}^{\infty} \frac{(-1)^n (\sqrt{2} a_1 a_2 u \mathbf{r}^\top \mathbf{h})^n}{n!} \mathbb{E}(V_1^{-n/2} V_2^{n\nu_2/2} e^{-\frac{1}{2}(\frac{a_1^2 \|\mathbf{h}\|^2}{V_1} + 2a_2^2 u^2 V_2^{\nu_2})}) \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n (\sqrt{2} a_1 a_2 u \mathbf{r}^\top \mathbf{h})^n}{n!} \mathbb{E}(V_1^{-n/2} e^{-\frac{1}{2}(\frac{a_1^2 \|\mathbf{h}\|^2}{V_1})}) \mathbb{E}(V_2^{n\nu_2/2} e^{-a_2^2 u^2 V_2^{\nu_2}}) \quad (4.8) \\ &= \sum_{n=0}^{\infty} b_{n, r, a_1, a_2, \nu_1, \nu_2}(\mathbf{h}, u), \end{aligned}$$

where

$$b_{n, r, a_1, a_2, \nu_1, \nu_2}(\mathbf{h}, u) = \frac{(-1)^n (\sqrt{2} a_1 a_2 u \mathbf{r}^\top \mathbf{h})^n}{2^\nu \Gamma(\nu) \Gamma(1 + \frac{1}{\nu_2}) n!} D_{\nu - \frac{n}{2}}(a_1 \|\mathbf{h}\|) G_{\nu_2, \frac{n}{2}}(a_2 |u|), \quad (4.9)$$

and

$$D_\alpha(z) = \int_0^\infty v^{\alpha-1} e^{-\frac{1}{2}(\frac{z^2}{v} + v)} dv, \quad \alpha \in \mathbb{R}, \quad (4.10)$$

and

$$G_{\nu_2, \tau}(z) = \int_0^\infty v^{\nu_2 \tau} e^{-v^{\nu_2}(1+z^2)} dv, \quad \nu_2, \tau \in \mathbb{R}. \quad (4.11)$$

The first and second equation in 4.8 is trivial because of the independence of  $V_1$  and  $V_2$ . For the third equation, there is:

$$\begin{aligned} \mathbb{E}(V_1^{-\frac{n}{2}} e^{-\frac{a_1^2 \|\mathbf{h}\|^2}{2V_1}}) &= \int_0^\infty v^{-\frac{n}{2}} e^{-\frac{a_1^2 \|\mathbf{h}\|^2}{2v}} \frac{1}{\Gamma(\nu)2^\nu} v^{\nu-1} e^{-\frac{v}{2}} dv \\ &= \frac{1}{\Gamma(\nu)2^\nu} \int_0^\infty v^{\nu-\frac{n}{2}-1} e^{-\frac{1}{2}(\frac{a_1 \|\mathbf{h}\|^2}{v} + v)} dv \\ &= \frac{1}{\Gamma(\nu)2^\nu} D_{\nu-\frac{n}{2}}(a_1 \|\mathbf{h}\|) \end{aligned}$$

$$\begin{aligned} \mathbb{E}(V_2^{\frac{n\nu_2}{2}} e^{-a_2^2 u^2 V_2^{\nu_2}}) &= \int_0^\infty v^{\frac{\nu_2 n}{2}} e^{-a_2^2 u^2 v^{\nu_2}} e^{-v^{\nu_2}} \frac{1}{\Gamma(1 + \frac{1}{\nu_2})} dv \\ &= \frac{1}{\Gamma(1 + \frac{1}{\nu_2})} \int_0^\infty v^{\frac{\nu_2 n}{2}} e^{-v^{\nu_2}(1+a_2^2 u^2)} dv \\ &= \frac{1}{\Gamma(1 + \frac{1}{\nu_2})} G_{\nu_2, \frac{n}{2}}(a_2 |u|) \end{aligned}$$

Thus  $\mathcal{N}_{r, a_1, a_2, \nu_1, \nu_2}$  can be represented as the summation of a series of  $b_{n, r, a_1, a_2, \nu_1, \nu_2}(\mathbf{h}, u)$ .

Remark that the difference between  $D_\alpha(z)$  and the integral in 3.3 is that we allow  $\alpha \leq 0$  in the definition of  $D_\alpha$ , which is not contained in 3.3. Values of  $D_\alpha$  when  $\alpha \leq 0$  cannot be directly obtained by Corollary 1 3.2. If  $\alpha > 0$ , then  $D_\alpha(z)$  is well-defined for any  $z \in \mathbb{R}$ . If  $\alpha \leq 0$ , then  $D_\alpha(z)$  is only well-defined for  $z \neq 0$ . In order to have a well-defined Taylor expansion in 4.8, we need an approach to compute  $D_\alpha(z)$  for any  $\alpha \leq 0$  with  $z \neq 0$  and also a way to justify the convergence rate of the Taylor expansion.

**Theorem 4.2.1.** *If  $\alpha > 0$ , then  $D_\alpha(z) = 2^\alpha \Gamma(\alpha) M_\alpha(|z|)$  for all  $z \in \mathbb{R}$ . If  $\alpha < 0$  and  $|z| > 0$ , then  $D_\alpha(|z|) = 2^{|\alpha|} \Gamma(|\alpha|) |z|^{2\alpha} M_{|\alpha|}(|z|)$ . If  $\alpha = 0$  and  $|z| > 0$ , then  $D_0(|z|) = 4|z|^{-2} [M_2(|z|) - M_1(|z|)]$ .*

*Proof.* The conclusion for  $\alpha > 0$  can be directly implied by 3.3. If  $\alpha < 0$  and  $z \neq 0$ , using variable transformation  $w = z^2/v$  in the integral expression of  $D_\alpha(z)$ , there is

$$D_\alpha(z) = |z|^{2\alpha} \int_0^\infty w^{-\alpha-1} e^{-\frac{1}{2}(\frac{z^2}{w} + w)} dw = z^{2\alpha} \int_0^\infty w^{|\alpha|-1} e^{-\frac{1}{2}(\frac{z^2}{w} + w)} dw = |z|^{2\alpha} D_{|\alpha|}(|z|), \quad (4.12)$$

implying the conclusion for  $\alpha < 0$ . For  $\alpha = 0$ , there is

$$D_0(z) = \frac{2}{z^2} \int_0^\infty v e^{-\frac{v}{2}} de^{-\frac{z^2}{2v}} = \frac{2}{z^2} \left[ \int_0^\infty \left( \frac{v}{2} - 1 \right) e^{-\frac{1}{2} \left( \frac{z^2}{v} + v \right)} dv \right] = \frac{2}{z^2} D_2(z) - \frac{1}{z^2} D_1(z).$$

□

**Corollary 4.2.1.** *If  $\alpha \leq 0$  and  $\beta > |\alpha|$  then  $|z|^{2\beta} D_\alpha(z)$  is well-defined and continuous in all  $z \in \mathbb{R}$ .*

*Proof.* If  $\alpha < 0$  then  $|z|^{2\beta} D_\alpha(z) = 2^{|\alpha|} \Gamma(|\alpha|) |z|^{2(\beta+\alpha)} M_{|\alpha|}(|z|)$  implying that it is continuous in  $z \in \mathbb{R}$ . Write

$$|z|^{2\beta} D_0(z) = |z|^{2\beta} \int_0^1 v^{-1} e^{-\frac{1}{2} \left( \frac{z^2}{v} + v \right)} dv + |z|^{2\beta} \int_1^\infty e^{-\frac{1}{2} \left( \frac{z^2}{v} + v \right)} dv. \quad (4.13)$$

The second term on the right hand side of 4.13 is continuous in  $z \in \mathbb{R}$ . For any  $\beta > \gamma > 0$ , the first term is dominated by

$$|z|^{2\beta} \int_0^\infty v^{-\gamma-1} e^{-\frac{1}{2} \left( \frac{z^2}{v} + v \right)} dz = |z|^{2\beta} D_{-\gamma}(z),$$

which is continuous in all  $z \in \mathbb{R}$ . Letting  $\gamma \rightarrow 0$  and using the Dominated Convergence Theorem, we conclude that the first term on the right hand side of 4.13 is also continuous in all  $z \in \mathbb{R}$ . □

**Remark:**  $G_{\nu_2, \tau}(z)$  can be represented using Gamma functions:

$$G_{\nu_2, \tau}(z) = \int_0^\infty v^{\nu_2 \tau} e^{-v^{\nu_2} (1+z^2)} dv. \quad (4.14)$$

Using variable transformation  $t = v^{\nu_2}$ , we have  $v = t^{\frac{1}{\nu_2}}$  and  $dv = \frac{1}{\nu_2} t^{\frac{1}{\nu_2}-1}$ . Thus

$$\begin{aligned} G_{\nu_2, \tau}(z) &= \int_0^\infty t^\tau e^{-t(1+z^2)} \frac{1}{\nu_2} t^{\frac{1}{\nu_2}-1} dt \\ &= \frac{1}{\nu_2} \int_0^\infty t^{\tau+\frac{1}{\nu_2}-1} e^{-t(1+z^2)} dt \\ &= \frac{1}{\nu_2} \Gamma\left(\tau + \frac{1}{\nu_2}\right) (1+z^2)^{-\left(\tau+\frac{1}{\nu_2}\right)}. \end{aligned} \quad (4.15)$$

Similarly, the derivative and second order derivative of  $G_{\nu_2, \tau}(z)$  can also be represented using Gamma functions:

$$\frac{dG_{\nu_2, \tau}(z)}{dz} = -\frac{2}{\nu_2} \left(\tau + \frac{1}{\nu_2}\right) \Gamma\left(\tau + \frac{1}{\nu_2}\right) z(1+z^2)^{-\left(\tau + \frac{1}{\nu_2} + 1\right)}. \quad (4.16)$$

$$\begin{aligned} \frac{d^2G_{\nu_2, \tau}(z)}{dz^2} &= -\frac{2}{\nu_2} \left(\tau + \frac{1}{\nu_2}\right) \Gamma\left(\tau + \frac{1}{\nu_2}\right) (1+z^2)^{-\left(\tau + \frac{1}{\nu_2} + 1\right)} + \\ &\quad \frac{4}{\nu_2} \left(\tau + \frac{1}{\nu_2}\right) \left(\tau + \frac{1}{\nu_2} + 1\right) \Gamma\left(\tau + \frac{1}{\nu_2}\right) z^2 (1+z^2)^{-\left(\tau + \frac{1}{\nu_2} + 2\right)}. \end{aligned}$$

Let  $\mathbf{h}_0 = \mathbf{h}/\|\mathbf{h}\|$ ,  $u_0 = u/|u|$ , and  $\mathbf{r}_0 = \mathbf{r}/\|\mathbf{r}\|$ . Then the term  $b_{n, \mathbf{r}, a_1, a_2, \nu, \nu_2}(\mathbf{h}, u)$  can be also written as:

$$b_{n, \mathbf{r}, a_1, a_2, \nu, \nu_2}(\mathbf{h}, u) = \frac{(-1)^n (\sqrt{2} u_0 \mathbf{r}_0^\top \mathbf{h}_0)^n}{2^\nu \Gamma(\nu) \Gamma(1 + \frac{1}{\nu_2}) n!} \|\mathbf{r}\|^n d_{n, \nu}(a_1 \|\mathbf{h}\|) g_{n, \nu_2}(a_2 |u|), \quad (4.17)$$

where

$$d_{n, \nu}(z) = |z|^n D_{\nu - \frac{n}{2}}(|z|), \quad z \in \mathbb{R}, \quad (4.18)$$

and

$$g_{n, \beta}(z) = |z|^n G_{\beta, \frac{n}{2}}(|z|), \quad z \in \mathbb{R}. \quad (4.19)$$

If  $2\nu_1$  is not an integer, it is enough to consider

$$d_{n, \nu}(z) = 2^{|\nu - \frac{n}{2}|} \Gamma(|\nu - \frac{n}{2}|) |z|^{2\nu \wedge n} M_{|\nu - \frac{n}{2}|}(|z|), \quad z \in \mathbb{R} \quad (4.20)$$

in the computation of 4.17; otherwise, one also needs to consider

$$d_{n, \frac{n}{2}}(z) = 4|z|^{n-2} [M_1(|z|) - M_2(|z|)], \quad z \in \mathbb{R}. \quad (4.21)$$

It is enough to consider both 4.20 and 4.21 in the computation of  $\mathcal{N}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2}$ . Also, the Taylor expansion converges exponentially fast if  $\|\mathbf{r}\| < 1$ .

**Theorem 4.2.2.** *The summation of the Taylor series*

$$\sum_{n=0}^{\infty} b_{n,r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u) = \sum_{n=0}^{\infty} \frac{(-1)^n (\sqrt{2}u_0 \mathbf{r}_0^\top \mathbf{h}_0)^n}{2^\nu \Gamma(\nu) \Gamma(1 + \frac{1}{\nu_2}) n!} \|\mathbf{r}\|^n d_{n,\nu}(a_1 \|\mathbf{h}\|) g_{n,\nu_2}(a_2 |u|) \quad (4.22)$$

absolutely uniformly converges in  $\mathbf{h} \in \mathbb{R}^d, u \in \mathbb{R}, \mathbf{r} \in \mathbb{R}^d$  exponentially fast if  $\|\mathbf{r}\| \leq 1 - \epsilon$  for any  $\epsilon \in (0, 1)$ .

*Proof.* If  $n > 2\nu_1$ , then

$$d_{\nu_1 - \frac{n}{2}}(a_1 \|\mathbf{h}\|) = 2^{\frac{n}{2} - \nu_1} \Gamma(\frac{n}{2} - \nu_1) (a_1 \|\mathbf{h}\|)^{2\nu_1} M_{\frac{n}{2} - \nu_1}(a_1 \|\mathbf{h}\|),$$

and

$$g_{\nu_2, \frac{n}{2}}(a_2 |u|) = (a_2 |u|)^{\frac{n}{2}} \frac{1}{\nu_2} \Gamma(\frac{n}{2} + \frac{1}{\nu_2}) (1 + (a_2 |u|)^2)^{-(\frac{n}{2} + \frac{1}{\nu_2})} \leq \frac{1}{\nu_2} \Gamma(\frac{n}{2} + \frac{1}{\nu_2}) (1 + (a_2 |u|)^2)^{-\frac{1}{\nu_2}}.$$

Implying that

$$b_{n,r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u) \leq \frac{(-1)^n (u_0 \mathbf{r}_0^\top \mathbf{h}_0)^n (a_1 \|\mathbf{h}\|)^{2\nu_1} [1 + (a_2 |u|)^2]^{-\frac{1}{\nu_2}} 2^n \mathbf{r}^n \Gamma(\frac{n}{2} - \nu_1) \Gamma(\frac{n}{2} + \frac{1}{\nu_2})}{2^{2\nu} \Gamma(\nu_1) \Gamma(1 + \frac{1}{\nu_2}) n!} M_{\frac{n}{2} - \nu_1}(a_1 \|\mathbf{h}\|).$$

Thus

$$b_{n,r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u) \leq \frac{(-1)^n (a_1 \|\mathbf{h}\|)^{2\nu_1} [1 + (a_2 |u|)^2]^{-\frac{1}{\nu_2}}}{2^{2\nu_1} \Gamma(\nu_1) \Gamma(1 + \frac{1}{\nu_2})} c_n,$$

where

$$c_n = \frac{2^n \mathbf{r}^n \Gamma(\frac{n}{2} - \nu_1) \Gamma(\frac{n}{2} + \frac{1}{\nu_2})}{n!}.$$

It is enough to show the uniform converges of  $\sum_{n=[2\nu_1]+1}^{\infty} c_n$  in  $r \in [0, 1 - \epsilon]$  for any  $0 < \epsilon < 1$ . Using Stirling's approximation that  $\Gamma(z) \approx \sqrt{2\pi}e^{-z}z^{z-\frac{1}{2}}(1 + o(1))$  for sufficiently large  $z$ , there is

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{c_{n+1}}{c_n} &= \lim_{n \rightarrow \infty} \frac{2r\Gamma(\frac{n+1}{2} - \nu_1)\Gamma(\frac{n+1}{2} + \frac{1}{\nu_2})}{(n+1)\Gamma(\frac{n}{2} - \nu_1)\Gamma(\frac{n}{2} + \frac{1}{\nu_2})} \\ &= \lim_{n \rightarrow \infty} \frac{2r(\frac{n+1}{2} - \nu_1)^{\frac{1}{2}}(\frac{n+1}{2} + \frac{1}{\nu_2})^{\frac{1}{2}}}{(n+1)e}. \end{aligned}$$

□

Theorem 4.2.1 provides a way to compute  $b_{n,r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$  for any given  $n$ . A numerical algorithm is obtained if 4.8 is employed. Corollary 2 concludes that  $b_{n,r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$  is continuous in all  $\mathbf{h}$  and  $u$  for any  $n$ . Theorem 4.2.2 concludes that 4.8 is valid and its right hand side is absolutely continuous in  $\mathbf{h}, u$  and  $\mathbf{r}$  is  $\|\mathbf{r}\| < 1 - \epsilon$  for any given  $\epsilon > 0$ . The expansion uniformly converges to a separable model as  $\|\mathbf{r}\| \rightarrow 0$ . The convergence rate of 4.8 is exponentially fast, therefore the algorithm based on Taylor expansion approach is efficient.

Using the Taylor expansion of  $\mathcal{N}_{r,a_1,a_2,\nu_1,\nu_2}$ , the scaling parameters  $a_1, a_2$  and the space-time interaction parameter  $\mathbf{r}$  in the correlation models can be estimated via the profile likelihood function. And the smoothness parameters  $\nu_1$  and  $\nu_2$  can be estimated using the two dimensional golden section search.

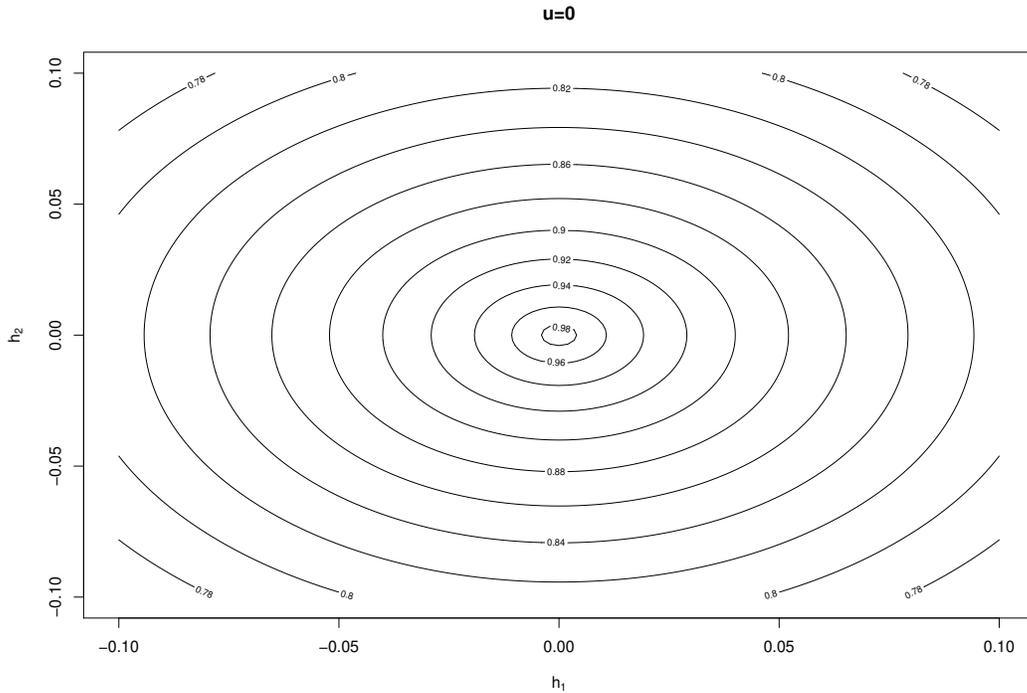
### 4.3 The Case When $d = 2$

We specify  $\mathcal{N}_{r,a_1,a_2,\nu_1,\nu_2}$  in the case when  $d = 2$ , in order to study its properties in geostatistics applications. Let  $\mathbf{r} = (r_1, r_2)$ , and  $\mathbf{h} = (h_1, h_2)$  when  $d = 2$ . If  $\|\mathbf{r}\|$  and  $\|\mathbf{h}\|$  are positive, then  $\mathbf{h}_0 = (h_{01}, h_{02}) = (h_1/\|\mathbf{h}\|, h_2/\|\mathbf{h}\|)$  and  $\mathbf{r}_0 = (r_{01}, r_{02}) = (r_1/\|\mathbf{r}\|, r_2/\|\mathbf{r}\|)$  are well defined. Equation 4.3 becomes

$$\mathcal{N}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u) = \mathbb{E}\left[e^{-\frac{\sqrt{2}a_1a_2u_0\|\mathbf{r}\|\|\mathbf{h}\|(r_{01}h_{01}+r_{02}h_{02})V_2^{\nu_2/2}}{\sqrt{V_1}} e^{-\frac{1}{2}(\frac{a_1^2\|\mathbf{h}\|^2}{V_1}+2a_2^2u^2V_2^{\nu_2})}}\right]. \quad (4.23)$$

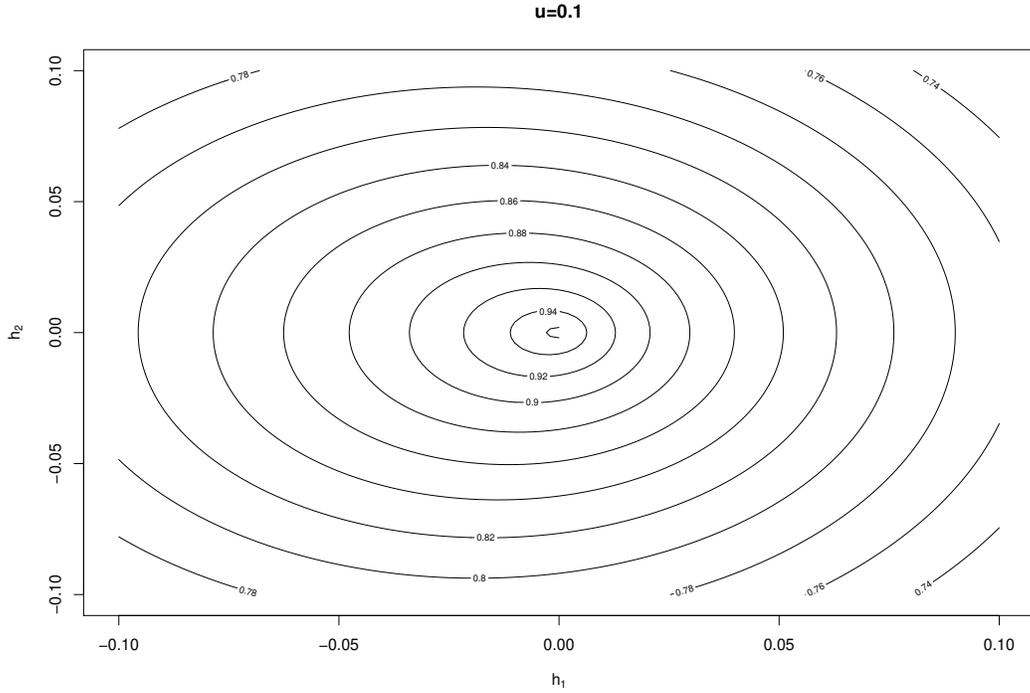
If  $u_0 > 0$ , then given  $|u|$  and  $\|\mathbf{h}\|$ ,  $\mathcal{N}_{\mathbf{r},a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$  is maximized at  $\mathbf{h}_0 = \mathbf{r}_0$  and minimized at  $\mathbf{h}_0 = -\mathbf{r}_0$ . If  $\mathbf{h}_0$  and  $\mathbf{u}_0$  are vertical, then  $\mathcal{N}_{\mathbf{r},a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u) = M_2(\mathbf{h}|\nu_1, a_1)C(u|\nu_2, a_2)$ . This is satisfied if  $r_{01}h_{01} + r_{02}h_{02} = 0$ .

To study the performance of  $\mathcal{N}$  in the spatial domain, we use 4.17 to numerically compute the values of the correlation function on the left hand side of 4.23 for selected  $u$  when  $\mathbf{r} = (0.5, 0)$ ,  $a_1 = a_2 = 1$ , and  $\nu_1 = \nu_2 = 0.35$  (Figure 4.1 to Figure 4.4). It shows that the right hand side of 4.23 reduces to  $M_2(\mathbf{h}|\nu_1, a_1)$  is  $u = 0$  (Figure 4.1). Similar to what we have seen in Figure 3.3 to Figure 3.6, the contour plots are more symmetric for small positive  $u$ , and the values of the correlation function strictly decreases in  $|u|$  or  $\|\mathbf{h}\|$  increases along a certain direction. The decreasing speed is maximized at the positive direction of  $\mathbf{r}$  and minimized at the negative direction of  $\mathbf{r}$ . These contour plots share many properties with those of the NFSST Matérn model.



**Figure 4.1.** Contour plots for  $\mathcal{N}_{\mathbf{r},a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$  as functions of  $\mathbf{h} = (h_1, h_2)$ , with  $u = 0$ ,  $\mathbf{r} = (0.5, 0)$ ,  $a_1 = a_2 = 1$ , and  $\nu_1 = \nu_2 = 0.35$ .

In Figure 4.5 to Figure 4.8, we change the smoothness parameter  $\nu_2$  in  $\mathcal{N}_{\mathbf{r},a_1,a_2,\nu_1,\nu_2}$ , from 0.35 to 3. The contour plots become denser and more similar to those in Figure 3.3 to

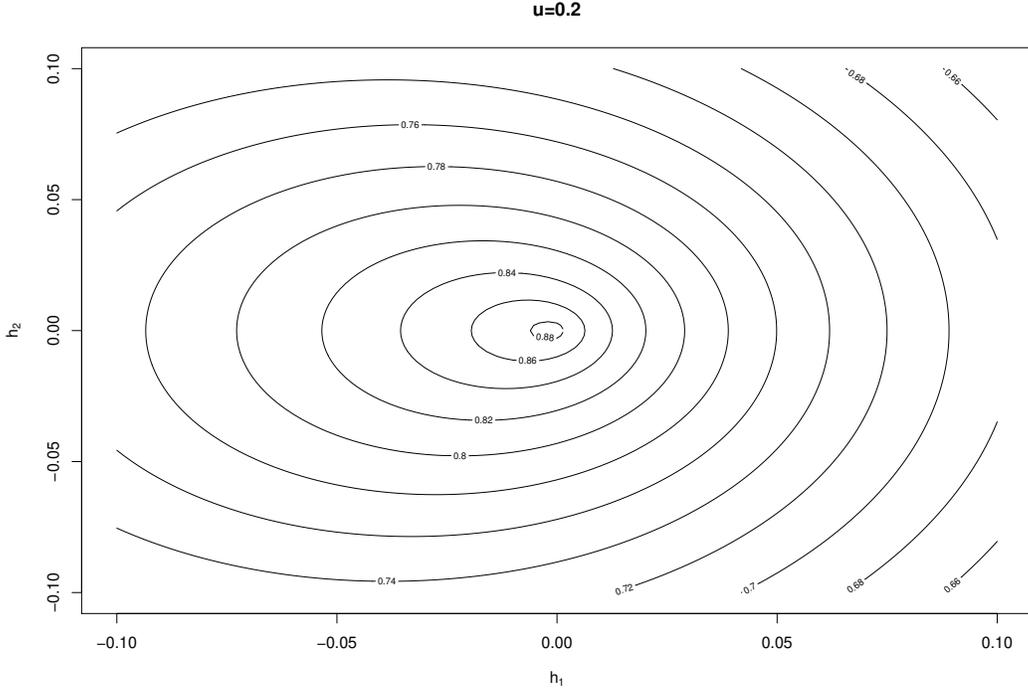


**Figure 4.2.** Contour plots for  $\mathcal{N}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2}(\mathbf{h}, u)$  as functions of  $\mathbf{h} = (h_1, h_2)$ , with  $u = 0.1$ ,  $\mathbf{r} = (0.5, 0)$ ,  $a_1 = a_2 = 1$ , and  $\nu_1 = \nu_2 = 0.35$ .

Figure 3.6. In fact, by toning the scale and smoothness parameters of the temporal Cauchy correlation function in  $\mathcal{N}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2}$ , it may have very close values with the NFSST Matérn model.

To compare, we also use 4.17 to study the performance of  $\mathcal{N}$  in the spatial domain when  $\mathbf{r}$  varies. We numerically compute the correlation function on the left hand side of 4.23 for selected  $\mathbf{r}$  when  $u = 0.5$ ,  $a_1 = a_2 = 1$ , and  $\nu_1 = \nu_2 = 0.35$  (Figure 4.9 to Figure 4.12). By theorem 4.1.1, the curves in Figure 4.9 are isotropic. In Figure 4.10) to Figure 4.12, the curves are not isotropic, and the magnitude of anisotropy increases as  $\|\mathbf{r}\|$  increases.

We can also use polar coordinates to represent the Taylor expansion of  $\mathcal{N}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2}(\mathbf{h}, u)$ . Similar to the approach in section 3.4, let the angle of  $\mathbf{r}$  be  $\sigma$  and the angle of  $\mathbf{h}$  be  $\eta$ ,  $\omega, \eta \in (-\pi, \pi]$ . Then we have  $\mathbf{r} = (r_1, r_2) = (\|\mathbf{r}\| \cos \omega, \|\mathbf{r}\| \sin \omega)$ ,  $\mathbf{h} =$



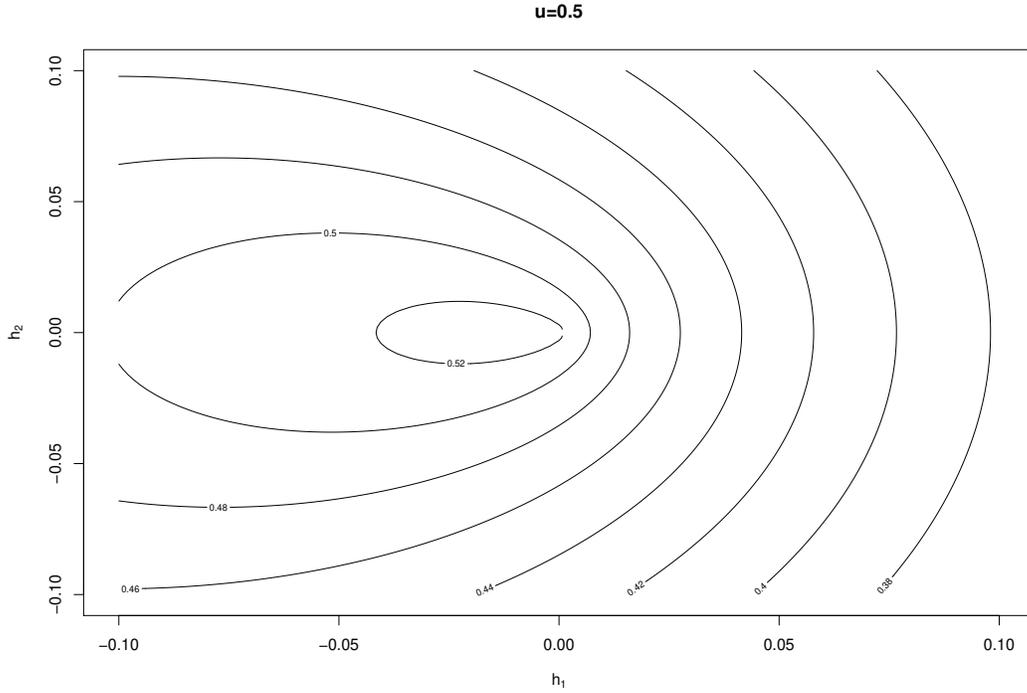
**Figure 4.3.** Contour plots for  $\mathcal{N}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2}(\mathbf{h}, u)$  as functions of  $\mathbf{h} = (h_1, h_2)$ , with  $u = 0.2$ ,  $\mathbf{r} = (0.5, 0)$ ,  $a_1 = a_2 = 1$ , and  $\nu_1 = \nu_2 = 0.35$ .

$(h_1, h_2) = (\|\mathbf{h}\| \cos \eta, \|\mathbf{h}\| \sin \eta)$ , and  $\mathbf{r}^\top \mathbf{h} = r_1 h_1 + r_2 h_2 = \|\mathbf{r}\| \|\mathbf{h}\| \cos(\omega - \eta)$ , which yields  $\cos(-\eta) = (r_1 h_1 + r_2 h_2) / (\|\mathbf{r}\| \|\mathbf{h}\|)$  and

$$b_{n, \mathbf{r}, a_1, a_2, \nu_1, \nu_2}(\mathbf{h}, u) = \frac{[-\sqrt{2}(u) \cos(\omega - \eta)]^n}{2^{\nu_1} \Gamma \nu_1 \Gamma 1 + \frac{1}{\nu_2} n!} \|\mathbf{r}\|^n d_{n, \nu_1}(a_1 \|\mathbf{h}\|) g_{n, \nu_2}(a_2 |u|). \quad (4.24)$$

If  $u$  and  $\|\mathbf{h}\|$  are positive, then  $\mathcal{N}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2}(\mathbf{h}, u) < M_2(\mathbf{h} | \nu_1, a_1) C(u | \nu_2, a_2)$  if  $|\omega - \eta| < \pi/2$ . And  $\mathcal{N}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2}(\mathbf{h}, u) > M_2(\mathbf{h} | \nu_1, a_1) C(u | \nu_2, a_2)$  if  $|\omega - \eta| > \pi/2$ . Using the polar expression of  $b_{n, \mathbf{r}, a_1, a_2, \nu_1, \nu_2}(\mathbf{h}, u)$  in 4.24, we can also interpret Figure 4.1 to Figure 4.4 according to a rotation of  $\mathbf{r}$ : if  $\mathbf{r}$  is rotated by an orthogonal transformation then the corresponding contour plot of  $\mathcal{N}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2}(\mathbf{h}, u)$  is also rotated with the same angle of the transformation.

To study the performance of  $\mathcal{N}$  in the spatiotemporal domain, we use 4.24 to numerically compute the values of  $\mathcal{N}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2}(\mathbf{h}, u)$ . We keep  $\mathbf{r} = (0.5, 0)$ ,  $a_1 = a_2 = 1$ ,  $\nu_1 = \nu_2 = 0.35$  such that  $\omega = 0$  and  $\cos(\omega - \eta) = \cos \eta$  in 4.24. As Figure 4.13 to Figure 4.16 show, the value of the correlation function is symmetric about zero in  $u$  only when  $\eta = \pi/2$ ,

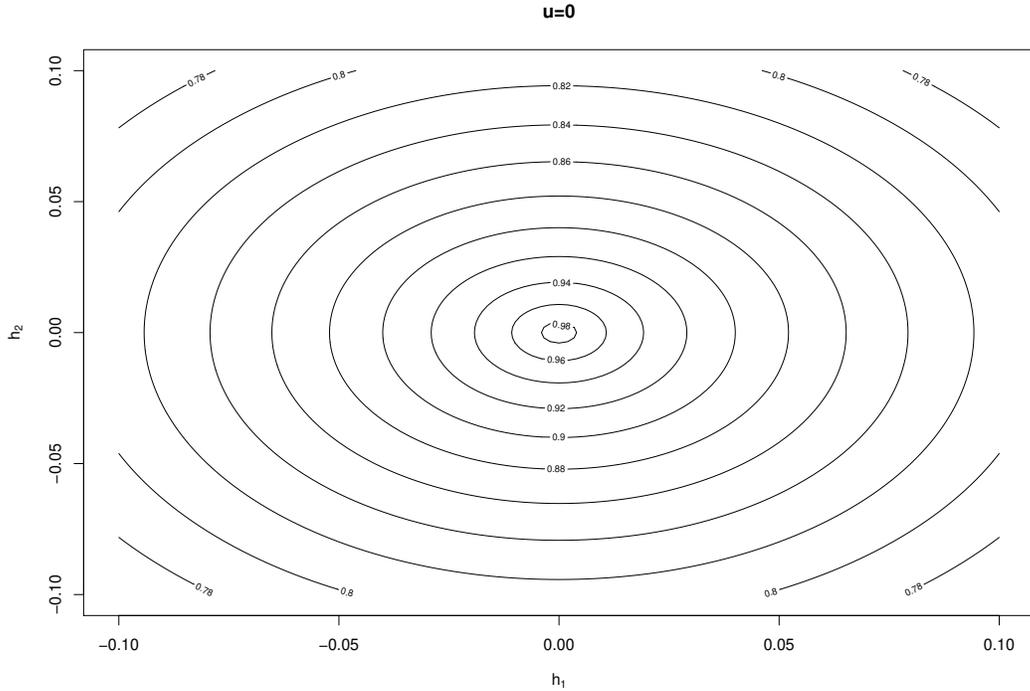


**Figure 4.4.** Contour plots for  $\mathcal{N}_{\mathbf{r},a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$  as functions of  $\mathbf{h} = (h_1, h_2)$ , with  $u = 0.5$ ,  $\mathbf{r} = (0.5, 0)$ ,  $a_1 = a_2 = 1$ , and  $\nu_1 = \nu_2 = 0.35$ .

which indicates  $\cos(\eta) = 0$  and  $\mathcal{N}_{\mathbf{r},a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$  is separable. When  $\eta \in (\pi/2, \pi]$ , because  $\cos(\eta) = -\cos(\pi - \eta)$ , we can derive the contour plots for  $\eta \in (\pi/2, \pi]$  by reflecting the plots for  $\eta \in [0, \pi/2]$ .

The performance of  $\mathcal{N}$  is evaluated in Figure 4.17 to Figure 4.20 where we let  $\mathbf{r}$  change and let  $\mathbf{h}$  be fixed at  $\mathbf{h} = (0.5, 0)$ , while the rest parameters keep unchanged. It shows that values of  $\mathcal{N}_{\mathbf{r},a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$  decreases in  $\|\mathbf{r}\|$  if  $\eta < \pi/2$  and  $u$  is positive. Values of  $\mathcal{N}_{\mathbf{r},a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$  does not change with  $\|\mathbf{r}\|$  if  $\eta = \pi/2$ , which is clear due to 4.8.

Similar as the NFSST Matérn model,  $\mathcal{N}_{\mathbf{r},a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$  is symmetric about  $\mathbf{r}$  in the spatial domain. For  $\mathbf{h}_1, \mathbf{h}_2 \in \mathbb{R}^2$  with the same mode, if they are symmetric about  $\mathbf{r}$ , then  $\mathcal{N}_{\mathbf{r},a_1,a_2,\nu_1,\nu_2}(\mathbf{h}_1, u) = \mathcal{N}_{\mathbf{r},a_1,a_2,\nu_1,\nu_2}(\mathbf{h}_2, u)$ ; otherwise,  $\mathcal{N}_{\mathbf{r},a_1,a_2,\nu_1,\nu_2}(\mathbf{h}_1, u) > \mathcal{N}_{\mathbf{r},a_1,a_2,\nu_1,\nu_2}(\mathbf{h}_2, u)$  if the angle between  $\mathbf{h}_1$  and  $\mathbf{r}$  is less than the angle between  $\mathbf{h}_2$  and  $\mathbf{r}$ . Particularly, for positive  $u$ , the space-time correlation is strongest if the direction of the space change follows the opposite direction of  $\mathbf{r}$ , and it is weakest if the direction of the space change is the same as  $\mathbf{r}$ .

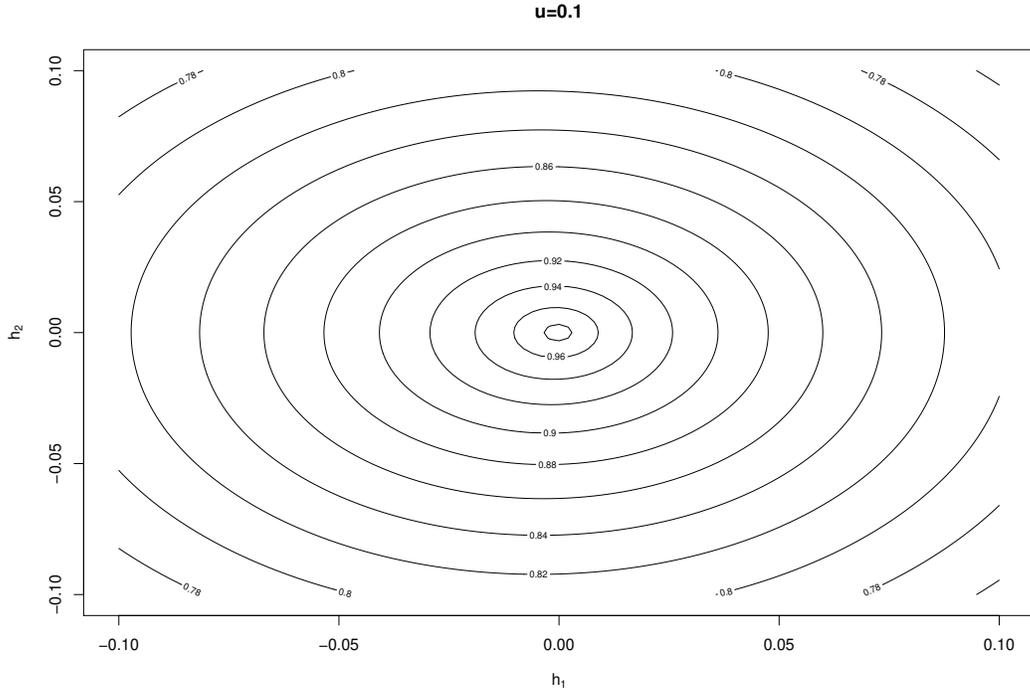


**Figure 4.5.** Contour plots for  $\mathcal{N}_{\mathbf{r},a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$  as functions of  $\mathbf{h} = (h_1, h_2)$ , with  $u = 0$ ,  $\mathbf{r} = (0.5, 0)$ ,  $a_1 = a_2 = 1$ , and  $\nu_1 = 0.35, \nu_2 = 3$ .

#### 4.4 Discussion

In this section we have some discussion on the comparison between NSFFT Matérn correlation model and NFSST Matérn-Cauchy model. Although these two models are closed related and they share many similar properties, the temporal Cauchy margin gives the model some unique advantages compared with temporal Matérn margin.

Firstly, the computation efficiency of the NSFFT Matérn-Cauchy model is much higher than that of the NSFFT Matérn model. As discussed in Section 4.2, using the Taylor expansion of  $\mathcal{N}_{\mathbf{r},a_1,a_2,\nu_1,\nu_2}$ , the scaling parameters  $a_1, a_2$  and the space-time interaction parameter  $\mathbf{r}$  in the correlation models can be estimated via maximizing the profile likelihood function. In order to maximize the profile likelihood function via Newton-Raphson method, we need to calculate the derivatives and second order derivatives of 4.17. The spatial Matérn component  $d_{n,\nu_1}(a_1\|\mathbf{h}\|)$  in 4.17 contains the modified Bessel functions of the second kind by 3.3, while the temporal Cauchy component  $g_{n,\nu_2}(a_2\|u\|)$  contains Gamma functions by 4.15. The



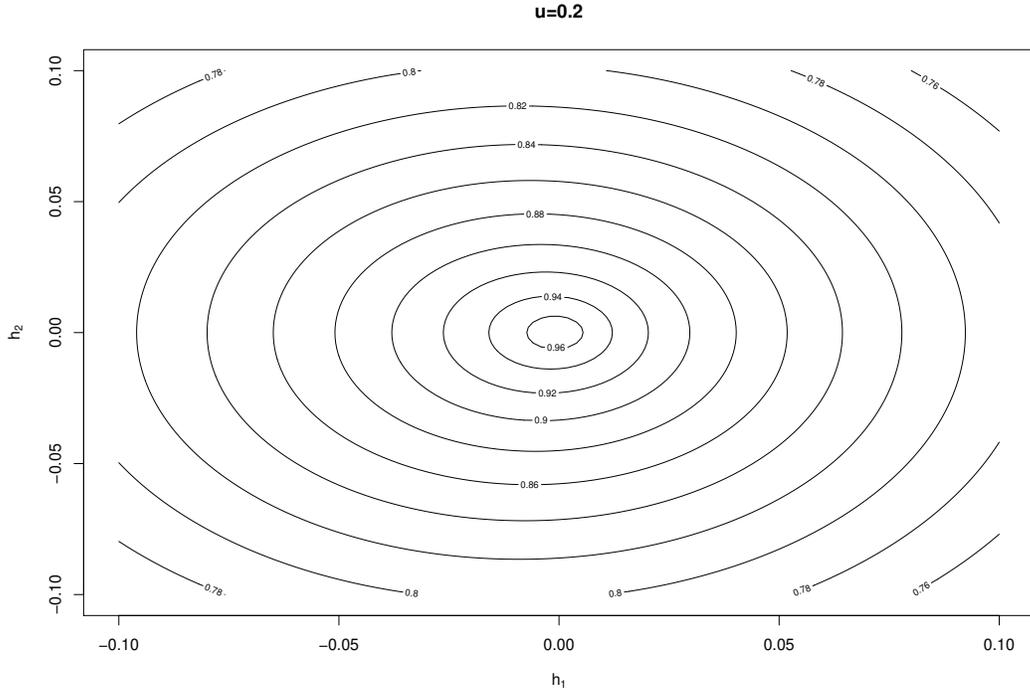
**Figure 4.6.** Contour plots for  $\mathcal{N}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2}(\mathbf{h}, u)$  as functions of  $\mathbf{h} = (h_1, h_2)$ , with  $u = 0.1$ ,  $\mathbf{r} = (0.5, 0)$ ,  $a_1 = a_2 = 1$ , and  $\nu_1 = 0.35, \nu_2 = 3$ .

computation complexity of the modified Bessel functions of the second kind is much higher than that of Gamma functions. When doing numerical computation using R for a Shandong province temperature data set, which will be introduced in Chapter 5, the average elapsed time for each iteration of updating  $\boldsymbol{\theta} = (a_1, a_2, \mathbf{r})$  using the NFSST Matérn-Cauchy model was 7.5 hours, almost halved the average elapsed time of the NFSST Matérn model, which was 14.0 hours.

Secondly, as mentioned in Section 3.2, the NFSST Matérn-Cauchy model may better fit spatiotemporal processes with temporal long-range dependence. The heavy tails of Cauchy distributions lead to temporal LRD properties of corresponding Matérn-Cauchy correlation functions. More specifically, if the smoothness parameter  $\nu_2$  in a Matérn Cauchy correlation function

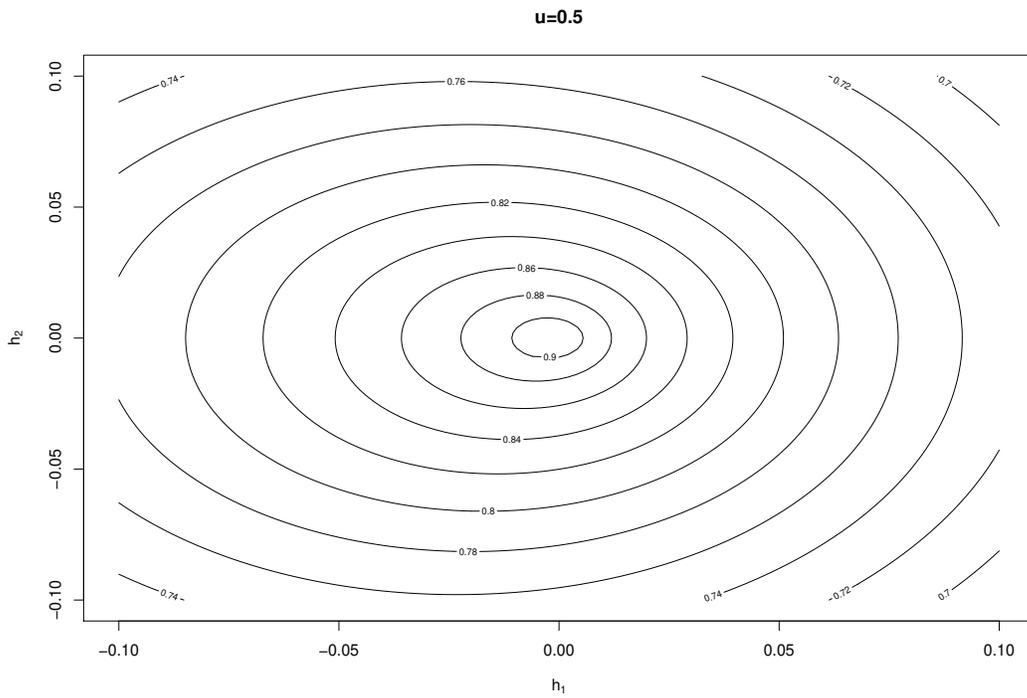
$$C(u) = \frac{1}{(1 + a^2 u^2)^{1/\nu_2}}, u \in \mathbb{R}, a, \nu_2 > 0 \quad (4.25)$$

is greater or equal to 2, its temporal margin has long-range dependence property.

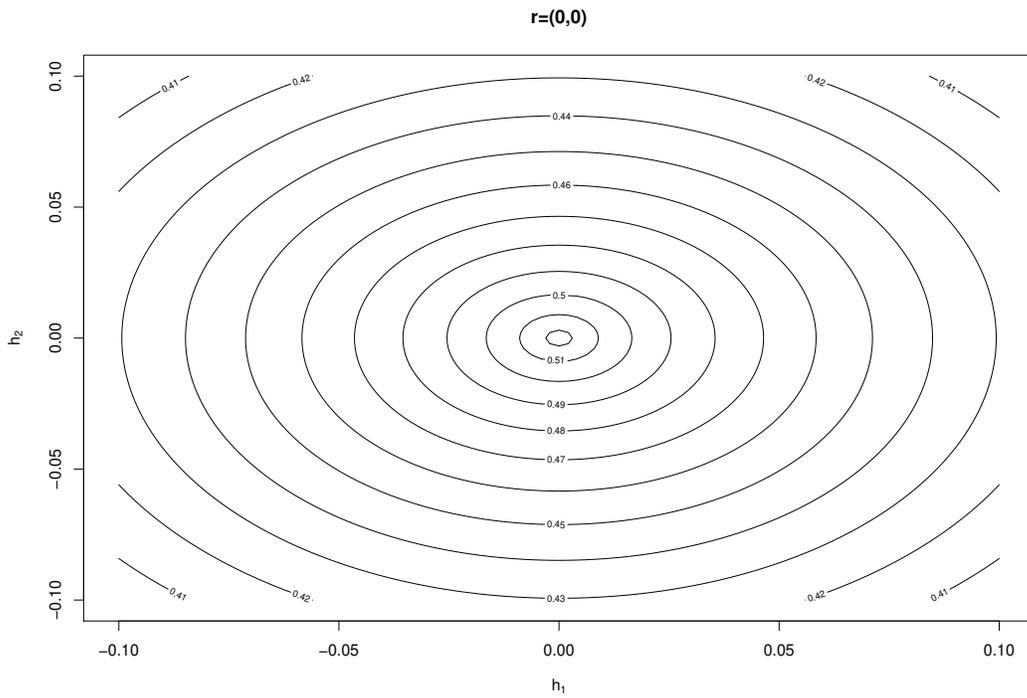


**Figure 4.7.** Contour plots for  $\mathcal{N}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2}(\mathbf{h}, u)$  as functions of  $\mathbf{h} = (h_1, h_2)$ , with  $u = 0.2$ ,  $\mathbf{r} = (0.5, 0)$ ,  $a_1 = a_2 = 1$ , and  $\nu_1 = 0.35, \nu_2 = 3$ .

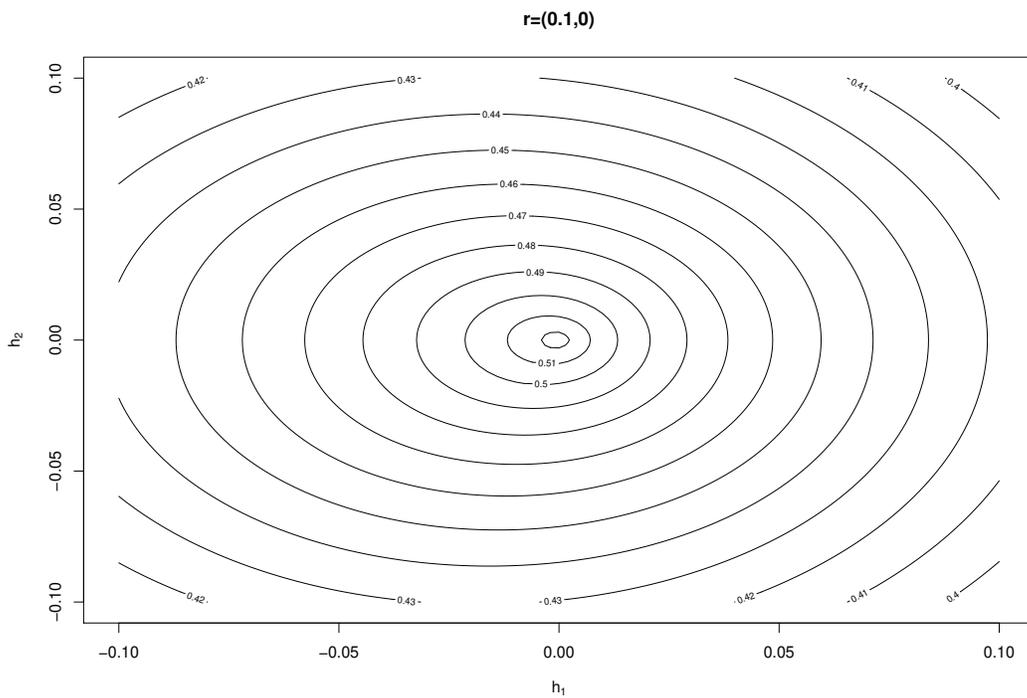
There have been rich studies on LRD climate processes. In these studies, the climate processes are studied as time series observed at different locations. In our model, we view the climate processes as space-time processes, which allows us to gain insight into the interaction between space and time, as well as the long-term statistical behavior of the climate systems. We study space-time processes whose temporal margin processes have LRD properties. If such space-time processes are fitted using the NFSST Matérn-Cauchy model, their temporal smoothness parameters  $\nu_2$  should be greater or equal to 2.



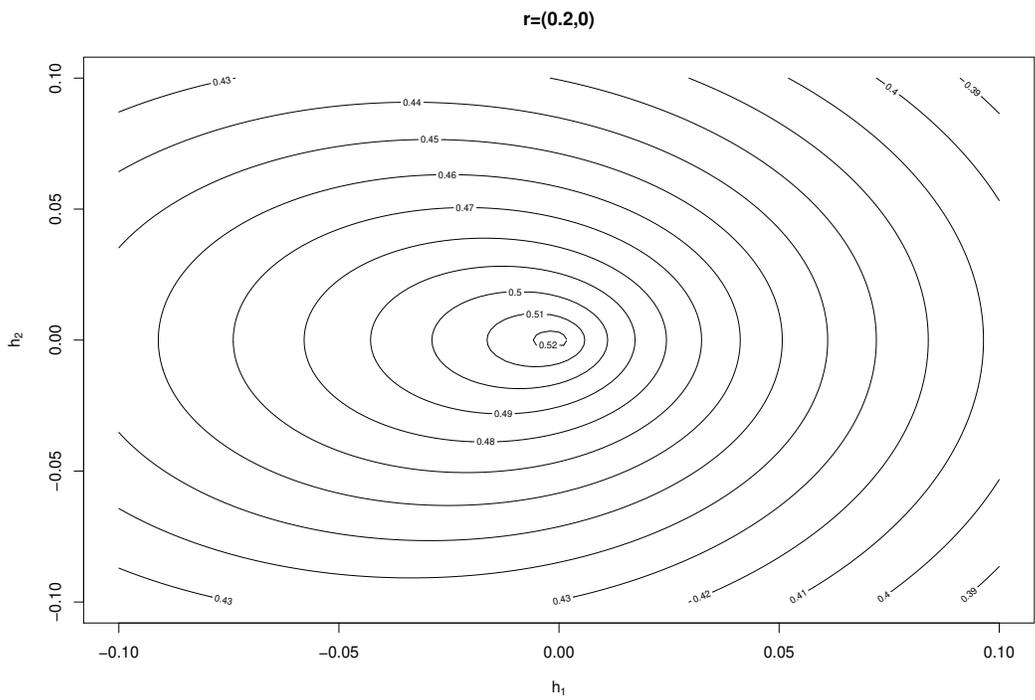
**Figure 4.8.** Contour plots for  $\mathcal{N}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2}(\mathbf{h}, u)$  as functions of  $\mathbf{h} = (h_1, h_2)$ , with  $u = 0.5$ ,  $\mathbf{r} = (0.5, 0)$ ,  $a_1 = a_2 = 1$ , and  $\nu_1 = 0.35, \nu_2 = 3$ .



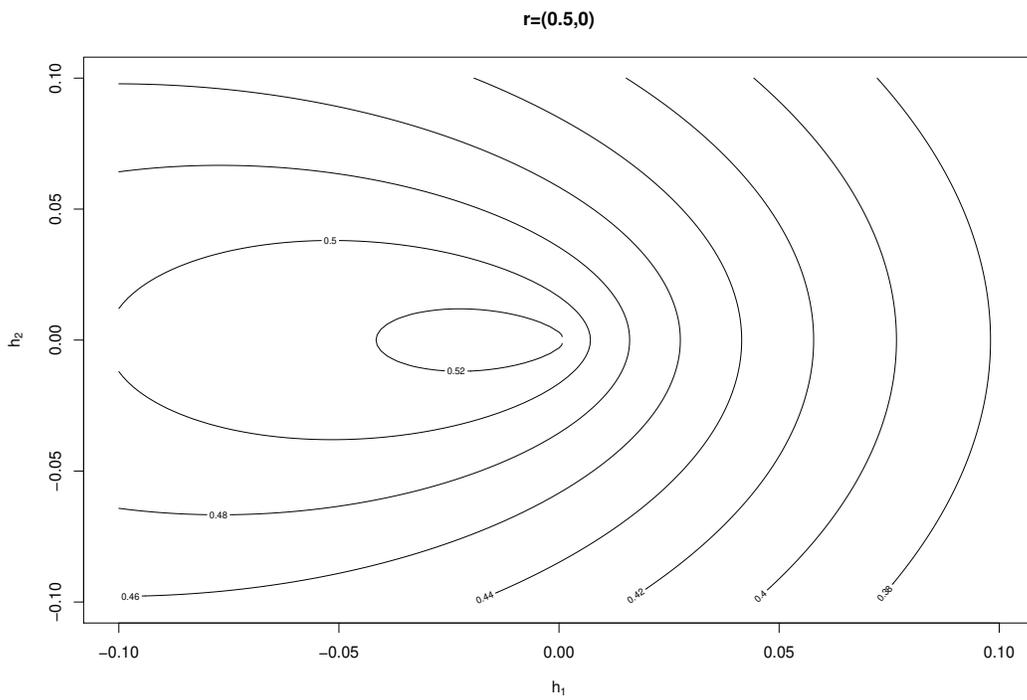
**Figure 4.9.** Contour plots for  $\mathcal{N}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2}(\mathbf{h}, u)$  as functions of  $\mathbf{h} = (h_1, h_2)$ , with  $\mathbf{r} = (0, 0)$ ,  $u = 0.5$ ,  $a_1 = a_2 = 1$ , and  $\nu_1 = \nu_2 = 0.35$ .



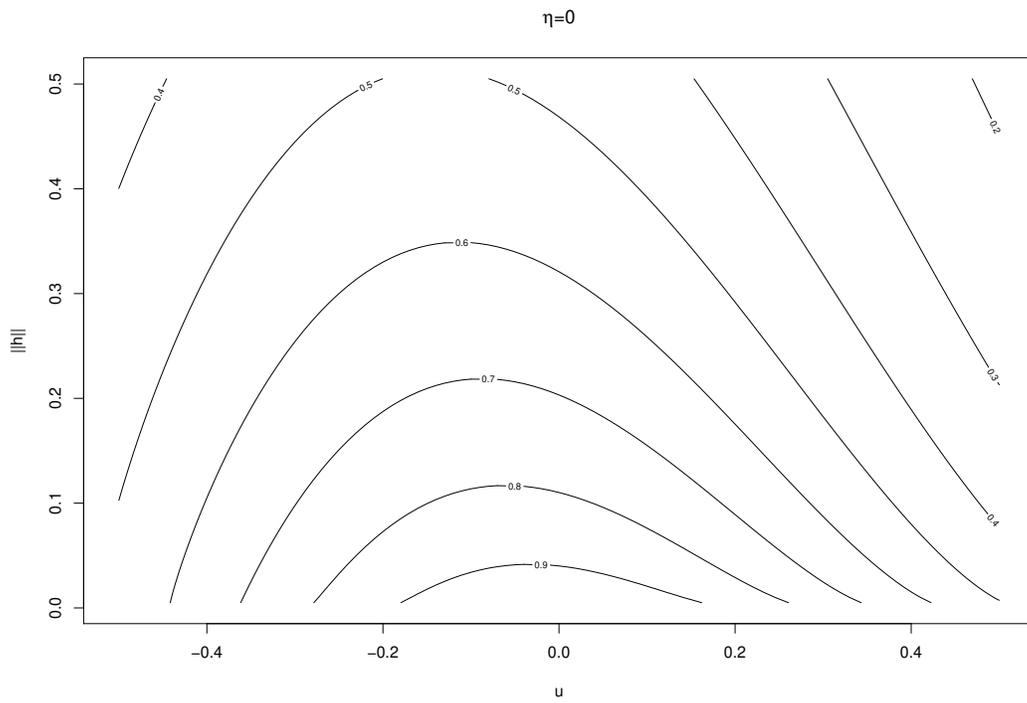
**Figure 4.10.** Contour plots for  $\mathcal{N}_{\mathbf{r},a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$  as functions of  $\mathbf{h} = (h_1, h_2)$ , with  $\mathbf{r} = (0.1, 0)$ ,  $u = 0.5$ ,  $a_1 = a_2 = 1$ , and  $\nu_1 = \nu_2 = 0.35$ .



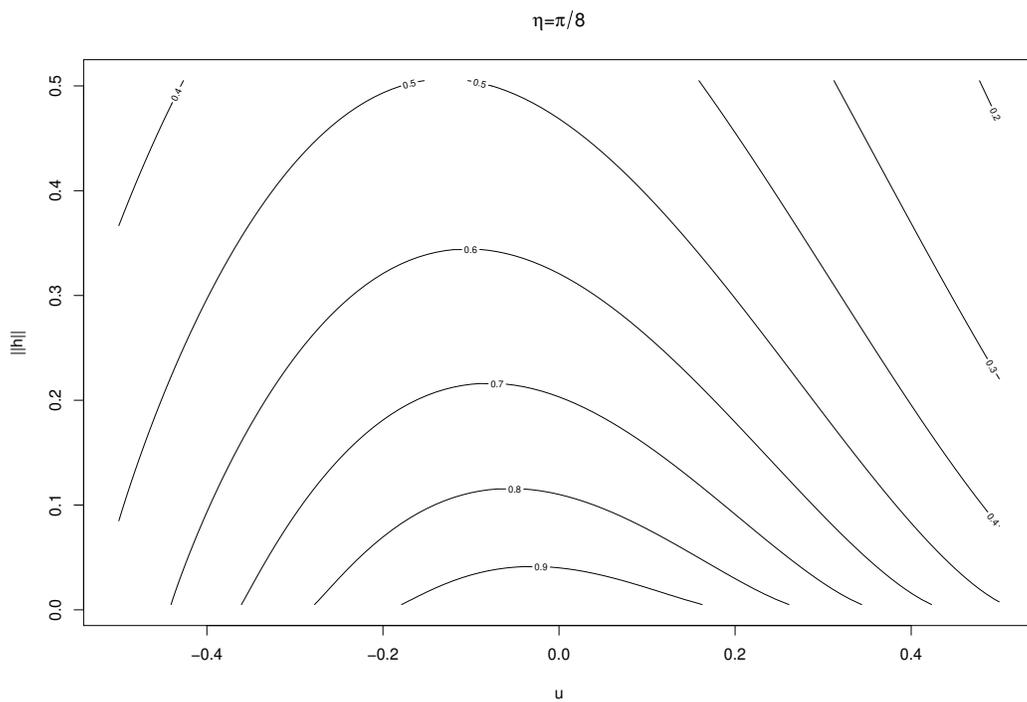
**Figure 4.11.** Contour plots for  $\mathcal{N}_{\mathbf{r},a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$  as functions of  $\mathbf{h} = (h_1, h_2)$ , with  $\mathbf{r} = (0.2, 0)$ ,  $u = 0.5$ ,  $a_1 = a_2 = 1$ , and  $\nu_1 = \nu_2 = 0.35$ .



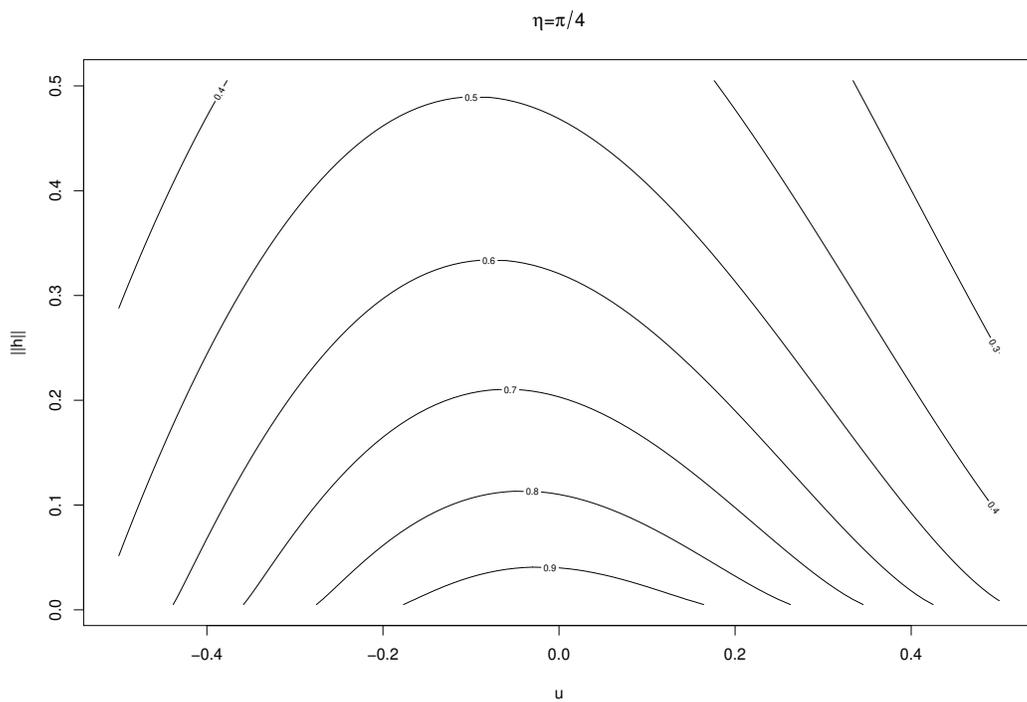
**Figure 4.12.** Contour plots for  $\mathcal{N}_{\mathbf{r},a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$  as functions of  $\mathbf{h} = (h_1, h_2)$ , with  $\mathbf{r} = (0.5, 0)$ ,  $u = 0.5$ ,  $a_1 = a_2 = 1$ , and  $\nu_1 = \nu_2 = 0.35$ .



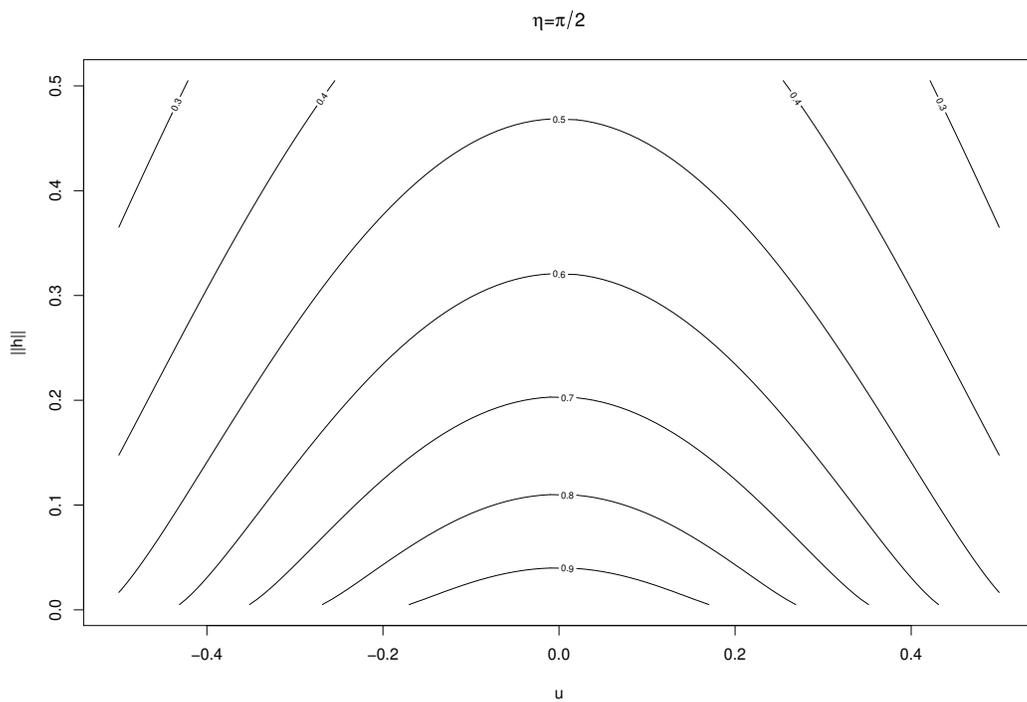
**Figure 4.13.** Contour plots for  $\mathcal{N}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$  as functions of  $\|\mathbf{h}\|$  and  $u$ , with  $\eta = 0$ ,  $\mathbf{r} = (0.5, 0)$ ,  $a_1 = a_2 = 1$ , and  $\nu_1 = \nu_2 = 0.35$ , where  $\eta$  is the angle between  $\mathbf{h}$  and the horizontal axis.



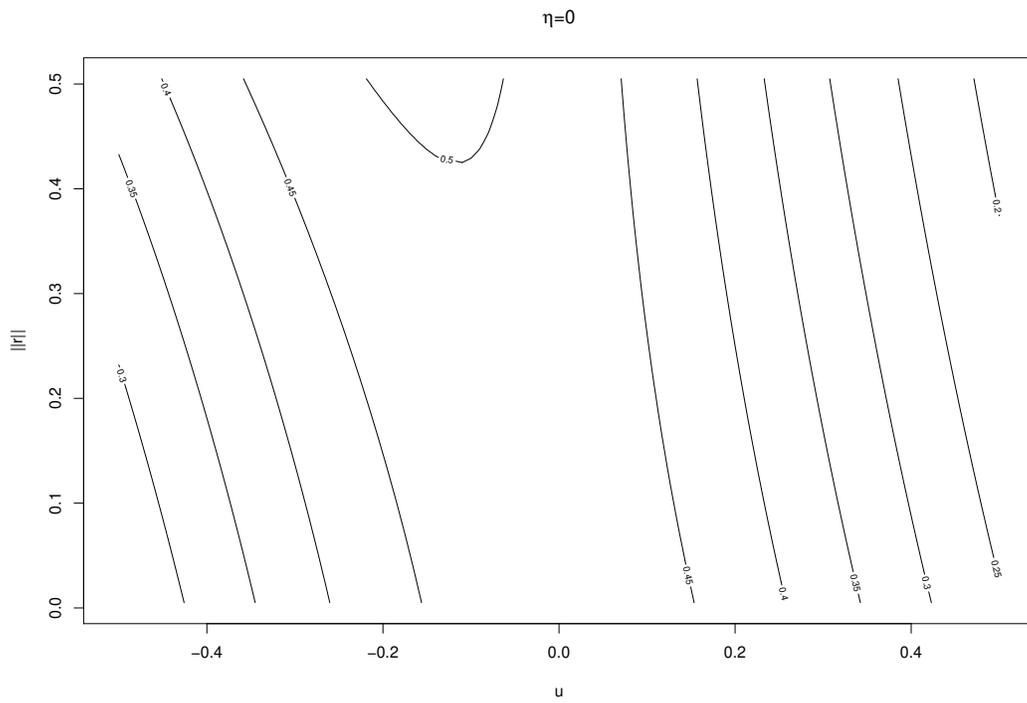
**Figure 4.14.** Contour plots for  $\mathcal{N}_{\mathbf{r},a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$  as functions of  $\|\mathbf{h}\|$  and  $u$ , with  $\eta = \pi/8$ ,  $\mathbf{r} = (0.5, 0)$ ,  $a_1 = a_2 = 1$ , and  $\nu_1 = \nu_2 = 0.35$ , where  $\eta$  is the angle between  $\mathbf{h}$  and the horizontal axis.



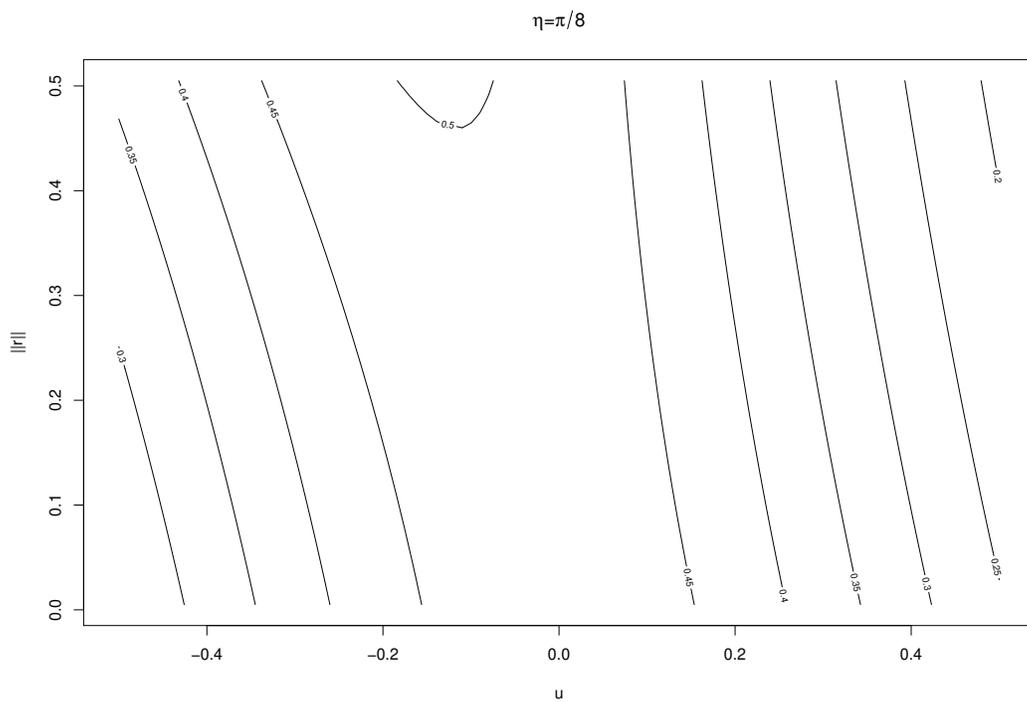
**Figure 4.15.** Contour plots for  $\mathcal{N}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2}(\mathbf{h}, u)$  as functions of  $\|\mathbf{h}\|$  and  $u$ , with  $\eta = \pi/4$ ,  $\mathbf{r} = (0.5, 0)$ ,  $a_1 = a_2 = 1$ , and  $\nu_1 = \nu_2 = 0.35$ , where  $\eta$  is the angle between  $\mathbf{h}$  and the horizontal axis.



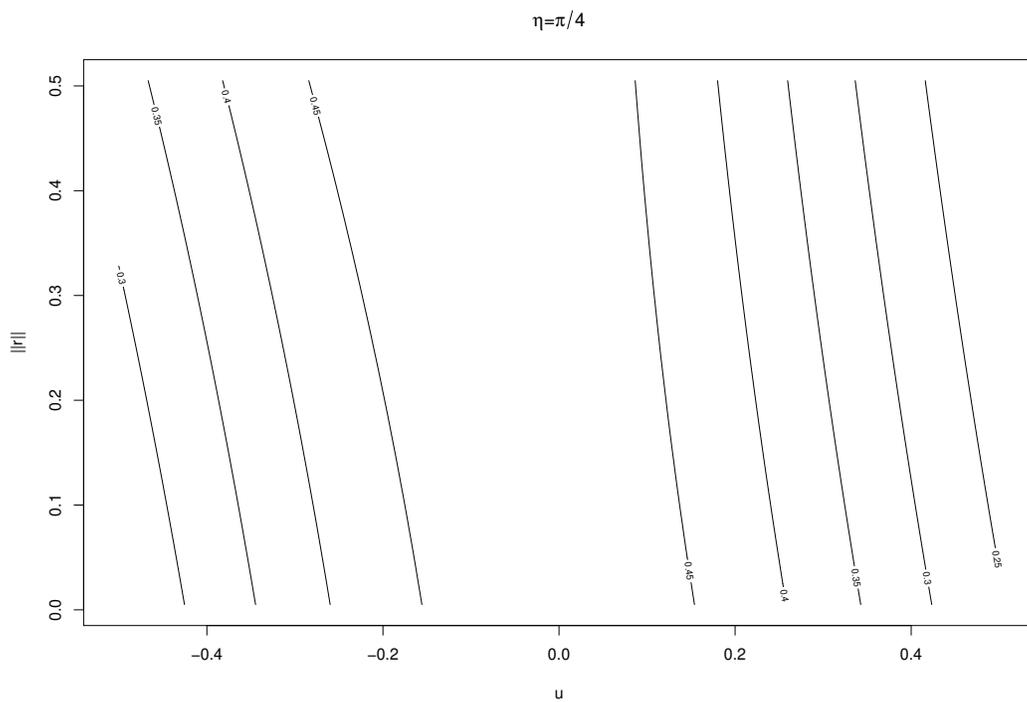
**Figure 4.16.** Contour plots for  $\mathcal{N}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2}(\mathbf{h}, u)$  as functions of  $\|\mathbf{h}\|$  and  $u$ , with  $\eta = \pi/2$ ,  $\mathbf{r} = (0.5, 0)$ ,  $a_1 = a_2 = 1$ , and  $\nu_1 = \nu_2 = 0.35$ , where  $\eta$  is the angle between  $\mathbf{h}$  and the horizontal axis.



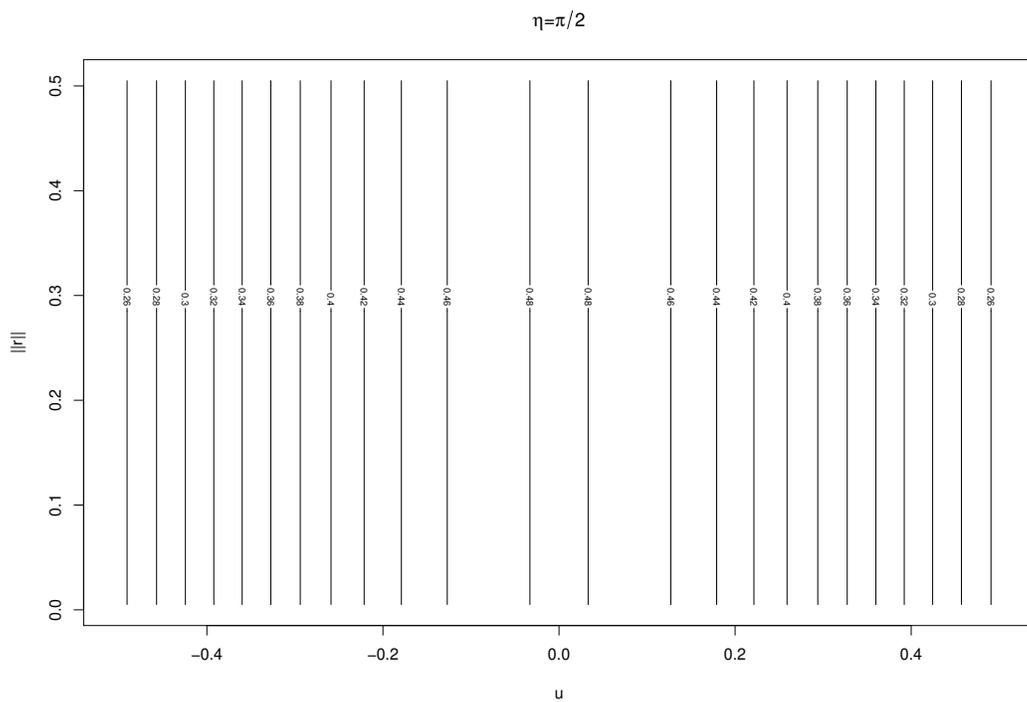
**Figure 4.17.** Contour plots for  $\mathcal{N}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$  as functions of  $\|\mathbf{r}\|$  and  $u$ , with  $\eta = 0$ ,  $\mathbf{h} = (0.5, 0)$ ,  $a_1 = a_2 = 1$ , and  $\nu_1 = \nu_2 = 0.35$ , where  $\eta$  is the angle between  $\mathbf{r}$  and the horizontal axis.



**Figure 4.18.** Contour plots for  $\mathcal{N}_{r, a_1, a_2, \nu_1, \nu_2}(\mathbf{h}, u)$  as functions of  $\|\mathbf{r}\|$  and  $u$ , with  $\eta = \pi/8$ ,  $\mathbf{h} = (0.5, 0)$ ,  $a_1 = a_2 = 1$ , and  $\nu_1 = \nu_2 = 0.35$ , where  $\eta$  is the angle between  $\mathbf{r}$  and the horizontal axis.



**Figure 4.19.** Contour plots for  $\mathcal{N}_{r, a_1, a_2, \nu_1, \nu_2}(\mathbf{h}, u)$  as functions of  $\|\mathbf{r}\|$  and  $u$ , with  $\eta = \pi/4$ ,  $\mathbf{h} = (0.5, 0)$ ,  $a_1 = a_2 = 1$ , and  $\nu_1 = \nu_2 = 0.35$ , where  $\eta$  is the angle between  $\mathbf{r}$  and the horizontal axis.



**Figure 4.20.** Contour plots for  $\mathcal{N}_{\mathbf{r},a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$  as functions of  $\|\mathbf{r}\|$  and  $u$ , with  $\eta = \pi/2$ ,  $\mathbf{h} = (0.5, 0)$ ,  $a_1 = a_2 = 1$ , and  $\nu_1 = \nu_2 = 0.35$ , where  $\eta$  is the angle between  $\mathbf{r}$  and the horizontal axis.

## 5. Numerical Results

As mentioned earlier, there are rich climate and geophysical processes under the influence of monsoon and ocean currents. And traditional fully-symmetric models are not realistic for these processes. In this chapter we fit three meteorological processes with our NFSST Matérn Cauchy model, and give analysis on fitting results.

### 5.1 Shandong Temperature Data

We applied our Matérn- Cauchy model to the Chinese daily temperature data in this section. The data were provided by the Climatic Data Center, National Meteorological Information Center, China Meteorological Administration. They contained the average daily temperature, the lowest daily temperature, and the highest daily temperature at 756 weather stations from 1951 to 2007 in all of the nine climatic zones in China. In order to avoid the consideration of climate zones and seasonal patterns, we decided to focus on our analysis for data within a single month and a single province.

We extracted the daily highest temperature data in July 2007 in Shandong province. Shandong province is located in Northern China, extended from  $34.37^\circ$  to  $39.38^\circ$  latitude north and  $114.32^\circ$  to  $122.72^\circ$  longitude east. Its area is about 157.8 thousand square kilometers. After extraction, the data set contained 20 stations in the province from the first day to the last day in July 2007.

After excluding one missing value, it had 619 observations of daily highest temperature values. The impact of altitude was removed by regressing the daily highest temperature across sites on their monthly averages.

Let  $Y(\mathbf{s}, t)$  be the daily highest temperature relative to its site average. We used a geostatistical model

$$Y(\mathbf{s}, t) = \mu + Z(\mathbf{s}, t) + \epsilon(\mathbf{s}, t)$$



**Figure 5.1.** Locations of Weather Stations in Shandong Province, China in July 2007.

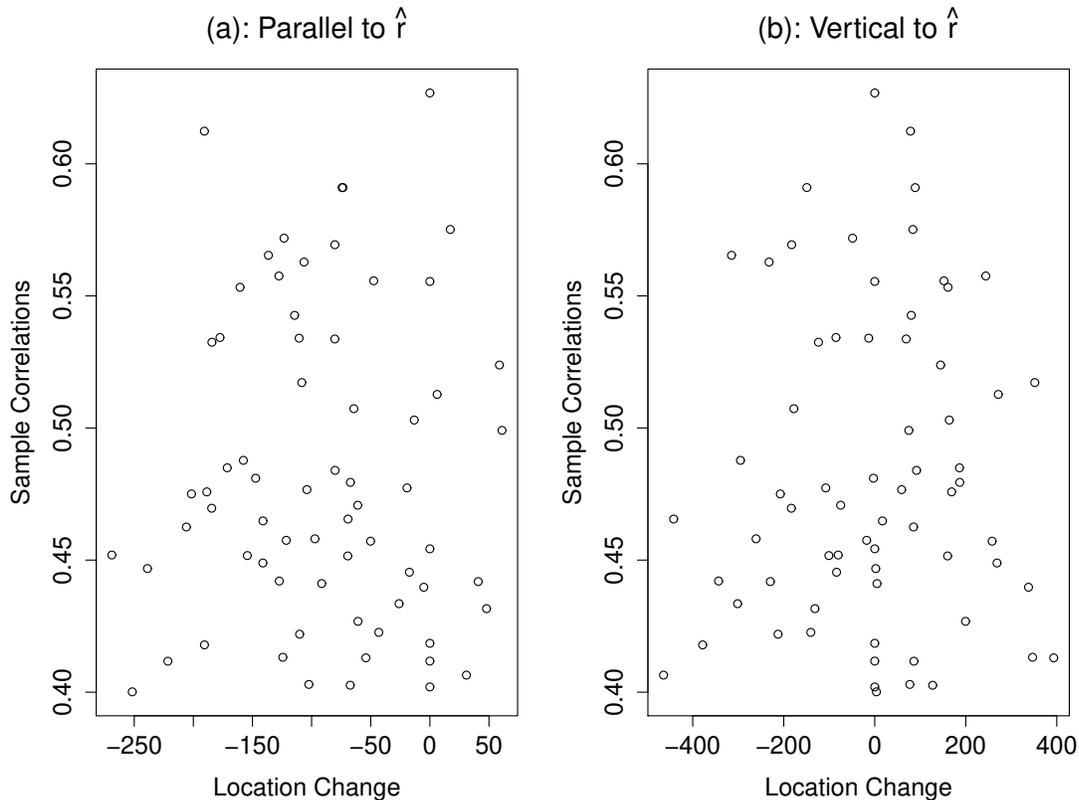
to analyze the spatiotemporal process  $Y(\mathbf{s}, t)$ , where  $\epsilon(\mathbf{s}, t)$  was a white noise process and  $Z(\mathbf{s}, t)$  was a mean zero stationary Gaussian process. The covariance function of  $Z(\mathbf{s}, t)$  was modeled as

$$E[Z(\mathbf{s}, t)Z(\mathbf{s} + \mathbf{h}, t + u)] = \tau^2 \mathcal{N}_{\mathbf{r}, a_1, a_2, \nu_1, \nu_2}(\mathbf{h}, u),$$

where  $a_1$  and  $a_2$  were the scale parameters and  $\nu_1$  and  $\nu_2$  were the smoothness parameters for the space and time processes, respectively. We used  $\boldsymbol{\theta} = (\nu_1, a_1, \nu_2, a_2, r_1, r_2)$  to represent the correlation parameters and  $\eta = \frac{\tau^2}{\tau^2 + \sigma^2}$  to represent the nugget effect parameter, where  $\sigma^2$  was the variance parameter in the white noise process. We estimated  $\eta$  and  $\boldsymbol{\theta}$  by the profile likelihood approach. Applying a Newton-Raphson algorithm, the estimates were  $\hat{\eta} = 0.92931$  and  $\hat{\boldsymbol{\theta}} = (0.2646, 0.0009794, 0.1977, 0.4816, -0.1867, 0.3728)$ . Because the estimates of the correlation parameters were  $\hat{\mathbf{r}} = (0.1867, 0.3728) \neq \mathbf{0}$ , we concluded that the model was nonseparable. And because of  $\hat{\nu}_2 = 0.1977 < 2$ , we concluded that there was no temporal long-range dependence in this process.

To conclude our model is appropriate in the analysis of the data set, we fixed the time difference to be 1 and let the space difference varied, and plotted the sample correlations along the direction parallel to  $\mathbf{r}$  and along the counterclockwise direction vertical to  $\mathbf{r}$ . We

found that most of large  $\hat{C}(\mathbf{s}_i - \mathbf{s}_j, 1)$  values were at the opposite direction of  $\mathbf{r}$ , indicating that the sample correlations were not symmetric along  $\mathbf{r}$ . Along the direction vertical to  $\mathbf{r}$ , we found that large  $\hat{C}(\mathbf{s}_i - \mathbf{s}_j, 1)$  values were almost symmetric. Therefore, we concluded that our model was more appropriate than a fully symmetric model in the analysis of the data set.



**Figure 5.2.** Values of  $\hat{C}(\mathbf{s}_i - \mathbf{s}_j, 1)$  when they are greater than 0.4 corresponding to the direction parallel to  $\hat{\mathbf{r}}$  (left) and the counterclockwise direction vertical to  $\hat{\mathbf{r}}$  (right).

## 5.2 Irish Wind Data

We also applied our NFSST Matérn-Cauchy model to the Irish wind speed data set, which is available in the R package *gstat*. The data set contained the daily average wind speeds from 1961 to 1978 at 12 stations in Ireland. As recommended by Gneiting (2002) [5], we removed Rosslare in our analysis. The whole period of the data has been previously

analyzed by many authors (Fonseca and Steel, 2011 [26]; Gneiting, 2002; Stein, 2005 [7], e.g.). In order to understand the non-fully symmetry, we focused on the wind speed in July 1961. The final data contained 11 stations in 31 days. The size of the data was  $n = 11 \times 31 = 341$ . Following Haslett and Raftery (1989) [46], we used the square root transformation of the wind speed because this transformation made the data nearly Gaussian. After that, we obtained the response values by removing the station averages.

Using similar procedure we displayed in the previous data analysis, let  $Y(\mathbf{s}, t)$  be the square root of wind speed. We used the model

$$Y(\mathbf{s}, t) = \mu + Z(\mathbf{s}, t) + \epsilon(\mathbf{s}, t)$$

to fit the process. And let the covariance function of  $Z(\mathbf{s}, t)$  be NFSST Matérn-Cauchy:

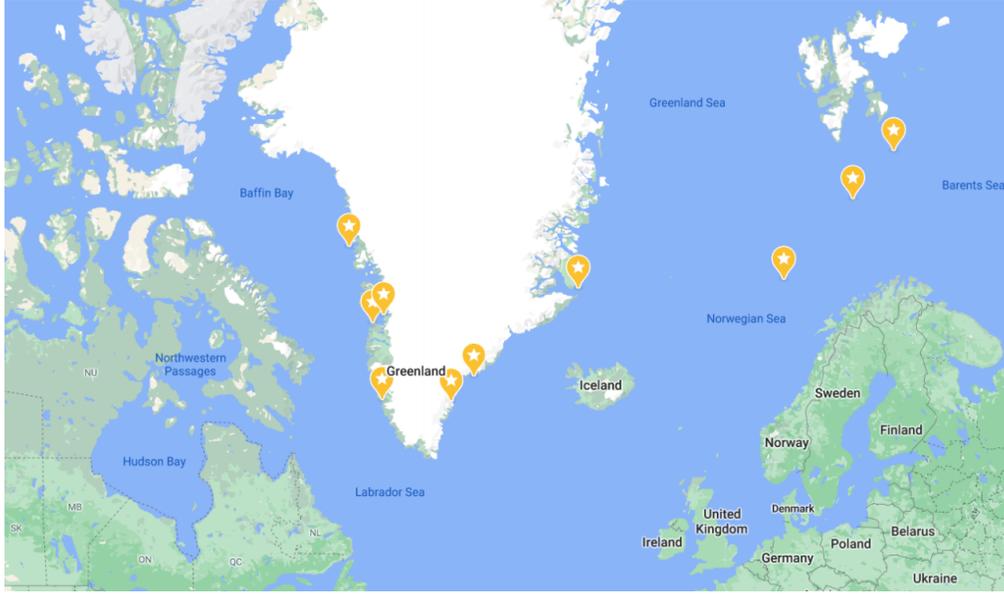
$$E[Z(\mathbf{s}, t)Z(\mathbf{s} + \mathbf{h}, t + u)] = \tau^2 \mathcal{N}_{r, a_1, a_2, \nu_1, \nu_2}(\mathbf{h}, u).$$

Let  $\boldsymbol{\theta} = (\nu_1, a_1, \nu_2, a_2, r_1, r_2)$  and  $\eta$  be the nugget effect, their MLEs were  $\hat{\eta} = 0.9459$  and  $\hat{\boldsymbol{\theta}} = (0.6806, 0.001653, 0.5606, 0.7979, -0.6286, 0.3167)$ . The estimated model was nonseparable as  $\hat{r} \neq 0$ . And it had no temporal long-range dependence since  $\hat{\nu}_2 = 0.5606 < 2$

### 5.3 North Atlantic Ocean Temperature Data

We fitted a North Atlantic Ocean Temperature data set, which is available at <http://rim-frost.no/> with our NFSST Matérn-Cauchy model. The data set contained the annual average temperature from 1961 to 2010 at 10 different stations in North Atlantic Ocean, some of them located around Greenland, while others located in Norwegian sea. Knowing that some climate time-series processes within this geographical region were long-range dependent, we investigated further into the temporal LRD of our model. The stations studied were Angmassalik (Greenland-S.E.), POS : (65.60, -37.63), Bjørnøya (Norway-North Atlantic), POS : (74.5, 19.0), Egedsminde (Aasiaat-Greenlan West), POS : (68.70, -52.75), Hopen (Norway-North Atlantic), POS : (76.30, 25.0), Illulisat (Jacobshavn-Greenland West), POS : (69.12, -51.10), Ittoqqortoormiit (Greenland EAST) POS : (70.50, -22.00), Jan Mayen

(Norway-North Atlantic), POS : (70.93, 8.67), Nuuk - Godthøp (Greenland-South West), POS : (64.10, -51.35), Tasiilaq (Greenland-East), POS : (64, -41), Upernavik (Greenland-North West), POS : (72.49, -56.20).



**Figure 5.3.** Locations of Weather Stations in North Atlantic Ocean.

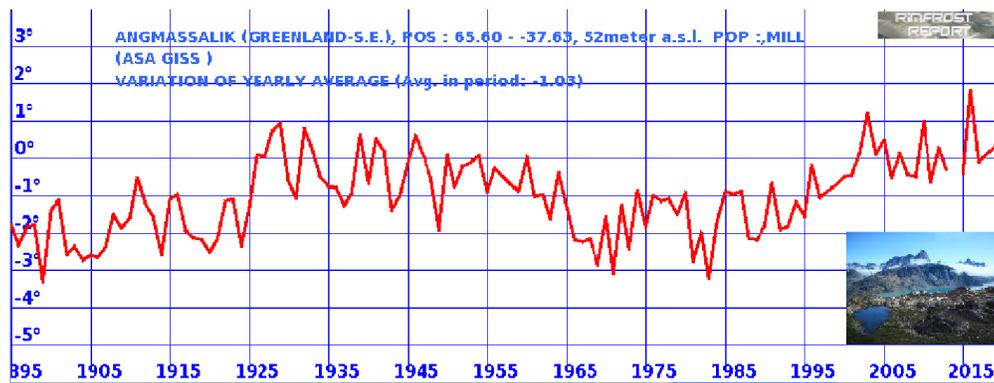
After 2 missing values were removed, the data size was  $n = 10 \times 50 - 2 = 498$ . Firstly we viewed the data as time series at each station, and estimated the Hurst exponents of the temperature processes at each location. All those estimated exponents were great than 0.5, indicating the existence of LRD. For instance, figure 5.4 shows the temperature process at station Angmassalik. The estimated R over S Hurst exponent of this process was 0.76.

Then we fitted the data as a spatio-temporal process using the NFSST Matérn-Cauchy model

$$Y(\mathbf{s}, t) = \mu + Z(\mathbf{s}, t) + \epsilon(\mathbf{s}, t),$$

where  $Y(\mathbf{s}, t)$  is the annual average temperature.  $E[Z(\mathbf{s}, t)Z(\mathbf{s}+\mathbf{h}, t+u)] = \tau^2 \mathcal{N}_{r, a_1, a_2, \nu_1, \nu_2}(\mathbf{h}, u)$ .

Let  $\boldsymbol{\theta} = (\nu_1, a_1, \nu_2, a_2, r_1, r_2)$  and  $\eta$  be the nugget effect, the MLEs were  $\hat{\eta} = 0.9452$  and  $\hat{\boldsymbol{\theta}} = (0.7893, 0.0003199, 2.7238, 2.6711, 0.03384, 0.1356)$ . There seemed to be slight non-fully symmetry towards northeast, and the temporal margin Cauchy correlation had LRD since  $\hat{\nu}_2 = 2.7238 > 2$ .



**Figure 5.4.** Yearly average temperature process at Angmassalik, source <http://rimfrost.no/>.

## 6. Conclusion and Future Work

### 6.1 Conclusion

In this dissertation, we review the methodology of geostatistics and the analysis of space-time processes. With Bochner's representation, one can construct space-time correlation functions using existing spatial marginal and temporal marginal correlation functions. Inspired by the Non-Fully Symmetric Space-Time (NFSST) Matérn model by Zhang, T. and Zhang, H. (2015) [1], we construct the non-fully symmetric space-time Matérn-Cauchy correlation model, with spatial Matérn marginal correlations and temporal Cauchy marginal correlations. The theoretical and computational properties of the Matérn-Cauchy model are studied. And three meteorological space-time processes are fitted with Matérn-Cauchy correlation functions, including one with temporal long range dependence property.

The numerical computation of the NFSST Matérn-Cauchy model is much faster compared with the NFSST Matérn model, since its temporal component does not contain modified Bessel functions of the second kind. Meanwhile it has similar estimate of the parameters when applied to meteorological space-time data sets. Moreover, it handles space-time processes with temporal long-range dependence, which were traditionally studied merely as time series. Our model allows us to gain insight into the interaction of spatial correlated long-term statistical behavior of climate systems.

### 6.2 Future Work

There are two interesting topics we want to further study on. The first one is about the ridge problem 3.18 proposed by Stein (2005) [7]. Zhang, T. and Zhang, H. in 2015 [1] proved that the NFSST Matérn model may overcome this issue when  $\nu_2 \in (0, 1/2)$  in  $\mathcal{M}_{r,a_1,a_2,\nu_1,\nu_2}(\mathbf{h}, u)$ . However, we have not reached similar conclusions in the NFSST Matérn-Cauchy model. We would like to find a sufficient condition under which the NFSST Matérn-Cauchy model does not satisfy 3.17, thus overcome the ridge problem.

The second problem is about the extension of the definition of space-time long-range dependence. Our NFSST Matérn-Cauchy model can fit some space-time processes with

temporal marginal LRD, as defined in 3.6. When the spatial lag  $\mathbf{h}$  is fixed,  $\{(\mathbf{h}, u); \mathbf{h} = \mathbf{h}_0\}$  is a subspace of  $\{(\mathbf{h}, u); \mathbf{h} \in \mathbb{R}^d, u \in \mathbb{R}\}$ . In other words, the temporal marginal process of a space-time process is a sub-process on a subspace of  $\mathbb{R}^d \times \mathbb{R}$ . We wonder if we can extend the definition of LRD to sub-spaces like  $L(\mathbf{h}, u) = 0$ , such that  $\int_{L(\mathbf{h}, u)=0} C(\mathbf{h}, u) du = \infty$ , where  $L$  is a function of  $\mathbf{h}$  and  $u$ , and  $C(\mathbf{h}, u)$  is the corresponding correlation function of the space-time process. This extension may further reveal the interaction between space and time.

We also plan to upload a R package to CRAN, including the fitting and kriging of both NFSST Matérn model and NFSST Matérn-Cauchy model.

## REFERENCES

- [1] T. Zhang and H. Zhang, *Non-Fully Symmetric Space-Time Matérn Covariance Functions*. 2015.
- [2] G. Mathéron, *Applications des méthodes statistiques 'a l'évaluation des gisements*. Annales des Mines, 12, 50-75, 1955.
- [3] A. E. Gelfand, P. Diggle, P. Guttorp, and M. Fuentes, *Handbook of Spatial Statistics*, First. Chapman Hall/CRC Handbooks of Modern Statistical Methods, 2010.
- [4] B. Matérn, *Spatial variation*. Meddelanden fran Statens Skogsforsknings Institut, Stockholm. Band 49, No. 5, 1960.
- [5] T. Gneiting, *Nonseparable, stationary covariance functions for space-time data*. Journal of the American Statistical Association, 97, 590-600, 2002.
- [6] L. Scaccia and R. J. Martin, *Testing axial symmetry and separability of lattice processes*. Journal of Statistical Planning and Inference, 131, 19-39, 2005.
- [7] M. L. Stein, *Space-Time Covariance Functions*. Journal of the American Statistical Association, 100, 310-321, 2005.
- [8] B. H. Briggs, *On the analysis of moving patterns in geophysics - I. Correlation analysis*. Journal of Atmospheric and Terrestrial Geophysics, 30, 1777-1788, 1968.
- [9] B. Matérn, *Spatial Variation*, Second. Berlin: Springer-Verlag, 1986.
- [10] M. S. Handcock and M. L. Stein, *A Bayesian analysis of Kriging*. Technometrics, 35, 403-410, 1993.
- [11] M. L. Stein, *Space-Time Covariance Functions*. Journal of the American Statistical Association, 100, 310-321, 2005.
- [12] T. Guttorp and T. Gneiting, *Studies in the history of probability and statistics XLIX: On the Matérn correlation family*. Biometrika, 93, 989-995, 2006.
- [13] D. Lee and G. Shaddick, *Spatial modeling of air pollution in studies of its short-term health effects*. Biometrics, 66, 1238-1246, 2010.
- [14] G. North, J. Wang, and M. Genton, *Correlation models for temperature fields*. Correlation models for temperature fields. Journal of Climate, 24, 5850-5862, 2011.

- [15] N. Cressie and H. C. Huang, *Classes of nonseparable, spatio-temporal stationary covariance functions*. Journal of the American Statistical, 94, 1330-1340, 1999.
- [16] P. C. Kyriakidis and A. G. Journel, *Geostatistical space-time models: a review*. Mathematical Geology, 31, 651-684, 1999.
- [17] S. De Iaco, D. Myers, and D. Posa, *Space-time analysis using a general product-sum model*. Statistics and Probability Letters, 52, 21-28, 2001.
- [18] C. Ma, *Spatio-temporal covariance functions generated by mixtures*. Mathematical Geology, 34, 965-975, 2002.
- [19] C. Ma, *Spatiotemporal stationary covariance models*. Journal of Multivariate Analysis, 86, 97-107, 2003.
- [20] J. Mateu, E. Porcu, and P. Gregori, *Recent advances to model anisotropic space-time data*. Statistical Methods and Applications, 17, 209-223, 2008.
- [21] E. Porcu, P. Gregori, and J. Mateu, *Nonseparable stationary anisotropic space-time covariance functions*. Stochastic Environmental Research and Risk Assessment, 21, 113-122, 2006.
- [22] E. Porcu, J. Mateu, and F. Saura, *New classes of covariance and spectral density functions for spatio-temporal modelling*. Stochastic Environmental Research and Risk Assessment, 22, 65- 79, 2008.
- [23] C. Ma, *Linear combinations of space-time covariance functions and variograms*. IEEE Transactions on signal processing, 53, 857-864, 2005.
- [24] A. Rodrigues and P. Diggle, *A class of convolution-based models for spatiotemporal processes with non-separable covariance structure*. Scandinavian Journal of Statistics, 37, 553- 567, 2010.
- [25] A. Kolovos, G. Christakos, D. Hristopulos, and M. Serre, *Methods for generating non-separable spatiotemporal covariance models with potential environmental applications*. Advances in Water Resources, 27, 815-830, 2004.
- [26] T. C. O. Fonseca and M. F. J. Steel, *A general class of nonseparable space-time correlation functions*. Environmentrics, 22, 224-242, 2011.
- [27] P. Brown, K. Karesen, and G. O. R. S. Tonellato, *Blur-generated nonseparable space-time models*. Journal of Royal Statistical Society Series B, 62, 847-860, 2000.

- [28] M. Fuentes, *Testing for separability of spatial-temporal covariance functions*. Journal of Statistical Planning and Inference, 136, 447-466, 2006.
- [29] B. Li, M. Genton, and M. Sherman, *A nonparametric assessment of properties of space-time covariance functions*. Journal of the American Statistical Association, 102, 736-744, 2007.
- [30] M. Mitchell, M. G. Genton, and M. Gumpertz, *Testing for separability of space-time covariances*. Environmetrics, 16, 819-831, 2005.
- [31] S. De Iaco, D. Posa, and D. Myers, *Characteristics of some classes of Space-time covariance functions*. Journal of Statistical Planning and Inference, 143, 2002-2015, 2013.
- [32] S. De Iaco and D. Posa, *Positive and negative non-separability for space-time covariance models*. Journal of Statistical Planning and Inference, 143, 378-391, 2013.
- [33] X. Shao and B. Li, *A tuning parameter free test for properties of spacetime covariance functions*. Journal of Statistical Planning and Inference, 139, 4031-4038, 2009.
- [34] T. Gneiting, M. G. Genton, and P. Guttorp, *Geostatistical space-time models, stationarity, separability, and full symmetry*. In Statistical Methods for Spatiotemporal Systems, Finkenstadt, B., Held, L., and Lsham, V. (eds). Chapman Hall/CRC, 151-175, 2007.
- [35] S. De Iaco, D. Myers, and D. Posa, *Nonseparable space-time covariance models: Some parametric families*. Mathematical Geology, 34, 23-41, 2002.
- [36] C. Ma, *Families of spatio-temporal stationary covariance models*. Journal of Statistical Planning and Inference, 116, 489-501, 2003.
- [37] S. Bochner, *Harmonic Analysis and the Theory of Probability*. Berkeley and Los Angeles: University of California Press, 1955.
- [38] H. Cramér and M. R. Leadbetter, *Stationary and Related Stochastic Processes. Sample Function Properties and Their Applications*. New York: John Wiley Sons, 1967.
- [39] P. R. Rider, *Generalized Cauchy distributions*. Annals of the Institute of Statistical Mathematics, 1957, 9(1):215-223, 1957.
- [40] J. Miller and J. Thomas, *Detectors for discrete-time signals in non-Gaussian noise*. IEEE Transactions on Information Theory, 1972, 18(2):241-250, 1972.

- [41] R. E. Carrillo, K. E. Barner, and T. C. Aysal, *Robust Sampling and Reconstruction Methods for Sparse Signals in the Presence of Impulsive Noise*. IEEE Journal of Selected Topics in Signal Processing, 2010, 4(2):392-408, 2010.
- [42] C. Granger and R. Joyeux, *An introduction of long memory time series models*. Journal of Time Series Analysis, 1980, 4(1):221-228, 1980.
- [43] H. E. Hurst, *Long-term storage capacity of reservoirs*. Transactions of the American Society of Civil Engineers. 116: 770, 1951.
- [44] J. Beran, *Statistics for Long-Memory Processes*. Monographs on Statistics, Applied Probability, Vol. 61. New York: Chapman, and Hall/CRC, 1994.
- [45] J. Franke, *Long-Range Dependence in Millennial-Scale Climate Models*. 98th Annual Meeting of American Meteorological Society, 2018.
- [46] J. Haslett and A. E. Raftery, *Space-time modeling with long-memory dependence: assessing Ireland's wind power resources (with discussion)*. Applied Statistics, 38, 1-50, 1989.

## VITA

Zizhuang Wu was born in 1992 in Beijing, China. He received his bachelors degree in School of Mathematical Sciences, Peking University, China in 2014. Then he joined the Statistics Department at Purdue University for graduate study. He received his master degree in Probability and Statistics in 2016 and earned a doctoral degree in Statistics in 2021. His research interests include spatial statistics, time-series analysis and machine learning. His Ph.D. research focuses on construction of space-time covariance functions in geostatistics. During his graduate career, he served as a teaching assistant for various of graduate and undergraduate level courses in Statistics Department. He had summer internship experience at Genentech, a bio-pharmaceutical company. He would like to pursue a professional career as a statistical scientist in pharmaceutical industry.