UNIVERSIDAD DE INVESTIGACIÓN DE TECNOLOGÍA EXPERIMENTAL YACHAY

Escuela de Ciencias Matemáticas y Computacionales

TÍTULO: Kriging and Kalman filter to estimate dynamic spatio-temporal models

Trabajo de integración curricular presentado como requisito para la obtención del título de Matemático

Autor:

Becerra Loaiza Inti Israel

Tutor:

PH.D. Infante Quirpa Saba Rafael

Urcuquí, febrero de 2020

SECRETARÍA GENERAL
(Vicerrectorado Académico/Cancillería)
ESCUELA DE CIENCIAS MATEMÁTICAS Y COMPUTACIONALES
CARRERA DE MATEMÁTICA
ACTA DE DEFENSA No. UITEY-ITE-2020-00010-AD

En la ciudad de San Miguel de Urcuquí, Provincia de Imbabura, a los 12 días del mes de marzo de 2020, a las 14:00 horas, en el Aula CHA-01 de la Universidad de Investigación de Tecnología Experimental Yachay y ante el Tribunal Calificador, integrado por los docentes:

Presidente Tribunal de Defensa	Dr. AMARO MARTIN, ISIDRO RAFAEL , Ph.D.
Miembro No Tutor	Dr. MANZANILLA MORILLO, RAUL , Ph.D.
Tutor	Dr. INFANTE QUIRPA, SABA RAFAEL , Ph.D.

Se presenta el(la) señor(ita) estudiante BECERRA LOAIZA, INTI ISRAEL, con cédula de identidad No. 1103912620, de la ESCUELA DE CIENCIAS MATEMÁTICAS Y COMPUTACIONALES, de la Carrera de MATEMÁTICA, aprobada por el Consejo de Educación Superior (CES), mediante Resolución RPC-SO-15-No.174-2015, con el objeto de rendir la sustentación de su trabajo de titulación denominado: Kriging and Kalman filter to estimate dynamic spatio-temporal models , previa a la obtención del título de MATEMÁTICO/A.

El citado trabajo de titulación, fue debidamente aprobado por el(los) docente(s):

Tutor	Dr. INFANTE QUIRPA, SABA RAFAEL , Ph.D.

Y recibió las observaciones de los otros miembros del Tribunal Calificador, las mismas que han sido incorporadas por el(la) estudiante.

Previamente cumplidos los requisitos legales y reglamentarios, el trabajo de titulación fue sustentado por el(la) estudiante y examinado por los miembros del Tribunal Calificador. Escuchada la sustentación del trabajo de titulación, que integró la exposición de el(la) estudiante sobre el contenido de la misma y las preguntas formuladas por los miembros del Tribunal, se califica la sustentación del trabajo de titulación con las siguientes calificaciones:
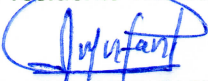
Tipo	Docente	Calificación
Tutor	Dr. INFANTE QUIRPA, SABA RAFAEL , Ph.D.	10,0
Presidente Tribunal De Defensa	Dr. AMARO MARTIN, ISIDRO RAFAEL , Ph.D.	10,0
Miembro Tribunal De Defensa	Dr. MANZANILLA MORILLO, RAUL , Ph.D.	10,0

Lo que da un promedio de: 10 (Diez punto Cero), sobre 10 (diez), equivalente a: APROBADO

Para constancia de lo actuado, firman los miembros del Tribunal Calificador, el/la estudiante y el/la secretario ad-hoc.
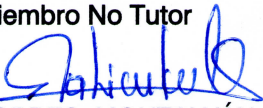
BECERRA LOAIZA, INTI ISRAEL
Estudiante

Dr. AMARO MARTIN, ISIDRO RAFAEL , Ph.D.
Presidente Tribunal de Defensa

Dr. INFANTE QUIRPA, SABA RAFAEL , Ph.D.
Tutor

Dr. MANZANILLA MORILLO, RAUL , Ph.D.
Miembro No Tutor

TORRES MONTALVÁN, TATIANA BEATRIZ
Secretario Ad-hoc

Autoría

Yo, **INTI ISRAEL BECERRA LOAIZA**, con cédula de identidad 1103912620, declaro que las ideas, juicios, valoraciones, interpretaciones, consultas bibliográficas, definiciones y conceptualizaciones expuestas en el presente trabajo; así cómo, los procedimientos y herramientas utilizadas en la investigación, son de absoluta responsabilidad del autor del trabajo de integración curricular. Así mismo, me acojo a los reglamentos internos de la Universidad de Investigación de Tecnología Experimental Yachay.

Urcuquí, febrero de 2020.

Inti Israel Becerra Loaiza
CI: 1103912620

Autorización de publicación

Yo, **INTI ISRAEL BECERRA LOAIZA**, con cédula de identidad 1103912620, cedo a la Universidad de Tecnología Experimental Yachay, los derechos de publicación de la presente obra, sin que deba haber un reconocimiento económico por este concepto. Declaro además que el texto del presente trabajo de titulación no podrá ser cedido a ninguna empresa editorial para su publicación u otros fines, sin contar previamente con la autorización escrita de la Universidad.

Asimismo, autorizo a la Universidad que realice la digitalización y publicación de este trabajo de integración curricular en el repositorio virtual, de conformidad a lo dispuesto en el Art. 144 de la Ley Orgánica de Educación Superior.

Urcuquí, febrero de 2020.

Inti Israel Becerra Loaiza

CI: 1103912620

Dedication

To my parents Julio and Teresa.
To my brothers Karina, Stalin and Vanesa.
To my beloved Mónica.

Inti Israel Becerra Loaiza

Acknowledgements

Foremost, I would like to express my sincere gratitude to my family, my parents, sisters and brother for being my fundamental pillar and having supported me unconditionally during all this time.

I thank a special person, my beloved Moni, for her continuous and inexhaustible love, support and understanding during my pursuit of bachelor's degree. You were with me at all thouse hard times. Somtimes, I thought that obtein this would be impossible. You helped me to keep on target.

Special thanks to my advidser, Prof. Saba Infante, for his continuous support in this research. His motivation, enthusiasm, and immense knowledge were my motivation for being perseverant in every step of this thesis.

Besides my advisor, I would like to thank the rest of my thesis committee: Prof. Rafael Amaro, and Prof. Raul Manzanilla, for their encouragement, advices, and hard questions.

Last but not the least, I would like to thank my friends, for the good moments, sleepless nights, and all the fun we had.

Inti Israel Becerra Loaiza

Resumen

En la actualidad existe un interés creciente en comprender las dinámicas de ciertos procesos físicos y biológicos que son parcialmente observados y que se generan a gran escala en espacio y tiempo. Los modelos estadísticos espacio temporales se vienen utilizando cada vez más en una amplia variedad de disciplinas científicas tales como el mapeo de enfermedades en determinadas regiones, la interpretación de trazas sísmicas en la industria petrolera, análisis de redes de sensores robóticos y el monitoreo de estaciones meteorológicas entre otras aplicaciones. Esta metodología es apropiada para describir y predecir los procesos espacialmente explícitos que evolucionan en el tiempo. En este trabajo se utilizó una metodología bayesiana que involucra la combinación de un filtro Kriging universal y el filtro de Kalman. Las superficies de predicción espacial del modelo se construyeron usando el algoritmo de Kriging y los efectos temporales se estimarán por el algoritmo de filtro de Kalman. El Kriging proporciona un enfoque de estimación exitoso desde el punto de vista de la estadística espacial y el filtro de Kalman facilita un procedimiento recursivo bien establecido para la estimación de los modelos en la forma espacio estado. Se utilizarón algunas medidas de bondad de ajuste para validar las predicciones del modelo. La metodologia fue ilustrada usando series de tiempo de 30 años de tres estaciones meteorológicas del Ecuador. La estructura unificada del modelo permite hacer predicciones sobre la temperatura, las precipitaciones y la humedad en las 3 provincias analizadas obteniendo buenos ajustes. Comprender los patrones espaciales y tendencias puede ayudar a evualar politicas que contribuyan a la reducción del cambio climatico. Se usó la raiz del error cuadratico medio como medida de bondad de ajuste para medir la calidad de estimacion del algoritmo obteniendose resultados satisfactorios.

Palabras clave: Filtro Kriging Kalman, filtro Kriging Universal, Filtro de Kalman, modelos espacio temporales

Abstract

Nowadays there is a growing interest in understanding the dynamics of certain physical and biological processes. Those concepts are being partially observed and are generated on a large space and time scale. The use of spatio-temporal statistical models have been increase in a wide variety of scientific disciplines such as mapping diseases in certain regions, interpretation of oil industry seismic traces, robotic sensor networks analysis, and monitoring of stations meteorological among other applications. This methodology is appropriate to describe and predict spatial explicit processes, which evolves over time. A Bayesian methodology that involves the combination of a universal Kriging filter and the Kalman filter were proposed. The spatial prediction surfaces of the model were constructed using the Kriging algorithm and the Kalman filter algorithm. It will be able to estimate the temporal effects. Kriging provides a successful estimation approach from the point of view of spatial statistics. On the other hand, the Kalman filter facilitates describe a well-established recursive procedure in order to estimate models in the form of space-state. Some measures of goodness of fit were used to validate model predictions. The methodology was illustrated using 30-year time series from three meteorological stations in Ecuador. The unified structure of the model allows predictions about temperature, precipitation and humidity in the 3 states analyzed to obtain good adjustments. Understanding spatial patterns and trends can help to evaluate policies which contribute to climate change reduction. The root mean square error was used as a measure of goodness of fit to measure the algorithm estimation quality and to get satisfactory results.

Keywords: Kriging Kalman filter, Universal Kriging filter, Kalman filter, spatio-temporal models.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In the last years, spatio-temporal modelling has developed due to the necessity of analyze the temporal evolution of spatial behavior of random magnitudes. Thouse magnitudes are important in a variety of studies developed in various applied areas such as Geo-statistics, Geophysics Environment, Hydrology, Meteorology, Biology and Medicine.

The approach that will form the basis of our space time modelling of the data is based on General State Space (GSS) model for Spatial-Temporal data (ST-GSS model) designed to model the evolution of spatial fields through time [1]. We will use the Kalman filter due to estimation is recursive. The approach combines Kriging and the Kalman filter. We call it the Kriged Kalman filter (KKF) [2]. Recall that

a) Kriging has provided a successful Weiner prediction approach in Spatial Statistics/- Geostatistics.

b) The Kalman filter gives a well-established recursive procedure for estimation in general state space models applied to time series.

Several studies have used Kriging and the Kalman filter, [2] used a KKF and give a specific implementation using pollution data. [3] built a model to predict spatial surface using the well-known method of Kriging for optimum spatial prediction and they analyzed the temporal effects using the models underlying the Kalman filtering method. [4] give a practical introduction into Kalman filtering and one of its by products, the

Ensemble Kalman Filter (EnKF) to do history matching of oil reservoirs. [5] used a Bayesian method of the kriging algorithm with the aim of predicting random Gaussian fields, taking into account the uncertainty in the covariance function. They analyzed the best predictor estimator unbiased within a Bayesian structure, considering the parameters in the structure of the covariance and the effect on the quality of real and predicted prediction. [6] assure that the EnKF is largely unknown in the statistics community and they aim to change it. Additionally, they pretend to entice more statisticians to work on this topic. In [7], a methodology was proposed to design a distributed estimation algorithm that enables a robotic sensor network taking successive measurements of a dynamic physical process modeled as a spatio-temporal random field to obtain consistent and statistically sound representations of the spatial field. A robust KKF model is used in [8] that explicitly accounts for presence of measurement outliers. Exploiting outlier sparsity, a l1-regularized estimator that jointly predicts the spatialtemporal process at unmonitored locations, while identifying measurement outliers. In [9] developed an inversion technique that combines the reduced basis (RB) method and the ensemble Kalman filter to solve state parameter identification problems for large-scale nonlinear dynamical systems arising from the discretization of nonlinear time-dependent PDEs. Kriging algorithm is used [10] because it method produced the most accurate results of daily mean wind speeds, it was proved in previous studios, and he used the Kalman filter because it does not require extensive database management and yields significant model improvement. Some notions of Bayesian analysis with emphasis on Bayesian modeling and calculation are reviewed in [11], [12]. Additionally, a general hierarchical model for times series analysis is presented and discussed. [13], [14] and [15] have concepts for spacial data that can be used for spatio-temporal modeling, exploratory data analysis, and statistical inference (estimation, prediction, uncertainty quantification). The maximum likelihood principle is used in [16], [17] and [18] for models to time series with missing data observations or incomplete data. [19] used the Kriging Kalman filter to obtain global temperature predictions given local temperature measurements. By solving the heat transfer partial differential equation driving the wildfire evolution, it is shown that the spatio-temporal

mean temperature process associated with a wildfire evolving in a finite spatial domain under certain prescribed conditions can be approximated by a Fourier series. The Kriging Kalman filter is used to predict the wildfire temperature evolution and it is compared to that of standard Gaussian process regression. A software developed in Matlab based on the Kriged Kalman Filter model for dynamic spatio-temporal interpolation of Global Navigation Satellite System (GNSS) missing data is presented in [20]. The users can load source GNSS data, set parameters, view the interpolated series and save the final results.

Other related works are: [21] used a dynamic temporal space model that allowed inference about precipitation states in weather stations at Venezuela. In that work, Markov Chains Monte Carlo algorithm were used, in particular, the Gibs algorithm is used to complete the missing data, they also used a sequential Monte Carlo algorithm known as the parallelized ensemble Kalman filter to make efficient estimates at weather stations. In [22] proposed a method based on a sensor placement method for spatio-temporal field estimation based on a kriged Kalman filter (KKF) using a network of static or mobile sensors. The developed work dynamically designs the optimal to place the sensors. They combine the estimation error with minimization problem with a sparsity-enforcing penalty. They used a general structure of covariance matrices. Adittionally, the KKF is used to estimate the field using the measurements from the selected sensing locations.

Due to the stated in the previous paragraphs, this thesis considers the analysis and modelling of daily data from 1987 to 2017 of temperature, precipitation and relative humidity of three states of Ecuador (Imbabura, Carchi and Pichincha). We use Kriging Kalman Filter to predict the temporal space evolution of environmental phenomena mentioned before.

1.1   Problem Statement

Suppose that the data $x(s_1, t_1), \cdots, x(s_N, t_N)$ are obtained from a continuous spatial process $x(\mathbf{s}, t)$, where $\mathbf{s} = (x, y) \in D \subset \mathbb{R}^2$, and $t = \{1, 2, \cdots\}$ is a discrete time index. Suppose that the observable process has a measurement error component expressed

through an observation equation

$$x(\mathbf{s}, t) = \mu(\mathbf{s}, t) + \epsilon(\mathbf{s}, t), \qquad \epsilon(s, t) \sim N(0, k(s, t)) \tag{1.1}$$

where $\mu(s, t)$ can be seen as a softer process than $x(s, t)$ and $\epsilon(s, t)$ is an error component. The objective is predict the process $\mu(\cdot, \cdot)$ in all spatial locations and at points of time of interest, regardless of where and when the data $x(s_1, t_1), \cdots, x(s_N, t_N)$ are observed. The component $\mu(s, t)$ can be expressed as a linear combination that varies in the time $\alpha(t)$ of spatial fields $h(s)$, which we can call it common fields of the space-state model

$$\mu(\mathbf{s}, t) = h_1(\mathbf{s})\alpha_1(t) + h_2(\mathbf{s})\alpha_2(t) + \cdots + h_p(s)\alpha_p(t) = \mathbf{h}(\mathbf{s})^T \boldsymbol{\alpha}(t) \tag{1.2}$$

where

$$\mathbf{h}(\mathbf{s})^T = (h_1(\mathbf{s}), h_2(\mathbf{s}), \cdots, h_p(\mathbf{s})) \qquad and$$

$$\boldsymbol{\alpha}(t) = \begin{pmatrix} \alpha_1(t) \\ \alpha_2(t) \\ \vdots \\ \alpha_p(t) \end{pmatrix}$$

The vector $\boldsymbol{\alpha}(t)$ represents the state equation of (1.2), this is

$$\boldsymbol{\alpha}(t) = P\boldsymbol{\alpha}(t - 1) + K\boldsymbol{\eta}(t), \qquad \boldsymbol{\eta}(t) \sim N(0, \Sigma_{\boldsymbol{\eta}}) \tag{1.3}$$

It is assumed that p, K and $\Sigma_{\boldsymbol{\eta}}$ are known.

1.2   Objectives

1.2.1   General Objective

To formulate a stochastic estimation process for predicting the temporal space evolution of some environmental phenomena such as: temperature, precipitation and relative humidity, in three provinces of the Republic of Ecuador.

1.2.2   Specific Objectives

The next specific objectives will be followed in order to achieve the main goal.

- To promote a statistical model witch considers the temporal space variation for predicting the dynamics of the environmental phenomena studied.

- To use the methodology based on the Bayesian Kriged Kalman filter to estimate the parameters of the model.

- To obtain daily forecast of environmental phenomenon of Pichincha, Carchi and Imbabura ecuadorian provinces.

1.3   Justification

The approach on this model is developed on the general structure of space-state models for space-time data. It is designed to model the evolution of spatial fields through the time [1]. This research proposes to implement and validate the Kriging filter methodology in combination with the Kalman filter in order to estimate and predict some unknown variables or states of the climate, such as, temperature, rainfall and relative humidity in three provinces of Ecuador (Imbabura, Carchi and Pichincha). The objective is to make contributions for environmental organizations which are in charge of studying climate change in the Republic of Ecuador.

The Kalman filter will be used because the estimation is recursive. The approach proposed a mix of the Universal Kriging filter and the Kalman filter. We call it the

Kriging Kalman filter (KKF) [2]. It is important to note that the Kriging filter has been used successfully for prediction in Spatial Statistics also called Geostatistics. On the other hand, the Kalman filter provides a well-established recursive procedure for estimation in space-state models applied to time series.

1.4 Contribution

We are going to apply modern methodologies that allow:

- To reduce the problem dimension.

- To reduce the computational cost.

- To select models that allow spatial and temporal predictions of temperature, precipitation and humidity states quickly, efficiently and accurately.

- To compare and improve theoretical models with respect to empirical models.

- To build models that describe important climate aspects.

1.5 Thesis overview

This work is divided into 6 main Chapters named as follows: Introduction, Theoretical Framework, Methodology, Data description, Result, and Conclusions.

Chapter 1, the problem statement, objectives, the justification of this work and the scientific contributions of this research are presented.

Chapter 2, the statistical techniques that are used in spatial analysis are reviewed. Also, we discuss the fundamental geo-statistical methodologies to analyse spatial and spatio-temporal data in particularly the approaches we have used in this thesis.

The estimation for the KKF is presented in Chapter 3. At the begining, we will review the Kalman Filter and its use to estimate the states $\alpha(t)$ given the data, assuming the parameters and common fields to be specified. Next, the method of maximum likelihood estimation (MLE) and its interpretation using expectation–maximization (EM) algorithm

will be discussed. Finally, we are going to analize the implementation of the algorithms that incorporates both modeling spatial in order to define the set of common fields leading to H, and the MLE and the Kalman filter.

Chapter 4 gives a description of available data which we use in this thesis. We describe data preparation, editing and cleaning. It is necessary for the raw metereological data that we obtained from the national institute of meteorology and hydrology (INAMHI). We also present summary statistics for all data sets after getting into analysable form.

Chapter 5 gives the experimental results.

In Chapter 6, the conclusions obtained from this work are presented and introduced some ideas that can be extended for future work.

Chapter 2

Theoretical Framework

In this chapter we present some useful to deal with our problem. The statistical techniques which are used in spatial analysis are reviewed. Also, we discuss the fundamental geo-statistical methodologies to analyse spatial and spatio-temporal data specially the approaches we have used in this thesis.

2.1 Spatial Data

Spatially dependent data are often classified into three major types see e.g., [23]. These are: (i) point-referenced data (ii) point pattern data, and (iii) areal data. Below we discuss these three types of data.
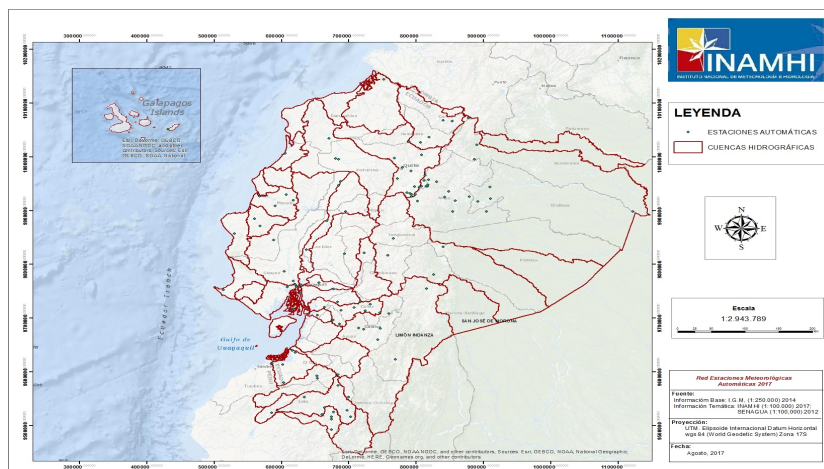
Figure 2.1: The map showing meteorological stations in Ecuador.

2.1.1   Point-referenced Data

In point-referenced data (also known as geostatistical data) the random observation $Z(s)$ is measured at a location $s \in S \subset \mathbb{R}^d$, and s varies continuously over the study region S. Theoretically, the number of locations in S is infinite. For example, see Figure 2.1 where the meteorological phenomena are monitored in several sites in Ecuador [24].

2.1.2   Point Pattern Data

In this type of spatial data, the study domain S is random and its index set gives the locations of random events that describe the observed spatial point patterns. An example of point pattern data is given in Figure 2.2, where the points represent locations of 3605 trees of the species Beilschmiedia pendula (Lauraceae) in a 1000 by 500 meter rectangular sampling region in the tropical rain forest of Barro Colorado Island [25].
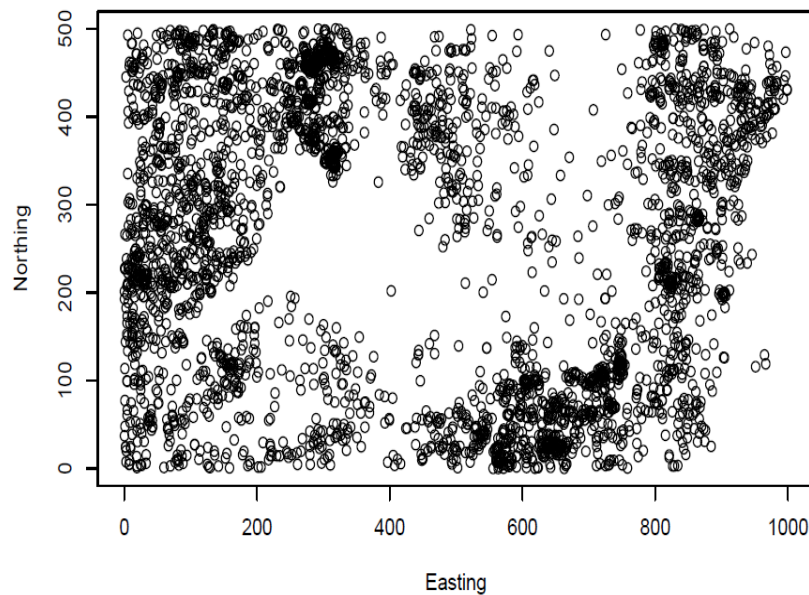
Figure 2.2: Example of point pattern data showing locations of trees in the rain forest of Barro Colorado Island.

2.1.3   Areal Data

In areal data, the study domain S is a fixed subset with regular or irregular shape, but partitioned into a finite number of areal units with well-defined boundaries.[24] For

example, Figure 2.3 shows the average 4th highest ozone concentration levels for the 33 states in the eastern US in 1997.
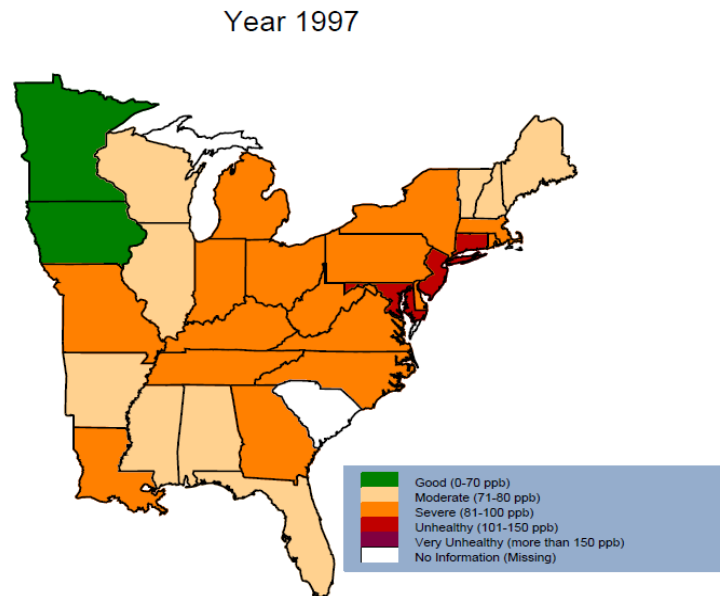
Figure 2.3: A choropleth map of the statewise average 4th highest ozone concentration levels in 1997.

2.2 Spatio-temporal data

Spatial data are traditionally thought of as random according to either geostatistical, areal or lattice, or point process (and sometimes random set) behavior. We think of geostatistical data as the kind where we could have observations of some variable or variables of interest (e.g., temperature and wind speed) at continuous locations over a given spatial domain, and where we seek to predict those variables at unknown locations in space. Lattice processes are defined on a finite or countable subset in space (e.g., grid nodes, pixels, polygons, small areas), such as the process defined by work-force indicators on a specific political geography (e.g., counties in the USA) over a specific period of time. A spatial point process is a stochastic process in which the locations of the points are random over the spatial domain, where these events can have attributes given in terms of marks (e.g., locations of trees in a forest are random events, with the diameter at breast height being the mark). Given the proliferation of various data sources and geographical

information system (GIS) software, it is important to broaden the perspective of spatial data to include not only points and polygons, but also lines, trajectories, and objects. It is also important to note that there can be significant differences in the abundance of spatial information versus temporal information. It should not be surprising that data from spatio-temporal processes can be considered from either a time-series perspective or a spatial-random-process perspective, as described in the previous paragraph [13].

2.3   Spatio-temporal models

Spatio-temporal data analysis is an emerging research area due to the development and application of novel computational techniques allowing for the analysis of large spatio-temporal databases. When the data are collected across time as well as space and has at least one spatial and one temporal property we can say that spatio-temporal models arise. An event in a spatio-temporal dataset describes a spatial and temporal phenomenon that exists at a certain time t and location x. Applications for spatio-temporal analysis include cases in the domains of biology, ecology, meteorology, medicine, transportation and forestry [13].

2.4   Dinamic spatio-temporal models

Dynamic modeling in the context of spatio-temporal data is simply the notion that we build statistical models that posit (either probabilistically or mechanistically) how a spatial process changes through time. It is inherently a conditional approach, in that we condition on knowing the past, and then we model how the past statistically evolves into the present. If the spatio-temporal phenomenon is what we call "stationary," we could take what we know about it in the present (and the past) and forecast what it will look like in the future.

Building spatio-temporal models using the dynamic approach is closer to how scientists think about the etiology of processes they study – that is, most spatio-temporal data really do correspond to a mechanistic real-world process that can be thought of as a

spatial process evolving through time.

The power of these models comes from established knowledge about the process's behavior, which may not be available for the problem at hand. In that case, one might specify more flexible classes of dynamic models that can adapt to various types of evolution, or turn to the descriptive approach and fit flexible mean and covariance functions to the data [13].

2.5   Covariograms and Semivariograms

2.5.1   Covariogram

We consider empirical spatio-temporal covariograms (and their close cousins, semivariograms) for measures of the joint spatio-temporal dependence. The characterizing of covariability in the spatio-temporal data as a function of specific lags in time and in space we are interested. Note that the lag in time is a scalar, but the lag in space is a vector (corresponding to the displacement between locations in d-dimensional space).

Consider the empirical spatio-temporal covariance function for various space and time lags. Here, we make an assumption that the first moment (mean) depends on space but not on time and that the second moment (covariance) depends only on the lag differences in space and time. Then the empirical spatio-temporal covariogram for spatial lag h and time lag τ is given by

$$\hat{C}_z(h;t) = \frac{1}{|N_s(h)|}\frac{1}{|N_t(\tau)|}\sum_{S_i,S_k \in N_s(h)}\sum_{t_j,t_l \in N_t(\tau)}\left(Z(s_i;t_j) - \hat{\mu}_{z,s}(s_i)\right)\left(Z(s_k:t_l)) - \hat{\mu}_{z,s}(S_k)\right)$$

(2.1)

where you will recall that $\hat{\mu}_{z,s}(s_i) = (1/T)\sum_{j=1}^{T}Z(s_i;t_j)$, $N_s(h)$ refers to the pairs of spatial locations with spatial lag within some tolerance of h, $N_t(\tau)$ refers to the pairs of time points with time lag within some tolerance of τ, and $|N(\cdot)|$ refers to the number of elements in $N(\cdot)$. Under isotropy, one often considers the lag only as a function of distance, $h = ||h||$, where $||\cdot||$ is the Euclidean norm [13].

2.5.2   Semivariogram

The semivariogram is defined as:

$$\gamma_z(s_i, s_k; t_j, t_l) = \frac{1}{2} var\left(Z(s_i; t_l) - Z(s_k; t_l)\right)$$

In the case where the covariance depends only on displacements in space and differences in time, this can be written as

$$\begin{aligned}
\gamma(h; \tau) &= \frac{1}{2} var\left(Z(s+h; t+\tau) - Z(s; t)\right) \\
&= C_z(0; 0) - cov\left(Z(s+h; t+\tau) - Z(s; t)\right) \\
&= C_z(0; 0) - C_z(h; \tau)
\end{aligned} \tag{2.2}$$

where $h = s_k - s_i$ is a spatial and $\tau = t_l - t_j$ is a temporal lag. Now, equation (2.2) does not always hold. It is possible that γ_z is a function of spatial lag h and temporal lag τ, but there is no stationary covariance function $C_z(h; \tau)$. We generally try to avoid these models of covariability by fitting trend terms that are linear and/or quadratic in spatio-temporal coordinates.
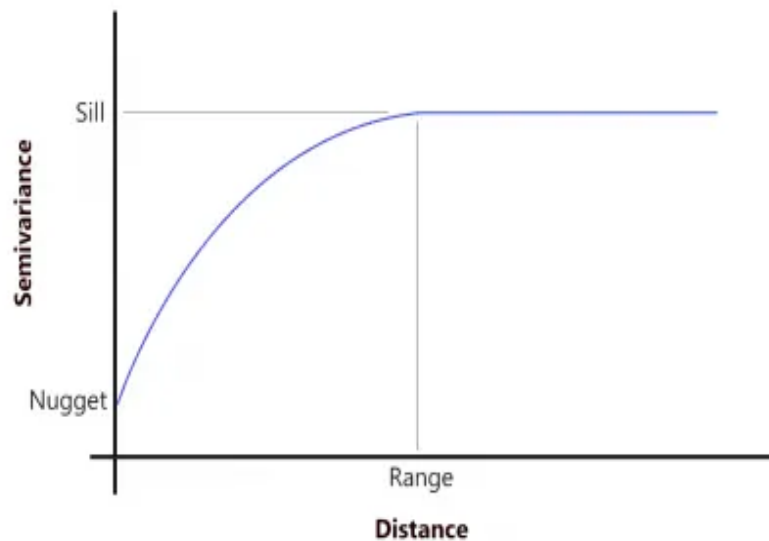
Figure 2.4: Semivariogram.

If the covariance function of the process is well defined, then the semivariogram is generally characterized by the nugget effect, the sill, and the partial sill. The nugget effect is given by $\gamma_z(h; \tau)$ when $h \to 0$ and $\tau \to 0$, while the sill is $\gamma_z(h; \tau)$ when $h \to \infty$ and $\tau \to \infty$. The partial sill is the difference between the sill and the nugget effect. The diagram below shows these components of a semivariogram as a function of spatial distance $\|h\|$.[13]

2.6   Hierarchical Statistical Models

Hierarchical modeling is based on the basic fact from probability theory that a collection of random variables with joint distribution can be decomposed into a marginal distribution and a series of conditional distributions. That is, if A, B, C are random variables, then we can write the joint distribution in terms of factorizations, such as $[A, B, C] = [A|B, C][B|C][C]$, where the bracket notation $[C]$ refers to a probability distribution for C, and $[B|C]$ refers to the conditional probability distribution of B given C, etc. For a spatial process, the joint distribution describes the stochastic behavior of the spatially referenced data and parameters. This can be difficult (if not impossible) to specify for many problems. It is often much easier to specify the distribution of the relevant conditional models (e.g., conditioning the observed data on the true process and parameters, etc.). In this case, the product of a series of relatively simple conditional models leads to a joint distribution that can be quite complex.

For complicated processes in the presence of data, it is useful to approach the problem by breaking it into three primary stages [11]:

1. Data Model: $[data|process, parameters]$

2. Process Model: $[process|parameters]$

3. Parameter Model: $[parameters]$

The first stage is concerned with the observational process or "data model," this model or process specifies the distribution of the data given the process of interest as well as

the parameters that describe the data model. The second stage describes a distribution, conditional on other parameters, for the process,. The last stage referees to the uncertainty in the parameters by endowing them with distributions. In general, each of these stages may have multiple substages.

Ultimately, we are interested in the "posterior" distribution, i.e., the distribution of the process and parameters updated by the data. This is obtained by Bayes' rule in which the posterior distribution is proportional to the product of the data, process, and parameter distributions:

$$[process, parameters|data] \propto [data|process, parameters] \times [process|parameters][parameters]$$

where the normalizing constant represents the integral of the right-hand side with respect to the process and parameters. This formula serves as the basis for hierarchical Bayesian analysis [14].

2.7   Kriging

We suppose that a spatially distributed variable is of interest, which in theory is defined at every point over a bounded study region of interest, $S \subset \mathbb{R}^d$, where $d = 2, 3$. Also, We suppose further that this variable has been observed at each of distinct points in S, and that from the observations we wish to make inferences about the process that governs how this variable is distributed spatially and about values of the variable at locations where it was not observed [14].

The method used to achieve these objectives is known as kriging, it is very similar to a multiple linear regression applied to a spatial context, where the random variables $Z(s)$ act as regressor variables, and the random variable at the point where the prediction is interesting, $Z(s_0)$, serves as the dependent variable. This is a local estimation technique which has the quality of being the best unbiased linear estimator of Z. [26]

In the classical Geostatistic analysis, the final stage is the prediction of values of $Z(s)$ at desired locations, perhaps even at all points. The South African mining engineer Danie

Krige, who was the first to develop and apply them by what, all methods dedicated to this purpose are called kriging.

The kriging predictor depends of the model adopted for the random function $Z(s)$: In general, $Z(s)$ is usually split into a trend component and a residual component, as expressed in the equation

$$Z(s) = \mu(s) + \epsilon(s) \tag{2.3}$$

where $\mu(s) = E[Z(s)]$ is the average function (the mean is assumed to be constant) and $\epsilon(s)$ is the residual component of which the variogram or the covariogram is supposed to be known.

The kriging variants depend on the model adopted for the trend $\mu(s)$, following describes Simple, Ordinary and Universal kriging. In [15] this method is found in greater detail and also others such as: indicator, lognormal, transgaussian, robust kriging and cokriging for the multinomial case.

2.7.1 Simple Kriging

Let the stochastic response $\epsilon(s)$ (point-referenced data) at site s be strictly stationary, so that it is written as:

$$Z(s) = \mu(s) + \epsilon(s) \tag{2.4}$$

where, $\mu(s)$ is a known function and $\epsilon(s)$ is the spatial error process and assumed it to be Gaussian with mean zero and covariance matrix Σ. Let $Z(s) = (z(s_1), ..., z(s_n))^T$. The mean and variance of the process is written as, $E(Z(s)) = \mu(s)$ and $Var(Z(s)) = \Sigma$. To obtain prediction at unknown site s_0, we can estimate the optimal prediction $Z(s_0)$ as:

$$\hat{Z}(s_0) = \hat{\mu}(s_0) + C^T \Sigma^{-1}(Z(s) - \mu(s))$$

where, $C^T = cov(Z(s), Z(s_0))$. This type of kriging is known as simple kriging.[24]

2.7.2 Ordinary Kriging

Assume that the mean process $\mu(s) = \mu$ is known and does not vary with spatial locations s, hence the model in equation (2.4) is written as:

$$Z(s) = \mu(s) + \epsilon(s)$$

The estimated optimal prediction $Z(s_0)$ at site s_0 is known as the ordinary kriging, and is written as:

$$\hat{Z}(s_0) = \hat{\mu} + C'\Sigma^{-1}(Z(s) - \mu)$$

where, $\hat{\mu} = (1'\Sigma^{-1}1)^{-1}1'\Sigma^{-1}Z(s)$, and 1 is a vector with all elements equal to 1.[24]

2.7.3 Universal Kriging

Assume the mean process $\mu(s)$ is unknown and it varies over space in the linear regression form $\mu(s) = X(s)'\beta$, and the covariance function Σ is known as in the model (2.4). The model is written as:

$$Z(s) = X^T(s)\beta + \mu(s)$$

where, $\mu(s) = (\mu(s_1), \cdots, \mu(s_n))'$ and $\mu(s) \sim N(0, \Sigma)$, $Z(s) = (Z(s_1), \cdots, Z(s_n))'$, $\beta = (\beta_1, ..., \beta_p)'$ is the p parameters and $X^T(s)$ is the pxn covariate matrix.

Hence, we can estimate the optimal prediction at site s_0 as:

$$\hat{Z}(s_0) = X^T(s)\hat{\beta} + C'\Sigma^{-1}(Z(s) - X^T(s)\hat{\beta})$$

where, $\hat{\beta} = (X^T(s)\Sigma^{-1}X(s))^{-1}X^T(s)\Sigma^{-1}Z(s)$.

This type of kriging is known as the universal kriging [24].

2.8   Kalman Filter

Kalman filters are used to estimate states based on linear dynamical systems in state space format [27]. The process model defines the evolution of the state from time $k-1$ to time k as:

$$\boldsymbol{x}_k = F\boldsymbol{x}_{k-1} + B\boldsymbol{u}_{k-1} + \boldsymbol{w}_{k-1} \tag{2.5}$$

where F is the state transition matrix applied to the previous state vector $\boldsymbol{x}_{k-1}$, B is the control input matrix applied to the control vector $\boldsymbol{u}_{k-1}$, and $\boldsymbol{w}_{k-1}$ is the process noise vector that is assumed to be zero-mean Gaussian with the covariance Q, i.e. - $\boldsymbol{w}_{k-1} \sim N(0, Q)$. The process model is paired with the measurement model that describes the relationship between the state and the measurement at the current time step k as:

$$\boldsymbol{z}_k = H\boldsymbol{x}_k + \boldsymbol{\nu}_k \tag{2.6}$$

where $\boldsymbol{z}_k$ is the measurement vector, H is the measurement matrix, and $\boldsymbol{\nu}_k$ is the measurement noise vector that is assumed to be zero-mean Gaussian with the covariance R, i.e., $\boldsymbol{\nu}_k \sim N(0, R)$ Note that sometimes the term "measurement" is called "observation" in different literature.

The role of the Kalman filter is to provide estimate of $\boldsymbol{x}_k$ at time k, given the initial estimate of $\boldsymbol{x}_0$, the series of measurement, $\boldsymbol{z}_1, \boldsymbol{z}_1, \cdots, \boldsymbol{z}_k$ and the information of the system described by F, B, H, Q, and R. Note that subscripts to these matrices are omitted here by assuming that they are invariant over time as in most applications. Although the covariance matrices are supposed to reflect the statistics of the noises, the true statistics of the noises is not known or not Gaussian in many practical applications [28]. Therefore, Q and R are usually used as tuning parameters that user can adjust to get desired performance.

Kalman filter algorithm consists of two stages: prediction and update. Note that terms "prediction" and "update" are often called "propagation" and "correction," respectively, in different literature. The Kalman filter algorithm is summarized as follows:

Prediction	
Predicted state estimate	$\boldsymbol{x}_k^- = F\hat{\boldsymbol{x}}_{k-1}^+ + B\boldsymbol{u}_{k-1}$
Predicted error covariance	$P_k^- = FP_{k-1}^+F^T + q$

Update:	
Measurement residual	$\hat{\boldsymbol{y}} = \boldsymbol{z}_k - H\hat{\boldsymbol{x}}_k^-$
Kalman gain	$K_k = P_k^-H^T(R + HP_k^-H^T)^{-1}$
Updated state estimate	$\boldsymbol{x}_k^+ = \boldsymbol{x}_k^- + K_k\hat{\boldsymbol{y}}$
Updated error covariance	$P_k^+ = (I - K_kH)P_k^-$

Table 2.1: The Kalman filter algorithm

In the above equations, the hat operator ($\hat{}$) means an estimate of a variable. That is, $\hat{\boldsymbol{x}}$ is an estimate of $\boldsymbol{x}$. The superscripts – and + denote predicted (prior) and updated (posterior) estimates, respectively.

The predicted state estimate is evolved from the updated previous updated state estimate. The new term P is called state error covariance. It encrypts the error covariance that the filter thinks the estimate error has. Note that the covariance of a random variable $\boldsymbol{x}$ is defined as $cov(\boldsymbol{x}) = \mathbb{E}\left[(\boldsymbol{x} - \hat{\boldsymbol{x}})(\boldsymbol{x} - \hat{\boldsymbol{x}})^T\right]^T$ where $\mathbb{E}$ denotes the expected (mean) value of its argument. One can observe that the error covariance becomes larger at the prediction stage due to the summation with Q, which means the filter is more uncertain of the state estimate after the prediction step.

In the update stage, the measurement residual $\tilde{\boldsymbol{y}}_k$ is computed first. The measurement residual, also known as innovation, is the difference between the true measurement,$\boldsymbol{z}_k$, and the estimated measurement,$H\hat{\boldsymbol{x}}_k^-$. The filter estimates the current measurement by multiplying the predicted state by the measurement matrix. The residual, $\tilde{\boldsymbol{y}}_k$, is later then multiplied by the Kalman gain,K_k, to provide the correction, $K_k\tilde{\boldsymbol{y}}_k$, to the predicted estimate $\hat{\boldsymbol{x}}_k^-$. After it obtains the updated state estimate, the Kalman filter calculates the updated error covariance, P_k^+, which will be used in the next time step. Note that the updated error covariance is smaller than the predicted error covariance, which means the filter is more certain of the state estimate after the measurement is utilized in the update stage.

We need an initialization stage to implement the Kalman filter. As initial values, we need the initial guess of state estimate, $\hat{\boldsymbol{x}}_0^+$, and the initial guess of the error covariance matrix, P_0^+. Togetherwith Q and R, $\hat{\boldsymbol{x}}_0^+$ and P_0^+ play an important role to obtain desired performance. There is a rule of thumb called "initial ignorance", which means that the user should choose a large P_0^+ for quicker convergence. Finally, one can obtain implement a Kalman filter by implementing the prediction and update stages for each time step, $k = 1, 2, 3, \ldots$, after the initialization of estimates.

Note that Kalman filters are derived based on the assumption that the process and measurement models are linear, i.e., they can be expressed with the matrices F, B, and H, and the process and measurement noise are additive Gaussian. Hence, a Kalman filter provides optimal estimate only if the assumptions are satisfied [28].

Chapter 3

Methodology

Our estimation for the KKF consists of several parts: first, we are going to review the Kalman Filter in scene of spatio-temporal models and its use to estimate the states $\alpha(t)$ given the data. We will assume the parameters and common fields to be specified. Second, we obtein the common fields with the method of maximum likelihood estimation (MLE). Also,we use that interpretation to get EM algorithm of the parameters $P, \Sigma_{\boldsymbol{\eta}}$, and Σ_{ϵ} data about the common fields, the matrix H. Additionally, we discuss implementation of the algorithms. For that reason we review some overall strategic concerns and turn to specifics of the implementation of the full algorithm that incorporates both spatial modeling to define the set of common fields leading to H, the MLE, and the Kalman filter.

3.1   Principal fields

The systematic spatial component in $x(\mathbf{s}, t)$ is modeled as linear combinations that vary in the time of p common fields, comprising q trend fields and $r = p - q$ main fields, which is written as:

$$\mathbf{h}(\mathbf{s})^T = (h_1(\mathbf{s}), \cdots, h_q(\mathbf{s}), h_{q+1}(\mathbf{s}), \cdots, h_{q+r}(\mathbf{s})) \tag{3.1}$$

the fields are found considering components of common tendency, and common spatial dependence through time. Consider the spatial linear model

$$x(\mathbf{s}) = f^T(\mathbf{s})\beta + \zeta(\mathbf{s}) \tag{3.2}$$

where $\beta = (\beta_q, \cdots, \beta_q)^T$, $f(\mathbf{s}) = (f_1(\mathbf{s}), \cdots, f_q(\mathbf{s}))^T$ y los $f_j(\mathbf{s})$ $(j = 1, \cdots, q)$ are fields of tendency of a polynomial form given in the coordinates of s. Let

$$cov(\zeta(\mathbf{s}), \zeta(\mathbf{s}')) = \sigma_\zeta(\mathbf{s}, \mathbf{s}') \tag{3.3}$$

a positive definite conditional function. Suppose, that the observations $\mathbf{x} = (x_1, \cdots, x_m)$ are taken in m locations, $m \geq p$, $\mathbf{s}_1^*, \cdots, \mathbf{s}_m^*$. The common spatial dependence through time is expressed as the set of krigring predectores, possibly with restrictions because it is an array of not complete range for the set of observations on these sites.

We write $\boldsymbol{\sigma}_\zeta(\mathbf{s})$ for the m-vector of covariances with i-th element $\sigma_\zeta(\mathbf{s}_i^*, \mathbf{s})$, and we write Σ_ζ as the matrix of covariance $m \times m$ with $(\Sigma_\zeta)_{ij} = \sigma_\zeta(\mathbf{s}_i^*, \mathbf{s}_j^*)$. Let f_j is the m-vector of j-th field trend in the sites, are the i-th element $f_j(\mathbf{s}_i^*)$.

The kriging predictor is

$$\widehat{x}(\mathbf{s}) = \mathbf{f}(\mathbf{s})^T A \mathbf{x} + \boldsymbol{\sigma}_\zeta(\mathbf{s})^T B \mathbf{x} \tag{3.4}$$

where, writing $F : m \times q$ for the trend matrix with ij-th element $f_j(\mathbf{s}_i^*)$,

$$A = (F^T \Sigma_\zeta^{-1} F)^{-1} F^T \Sigma_\zeta^{-1} \tag{3.5}$$

is the trend matrix, and

$$B = \Sigma_\zeta^{-1} - \Sigma_\zeta^{-1} F (F^T \Sigma_\zeta^{-1} F)^{-1} F^T \Sigma^{-1} \tag{3.6}$$

is the partial information matrix or generalized bending energy matrix, where we assume that Σ_ζ and $F^T \Sigma_\zeta^{-1} F$ be non-singular. It is assumed that the columns of F are

linearly independent. It is considered the spectral decomposition of B.

$$B = UDU^T, \qquad B\mathbf{u}_i = d_i\mathbf{u}_i \qquad (3.7)$$

where $U = (\mathbf{u}_1, \cdots, \mathbf{u}_m)$ y $D = diag\{d_1, \cdots, d_m\}$. At least q of the eigenvalues d_i are iqual to zero. Writes the eigenvalues in non-drecreasing order, $0 = d_1 = d_2 = \cdots = d_q < d_{q+1} < d_{q+2} < \cdots < d_m$. Energy matrix B has maximum rank $m - q$, and the vectors of B are orthogonal to the columns of F, $BF = 0$ i.e. $\mathbf{f}_j^T\mathbf{u}_i = 0$, para $j = 1, \cdots, q$ y $q + 1, \cdots, m$.

Suppose that the observations taken at the chosen sites coincide with the i-th eigenvector of B, $i = 1, \cdots, m$. The Kriging predictor given in (3.4) is written as:

$$\widehat{x}(\mathbf{s})_{x=u_t} = \mathbf{f}(\mathbf{s})^T A\mathbf{u}_i + \boldsymbol{\sigma}_\zeta(\mathbf{s})^T B\mathbf{u}_i$$
$$= \mathbf{f}(\mathbf{s})^T A\mathbf{u}_i + d_i\boldsymbol{\sigma}_\zeta(\mathbf{s})^T \mathbf{u}_i \qquad (3.8)$$

Any m-vector of x observations is the linear conbination $U^T x$ of $\mathbf{u}_i$, and the predictor Kriging $\widehat{x}(\mathbf{s})$ is the linear combination of $U^T x$ of the $\widehat{x}(\mathbf{s})_{x=\mathbf{u}_i}$ in (3.8). Therefore $\widehat{x}(\mathbf{s})$ is the linear combination of trend fields $f_j(\mathbf{s})$ $(j = 1, \cdots, q)$ and the $m - q$ main fields $\boldsymbol{\sigma}_\zeta(\mathbf{s})^T\mathbf{u}_i$, $(i = q + 1, \cdots, m)$ evaluated in $\mathbf{s}$. These functions encompass the appearance of all Kriging solutions with observations in the m given sites, and the specified trend fields and covariances.

In practice the main r fields used in the KKF model may include either all $r = m - q$ or a subset$r < m - q$ of the main fields. So

$$p = q + r \leq m$$

[29] notes, for thin plates splines, the association of small eigenvalues of the energy matrix with large scalar variation.

Let $\{j_1, \cdots, j_r\}$ denote a subset of size $r \leq m - q$ of $\{q + 1, \cdots, m\}$. Then the average

component in the observation equation of the KKF model has the form

$$\mu(\mathbf{s}, t) = \sum_{j=1}^{q} h_j(\mathbf{s})\alpha_j(t) + \sum_{k=1}^{r} h_{q+k}(\mathbf{s})\alpha_{q+k}(t) \tag{3.9}$$

where

$$h_j(\mathbf{s}) = f_j(\mathbf{s}), \qquad\qquad j = 1, \cdots, q \tag{3.10}$$

$$h_{q+k}(\mathbf{s}) = \boldsymbol{\sigma}_\zeta(\mathbf{s})^T \mathbf{u}_{j_k}, \qquad k = 1, \cdots, r \tag{3.11}$$

When the complement set of main fields is used with non-zero eigenvalues is using $r = m - q$, then we write $j_k = q + k$.

The Kriging predictor depends on the covariogram, the spatial pattern of the points and the trend fields in the matrix $\mathbf{F}$.

Therefore, for the KKF we need develop methods to determine a suitable set, possibly optimal set of sites, while we choose and adjust an appropriate covariogram model identifying an $\mathbf{F}$ matrix, and we choose all the subsets of main fields.

3.2   Specifying the temporal component

In this section, we consider several specifications of the temporary component of the ST-GSS model. This we leads in some way to specify the transition matrix $\mathbf{P}$ of order $p \times p$ of the system equations, so that, for particular options of the temporal component, $\mathbf{P}$ can be fixed or it can have a bit parameters instead of the nominal set of p^2 elements.

We can discuss two methods called autoregressive modeling and dynamic linear modeling through an increased Holt-Winters model.

The multivariate configuration increase considerably the complexity of model specification and selection, since in principle it can associate a univariate time series model in each common field and take the additional cross-correlation terms.

In our applications, normally a single structure is repeated for each common field and we take the cross-correlation terms to be zero.

3.2.1   Autoregressive specification

The structured space state form of the autoregressive model includes structural zeros in the parameter vector of the observation equation. These correspond to elements of the state vector, which although they are essential for the update, they do not directly select the effect of media component.

We write for each $\mathbf{s}$,

$$x(\mathbf{s}, t) = \mathbf{h}^\dagger(\mathbf{s})^T \boldsymbol{\alpha}(t) + \epsilon(\mathbf{s}, t), \qquad (3.12)$$

where for an autoregressive model of order u, $AR(u)$,

$$\boldsymbol{h}^\dagger(\mathbf{s})^T = (h_1(\mathbf{s}), \mathbf{0}_{u-1}^T, h_2(\mathbf{s}), \mathbf{0}_{u-1}^T, \cdots, h_{q+r}(\mathbf{s}), \mathbf{0}_{u-1}^T) \qquad (3.13)$$

an $(q+r)u$ vector, denoting by $\mathbf{0}_{u-1}^T$ a row vector of $(u-1)$ zeros.

Note that the simplest AR(u) model would have a single vector of parameters, $\mathbf{h}^\dagger(\mathbf{s})^T = (h_1(\mathbf{s}), \mathbf{0}_{u-1}^T)$, length u. The different fields of zeros $h_1(\mathbf{s}), \cdots, h_{q+r}(\mathbf{s})$ comprise q trend fields and r main fields. The quantity p denotes the dimension of the state vector of the GSS model, that is, the lengths of $\mathbf{h}^\dagger(\mathbf{s})$ and $\boldsymbol{\alpha}(t)$, and $p = (q+r)u$.

From the equation (1.2), we say that there are p common fields that comprise q trend fields, r main fields, and $p_0 = (q+r)(u-1)$ null fields with $p = q + r + p_0$.

The transition matrix $\mathbf{P}$ of state equation is a diagonal block $p \times p$ with $(q+r)$ identical blocks, every $u \times u$ and is written Q_u:

$$P = blockdiag\{Q_u, \cdots, Q_u\} \qquad (3.14)$$

where

$$\begin{pmatrix} \phi_p & 1 & 0 & 0 & \cdots & 0 \\ \phi_{p-1} & 0 & 1 & 0 & \cdots & 0 \\ \phi_{p-2} & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \phi_2 & 0 & 0 & 0 & \cdots & 1 \\ \phi_1 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \tag{3.15}$$

when $u = 1$, the AR(1) model, then $p_0 = 0$, $Q_u = [\phi]$ and

$$P = diag\,\{\phi, \cdots, \phi\} = \phi I_{q+r} \tag{3.16}$$

The transition matrix has a respective block diagonal form, this is not true when the components associated with each common field follows different autoregressive models. Then we can have, for example, when cuando $u = 1$,

$$P = diag\,\{\phi, \cdots, \phi_{q+r}\} \tag{3.17}$$

Note that this state-space form of the autoregressive model is not unique but is minimal [16].

3.2.2 Dynamic linear model

An alternative model is one based on the dynamic linear model [12]. The simplest dynamic linear model is the Holt-Winters model, based on slope and level. Consider the ST-GSS model with $h^\dagger(s)$ is equal to an AR ($u = 2$),

$$\mathbf{h}^\dagger(\mathbf{s})^T = (H_1(\mathbf{s}), 0, h_2(\mathbf{s}), 0, \cdots, h_{q+r}(\mathbf{s}), 0). \tag{3.18}$$

Matrix transition $\mathbf{P}$ is $p \times p$, $p = 2(q + r)$,

$$P = blockdiag\,\{Q_2^*, \cdots, Q_2^*\} \tag{3.19}$$

where

$$Q_2^* = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

This completely specifies the P matrix, reducing the number of parameters to be estimated.

For the Holt-Winters model, we have

$$\mu(\mathbf{s}, t) = h_1(\mathbf{s})\alpha_1(t) + h_2(\mathbf{s})\alpha_3(t) + \cdots + h_j(\mathbf{s})\alpha_{2j-1}(t) + \cdots + h_{q+r}(\mathbf{s})\alpha_{p-1}(t) \quad (3.20)$$

where each $\alpha_1(t), \alpha_3(t), \cdots, \alpha_{p-1}(t)$ follows a linear trajectory in the mean. Write $\mathbf{k}_{2j-1}$ and $\mathbf{k}_{2j}$ for the $(2j-1)'$th and $2j'$th rows of the innovation parameter matrix $\mathbf{K}$. Then

$$\alpha_{2j-1}(t) = \alpha_{2j-1}(t-1) + \alpha_{2j}(t-1) + \mathbf{k}_{2j-1}\boldsymbol{\eta}(t)$$
$$\alpha_{2j}(t) = \alpha_{2j}(t-1) + \mathbf{k}_{2j}\boldsymbol{\eta}(t) \quad (3.21)$$

that is $\alpha_{2j-1}(t)$ the scalar multiplier of $h_j(s)$, increases by $\alpha_{2j-1}(t-1)$ each unit of time up to stochastic innovations, with level $\alpha_{2j-1}(1)$ at time $t = 1$.

An interesting alternative is to incorporate separate spatial fields for level $h_L(s)$ and slope $h_S(s)$. With $p = 3$ common fields, includes a null field

$$\mathbf{h}^\dagger(\mathbf{s})^T = (h_L(\mathbf{s}), h_S(\mathbf{s}), 0)$$
$$\boldsymbol{\alpha}(t)^T = (\alpha_L(t), \alpha_S(t), \alpha_S'(t))$$

and transition matrix 3×3 given by

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \quad (3.22)$$

The base spatial field is the random multiple $\alpha_L(t)$ of $h_L(\mathbf{s})$, where $\alpha_L(t)$ follows Brownian motion centered at $\alpha_L(1)$. The contribution from the slope is the random multiple $\alpha_S(t)$ of $h_S(\mathbf{s})$, where $\alpha_S(t)$ follows its own scalar level plus slope model with base value $\alpha_S(1)$ and slope $\alpha'_S(t)$.

This pattern can be repeated to include several sets of 3 state parameters, and fields possibly related to trends and main fields. However, for the case above with $p = 3$, none of the common fields $h_L(\mathbf{s})$ and $h_S(\mathbf{s})$ will be either a trend field or a main field, except in very special cases. Effectively, the model can express $h_l(\mathbf{s})$ y $h_S(\mathbf{s})$ as a linear combination of the trend and the main fields. The coefficients of the linear combinations are converted into parameters in the observation equation. These fields are given by

$$\mathbf{h}(\mathbf{s})^T = (h_1(\mathbf{s}), h_2(\mathbf{s}), \cdots, h_{q+r}(\mathbf{s}))$$

$$h_L(\mathbf{s}) = \boldsymbol{\gamma}_L^T \mathbf{h}(\mathbf{s}) \tag{3.23}$$

$$h_S(\mathbf{s}) = \boldsymbol{\gamma}_S^T \mathbf{h}(\mathbf{s}) \tag{3.24}$$

The mean field at time t can be written

$$\mu(\mathbf{s}, t) = \mathbf{h}(\mathbf{s})^T \boldsymbol{\alpha}^+(t)$$

where

$$\boldsymbol{\alpha}^+(t) = \alpha_L(t)\gamma_L + \alpha_S(t)\gamma_S$$

Note that the above is essentially an ARIMA(0,2,2) model, a Holt-Winters model with the additional term comprising the base level $h_L(s)$.

3.3   The Kalman filter recursion

Now suppose that we have observations in n sites, $\mathbf{s}_i$, $i = 1, \cdots, n$ in T times, $t = 1, \cdots, T$. Write the t-th observation vector as

$$\mathbf{x}_t^T = (x(s_1, t), \cdots, x(s_n, t)), \tag{3.25}$$

and collect these into the $T \times n$ matrix of observations denoted $\mathbf{X}$ with t'th row x_t^T.

Suppose that the $n \times p$ parameter matrix is

$$H = \begin{pmatrix} \mathbf{h}^T(s_1) \\ \vdots \\ \mathbf{h}^T(s_n) \end{pmatrix} \tag{3.26}$$

The observation equation is

$$\mathbf{x}_t = H\boldsymbol{\alpha}(t) + \epsilon(t) \tag{3.27}$$

then the model in the space-state form is

$$\mathbf{x}_t = H\boldsymbol{\alpha}(t-1) + \epsilon(t); \qquad (O.E) \qquad \epsilon(t) \sim N(0, \Sigma_\epsilon)$$

$$\boldsymbol{\alpha}(t) = P\boldsymbol{\alpha}(t) + K\boldsymbol{\eta}_t; \qquad (S.E) \qquad \boldsymbol{\eta}_t \sim N(0, \Sigma_\eta)$$

The KKF model includes spatial parameters in the covariogram model and temporal parameters in the GSS model.

If the spatial parameters were known, maximum likelihood estimators may be obtained of the temporal parameters using established Kalman terms.

Let $\mathbf{a}_{t-1|t-1}, \mathbf{a}_{t|t-1}$, y $\mathbf{a}_{t|t}$ the estimators of $\boldsymbol{\alpha}(t-1), \boldsymbol{\alpha}(t)$, y $\boldsymbol{\alpha}(t)$ based on the observations available to times $t-1$ and t respectively. Let $C_{t-1|t-1}, C_{t|t-1}$ y $C_{t|t}$ the covariances of $\mathbf{a}_{t-1|t-1}, \mathbf{a}_{t|t-1}$, and $\mathbf{a}_{t|t}$ respectively.

Given $\mathbf{a}_{t-1|t-1}$ and $C_{t-1|t-1}$, and using the system equation (S.E) y observation equation (O.E), we obtain the equations for the Kalman filter states.

3.4   Maximum likelihood estimation

Let's suppose

$$E(\boldsymbol{\alpha}_{t-1}|x_{1:t-1}) = \mathbf{a}_{t-1|t-1} \tag{3.28}$$

$$Var(\boldsymbol{\alpha}_{t-1}|x_{1:t-1}) = C_{t-1|t-1} \tag{3.29}$$

it known at $t-1$ time. The optimal predictor of $\boldsymbol{\alpha}_t$ its associated mean squared error at $t-1$ time is given by

$$E(\boldsymbol{\alpha}_t|x_{1:t-1}) = \mathbf{a}_{t|t-1} \quad y \quad Var(\boldsymbol{\alpha}_t|x_{1:t-1}) = C_{t|t-1} \tag{3.30}$$

The equation (3.30) is known as the prediction equation and is obtained as follows

$$
\begin{aligned}
E(\boldsymbol{\alpha}_t|x_{1:t-1}) &= E(p\boldsymbol{\alpha}_{t-1} + k\boldsymbol{\eta}_t|x_{1:t-1}) \\
&= pE(\boldsymbol{\alpha}_{t-1}|x_{1:t-1}) + KE(\boldsymbol{\eta}_t|x_{1:t-1}) \\
&= p\mathbf{a}_{t-1|t-1} \\
&= \mathbf{a}_{t|t-1}
\end{aligned}
$$

$$
\begin{aligned}
Var(\boldsymbol{\alpha}_t|x_{1:t-1}) &= Var(p\boldsymbol{\alpha}_{t-1} + k\boldsymbol{\eta}_t|x_{1:t-1}) \\
&= pVar(\boldsymbol{\alpha}_{t-1}|x_{1:t-1})p^T + Var(k\boldsymbol{\eta}_t|x_{1:t-1}) \\
&= pC_{t-1|t-1}p^T + kVar(\boldsymbol{\eta}_t|x_{1:t-1})K^T \\
&= PC_{t-1|t-1}p^T + k\Sigma_{\boldsymbol{\eta}}k^T \\
&= C_{t|t-1}
\end{aligned}
$$

$$
\begin{aligned}
E(x_t|x_{1:t-1}) &= E(H\boldsymbol{\alpha}_t + \epsilon_t|x_{1:t-1}) \\
&= HE(\boldsymbol{\alpha}_t|x_{1:t-1}) + E(\epsilon_t|x1:t-1) \\
&= H\mathbf{a}_{t|t-1}
\end{aligned}
$$

$$Var(x_t|x_{1:t-1}) = Var(H\boldsymbol{\alpha}_t + \epsilon_t|x_{1:t-1})$$
$$= HVar(\boldsymbol{\alpha}_t|x_{1:t-1})H^T + Var(\epsilon_t|x_{1:t-1})$$
$$= HC_{t|t-1}H^T + \Sigma_\epsilon$$

$$Cov(\boldsymbol{\alpha}_t, x_t|x_{1:t-1}) = Cov\left\{\boldsymbol{\alpha}_t, E(x_t|\boldsymbol{\alpha}_t)|x_{1:t-1}\right\}$$
$$= Cov\left\{\boldsymbol{\alpha}_t, E(H\boldsymbol{\alpha}_t + \epsilon_t|\boldsymbol{\alpha}_t)|x_{1:t-1}\right\}$$
$$= Cov\left\{\boldsymbol{\alpha}_t, H\boldsymbol{\alpha}_t|x_{1:t-1}\right\}$$
$$= Var(\boldsymbol{\alpha}_t|x_{1:t-1})H^T$$
$$= C_{t|t-1}H^T$$

Using the Bayes linear adjustment,

$$E(\boldsymbol{\alpha}_t|x_{1:t-1}) = E(\boldsymbol{\alpha}_t|x_{1:t-1}) + cov(\boldsymbol{\alpha}_t, x_t|x_{1:t-1})Var^{-1}(x_t|x_{1:t-1})[x_t - E(x_t|x_{1:t-1})]$$
(3.31)

$$Var(\boldsymbol{\alpha}_t|x_{1:t}) = Var(\boldsymbol{\alpha}_t|x_{1:t-1}) - Cov(\boldsymbol{\alpha}_t, x_t|x_{1:t-1})Var^{-1}(x_t|x_{1:t-1})Cov^T(\boldsymbol{\alpha}_t, x_t|x_{1:t-1})$$
(3.32)

We get optimal state estimator update equations x_t, given information $x_1 : t = (x_1, \cdots, x_t)$ at time t, that is

$$\mathbf{a}_{t|t} = \mathbf{a}_{t|t-1} + C_{t|t-1}H^T[HC_{t|t-1}H^T + \Sigma_\epsilon]^{-1}[x_t - H\mathbf{a}_{t|t-1}] \tag{3.33}$$

$$C_{t|t} = C_{t|t-1} - C_{t|t-1}H^T(HC_{t|t-1}H + \Sigma_\epsilon)^{-1}[C_{t|t-1}H^T]^T \tag{3.34}$$

An initial estimator is required $\mathbf{a}_{0|0}$ and $C_{0|0}$. Now an approach is described using the EM algorithm following [30] and [31]. Suppose that the initial state $\mathbf{a}_{0|0}$ is taken from a Gaussian distribution $(\boldsymbol{\mu}_0, \Sigma_0)$, y ϵ_t y $\boldsymbol{\eta}_t$ are jointly Gaussian. The complete data $(\mathbf{a}_{0|0}, \boldsymbol{\alpha}(1), \cdots, \boldsymbol{\alpha}(T), \mathbf{x}_1, \cdots, \mathbf{x}_T)$ are considered, but of course, $\boldsymbol{\alpha}(t)$ are not observed. Suppose that the innovation parameter matrix $\mathbf{K}$ is equal to identity matrix, $K = I_p$.

The likelihood of complete data can be written

$$
\begin{aligned}
P(\alpha_0, \alpha_1, \cdots, \alpha_T, x_1, \cdots, x_T) = \prod_{t=1}^{T} & \frac{1}{|\Sigma_0|^{\frac{1}{2}}} exp\left\{ -\frac{1}{2}(a_{0|0} - \mu_0)^T \Sigma^{-1}(a_{0|0} - \mu_0) \right\} \\
& \frac{1}{\Sigma_\eta^{\frac{T}{2}}} exp\left\{ -\frac{1}{2}(\alpha_t - \rho\alpha_{T-1})^T \Sigma_\eta^{-1}(\alpha_t - \rho\alpha_{t-1}) \right\} \quad (3.35) \\
& \frac{1}{\Sigma_\epsilon^{\frac{T}{2}}} exp\left\{ -\frac{1}{2}(x_t - H\alpha_t)^T \Sigma_\epsilon^{-1}(x_t - H\alpha_t) \right\}
\end{aligned}
$$

Taking logarithm, we have left

$$
\begin{aligned}
\ln L = \ln(P(\alpha_0, \alpha_1, \cdots, \alpha_T, x_1, \cdots, x_T)) = & -\frac{1}{2}\ln|\Sigma_0| - \frac{1}{2}(a_{0|0} - \mu_0)^T \Sigma^{-1}(a_{0|0} - \mu_0) \\
& -\frac{T}{2}\ln|\Sigma_\eta| - \frac{1}{2}\sum_{t=1}^{T}\left[(\alpha_t - \rho\alpha_{t-1})^T \Sigma_\eta^{-1}(\alpha_t - \rho\alpha_{t-1})\right] \\
& -\frac{T}{2}\ln|\Sigma_\epsilon| - \sum_{t=1}^{T}\left[(x_t - H\alpha_t)^T \Sigma_\epsilon^{-1}(x_t - H\alpha_t)\right]
\end{aligned}
$$

$$(3.36)$$

Our objective is to maximize lnL with respect to $\mu_0, \Sigma_0, \rho, \Sigma_\eta$ and Σ_ϵ. We apply the EM algorithm conditionally with respect to the observed series $x_1, \cdots, x_T$ to maximize (3.36), where α_t is considered a lost data.

This result is proved in [17], it leads to a sequence of non-decreasing likelihoods. One difficulty is that there is a simple compound term in (3.36) in μ_0 and Σ_0, so that both cannot be estimated. Two alternatives are a) fix Σ_0 and estimate μ_0 b) consider μ_0 as fixed and set $\Sigma_0 = 0$. This problem is discussed in [32] and was reorganized by [31]. The estimate in the time domain then gives a estimates of the parameters $\rho, \Sigma_\eta, \Sigma_\epsilon$ and the states α_t.

For the missing values, both $\mathbf{a}_{t|t}$ and $C_{t|t}$ are computed using the t'th observation $\mathbf{x}_t$. When $\mathbf{x}_t$ have missing values we omit those rows in H and calculate $\mathbf{a}_{t|t}$ and $C_{t|t}$ relying on the non-missing data. An adjustment to the recursion formulas and the likelihood is required. This is in line with the usual procedure about missing values in KF application [18].

3.5   Implementation of the algorithm

For implementation of the full algorithm is necessary estimate a set of common fields. The objective in estimating the common fields is to capture as much as possible the spatial variation at each site and through the sites, as a linear combination of the common fields. These linear combinations or state vectors, are constrained by the ST-GSSS model, that is, the KKF model.

Given the choice of common fields, we choose the structure of a temporal model, possibly of a family of related temporal models, and proceed to estimate parameters for a selected model. This also leads us to estimation of Σ_ϵ, the observation error, summarized most appropriately for spatial prediction as a spatial covariogram (or variogram).

For data $x(s,t), s \in \{s_1, \cdots, s_n\}, t \in \{1, \cdots, T\}$, given values of the temporal parameters, $\boldsymbol{\theta}$, we can maximize lnL with respect to ϕ and given values of the spatial parameters, ϕ, we can maximize lnL with respect to $\boldsymbol{\theta}$. This suggests an iterative two stage estimation process.

Algorithm I

Step 1. Propose an $n \times P$ matrix $\mathbf{H}$, by estimating a covariogram for Σ_ζ from the data, and combining columns containing values of q trend fields at the s_i together with columns containing the first $r = p - q$ principal fields $\mathbf{u}_{q+j}$, $j = 1, \cdots, r$, obtained from the generalized bending energy matrix B, with normative sites $\mathbf{s}_i^* = \mathbf{s}_i, i = 1, \cdots, m = n$.

Step 2. Estimate, using the Kalman filter in the EM algorithm, the parameters $P, \Sigma_\epsilon, \Sigma_\eta$ and the initial condition $\boldsymbol{\alpha}(0) \sim N(\boldsymbol{\mu}_0, \Sigma_0)$.

Step 3. Using the estimated parameters from the EM algorithm maximize the likelihood with respect to the covariogram parameters.

Repeat Step 2 and Step 3 until convergence is achieved.

Algorithm II

Steps 1 and 2 are the same of algorithm I

Step 3. Compute the covariance matrix Σ_ζ by substituting the estimated parameters into the formula for the covariance at time t for t large.

$$cov(X(t,s), X(t,s')) = \mathbf{h}(s)P^t\Sigma_0 P^{tT}\mathbf{h}(s')^T + \sum_{i=0}^{t-1} \mathbf{h}(s)P^i\Sigma_\eta P^{iT}\mathbf{h}(s')^T + \\ cov(\epsilon(s), \epsilon(s')) \tag{3.37}$$

If the eigenvalues of P are less than unity, then the first term on the right of (3.37) is negligible when t is large. Note that the two covariances in the formula are respectively Σ_ζ and Σ_ϵ.

Step 4. Compute the H matrix using the new covariance structure.

Repeat Steps 2-4 until convergence is achieved. [2]

Chapter 4

Data Description

We have already mentioned in Chapter 1 that our main interest is to model and analyse the daily temperature, precipitation and humidity data in a study region in the north of Ecuador.

Figure 4.1: A plot of 3 meteorological monitoring sites in the study region.

There are a large number of meteorological monitoring sites in Ecuador. However, they have problems when the data is collected in the majority of those sites.For that reason, A lot of missing data arises throw the recoleccion. In addition, there are time intervals where we can find outliers. In this chapter we provide details regarding the

procedures that we have adopted for data cleaning and editing so that the data can be readily used for modelling purposes.

We obtain, after cleaning, 10958 data. One for day since January 1988 to December 2017 from 3 locations see Figure 4.1. In this chapter we also provide a summary statistics and graphical displays to describe this large data set. This summary statistics will be used in model based analysis in the next chapter.

Moreover, we have also obtained the forecasts of temperature, precipitation and humidity on the study region. We present summaries of these forecasts and compare with them of the observed data.

In our study we use the meteorological variables obtained from The National Institute for Meteorology and Hydrology (INAMHI). We provide the data processing and a summary statistics of these meteorological variables.

4.1   Data Preparation, Editing and Cleaning

The INAMHI collects daily meteorological data of all Ecuador covering the 24 provinces. We consider a part of the north Ecuador as our study region (see Figure 4.1), where we finally have data from 3 sites.

In this work a daily study of the observed data is carried out. Additionally, in some years there are a missing observations. Missing data were completed using the mean of the respective month, for example, if the temperature data of February 5, 1990 was missing, the mean of February is calculated and this value replaces the missing data.

4.2   Descriptive Statistics

In this section we discuss some summary statistics and graphical displays of the data set we prepared in the previous section. Recall that we have daily data from 3 monitoring sites(= n) for 365 (= T) days in a year (January 1 to December 31) for 30 (= r) years(1988 to 2017). Out of these 32874 (= nrT) possible daily observations per variable (3 variables), hence we have 98622 daily observations, 1374 (i.e., 1.39%) are missing, 348

possible monthly observations, and 30 possible yearly observations.

The Pichincha province contain a higher percentage (53%) of missing observations compared to the other provinces. This is possibly because of human causes or to failures in measuring equipment.

Province	Min	Mean	Median	Max	sd
Pichincha	9.90	15.12	15.10	20.00	1.29
Imbabura	6.00	10.46	10.50	13.80	0.95
Carchi	7.30	12.35	12.30	17.50	1.14

Table 4.1: Summary statistics for daily maximum temperature in °C

We can observe from Table 4.1 that temperature varies from 6.00 °C in Imbabura to 20.00 °C in Pichincha. Addicionaly, Pichincha have a 15.12 °C and 15.10 °C of mean and median respectively, it is higher than others provinces.

Province	Min	Mean	Median	Max	sd
Pichincha	0.00	3.02	0.00	75.00	6.64
Imbabura	0.00	3.43	0.00	140.50	7.44
Carchi	0.00	2.74	0.40	103.50	5.57

Table 4.2: Summary statistics for daily maximum precipitation in mm

From Table 4.2 we can observe that precipitation varies from 0.00 mm in all province to 140.50 mm in Imbabura. Moreover, Carchi have 2.74 mm of mean and 0.40 mm of median, if compares with others provinces, Carchi have lowest mean and higher median.

Province	Min	Mean	Median	Max	sd
Pichincha	35.00	70.33	72.00	97.00	10.31
Imbabura	48.00	87.61	89.00	99.00	6.16
Carchi	48.00	80.77	81.00	99.00	6.04

Table 4.3: Summary statistics for daily Relative Humidity in %

Table 4.3 shows the statistics for daily Relative Humedity, it varies from 35 % in Pichincha to 99 % in Imbabura and Carchi. Also, Imbabura have 87.61 % and 89% of mean and median respectively, it is higher than others provinces.

Tables 4.4, 4.5,and 4.6 shows the correlation between meteorological variables used in this thesis.

	Temperature	Precipitation	Humidity
Temperature	1.00	-0.34	-0.68
Precipitation	-0.34	1.00	0.41
Humidity	-0.68	0.41	1.00

Table 4.4: Correlation matrix of daily variables of Pichincha

	Temperature	Precipitation	Humidity
Temperature	1.00	-0.09	-0.36
Precipitation	-0.09	1.00	0.26
Humidity	-0.36	0.26	1.00

Table 4.5: Correlation matrix of daily variables of Imbabura

We observe in tables 4.4 and 4.5 that precipitation has a positive correlation with relative humidity, whereas precipitation and relative humidity show negative correlation with temperature.

	Temperature	Precipitation	Humidity
Temperature	1.00	0.05	-0.27
Precipitation	0.05	1.00	0.23
Humidity	-0.27	0.23	1.00

Table 4.6: Correlation matrix of daily variables of Carchi

Table 4.6 shows that temperature show negative correlation with relative humidity, while precipitation has a positive correlation with temperature and relative humidity.The box-plot of temperatures values by year are given in Figure 4.2. Here, we can observe that on average the temperature in Pichincha (figure 4.2 (a)) is lowest in 1989 and 1999, whereas it is highest in 2015. In figure 4.2 (b), we can observe that the temperature average in Imbabura is lowest in 1999 and highest in 2016. Additionally, in figure 4.2 (c), we can see that the temperature average in Carchi is lowest in 1989 and highest in 2015.

a) Pichincha
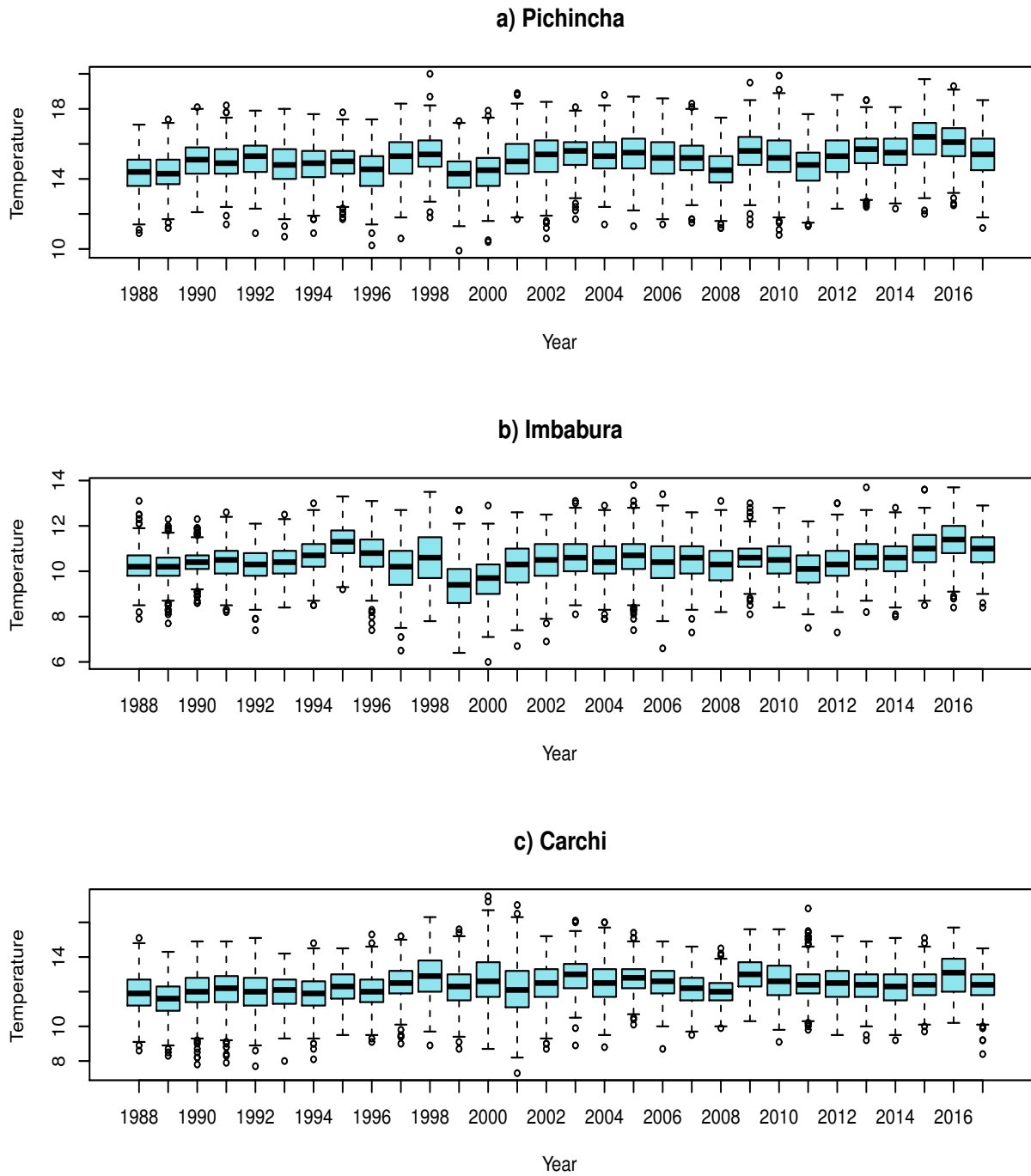
b) Imbabura

c) Carchi

Figure 4.2: Temperature box-plot: (a) Pichincha, (b) Imbabura, (c) Carchi.

Figure 4.3 shows the precipitation for different years. Here, we can observe that on average precipitation in study region is similar for all years. Additionally, we can observe a lot of outliers in these time period.
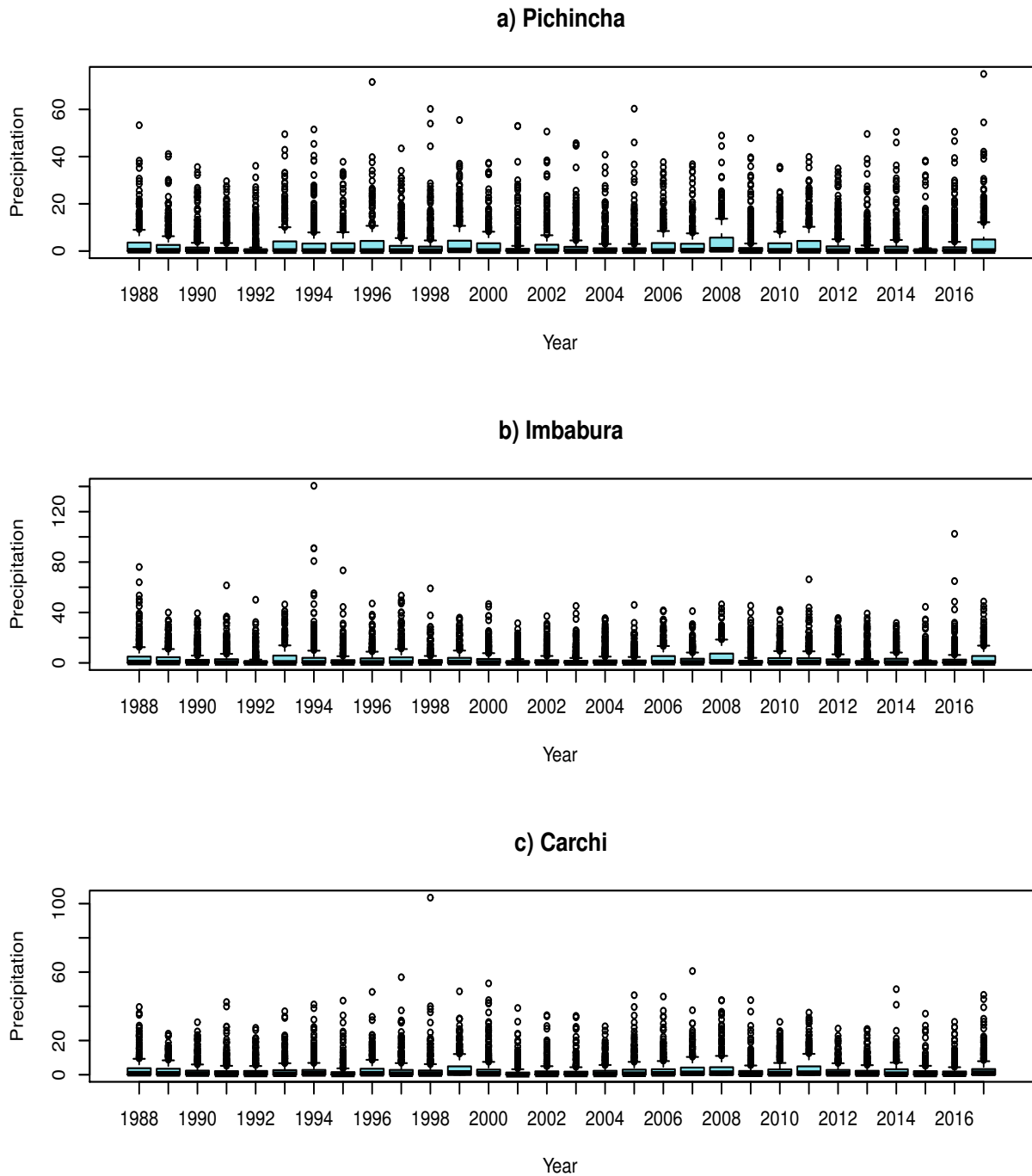
Figure 4.3: Precipitation box-plot: (a) Pichincha, (b) Imbabura, (c) Carchi.

The box-plot (Figure 4.4) shows the relative humidity for different years. Here, we can observe that on average the relative humidity in Pichincha (figure 4.4 (a)) is lowest in 2001 and 2002, whereas it is highest in 2011 and 2012. In figure 4.4 (b), we can observe

that the relative humidity average in Imbabura is lowest in 1999 and highest in 2016. Additionally, in figure 4.4 (c), we can see that the relative humidity average in Carchi is lowest in 2004, 2005 and 2007, and highest from 1997 to 2000.
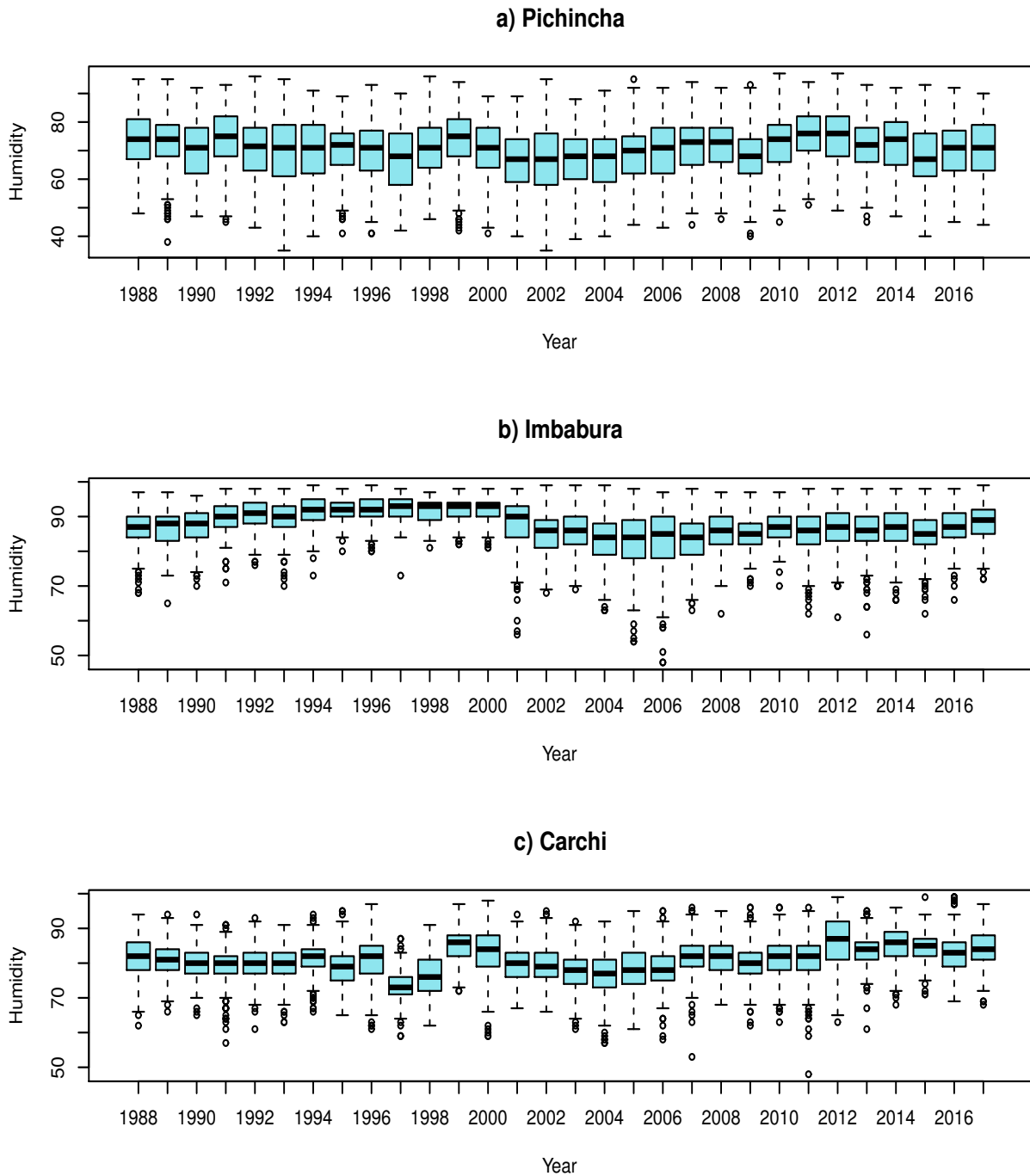
Figure 4.4: Relative humidity box-plot: (a) Pichincha, (b) Imbabura, (c) Carchi.
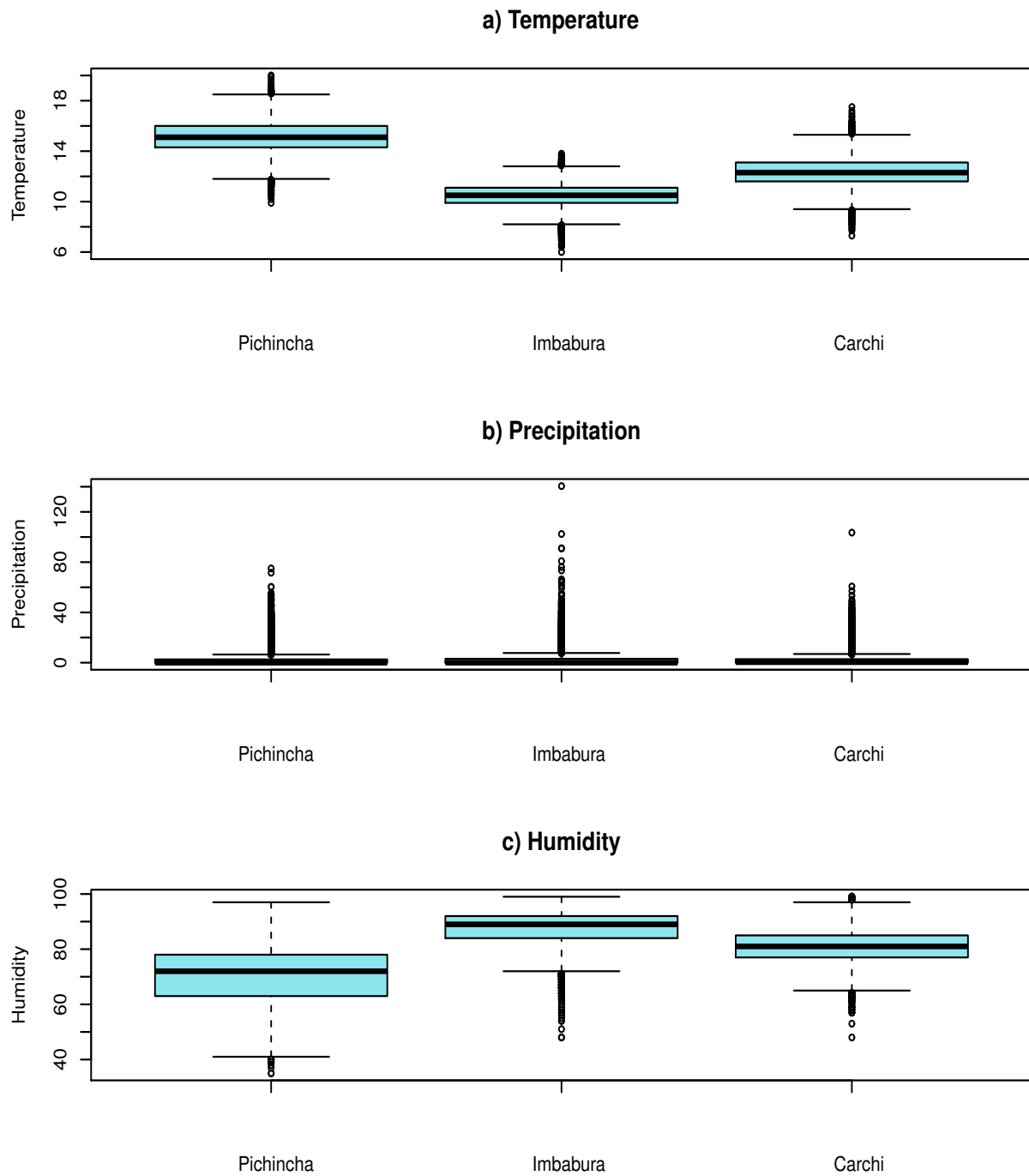
Figure 4.5: Box-plot of the three meteorological variables by provinces: (a) temperature (°C), (b) precipitation (mm), (c) humidity (%).

Figure 4.5 represents the box-plot of temperature (a), precipittion (b) and relative humidity (c) of the different provinces within our study region of Ecuador. Pichincha

have higher temperature levels and Imbabura lowest levels. The precipitation is similar in all provinces. Finally, relarive humidity is lowest in Pichincha and highest in Imbabura.

Chapter 5

Results

The data for this work were provided by the national institute of meteorology and hydrology (INAMHI). They comprise 10958 daily observations of temperature, precipitation and humidity at 3 monitored sites in three provinces of the Republic of Ecuador (Carchi, Imbabura and Pichincha). There are almost 30 years of readings for the period since 1st January 1988 to 31th December 2017. We obtained semi-variograms for the spatial part and interpolation. Also, filter data for temporal part of the three parameters under study was performed. The results obtained are presented below.

5.1   Semi-Variograms

Semivariogram is a function that relates semivariance to the vector $\boldsymbol{h}$ known as lag, which denotes the distance and direction of any pair of vectors.

The semivariogram of the data described above is shown in the Figure 5.1. For the empirical semivariogram, constant trend was taken and the data were fitted with an exponential model. The function adjustment allows to extract a series of parameters that are going to be used for spatial interpolation (kriging) and that define the degree and scale of spatial variation. These parameters are the range, nugget, and sill. The parameters obtained are as follows:

The range is the distance at which the semi-variance ceases to increase, therefore, it

Model:	Exponential
Range:	0.00285
Sill:	$1.98e^{-29}$
Nugget:	0
h:	$\begin{bmatrix} 1.25 & 1.75 & 2.75 \end{bmatrix}$

Table 5.1: Semi-variogram data

indicates the distance from which the samples are spatially independent of each other [33]. In this case, the range is 0.00285 km, the points of our semivariogram are after the range, therefore, all points are independent. In other words, as the distance of the sampling points increases, there is no longer a relationship between them. Variance begins to stabilize and sample values are not related to each other.

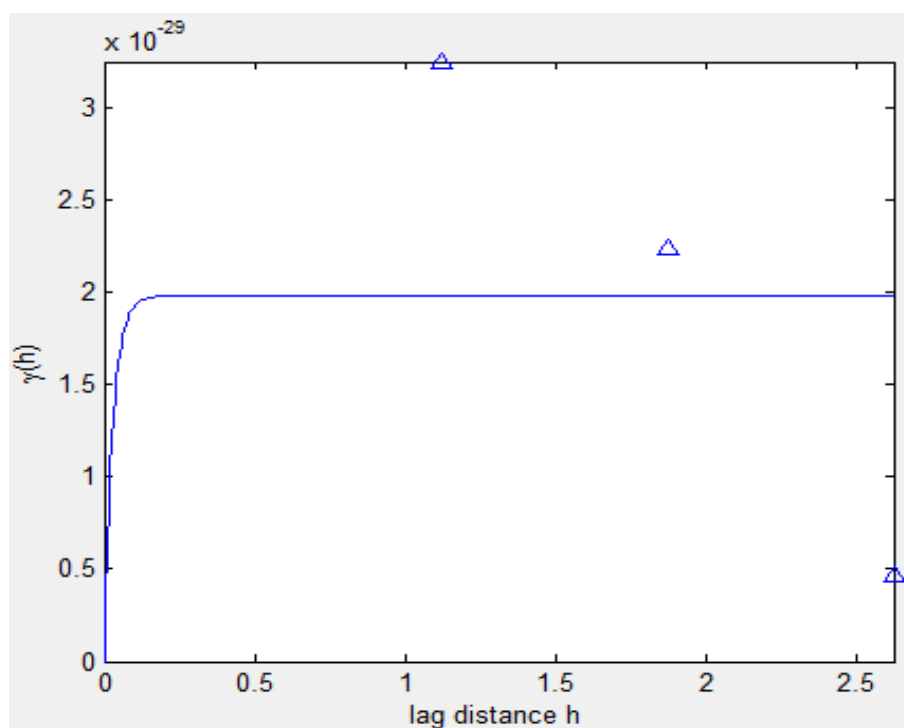Figure 5.1: Semi-variogram of the study region.

Nugget is the variance not explained by the model, and it is calculated as the interception with the Y axis. It is also known as the variance error since the variance of two points separated by 0 meters (the interception with the Y axis) should be zero . That is why this variance is usually indicating variability at a lower scale than the one sampled.

In addition, analytical or sampling errors also contribute to the appearance of error variance. Nugget for our case is zero (see Table 5.1 and Figure 5.1). Note that in figure 5.1 the values for the experimental variogram appear to vary around a constant, from an approximate distance of 0.0285 km (range). This constant value is identified as sill , which in this case is approximately $1.98e^{-29}$ semi-variance units.

5.2   Filter Data

The effect interpolation and filter data by KKF is closely related to the precision of spatial field $\mathbf{H}$, the fitted semi-variogram model reflects this precision. Since the semi-variogram model did not fit properly, and the use of an inaccurate space field $\mathbf{H}$ will lead to the filter divergence, a unitary matrix for $\mathbf{H}$ was used to perform the filtering and interpolation of the data.

All computations were carried out in MATLAB on a Windows 10 system with a Core I5 Intel processor with 2.27 GHz and 4 GB RAM memory. The computations related to the 'EM + Kalman filter' in wich 20 iterations were performed, costed 743 (12.48 min) seconds.

The original time series and interpolation effects of temperature for the 3 sites is displayed in Figure 5.2. We could see that the interpolation performed for Pichincha and Carchi provinces is better than Imbabura province.

In Figure 5.3 we can observe the original data time series of precipitation and their interpolation effects for the 3 sites in study region. In this case, the Carchi Interpolation is better than Pichincha and Imbabura interpolation. In addition, many outliers values can be observed in all series.

Relative humidity time series with its original data and its interpolated data can be seen in the Figure 5.4. In the year 2002, there is a sudden change in the time series for the Imbabura province. The data interpolation is better in Imbabura and Carchi provinces than Pichincha, this may be because in the Pichincha province there were more missing data than other provinces.
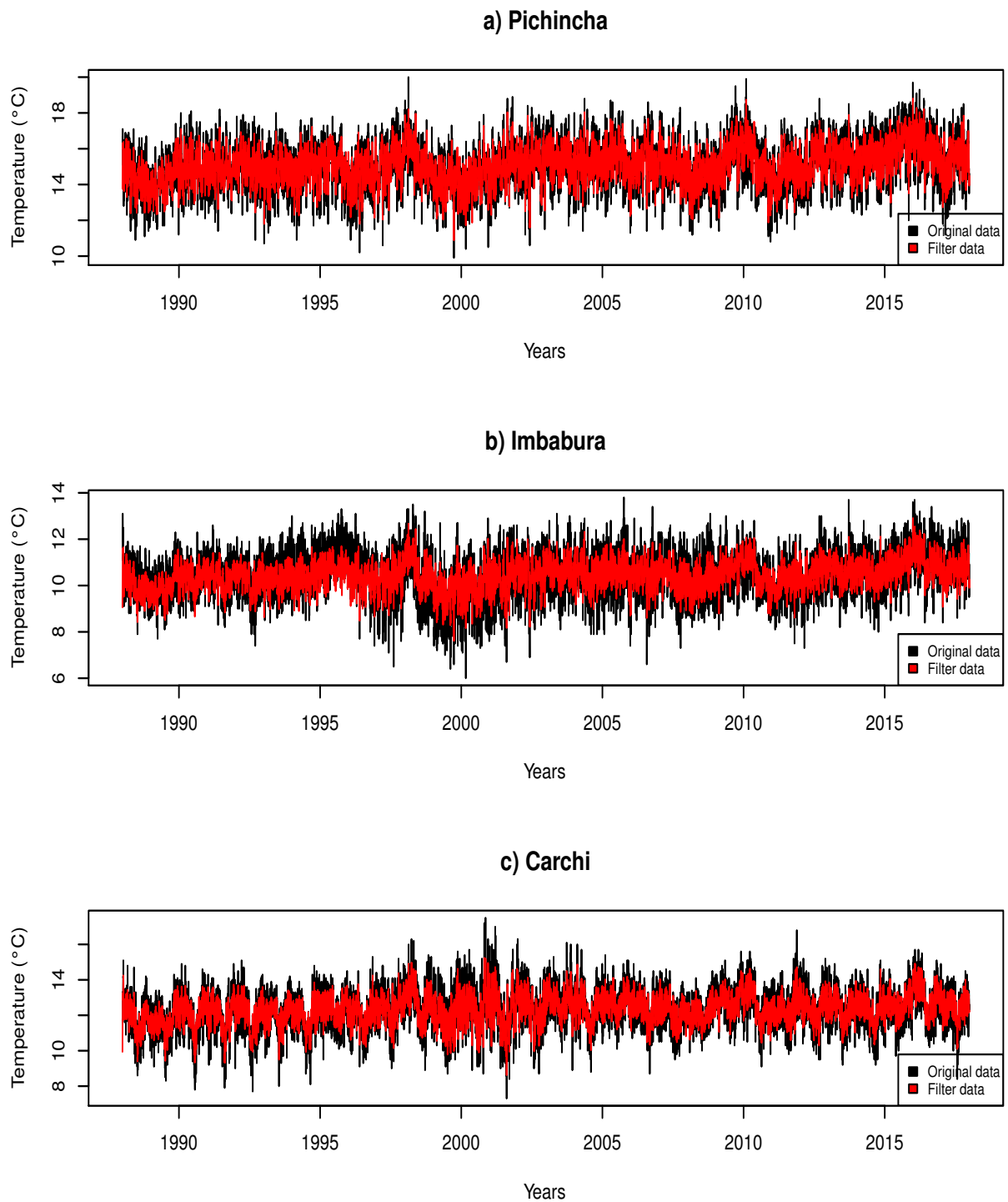
Figure 5.2: Temperature time series of study region: a)Pichincha, b)Imbabura, c)Carchi.

a) Pichincha

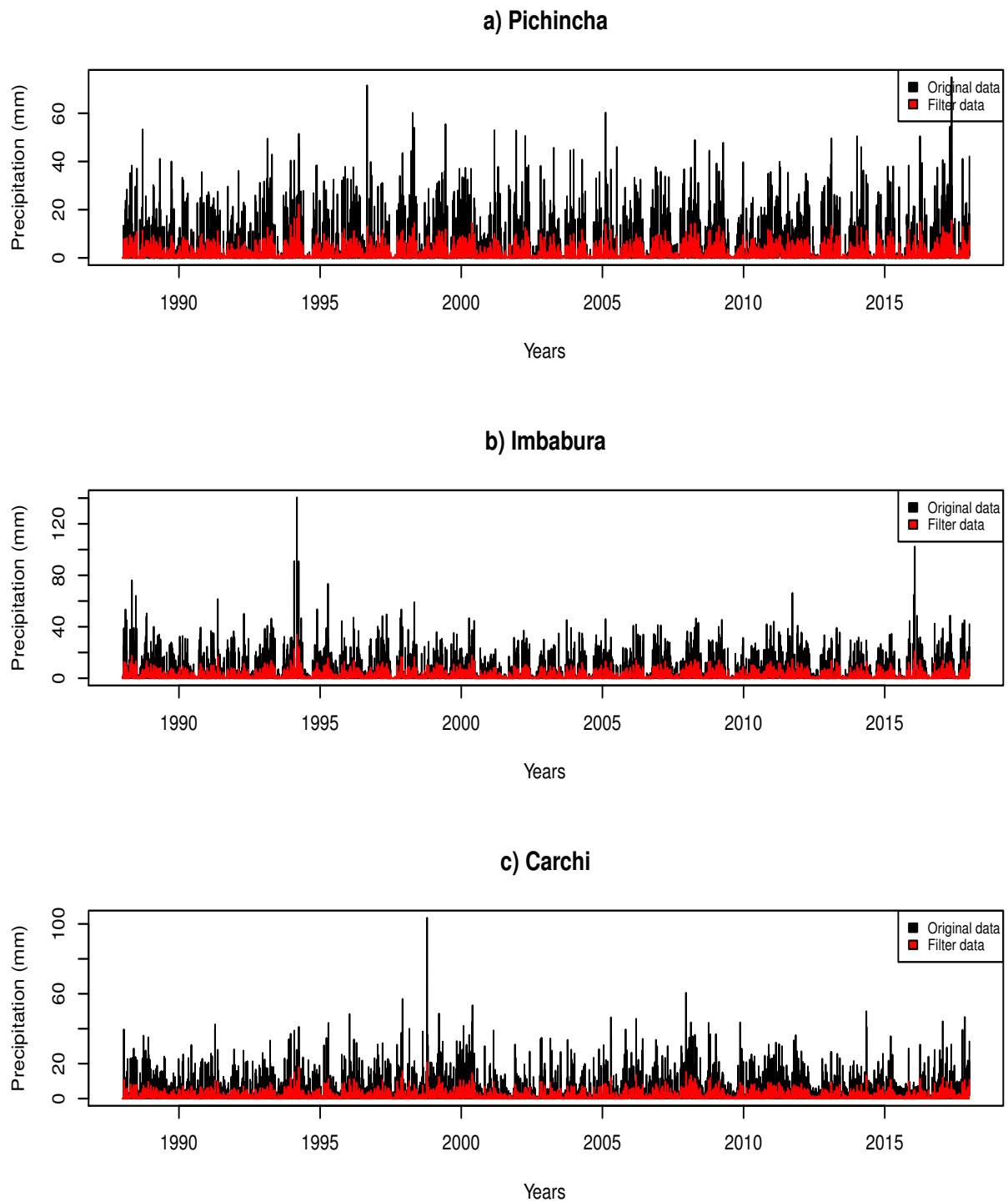**b) Imbabura**

c) Carchi

Figure 5.3: Precipitation time series of study region: a)Pichincha, b)Imbabura, c)Carchi.

Figure 5.4: Relative Humidity time series of study region: a)Pichincha, b)Imbabura, c)Carchi.
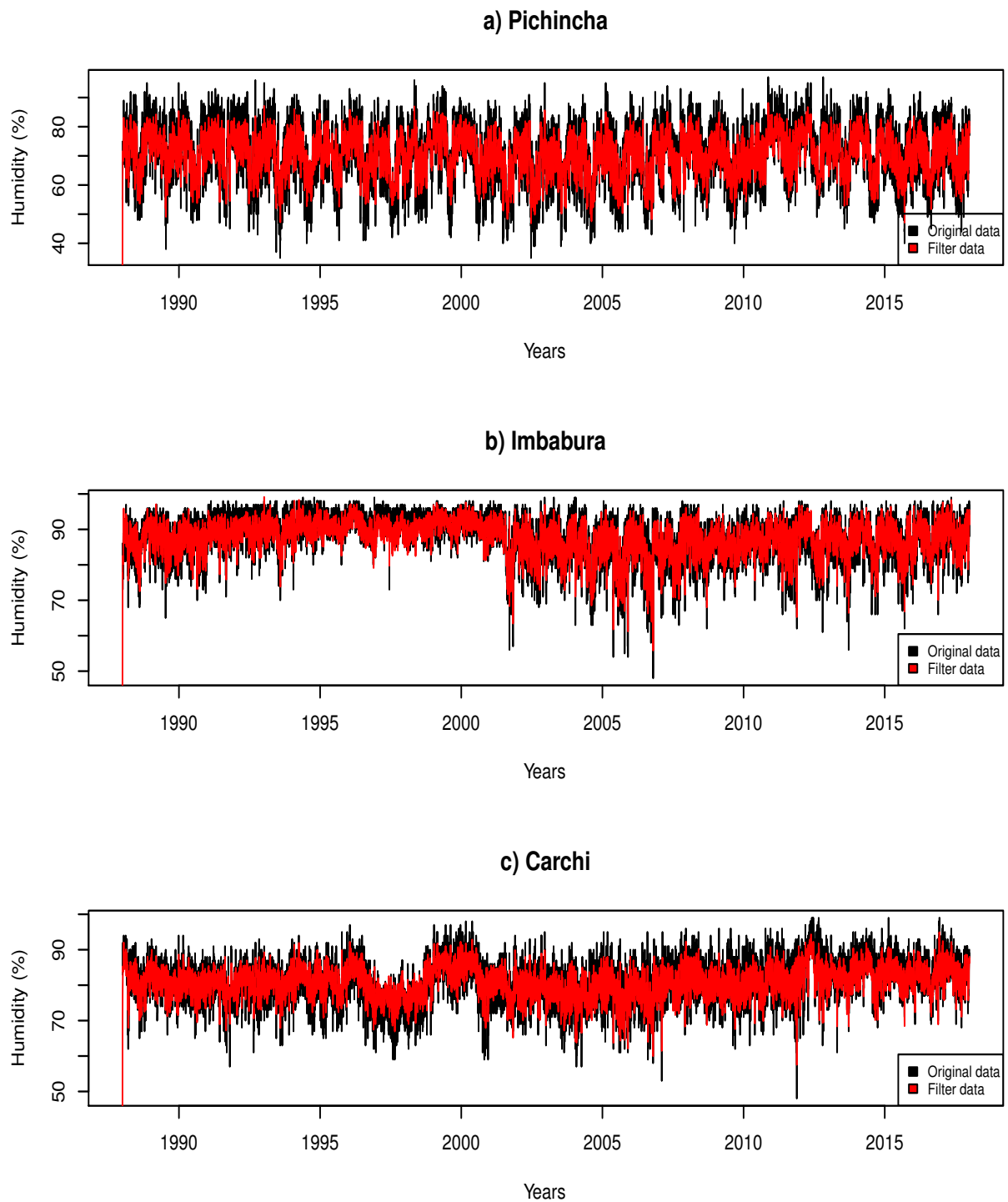
Province	Min	Mean	Median	Max	SD
Pichincha	10.88	15.09	15.11	18.81	0.98
Imbabura	7.61	10.44	10.45	12.91	0.61
Carchi	8.62	12.33	12.34	15.24	0.79

Table 5.2: Summary statistics for daily maximum temperature in °C

Province	Min	Mean	Median	Max	SD
Pichincha	0.00	1.49	0.59	22.20	2.07
Imbabura	0.00	1.79	0.67	33.83	2.38
Carchi	0.00	1.29	0.62	20.84	1.73

Table 5.3: Summary statistics for daily maximum precipitation in mm

Province	Min	Mean	Median	Max	SD
Pichincha	28.06	70.23	71.04	88.15	6.95
Imbabura	43.74	87.48	88.23	99.13	4.85
Carchi	43.47	80.63	80.80	94.93	4.31

Table 5.4: Summary statistics for daily Relative Humidity in %

In Tables 5.2, 5.3 and 5.3 is presented a statistical summary of forecast data temperature, precipitation and humidity

Province	Temperature	Precipitation	Humidity
Pichincha	0.54	5.20	4.32
Imbabura	0.58	5.75	2.52
Carchi	0.53	4.42	2.94

Table 5.5: RMSE of 3 meteorological variables on study region

In Table 5.5, we can observe that the root mean square error (RMSE) between the real data and the KKF data has a low error in the temperature variable compared to the other variables studied. This is because there are outliers data in precipitation and humidity time series that are not present in the temperature time series.

Additionally, humidity error in Pichincha is greater than Imbabura and Carchi. This may be due to missing data, in Pichincha there were more missing data in humidity variable than other provinces.

5.3   Forecasting

Temperature, precipitation and humidity forecasting was made for the next 3 days. We obtained the following results presented in the Table 5.6

		Day 1	Day 2	Day 3
Pichincha	Temperature	14.53	14.55	14.55
	Precipitation	6.64	3.39	1.73
	Humidity	78.50	77.45	76.71
Imbabura	Temperature	10.19	10.14	10.10
	Precipitation	7.56	3.88	1.98
	Humidity	93.97	93.98	93.90
Carchi	Temperature	12.14	12.06	11.99
	Precipitation	5.66	2.90	1.48
	Humidity	87.46	87.09	86.80

Table 5.6: The forecast of next 3 days.

The results obtained in temperature and humidity forecast are similar for the 3 days. On the other hand, the precipitation results decrease rapidly.

Chapter 6

Conclusions and Future work

In this thesis we used Kriging and Kalman filter (KKF) to analyse and obtain long-term trends in spatio-temporal data. With this in mind, we use spatio-temporal data obtained from 3 meteorological station in 3 Ecuador provinces (Pichincha, Imbabura and Carchi) for 30 years. As a part of this analysis, we have done editing and cleaning of the raw meteorological data that they have been obtained from the INAMHI (see Chapter 4). We model meteorological data using Kriging and Kalman filter. We had issues of choosing appropriate semi-variogram modelling strategies for analysing spatial components because of the semi-variogram model did not fit properly. Also, a unitary matrix for **H** was used to perform the filtering and interpolation of the data. (see Chapter 5). Some conclusions are as follows:

- There was a problem performing the semivariogram and therefore spatial fields could not be obtained, this may have been due to the few available sites for this study (only three). We must have more sites where the data is obtained.

- The semi-variogram model did not fit properly, a unitary matrix for **H** was used to perform interpolation and filtering to obtend results. It can be seen that we have very good results doing this way.

- An exploratory study was conducted to understand the data dynamics.

- A time space model was used to model time series of precipitation, humidity and

temperature of some meteorological stations in Ecuador for obtaining satisfactory results.

- The KKF algorithm was implemented to estimate the unknown states and parameters in the model.

- RMSE was used as a measure of goodness of fit to calibrate the KKF quality estimation. The greatest error was obtained for precipitation variable. One of the reasons could be the data interval is greater than other variables (Min = 0 mm, max = 140.50mm). Moreover, many outliers are present in the time series which can affect the model.

In a future work, we hope to implement the model on more meteorological stations around Ecuador to be able to estimate in a better way the semi-variogram in orden to obtain good results in the spatial componet. Therefore, it could be possible to interpolate the results to other points in space and make prediction maps. Additionally, we can give more realistic filter and interpolation results. Other filters such as an extended Kalman Filter or some variants of the particle filters also can be used to consider non-linear and non-Gaussian cases.

References

[1] C. Goodall and K. V. Mardia, "Challenges in multivariate spatio-temporal modeling," in *Proceedings of the XVIIth international biometric conference*, 1994, pp. 1–17.

[2] K. V. Mardia, C. Goodall, E. J. Redfern, and F. J. Alonso, "The kriged kalman filter," *Test*, vol. 7, no. 2, pp. 217–282, 1998.

[3] S. K. Sahu and K. V. Mardia, "A bayesian kriged kalman model for short-term forecasting of air pollution levels," *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, vol. 54, no. 1, pp. 223–244, 2005.

[4] A. Almendral-Vazquez and A. R. Syversveen, "The ensemble kalman filter-theory and applications in oil industry," *Norsk Regnesentral, Norwegian Computing Centre*, 2006.

[5] M. S. Handcock and M. L. Stein, "A bayesian analysis of kriging," *Technometrics*, vol. 35, no. 4, pp. 403–410, 1993.

[6] Y. Chen, D. S. Oliver, and D. Zhang, "Data assimilation for nonlinear problems by ensemble kalman filter with reparameterization," *Journal of Petroleum Science and Engineering*, vol. 66, no. 1-2, pp. 1–14, 2009.

[7] J. Cortés, "Distributed kriged kalman filter for spatial estimation," *IEEE Transactions on Automatic Control*, vol. 54, no. 12, pp. 2816–2827, 2009.

[8] B. Baingana, E. Dall'Anese, G. Mateos, and G. B. Giannakis, "Robust kriged kalman filtering," in *2015 49th Asilomar Conference on Signals, Systems and Computers.* IEEE, 2015, pp. 1525–1529.

[9] S. Pagani, A. Manzoni, and A. Quarteroni, "A reduced basis ensemble kalman filter for state/parameter identification in large-scale nonlinear dynamical systems," Technical Report, École polytechnique fédérale de Lausanne, Tech. Rep., 2016.

[10] A. Samalot, "Combined universal kriging and kalman filter techniques to improve wind speed prediction for northeastern us," 2017.

[11] L. M. Berliner, "Hierarchical bayesian time series models," in *Maximum entropy and Bayesian methods.* Springer, 1996, pp. 15–22.

[12] P. J. Harrison and C. F. Stevens, "Bayesian forecasting," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 38, no. 3, pp. 205–228, 1976.

[13] C. K. Wikle, A. Zammit-Mangion, and N. Cressie, *Spatio-temporal Statistics with R.* CRC Press, 2019.

[14] A. E. Gelfand, P. Diggle, P. Guttorp, and M. Fuentes, *Handbook of spatial statistics.* CRC press, 2010.

[15] N. Cressie, "C, 1993: Statistics for spatial data. rev. ed."

[16] H. Akaike, "Information theory and an extension of the maximum likelihood principle,[w:] proceedings of the 2nd international symposium on information, bn petrow, f," *Czaki, Akademiai Kiado, Budapest*, 1973.

[17] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the em algorithm," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 39, no. 1, pp. 1–22, 1977.

[18] R. H. Jones, "Maximum likelihood fitting of arma models to time series with missing observations," *Technometrics*, vol. 22, no. 3, pp. 389–395, 1980.

[19] N. K. K. C. Phan, "Kriged kalman filtering for predicting the wildfire temperature evolution," Ph.D. dissertation, 2014.

[20] N. Liu, W. Dai, R. Santerre, and C. Kuang, "A matlab-based kriged kalman filter software for interpolating missing data in gnss coordinate time series," *GPS Solutions*, vol. 22, no. 1, p. 25, 2018.

[21] L. Sánchez, S. Infante, V. Griffin, and D. Rey, "Spatio-temporal dynamic model and parallelized ensemble kalman filter for precipitation data," *Brazilian Journal of Probability and Statistics*, pp. 653–675, 2016.

[22] V. Roy, A. Simonetto, and G. Leus, "Spatio-temporal field estimation using kriged kalman filter (kkf) with sparsity-enforcing sensor placement," *Sensors*, vol. 18, no. 6, p. 1778, 2018.

[23] S. Banerjee, B. P. Carlin, and A. E. Gelfand, *Hierarchical modeling and analysis for spatial data.*   Chapman and Hall/CRC, 2014.

[24] K. S. Bakar, "Bayesian analysis of daily maximum ozone levels," Ph.D. dissertation, University of Southampton, 2012.

[25] R. Condit, *Tropical forest census plots: methods and results from Barro Colorado Island, Panama and a comparison with other plots.*   Springer Science & Business Media, 1998.

[26] M. Á. Reyes Cortés, "Estimación paramétrica y no paramétrica de la tendencia en datos con dependencia espacial. un estudio de simulación." 2010.

[27] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Journal of basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.

[28] Y. Kim and H. Bang, "Introduction to kalman filter and its applications," in *Kalman Filter.*   IntechOpen, 2018.

[29] F. L. Bookstein, "Principal warps: Thin-plate splines and the decomposition of deformations," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 11, no. 6, pp. 567–585, 1989.

[30] R. H. Shumway and D. S. Stoffer, "An approach to time series smoothing and forecasting using the em algorithm," *Journal of time series analysis*, vol. 3, no. 4, pp. 253–264, 1982.

[31] M. W. Watson and R. F. Engle, "Alternative algorithms for the estimation of dynamic factor, mimic and varying coefficient regression models," *Journal of Econometrics*, vol. 23, no. 3, pp. 385–400, 1983.

[32] A. C. Harvey, *Forecasting, structural time series models and the Kalman filter*. Cambridge university press, 1989.

[33] M. d. R. P. Arjones, "Heterogeneidad espacial de nutrientes del suelo en ecosistemas terrestres," Ph.D. dissertation, Universidade de Vigo, 2006.

Appendices

Appendix A

Algorithm Code

The algorithms in this section were obtained from [20]. Some changes were made to adjust them to this work.

A.1 Spatial field algorithm

```matlab
function H = SpatialFiled(x,y,S,proportion)

% Description:
%   Compute Spatial Filed of Kriged Kalman Filter
% Input:
%   x - array with longitude coordinates.
%   y - array with latitude coordinates.
%   S - Parameter of Semi-Variogram Value Fit
% Outout:
%   H - Spatial Filed
%------------------------------------------------------------%
% error checking
if size(y,1) ~= size(x,1)
    error('x and y must have the same number of rows!');
```

```matlab
15  end
16  R = ComputeCov(S,x,y);
17
18  % trend field
19  F = feval(S.trendfun,x,y);
20  q = size(F,2);
21  n = size(R,1);
22  InvR = eye(n)/R;
23  A = (F'*InvR*F)\F'*InvR;%Trend Desigen Matrix
24  B = InvR - InvR*F*A;%bending energy matrix
25  [U,E] = svd(B);%spectral decomposition of B
26  U=fliplr(U);E=fliplr(rot90(E));
27  e = diag(E);%eigen value
28
29  % dimensionality reduction
30  [e,ij] = sort(e,'ascend');
31  ij = n*repmat(ij'-1,n,1)+repmat((1:n)',1,n);
32  U = U(ij);
33  for i = 1:n
34      if((sum(e(1:i))/sum(e)) >= proportion)
35          p = i;
36          break;
37      end
38  end
39
40  % construct spatial filed
41  e = e(1:p,:);
42  E = repmat(e',n,1);
43  U = U(:,1:p);
```

```matlab
44 H2 = (R*U);%principle fields
45 H2(:,1:q) = [];
46 H = [F,H2];
47
48 function R = ComputeCov(S,x,y,xi,yi)
49 % Description:
50 %    Compute covariance matrix according to Semi-Variogram
51 %    function
52 % Input:
53 %    S - Parameter ofSemi-Variogram function
54 %    x,y - konwn point coordinates
55 %    xi,yi - unkonwn point coordinates
56 % Output:
57 %    R - covariance matrix
58 %--------------------------------------------------------------%
59 % check Input value
60 if nargin ==3
61     xi = x;  yi = y;
62 elseif nargin == 5
63 else
64     error('Number of Input Parameter are not correct!');
65 end
66 if length(x) ~= length(y)
67     error('x and y must have the same number of rows');
68 end
69 N1 = length(x);%number of konwn point
70 N2 = length(xi);%number of unkonwn point
71
72 %compute covariance matrix
```

```matlab
73  mzmax = N1*N2;
74  ij = nan(mzmax,2);
75  d = nan(mzmax,1);
76  ll = 0;
77  for k=1:N1
78      ll = ll(end)+(1:N2);
79      ij(ll,:) = [repmat(k,N2,1) (1:N2)'];
80      d(ll,:) = sqrt((repmat(x(k,1),N2,1) - xi).^2 +...
81          (repmat(y(k,1),N2,1) - yi).^2);
82  end
83  r = zeros(mzmax,1);
84
85  % range of Semi-Variogram function
86  if strcmp(S.model, 'spherical')
87      a = S.range;
88  elseif strcmp(S.model, 'exponential')
89      a = S.range;
90  elseif strcmp(S.model, 'gaussian')
91      a = S.range;
92  end
93  % a = S.range;
94  idx = d <= a;
95  b = [S.range,S.sill];
96  r(idx) = feval(S.func,b,a) - feval(S.func,b,d(idx));
97  R = sparse(ij(idx,1),ij(idx,2),r(idx),N1,N2);
```

A.2   Variogram algorithm

```matlab
function S = variog(x,y,trendfun,nrbins,maxdist)
% Description:
%   variog calculates the data experimental variogram
%   based variogram m-file
% Input:
%   x - array with [longitude latitude] coordinates,
%       a m*2 matrix. m is number of dimensions site.
%   y - GNSS series, a T*m matrix.T is number of dimensions
%       time.
%   trendfun - choosed trend function to remove data trend
%              ,mostly used function is constant, linear etc.
%   nrbins - number bins the distance should be grouped into
%            (default = 20)
%   maxdist - maximum distance for variogram calculation
%             (default = maximum distance in the dataset / 2)
% Output:
%   S - structure array with distance and gamma - vector
%------------------------------------------------------------%
% error checking
if size(y,2) ~= size(x,1)
    error('x and y must have the same number of rows!');
end

% change coordinate into normal distribution
x1 = x(:,1); y1 = x(:,2);
mxlength = length(x1);
mx1 = mean(x1); sx1 = std(x1);
my1 = mean(y1); sy1 = std(y1);
```

```matlab
29    x1 = (x1 - repmat(mx1,mxlength,1))./repmat(sx1,mxlength,1);
30    y1 = (y1 - repmat(my1,mxlength,1))./repmat(sy1,mxlength,1);
31    x = [x1,y1];
32
33    if isempty(nrbins)
34        nrbins = 20;
35    end
36    if isempty(maxdist)
37        minx = min(x,[],1);
38        maxx = max(x,[],1);
39        maxdist = sqrt(sum((maxx-minx).^2))/2;
40    end
41
42    % remove trend
43    T = size(y,1);
44    num_point = size(x,1);
45    res = nan(size(y));
46    for t=1:T
47        idex = any(isnan(y(t,:)'),2); %check NaN value
48        ytmp = y(t,:)';%restore time t value
49        xtmp = x;
50        ytmp(idex,:) = [];%delete NaN value
51        xtmp(idex,:) = [];
52        F = feval(trendfun,xtmp(:,1),xtmp(:,2));%Trend Design
53                                                %Matrix
54        beta = (F'*F)\F'*ytmp;
55        delta = ytmp - F*beta;
56        res(t,~idex) = delta;
57    end
```

```matlab
58
59  % calculate the exp-variogram value
60  num = num_point*(num_point+1)/2;
61  d = nan(num,1);%distance
62  val = nan(T,num);
63  ll = 0;
64  for k = 1:num_point
65      ll = ll(end)+(1:num_point+1-k);
66      dis = repmat(x(k,:),num_point+1-k,1) - x(k:num_point,:);
67      d(ll,:) = sqrt(dis(:,1).^2 + dis(:,2).^2);
68      val(:,ll) = (repmat(res(:,k),1,num_point+1-k)...
69          - res(:,k:num_point)).^2;
70  end
71
72  d(d>=maxdist,:) = [];
73  val(:,d>=maxdist)=[];
74  edges = linspace(0,maxdist,nrbins+1);
75  Sval = zeros(nrbins,1);
76  Sd = zeros(nrbins,1);
77  numcount = zeros(nrbins,1);
78  for i=1:nrbins
79      semi_varig_tmp = val(:,d>edges(i)& d<=edges(i+1));
80      idex = isnan(semi_varig_tmp);
81      numcount(i,1) = sum(sum(~idex));
82      semi_varig_tmp = semi_varig_tmp(:);
83      semi_varig_tmp(idex) = [];
84      std_semi_varig_tmp = std(semi_varig_tmp);
85      idex = abs(semi_varig_tmp)>3*std_semi_varig_tmp;
86      semi_varig_tmp(idex) = [];
```

```matlab
87 %       semi_varig_tmp = sort(semi_varig_tmp);
88 %       istart = ceil(0.3*length(semi_varig_tmp));
89 %       iend = ceil(0.8*length(semi_varig_tmp));
90 %       semi_varig_tmp(:,[1:istart iend:end]) = [];
91     Sval(i,1) = mean(semi_varig_tmp)/2;
92     Sd(i,1) = (edges(i)+edges(i+1))/2;
93 end
94 S.distance = Sd;
95 S.val = Sval;
96 S.num = numcount;
```

A.3   Expectation Maximization (EM) algorithm with Kriged Kalman filter

```matlab
function obs = EMEst_filter(Zt,Ht,F,alpha0,P0,R,Q,iters...
                            ,direction)
% Description:
%    Expectation Maximization(EM) Algorithm with Kriged Kalman
%    Filter to filter observed data and interpolate missing
%    data
% Input:
%    Zt - observed values, a n*m matrix, n is the length of
%         observed time, m is the length of observed site.
%    Ht - Spatial Filed of Kriged Kalman Filter or a unit
%         matrix
%    F - state transition matrix
%    alpha0 - initial state vector of Kriged Kalman Filter
%    P0 - covariance matrix of initial state vector
%    R - observation noise covariance matrix
%    Q - system state noise covariance matrix
%    iters - EM Algorithm iteration number
% Output:
%    obs - filtered and interpolated missing data
%------------------------------------------------------------%
[n,m] = size(Zt);%size of observation matrix
p = size(Ht,2);%dimension of spatial filed
if direction == 1
    str_direciton = 'EM Temperature:  ';
elseif direction == 2
    str_direciton = 'EM Precipitation:  ';
```

```matlab
27  elseif direction == 3
28      str_direciton = 'EM Humidity:  ';
29  else
30      error('Error in the data!');
31  end
32  obs = zeros(n,m);
33  for j=1:iters
34      alpha_t = alpha0;
35      Pt = P0;
36      A = zeros(p,p);
37      B = zeros(p,p);
38      C = zeros(p,p);
39      Rn = zeros(m,m);
40
41      str_process = sprintf('the %dth Kalman Filter,totally...
42          %d/%d',j,j,iters);
43      str = [str_direciton,str_process];
44      h = waitbar(0,str);
45      for i=1:n
46          if ~ishandle(h)
47              obs = [];
48              return;
49          end
50          waitbar(i/n,h);
51          %change missing value and design matrix into 0
52          II = any(isnan(Zt(i,:)'),2);
53          Z = Zt(i,:)';  Z(II,:)=0;
54          H = Ht;    H(II,:)=0;
55
```

```matlab
56          Rc = R;
57          II_F = ~II;
58          II = double(II);
59          II_F = double(II_F);
60          idex = II'*II + II_F'*II_F;
61          idex = ~(logical(idex));
62          Rc(idex) = 0;
63
64          %Kalman Filter
65          alpha_ = F*alpha_t;%one-step forecast state value
66          P_ = F*Pt*F' + Q;%covariance matrix of one-step
67                           %forecast value
68          K = P_*H'/(H*P_*H' + Rc);%gain matrix
69          alpha_t1 = alpha_ + K*(Z - H*alpha_);%filtered state
70                                               %value
71          Pt1 = P_ - K*H*P_;%covariance matrix of filtered
72                            % statevalue
73          Ptt1 = (eye(p) - K*H)*F*Pt;%covariance matrix between
74                                     %forecast filtered state value
75          obs(i,:) = (Ht*alpha_t1)';
76
77          A = A + Pt + alpha_t*alpha_t';
78          B = B + Ptt1 + alpha_t1*alpha_t';
79          C = C + Pt1 + alpha_t1*alpha_t1';
80
81          Rc = zeros(m,m);
82          II = double(II); II = II*II'; II = logical(II);
83          Rc(II) = R(II);
84          Rn = Rn +...
```

```matlab
85              ((Z-H*alpha_t)*(Z-H*alpha_t)'+...
86              H*Pt*H' + Rc)/n;
87
88          alpha_t = alpha_t1;
89          Pt = Pt1;
90      end
91      close(h)
92
93      %update new Kalman parameter
94      F = B/A;
95      Q = (C - F*B')/n;
96      R = Rn;
97 end
```