

INTERPOLATION OF IONOSPHERIC MODALITIES USING KRIGING, CO-KRIGING AND SPATIO-TEMPORAL KRIGING

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE
OF BILKENT UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR
THE DEGREE OF
MASTER OF SCIENCE
IN
ELECTRICAL AND ELECTRONICS ENGINEERING

By
Sadia Khaf
June 2018

INTERPOLATION OF IONOSPHERIC MODALITIES USING
KRIGING, CO-KRIGING AND SPATIO-TEMPORAL KRIGING

By Sadia Khaf

June 2018

We certify that we have read this thesis and that in our opinion it is fully adequate,
in scope and in quality, as a thesis for the degree of Master of Science.

Orhan Arıkan(Advisor)

Feza Arıkan

Sinan Gezici

Approved for the Graduate School of Engineering and Science:

Ezhan Karaşan
Director of the Graduate School

ABSTRACT

INTERPOLATION OF IONOSPHERIC MODALITIES USING KRIGING, CO-KRIGING AND SPATIO-TEMPORAL KRIGING

Sadia Khaf

M.S. in Electrical and Electronics Engineering

Advisor: Orhan Arıkan

June 2018

Long distance communication and navigation systems operating in the HF band use interacting signals as they travel through the ionosphere. It is important to accurately model ionospheric behavior to increase the performance of these systems. Delays occurring in the signals depend on the refractivity which is a function of frequency of the signals, and the electron density on the signal path at the time of propagation. Depending on the change in the solar activities, the electron distribution in the ionosphere changes spatially and temporally. The change in ionosphere can be tracked by various parameters and the space-time distribution of these parameters. Total Electron Content (TEC), the total number of electrons in a cylinder with one meter square cross-sectional area over a ray path is used as an important descriptor for the ionosphere. It is possible to generate TEC maps with high spatial resolution using the information obtained by processing the GPS satellite signals by constantly operating reference stations (CORS) GPS receivers. In particular, there are two other parameters that are used in HF communication and direction finding applications: foF2, which is the highest plasma frequency of foF2 layer, and hmF2, which is the height of maximum ionization. Sensitive foF2 and hmF2 measurements can be made by ionosonde systems. However, these systems are highly sparser than TEC measurements. For this reason, the resolution of the foF2 and hmF2 maps is less than the TEC maps. In this study, we propose a space-time mapping technique based on Co-Kriging which is used in conjunction with TEC data, that is correlated to these parameters, to increase the resolutions of foF2 and hmF2 maps. The performance of the proposed technique is compared with the alternatives and the increase in performance achieved is described statistically.

Keywords: Ionosphere, TEC, Kriging, Co-Kriging, Spatio-temporal Kriging.

ÖZET

İYONKÜRE DEĞİŞKENLENLERİN KRİGLEME, EŞ-KRİGLEME VE UZAY-ZAMAN KRİGLEME TEKNİKLERİ KULLANILARAK ARADEĞERLEMESİ

Sadia Khaf

Elektrik-Elektronik Mühendisliği, Yüksek Lisans

Tez Danışmanı: Orhan Arıkan

Haziran 2018

HF bandında çalışan uzak mesafeli iletişim sistemleri ve navigasyon sistemleri iyonküre içinden geçerken etkileşen sinyaller kullanmaktadırlar. Sistemlerin başarımını artırmak için iyonkürenin hassas olarak modellenmesi önem taşımaktadır. Sinyaller üzerinde oluşan gecikmeler sinyallerin frekansına ve yayılım anında sinyal yolu üzerindeki elektron yoğunluğuna bağlıdır. Özellikle Güneş aktivitelerindeki değişime bağlı olarak İyonküredeki elektron dağılımı yerel ve zamansal değişim göstermektedir. İyonküredeki değişim çeşitli parametreler ve bu parametrelerin uzay-zamansal dağılımı ile takip edilmektedir. Toplam Elektron İçeriği (TEİ: yer yüzeyinden uzaya doğru dik olarak yükselen bir metre kare kesit alanına sahip bir silindir içinde kalan toplam elektron sayısı) haritaları iyonküre için önemli bir betimleyici olarak kullanılmaktadır. GPS uydu sinyallerinin yeri belli hassas GPS alıcıları tarafından işlenmesi ile elde edilen bilgiler kullanılarak uzamsal sıklığı yüksek TEİ haritaları oluşturulması mümkündür. Özellikle HF iletişim ve yön bulma uygulamalarında kullanılan iki ayrı parametre daha vardır: yeryüzeyine dik olarak yapılan HF yayımının geriye yansıtılabildiği en yüksek frekans olan f_oF_2 ve bu frekanstaki yansımanın yerden yüksekliği olan h_mF_2 . Hassas f_oF_2 ve h_mF_2 ölçümleri iyonosonda sistemleri tarafından yapılabilmektedir. Ancak bu sistemler TEİ ölçüm istasyonlarına göre daha seyrek olarak bulunmaktadır. Bu nedenle TEİ haritalarına göre, f_oF_2 ve h_mF_2 haritalarının çözünürlükleri daha az olmaktadır. Bu tezde sunulan çalışmada f_oF_2 ve h_mF_2 haritalarının çözünürlüklerini artırmak amacıyla bu parametrelerle ilintisi olan TEİ verilerinin birlikte kullanıldığı eş-Kriglemeye dayalı bir uzay-zaman haritalama tekniği önerilmiştir. Önerilen tekniğin başarımı alternatifleriyle kıyaslanmış ve sağlanan başarımların istatistiksel olarak betimlenmiştir.

Anahtar sözcükler: iyonküre, TEİ, Krigleme, Eş-Krigleme, Uzay-Zaman Krigleme.



Acknowledgement

I dedicate my work to Professor Orhan Arikan, who made this work possible by his constant guidance and support,

To mom, dad, who supported my goals despite our conservative background and made me the first girl in my family and neighborhood to become an engineer, to study abroad,

And to all my teachers who made me who I am today.

I also acknowledge that this research is funded by TÜBİTAK projects 114E541 and 115E915.

Contents

1	Introduction	1
1.1	The Structure of the Ionosphere	1
1.1.1	The Empirical Model Describing Ionosphere	5
1.2	Review of Existing Kriging and Co-Kriging Methods	6
1.3	Review of the Methods used in Ionospheric studies	8
1.4	Sources of Ionospheric Data	10
1.5	Geomagnetic Storms in Ionosphere	11
1.6	Correlation between Ionospheric Modalities foF2, hmF2 and TEC	13
1.7	The Grid Schemes for Ionospheric Mapping Techniques	14
1.7.1	No Grid: Anisotropic Co-Kriging	14
1.7.2	Coarse Grid: Inverse Square Distance	14
1.7.3	Coarse Grid: Ordinary Kriging	15
1.8	Ionospheric Mapping/ Interpolation Techniques	15

1.8.1	Kriging	15
1.8.2	Co-Kriging	17
1.8.3	Spatio-temporal Kriging	18
1.9	The Map Validation Process	19
1.10	The Flowchart of Co-Kriging	20
2	Mathematical Formulation of the Interpolation Problem	25
2.1	Estimation Using Inverse Square Distance (ISD)	27
2.2	Estimation Using Ordinary Kriging	27
2.2.1	Theoretical Variogram Models	29
2.3	Co-Kriging	31
2.4	Spatio-temporal Kriging	36
2.4.1	Spatio-temporal Variogram	38
3	Simulations of the Proposed Techniques	42
3.1	Region of Interest for Mapping Techniques	43
3.2	Visualizing Correlations between Ionospheric Modalities	46
3.3	Kriging	46
3.4	Co-Kriging	47
3.5	Spatio-temporal Kriging	48

3.5.1	Estimation (Maps Inside the Temporal Extent)	49
3.5.2	Prediction (Maps Outside the Temporal extent)	49
3.5.3	Temporal Estimates - Inside Spatial Extent	50
3.5.4	Temporal Estimates - Outside Spatial Extent	50
4	Discussion	74
4.1	Kriging vs Co-Kriging	74
4.2	Kriging vs Spatio-temporal Kriging	75
5	Conclusion	87
A	Earthquakes Tables	97
B	Correlation Tables	99
C	GPS Stations Over Europe	101

List of Figures

1.2	hmF2 map (km), obtained from IONOLAB website, generated by utilizing IRI-Plas model for discrete locations in the world for 11 May 2013, 0 UT	22
1.3	Correlation of TEC with foF2 and hmF2 over the month of October 2016, at 12 UT, showing a quiet month with two geomagnetically disturbed days (October 12 and October 23).	22
1.4	Flowchart describing work sequence with respect to all four phases involved.	24
2.1	Grid Structure defined between the initial coordinates (θ_i and ϕ_i) and final coordinates (θ_f and ϕ_f), with latitude spacing Δ_θ , and longitude spacing Δ_ϕ	41
3.1	GPS stations over Europe (filled circles), R_E is the region bounded by dash dot box , R_I is the region bounded by the solid box, and R_O is the region between the dashed box and the solid box.	52
3.2	Ionosonde stations over Europe indicated by right triangles, station numbers correspond to stations given in Table 3.3.	53

3.3	GPS stations over Europe (indicated by left triangles) where D_{iz} for foF2 is available, station numbers correspond to station numbers given in Table 3.4.	54
3.4	Selected locations for primary and secondary modalities in D_{iri} , triangles indicate the primary data $D_{iri-foF2}$, squares indicate the secondary data $D_{iri-TEC}$, and the filled circles indicate the locations selected for testing	56
3.5	$D_{iz-foF2}$, D_{iz-TEC} , and $D_{iz-hmF2}$ on October 01, 2016 from station 1 to station 17 at 12 UT, station numbers correspond to stations in Table 3.4.	56
3.6	Correlation of TEC with foF2 and hmF2 over the month of October 2016, at 12 UT, showing a quiet month with two geomagnetically disturbed days (October 12 and October 23).	57
3.7	foF2 map generated by IONOLAB-MAP, using Universal Kriging over D_{iz} in region R_E^I , for October 04, 2016 at 12 UT.	58
3.8	foF2 map generated by IONOLAB-MAP, using Universal Kriging over D_{iri} in region R_E^I , for April 22, 2009 at 12 UT.	59
3.9	Error map for IONOLAB-MAP, over D_{iri} in region R_E^I , for April 22, 2009 at 12 UT.	60
3.10	foF2 map generated by Co-Kriging, over D_{iz} in region R_E^I , for March 12, 2010 at 12 UT.	61
3.11	foF2 map generated by Co-Kriging, over D_{iri} in region R_E^I , for April 22, 2009 at 12 UT.	62
3.12	Error map for Co-Kriging, over D_{iri} in region R_E^I , for April 22, 2009 at 12 UT.	63

3.13 foF2 map generated by Co-Kriging, over D_{snd} in region R_E^I , for March 12, 2010 at 12 UT.	64
3.14 Experimental Spatio-temporal Variogram (January 23, 2010 0 UT to January 24 15 UT) $D_{ion-TEC}$ over region R_I^I	64
3.15 Weights associated with station 86 for prediction at station 74, against temporal distance from January 19, 2010 0 UT to January 20, 2010 15 UT.	65
3.16 TEC map generated by Spatio-temporal Kriging STK, over $D_{ion-TEC}$ in region R_I^I and R_O^I , for January 19, 2010 at 12 UT. . .	66
3.17 TEC map generated by Spatio-temporal Kriging STK-SB, over $D_{ion-TEC}$ in region R_I^I and R_O^I , for January 19, 2010 at 12 UT. . .	67
3.18 TEC map generated by Spatio-temporal Kriging STK, over $D_{ion-TEC}$ in region R_I^O and R_O^O using data only in R_I^I , for January 19, 2010 at 17 UT.	68
3.19 TEC map generated by Spatio-temporal Kriging STK-SB, over $D_{ion-TEC}$ in region R_I^O and R_O^O using data only in R_I^I , for January 19, 2010 at 17 UT.	69
3.20 TEC temporal interpolation by Spatio-temporal Kriging, over $D_{ion-TEC}$ on station 86 in R_I , for January 19, 2010.	70
3.21 TEC temporal interpolation by Spatio-temporal Kriging, over $D_{ion-TEC}$ on station 86 in R_I , for January 20, 2010.	71
3.22 TEC temporal interpolation by Spatio-temporal Kriging, over $D_{ion-TEC}$ on station 98 in R_O , for January 19, 2010.	72
3.23 TEC temporal interpolation by Spatio-temporal Kriging, over $D_{ion-TEC}$ on station 98 in R_O , for January 20, 2010.	73

4.1	GPS Stations (indicated by filled circles) used as test stations in regions R_I and R_O	79
4.2	RMSE for Spatio-temporal Kriging STK from January 01, 2010 to March 31, 2010, indicated by colored circles, at stations indicated by station numbers that correspond to Table C.1 a)RMSE over $D_{ion-TEC}$ for R_I^I and R_O^I b) RMSE over $D_{ion-TEC}$ for R_I^O and R_O^O	81
4.3	a) RMSE for Spatio-temporal Kriging STK-SB from January 01, 2010 to March 31, 2010, indicated by colored circles, at stations indicated by station numbers that correspond to Table C.1 a)RMSE over $D_{ion-TEC}$ for R_I^I and R_O^I b) RMSE over $D_{ion-TEC}$ for R_I^O and R_O^O	82
4.4	RMSE for IONOLAB-MAP Kriging (Universal Kriging with Linear Trend in this region) from January 01, 2010 to March 31, 2010, over $D_{ion-TEC}$ for R_I^I and R_O^I	83
4.5	a) RMSE against spatial distance in (km) from the center of the region R_I , for Spatiotemporal Kriging STK, Spatiotemporal Kriging STK-SB and IONOLAB-MAP Universal Kriging in the region R^I b) RMSE against spatial distance in (km) from the center of the region R_I , for Spatiotemporal Kriging STK, Spatiotemporal Kriging STK-SB and IONOLAB-MAP Universal Kriging in the region R^O The * indicates that these plots are the predictions in region R^O	84
4.6	RMSE against spatial and temporal distances h and h_t from the center of the region R_I^I , (o-) represents the RMSE for the region R^I and (*-) represents the RMSE for the region R^O	85
4.7	Error Density indicated by solid line for STK, dashed line for STK-SB, and dash dot line for IONOLAB-MAP Kriging over $D_{ion-TEC}$ from January 01, 2010 to March 31, 2010 a) in the region R^I , b) in the region R^O	86

Chapter 1

Introduction

In this Chapter, we introduce the structure of ionosphere and some basic Physics behind ionospheric modeling. Later, we discuss an empirical model that is widely used to describe the behavior of ionosphere. We describe the problem of mapping the ionosphere, its importance and the most commonly used methods for this purpose. We discuss the history of some of the commonly used mapping and interpolation techniques for other applications first and then describe the development of such methods specifically for ionospheric studies. The state of current works in ionospheric interpolation is presented and the contribution of this thesis towards extending those methods is described. At the end of this chapter, we presented the over all work flow of the methods that will be used for ionospheric interpolation in this study.

1.1 The Structure of the Ionosphere

The layer of earth's atmosphere that lies above mesosphere and extends from 60 km to 1100 km is called ionosphere. Due to high energy solar and cosmic rays the atoms in this layer are stripped off of one or more electrons and called "ionized". Hence, this layer contains a high concentration of ions and free electrons. At

any particular time, only half of earth's atmosphere faces the sun and although the cosmic rays ionize the atoms at night too but not as much as the sun does. This is the reason ionosphere is much less charged at night. Ionosphere reflects the frequencies roughly 30 MHz and below that but for higher frequencies it adds delays in their path. Ionosphere and the phenomena occurring in it are important for many reasons: top among them being the radio communication and navigation applications. This layer influences radio wave propagation to distant places on earth and between satellites and earth hence accurately modeling the behavior of ionosphere becomes of crucial significance. All the military missile guidance and drone applications rely on GPS and inaccuracies in modeling ionospheric behavior can lead to disastrous effects. Some of these errors and inaccuracies are discussed in [1].

Ionosphere has varying concentrations of electrons at different heights and it behaves differently in various regions with respect to height. Such phenomenon divides it into three main layers D, E and F. There are no strictly defined altitudes for these layers because ionosphere keeps changing behavior with respect to place and time but we can consider the D layer to be the lowest layer extending from 60 km to 85 km and present only at day time. E layer extends from 85 km to 140 km and is always present both during day and night. The F layer is the highest or outer most layer and is considered to be present from 40 km to 500 km [2]. Most of the communication is usually performed over E and F layers, the reason being that the D layer absorbs most of the signal and introduces high attenuation. During the day time F layer exhibits a further division into two distinct layers and we call them F1 and F2 layer [3]. At night time ionosphere becomes less charged and D layer disappears [4]. In addition to this day and night variability, there are trends in electron density behavior that have cycles with periods of a month, year, season and sunspot cycle. Generally ionosphere follows a 24 hour cycle but the effects of ionization also follow sun's rotation period which is around 27 days. On top of that, they also follow solar activity period which is 11 years. These patterns can be disturbed during high solar activity times and can behave different than expected. On earth, different behavior is observed in different regions, these regions being the polar, midlatitude and equatorial regions. The

midlatitude region is considered the most stable one.

The F layer in the ionosphere is of highest significance and in that, F2 is the more important one. This layer has the highest electron density and it bends the HF waves and sends them back to earth. Various factors affect the amount of bending a wave undergoes while traveling through a medium, one of the most important factors being the relative dielectric constant, ϵ_r of plasma:

$$\epsilon_r = 1 - \frac{N_e e^2}{\omega^2 m \epsilon_o} = 1 - \frac{\omega_p^2}{\omega^2} , \quad (1.1)$$

where N_e is the number of electrons per unit volume (typically 1 m^3 but some works use 1 cm^3), e and m are charge and mass of electron respectively and ϵ_o is the permittivity of free space. Here ω_p is the plasma frequency:

$$\omega_p = \sqrt{N_e e^2 / m \epsilon_o} . \quad (1.2)$$

$$\omega_p = 2\pi f_p \cong 9\sqrt{N_e} . \quad (1.3)$$

Since the medium dielectric constant, ϵ_r , is frequency dependent, the medium refractive index $n = \sqrt{\epsilon_r}$ will also be frequency dependent. Electrons face high amount of collisions with ions and molecules at lower heights and that acts as a damping force. The law of conservation of momentum suggests that damping force $-\nu m v$ to be added to the force equation. Taking this effect into account, the dielectric constant becomes;

$$\epsilon_r = 1 - \frac{\omega_p^2}{\omega(\omega - j\nu)} , \quad (1.4)$$

$$\epsilon_r = \begin{cases} < 1 & , \quad \omega > \omega_p \\ 0 & , \quad \omega = \omega_p \\ < 0 & , \quad \omega < \omega_p \end{cases}$$

The wave propagates through a medium with e^{-jkr} where k is the wave propagation constant given as;

$$k = \frac{2\pi}{\lambda} = \omega \sqrt{\epsilon \mu} = \omega \sqrt{\epsilon_o \epsilon_r \mu_o} . \quad (1.5)$$

When $\omega < \omega_p$, $\epsilon_r < 0$ and the wave propagation constant, k is purely imaginary. This makes the wave decay exponentially. Hence, the wave travels through plasma

until it reaches a height where electron density is high enough to make the plasma frequency ω_p comparable to wave frequency ω . At that point ϵ_r and n become zero so the wave cannot travel further and to conserve the energy it bounces back. The maximum height a wave travels through before being reflected back is called height of maximum frequency and for F2 layer this height is represented as hmF2. Critical frequency is the highest frequency that will be returned to earth from a given layer of ionosphere. For F2 layer, this is called foF2.

The electron density is an important parameter to be observed in ionospheric studies since it is the main modality that represents the changes in ionosphere [5]. HF signals are transmitted from the ground to ionosphere for long distance communications and the signals with higher frequencies pass through the ionosphere and have delays in proportion to the frequency. These delays are a major reason behind errors in satellite navigation and communication applications. As a result, we need to know the frequencies that can be used for communications and the delays that they will suffer and both of these things depend on the electron density in the ionosphere at any particular time and place. Total Electron Content (TEC) is the line integral of electron density N_e on a ray path and this is used as one of the main observable parameters to characterize ionosphere's behavior. TEC corresponds to the total number of free electrons in a cylinder of 1 m^2 cross sectional area between the satellite and the receiver:

$$TEC(x, y, z, t) = \int_L N_e(l) dl . \quad (1.6)$$

It is measured in TECU which is 10^{16} electrons/ m^2 . Various methods are used for its estimation in ionosphere, such as, beacon satellite measurements, Faraday rotation method etc. Some of them are mentioned in but not limited to [6], [7], [8].

The measurement data for foF2 and hmF2 are collected globally at very sparse locations, hence the task of generating accurate global hmF2 and foF2 maps using sparse data points is practically important. The idea of using denser TEC data and the correlation between TEC and these modalities to generate high resolution foF2 and hmF2 maps is of interest. We investigate various techniques for generating these maps and do a comparative analysis of them. Some sophisticated

statistical mapping techniques are developed that take that secondary information into account for generating more accurate maps. Some part of the foF2 and hmF2 data used in this work is generated using IRI-Plas model and details about that can be found in the next Section.

1.1.1 The Empirical Model Describing Ionosphere

International Reference Ionosphere (IRI) is a climatic and empirical model of the ionosphere. This model takes physical properties of ionosphere into account as well as the huge amount of data collected over the years [9]. It is sponsored by the COmmittee on SPACE Research (COSPAR) and the International Union of Radio Science (URSI). This model was first developed in late sixties and it has been evolving ever since. Continuous efforts are put into its improvement. It is updated every year during special IRI Workshops. The website can be reached at <http://irimodel.org/>

International Reference Ionosphere extended to Plasma sphere (IRI-Plas) is an extended version of the IRI that incorporates TEC, F2 layer critical frequency (foF2) and maximum ionization height in computation of electron density up to the GPS orbital height of 20,000 km [10], [11]. There is a scale parameter set involved which is updated to compensate for changes in the topside of ionosphere [12]. Since the GPS satellites are orbiting the Earth at around 20,200 km altitude, IRI-Plas model generates data very close to actual measurements by the ionosonde stations on geomagnetically quiet days [13]. We use this model for generation of foF2 and hmF2 data. The Figures ?? and 1.2 were obtained from the online source http://www.ionolab.org/index.php?page=iriplasopt_old&language=en and are generated using the IRI-Plas model presented in [14].

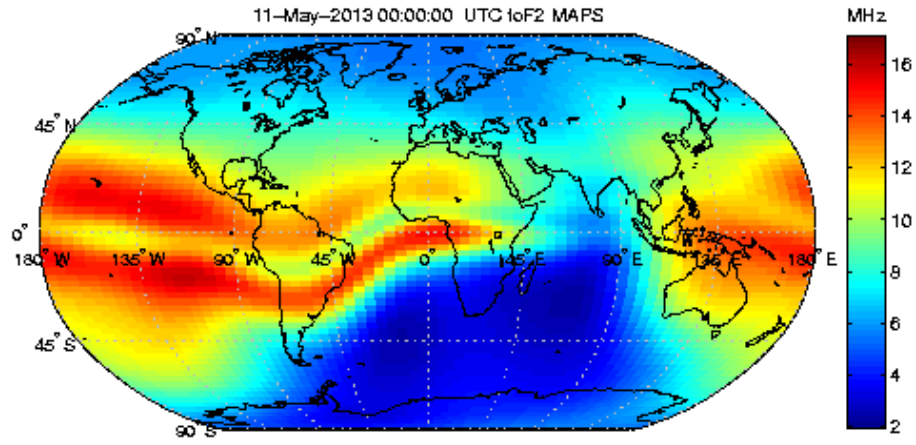


Figure 1.1: foF2 map (MHz), obtained from IONOLAB website, generated by utilizing IRI-Plas model for discrete locations in the world for 11 May 2013, 0 UT [14].

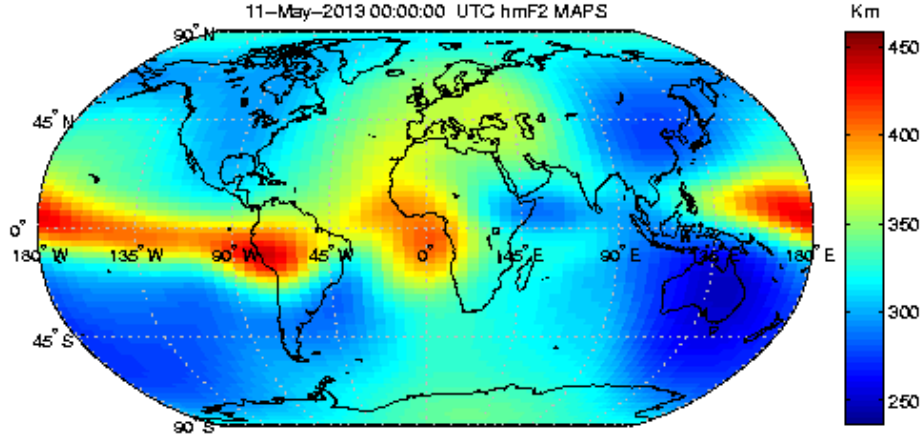


Figure 1.2: hmF2 map (km), obtained from IONOLAB website, generated by utilizing IRI-Plas model for discrete locations in the world for 11 May 2013, 0 UT [14].

In this Section, the structure of the ionosphere and the empirical model describing it was discussed. In the next Section, review of some geostatistical interpolation methods is presented.

1.2 Review of Existing Kriging and Co-Kriging Methods

The idea of Kriging has been around in the literature since 1951. It was first introduced by Danie G. Krige in his Master's thesis. The method became popular

and was named after him. He was originally trying to estimate gold distribution in South Africa by using a number of boreholes in the region of interest [15]. Georges Matheron developed the theoretical basis for this method in 1960 using the thesis presented by Krige and the method has been used in geostatistics ever since that. A very early comparison of Kriging with Polynomial Interpolation Process was presented by G. Matheron in 1967 [16].

Later in 1970, some variants of Kriging were discussed by A. Marechal in [17] and they mentioned certain conditions under which the equations proposed by Matheron could be simplified. They also discussed Universal Kriging, Nugget Effect in interpolation and used Kriging in one of the popular forms still used today.

The interpolation technique started gaining popularity outside the mining field and was used in hydrosociences studies by Delhomme in 1978 [18]. The author presents some case-studies in data input for numerical models, automatic contouring, and average precipitation estimation over a given catchment area (the area from which rainfall flows into a river, lake, or reservoir). This formed the basis for use of Kriging in estimation of rainfall precipitation and the formulation was later used by many others for the same purpose [19], [20], and [21].

The idea of Co-Kriging has been applied by practitioners at the Centre de Géostatistiques (France) for estimation of rainfall [22]. They use rainfall together with elevation at that point for Co-Kriging. They assume a bivariate variogram model between rainfall and elevation and use weighted average to calculate rainfall at unsampled points. The application of Co-Kriging on ionospheric data could not be found in the literature. This thesis contributes towards implementing Co-Kriging algorithm on ionospheric data.

An interesting modality for appraisers and real estate companies is house prices. Kriging and Co-Kriging is used by Jorge Chica-Olmo for the prediction of house prices based on their locations. They take house price as variable of interest for Kriging and use secondary variables like surface area for Co-Kriging to generate house price maps for an area [23].

Extending the idea of Kriging and Co-Kriging from spatial domain to spatiotemporal domain has been studied by Cideris and Gabella [24] and they applied their algorithm on radar and rain gauge data over Switzerland. They introduce the idea of using data for more than one time instant by aggregating over a period of one hour. Such aggregation is not very useful for ionospheric data though because of the rapidly varying temporal data and spatio-temporal variation has to be studied to be able to generate spatio-temporal maps. The idea of spatio-temporal correlation and spatio-temporal covariance models has been studied by H. Liu and B. Yang for Landsat ETM+ NDVI images over the Maumee River Basin [25]. The application of spatio-temporal interpolation using Co-Kriging for ionospheric data could not be found in the previous works.

Time series data and role of covariance functions is important in spatio-temporal modeling. An application of Gaussian Processes, Bayesian learning and a class of covariance models is studied by S. Roberts and M. Osborne in the paper *Gaussian Processes for Time series Modeling*. In that paper, they discuss the importance of Bayesian modeling in time series data and non-parametric models [26]. This forms our basis for understanding time-series data. They also discuss in detail the role of various covariance functions in regression problems. Finally, they mention advantages of using active data selection in case of availability of multiple data streams and failure of one or more sensor data. This idea of active selection is useful in missing data scenarios.

The application of Machine learning for training a Gaussian process is explored by Carl E. Rasmussen [27]. In this paper, he discusses the use of parametric models for mean and covariance functions and how they affect the posterior distribution. They use three parameters for mean and three parameters for covariance function and assume an underlying quadratic function. They conclude by mentioning the advantage of this approach which is that the trade off between the penalty and the data-fit (Occam's razor) is automatic.

In this Section, we saw some uses of Kriging and Co-Kriging methods for some general applications. In the next Section, we study their application specifically in ionospheric studies.

1.3 Review of the Methods used in Ionospheric studies

A comparative analysis of Kriging with an alternative approach for detecting geomagnetic anomalies over Galapagos is studied by Caress in 1989 [28]. They assumed the availability of power spectral density of the field from which the anomaly arises and propose a grid interpolation scheme based on that.

The use of Kriging for improving global ionospheric vertical TEC estimates is analyzed by R. Orús in [29] in 2005. They discuss old University of Catalonia (UPC) maps, called the Global Ionospheric Maps (GIMs), created using GPS data and discuss improving upon them using Kriging. They present a comparative RMSE analysis of how Kriging improves the UPC GIMs.

Dr. Riccardo Notarpietro in the book *Mitigation of Ionospheric Threats to GNSS: an Appraisal of the Scientific and Technological Outputs of the TRANSMIT Project* dedicates a chapter on Kriging of TEC data from ionosphere and explains various semi-variogram models [30]. The chapter limits its content to Kriging of TEC only and does not mention Co-Kriging.

The application of Kriging on ionospheric data and observing the differences between calm days and geomagnetically disturbed days is presented by F. Arıkan and I. Sayin in [31] in 2008. Random Field Priors are also discussed in this paper and error measures for Kriging and Random Field Priors are analyzed over synthetically generated TEC data.

A comprehensive list of techniques applied over ionospheric data can be found in [32]. In addition to Ordinary Kriging and Universal Kriging, this work also discusses the application of Neural Networks in the study of ionospheric parameters. The advantages and disadvantages of using each of the techniques is discussed and sophisticated geostatistical techniques are also compared with some relatively simpler methods such as Inverse Square Distance Weighting method or Cubic Spline method. Performance of these methods is analyzed over synthetic

data and later these methods are applied over the actual TEC data prepared by IONOLAB.

TEC mapping over midlatitude regions using GPS Network sensors is studied by M. N. Deviren, F. Arıkan and O. Arıkan in [33]. After testing several variogram functions, Matérn variogram is chosen for final implementation and the optimization algorithm chosen for parameter estimation is Particle Swarm Optimization. The work in this thesis is based on this and can be considered an extension of this. This thesis extends the capabilities of the softwares developed by IONOLAB in this study and enables the use of secondary modalities in interpolation as well as use of data for multiple time stamps instead of one.

In IONOLAB studies, Ordinary Kriging (which assumes constant mean only in the neighborhood of the point of interest) and Universal Kriging (which assumes a polynomial trend model, such as linear) methods are implemented on synthetic data sets which are assumed to be the possible TEC distributions. These data sets are obtained by adding spatially correlated noise to a trend function. The performance of interpolation with various surface functions, regular and irregular sampling patterns, number of sample points and variogram fitting functions is investigated with both Ordinary Kriging and Universal Kriging methods. We extend this work to Co-Kriging of foF2 and hmF2 with TEC using multivariate variograms. We also incorporate temporal correlations in the data and present Spatio-temporal Kriging as an extension of Kriging for TEC data.

1.4 Sources of Ionospheric Data

The data used in this research comes from multiple sources. We use the sparsely available foF2 and hmF2 data as our primary data and treat the denser TEC data as our secondary data mainly used to add more resolution to foF2 and hmF2 maps. The TEC data used in this work comes from IONOLAB (www.ionolab.org). IONOLAB-TEC is prepared using the Regularized Estimation (Reg-Est) algorithm that takes all the available GPS signals from all satellites and combines

them using the Least Square method along with two step regularization [34]. The data can be downloaded from IONOLAB website using the online tool, in high resolution (30 second), or robust TEC estimates with a 2.5 minute resolution for faster processing [35], [36]. This data will be referred to as $D_{ion-TEC}$ throughout this thesis.

The data for foF2 and hmF2 is used in multiple formats. The first type of data comes from IZMIRAN institute <http://www.izmiran.ru/>. This data has 56 global stations collecting hourly values for these modalities and providing co-located estimates of TEC, foF2, and hmF2. We restrict the region over Europe and use 17 stations from this data that are located in this region. This data will be referred to as D_{iz-TEC} , $D_{iz-foF2}$, and $D_{iz-hmF2}$ for TEC, foF2 and hmF2 data respectively. This data does not come on a regular grid so we apply the grid schemes described in Section 1.7 on this data before using it for Co-Kriging.

Another source of data for these modalities is IRI-Plas-MAP model. The foF2, and hmF2 values are produced by the IRI-Plas 2013 model (provided by Dr. T.L. Gulyaeva) using GIM-TEC maps as external TEC inputs. Two different temporal resolutions are available for this data on ionolab website [37]. For the dates prior to the year 2011, only Jet Propulsion Laboratory (JPL) GIM-TEC maps (having two-hour time resolution) are available for processing. For the dates later than 2010, Polytechnical University of Catalonia (gAGE/UPC) UHR GIM-TEC maps (having one-hour time resolution) are also available for processing. As an IRI-Plas-MAP archive foF2 and hmF2 maps generated using IRI-Plas 2011 are readily available for downloading from the dates between 01-Jan-1999 and 11-May-2013. Two of these maps are shown previously in Figures ??, and 1.2. These maps are available in IONEX as well as .png formats. The data from this source will be referred to as $D_{iri-TEC}$, $D_{iri-foF2}$, and $D_{iri-hmF2}$ for TEC, foF2 and hmF2 data respectively.

The main source of data for foF2 and hmF2 is the ionosonde network. Ionosonde technology was first developed in 1925. Ionosodes are special radars for monitoring the ionosphere, consisting mainly of transmitters and receivers

that send a range of frequencies vertically upward and record the echo of reflected back frequencies. This process is called sounding and is used to determine the appropriate frequencies for communication. The results of this sounding are stored/displayed in forms of ionograms [38]. This is an expensive technology and there are far fewer stations for such sounding than there are for TEC measurements. We use programs developed by IONOLAB to get ionosonde data from ionosonde network servers [39]. This data has 15 minutes temporal resolution. The data from this source will be referred to as $D_{snd-foF2}$, and $D_{snd-hmF2}$ for foF2 and hmF2 data respectively.

1.5 Geomagnetic Storms in Ionosphere

The most frequent behavior of ionosphere is usually geomagnetically quiet but there can be unexpectedly high spikes in amount of ionization sometimes. Such conditions totally change the normal behavior of ionosphere and amount of TEC becomes unexpectedly large or small. Changes in electron density during these times also cause hmF2 and foF2 to change. There are some indices that can be used to categorize a day as a calm day or a geomagnetic storm day. These are Wp, Dst, AE, Kp, and Ap indices [40] and their daily values can be downloaded from the following sources:

<http://www.izmiran.ru/ionosphere/weather/storm/>

<http://wdc.kugi.kyoto-u.ac.jp/dstdir/>, and

<http://wdc.kugi.kyoto-u.ac.jp/kp/>

The unexpectedly high values for these can be an indication of a geomagnetic storm. The tables for these indices for some quiet and disturbed days can be found in [41].

Table 1.1: Geomagnetically Quiet Days.

Quiet Days		
Day	Month	Year
22	April	2009
12	March	2010
15	April	2011
12	June	2011
01	September	2011
21	September	2011
25	December	2011
15	August	2012

Quiet days are the days with Wp index between 2 – 3 and Dst, AE, Kp, Ap indices within their normal bounds. Also there are no earthquakes magnitude 5.5+ in 3 days before or after these dates. Table 1.1 lists the quiet days used in this study.

Table 1.2: Geomagnetically Disturbed Days.

Geomagnetically Disturbed Days		
Day	Month	Year
5	February	2011
10	March	2011
28	May	2011
15	February	2012

Disturbed days used in this study are listed in Table 1.2. These are the dates with Wp index 4+, and extreme values for Dst, AE, Ap and Kp values.

Table 1.3: Disturbed Days with Earthquake.

Disturbed followed by EQ Days			
Day	Month	Year	Earthquake that followed
6	August	2011	Magnitude 5.6 on 11 August 2011 in Southern Xinjiang, China
1	November	2011	Magnitude 5.6 on 09 November 2011 in Eastern Turkey
13	April	2012	Magnitude 5.6 on 16 April 2012 in Southern Greece

Table 1.3 lists the disturbed days used in this study that were followed by an earthquake. These are the days when an earthquake magnitude 5.5+ was observed within a week of a strong geomagnetic activity. The details about earthquakes that happened near the days used in this study are given in appendix A Table A.1

We also used continuous periods of time in this study. These are the days that can belong to any of the above categories but are continuous in time. They are used for spatio-temporal analysis and data for more than one day is to be used for smoothing/prediction using Spatio-temporal Kriging. Initially, we used a period of 10 days from January 15, 2010 to January 25, 2010 for the performance analysis of Spatio-temporal Kriging technique. These days include two disturbed days January 15 and January 24 and remaining quiet days. Later, a period of three months from January 01, 2010 to March 31, 2010 was used for computations of RMSE. The catalog of geomagnetically quiet and disturbed ionospheric days over Europe can be found at http://rwc.cbk.waw.pl/rwc/q_d_days.ct1.

1.6 Correlation between Ionospheric Modalities foF2, hmF2 and TEC

We observe periodically correlated data so we chose a period of 24 hours and see the relationship over a month. Pearson's correlation coefficient was used for computing such correlation. For the vectors $\mathbf{X}$ and $\mathbf{Y}$, Pearson's correlation coefficient is computed as,

$$\rho = \frac{\sum_{i=1}^n (X_i - \mu_X)(Y_i - \mu_Y)}{\left\{ \sum_{i=1}^n (X_i - \mu_X)^2 \sum_{j=1}^n (Y_j - \mu_Y)^2 \right\}^{1/2}} \quad (1.7)$$

where,

$$\mu_X = \frac{1}{n} \sum_{i=1}^n (X_i)$$

and

$$\mu_Y = \frac{1}{n} \sum_{i=1}^n (Y_i)$$

and n is the length of vectors $\mathbf{X}$ and $\mathbf{Y}$, such that $1 \leq i \leq n$. Also,

$$-1 \leq \rho \leq +1 ,$$

where -1 and $+1$ indicate perfectly negative and perfectly positive correlation between $\mathbf{X}$ and $\mathbf{Y}$.

We compute this coefficient using D_{iz-TEC} , $D_{iz-foF2}$, and $D_{iz-hmF2}$ from October 01, 2016 to October 31, 2016 at every hour. n in this case represents the total number of data locations. ρ for every hour in every day, computed over all available stations for D_{iz} is given in Tables in Appendix B.

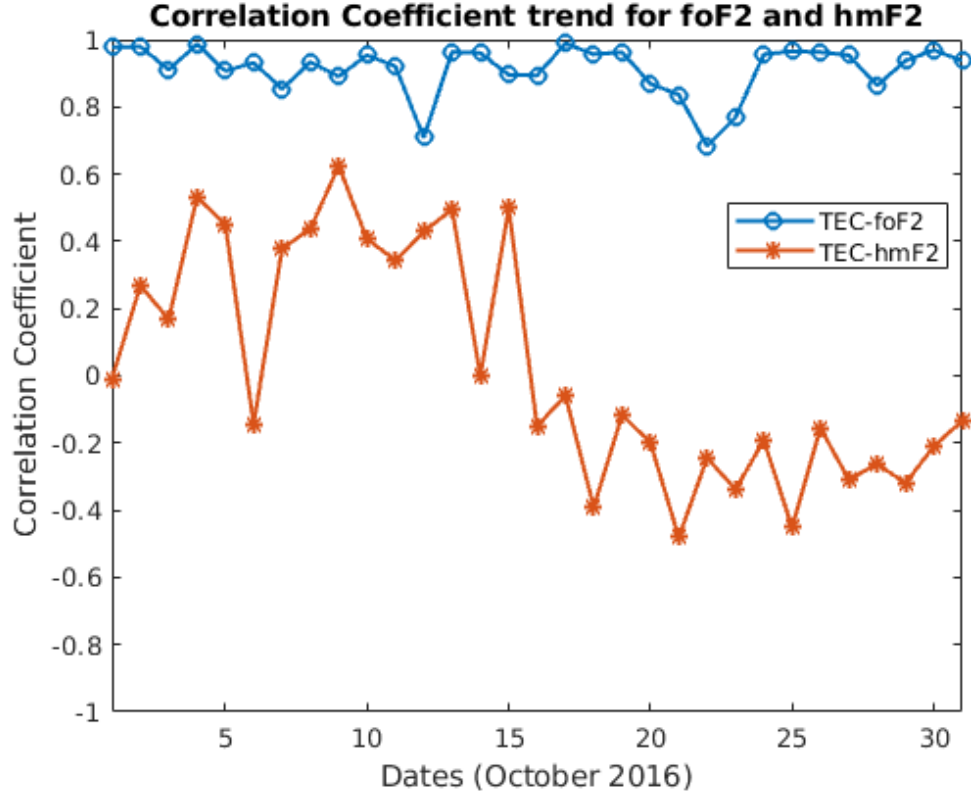


Figure 1.3: Correlation of TEC with foF2 and hmF2 over the month of October 2016, at 12 UT, showing a quiet month with two geomagnetically disturbed days (October 12 and October 23).

The Figure 1.3 shows the correlation of foF2 with TEC and correlation of hmF2 with TEC at 12 UT from October 01, 2016 to October 31, 2016. A strong positive correlation is observed for foF2. The variation in correlation at other hours of the day can be observed in Appendix B.

In the following Section, we define the schemes that will be used to bring data to regular grid.

1.7 The Grid Schemes for Ionospheric Mapping Techniques

The observation data that we have is on irregular points and there is no defined uniform distance between sample points. Moreover, the data for TEC and the data for hmF2 or foF2 may not be available at the same points always since we have more TEC data. This calls for some mechanism to use the available data effectively. We propose three types of techniques here. The final output of all these schemes is fine grid maps with 1 degree resolution for both longitude and latitude.

1.7.1 No Grid: Anisotropic Co-Kriging

The idea here is to use the original data at co-located points for calculation of semi-variogram and cross-variogram parameters and then using all the remaining data for calculating semi-variograms and cross-variograms, keeping those parameters constant. The advantage of this method is that it is robust and computationally less complex and no preprocessing is required on the data. The disadvantage is that it uses only a small portion of actual available data for parameter estimation which might not represent the realistic trends in the data.

1.7.2 Coarse Grid: Inverse Square Distance

To bring the hmF2, foF2 and TEC data to the same grid for Co-Kriging we estimate these values at a regular coarse grid 5 degrees apart in longitude and 2.5 degrees apart in latitude. We use Inverse Square Distance (ISD) mapping which is one of the simplest interpolation technique where distance of all map points from each data point is calculated. Each map point is then calculated as a weighted sum of all the data points with weights being inverse square distance between the map point and the data point. To keep the estimate unbiased, this

sum is divided by sum of weights. The advantage of this method is reduced computational complexity. The disadvantage is that ISD might introduce higher errors for storm days.

1.7.3 Coarse Grid: Ordinary Kriging

Ordinary Kriging is used separately on TEC and on hmF2 or foF2 to generate their coarse resolution maps and these maps are then used as an input for our Co-Kriging system. We use the same 5 degree longitude and 2.5 degree latitude resolution for generating these maps. The advantage includes better estimates than ISD but it comes at the cost of higher computational complexity. Once the data for primary and secondary modalities is available at the same grid points, the task of doing Co-Kriging becomes confined to only co-located Co-Kriging or isotopic Co-Kriging. We use coarse grid data to estimate semi-variogram and cross-variogram parameters and then generate fine grid of 0.1 degree resolution maps.

1.8 Ionospheric Mapping/ Interpolation Techniques

In this Section, we discuss the mapping techniques that we use in simulations over ionospheric data. A theoretical background for these methods is provided here whereas the detailed mathematical formulation will be discussed in the next chapter.

1.8.1 Kriging

Kriging has been used as a mapping technique in meteorology, geology, and mining and it has many versions such as Simple Kriging, Ordinary Kriging, and Universal

Kriging. Kriging is sometimes also referred to as Best Linear Unbiased Estimator (BLUE) and as the name suggests, it is a linear estimator [42]. Kriging calculates the values at points of interest as a weighted sum of the values at other points and its variations are mainly based on assumptions about underlying mean function of the distribution of data.

Due to its popularity among geostatistical techniques, Ordinary Kriging [43] is used to generate maps in this study. Ordinary Kriging also uses distance between data points to explain their correlation [44]. It calculates value at unknown location as a weighted sum of value at other data points and these weights are calculated using semi-variograms.

1.8.1.1 Variogram

Variogram is a representation of correlation between data points as a function of distance. We calculate experimental variogram and then fit a parametric model to it [45] and using the parameters found, we calculate variogram at every map point [46]. We then calculate the weights from these variogram values and generate TEC maps. Variograms are defined by several parameters:

Sill: The variance value at which the variogram levels off. Also, used to refer to the *amplitude* of a certain component of the variogram.

Range: The lag distance at which the variogram reaches the sill value. Presumably, autocorrelation is essentially zero beyond the range.

Nugget: In theory, the variogram value at the origin (0 lag) should be zero. If it is significantly different from zero for lags very close to zero, then this variogram value is referred to as the nugget. The nugget represents variability at distances smaller than the typical sample spacing, including measurement error [47].

In this Section, we briefly discussed the theory of Kriging. In the next Section, we describe the theoretical background of Co-Kriging.

1.8.2 Co-Kriging

The most sophisticated technique used in this study so far is Co-Kriging, which uses secondary correlated data. We observe that TEC, hmF2 and foF2 are correlated with each other and this correlation can help us generate high resolution maps of foF2 and hmF2. Co-Kriging is actually a multivariate version of Kriging that takes point maps of primary and secondary modality and generates high resolution maps for the sparsely sampled modality. The advantage of using TEC data for generating hmF2 or foF2 maps is that we can use the denser TEC network along with the sparse hmF2 or foF2 values and still generate high resolution maps. The details about bringing the data on the same grid as the one for TEC are given in the next Section. Since our primary and secondary modalities are not in the same units, some preprocessing of the secondary modalities is needed to be able to perform mathematical operations on them. We standardize them first and then use in the mapping algorithm.

The main advantage of Co-Kriging comes in the scenarios where [48]

- a) A modality has a distribution with low spatial continuity but it correlates with another modality that shows relatively high continuity. The advantage is higher if the first modality has lesser observations.
- b) A modality is poorly sampled or can be observed only at very sparse locations but it correlates with another modality that is sampled at denser locations. The advantage can be higher if the denser modality is also more precise and this can help take advantage of secondary modality to improve estimations of primary modality.

We use Co-Kriging mainly due to the second reason because our data for hmF2 and foF2 are sparse whereas data for TEC is relatively much dense.

1.8.2.1 Cross-variogram

Variogram and cross-variogram are calculated for primary and secondary modalities and weights are calculated in the same way for estimation of map points. Cross-variograms are a measure of correlation between primary and secondary

modalities as a function of distance. Cross-variogram can be thought of as a multivariate version of semi-variogram. They are often used to describe correlation between the two modalities where one of them is sparsely sampled and another one is densely sampled, leading to their correlation being the key help factor in interpolation of sparse modality. We have to be careful while calculating cross-variogram since our modalities need to be in the same units [47].

In this Section, we described the theory of Co-Kriging and the motivation behind using it. We also discussed the scenarios where Co-Kriging will be most beneficial. In the next Section, we discuss Spatio-temporal Kriging as an extension of Kriging.

1.8.3 Spatio-temporal Kriging

Spatio-temporal Kriging is an extension of Kriging that assumes strong temporal correlations in time-series data. The interpolation is taken from 2 dimensional S domain to 3 dimensional $S \times T$ domain. All the assumptions about spatial correlations in the data that hold for Kriging also hold for Spatio-temporal Kriging. The spatial fields are represented as $Z(S)$ so we represent the spatio-temporal field as $Z(s, t)$ where t in our case represents the date and time. We implement a localized version of Spatio-temporal Kriging, i.e. we restrict the number of neighboring points used in the estimations to reduce the computational complexity and to make the process more efficient [49].

1.8.3.1 Spatio-temporal Variogram

Spatial correlations are analyzed using variograms and spatio-temporal correlations can be represented using spatio-temporal variograms. Such variograms involve calculations over spatial as well as temporal distances.

Unlike spatial variogram, spatio-temporal variogram is not a line but a surface. One axis represents the spatial distances while the other represents temporal

axis and variogram values are plotted as a surface against these axes. Usually the variogram has a low (close to zero) value at origin, indicating a strong correlation between points that are closest to each other in space and time [50]. On the spatial axis, the variogram generally shows a monotonically increasing trend due to decreasing correlation as the points become far apart but on the temporal axis we can see some sinusoidal trends due to periodic nature of ionospheric data. The technique can be used for other data sets as well though that have decreasing correlation against increasing time separations.

Variogram models are fitted to the experimental variograms and details about such models can be found in the formulation Section 2.4.1

1.9 The Map Validation Process

One way to validate the results is doing cross validation. Cross-validation is a reliable performance criteria because it gives us some insight into how the algorithm will perform on the data it has never seen before because the core concept behind cross validation is not using the entire available data for training but leaving some for validation and testing instead. In that manner the left out data will be *new* for the algorithm when it comes to testing the performance. Kriging involves fitting theoretical variogram models to experimental variograms and this model fitting finds the suitable set of parameters to be used in Kriging estimations. We use some part of data as validation data, create multiple sets of parameters using model fitting and then test these fittings on validation data. The set of parameters that performs best on validation data is selected for generating maps and predictions on test points.

Unfortunately, the drawback of such validation and testing is, you need to have sufficient data so you can use some for training and leave some out for validation and testing. For D_{iz} and D_{snd} , while generating foF2 or hmF2 maps, we cannot afford to leave out any data since we already suffer from sparsity of data. One way to overcome that is leave-one-out cross validation but the

computational cost becomes too high. Hence for the sake of performance test of Kriging and Co-Kriging we use IRI-Plas model to generate foF2 and hmF2 data so that we can use some to train the system and leave some for validation and some for test. A common ratio used for such divisions is 60% for training, 20% for validation, and 20% for testing. For Spatio-temporal data we have sufficiently dense data for TEC available and using a number of stations for validation and testing does not result in too small training data. The validation and test stations for Spatio-temporal Kriging are selected such that some test stations are close to the boundary of training region and some are far away from the training neighborhood used (details in Section 3.5). Such selection of test points enables us to see how far we can extrapolate within reasonable margin of error using these algorithms.

After validation we test the selected parameters and methods on test data to provide confidence measures. RMSE is used as a measure of performance and comparison tables are provided in the Section 4.1

1.10 The Flowchart of Co-Kriging

The standard procedure that we use can be summarized in four steps i.e. Input data, Model fitting phase, validation and test phase, and finally map generation phase. Each phase consists of one or more steps involved that are described in the flowchart and algorithm below.

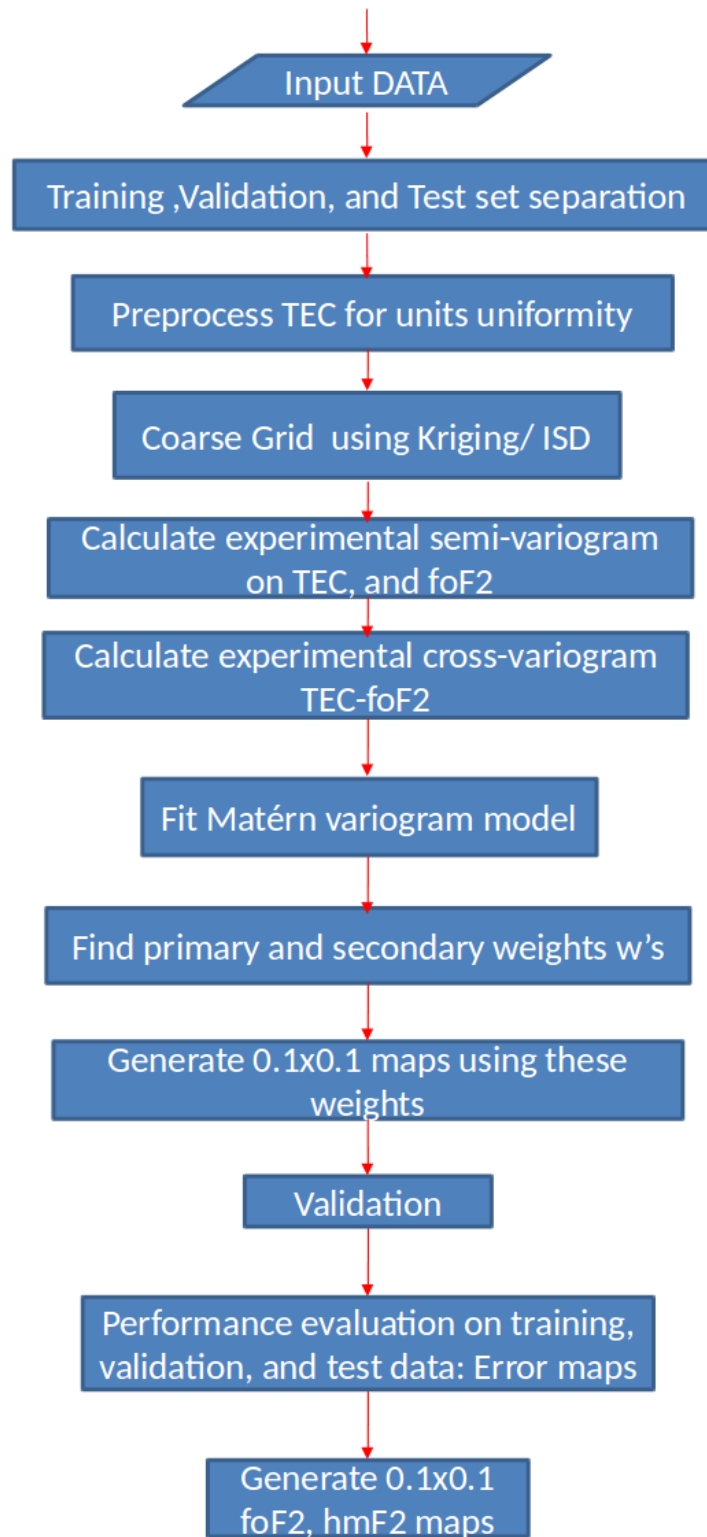


Figure 1.4: Flowchart describing work sequence with respect to all four phases involved.

Algorithm 1: Co-Kriging Algorithm

Input: TEC,foF2/ hmF2

Output: map,error-map

- 1 Separate data into training, validation and test sets according to scheme in Section 1.9
 - 2 Preprocessing:
 - 3 **a:** Preprocess $D_{ion-TEC}$ for units uniformity according to Equation 2.28
 - b:** Estimate $D_{ion-TEC}$ on a $2.5^\circ \times 5^\circ$ grid using Equations 2.6, 2.27
 - c:** Estimate foF2/ hmF2 on a $2.5^\circ \times 5^\circ$ grid using Equations 2.6, 2.27
 - 4 Calculate experimental semi-variogram and cross-variogram according to Equations 2.16, and 2.38
 - 5 Fit theoretical Matérn variogram model given in Equation 2.20
 - 6 Calculate weights for foF2/ hmF2 and TEC according to Equation 2.49
 - 7 Validation
 - 8 Generate fine grid $0.1^\circ \times 0.1^\circ$ resolution map according to Equation 2.51
- Performance analysis: Generate training error and test error maps
-

In this Chapter, we have discussed the nature of ionosphere, its behavior over space and time and the importance of modeling such spatio-temporal behavior. We discussed the literature review on existing techniques and described how this thesis will contribute towards extending those techniques for ionospheric interpolation. We briefly described the theory behind the methods that we are going to implement over ionospheric data in this Chapter. We conclude this Chapter with the work flow of the proposed techniques. In the next Chapter, we discuss the mathematical formulation of these methods.

Chapter 2

Mathematical Formulation of the Interpolation Problem

In this chapter, we discuss a number of ways using which we can generate maps of ionospheric modalities. First, we discuss the problem of data for these modalities not being available at the same locations and how to tackle that using gridding schemes for Co-Kriging. Later, we provide the mathematical formulations of the interpolation schemes proposed. The Kriging and Co-Kriging formulation of [47] is used but the missing steps in his derivation of Co-Kriging are retraced and provided in detail in this chapter. To keep the notations of this formulation in line with the previous works of IONOLAB, all notations are modified to match the notations in [31], [33], [51], and [52].

This work is an extension of previous IONOLAB work on Ordinary Kriging and Universal Kriging by M. N. Deviren [52]. The software developed by IONOLAB is named IONOLAB-MAP and uses Ordinary Kriging or Universal Kriging with linear trend algorithm depending on the region of interest for interpolation. The software takes date, time, data and desired coordinates as input and gives the option to augment GPS data with GIMs. The interpolation is done using the Matérn variogram function and particle swarm optimization and the plots are stored in the user defined directory. This thesis extends the work mentioned

above to Co-Kriging and Spatio-temporal Kriging.

Ionospheric modalities such as TEC, hmF2, foF2 can be modeled as random fields with underlying trend function. These modalities vary in both space and time and we use $\mathbf{x} = [\theta \ \phi]^T$ to represent the spatial location of a data point or an estimation point where θ , and ϕ are latitude and longitude respectively. $Z(\mathbf{x}) = Z(\theta, \phi)$ represents an observation at latitude θ and longitude ϕ and $\hat{Z}(\mathbf{x})$ is the estimate of the predicted value at that point $\mathbf{x}$.

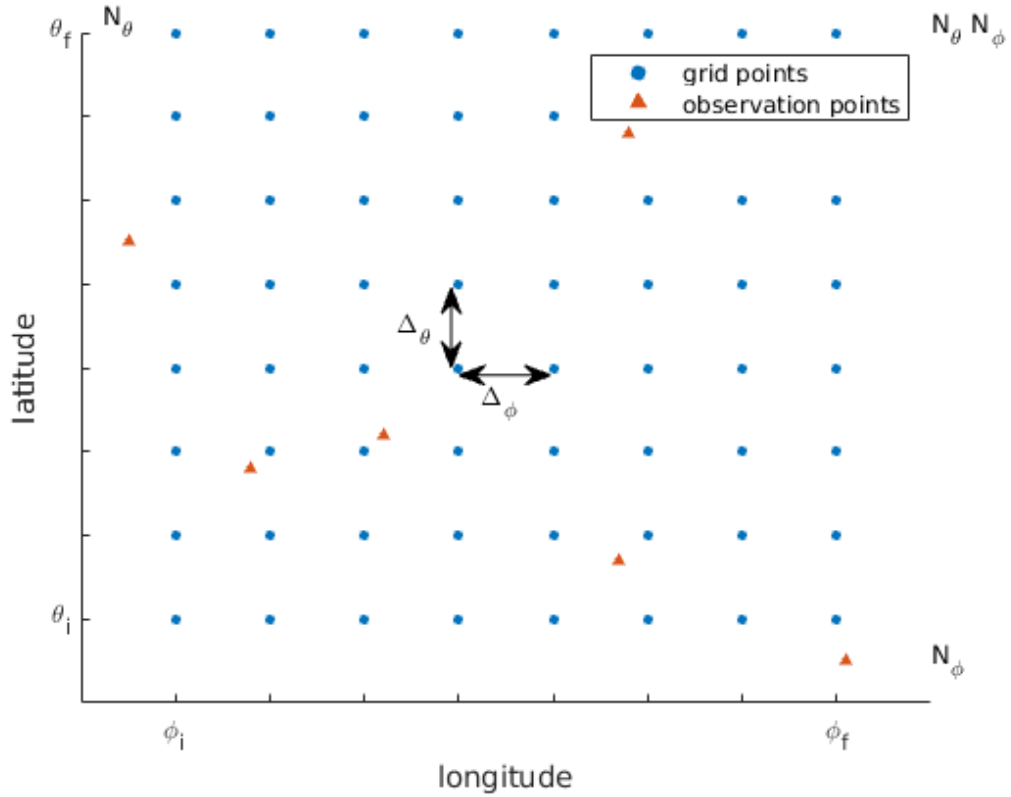


Figure 2.1: Grid Structure defined between the initial coordinates (θ_i and ϕ_i) and final coordinates (θ_f and ϕ_f), with latitude spacing Δ_θ , and longitude spacing Δ_ϕ .

The region of interest is divided into uniformly spaced points with latitude spacing Δ_θ and longitude spacing Δ_ϕ , such that,

$$\Delta_\theta = \frac{\theta_f - \theta_i}{N_\theta}, \quad (2.1)$$

and,

$$\Delta_\phi = \frac{\phi_f - \phi_i}{N_\phi} . \quad (2.2)$$

Any point (θ, ϕ) on the grid can be reached using,

$$\theta = \theta_i + n_\theta \Delta_\theta , \quad (2.3)$$

$$\phi = \phi_i + n_\phi \Delta_\phi , \quad (2.4)$$

where,

$$1 \leq n_\theta \leq N_\theta ,$$

$$1 \leq n_\phi \leq N_\phi ,$$

Every grid point is assigned an index l , such that,

$$l = n_\theta + (n_\phi - 1)N_\theta , \quad (2.5)$$

and,

$$1 \leq l \leq N_\theta N_\phi .$$

Such a distribution forms a grid shown in Figure 2.1 with N_θ latitude and N_ϕ longitude points. At each grid point l , the estimate $\hat{Z}(\mathbf{x}_l)$ is calculated from N_a neighboring observation points $Z(\mathbf{x}_{n_a})$ such that $1 \leq n_a \leq N_a$. In the following sections, we discuss some schemes for such estimation.

2.1 Estimation Using Inverse Square Distance (ISD)

For generating the coarse grid hmF2 and foF2 maps, we prefer Inverse Square Distance over simpler methods like nearest neighbor because the later method generates block effects or patches and to avoid them we can work with inverse square distance technique. In this method, every map point is calculated as a weighted sum of all the observation points with the closest observations being assigned the highest weight. This gives us a smooth map.

$$\hat{Z}(\mathbf{x}_l) = \frac{1}{\sum_{n_a=1}^{N_a} w_{l;n_a}} \sum_{n_a=1}^{N_a} w_{l;n_a} Z(\mathbf{x}_{n_a}) , \quad (2.6)$$

where N_a is the total number of observation points used, $\mathbf{x}_l$ is the grid location where value is to be estimated and $\mathbf{x}_{n_a}$ is the location of the observation point. These weights are essentially the inverse square distance between the two points.

$$w_{l;n_a} = \frac{1}{\|\mathbf{x}_l - \mathbf{x}_{n_a}\|^2} . \quad (2.7)$$

2.2 Estimation Using Ordinary Kriging

Ordinary Kriging (OK) is a geostatistical technique that calculates the values at the grid points from the nearby values. It assumes a constant mean as the underlying trend function $\mu(\mathbf{x}) = \mu$. This constant is unknown and the following conditions should be fulfilled in order to work with ordinary Kriging.

$$E\{Z(\mathbf{x}) - Z(\mathbf{x} + h)\} = 0 . \quad (2.8)$$

$$\text{var}\{Z(\mathbf{x}) - Z(\mathbf{x} + h)\} = 2\gamma_{ZZ}(h) . \quad (2.9)$$

where $\gamma_{ZZ}(h)$ is variogram between points that are h distance apart. Ordinary Kriging predictor is a linear combination of data values at other points, i.e.

$$\hat{Z}(\mathbf{x}_l) = \sum_{n_a=1}^{N_a} w_{l;n_a} Z(\mathbf{x}_{n_a}) , \quad (2.10)$$

where l is the index of the predicted value at grid point $\mathbf{x}_l$, N_a is the total number of neighboring observation points used in the prediction, and $w_{l;n_a}$ is the weight of the n_a^{th} observation at grid index l . For unbiasedness,

$$E\{\hat{Z}(\mathbf{x}_l)\} = E\left\{\sum_{n_a=1}^{N_a} w_{l;n_a} Z(\mathbf{x}_{n_a})\right\} , \quad (2.11)$$

which employs,

$$\forall l, \sum_{n_a=1}^{N_a} w_{l;n_a} = 1 .$$

Following the derivation in [52], predictor variance can be written as,

$$\begin{aligned}
F(w_l) &= \text{var}[Z(\mathbf{x}_l) - \hat{Z}(\mathbf{x}_l)] \\
&= E\{[Z(\mathbf{x}_l) - \hat{Z}(\mathbf{x}_l)]^2\} \\
&= E\left\{\left[Z(\mathbf{x}_l) - \sum_{n_a=1}^{N_a} w_{l;n_a} Z(\mathbf{x}_{n_a})\right]^2\right\}.
\end{aligned} \tag{2.12}$$

We seek to minimize the predictor variance, $F(w_l)$,

$$\begin{aligned}
F(w_l) &= E\{Z^2(\mathbf{x}_l) - 2 \sum_{n_a=1}^{N_a} w_{l;n_a} Z(\mathbf{x}_l) Z(\mathbf{x}_{n_a}) \\
&\quad + \sum_{n_a=1}^{N_a} \sum_{n_b=1}^{N_a} w_{l;n_a} w_{l;n_b} Z(\mathbf{x}_{n_a}) Z(\mathbf{x}_{n_b})\}.
\end{aligned} \tag{2.13}$$

Adding and subtracting $\sum_{n_a=1}^{N_a} w_{l;n_a} Z^2(\mathbf{x}_{n_a})$, Equation 2.13 becomes,

$$\begin{aligned}
F(w_l) &= E\{Z^2(\mathbf{x}_l) - 2 \sum_{n_a=1}^{N_a} w_{l;n_a} Z(\mathbf{x}_l) Z(\mathbf{x}_{n_a}) + \sum_{n_a=1}^{N_a} w_{l;n_a} Z^2(\mathbf{x}_{n_a}) \\
&\quad - \sum_{n_a=1}^{N_a} w_{l;n_a} Z^2(\mathbf{x}_{n_a}) + \sum_{n_a=1}^{N_a} \sum_{n_b=1}^{N_a} w_{l;n_a} w_{l;n_b} Z(\mathbf{x}_{n_a}) Z(\mathbf{x}_{n_b})\}.
\end{aligned}$$

Using the fact that $\sum_{n_a=1}^{N_a} w_{l;n_a} = 1$

$$\begin{aligned}
F(w_l) &= E\left\{\sum_{n_a=1}^{N_a} w_{l;n_a} [Z(\mathbf{x}_l) - Z(\mathbf{x}_{n_a})]^2 - \sum_{n_a=1}^{N_a} w_{l;n_a} \frac{Z^2(\mathbf{x}_{n_a})}{2} \right. \\
&\quad \left. - \sum_{n_a=1}^{N_a} w_{l;n_a} \frac{Z^2(\mathbf{x}_{n_a})}{2} + \sum_{n_a=1}^{N_a} \sum_{n_b=1}^{N_a} w_{l;n_a} w_{l;n_b} Z(\mathbf{x}_{n_a}) Z(\mathbf{x}_{n_b})\right\},
\end{aligned}$$

completing squares, Equation 2.13 becomes,

$$\begin{aligned}
F(w_l) &= E\left\{\sum_{n_a=1}^{N_a} w_{l;n_a} [Z(\mathbf{x}_l) - Z(\mathbf{x}_{n_a})]^2 \right. \\
&\quad \left. - \sum_{n_a=1}^{N_a} \sum_{n_b=1}^{N_a} w_{l;n_a} w_{l;n_b} \frac{[Z(\mathbf{x}_{n_a}) - Z(\mathbf{x}_{n_b})]^2}{2}\right\}.
\end{aligned} \tag{2.14}$$

We know that,

$$2\gamma_{ZZ}(\mathbf{x}_{n_a}, \mathbf{x}_{n_b}) = E\{[Z(\mathbf{x}_{n_a}) - Z(\mathbf{x}_{n_b})]^2\}. \tag{2.15}$$

The experimental variogram is calculated as

$$\gamma_{ZZ;e}(h) = \frac{1}{2N(h)} \sum_{n_a=1}^{N(h)} [Z(\mathbf{x}_{n_a}) - Z(\mathbf{x}_{n_a} + h)]^2, \quad (2.16)$$

where $N(h)$ is the number of observation pairs that are h Euclidean distance apart. Next, we fit a theoretical variogram model to the experimental variogram.

2.2.1 Theoretical Variogram Models

Some theoretical variogram models are given in [53]. The exponential variogram model is defined through parameter set $\mathbf{P}_e = [\alpha_{1,e} \ \alpha_{2,e}]^T$ as,

$$\hat{\gamma}^E(h, \mathbf{P}_e) = \alpha_{1,e}^2 \left(1 - \exp\left(\frac{-h}{\alpha_{2,e}}\right) \right) \quad (2.17)$$

Another variogram model is Gaussian model, defined through parameter set $\mathbf{P}_g = [\alpha_{1,g} \ \alpha_{2,g}]^T$ as,

$$\hat{\gamma}^G(h, \mathbf{P}_g) = \alpha_{1,g}^2 \left(1 - \exp\left(\frac{-h^2}{\alpha_{2,g}}\right) \right) \quad (2.18)$$

The Spherical variogram model using parameter set $\mathbf{P}_s = [\alpha_{1,s} \ \alpha_{2,s}]^T$

$$\hat{\gamma}^S(h, \mathbf{P}_s) = \alpha_{1,s}^2 \left(1.5 \frac{h}{\alpha_{2,s}} - 0.5 \frac{h^3}{\alpha_{2,s}^3} \right) \quad (2.19)$$

This Matérn family actually represents multiple variogram models at the same time. The flexibility of choosing the shape parameter allows it to take the form of any other variogram model so instead of using one particular model, we work with Matérn family with parameters $\mathbf{P}_m = [\alpha_1 \ \alpha_2 \ \alpha_3]^T$.

$$\hat{\gamma}^M(h, \mathbf{P}_m) = \alpha_1^2 \left(1 - \frac{1}{2^{\alpha_2-1} \Gamma(\alpha_3)} \left(\frac{h}{\alpha_2} \right)^{\alpha_3} K_{\alpha_3} \left(\frac{h}{\alpha_2} \right) \right), \quad (2.20)$$

where K_{α_3} is the modified Bessel function of the second kind and α_3 is the shape smoothness parameter [54] and its value can range between 0 to ∞ . Depending on the value of α_3 , the Matérn variogram corresponds to other models. For $\alpha_3 = 0.5$, this model reduces to the exponential model in Equation 2.17, and for $\alpha_3 = \infty$, this model reduces to the Gaussian model in Equation 2.18.

Replacing the variogram definition of Equation 2.15 in Equation 2.14

$$F(w_l) = 2 \sum_{n_a=1}^{N_a} w_{l;n_a} \gamma_{ZZ}(\mathbf{x}_l, \mathbf{x}_{n_a}) - \sum_{n_a=1}^{N_a} \sum_{n_b=1}^{N_a} w_{l;n_a} w_{l;n_b} \gamma_{ZZ}(\mathbf{x}_{n_a}, \mathbf{x}_{n_b}) . \quad (2.21)$$

We want to minimize $F(w_l)$ subject to $K(w_l)$ such that $K(w_l) = \sum_{n_a=1}^{N_a} w_{l;n_a} - 1$, which is the unbiasedness condition for the predictor. We use Lagrangian formulation for this purpose as

$$L(w_l) = F(w_l) - \lambda K(w_l) . \quad (2.22)$$

where λ is a Lagrange multiplier. Differentiating L with respect to $w_{l;n_a}$,

$$\frac{dL(w_l)}{dw_{l;n_a}} = 2\gamma_{ZZ}(\mathbf{x}_l, \mathbf{x}_{n_a}) - 2 \sum_{n_b=1}^{N_a} w_{l;n_b} \gamma_{ZZ}(\mathbf{x}_{n_a}, \mathbf{x}_{n_b}) - \lambda , \quad (2.23)$$

for $n_a = 1 \dots N_a$.

Now, differentiating L with respect to λ ,

$$\frac{dL(w_l)}{d\lambda} = - \sum_{n_a=1}^{N_a} w_{l;n_a} + 1 . \quad (2.24)$$

$$\underbrace{\begin{bmatrix} \gamma_{ZZ}(1,1) & \gamma_{ZZ}(1,2) & \cdots & \gamma_{ZZ}(1,N_a) & 1 \\ \gamma_{ZZ}(2,1) & \gamma_{ZZ}(2,2) & \cdots & \gamma_{ZZ}(2,N_a) & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \gamma_{ZZ}(N_a,1) & \gamma_{ZZ}(N_a,2) & \cdots & \gamma_{ZZ}(N_a,N_a) & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{bmatrix}}_{\hat{\mathbf{F}}} \underbrace{\begin{bmatrix} w_{l;1} \\ w_{l;2} \\ \vdots \\ w_{l;N_a} \\ \lambda/2 \end{bmatrix}}_{\mathbf{w}} = \underbrace{\begin{bmatrix} \gamma_{ZZ}(l,1) \\ \gamma_{ZZ}(l,2) \\ \vdots \\ \gamma_{ZZ}(l,N_a) \\ 1 \end{bmatrix}}_{\hat{\mathbf{F}}_1} . \quad (2.25)$$

The symbols $\hat{\mathbf{F}}$, $\mathbf{w}$, and $\hat{\mathbf{F}}_1$ placed underneath the matrices in Equation 2.25 will be used for these matrices. The matrix $\hat{\mathbf{F}}$ is usually invertible and we find the values for estimates of $\mathbf{w}$'s and determine $\hat{Z}(\mathbf{x}_l)$ as

$$\hat{\mathbf{w}} = (\hat{\mathbf{F}}^T \hat{\mathbf{F}})^{-1} \hat{\mathbf{F}}^T \hat{\mathbf{F}}_1 . \quad (2.26)$$

$$\hat{Z}(\mathbf{x}_l) = \sum_{n_a=1}^{N_a} \hat{w}_{l;n_a} Z(\mathbf{x}_{n_a}) . \quad (2.27)$$

In this section we discussed the estimation of a modality using Ordinary Kriging. The next section describes estimation of a modality using Co-Kriging.

2.3 Co-Kriging

Co-Kriging is a multivariate version of Kriging that uses one modality as primary and another one as secondary to generate high resolution maps for the primary modality. The data used for Co-Kriging is assumed to have been pre-processed to bring both primary and secondary modalities to the same grid and same units. The modality to be estimated is represented as $\hat{Z}(\mathbf{x})$ whereas primary and secondary data used is represented as $Z(\mathbf{x})$ and $Y(\mathbf{x})$ respectively. For the sake of differentiating between original data and pre-processed data and for the sake of continuation in the notation, we use $\tilde{Y}(\mathbf{x})$ for the preprocessed data. Since the modalities have different units, we need to be careful with our mathematical operations that we perform on them. One way to bring them in the same units is doing the following on the secondary modality.

$$\tilde{Y}(\mathbf{x}) = \frac{Y(\mathbf{x}) - \mu_Y}{\sigma_Y} \sigma_Z, \quad (2.28)$$

where μ_Y is the mean of the observed secondary data, calculated as,

$$\mu_Y = \frac{1}{N_y} \sum_{n_y=1}^{N_y} Y(\mathbf{x}_{n_y}) \quad (2.29)$$

where, $1 \leq n_y \leq N_y$. N_y is the total number of observation points for secondary modality. In Equation 2.28, σ_Y is the standard deviation of the observed secondary data, and σ_Z is the standard deviation of the observed primary data, calculated as,

$$\sigma_Y = \sqrt{\frac{\sum_{n_y=1}^{N_y} (Y(\mathbf{x}_{n_y}) - \mu_Y)^2}{N_y - 1}}, \quad (2.30)$$

and,

$$\sigma_Z = \sqrt{\frac{\sum_{n_z=1}^{N_z} (Z(\mathbf{x}_{n_z}) - \mu_Z)^2}{N_z - 1}}, \quad (2.31)$$

where,

$$1 \leq n_z \leq N_z .$$

N_z is the total number of observation points for primary modality. The predictor in this case becomes,

$$\hat{Z}(\mathbf{x}_l) = \sum_{n_y=1}^{N_y} w_{\tilde{Y};l,n_y} \tilde{Y}(\mathbf{x}_{n_y}) + \sum_{n_z=1}^{N_z} w_{Z;l,n_z} Z(\mathbf{x}_{n_z}) , \quad (2.32)$$

and $w_{\tilde{Y};l,n_y}$ and $w_{Z;l,n_z}$ are the corresponding weights. For the predictor to be unbiased,

$$E\{\hat{Z}(\mathbf{x}_l)\} = E\left\{ \sum_{n_y=1}^{N_y} w_{\tilde{Y};l,n_y} \tilde{Y}(\mathbf{x}_{n_y}) + \sum_{n_z=1}^{N_z} w_{Z;l,n_z} Z(\mathbf{x}_{n_z}) \right\} . \quad (2.33)$$

which gives us the following sufficient conditions, as given in [47].

$$\sum_{n_z=1}^{N_z} w_{Z;l,n_z} = 1 . \quad (2.34)$$

$$\sum_{n_y=1}^{N_y} w_{\tilde{Y};l,n_y} = 0 . \quad (2.35)$$

Our objective in this derivation is to minimize the predictor variance with respect to some constraints for unbiasedness. The predictor variance $F(w_l)$, given in Equation 2.12 in Co-Kriging context becomes,

$$\begin{aligned} F(w_l) &= E\{[Z(\mathbf{x}_l) - \hat{Z}(\mathbf{x}_l)]^2\} \\ &= E\left\{ \left[Z(\mathbf{x}_l) - E\left\{ \sum_{n_y=1}^{N_y} w_{\tilde{Y};l,n_y} \tilde{Y}(\mathbf{x}_{n_y}) - \sum_{n_z=1}^{N_z} w_{Z;l,n_z} Z(\mathbf{x}_{n_z}) \right\} \right]^2 \right\} \\ &= E\{Z^2(\mathbf{x}_l) + \sum_{n_y=1}^{N_y} \sum_{n_c=1}^{N_y} w_{\tilde{Y};l,n_y} w_{\tilde{Y};l,n_c} \tilde{Y}(\mathbf{x}_{n_y}) \tilde{Y}(\mathbf{x}_{n_c}) \\ &\quad + \sum_{n_z=1}^{N_z} \sum_{n_d=1}^{N_z} w_{Z;l,n_z} w_{Z;l,n_d} Z(\mathbf{x}_{n_z}) Z(\mathbf{x}_{n_d}) - 2 \sum_{n_y=1}^{N_y} w_{\tilde{Y};l,n_y} \tilde{Y}(\mathbf{x}_{n_y}) Z(\mathbf{x}_l) \\ &\quad + 2 \sum_{n_y=1}^{N_y} \sum_{n_z=1}^{N_z} w_{\tilde{Y};l,n_y} w_{Z;l,n_z} \tilde{Y}(\mathbf{x}_{n_y}) Z(\mathbf{x}_{n_z}) - 2 \sum_{n_z=1}^{N_z} w_{Z;l,n_z} Z(\mathbf{x}_{n_z}) Z(\mathbf{x}_l)\} \end{aligned} \quad (2.36)$$

where n_c and n_d are just alternative indices used while opening square, and

$$1 \leq n_c \leq N_y, \quad 1 \leq n_d \leq N_z .$$

Using the fact that $\sum_{n_z=1}^{N_z} w_{Z;l,n_z} Z(\mathbf{x}_{n_z}) = 1$ and $\sum_{n_y=1}^{N_y} w_{\tilde{Y};l,n_y} = 0$, following the derivation in [55], we use completing square method to rearrange the terms in order to form the definition of variogram and cross variogram. Following such rearrangement, Equation 2.36 becomes,

$$\begin{aligned} F(w_l) = & E \left\{ - \sum_{n_y=1}^{N_y} \sum_{n_c=1}^{N_y} w_{\tilde{Y};l,n_y} w_{\tilde{Y};l,n_c} \frac{[\tilde{Y}(\mathbf{x}_{n_y}) - \tilde{Y}(\mathbf{x}_{n_c})]^2}{2} \right. \\ & + 2 \sum_{n_y=1}^{N_y} w_{\tilde{Y};l,n_y} \frac{[Z(\mathbf{x}_l) - \tilde{Y}(\mathbf{x}_{n_y})]^2}{2} + 2 \sum_{n_z=1}^{N_z} w_{Z;l,n_z} \frac{[Z(\mathbf{x}_l) - Z(\mathbf{x}_{n_z})]^2}{2} \\ & - 2 \sum_{n_y=1}^{N_y} \sum_{n_z=1}^{N_z} w_{\tilde{Y};l,n_y} w_{Z;l,n_z} \frac{[\tilde{Y}(\mathbf{x}_{n_y}) - Z(\mathbf{x}_{n_z})]^2}{2} \\ & \left. - \sum_{n_z=1}^{N_z} \sum_{n_d=1}^{N_z} w_{Z;l,n_z} w_{Z;l,n_d} \frac{[Z(\mathbf{x}_{n_z}) - Z(\mathbf{x}_{n_d})]^2}{2} \right\} . \end{aligned}$$

All the terms in the equation above are now in the form of complete squares and we can replace that with the definition of cross-variogram [55]:

$$\begin{aligned} 2\gamma_{\tilde{Y}Z}(\mathbf{x}_{n_y}, \mathbf{x}_{n_z}) &= \text{var}[\tilde{Y}(\mathbf{x}_{n_y}) - Z(\mathbf{x}_{n_z})] \\ &= E\{[\tilde{Y}(\mathbf{x}_{n_y}) - Z(\mathbf{x}_{n_z})]^2\} \end{aligned} \quad (2.37)$$

We calculate the experimental cross-variogram as,

$$\gamma_{\tilde{Y}Z;e}(h) = \frac{1}{2N(h)} \sum_{n_a=1}^{N(h)} [\tilde{Y}(\mathbf{x}_{n_a}) - Z(\mathbf{x}_{n_a} + h)]^2 , \quad (2.38)$$

where $N(h)$ is the number of observation points h distance apart. We fit the same theoretical model as used in Ordinary Kriging, Equation 2.20. $\hat{\gamma}_{ZZ}$ shows an estimate of the experimental variogram $\gamma_{ZZ;e}$, obtained through model fitting.

Making this replacement,

$$\begin{aligned}
F(w_l) = & - \sum_{n_y=1}^{N_y} \sum_{n_c=1}^{N_y} w_{\tilde{Y};l,n_y} w_{\tilde{Y};l,n_c} \hat{\gamma}_{\tilde{Y}\tilde{Y}}(\mathbf{x}_{n_y}, \mathbf{x}_{n_c}) \\
& + 2 \sum_{n_y=1}^{N_y} w_{\tilde{Y};l,n_y} \hat{\gamma}_{Z\tilde{Y}}(\mathbf{x}_l, \mathbf{x}_{n_y}) + 2 \sum_{n_z=1}^{N_z} w_{Z;l,n_z} \hat{\gamma}_{ZZ}(\mathbf{x}_l, \mathbf{x}_{n_z}) \\
& - 2 \sum_{n_y=1}^{N_y} \sum_{n_z=1}^{N_z} w_{\tilde{Y};l,n_y} w_{Z;l,n_z} \hat{\gamma}_{\tilde{Y}Z}(\mathbf{x}_{n_y}, \mathbf{x}_{n_z}) \\
& - \sum_{n_z=1}^{N_z} \sum_{n_d=1}^{N_z} w_{Z;l,n_z} w_{Z;l,n_d} \hat{\gamma}_{ZZ}(\mathbf{x}_{n_z}, \mathbf{x}_{n_d}) .
\end{aligned}$$

subject to the unbiasedness conditions,

$$J_1(w_{Z;l,n_z}) = \sum_{n_z=1}^{N_z} w_{Z;l,n_z} - 1 ,$$

and,

$$J_2(w_{\tilde{Y};l,n_y}) = \sum_{n_y=1}^{N_y} w_{\tilde{Y};l,n_y} .$$

The cost function $M(w_l)$ is defined as,

$$M(w_l) = F(w_l) - \lambda_1 J_1(w_{Z;l,n_z}) - \lambda_2 J_2(w_{\tilde{Y};l,n_y}) . \quad (2.39)$$

Differentiating with respect to $w_{\tilde{Y};l,n_y}, w_{Z;l,n_z}, \lambda_1$ and λ_2 gives

$$\begin{aligned}
\forall n_y, \quad \frac{dM(w_l)}{dw_{\tilde{Y};l,n_y}} = & - 2 \sum_{n_c=1}^{N_y} w_{\tilde{Y};l,n_c} \hat{\gamma}_{\tilde{Y}\tilde{Y}}(\mathbf{x}_{n_y}, \mathbf{x}_{n_c}) + 2 \hat{\gamma}_{Z\tilde{Y}}(\mathbf{x}_l, \mathbf{x}_{n_y}) \\
& - 2 \sum_{n_z=1}^{N_z} w_{Z;l,n_z} \hat{\gamma}_{\tilde{Y}Z}(\mathbf{x}_{n_y}, \mathbf{x}_{n_z}) - \lambda_1 ,
\end{aligned} \quad (2.40)$$

$$\begin{aligned}
\forall n_z, \quad \frac{dM(w_l)}{dw_{Z;l,n_z}} = & 2 \hat{\gamma}_{ZZ}(\mathbf{x}_l, \mathbf{x}_{n_z}) - 2 \sum_{n_y=1}^{N_y} w_{\tilde{Y};l,n_y} \hat{\gamma}_{\tilde{Y}Z}(\mathbf{x}_{n_y}, \mathbf{x}_{n_z}) \\
& - 2 \sum_{n_d=1}^{N_z} w_{Z;l,n_d} \hat{\gamma}_{ZZ}(\mathbf{x}_{n_z}, \mathbf{x}_{n_d}) - \lambda_2 ,
\end{aligned} \quad (2.41)$$

$$\frac{dM(w_l)}{d\lambda_1} = \sum_{n_z=1}^n w_{Z;l,n_z} + 1 . \quad (2.42)$$

$$\frac{dM(w_l)}{d\lambda_2} = \sum_{n_y=1}^{N_y} w_{\tilde{Y};l,n_y} . \quad (2.43)$$

To represent the system of equations in a matrix form, let $\mathbf{G}_{ZZ}$, $\mathbf{G}_{Z\tilde{Y}}$, $\mathbf{G}_{\tilde{Y}Z}$, and $\mathbf{G}_{\tilde{Y}\tilde{Y}}$ be matrices, such that,

$$\mathbf{G}_{ZZ} = \begin{bmatrix} \hat{\gamma}_{ZZ}(\mathbf{x}_{n_1}, \mathbf{x}_{n_1}) & \cdots & \hat{\gamma}_{ZZ}(\mathbf{x}_{n_1}, \mathbf{x}_{N_z}) \\ \vdots & \ddots & \vdots \\ \hat{\gamma}_{ZZ}(\mathbf{x}_{N_z}, \mathbf{x}_{n_1}) & \cdots & \hat{\gamma}_{ZZ}(\mathbf{x}_{N_z}, \mathbf{x}_{N_z}) \end{bmatrix} . \quad (2.44)$$

$$\mathbf{G}_{Z\tilde{Y}} = \begin{bmatrix} \hat{\gamma}_{Z\tilde{Y}}(\mathbf{x}_{n_1}, \mathbf{x}_{n_1}) & \cdots & \hat{\gamma}_{Z\tilde{Y}}(\mathbf{x}_{n_1}, \mathbf{x}_{N_y}) \\ \vdots & \ddots & \vdots \\ \hat{\gamma}_{Z\tilde{Y}}(\mathbf{x}_{N_z}, \mathbf{x}_{n_1}) & \cdots & \hat{\gamma}_{Z\tilde{Y}}(\mathbf{x}_{N_z}, \mathbf{x}_{N_y}) \end{bmatrix} . \quad (2.45)$$

$$\mathbf{G}_{\tilde{Y}Z} = \begin{bmatrix} \hat{\gamma}_{\tilde{Y}Z}(\mathbf{x}_{n_1}, \mathbf{x}_{n_1}) & \cdots & \hat{\gamma}_{\tilde{Y}Z}(\mathbf{x}_{n_1}, \mathbf{x}_{N_z}) \\ \vdots & \ddots & \vdots \\ \hat{\gamma}_{\tilde{Y}Z}(\mathbf{x}_{N_y}, \mathbf{x}_{n_1}) & \cdots & \hat{\gamma}_{\tilde{Y}Z}(\mathbf{x}_{N_y}, \mathbf{x}_{N_z}) \end{bmatrix} . \quad (2.46)$$

$$\mathbf{G}_{\tilde{Y}\tilde{Y}} = \begin{bmatrix} \hat{\gamma}_{\tilde{Y}\tilde{Y}}(\mathbf{x}_{n_1}, \mathbf{x}_{n_1}) & \cdots & \hat{\gamma}_{\tilde{Y}\tilde{Y}}(\mathbf{x}_{n_1}, \mathbf{x}_{N_y}) \\ \vdots & \ddots & \vdots \\ \hat{\gamma}_{\tilde{Y}\tilde{Y}}(\mathbf{x}_{N_y}, \mathbf{x}_{n_1}) & \cdots & \hat{\gamma}_{\tilde{Y}\tilde{Y}}(\mathbf{x}_{N_y}, \mathbf{x}_{N_y}) \end{bmatrix} . \quad (2.47)$$

$$\underbrace{\begin{bmatrix} \mathbf{G}_{ZZ} & \mathbf{G}_{Z\tilde{Y}} & \underline{1}^T & \underline{0}^T \\ \mathbf{G}_{\tilde{Y}Z} & \mathbf{G}_{\tilde{Y}\tilde{Y}} & \underline{0}^T & \underline{1}^T \\ \underline{1} & \underline{0} & 0 & 0 \\ \underline{0} & \underline{1} & 0 & 0 \end{bmatrix}}_{\hat{\mathbf{r}}} \underbrace{\begin{bmatrix} \underline{w_Z} \\ \underline{w_{\tilde{Y}}} \\ \frac{\lambda_1}{2} \\ \frac{\lambda_2}{2} \end{bmatrix}}_{\mathbf{W}} = \underbrace{\begin{bmatrix} \mathbf{G}_{lZ} \\ \mathbf{G}_{l\tilde{Y}} \\ 1 \\ 0 \end{bmatrix}}_{\hat{\mathbf{r}}_l} , \quad (2.48)$$

where,

$$\underline{w_Z} = \begin{bmatrix} w_{Z;l,n_1} & \cdots & w_{Z;l,N_z} \end{bmatrix}^T ,$$

$$\underline{w_{\tilde{Y}}} = \begin{bmatrix} w_{\tilde{Y};l,n_1} & \cdots & w_{\tilde{Y};l,N_y} \end{bmatrix}^T ,$$

$$\mathbf{G}_{lZ} = \begin{bmatrix} \hat{\gamma}_{ZZ}(\mathbf{x}_{n_l}, \mathbf{x}_{n_1}) & \cdots & \hat{\gamma}_{ZZ}(\mathbf{x}_{n_l}, \mathbf{x}_{N_z}) \end{bmatrix}^T ,$$

$$\mathbf{G}_{l\tilde{Y}} = \begin{bmatrix} \hat{\gamma}_{Z\tilde{Y}}(\mathbf{x}_{n_l}, \mathbf{x}_{n_1}) & \cdots & \hat{\gamma}_{Z\tilde{Y}}(\mathbf{x}_{n_l}, \mathbf{x}_{N_y}) \end{bmatrix}^T ,$$

and $\underline{1}^T$ and $\underline{0}^T$ are the vectors of length $N_z + N_y$.

$$\hat{\mathbf{W}} = (\hat{\mathbf{\Gamma}}^T \hat{\mathbf{\Gamma}})^{-1} \hat{\mathbf{\Gamma}}^T \hat{\mathbf{\Gamma}}_1, \quad (2.49)$$

where,

$$\hat{\mathbf{W}} = \begin{bmatrix} \hat{w}_{Z;l,n_1} & \cdots & \hat{w}_{Z;l,N_z} & \hat{w}_{\tilde{Y};l,n_1} & \cdots & \hat{w}_{\tilde{Y};l,N_y} & \lambda_1/2 & \lambda_2/2 \end{bmatrix}^T \quad (2.50)$$

$$\hat{Z}(\mathbf{x}_l) = \sum_{n_y=1}^{N_y} \hat{w}_{\tilde{Y};l,n_y} \tilde{Y}(\mathbf{x}_{n_y}) + \sum_{n_z=1}^{N_z} \hat{w}_{Z;l,n_z} Z(\mathbf{x}_{n_z}) \quad (2.51)$$

This concludes the section on Co-Kriging. As apparant from the Equation 2.48, Co-Kriging involves much more numerical computations than Ordinary Kriging. The advantage of using Co-Kriging over Ordinary Kriging is discussed in Chapter 3. In the next Section, we discuss the formulation of Spatio-temporal Kriging.

2.4 Spatio-temporal Kriging

Spatio-temporal Kriging is an extension of Kriging that uses data from multiple time instants instead of just one to generate spatio-temporal estimations. Value at the point of interest is assumed to be a linear combination of value at the same point in different times and the values on neighboring points at multiple time stamps around the time of interest. Since using data for multiple time stamps significantly increases the processing involved, we limit the number of surrounding points included in such calculations by drawing a neighborhood limit. There are some scenarios that we investigate using this technique;

- Smoothing:
 1. High resolution maps inside the boundary of spatial and temporal extent of data used for training
 2. High resolution maps outside the boundary of spatial but inside temporal extent of data used for training
- Prediction:

1. Using past data to generate temporal maps for future inside spatial boundary of data
2. Using past data to generate temporal maps for future outside spatial boundary of data

In all these cases the general assumption holds i.e., value at any point in time and space is a linear combination of values at neighboring points in time and space. The estimate at grid location $\mathbf{x}_l$, and time of interest t_m is given by:

$$\hat{Z}(\mathbf{x}_l, t_m) = \sum_{n_z=1}^{N_z} \sum_{n_t=1}^{N_t} w_{l;m;n_z;n_t} Z(\mathbf{x}_{n_z}, t_{n_t}) , \quad (2.52)$$

where N_t is total number of time stamps used for estimation/ prediction, such that,

$$1 \leq n_t \leq N_t ,$$

and $w_{l;m;n_z;n_t}$ is the weight associated with n_z^{th} observation at time t_{n_t} for estimation at grid location $\mathbf{x}_l$ and time t_m . For unbiasedness,

$$E\{\hat{Z}(\mathbf{x}_l, t_m)\} = E\left\{\sum_{n_z=1}^{N_z} \sum_{n_t=1}^{N_t} w_{l;m;n_z;n_t} Z(\mathbf{x}_{n_z}, t_{n_t})\right\} ,$$

which employs,

$$\sum_{n_z=1}^{N_z} \sum_{n_t=1}^{N_t} w_{l;m;n_z;n_t} = 1 .$$

which results in the same system of linear equations as that of Ordinary Kriging in Equation 2.25. Notice the difference that instead of $(N_a + 1) \times (N_a + 1)$ we now have $(N_z N_t + 1) \times (N_z N_t + 1)$ variogram matrix because each time stamp for each location acts as another data point. Another difference from Ordinary Kriging lies in the way distance between pairs of data points is calculated. Since each point has time as third coordinate, all distances are now two columns where the first column represents the spatial distance between the points and the second column represents the temporal distance. The temporal distance between any two time instants t_{n_t} and t_{n_q} is calculated as;

$$h_t = |t_{n_t} - t_{n_q}| , \quad (2.53)$$

$$1 \leq n_q \leq N_t .$$

Next, we calculate experimental variograms and fit theoretical models.

2.4.1 Spatio-temporal Variogram

Just like for the Ordinary Kriging, we have to calculate experimental variograms and fit a theoretical model to solve for weights and calculate values at points of interest. The spatio-temporal variogram at spatial lag h_s and temporal lag h_t is calculated as;

$$2\gamma_{ZZ}(h, h_t) = \text{var}\{Z(\mathbf{x}, t) - Z(\mathbf{x} + h, t + h_t)\} . \quad (2.54)$$

After calculating experimental variogram values, all that's left is fitting an appropriate spatio-temporal model to it through some optimization routine to find the set of variogram model parameters involved. A few spatio-temporal models are investigated by Skøien and Blöschl in [56]. We first try their exponential model because some physical interpretation of the parameters involved is possible and also because the number of parameters is neither too small nor too large for modeling spatio-temporal data. This model, defined by set of parameters $\mathbf{P}_{st1} = [a \ b \ c \ d \ a_s \ b_s \ a_t \ b_t]^T$ is given as:

$$\gamma^{st1}(\mathbf{P}_{st1}, h, h_t) = a(1 - \exp(-((ch_t + h)/d)^b) + a_s h^{b_s} + a_t h_t^{b_t} . \quad (2.55)$$

This model takes care of the units of spatial and temporal distances for calculations. Since spatial distances are in units of kilometers and temporal distances are in units of hours, the parameter c has units kilometers per hour and this addition is correct. The parameter d represents the combined correlation length and is given in kilometers to make the over all term unit less. Looking closely, the first term in this model resembles the spatial exponential model for variogram and parameter a is analogous to variance term (also called sill). The second and third terms in the model represent spatial and temporal non-stationarity.

To find this set of parameters $\mathbf{P}_{st1}$, numerical optimization routines can be used to fit variogram model to experimental variogram. We also use 16 fold validation to select the best set of parameters over validation data and this significantly improves performance over test data.

Using the model in Equation 2.55 proves to be useful for interpolations within the period of one day because there is no significant temporal periodicity in data during such a small period and a monotonically increasing spatio-temporal variogram model is sufficient to model the trends in the data but beyond one day such model is not very useful. For modeling temporal periodicity in the data, we need to look at temporal variations as some form of sinusoids. A spatio-temporal model that incorporates temporal periodicity is proposed by Bilonick in [57]. It has $\mathbf{P}_{st2} = [C_0 \ C_P \ C_S \ a_s \ k_L]^T$ as model parameters. This model is better suited to ionospheric data and can be written as,

$$\gamma^{st2}(\mathbf{P}_{st2}, h, h_t) = \begin{cases} C_0 + 0.5C_P + k_L h, & \text{if } h_t = 0. \\ C_0 + C_P\{1 - 0.5 \cos(2\pi h_t)\} + k_L h \\ + C_S\{3h_t(2a_s)^{-1} - 2^{-1}h_t^3 a_s^{-3}\}, & \text{if } 0 < h_t < a_s. \\ C_0 + C_P\{1 - 0.5 \cos(2\pi h_t)\} + k_L h + C_S, & \text{if } h_t \geq a_s. \end{cases} \quad (2.56)$$

This is one of the earliest models used in spatio-temporal modeling and authors used it over Sulphate deposition data for New York State in 1980-1981 [57]. Some physical interpretation of parameters is also possible here. The Nugget effect is modeled by C_0 . $C_P\{1 - 0.5 \cos(2\pi h_t)\}$ models the temporal periodicity in the data using cosine function. The term $C_S\{3h_t(2a_s)^{-1} - 2^{-1}h_t^3 a_s^{-3}\}$ is used to model the underlying aperiodic trend in the data using a spherical variogram model. In the Bilonick (1985) paper, the trend is modeled as linear based on the appearance of the experimental variogram. Depending on the empirical variogram for various datasets, some alterations in the model can be made to accurately model the trends in the data.

In this chapter, we have discussed the mathematical formulation for Kriging, Co-Kriging and Spatio-temporal Kriging as well as mathematical definitions of the variogram models to be used in simulations. We followed the basic derivations

provided in the literature before and retraced some of their steps to provide complete derivations for the reader. In the following Chapter, we will provide the simulations over ionospheric data for the methods discussed in this Chapter.



Chapter 3

Simulations of the Proposed Techniques

In this chapter, we apply the formulation discussed in Chapter 2 on ionospheric data and obtain resulting maps. First, we select a region of interest for the simulations. Section 3.1 discusses the regions of interest for both spatial and spatio-temporal techniques and provides tables of station locations used in simulations. In Section 3.2, we describe the correlation between the modalities TEC-foF2 and TEC-hmF2 to emphasize the motivation for Co-Kriging. In Section 3.3, we provide the maps and error maps obtained by using Kriging, and in Section 3.3 we provide the corresponding maps and error maps obtained by using Co-Kriging. Section 3.5 describes the scenarios explored in Spatio-temporal analysis and provides the maps obtained by using Spatio-temporal Kriging. We conclude the chapter by discussing the scenarios for which Spatio-temporal Kriging is particularly useful.

3.1 Region of Interest for Mapping Techniques

We chose the region R_E from $\theta_i = 20^\circ$ N to $\theta_f = 80^\circ$ N and $\phi_i = -20^\circ$ E to $\phi_f = 60^\circ$ E and generate maps using the techniques discussed. This region corresponds to Europe on the map, covering about 15.9 million km^2 area, and is shown in Figure 3.1. Kriging and Co-Kriging maps are generated for this region and Root Mean Square Error (RMSE) is computed over left out stations in the region. A lexicographical grid l_s for all test points, and all time instants used for testing is defined as,

$$l_s = n_\beta + (n_\alpha - 1)N_\beta , \quad (3.1)$$

such that,

$$1 \leq n_\alpha \leq N_\alpha ,$$

$$1 \leq n_\beta \leq N_\beta ,$$

where N_α is the number of stations left out from the data for testing purpose and N_β is total number of time instants of test data. *Error* at l_s is computed as,

$$\text{Error}(l_s) = \hat{Z}(l_s) - Z(l_s) , \quad (3.2)$$

and the density of *Error* can be calculated as,

$$P_E(\text{Error}) = \sum_{l_s=1}^{N_\alpha N_\beta} R(\text{Error} - \text{Error}_{l_s; \sigma}) , \quad (3.3)$$

where $R(\cdot)$ is the Kernel used for density estimation and μ_{Error} and σ are the mean and standard deviation of Error. We use Gaussian Kernel defined as,

$$R(\text{Error}, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{(\text{Error} - \mu_{\text{Error}})^2}{\sigma^2} \right\} \quad (3.4)$$

where,

$$\mu_{\text{Error}} = \frac{1}{N_\alpha N_\beta} \sum_{l_s=1}^{N_\alpha N_\beta} \text{Error}(l_s) , \quad (3.5)$$

and

$$\sigma = \sqrt{\frac{\sum_{l_s=1}^{N_\alpha N_\beta} \text{Error}(l_s) - \mu_{\text{Error}}}{N_\alpha N_\beta}} . \quad (3.6)$$

Error at test station α at time t_{n_β} is computed as,

$$Error(\mathbf{x}_{n_\alpha}, t_{n_\beta}) = \hat{Z}(\mathbf{x}_{n_\alpha}, t_{n_\beta}) - Z(\mathbf{x}_{n_\alpha}, t_{n_\beta}) . \quad (3.7)$$

RMSE at a particular station α_o , for all time instants used, is computed as,

$$RMSE(\mathbf{x}_{n_{\alpha_o}}, t_{n_\beta}) = \sqrt{\frac{\sum_{t_{n_\beta}=1}^{N_\beta} (Error(\mathbf{x}_{n_{\alpha_o}}, t_{n_\beta}))^2}{N_\beta}} . \quad (3.8)$$

RMSE at a particular time t_{β_o} for all the stations is computed as,

$$RMSE(\mathbf{x}_{n_\alpha}, t_{n_{\beta_o}}) = \sqrt{\frac{\sum_{t_{n_\alpha}=1}^{N_\alpha} (Error(\mathbf{x}_{n_\alpha}, t_{n_{\beta_o}}))^2}{N_\alpha}} . \quad (3.9)$$

In case of IONOSONDE data (D_{snd}), and IZMIRAN data (D_{iz}), we only have data from seven and 17 stations respectively, and out of that we want to use some for training and some for testing. We use all seven stations in D_{snd} for generating foF2 maps and for performance analysis we use D_{iz} and IRI-Plas Data (D_{iri}). Still in case of D_{iz} , using such a small data forces us to use multiple divisions of training and test data to be able generate statistically significant conclusions. Inferring a conclusion from using one random division of points in training and test data will be insufficient for drawing a general conclusion so for the sake of generalization we use the data for 31 days instead of just one day. To further increase the reliability of the results, we generate 50 random divisions of training and test points over D_{iz} and use the same 50 divisions over the period of 31 days at 12 UT to calculate the final RMSE plots. Figures 3.1, and 3.2 show the station locations for GPS stations and ionosonde stations respectively.

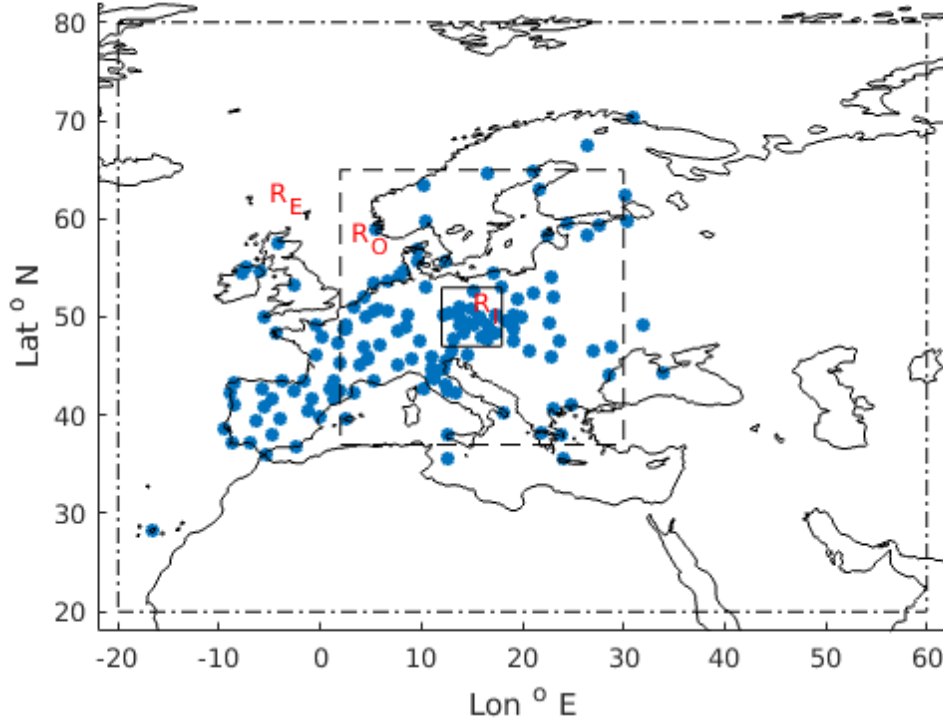


Figure 3.1: GPS stations over Europe (filled circles), R_E is the region bounded by dash dot box , R_I is the region bounded by the solid box, and R_O is the region between the dashed box and the solid box.

The entire region R_E is used for Kriging and Co-Kriging. For Spatio-temporal Kriging R_I is used for training and predictions are made both in R_I and R_O . R_I is defined as the region from $\theta_i = 12^\circ\text{N}$ to $\theta_f = 18^\circ\text{N}$ and $\phi_i = 47^\circ\text{E}$ to $\phi_f = 53^\circ\text{E}$, and it covers 286 thousand km^2 area over Central Europe. On the temporal axis, the region of interest is bounded by time t_i and t_f for training data such that,

$$t_i \leq t_{n_i} \leq t_f \quad (3.10)$$

and superscripts I and O are used with the regions to indicate whether a test point lies inside or outside the temporal extent of training data. Table 3.1 explains the notation used in Spatio-temporal Kriging for different regions.

Table 3.1: Regions based on their spatial and temporal location

	Temporally Inside	Temporally Outside
Spatially Inside	R_I^I	R_I^O
Spatially Outside	R_O^I	R_O^O

Table 3.2 lists the boundaries in terms of coordinates for these regions.

Table 3.2: Boundaries for Region of Interest.

Region	θ_i^o N	θ_f^o N	ϕ_i^o E	ϕ_f^o E	t (UT)
R_E^I	20	80	-20	60	12
R_I^I	47	53	12	18	12
R_E^O	20	80	-20	60	17
R_I^O	47	53	12	18	17

The region R_O is defined as, $R_O = R_E - R_I$ and covers the region between latitude 20°N to 47°N and 53°N to 80°N, and longitude -20°E to 12°E and 18°E to 60°E. We chose this region because we have denser observations available in this region compared to other parts of Europe and we have enough data available to leave some stations out for validation and some for testing purposes and provide a measure of performance. Test stations from all the regions are used for performance analysis.

Table C lists the GPS station numbers and station locations for $D_{ion-TEC}$. This data is used for Spatio-temporal Kriging and the station numbers can be used to identify the temporal predictions provided in Section 3.5.

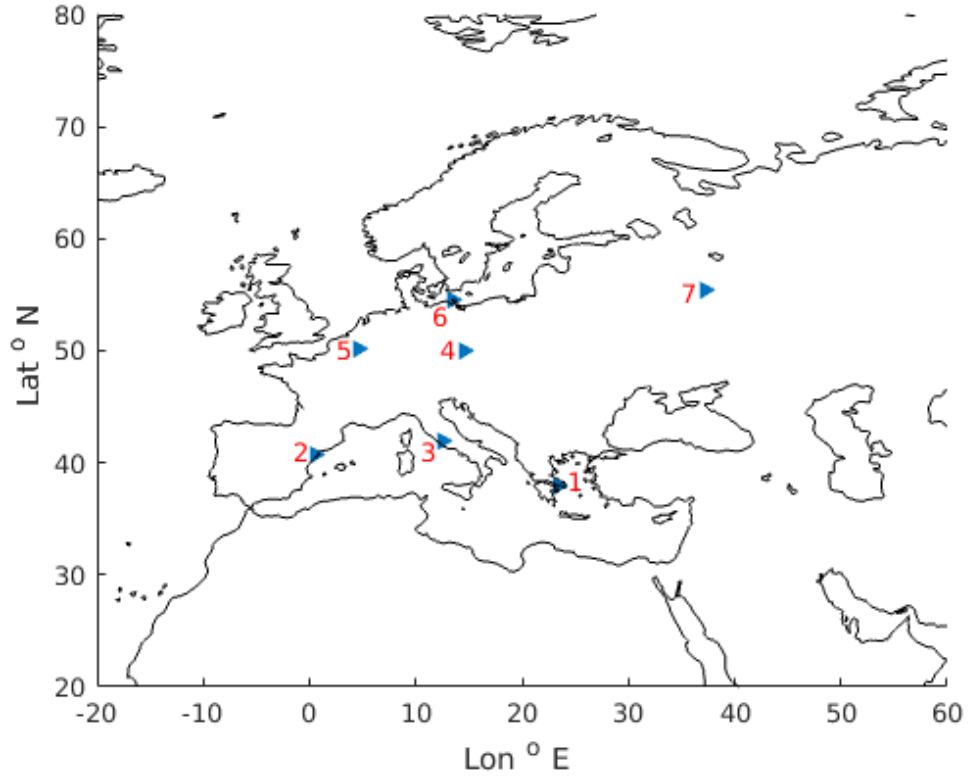


Figure 3.2: Ionosonde stations over Europe indicated by right triangles, station numbers correspond to stations given in Table 3.3.

Table 3.3 lists the station numbers and station locations for ionosonde stations over Europe. The same station numbers are used for D_{snd} throughout this work and can be seen on the map in Figure 3.2.

Table 3.3: Ionosode Station Numbers and locations for D_{snd} .

Station Number	Station Code	Lat ° N	Lon ° E	Location
1	AT138	38	23.5	Athens
2	EB040	40.8	0.5	Roquetes
3	RO041	41.9	12.5	Rome
4	PQ052	50	14.6	Pruhonic
5	DB049	50.1	4.6	Dourbes
6	JR055	54.6	13.4	Juliusruh
7	MO155	55.47	37.3	Moscow

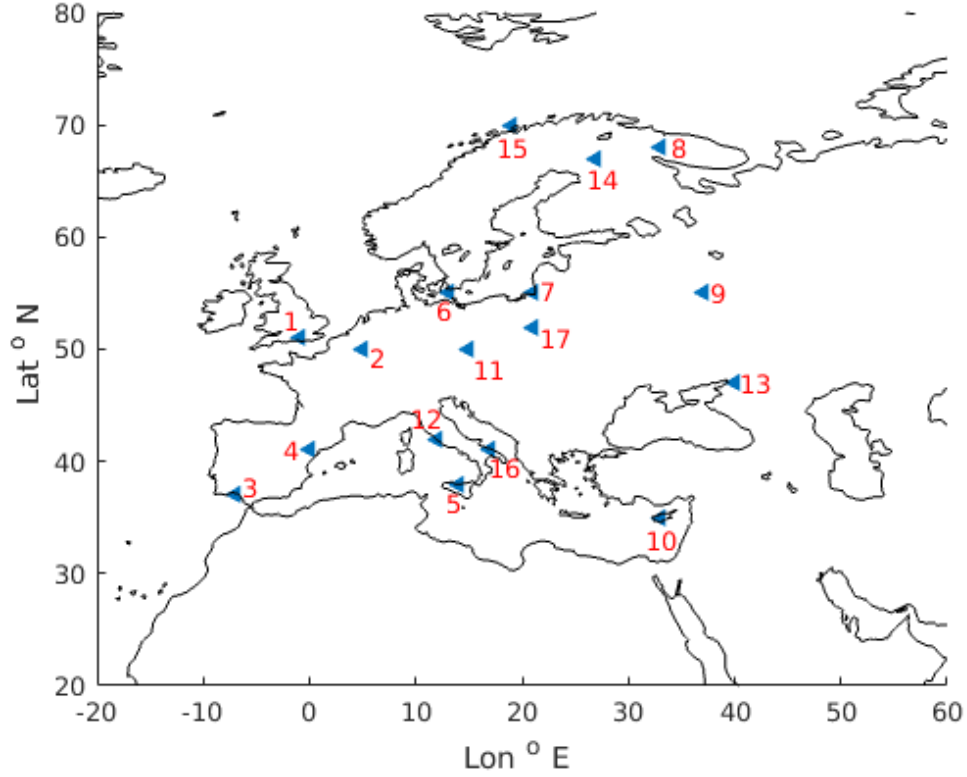


Figure 3.3: GPS stations over Europe (indicated by left triangles) where D_{iz} for foF2 is available, station numbers correspond to station numbers given in Table 3.4.

Table 3.4 lists the station numbers and station locations for D_{iz} . The same

station numbers are used for D_{iz} throughout this work and can be seen on the map in Figure 3.3.

Table 3.4: Station Numbers and locations for D_{iz}

Station Number	Lat ° N	Lon ° E	Station Name	Country
1	51	-1	Chilton	UK
2	50	5	Dourbes	Belgium
3	37	-7	Arenosillo	Spain
4	41	0	Tortosa	Spain
5	38	14	Gibilmanna	Italy
6	55	13	Juliusruh	Germany
7	55	21	Kaliningrad	Russia
8	68	33	Murmansk	Russia
9	55	37	Moscow	Russia
10	35	33	Nicosia	Cyprus
11	50	15	Pruhonice	Czech Rep.
12	42	12	Rome	Italy
13	47	40	Rostov	Russia
14	67	27	Sodankyla	Finland
15	70	19	Tromso	Norway
16	41	17	San Vito	Italy
17	52	21	Warsaw	Poland

Figure 3.4 shows the locations for primary and secondary modality as well as the test points selected from IRI-Plas data D_{iri} . For Kriging, we only use primary data and test it on test data. For Co-Kriging, we use both primary data $D_{iri-foF2}$ and secondary data $D_{iri-TEC}$.

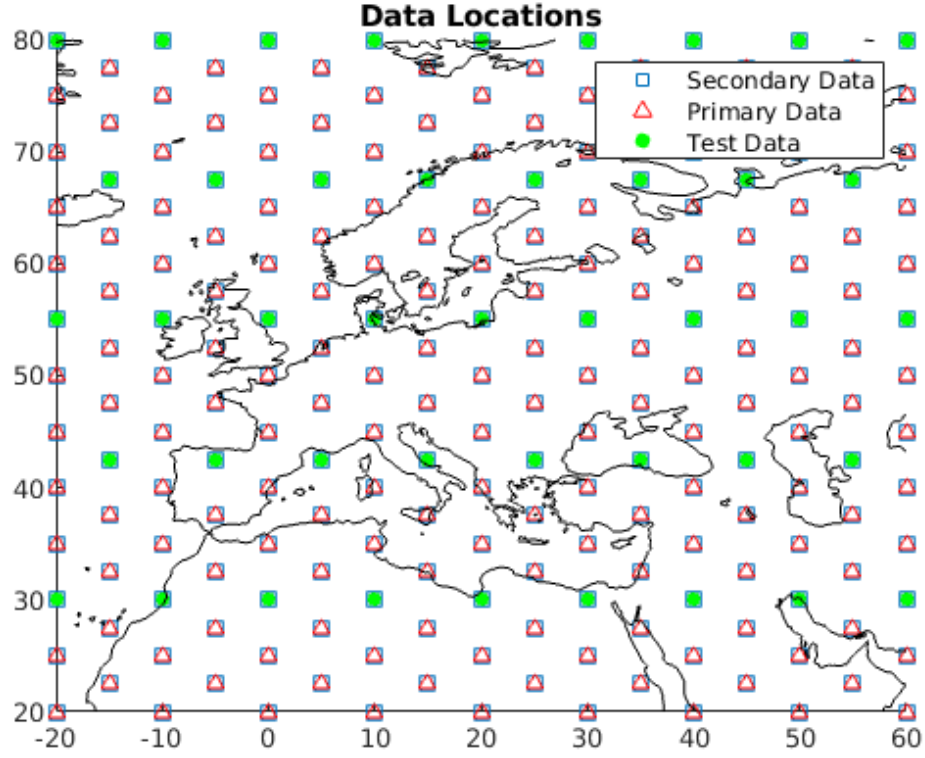


Figure 3.4: Selected locations for primary and secondary modalities in D_{iri} , triangles indicate the primary data $D_{iri-foF2}$, squares indicate the secondary data $D_{iri-TEC}$, and the filled circles indicate the locations selected for testing

D_{iri} enables us to leave sufficient data for as test sets and generalize performance of algorithms so we provide *Error* maps defined is Equation 3.2 for Kriging as well as for Co-Kriging in their respective Sections. In the next Section, we discuss some correlation trends in the data.

3.2 Visualizing Correlations between Ionospheric Modalities

We observe periodically correlated data so we chose a period of 24 hours and see the relationship over a month. The first plot shows the actual data values for foF2, TEC and hmF2 for D_{iz} for one of the days used in this work. The plot is for the 17 locations covered in D_{iz} over Europe. The second plot shows the correlation coefficient ρ at 12 UT from October 01, 2016 to October 31, 2016.

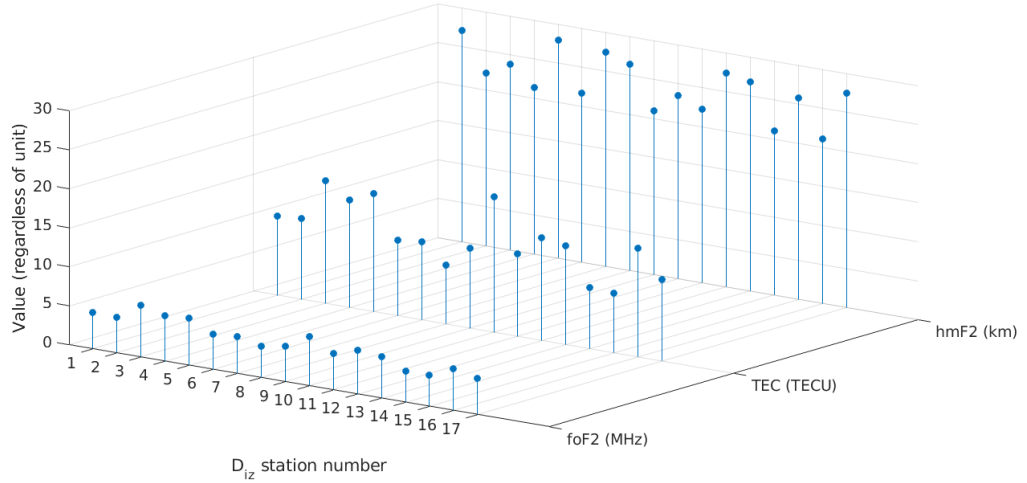


Figure 3.5: $D_{iz-foF2}$, D_{iz-TEC} , and $D_{iz-hmF2}$ on October 01, 2016 from station 1 to station 17 at 12 UT, station numbers correspond to stations in Table 3.4.

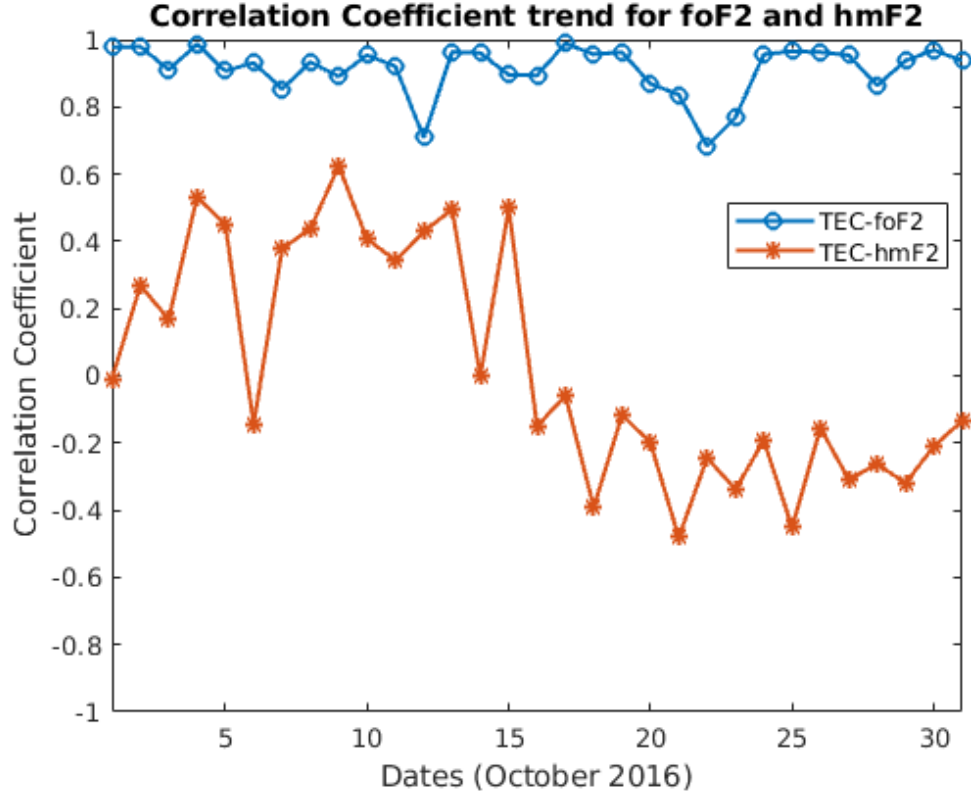


Figure 3.6: Correlation of TEC with foF2 and hmF2 over the month of October 2016, at 12 UT, showing a quiet month with two geomagnetically disturbed days (October 12 and October 23).

In this Section, we presented the correlation trends between ionospheric modalities. In the following Sections, we provide the maps and *Error* plots generated using Kriging and Co-Kriging for a comparative analysis.

3.3 Kriging

Kriging is performed over the European region R_E , over multiple data sets for multiple days. Depending on the number of observed values for that particular data set, one or more stations are left out for validation and test performance evaluation.

Kriging on D_{iz} is applied for a month (Oct. 2016) and one of the maps is shown in the Figure 3.7.

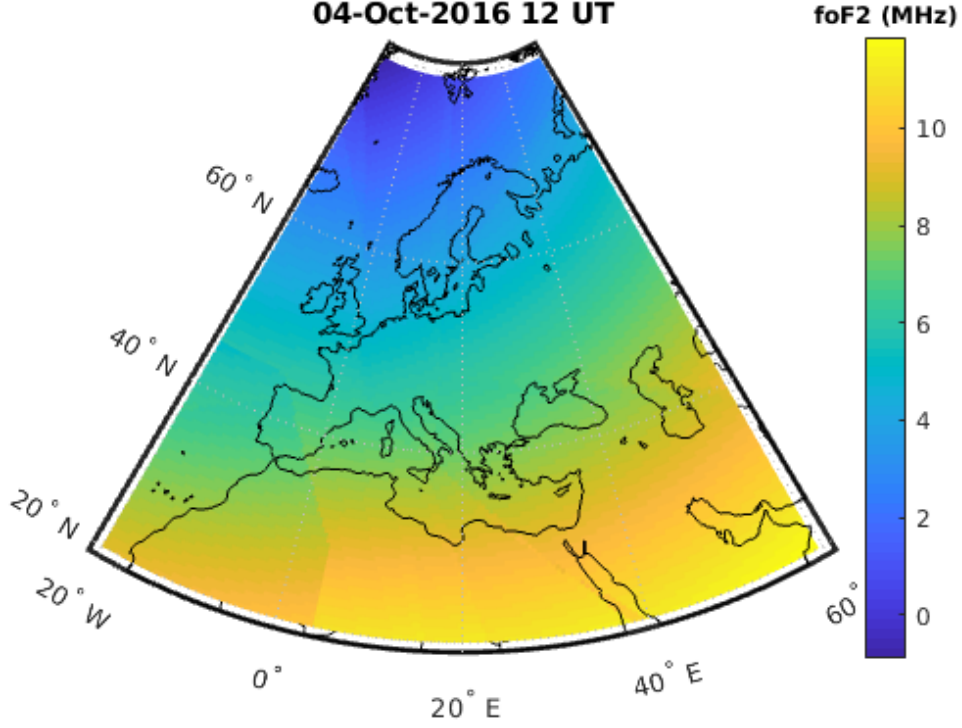


Figure 3.7: foF2 map generated by IONOLAB-MAP, using Universal Kriging over D_{iz} in region R_E^I , for October 04, 2016 at 12 UT.

Kriging on D_{iri} is applied for 15 days in Section 1.5, including some quiet and storm days as mentioned earlier and one of the maps is shown in the Figure 3.8. This is only to demonstrate a comparative analysis over large dataset with a number of test points selected at the same time. We use the region over Europe and strategically place test points to cover all types of regions e.g. close to center or close to boundary of data set.

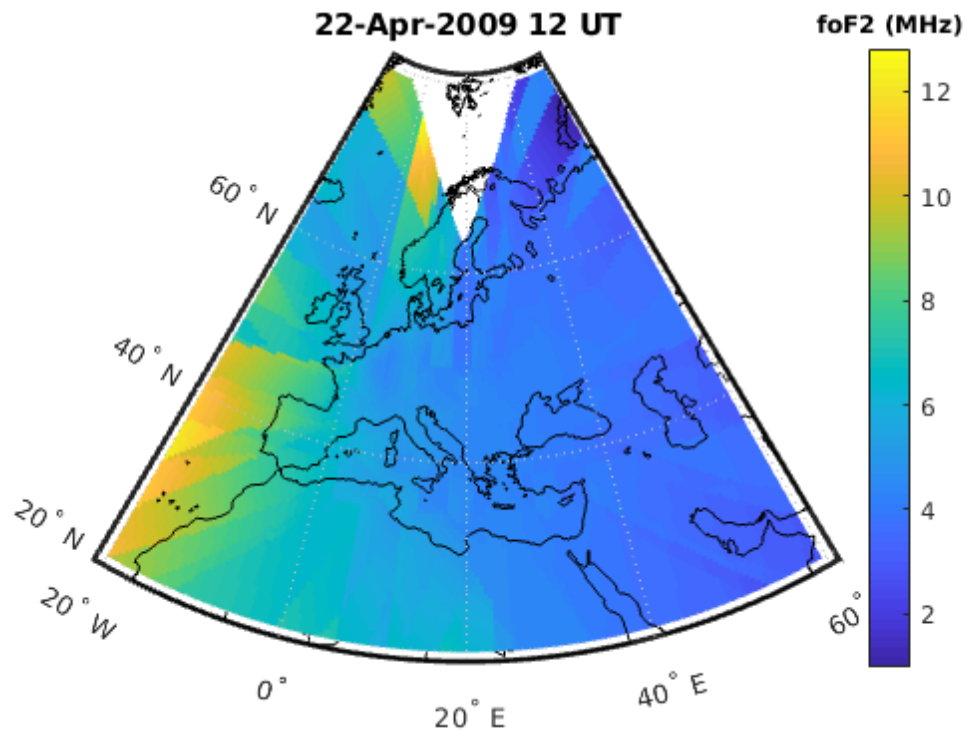


Figure 3.8: foF2 map generated by IONOLAB-MAP, using Universal Kriging over D_{iri} in region R_E^I , for April 22, 2009 at 12 UT.

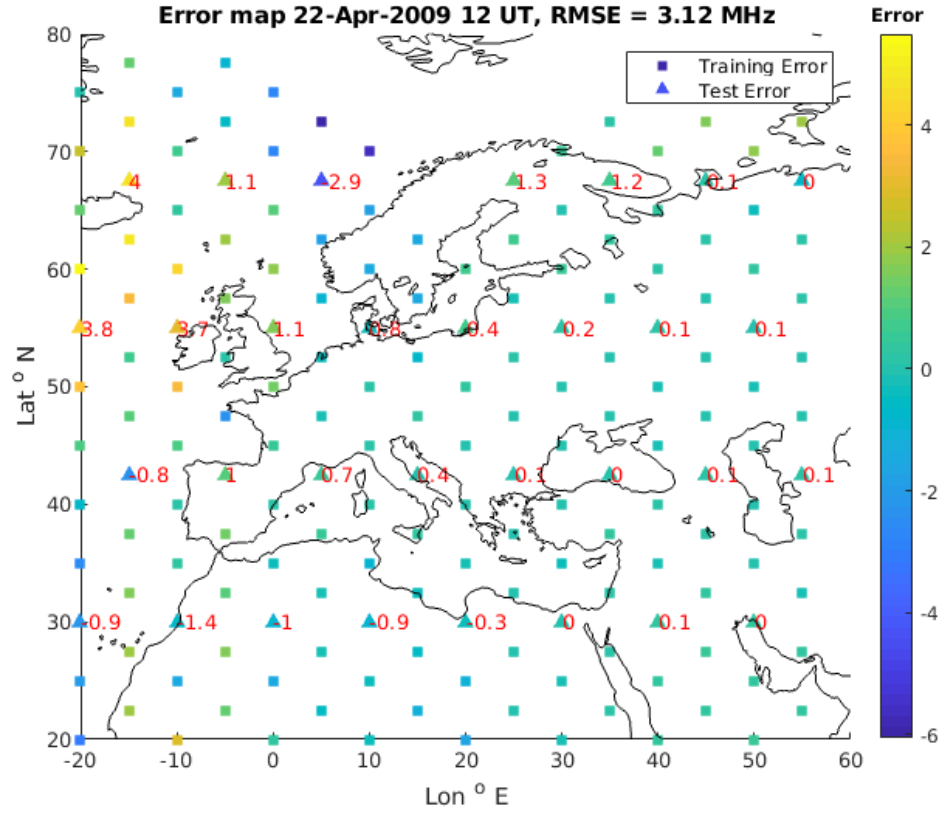


Figure 3.9: Error map for IONOLAB-MAP, over D_{iri} in region R_E^I , for April 22, 2009 at 12 UT.

In the next Section, we discuss the maps generated over multiple data sets using Co-Kriging.

3.4 Co-Kriging

Co-Kriging of foF2 with TEC is performed over multiple days to demonstrate the algorithm. As mentioned earlier, the region of interest for this part spans over Europe. D_{snd} for foF2 is available for just seven stations over Europe, and in case of D_{iz} , an interpolation of it is available at 17 places in Europe. Other than these two datasets, we also perform the algorithm over D_{iri} to evaluate performance

for multiple stations and make generalizations.

We have foF2 as our primary data and it is available at 17 locations in Europe in case of $D_{iz-foF2}$. The secondary data used for this part is TEC data prepared by IONOLAB D_{ionTEC} , available for 143 stations in the region of interest selected as shown earlier in Figures 3.3 and 3.1. Co-Kriging on D_{iz} is applied for a month (October, 2016) and one of the maps is shown in the Figure 3.10.

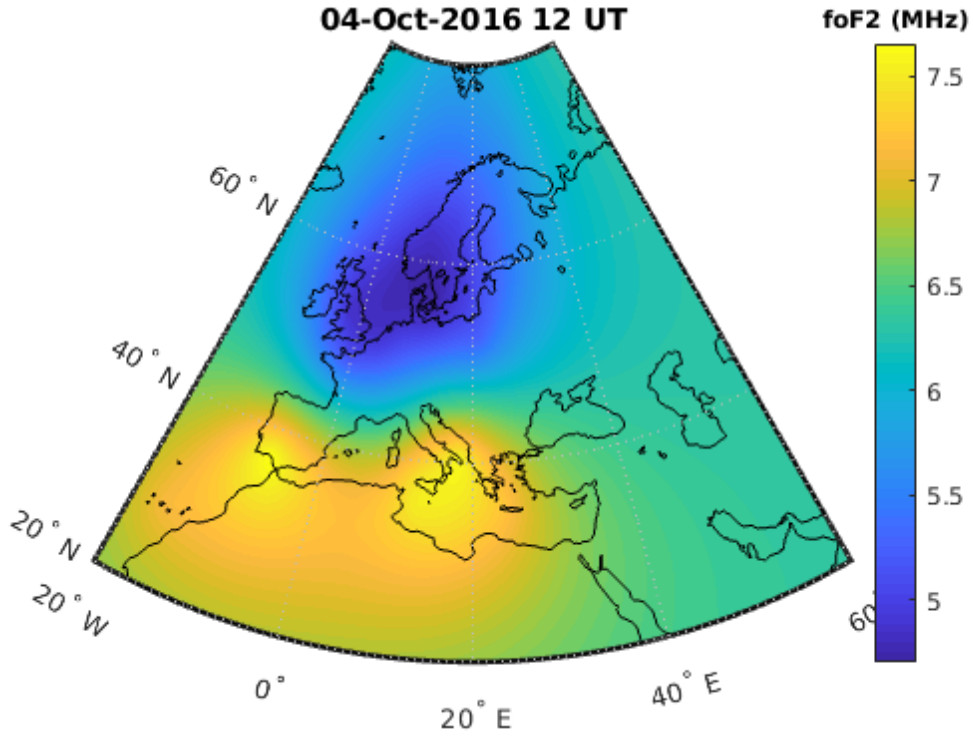


Figure 3.10: foF2 map generated by Co-Kriging, over D_{iz} in region R_E^I , for March 12, 2010 at 12 UT.

Co-Kriging is also applied on D_{iri} the same quiet and stormy 15 days and one of the maps is shown in the Figure 3.11. The train and test points are selected as we did for Kriging. Figure 3.4 shows the selected locations for primary and secondary modality as well as the test points selected. Notice that test points never overlap with primary modality but they can overlap with the secondary modality. This simulates a somewhat realistic scenario where you want to estimate the value of a

modality at a location where data for another correlated modality might already be available. Co-Kriging foF2 map can be seen in Figure 3.11 and the error map for this day for training and test data is shown in Figure 3.12.

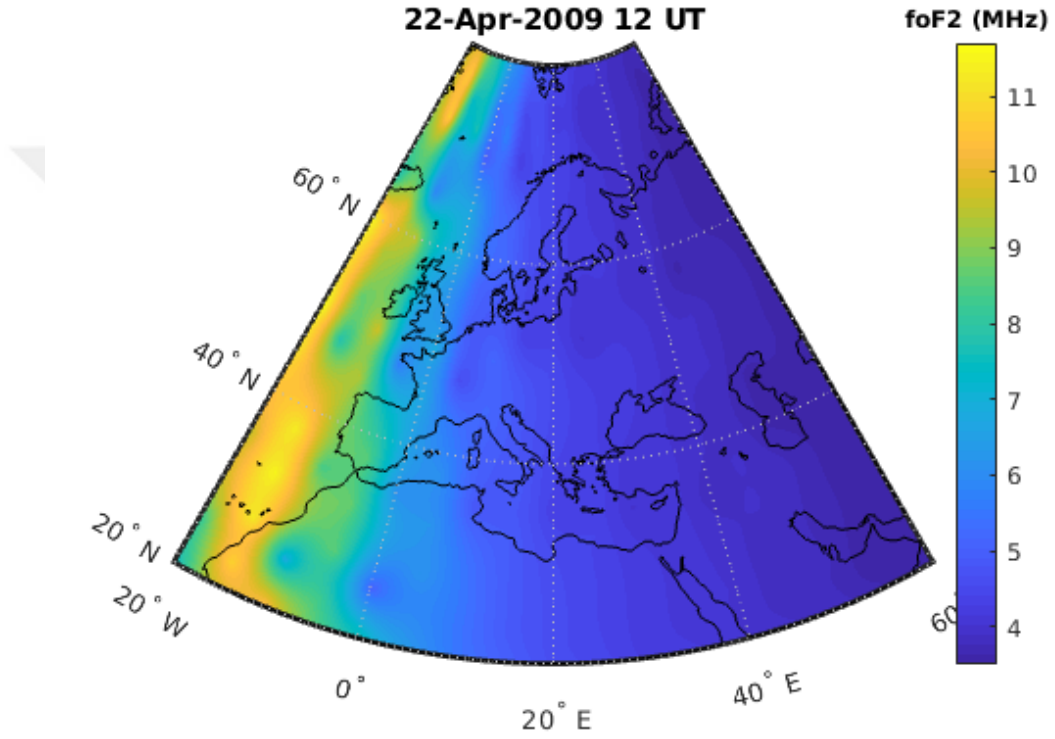


Figure 3.11: foF2 map generated by Co-Kriging, over D_{iri} in region R_E^I , for April 22, 2009 at 12 UT.

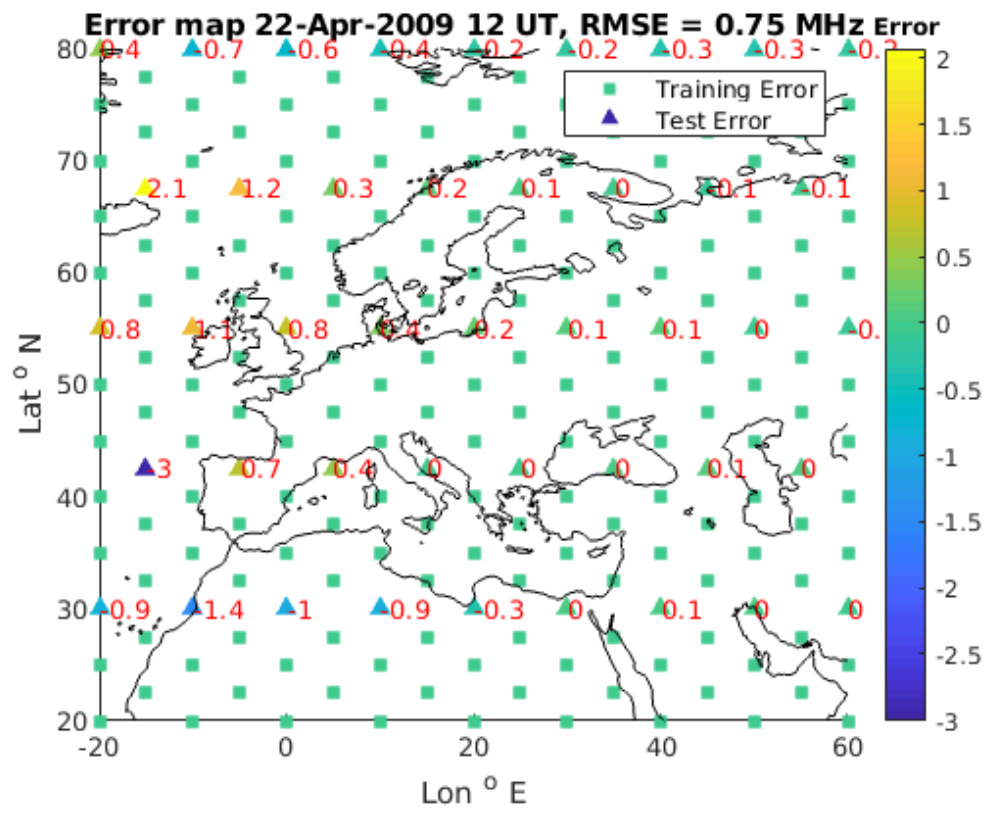


Figure 3.12: Error map for Co-Kriging, over D_{iri} in region R_E^I , for April 22, 2009 at 12 UT.

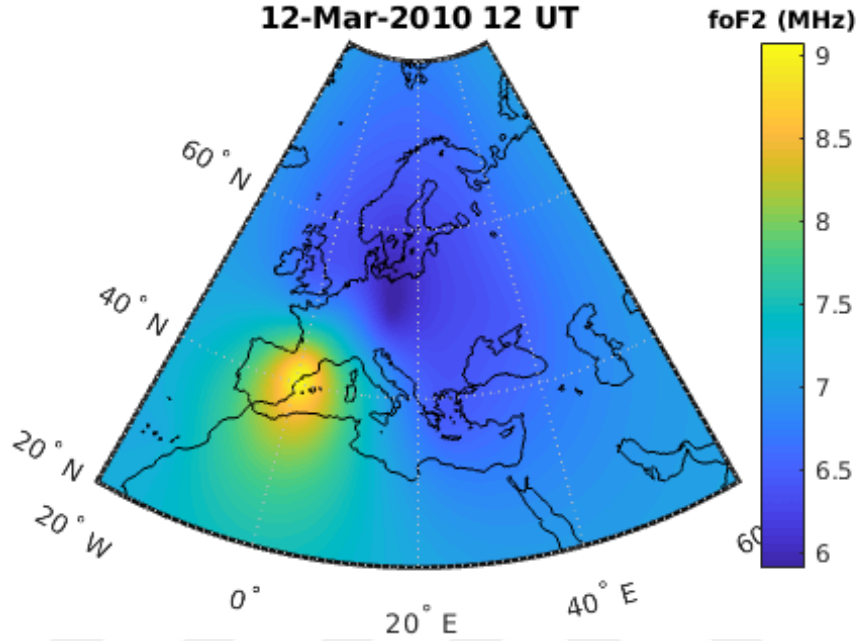


Figure 3.13: foF2 map generated by Co-Kriging, over D_{snd} in region R_E^I , for March 12, 2010 at 12 UT.

In case of D_{snd} , the locations where this data is available were provided previously in Figure 3.2. The map for foF2 generated by Co-Kriging using this data can be seen in Figure 3.13. All the available stations were used for Co-Kriging since only seven stations are available, and error measures are not computed for this data. To compensate for lack of error measures, the Co-Kriging algorithm is applied over D_{iz} and D_{iri} where we can afford leaving some data out for error measures and algorithm performance evaluation.

In the next Section, we discuss the spatial and temporal plots generated by Spatio-temporal Kriging.

3.5 Spatio-temporal Kriging

In this part, we use $D_{ion-TEC}$ for all the simulations. Spatio-temporal Kriging generates best results in localized settings so generate maps over the region R_I .

Using this region allows us to process more temporal data for each station and up to two days data can be processed at a time with each day having 2880 temporal observations (30 sec time resolution for each station). For the sake of performance evaluation, maps were also generated by using the Spatio-temporal Kriging approach by Skøien and Blöschl presented in [58] and will be referred to as STK-SB whereas the Spatio-temporal Kriging presented in this thesis in Section 2.4 will be referred to as STK.

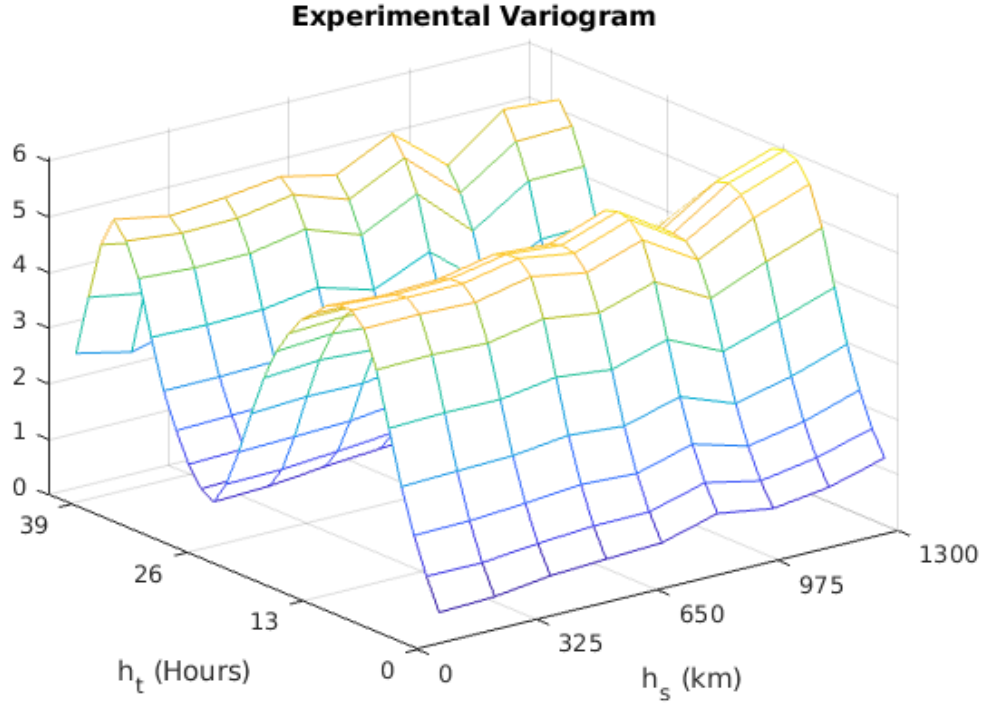


Figure 3.14: Experimental Spatio-temporal Variogram (January 23, 2010 0 UT to January 24 15 UT) $D_{ion-TEC}$ over region R_I^I .

Different variogram models were tried for Spatio-temporal Kriging, and initially MATLAB curve fitting tool was used manually in the experimental phases to view how each model fits to the experimental variogram calculated from the data. In the final implementation, the algorithm uses Particle Swarm Optimization (PSO) to find the model parameters. The experimental variogram calculated from $D_{ion-TEC}$ in the region R_I^I for a period of 37 hours (January 23, 2010, 0 UT to January 24, 15 UT) can be seen in Figure 3.14. Spatio-temporal Kriging

models the periodicity in the data and Figure 3.15 shows the weights associated with station 86 against temporal distance, for prediction at station 74.

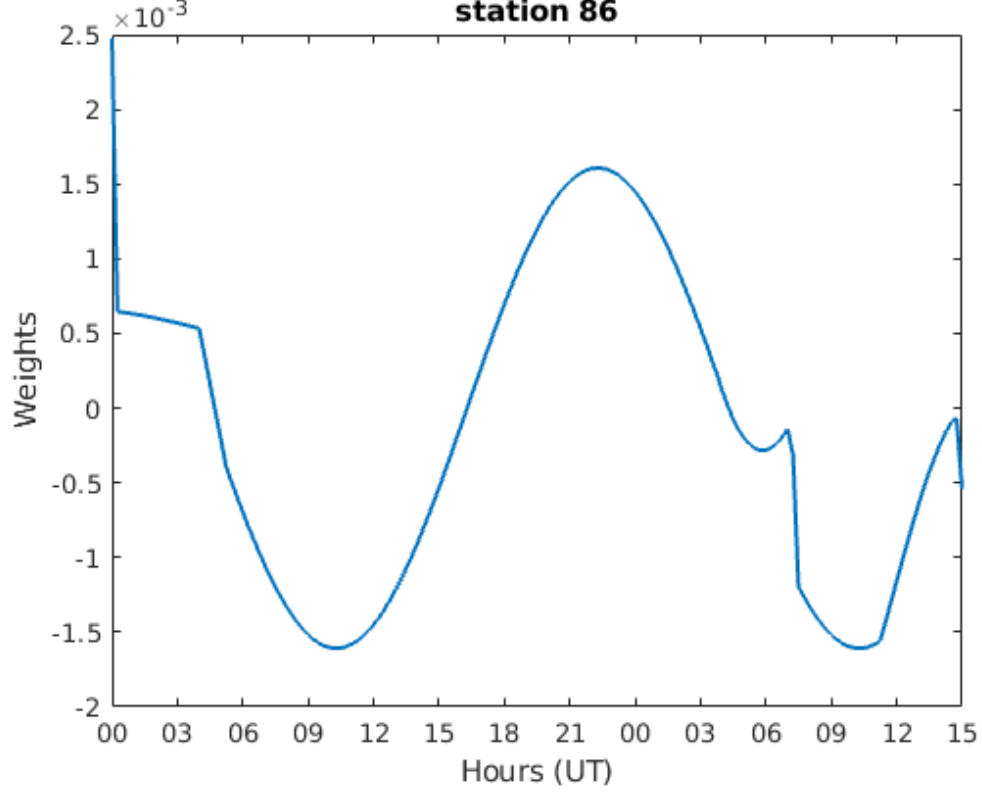


Figure 3.15: Weights associated with station 86 for prediction at station 74, against temporal distance from January 19, 2010 0 UT to January 20, 2010 15 UT.

The algorithm fits theoretical variogram to the experimental variogram during the validation step and the model parameters with least RMSE over validation data are selected automatically by the algorithm. The training data provided to the algorithm always comes from region R_I^I , the maps are generated for R^I and R^O . The temporal plots are generated for R_I^I , R_O^I , R_I^O , and R_O^O .

In the next Subsection, we provide the maps for the region R^I .

3.5.1 Estimation (Maps Inside the Temporal Extent)

The purpose of this technique is generating spatially dense estimates inside the temporal extent of data cube used. To understand that, we first have to understand the concept of data cube. We have data for some locations over the map and we can bound them inside a polygon (a rectangle in our case). Each of these stations has been collecting daily observations for many years so we have temporal data available for stations restricted by this rectangle for some years. Representing latitude and longitude as boundaries of rectangle and time the third dimension we have a data cube, boundaries of which are defined by minimum and maximum of θ , ϕ and t . Such division of regions was presented earlier in Table 3.1.

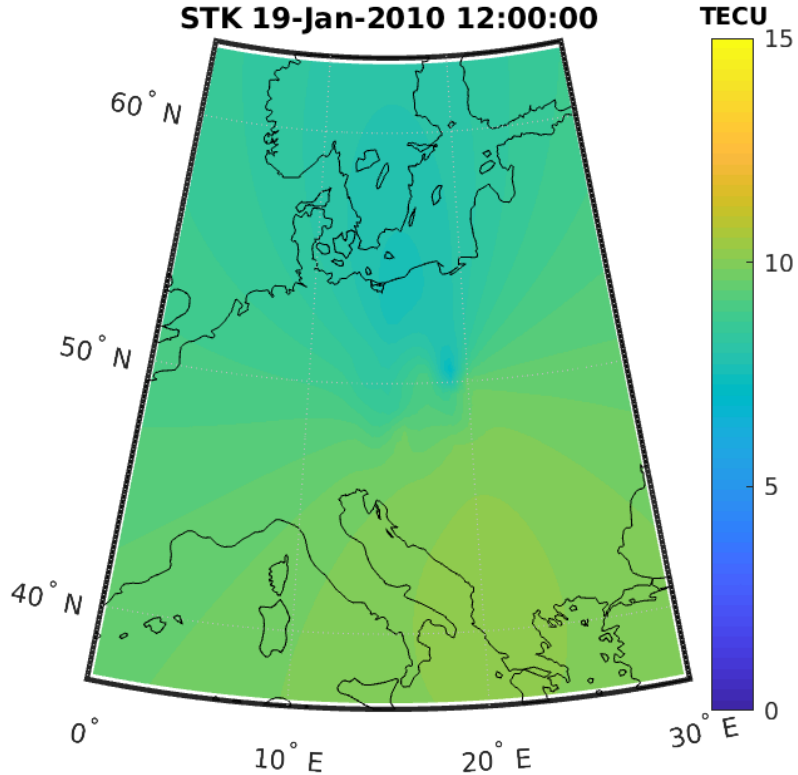


Figure 3.16: TEC map generated by Spatio-temporal Kriging STK, over $D_{ion-TEC}$ in region R_I^I and R_O^I , for January 19, 2010 at 12 UT.

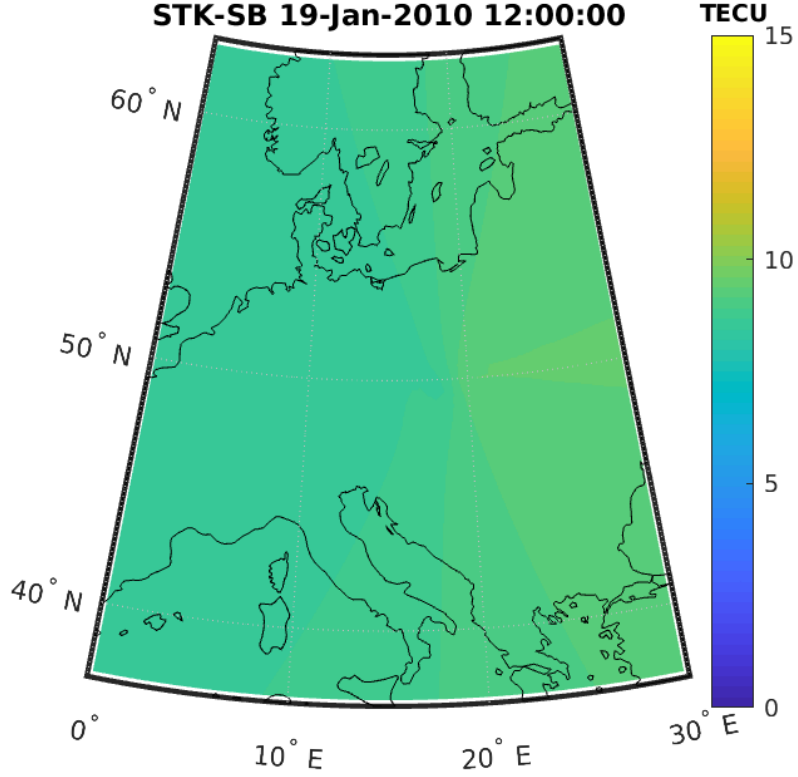


Figure 3.17: TEC map generated by Spatio-temporal Kriging STK-SB, over $D_{ion-TEC}$ in region R_I^I and R_O^I , for January 19, 2010 at 12 UT.

For the sake for comparison, we also generate these plots using the Spatio-temporal Kriging used in [58] (STK-SB). Figure 3.16 shows a map generated by using Spatio-temporal Kriging in Section 3.5, and Figure 3.17 shows a map generated by STK-SB. A comparison of RMSE is presented in Table 4.3.

In this Subsection, we presented the maps in the region R^I , In the next Subsection, we provide the maps for the region R^O .

3.5.2 Prediction (Maps Outside the Temporal extent)

Prediction refers to using *past* data to generate values for the *future*. Here, past and future represents the temporal axis of data and in this section we use data

for a specific number of days/hours, and then predict the modality on the hours after the last time stamp used in the training data. The objective of this method is to determine how far ahead in the future the algorithm is able to predict in terms of hours. Generally the algorithm generates predictions within reasonable margin for up to 2 hours.

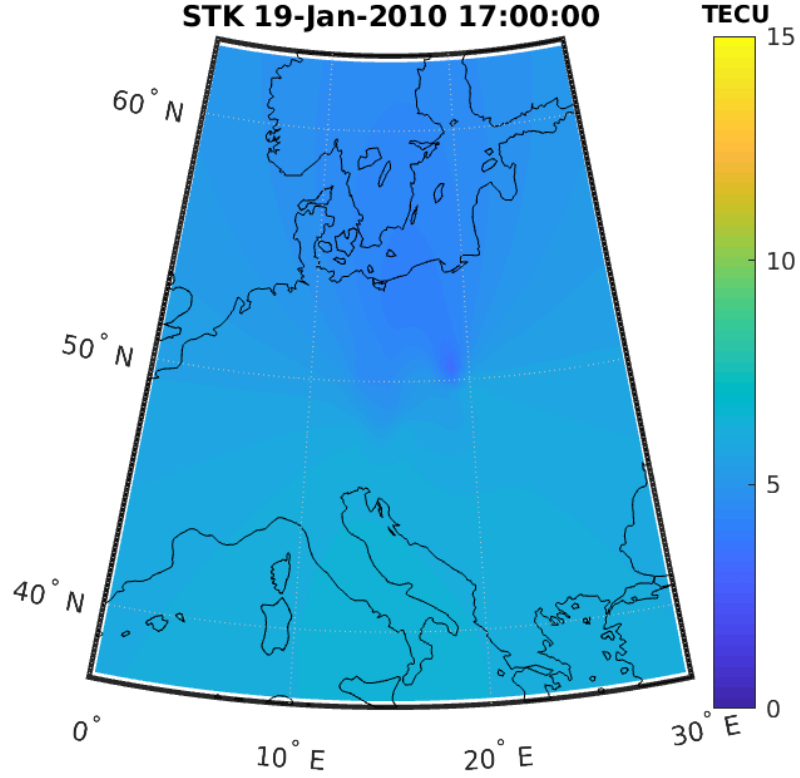


Figure 3.18: TEC map generated by Spatio-temporal Kriging STK, over $D_{ion-TEC}$ in region R_I^O and R_O^O using data only in R_I^I , for January 19, 2010 at 17 UT.

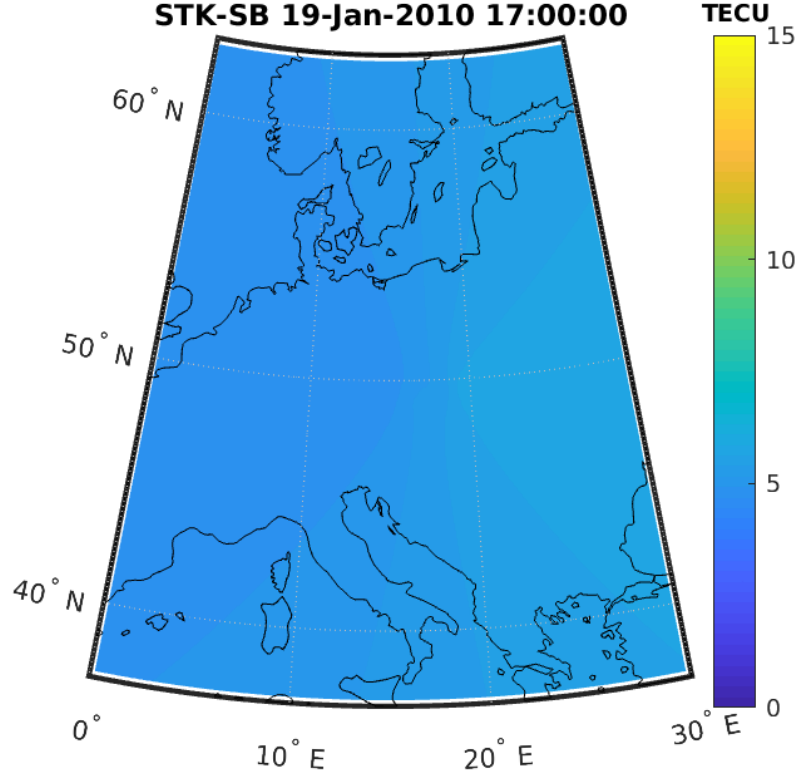


Figure 3.19: TEC map generated by Spatio-temporal Kriging STK-SB, over $D_{ion-TEC}$ in region R_I^O and R_O^O using data only in R_I^I , for January 19, 2010 at 17 UT.

In the next Subsection, we provide the temporal plots for the region R_I .

3.5.3 Temporal Estimates - Inside Spatial Extent

In this part we pick our test stations from region R_I and exclude them from training data set. We use data for 3 months, from January 01, 2010 to March 31, 2010 for the region R_I and leave three stations out for performance analysis. The comparison of RMSE for these cases is given in Table 4.3. At one time, data for upto 40 hours is used for generating spatial and temporal plots. Figure 3.21 shows one such temporal plot for one of the days in this period. For generating

this plot, data for 24 hours from the previous day and data for 15 hours from current day is used. Note that no data for the current day for the station of interest is included in parameter estimation and plot generation.

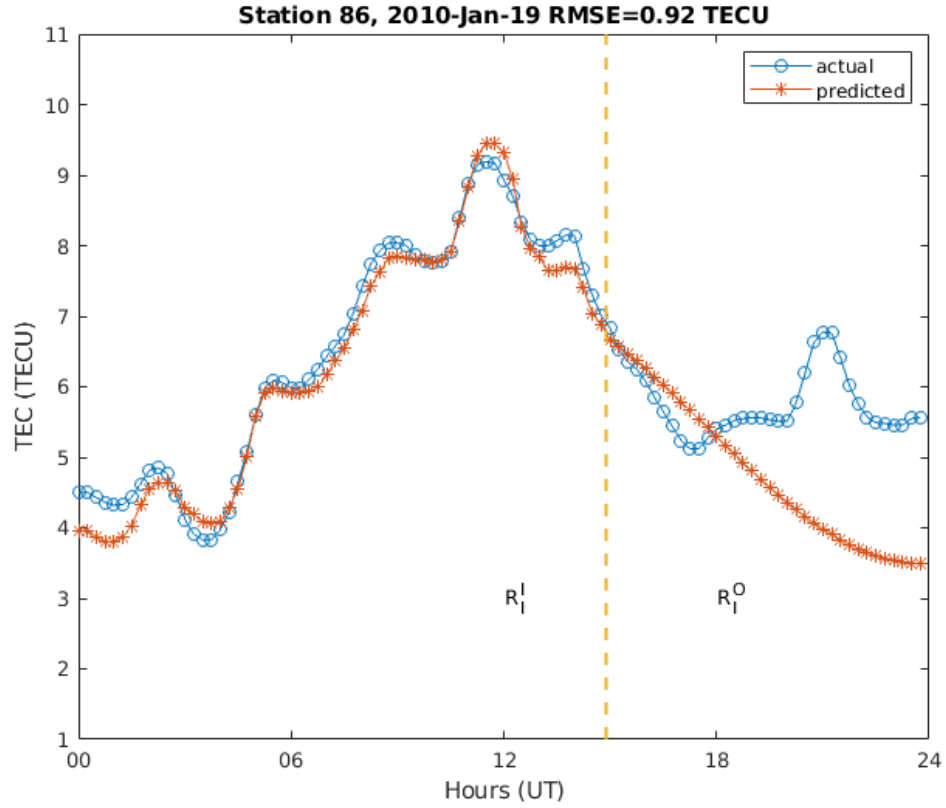


Figure 3.20: TEC temporal interpolation by Spatio-temporal Kriging, over $D_{ion-TEC}$ on station 86 in R_I , for January 19, 2010.

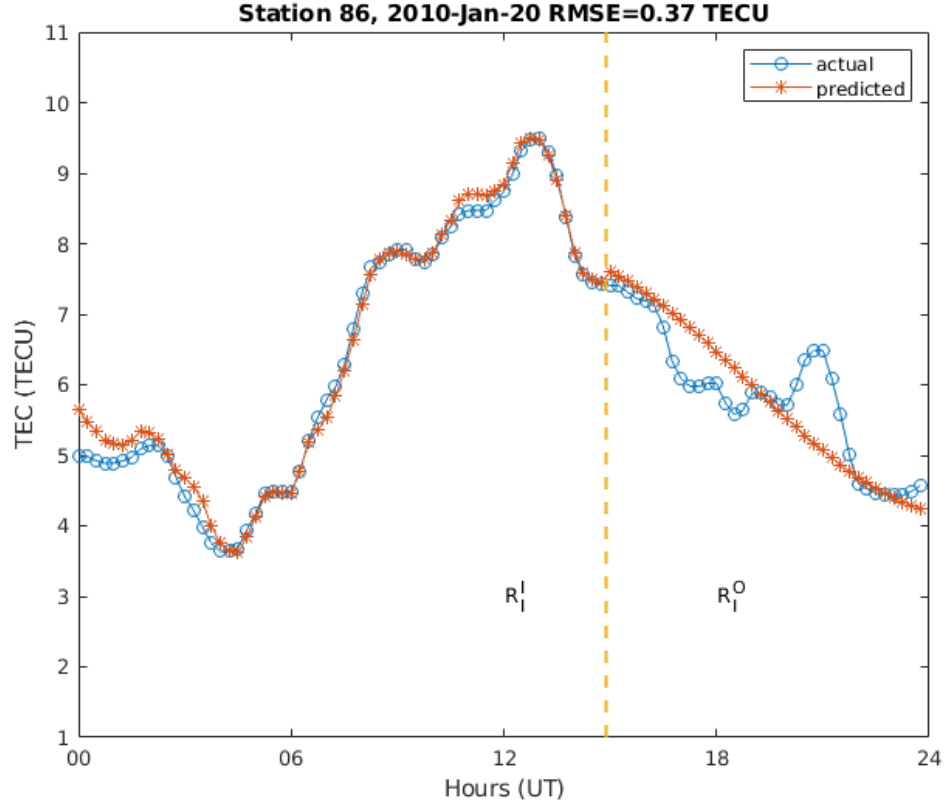


Figure 3.21: TEC temporal interpolation by Spatio-temporal Kriging, over $D_{ion-TEC}$ on station 86 in R_I , for January 20, 2010.

In the next Subsection, we provide the temporal plots for the region R_O .

3.5.4 Temporal Estimates - Outside Spatial Extent

The purpose of this section is to investigate how far the algorithm can go outside the spatial extent of data and still generate estimations within reasonable margin of error. We use training data in the region R_I but pick additional test stations from region R_O . The same data as used in the previous part is used for training and parameter estimation but the test stations are now chosen from region R_O . We generate predictions for nine stations in the region R_O and see how far away the predictions follow the actual data. The general trend is as expected, i.e. the

predictions degrade as we move away from the boundary but there are exceptions to the general rule also. See Table 4.3 for details.

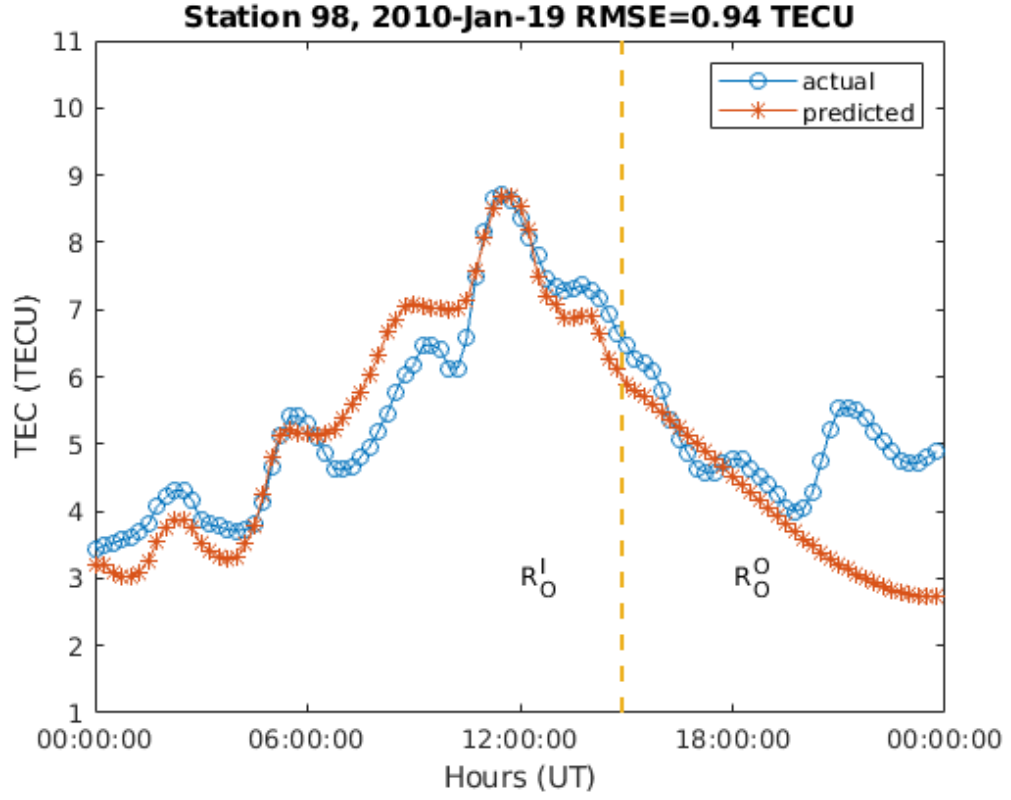


Figure 3.22: TEC temporal interpolation by Spatio-temporal Kriging, over $D_{ion-TEC}$ on station 98 in R_O , for January 19, 2010.

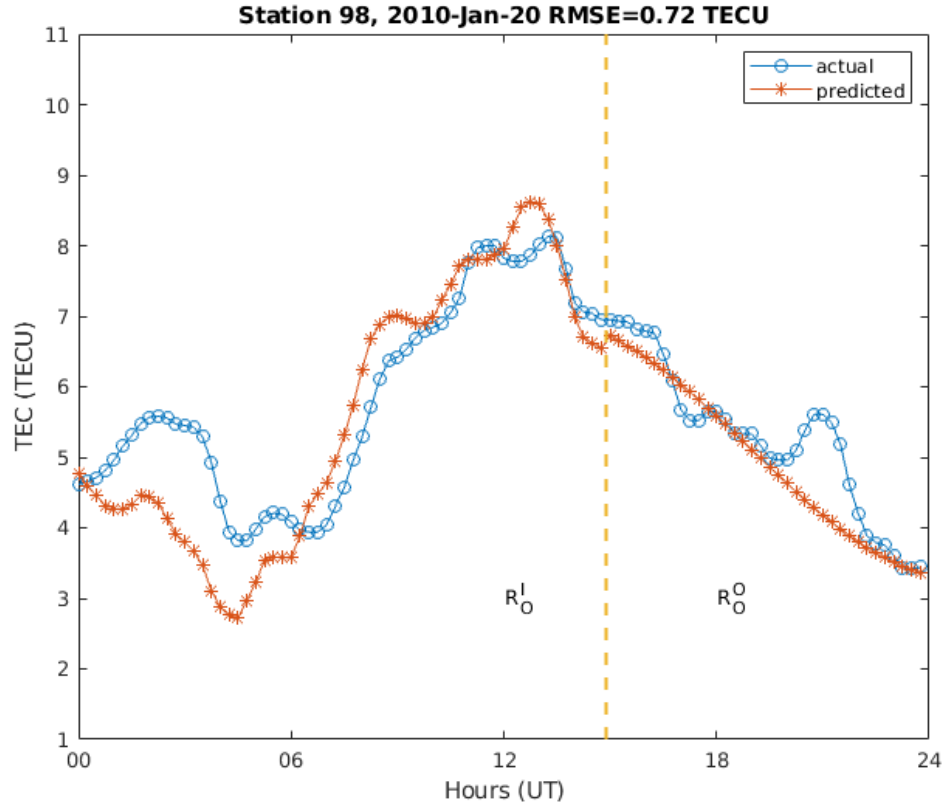


Figure 3.23: TEC temporal interpolation by Spatio-temporal Kriging, over $D_{ion-TEC}$ on station 98 in R_O , for January 20, 2010.

In this chapter we discussed how various interpolation techniques can be applied over ionospheric data and what kind of maps can be generated using these techniques. We also discuss the types of data appropriate for each technique and the usefulness of the techniques mentioned for some practical scenarios. We based this work on the earlier works of [52] for the IONOLAB and extended it from two dimensional estimations to three dimensional estimations and predictions.

Chapter 4

Discussion

This Chapter presents a comparison of error measures for various techniques discussed in this work. Spatial interpolation techniques are discussed first and a comparison of RMSE for different datasets is presented for these. Error comparison over actual ionospheric data as well as over model data generated using IRI-Plas model is provided for these techniques. Tables listing comparisons for the specific 15 geomagnetically quiet and disturbed days as well as tables listing RMSE over consecutive periods of time are provided. Later, error measures for Spatio-temporal Kriging are provided over TEC data.

4.1 Kriging vs Co-Kriging

In this section, we see error comparison for Co-Kriging with Universal Kriging. Table 4.1 shows the error values at station 11 (location 50°N, 15°E) for the month of October, 2016 over D_{iz} . The last row in the table shows RMSE for the month and as expected Co-Kriging generated lower RMSE than Kriging.

Table 4.1: Error Comparison at Station 11 in D_{iz} .

State	Date	Error for Kriging	Error for Co-Kriging
Quiet	10/01/16	-0.2074	-0.0760
Quiet	10/02/16	-0.0107	-0.1073
Quiet	10/03/16	0.4745	-0.1118
Quiet	10/04/16	-0.4069	-0.2480
Quiet	10/05/16	-1.8521	-0.3404
Quiet	10/06/16	0.8127	0.0493
Quiet	10/07/16	0.1292	-0.0167
Quiet	10/08/16	0.1673	-0.2632
Quiet	10/09/16	0.5850	0.6316
Quiet	10/10/16	0.1752	0.2338
Quiet	10/11/16	0.2629	0.2396
Disturbed	10/12/16	0.6839	0.2395
Quiet	10/13/16	0.9737	0.4237
Quiet	10/14/16	0.1347	0.2007
Quiet	10/15/16	0.6934	-0.2765
Quiet	10/16/16	0.0721	0.0745
Quiet	10/17/16	-0.8778	-0.0241
Quiet	10/18/16	-0.0914	-0.0147
Quiet	10/19/16	-0.0428	0.1961
Quiet	10/20/16	0.0450	0.0539
Quiet	10/21/16	-0.1645	-0.0866
Quiet	10/22/16	0.1024	0.0111
Disturbed	10/23/16	-0.0837	-0.2510
Quiet	10/24/16	-0.1669	0.1598
Quiet	10/25/16	0.4954	0.4593
Quiet	10/26/16	-0.2620	-0.0350
Quiet	10/27/16	-0.2347	-0.0622
Quiet	10/28/16	-0.2802	-0.2072
Quiet	10/29/16	-0.2752	0.1163
Quiet	10/30/16	0.0566	-0.0450
Quiet	10/31/16	-0.2648	0.0509
RMSE		0.5200	0.2211

The algorithm was also tested for the specific quiet and stormy days over IRI-Plas data and Table 4.2 shows that the test error for Co-Kriging is significantly lower than for Kriging. The trend remains the same for quiet days as well as for disturbed and disturbed followed by earthquake days.

Table 4.2: RMSE Comparison of IONOLAB-MAP Kriging and Co-Kriging over D_{iri} in region R_E .

State	Dates	RMSE Test Data		RMSE Training Data	
		Kriging	Co-Kriging	Kriging	Co-Kriging
Quiet	April 22, 2009	1.979	1.547	1.562	0.056
	March 12, 2010	1.310	0.947	1.070	0.081
	April 15, 2011	1.771	1.341	1.382	0.003
	June 12, 2011	1.164	0.889	1.018	0.089
	September 01, 2011	1.220	0.876	0.953	0.011
	September 21, 2011	1.265	0.867	0.980	0.023
	December 25, 2011	0.986	0.747	0.817	0.001
	August 15, 2012	1.069	0.814	0.895	0.018
Disturbed	February 05, 2011	0.957	0.735	0.759	0.009
	March 10, 2011	1.072	0.752	0.859	0.006
	May 28, 2011	1.316	1.022	1.098	0.581
	February 15, 2012	0.766	0.551	0.607	0.002
Disturbed+EQ	August 06, 2011	1.116	0.859	0.967	0.513
	November 01, 2011	0.737	0.523	0.621	0.004
	April 13, 2012	1.452	1.069	1.129	0.577

In the next Section, we discuss the performance of Spatio-temporal Kriging.

4.2 Kriging vs Spatio-temporal Kriging

To analyze the performance of Spatio-temporal Kriging, RMSE is computed over three months for strategically selected stations. Algorithm is applied over D_{ionTEC} in the region R_I^I and test stations are selected in regions R_I^I , R_I^O , and RO_O .

Figure 4.1 shows the GPS stations used as test points for RMSE computations. The station coordinates and station codes were provided earlier in Table C.

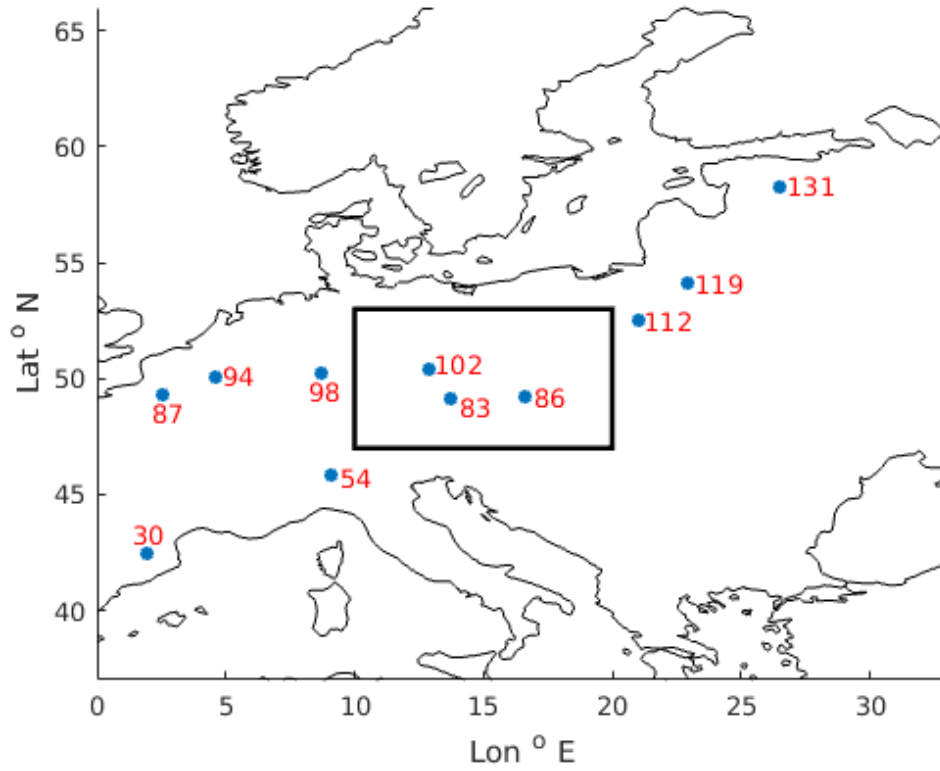


Figure 4.1: GPS Stations (indicated by filled circles) used as test stations in regions R_I and R_O .

Table 4.3 gives a comparison of RMSE for IONOLAB-MAP Kriging (Universal Kriging with Linear Trend), Spatio-temporal Kriging (STK) and Spatio-temporal Kriging (STK-SB). As expected, Spatio-temporal Kriging outperforms Kriging both inside and outside spatial and temporal extent of data used.

Table 4.3: RMSE over 3 months, January 01, 2010 to March 31, 2010.

		Location		RMSE (TECU)		
		Lat ° N	Lon ° E	STK	STK-SB	Kriging
R^I (12 UT)	R_I^I	49.21	16.59	0.1682	0.7150	0.2137
		50.36	12.89	0.4086	0.7918	0.3541
		49.13	13.72	0.6550	1.1335	0.6700
	R_O^I	62.39	30.1	3.1354	3.9168	3.9197
		58.27	26.47	2.6013	3.2207	2.1631
		54.1	22.93	1.3128	1.9578	1.9588
		52.48	21.04	1.1306	2.0225	1.8093
		50.09	4.59	0.6439	0.8752	1.4743
		49.26	2.51	1.0522	1.0409	1.8320
		50.22	8.73	0.4472	0.7405	0.9089
		42.48	1.97	1.5119	1.5091	2.6666
		45.8	9.1	1.0236	2.3806	2.9310
R^O (17 UT)	R_I^O	49.21	16.59	0.2851	1.0737	-
		50.36	12.89	0.8542	1.1116	-
		49.13	13.72	0.8782	0.9529	-
	R_O^O	62.39	30.1	3.9278	4.1975	-
		58.27	26.47	2.8833	3.7369	-
		54.1	22.93	1.7483	2.2164	-
		52.48	21.04	1.6975	2.1796	-
		50.09	4.59	0.9303	1.3171	-
		49.26	2.51	1.6216	1.1954	-
		50.22	8.73	1.1622	1.5494	-
		42.48	1.97	1.5181	2.2213	-
		45.8	9.1	1.0756	2.2052	-

The spatial trend of RMSE can be seen more easily in the form of Figures. As we move away from the spatial boundary of data used, the RMSE starts increasing gradually. We observe high errors at station 131. In addition to being farthest away from the training data, this station also lies in the high latitude

region and hence, displays higher error because the algorithm is provided training data in the midlatitude region only.

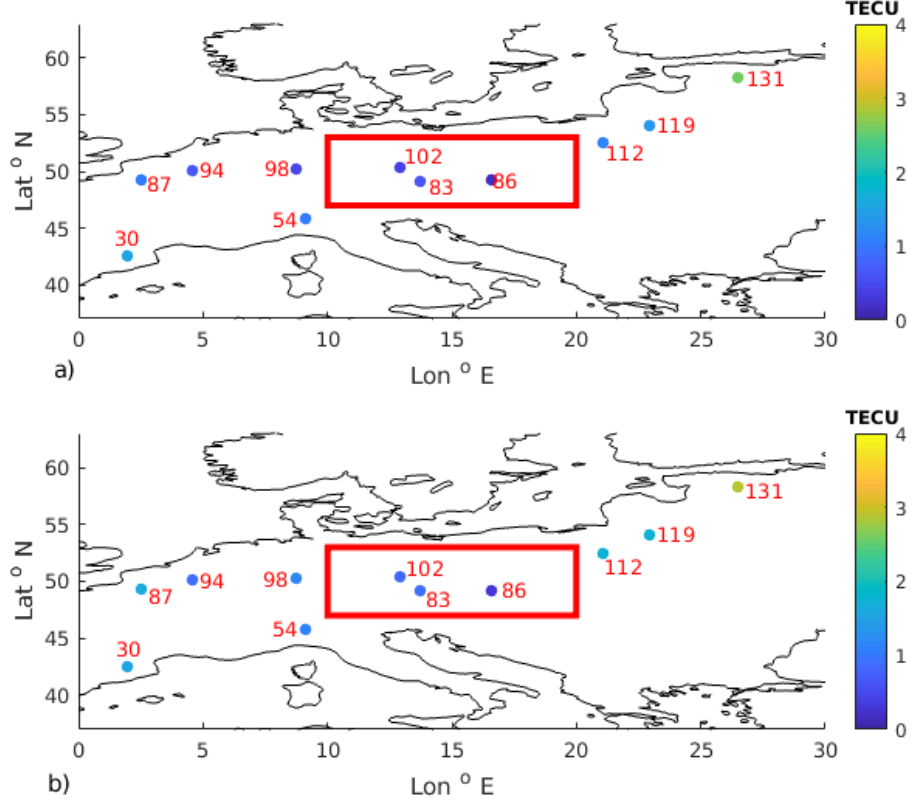


Figure 4.2: RMSE for Spatio-temporal Kriging STK from January 01, 2010 to March 31, 2010, indicated by colored circles, at stations indicated by station numbers that correspond to Table C a)RMSE over $D_{ion-TEC}$ for R_I^I and R_O^I b) RMSE over $D_{ion-TEC}$ for R_I^O and R_O^O .

Figure 4.2 shows the spatial boundary of region of interest. It marks the extent of data presented to the algorithm. Points located inside the region show a lower RMSE than the ones outside it, which is expected. Also, the comparison of the two subplots shows that predictions in the region R^O (two hours ahead of temporal boundary of training data) show higher RMSE, which is expected.

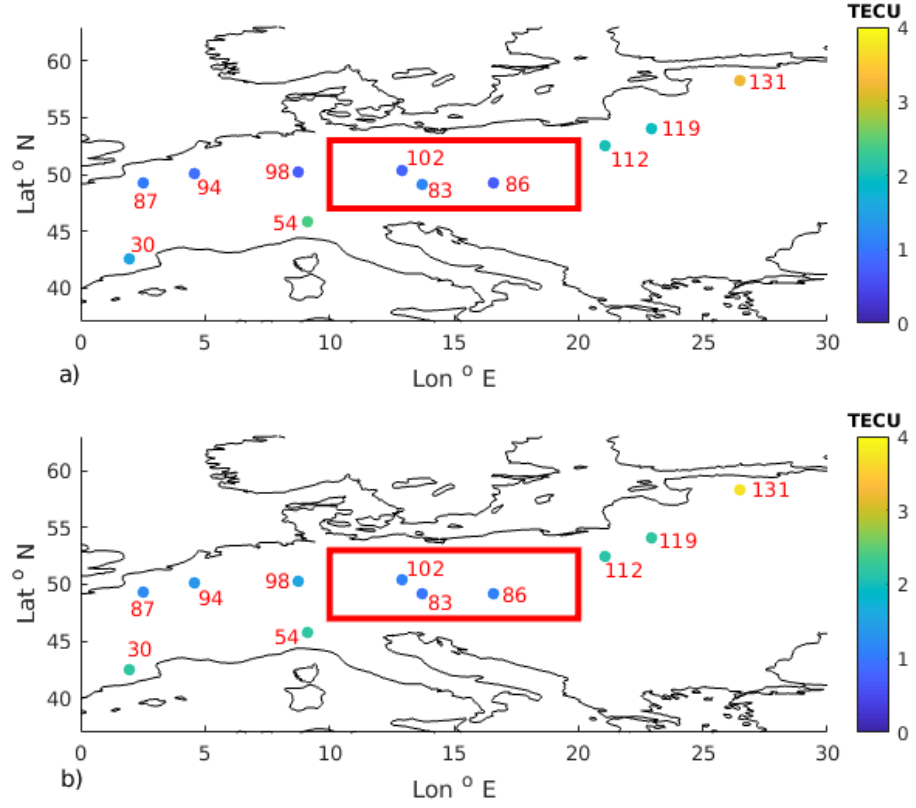


Figure 4.3: a) RMSE for Spatio-temporal Kriging STK-SB from January 01, 2010 to March 31, 2010, indicated by colored circles, at stations indicated by station numbers that correspond to Table C a) RMSE over $D_{ion-TEC}$ for R_I^I and R_O^I b) RMSE over $D_{ion-TEC}$ for R_I^O and R_O^O .

Figure 4.3 shows that STK-SB also follows the same RMSE spatial trend as that for STK but with slightly higher RMSE.

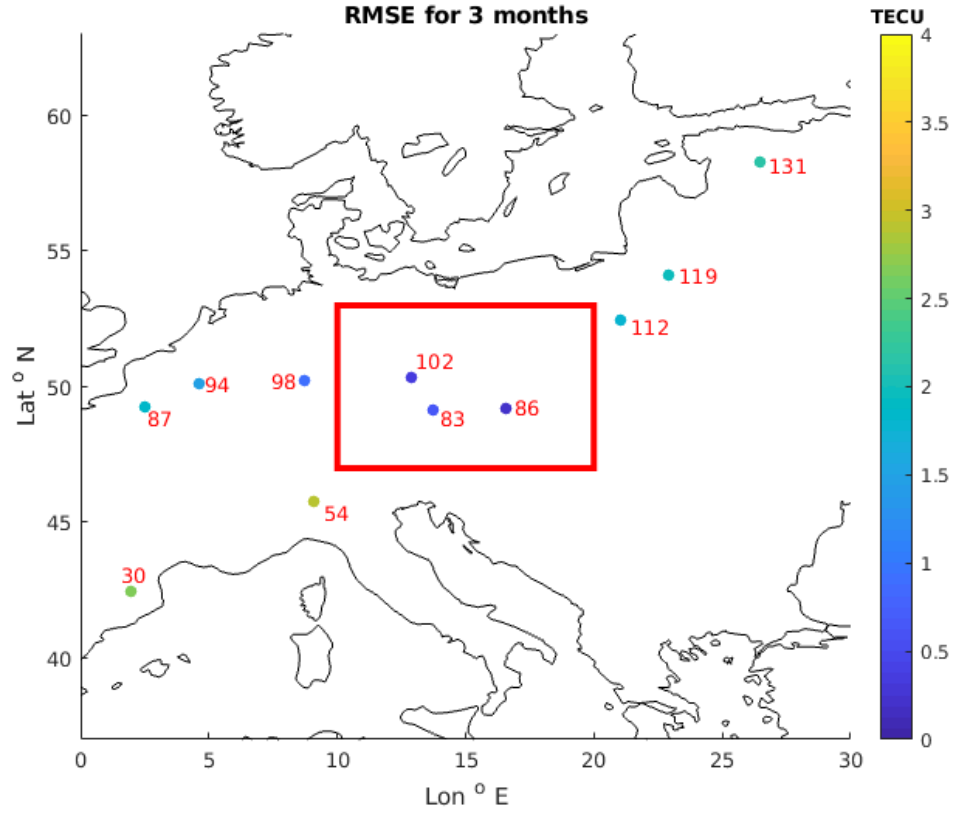


Figure 4.4: RMSE for IONOLAB-MAP Kriging (Universal Kriging with Linear Trend in this region) from January 01, 2010 to March 31, 2010, over $D_{ion-TEC}$ for R_I^I and R_O^I .

Finally, Figure 4.4 shows RMSE over same data using IONOLAB-MAP Universal Kriging. It performs reasonably well inside the spatial extent, in fact better for some cases but degrades more as we move outside the spatial boundary.

Figure 4.5 shows a comparison of RMSE for all the methods mentioned above against spatial distances. This Figure plots the RMSE in R^I as well as the RMSE in R^O (2 hours ahead RMSE) in the same plot. The lines with * represent the RMSE for R^O .

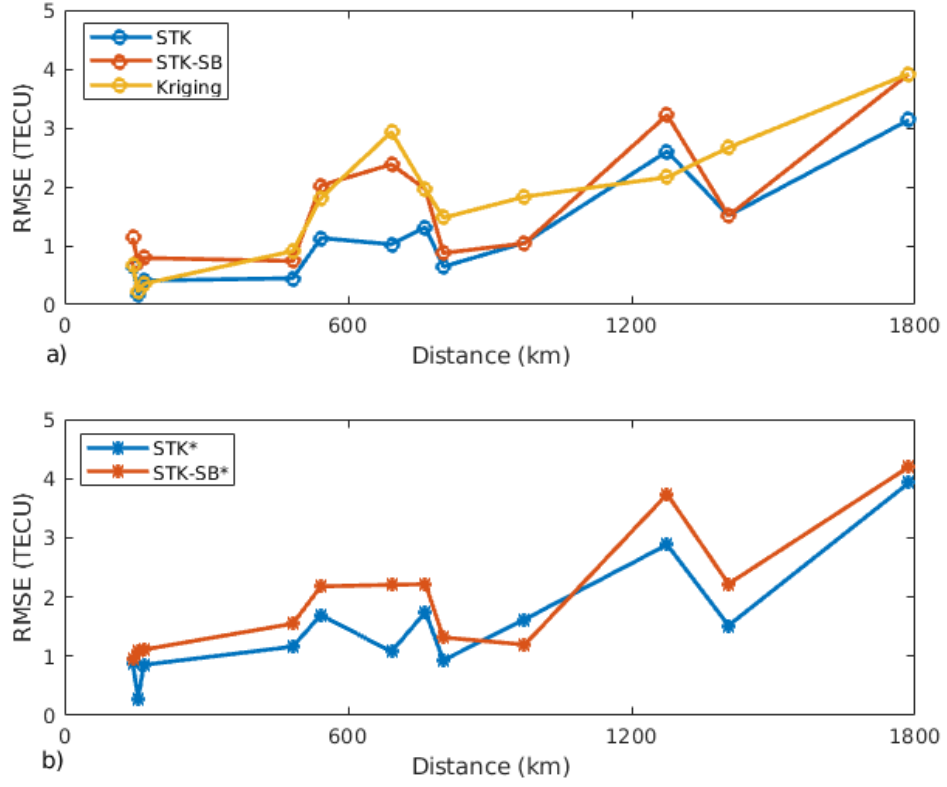


Figure 4.5: a) RMSE against spatial distance in (km) from the center of the region R_I , for Spatiotemporal Kriging STK, Spatiotemporal Kriging STK-SB and IONOLAB-MAP Universal Kriging in the region R^I b) RMSE against spatial distance in (km) from the center of the region R_I , for Spatiotemporal Kriging STK, Spatiotemporal Kriging STK-SB and IONOLAB-MAP Universal Kriging in the region R^O . The * indicates that these plots are the predictions in region R^O .

An increase in RMSE can be seen for the region R^O compared to the RMSE for R^I . This is expected because the algorithm was applied on data in R^I and extrapolated to R^O .

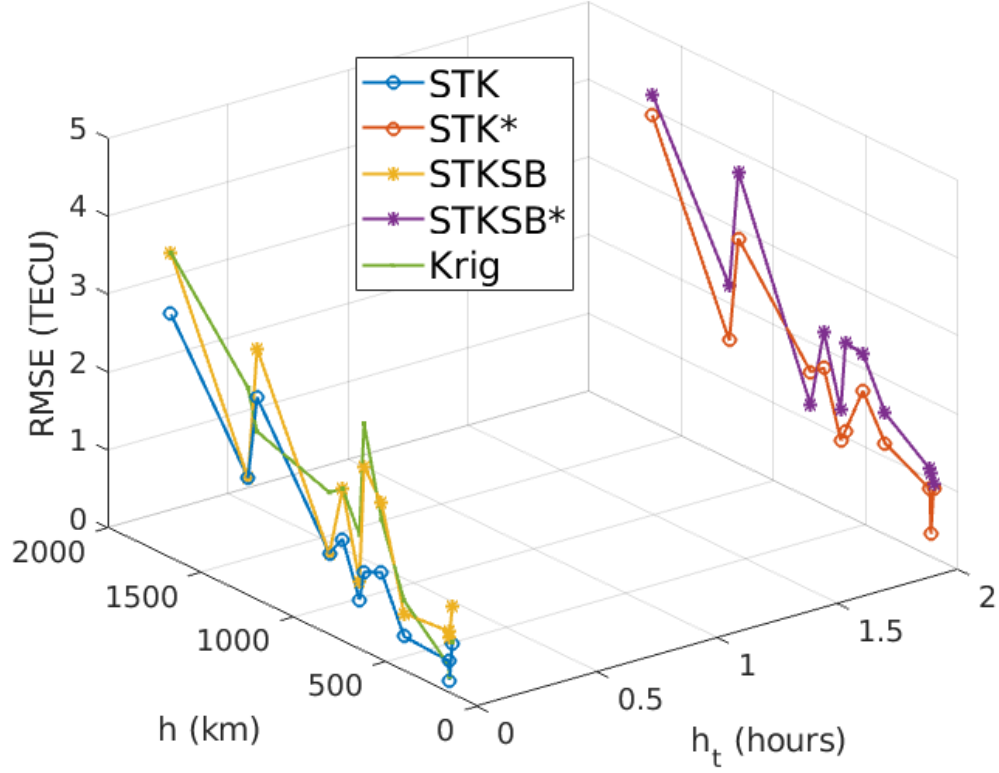


Figure 4.6: RMSE against spatial and temporal distances h and h_t from the center of the region R_I^I , (o-) represents the RMSE for the region R^I and (*-) represents the RMSE for the region R^O .

The spatial distance here for each test point is measured from the center of region R_I . Figure 4.6 shows increasing RMSE with increasing distance, as well as higher RMSE for 2 hours ahead predictions.

Figure 4.7 shows the density of Error computed using Equation 3.3, over test stations in all the regions.

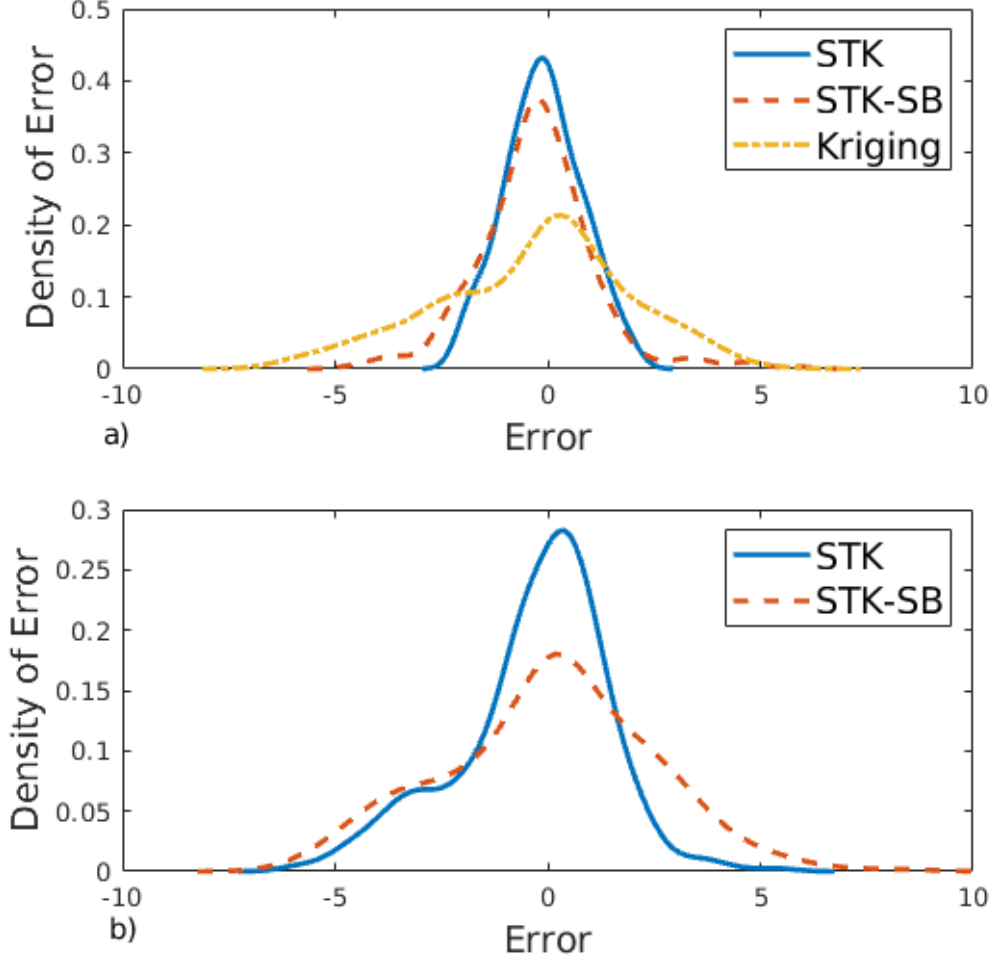


Figure 4.7: Error Density indicated by solid line for STK, dashed line for STK-SB, and dash dot line for IONOLAB-MAP Kriging over $D_{ion-TEC}$ from January 01, 2010 to March 31, 2010 a) in the region R^I , b) in the region R^O .

This Chapter discussed a comparative analysis of the Spatial and Spatio-temporal techniques used. We analyzed the maps over a number of datasets for different periods of time and observed the advantages of using Co-Kriging and Spatio-temporal Kriging over simpler interpolation methods. In the next Chapter, we provide concluding remarks and possible future works.

Chapter 5

Conclusion

The need for the interpolation of ionospheric modalities arises from the sparsity of the data available. Ionosphere surrounds the earth everywhere, whereas data for ionospheric modalities such as TEC, foF2, and hmF2 is collected only at sparse locations. The data for foF2 and hmF2 is even more sparse than the data for TEC because of the expensive nature of ionosondes but these modalities are very important for reliable satellite communication and an estimate of these modalities is needed on dense regular grids. In short, maps of these modalities are of importance for practical communication applications and the motivation behind this thesis was to investigate various methods for such interpolation and improve upon them in order to generate more reliable maps.

In order to start working on an interpolation problem, we first examined the most basic interpolation techniques that do not depend on any assumptions about the data available. Such interpolation schemes can be applied on any type of data available regardless of the nature of trends in the data about correlations and periodicity. One of the most basic techniques applied and examined was Inverse Squared Distance Weighting method. This technique is for spatial interpolations and is based on one simple assumption that locations closer to each other have somewhat similar data values and as we move away from these points, values start becoming more and more different from the values at the points where we

know the actual data. This method has been applied on almost every type of interpolation throughout the history of interpolation techniques.

The need for more advanced techniques than simply assigning the same values to all the points in a neighborhood or assigning them values based on their distance from the known value arises from underlying trends in the data. The assumption of spatial correlation based on the distance still holds. An examination of the ionospheric data over a period of time tells us that the data is not only spatially correlated but also has an approximately constant mean across the days at any specific time of the day. This assumption forms the basis of exploring unbiased estimators. Among the linear unbiased estimators, Kriging is the most widely used estimator. We examined the application of Kriging in past ionospheric studies and the performance of Kriging over ionospheric data motivates us to extend it further and improve upon it. We implement Kriging over multiple datasets and later compare it with other techniques. We observe that the data D_{iz} has a scaled version of actual values so we use a scaling factor of 0.1 to bring our data in the actual TEC range.

The idea of extending Kriging from using one modality for interpolation to using multiple modalities has been around in the literature before and we discussed some applications of it in the literature review. The motivation behind applying such idea on ionospheric data came from the strong correlation trends observed in the ionospheric modalities. The expression for Co-Kriging has been used in the literature before and partial derivations are also available. This thesis provides a complete derivation of Co-Kriging for the reader, based upon the derivation in [47].

The geostatistical techniques we apply on our data rely on variogram and cross-variogram calculations. Hence, they involve some multiplications and additions of the primary and the secondary modalities. Without appropriate data preparation, our results go beyond the range of physical measurements. For the sophisticated techniques like Co-Kriging, all the modalities being used must be in the same units. We therefore present the mathematical formulation for such data preparation in this study. We applied Co-Kriging method first on IRI-Plas model

data D_{iri} to see how it improves upon Kriging. To model a realistic scenario, data for TEC was used for more locations than the data for foF2 and performance was compared against Kriging. Use of multiple test locations and using data over multiple days including both quiet as well as disturbed days enabled us to generate statistically significant measures of error and judge the ability of Co-Kriging for improving upon Kriging. Having seen the improvements over Kriging over model data, we later applied Co-Kriging over co-located data in D_{iz} , where we could still spare some data for testing and observed the performance of Co-Kriging. Finally, we applied the algorithm on ionosonde data D_{snd} to generate high resolution foF2 maps using TEC data $D_{ion-TEC}$ prepared by IONOLAB as the secondary modality.

The motivation behind moving on to Spatio-temporal domain from spatial domain is obvious from the Spatio-temporal nature of the ionospheric data. We observed that this data is correlated not only spatially but also temporally and such Spatio-temporal correlation can help us generate Spatio-temporal maps. We studied the application of Spatio-temporal Kriging over rainfall estimation problem by Skøien and Blöschl in [58] and implemented their approach on $D_{ion-TEC}$. We observed two limitations in their approach and this thesis contributes towards improving upon their approach. The first limitation was seen in the way number of observed locations were treated individually and how their Spatio-temporal variograms were actually just a collection of temporal variograms for every pair of observed locations. We eliminate this limitation and take the full advantage of Spatio-temporal Kriging by using the Spatio-temporal variograms. The second limitation observed was the way Spatio-temporal correlations were modeled. Their approach works on the basic assumption that the correlation in the data decreases as we move further and further away both spatially and temporally. Such an approach cannot model temporal periodicity in the data (if there is any) because of a monotonically increasing function as a variogram. We improve upon this technique to incorporate the ability to model temporal periodicity by experimenting with various possible forms for such a model and finally arriving at the model proposed by Bilonick in [57]. We also introduce validation to further improve the performance over test data and extrapolate the maps both spatially

and temporally and the validation helps us improve test performance in all such scenarios. The major challenge we faced during the implementation of Spatio-temporal Kriging was the memory constraints on calculation of variograms. We dealt with that by using much lower temporal resolution (15 minutes) for the data than the data actually has (30 second). We did this by doing decimation on the data by a factor R after low pass filtering with a cut off frequency of π/R . These results can be improved further by allowing higher temporal resolutions. Another way to reduce the computational complexity is further restricting the region of interest to a smaller region.

Finally, it is notable to mention that Spatio-temporal Kriging is not the only way of ionospheric interpolation, neither it is the end of research in this domain. There might be many existing techniques in other disciplines such as machine learning that might be able to generate better temporal predictions but a comparison of that could not be studied. As a natural extension of Spatio-temporal Kriging, Spatio-temporal Co-Kriging can be studied in future and this work can be considered a basis for that. The greatest challenge in such implementation would be the memory constraints in computer systems.

Bibliography

- [1] P. M. Kintner and B. M. Ledvina, “The ionosphere, radio navigation, and global navigation satellite systems,” *Advances in Space Research*, vol. 35, no. 5, pp. 788–811, 2005.
- [2] K. Davies, *Ionospheric radio*. IET, 1990.
- [3] A. Komjathy, *Global ionospheric total electron content mapping using the Global Positioning System*. PhD thesis, University of New Brunswick, 1997.
- [4] T. Tascione, *Introduction to the Space Environment*. Intelligence Series, Krieger Pub., 1994.
- [5] I. K. Pawel Wielgosz, Dorota Grejner-Brzezinska, “Regional Ionosphere Mapping with Kriging and Multiquadric Methods,” *Journal of Global Positioning Systems*, vol. 2, no. 1, pp. 48–55, 2003.
- [6] T. Garner, T. Gaussiran Ii, B. Tolman, R. Harris, R. Calfas, and H. Gallagher, “Total electron content measurements in ionospheric physics,” *Advances in Space Research*, vol. 42, no. 4, pp. 720–726, 2008.
- [7] L. Kersley, D. Malan, S. E. Pryse, L. R. Cander, R. A. Bamford, A. Belehaki, R. Leitinger, S. M. Radicella, C. N. Mitchell, and P. S. Spencer, “Total electron content-a key parameter in propagation: measurement and use in ionospheric imaging,” *Annals of Geophysics*, 2004.
- [8] C. Ho, B. Wilson, A. Mannucci, U. Lindqwister, and D. Yuan, “A comparative study of ionospheric total electron content measurements using global

- ionospheric maps of GPS, TOPEX radar, and the Bent model,” *Radio Science*, vol. 32, no. 4, pp. 1499–1512, 1997.
- [9] D. Bilitza, D. Altadill, Y. Zhang, C. Mertens, V. Truhlik, P. Richards, L.-A. McKinnell, and B. Reinisch, “The international reference ionosphere 2012—a model of international collaboration,” *Journal of Space Weather and Space Climate*, vol. 4, p. A07, 2014.
 - [10] D. Bilitza and T. Gulyaeva, “Towards ISO standard earth ionosphere and plasmasphere model,” in *39th COSPAR Scientific Assembly*, vol. 39, p. 192, 2012.
 - [11] T. Gulyaeva, F. Arikan, and I. Stanislawski, “Inter-hemispheric imaging of the ionosphere with the upgraded IRI-Plas model during the space weather storms,” *Earth, Planets and Space*, vol. 63, no. 8, pp. 929–939, 2011.
 - [12] T. Gulyaeva, “Storm time behavior of topside scale height inferred from the ionosphere–plasmasphere model driven by the F2 layer peak and GPS-TEC observations,” *Advances in Space Research*, vol. 47, no. 6, pp. 913–920, 2011.
 - [13] T. L. Gulyaeva, F. Arikan, M. Hernandez-Pajares, and I. Stanislawski, “GIM-TEC adaptive ionospheric weather assessment and forecast system,” *Journal of Atmospheric and Solar-Terrestrial Physics*, vol. 102, pp. 329–340, 2013.
 - [14] F. Arikan, U. Sezen, T. Gulyaeva, and O. Cilibas, “Online, automatic, ionospheric maps: IRI-PLAS-MAP,” *Advances in Space Research*, vol. 55, no. 8, pp. 2106–2113, 2015.
 - [15] D. G. Krige, “A statistical approach to some basic mine valuation problems on the witwatersrand,” *Journal of the Southern African Institute of Mining and Metallurgy*, vol. 52, no. 6, pp. 119–139, 1951.
 - [16] G. Matheron, “Kriging or polynomial interpolation procedures,” *CIMM Transactions*, vol. 70, pp. 240–244, 1967.
 - [17] A. Marechal and J. Serra, “Random Kriging,” in *Geostatistics*, pp. 91–112, Springer, 1970.

- [18] J. Delhomme, “Kriging in the hydrosociences,” *Advances in Water Resources*, vol. 1, pp. 251–266, 1978.
- [19] J. Creutin and C. Obled, “Objective analyses and mapping techniques for rainfall fields: an objective comparison,” *Water Resources Research*, vol. 18, no. 2, pp. 413–431, 1982.
- [20] T. Lebel, G. Bastin, C. Obled, and J. Creutin, “On the accuracy of areal rainfall estimation: a case study,” *Water Resources Research*, vol. 23, no. 11, pp. 2123–2134, 1987.
- [21] B. P. Eynon and P. Switzer, “The variability of rainfall acidity,” *Canadian Journal of Statistics*, vol. 11, no. 1, pp. 11–23, 1983.
- [22] P. Goovaerts, “Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall,” *Journal of Hydrology*, vol. 228, no. 1, pp. 113–129, 2000.
- [23] J. Chica-Olmo, “Prediction of Housing Location Price by Multivariate Spatial Method: Co-Kriging,” *Journal of Real Estate Research*, vol. 29, no. 1, pp. 91–113, 2007.
- [24] I. Sideris, M. Gabella, R. Erdin, and U. Germann, “Real-time radar–rain-gauge merging using spatio-temporal Co-Kriging with external drift in the alpine terrain of Switzerland,” *Quarterly Journal of the Royal Meteorological Society*, vol. 140, no. 680, pp. 1097–1111, 2014.
- [25] H. Liu, B. Yang, and E. Kang, “Co-Kriging method for spatio-temporal assimilation of multi-scale satellite data,” in *Geoscience and Remote Sensing Symposium (IGARSS), 2015 IEEE International*, pp. 3314–3316, IEEE, 2015.
- [26] S. Roberts, M. Osborne, M. Ebdon, S. Reece, N. Gibson, and S. Aigrain, “Gaussian processes for time-series modelling,” *Phil. Trans. R. Soc. A*, vol. 371, no. 1984, p. 20110550, 2013.
- [27] C. E. Rasmussen, “Gaussian processes in machine learning,” in *Advanced lectures on machine learning*, pp. 63–71, Springer, 2004.

- [28] D. W. Caress and R. L. Parker, “Spectral interpolation and downward continuation of marine magnetic anomaly data,” *Journal of Geophysical Research: Solid Earth*, vol. 94, no. B12, pp. 17393–17407, 1989.
- [29] R. Orús, M. Hernández-Pajares, J. Juan, and J. Sanz, “Improvement of global ionospheric VTEC maps by using Kriging interpolation technique,” *Journal of Atmospheric and Solar-Terrestrial Physics*, vol. 67, no. 16, pp. 1598–1609, 2005.
- [30] G. D. F. Riccardo Notarpietro, Fabio Dovis and M. Aquino, *Mitigation of Ionospheric Threats to GNSS*. InTech, 2014.
- [31] I. Sayin, F. Arikan, and O. Arikan, “Regional TEC mapping with random field priors and Kriging,” *Radio Science*, vol. 43, no. 5, 2008.
- [32] F. Arikan, A. Yilmaz, O. Arikan, I. Sayin, M. Gurun, K. Akdogan, and S. Yildirim, “Space weather activities of IONOLAB group: TEC mapping,” in *EGU General Assembly Conference Abstracts*, vol. 11, p. 6962, 2009.
- [33] M. N. Deviren, F. Arikan, and O. Arikan, “Automatic regional mapping of total electron content using a GPS sensor network and isotropic Universal Kriging,” in *Information Fusion (FUSION), 2013 16th International Conference on*, pp. 1664–1669, IEEE, 2013.
- [34] F. Arikan, C. Erol, and O. Arikan, “Regularized estimation of vertical total electron content from global positioning system data,” *Journal of Geophysical Research: Space Physics*, vol. 108, no. A12, 2003.
- [35] H. Nayir, F. Arikan, O. Arikan, and C. Erol, “Total electron content estimation with reg-est,” *Journal of Geophysical Research: Space Physics*, vol. 112, no. A11, 2007.
- [36] F. Arikan, H. Nayir, U. Sezen, and O. Arikan, “Estimation of single station interfrequency receiver bias using GPS-TEC,” *Radio Science*, vol. 43, no. 4, 2008.
- [37] U. Sezen, O. Sahin, F. Arikan, and O. Arikan, “Estimation of hmF2 and foF2 Communication Parameters of Ionosphere F2-Layer Using GPS Data

- and IRI-Plas Model,” *IEEE Transactions on Antennas and Propagation*, vol. 61, no. 10, pp. 5264–5273, 2013.
- [38] Y. M. Lee, J. H. Jo, M. H. You, and C. O. Jeong, “Ionospheric electron density profiles derived from obliquely sounded HF radar,” in *Microwaves, Radar, and Wireless Communication (MIKON), 2014 20th International Conference on*, pp. 1–4, IEEE, 2014.
- [39] U. Sezen, O. Cilibas, F. Arikan, and T. Gulyaeva, “Online regional foF2 and hmF2 Maps from IRI-Plas: IONOLAB-MAP,” in *International Reference Ionosphere (IRI) Workshop*, 2013.
- [40] T. Gulyaeva and F. Arikan, “Statistical discrimination of global post-seismic ionosphere effects under geomagnetic quiet and storm conditions,” *Geomatics, Natural Hazards and Risk*, vol. 8, no. 2, pp. 509–524, 2017.
- [41] H. Tuna, *3D Electron Density Estimation in the Ionosphere by Using IRI-Plas model and GPS Measurements*. PhD thesis, Bilkent University, 2016.
- [42] N. Cressie, “The origins of Kriging,” *Mathematical geology*, vol. 22, no. 3, pp. 239–252, 1990.
- [43] P. Goovaerts, *Geostatistics for Natural Resources Evaluation*. Oxford University Press, 1997.
- [44] M. L. Stein, *Interpolation of Spatial Data, Some Theory for Kriging*. Springer Series in Statistics, Springer Science+Business Media, LLC, 1999.
- [45] M. Sherman, *Spatial Statistics and Spatio-Temporal Data, Covariance Functions and Directional Properties*. Wiley Series in Probability and Statistics, John Wiley and Sons Ltd, 2011.
- [46] B. Minasny and A. B. McBratney, “The matérn function as a general model for soil variograms,” *Geoderma*, vol. 128, no. 3-4, pp. 192–207, 2005.
- [47] N. A. C. Cressie, *Statistics for Spatial Data*. Wiley Series in Probability and Statistics, Wiley-Interscience, 1993.

- [48] A. Stein, F. D. van der Meer, and B. Gorte, *Spatial Statistics for Remote Sensing*, vol. 1. Springer Science & Business Media, 2006.
- [49] B. Gräler, M. Rehr, L. Gerharz, and E. Pebesma, “Spatio-temporal analysis and interpolation of PM10 measurements in Europe for 2009,” *ETC/ACM Technical Paper*, vol. 8, pp. 1–29, 2012.
- [50] Z. C. Zeng, L. Lei, S. Hou, F. Ru, X. Guan, and B. Zhang, “A regional gap-filling method based on spatiotemporal variogram model of CO_2 columns,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 6, pp. 3594–3603, 2014.
- [51] M. N. Deviren, F. Arikan, I. Sayin, and O. Koroglu, “Automatic TEC mapping using a GPS network and GIM-TEC,” in *Radio Science Conference (URSI AT-RASC), 2015 1st URSI Atlantic*, pp. 1–1, IEEE, 2015.
- [52] M. N. Deviren, “Estimation of space-time random field for Total Electron Content (TEC) over Turkey,” *MSc, Hacettepe University, Ankara, Turkey*, vol. 7, 2013.
- [53] J. P. Chilès and P. Delfiner, *Geostatistics: Modeling Spatial Uncertainty*. Wiley Series in Probability and Statistics, Wiley Online Library, 1999.
- [54] B. Minasny and A. B. McBratney, “The Matérn function as a general model for soil variograms,” *Geoderma*, vol. 128, no. 3, pp. 192–207, 2005.
- [55] N. Cressie, *Statistics for spatial data*. John Wiley & Sons, 2015.
- [56] J. O. Skøien and G. Blöschl, “Catchments as space-time filters: A joint spatio-temporal geostatistical analysis of runoff and precipitation,” *Hydrology and Earth System Sciences Discussions*, vol. 3, no. 3, pp. 941–985, 2006.
- [57] R. A. Bilonick, “The space-time distribution of sulfate deposition in the northeastern United States,” *Atmospheric Environment (1967)*, vol. 19, no. 11, pp. 1829–1845, 1985.
- [58] J. O. Skøien and G. Blöschl, “Spatiotemporal topological Kriging of runoff time series,” *Water Resources Research*, vol. 43, no. 9, 2007.

Appendix A

Earthquakes Tables

Table A.1: Earthquakes of magnitude 5.5M and above for the days mentioned in this work as well as for 7 days before and after the days mentioned.

Date	Lat ° N	Lon ° E	Magnitude	Place
April 0, 2009	42.334	13.334	6.3	Central Italy
April 07, 2009	42.275	13.464	5.5	Central Italy
March 08, 2010	38.864	39.986	6.1	Eastern Turkey
March 08,2010	38.709	40.051	5.6	Eastern Turkey
April 11, 2010	36.965	-3.542	6.3	Strait of Gibraltar
Feburary 28, 2011	34.98	25.42	5.7	Crete, Greece
April 01, 2011	35.662	26.56	6	Crete, Greece
July 19, 2011	40.081	71.41	6.1	Kyrgyzstan
August 11, 2011	39.955	77.028	5.6	Southern Xinjiang, China
September 22, 2011	39.785	38.842	5.5	Eastern Turkey
October 25, 2011	38.811	43.623	5.6	Eastern Turkey
November 09, 2011	38.429	43.229	5.6	Eastern Turkey
April 16, 2012	36.632	21.475	5.8	Southern Greece

Appendix B

Correlation Tables

Table B.1: Correlation of foF2 with TEC for October, 2016 from 0 UT to 23 UT.

		UT																								
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	
October, 2016	1	0.52	0.60	0.78	0.70	0.85	0.77	0.90	0.87	0.91	0.85	0.97	0.97	0.98	0.98	0.97	0.97	0.86	0.92	0.90	0.89	0.71	0.14	0.43	0.60	
	2	0.57	0.18	0.39	0.51	0.82	0.77	0.90	0.89	0.85	0.94	0.91	0.96	0.98	0.97	0.99	0.98	0.97	0.93	0.82	0.70	0.68	0.58	0.60		
	3	0.56	0.78	0.87	0.60	0.73	0.89	0.92	0.92	0.93	0.91	0.96	0.96	0.91	0.96	0.96	0.95	0.96	0.93	0.85	0.72	0.39	0.59	0.61	0.54	
	4	0.63	0.74	0.40	0.43	0.81	0.89	0.96	0.96	0.93	0.97	0.95	0.96	0.99	0.99	0.97	0.98	0.90	0.84	0.72	0.66	0.85	0.71	0.71	0.71	
	5	0.63	0.79	0.37	0.43	0.70	0.93	0.94	0.96	0.95	0.93	0.90	0.92	0.91	0.95	0.94	0.98	0.95	0.94	0.84	0.61	0.68	0.43	0.41	0.53	
	6	0.65	0.21	0.48	0.36	0.89	0.84	0.92	0.90	0.92	0.91	0.83	0.88	0.93	0.93	0.95	0.90	0.96	0.98	0.81	0.83	0.70	0.58	0.59	0.68	0.63
	7	0.45	0.60	0.59	0.50	0.28	0.91	0.89	0.92	0.88	0.89	0.94	0.92	0.85	0.89	0.97	0.95	0.96	0.84	0.77	0.72	0.67	0.67	0.59	0.72	
	8	0.72	0.86	0.77	0.56	0.87	0.94	0.93	0.91	0.80	0.85	0.92	0.92	0.93	0.92	0.94	0.96	0.95	0.85	0.75	0.72	0.77	0.76	0.74	0.66	
	9	0.56	0.54	0.77	0.66	0.82	0.87	0.91	0.93	0.90	0.86	0.85	0.91	0.89	0.91	0.96	0.95	0.97	0.79	0.80	0.72	0.85	0.74	0.58	0.04	
	10	0.46	0.51	0.47	0.71	0.86	0.91	0.92	0.94	0.87	0.90	0.88	0.95	0.95	0.95	0.98	0.97	0.95	0.90	0.80	0.71	0.39	0.50	0.64	0.72	
	11	0.16	0.12	0.57	0.57	0.58	0.89	0.92	0.93	0.89	0.84	0.91	0.86	0.92	0.93	0.91	0.96	0.88	0.84	0.74	0.73	0.36	0.41	0.48	0.59	
	12	0.46	0.73	0.75	0.72	0.87	0.85	0.86	0.86	0.86	0.83	0.82	0.67	0.71	0.75	0.91	0.96	0.95	0.83	0.64	0.60	0.20	0.04	0.42	0.56	
	13	0.41	0.54	0.31	0.39	0.61	0.87	0.87	0.89	0.89	0.84	0.77	0.66	0.93	0.96	0.91	0.92	0.96	0.85	0.70	0.59	0.44	0.91	0.89	0.66	0.66
	14	0.63	0.68	0.45	0.76	0.82	0.54	0.83	0.78	0.81	0.84	0.92	0.94	0.96	0.98	0.97	0.99	0.97	0.79	0.64	0.63	0.47	0.39	0.69	0.71	
	15	0.79	0.81	0.63	0.60	0.71	0.89	0.81	0.80	0.69	0.69	0.76	0.88	0.90	0.90	0.93	0.95	0.95	0.65	0.68	0.75	0.69	0.52	0.56	0.60	
	16	0.82	0.71	0.81	0.74	0.88	0.86	0.72	0.67	0.73	0.76	0.87	0.91	0.89	0.91	0.96	0.94	0.93	0.88	0.56	0.68	0.70	0.78	0.68	0.81	
	17	0.58	0.20	0.88	0.76	0.78	0.92	0.94	0.86	0.84	0.88	0.95	0.97	0.99	0.99	0.95	0.98	0.90	0.82	0.65	0.73	0.51	0.57	0.31	0.51	
	18	0.53	0.88	0.87	0.79	0.48	0.92	0.93	0.96	0.92	0.92	0.99	0.98	0.96	0.95	0.99	0.95	0.85	0.83	0.80	0.79	0.52	0.70	0.60	0.71	
	19	0.91	0.92	0.79	0.88	0.89	0.91	0.94	0.92	0.89	0.89	0.86	0.92	0.96	0.98	0.98	0.88	0.93	0.66	0.51	0.70	0.06	0.73	0.85	0.82	
	20	0.88	0.91	0.89	0.89	0.89	0.85	0.80	0.86	0.71	0.55	0.62	0.81	0.87	0.93	0.94	0.96	0.94	0.75	0.75	0.69	0.16	0.65	0.61	0.48	
	21	0.62	0.66	0.81	0.75	0.78	0.84	0.76	0.67	0.36	0.21	0.45	0.56	0.84	0.88	0.89	0.87	0.85	0.73	0.62	0.73	0.81	0.77	0.61	0.50	
	22	0.51	0.65	0.47	0.40	0.72	0.91	0.76	0.65	0.52	0.07	0.55	0.58	0.68	0.79	0.90	0.93	0.82	0.70	0.65	0.41	0.56	0.33	0.38	0.54	
	23	0.51	0.63	0.70	0.73	0.78	0.88	0.80	0.62	0.52	0.25	0.53	0.80	0.77	0.72	0.84	0.89	0.91	0.82	0.85	0.62	0.74	0.70	0.35	0.40	
	24	0.34	0.37	0.51	0.63	0.90	0.94	0.95	0.99	0.93	0.98	0.98	0.96	0.96	0.91	0.97	0.69	0.64	0.68	0.82	0.61	0.15	0.26	0.47	0.49	
	25	0.76	0.74	0.69	0.87	0.88	0.95	0.98	0.98	0.97	0.96	0.96	0.95	0.97	0.94	0.85	0.98	0.90	0.83	0.89	0.62	0.74	0.57	0.72	0.61	
	26	0.88	0.84	0.77	0.70	0.89	0.94	0.96	0.97	0.97	0.97	0.95	0.97	0.96	0.97	0.68	0.47	0.49	0.69	0.78	0.64	0.71	0.82	0.55	0.43	
	27	0.79	0.52	0.22	0.42	0.72	0.94	0.92	0.97	0.96	0.89	0.92	0.97	0.95	0.96	0.90	0.76	0.50	0.51	0.57	0.49	0.45	0.72	0.65	0.54	
	28	0.52	0.18	0.69	0.63	0.52	0.85	0.89	0.91	0.91	0.72	0.82	0.85	0.86	0.88	0.90	0.88	0.84	0.42	0.69	0.59	0.69	0.37	0.45	0.55	
	29	0.56	0.77	0.88	0.67	0.58	0.93	0.91	0.92	0.97	0.97	0.98	0.97	0.94	0.90	0.89	0.96	0.89	0.54	0.91	0.97	0.93	0.87	0.87	0.99	
	30	0.69	0.76	0.45	0.84	0.75	0.84	0.97	0.94	0.92	0.91	0.95	0.98	0.97	0.82	0.81	0.91	0.85	0.58	0.28	0.29	0.42	0.69	0.21	0.71	
	31	0.68	0.83	0.82	0.16	0.72	0.81	0.94	0.98	0.97	0.92	0.96	0.93	0.94	0.98	0.95	0.89	0.87	0.87	0.82	0.82	0.95	0.97	0.97	0.94	0.58

Table B.2: Correlation of hmF2 with TEC for October, 2016 from 0 UT to 23 UT.

	UT																							
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
October, 2016	1	0.02	-0.04	-0.17	-0.09	-0.58	-0.34	0.19	-0.09	0.02	-0.11	-0.16	-0.01	-0.14	-0.05	-0.12	-0.54	-0.39	-0.51	-0.74	-0.27	0.27	0.25	-0.06
	2	-0.34	0.29	-0.12	0.05	-0.60	-0.67	-0.22	-0.21	-0.15	-0.28	0.54	0.07	0.27	0.46	0.56	-0.22	-0.52	-0.46	-0.49	-0.39	-0.29	-0.49	-0.08
	3	0.02	-0.19	-0.57	-0.53	-0.58	-0.70	-0.59	-0.10	-0.22	0.27	0.44	0.10	0.17	0.33	0.46	-0.43	-0.55	-0.77	-0.66	0.09	-0.02	-0.26	-0.37
	4	-0.41	-0.22	-0.19	0.15	0.08	-0.34	-0.32	-0.01	0.15	0.17	0.32	0.19	0.53	0.54	0.15	-0.05	-0.58	-0.51	-0.81	-0.50	-0.48	-0.44	-0.57
	5	-0.37	-0.50	-0.42	0.16	-0.28	0.61	-0.53	-0.31	-0.26	-0.32	0.17	0.34	0.45	0.48	0.54	-0.08	-0.35	-0.65	-0.55	0.15	-0.30	-0.61	-0.01
	6	-0.18	-0.25	-0.27	-0.29	-0.40	-0.62	-0.40	-0.40	-0.44	-0.30	-0.26	0.06	-0.14	-0.24	-0.19	-0.37	-0.59	-0.78	-0.60	-0.70	-0.28	0.01	-0.03
	7	-0.28	-0.38	-0.43	0.58	-0.75	-0.70	-0.51	-0.33	-0.52	0.48	0.55	0.08	0.38	0.56	0.65	-0.30	-0.68	0.09	-0.69	-0.81	-0.33	-0.60	-0.72
	8	-0.40	-0.28	-0.06	-0.22	-0.40	-0.72	-0.63	-0.28	-0.08	0.47	0.54	0.20	0.44	0.59	-0.38	-0.60	-0.51	-0.26	-0.20	-0.23	-0.46	-0.57	-0.41
	9	0.45	-0.29	-0.27	0.36	-0.28	-0.41	-0.50	-0.30	0.37	-0.26	0.38	0.63	0.62	0.74	0.28	-0.49	-0.65	-0.42	-0.39	-0.45	-0.67	-0.61	-0.63
	10	-0.02	-0.23	-0.18	0.25	-0.60	-0.61	-0.66	-0.52	-0.61	-0.30	0.10	0.50	0.41	0.27	0.36	0.12	-0.50	-0.77	-0.74	-0.73	-0.13	-0.04	-0.43
	11	-0.65	-0.47	-0.08	-0.21	-0.43	-0.42	-0.59	-0.48	-0.05	-0.16	0.00	0.15	0.34	0.50	0.41	-0.23	-0.51	-0.48	-0.48	-0.14	-0.44	-0.53	-0.52
	12	-0.53	-0.50	-0.43	-0.16	-0.67	-0.70	-0.56	-0.37	-0.57	-0.63	-0.11	0.09	0.43	0.61	0.65	0.02	-0.49	-0.49	-0.52	-0.39	-0.24	0.02	-0.42
	13	-0.52	-0.40	-0.06	-0.21	-0.69	-0.61	-0.22	-0.16	-0.06	0.01	0.34	0.44	0.50	0.18	0.11	-0.38	-0.68	0.10	-0.28	-0.25	-0.29	-0.73	-0.41
	14	0.10	-0.18	0.24	-0.04	-0.36	-0.28	-0.09	0.44	0.22	0.18	0.27	-0.09	0.00	-0.06	-0.04	-0.15	-0.52	-0.37	-0.45	-0.56	-0.13	-0.63	-0.65
	15	-0.61	-0.49	-0.75	-0.53	-0.48	-0.43	-0.55	-0.14	-0.04	0.35	0.50	0.57	0.50	0.65	0.03	-0.26	-0.57	-0.51	-0.21	-0.20	-0.33	-0.36	0.01
	16	-0.67	-0.21	-0.45	-0.20	-0.23	-0.22	-0.28	-0.36	-0.32	-0.14	-0.01	-0.27	-0.15	-0.22	-0.58	-0.56	-0.63	-0.81	-0.83	-0.46	0.01	-0.21	-0.74
	17	-0.75	-0.37	-0.56	-0.48	-0.71	-0.33	-0.74	-0.65	-0.41	-0.25	-0.02	-0.45	-0.06	-0.45	-0.38	-0.70	-0.82	-0.60	-0.42	-0.66	-0.31	-0.55	-0.23
	18	-0.71	-0.75	-0.77	-0.58	0.50	-0.53	-0.41	-0.26	-0.29	-0.35	-0.48	-0.37	-0.39	-0.37	-0.67	-0.72	-0.84	-0.70	-0.62	-0.66	-0.48	-0.65	-0.53
	19	-0.78	-0.45	-0.04	0.33	-0.57	-0.55	-0.35	-0.08	-0.29	-0.31	0.09	-0.09	-0.12	-0.26	0.23	-0.62	-0.83	-0.75	-0.31	-0.31	-0.60	-0.55	-0.68
	20	-0.39	-0.53	-0.12	-0.47	-0.59	-0.63	-0.43	-0.39	-0.03	-0.34	-0.12	-0.01	-0.20	-0.10	-0.35	0.11	-0.69	-0.64	-0.44	-0.39	-0.50	-0.48	-0.72
	21	-0.38	-0.44	-0.60	-0.45	-0.46	-0.62	-0.46	-0.28	-0.59	-0.53	-0.45	-0.57	-0.48	-0.37	-0.40	-0.46	-0.51	-0.47	-0.21	-0.29	-0.42	-0.56	-0.41
	22	-0.20	-0.39	-0.40	-0.30	-0.41	-0.52	-0.40	-0.27	-0.02	-0.17	-0.01	-0.16	-0.24	-0.23	-0.25	-0.45	-0.49	-0.27	-0.04	-0.15	-0.21	-0.42	-0.39
	23	-0.01	-0.68	-0.53	-0.15	-0.48	-0.55	-0.41	-0.39	-0.37	-0.23	-0.17	-0.23	-0.34	-0.27	-0.39	-0.41	-0.39	-0.70	-0.25	-0.73	-0.58	-0.64	-0.82
	24	-0.56	-0.63	-0.68	-0.53	-0.56	-0.68	-0.63	-0.47	-0.32	-0.49	-0.25	-0.34	-0.19	-0.39	-0.33	-0.56	-0.73	-0.43	-0.71	-0.63	-0.14	-0.83	-0.61
	25	-0.74	-0.51	-0.63	-0.84	-0.67	-0.41	-0.51	-0.60	-0.44	-0.45	-0.43	-0.25	-0.45	-0.50	-0.79	-0.58	-0.74	-0.44	-0.70	-0.52	0.21	-0.65	-0.65
	26	-0.38	-0.53	-0.74	-0.78	-0.71	-0.70	-0.67	-0.43	-0.52	-0.46	-0.29	-0.22	-0.16	-0.14	-0.40	-0.62	-0.60	-0.35	-0.74	-0.44	-0.79	-0.74	-0.68
	27	-0.41	0.18	-0.68	-0.64	-0.12	-0.72	-0.51	-0.41	-0.44	-0.06	-0.17	-0.48	-0.31	-0.42	-0.53	-0.62	-0.65	-0.52	-0.70	-0.37	-0.39	-0.74	-0.49
	28	-0.72	-0.21	-0.50	-0.59	-0.59	-0.73	-0.45	-0.36	-0.46	-0.33	-0.54	-0.21	-0.26	-0.39	-0.30	-0.53	-0.44	-0.50	-0.42	-0.49	-0.71	-0.88	-0.48
	29	-0.62	-0.67	-0.33	-0.61	-0.38	-0.50	-0.65	-0.75	-0.36	-0.14	-0.28	-0.38	-0.32	-0.36	-0.65	-0.80	-0.84	-0.82	-0.91	-0.72	-0.80	-0.84	-0.95
	30	-0.66	-0.49	-0.50	-0.79	-0.69	-0.64	-0.73	-0.50	-0.32	-0.18	-0.17	0.03	-0.21	-0.16	-0.64	-0.50	-0.40	-0.54	-0.47	-0.54	-0.49	-0.42	-0.17
	31	-0.39	-0.48	-0.75	-0.76	-0.54	-0.89	-0.80	-0.63	-0.32	-0.23	-0.32	-0.13	-0.13	-0.08	-0.28	-0.36	-0.55	-0.62	-0.73	-0.63	-0.19	-0.95	-0.92

Appendix C

GPS Stations Over Europe

Table C.1: GPS Station Numbers and locations for $D_{ion-TEC}$.

Station Number	Station Code	Lat ° N	Lon ° E
1	izan	28.31	-16.5
2	lamp	35.5	12.61
3	tuc2	35.53	24.07
4	ceul	35.89	-5.31
5	alme	36.85	-2.46
6	lago	37.1	-8.67
7	huel	37.2	-6.92
8	coba	37.92	-4.72
9	milo	38.01	12.58
10	noa1	38.05	23.86
11	pat0	38.28	21.79
12	casc	38.69	-9.42
13	cace	39.48	-6.34
14	mall	39.55	2.62
15	sons	39.68	-3.96
16	borr	39.91	-0.08

Table C.1 continued from previous page

Station Number	Station Code	Lat ° N	Lon ° E
17	usal	40.33	18.11
18	teru	40.35	-1.12
19	aut1	40.57	23
20	sala	40.95	-5.5
21	gaia	41.11	-8.59
22	duth	41.14	24.92
23	bell	41.6	1.4
24	zara	41.63	-0.88
25	vala	41.7	-4.71
26	vigo	42.18	-8.81
27	creu	42.32	3.32
28	aqui	42.37	13.35
29	rioj	42.46	-2.5
30	lliv	42.48	1.97
31	untr	42.56	12.67
32	leon	42.59	-5.65
33	esco	42.69	0.98
34	elba	42.75	10.22
35	unpg	43.12	12.36
36	acor	43.36	-8.4
37	scoa	43.4	-1.68
38	cant	43.47	-3.8
39	axpv	43.49	5.33
40	tlmf	43.57	1.38
41	igmi	43.8	11.21
42	prat	43.89	11.1
43	cost	44.16	28.66
44	ktvl	44.39	33.97
45	bolg	44.5	11.36
46	msel	44.52	11.65

Table C.1 continued from previous page

Station Number	Station Code	Lat ° N	Lon ° E
47	mops	44.63	10.95
48	gari	44.68	12.25
49	puyv	45.04	3.88
50	tori	45.06	7.66
51	eglt	45.4	2.05
52	ven1	45.43	12.35
53	mdor	45.8	4.81
54	como	45.8	9.1
55	deva	45.88	22.91
56	sjdv	45.89	4.68
57	rove	45.89	11.04
58	gsr1	46.05	14.54
59	chiz	46.13	-0.41
60	oros	46.55	20.67
61	zouf	46.56	12.97
62	baca	46.56	26.91
63	autn	46.95	4.29
64	igeo	47.03	28.84
65	bscn	47.25	5.99
66	vfch	47.29	1.72
67	bute	47.48	19.06
68	pfa2	47.52	9.78
69	baia	47.65	23.56
70	sprn	47.68	16.58
71	sbg2	47.8	13.11
72	trf2	47.93	15.86
73	man2	48.02	0.16
74	linz	48.31	14.28
75	mop2	48.37	17.27
76	guip	48.44	-4.41

Table C.1 continued from previous page

Station Number	Station Code	Lat ° N	Lon ° E
77	entz	48.55	7.64
78	bbys	48.75	19.15
79	smne	48.84	2.43
80	mlvl	48.84	2.59
81	karl	49.01	8.41
82	kunz	49.11	15.2
83	vaco	49.13	13.72
84	caen	49.18	-0.46
85	smla	49.2	31.87
86	tubo	49.21	16.59
87	crei	49.26	2.51
88	ctab	49.41	14.68
89	usdl	49.43	22.59
90	cfrm	49.68	18.35
91	zywi	49.69	19.21
92	cpar	50.04	15.78
93	kraw	50.07	19.92
94	dour	50.09	4.59
95	newl	50.1	-5.54
96	crak	50.1	13.73
97	pous	50.14	12.3
98	klop	50.22	8.73
99	badh	50.23	8.61
100	kato	50.25	19.04
101	bisk	50.26	17.43
102	marj	50.36	12.89
103	eusk	50.67	6.76
104	ware	50.69	5.25
105	eijs	50.76	5.68
106	clib	50.77	15.06

Table C.1 continued from previous page

Station Number	Station Code	Lat ° N	Lon ° E
107	dent	50.93	3.4
108	dres	51.03	13.73
109	lodz	51.78	19.46
110	delf	51.99	4.39
111	bpdl	52.04	23.13
112	bogo	52.48	21.04
113	gwwl	52.74	15.21
114	hobu	53.05	10.48
115	bydg	53.13	17.99
116	dare	53.34	-2.64
117	ters	53.36	5.22
118	borj	53.58	6.67
119	swki	54.1	22.93
120	helg	54.17	7.89
121	enis	54.4	-7.64
122	redz	54.47	17.12
123	belf	54.58	-5.93
124	hoer	54.76	8.29
125	foyl	54.98	-7.34
126	smid	55.64	9.56
127	budp	55.74	12.5
128	suld	56.84	9.74
129	invr	57.49	-4.22
130	kure	58.26	22.51
131	tora	58.27	26.47
132	stas	59.02	5.6
133	toil	59.42	27.54
134	suur	59.46	24.38
135	oslo	59.74	10.37
136	pulk	59.77	30.33

Table C.1 continued from previous page

Station Number	Station Code	Lat ° N	Lon ° E
137	joen	62.39	30.1
138	vaas	62.96	21.77
139	trds	63.37	10.32
140	vil0	64.7	16.56
141	ske0	64.88	21.05
142	soda	67.42	26.39
143	vars	70.34	31.03