İSTANBUL TECHNICAL UNIVERSITY ★ INSTITUTE OF SCIENCE AND TECHNOLOGY

MODELLING AND PARAMETER ESTIMATION FOR STOCHASTIC PROCESSES SUPERIMPOSED ON PERIODIC SIGNALS

M.S. Thesis by
Eti MİZRAHİ, R.A.

(509991080)

Date of submission : 12 June 2001

Date of defence examination: 26 June 2001

Supervisor (Chairman): Prof. Dr. Ayşe H. BİLGE

Members of the Examining Committee: Assoc. Prof. Dr. Cevdet CERİT (İ.T.Ü)

Prof. Dr. Yurdanur TULUNAY (İ.T.Ü)

JUNE 2001

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ FEN BİLİMLERİ ENSTİTÜSÜ

PERİODİK SİNYALLER ÜZERİNE EKLİ STOKASTİK SÜREÇLERDE MODELLEME VE PARAMETRE TAHMİNİ

YÜKSEK LİNANS TEZİ

Araş. Gör. Eti MİZRAHİ
(509991080)

Tezin Enstitüye Verildiği Tarih : 12 Haziran 2001
Tezin Savunulduğu Tarih : 26 Haziran 2001

Tez Danışmanı : **Prof. Dr. Ayşe H. BİLGE**

Diğer Jüri Üyeleri **Doç. Dr. Cevdet CERİT (İ.T.Ü)**

Prof.Dr. Yurdanur Tulunay (İ.T.Ü)

HAZİRAN 2001

ACKNOWLEDGEMENTS

I would like to express all my sincere gratitudes to my supervisor Prof.Dr. Ayşe H.Bilge, for her guidence and her support . I would like to express my deep regards to Prof.Dr. Yurdanur Tulunay for her guidence, her help and kindness providing me the CD-Rome of COST251 which contains all the data used in this study. I would like to thank Ass.Prof. Cevdet Cerit for all his help in statistics. I also appreciate all my family, my husband Jeki, my sun Vedat, my daughter Leyla and my parents for their support. It is a pleasure to acknowledge Prof. Dr. Saadet Erbay for her support to continue my studies.

JUNE 2001 E. MİZRAHİ

CONTENTS

ABREVIATIONS

HF	: High frequency
UT	: Universal Time (as an agreement is Greenwich time)
LT	: Local Time (computed according to the longitude)
COST	: European Cooperation in the Field of Scientific and Technical Research
FFT	: Fast Fourier Transform
DFT	: Discrete Fourier Transform
Mhz	: Mega Hertz
N	: North
E	: East

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

LIST OF SYMBOLS

E	: Experiment
S	: Sample space
E	: Event
ζ	: Outcome of the experiment
P(E)	: The probability of the event
$\mathbf{F_x(x)}$	: The probability distribution function
$\mathbf{F_{xy}(x,y)}$	: The joint probabilty distribution function
f(x)	:The probability density function
σ^2	: Variance
$\mathbf{m}_n$	: The n'th moment of a random variable
$\mathbf{m_1}$	: First moment of a random variable
η	: The expected value of a random variable
μ_k	: The k'th central moment of a random variable
r	: The covariance
G(x)	: The Gaussian distribution
$\mathbf{R_{xx}}$	: The autocorrelation of a process
$\mathbf{C_{xx}}$	: The autocovariance of a process
$\mathbf{S_{xx}}$	: The Fourier Transform of the aotocorrelation
foF2	: The ionospheric critical frequency
ΔfoF2	: The deviations from monthly medians of the ionospheric critical frequency
$\mathbf{\Delta f^+}$	: The positive deviations from monthly medians of foF2

Δf	: The negative deviations from monthly medians of foF2
$(\Delta f)_c$	: Upper decile of Δf, a function of station, L.T. and R_{12}
$(\Delta f)_m$	: Upper decile of Δf computed with $(R_{12})_{median}$, as function of station and local time
R_{12}	: Solar index, an averaged sunspot number
$(R_{12})_{median}$	: The median value of R_{12}
A	: Constant of the linear model, coefficient of R_{12}
B	: Constant of the linear model, the intercept of the linear equation

ÖZET

PERİODİK SİNYALLER ÜZERİNE EKLİ STOKASTİK SÜREÇLERDE MODELLEME VE PARAMETRE TAHMİNİ

Bu çalışmada, 1958-1998 yılları arasında Avrupa'daki 20 ionosond istasyonundan alınan ionosferin F2 tabakasının kritik frekansı olan foF2'nin saatlik değerleri ile bu verilerin aylık medianlarından sapmaları (ΔfoF2) incelenmiştir. İonosferik kritik frekans verisi foF2 hem stokastik hem de deterministik değişkenlere bağlılık göstermektedir. Spektral analiz kullanılarak orijinal verinin ve aylık medyanların ana deterministik bileşenlerinin 11 aylık güneş periyodunun harmonikleri, 1 yıl ve harmonikleri, 27 günlük güneş periyodu ve ay periyodunun 2. harmoniği (12s 50.49d) olduğu gösterilmiştir. Aylık medyanların verinin deterministik kısmı için yeterli bir yaklaşım olduğu varsayılarak, aradaki fark, yani ΔfoF2, yaklaşık stokastik bir süreç olarak incelenmiştir. Aylık medyanlardan sapmaların kuvvet spektrumunun yaklaşık beyaz gürültü niteliğinde olduğu gözlenmiştir. Bu gözlem, aylık medyanların verinin deterministik kısmını temsil edebileceği varsayımını da desteklemektedir.

Ayrıca ΔfoF2'nın dağılım fonksiyonunun Gaussian dağılıma ne ölçüde uyduğu araştırılmıştır. Bu amaçla, çeşitli istasyon ve saatler için aylık medyanlardan sapmaların histogramları oluşturulmuş ve aylık medyanlardan sapmaların beklenen değerinin sıfira yakın ($\mu<0.5$) ve %10 ile %20 arası hata ile Gaussian olduğu gözlenmiştir. Standart sapmalar gündüz saatlerinde daha büyük olup 5-7 arasında değerler almaktadır. ΔfoF2'nin pozitif (Δf^+) ve negatif (Δf^-) değerlerinin mevsimsel bağımlılığı spektral analiz ile incelenmiş olup, alçak enlemlerde negatif sapmaların mevsimsel bağımlılığının pozitiflere göre daha belirgin olduğu, buna karşın yüksek enlemlerde her iki sapmanında mevsimsel bağımlılığı olduğu gözlenmiştir. Ayrıca pozitif sapmalarda ekinoks süresince modülasyonların kaybolduğu görülmüştür. Bu gözlemler sonucu, pozitif ve negatif

sapmalar ayrı ayrı incelenmiştir. Daha sağlıklı modeller elde etmek için model parametrelerinin çeşitli fiziksel etmenlere bağımlılığı araştırılmış ve ΔfoF2'nin pozitif (Δf^{+}) ve negatif (Δf^{-}) değerleri R_{12}, yerel saat ve mevsim bağımlılıkları açılarından ayrı ayrı incelenmiştir. Alçak enlemlerde pozitif sapmalar R_{12}'den bağımsız iken yüksek enlemlerde her iki sapmanın da R_{12} bağımlılığı gözlenmiştir.

Modelleme aşamasında ise, yapılacak olan model R_{12}'ye bağlı lineer bir model olacağından sadece negatif sapmalar ile çalışılmıştır. Her istasyon ve her saat için negatif sapmaların %90 güven aralığı ile R_{12}'ye bağlı lineer bir model oluşturulmuştur.

SUMMARY

MODELLING AND PARAMETER ESTIMATION FOR STOCHASTIC PROCESSES SUPERIMPOSED ON PERIODIC SIGNALS

In this work, the hourly ionospheric critical frequency denoted by foF2 and the deviations of hourly ionospheric critical frequency from monthly medians denoted by ΔfoF2, for 20 stations in Europe during the periode 1958-1998 are studied. The ionospheric critical frequency foF2 data has deterministic and stochastic components. Spectral analysis is used to show that the deterministic components (for each hour) are the harmonics of 11 year (solar cycle), 365 days and harmonics, 27 days and the second harmonic of lunar rotation period L_2 (12h 50.49m), for full data and for the deviations from monthly medians. Assuming that the monthly medians are the representation of the deterministic components of the original data, the difference between the original data and the monthly medians (ΔfoF2) is studied as an approximate stochastic process. The assumption is supported by the power spectrum of the deviations from monthly medians being almost white noise at the high end of the frequency spectrum, but with deterministic low frequency components.

It is of interest to know to what extend the deviations from monthly medians are Gaussian. For this purpose the histograms of the deviations from monthly medians for various stations and hours are obtained. It is shown that the deviations from monthly medians is nearly zero mean (mean<0.5) and approximately Gaussian with an error ranging between %10 to % 20. The standard deviations are larger for daylight hours and lie in the range 5-7. Distinctive behaviours of the positive deviations Δf^+ and negative deviations Δf^- are studied. Then qualitative observations are reported. The spectral analysis of negative and positive deviations studied separetely is used to show the seasonal dependency of negative deviations is much more visible then the seasonal

dependency of positive deviations, specially at low latitudes. At high latitudes, both positive and negative deviations have seasonal dependency. Modulations at equinox time disapear at positive deviations.

These observations lead to study negative deviations and positive deviations separetely. Thus the upperdeciles (ninth deciles) for each hour, station and year are computed. It is observed that positive deviations are almost independent from the solar index R_{12} at low latitudes. But at high latitudes it depends linearly on R_{12}. The negative deviations have a R_{12} dependency at all latitudes. Based on this results a linear model is develped only for negative deviations consisting of linear fits to Δf versus R_{12} for each station and hour, for the %90 confidence interval.

CHAPTER 1. INTRODUCTION

The aim of the present study is to apply deterministic and statistical tools to the problem of modeling and parameter estimation of a process with deterministic and stochastic components at various timescales. The critical frequency of the ionospheric F2 layer, denoted by foF2 is such a process for which reliable data is available over a timespan of about 40 years and for a large number of ionosonde stations over Europe.

The difficulty in modeling foF2 variations is that data has deterministic components at the timescales of years, months and hours. Unpredictable variations influenced by ionospheric conditions are superimposed on these periodic components. However these unpredictable variations are not completely random, because the type and the magnitude of the disturbance depends on the periodic variations above, and on other physical parameters such as the geographic location of the station.

Our strategy in developing a model is to consider the data as determinitistic at the first approximation, at the scale of years and months, determine the periodicities in the data by Fourier analysis, as presented in Chapter 3.2, and use a previously developed model **Baykal (1998)** to conclude that the monthly median values of foF2 for each hour represent the determinisitic variations within 3-4% relative error in the l_2 norm.

Then, the difference of foF2 and the monthly medians for each hour denoted by ΔfoF2, is the process that we consider stochastic and study in detail. The ΔfoF2 process considered as a time series sampled each hour, still has deterministic components again at the timescales of years, months and hours. Thus we have the option to consider it as a stochastic process whose statistical properties change in time, but the change in these properties are tied to deterministic effects. This would lead to work with nonstationary processes and previous work **Bilge and Tulunay(2000)** has shown that models based on data from time intervals over which the process is nearly stationary are more succesful. Thus we split the data to years, and consider data from each station, each year and each hour as stochastic process sampled daily. The statistical properties of these samples are analyzed, and the dependency of the statistical parameters on physical conditions are studied.

The aim of the work related to foF2 process is to develop working models to be used for forecasting foF2 within one hour. In this respect, the extreme magnitudes of the deviations are crucial and with this aim, we develop a quantitative model for the upper deciles of the negative values of ΔfoF2.

The plan of our work is as follows. In Chapter 2, we give basic definitions and properties of random variables and stochastic proceeses. In Chapter 3, we concentrate on the deterministic aspects of the data: We give a detailed description of the data, we use Fourier analysis to obtain its main periodic components, and we review a model using polynomial fits and trigonometric expansions. In Chapters 4 to 6, we concentrate on the deviations from monthly medians of foF2, i.e. ΔfoF2, which represent the stochastic components. In Chapter 4, we consider yearly data from each station for each hour, consisting of 365 daily values. We first consider each time series as a realization of some stochastic process, study its power spectrum and conclude that these samples are nearly white noise with some low frequency power. However because of deterministic effects, it is not possible to represent them as realizations of the same stochastic process. In order to categorize them, we consider each set of 365 daily values as samples of a random variable, and determine to what extend these random variables are Gaussian and how their variances depend on physical parameters. In Chapter 5, we study qualitative aspects of the dependence of the statistical properties on physical parameters, and determine the key factors by which samples can be considered as realizations of the "same" stochastic process. We conclude that, for prediction purposes, the dependency of the negative deviations from monthly medians on the sunspot numbers is the most crucial effect, and in Chapter 6, we obtain a quantitative model for the upper deciles of negative deviations from monthly medians in terms of the sunspot numbers.

CHAPTER 2. STOCHASTIC PROCESSES

In this chapter we shall give basic mathematical concepts related to stochastic processes. In Chapter 2.1 we define the probability density and distribution functions of a random variable and in Chapter 2.2 we describe stochastic processes.

2.1. Basic defininitions, random variables

The sample space

All possible outcomes of an experiment constitute a set that we call the "sample space" and denote by S. The elements of S, i.e. the outcomes are denoted by ζ. Any subset of the sample space S is known as an event and it is denoted by E. If in a certain trial the outcome ζ of the experiment is contained in the set E, then we say that the event E has occurred. Two events E1 and E2 are mutually exclusive if the intersection of the sets E1 and E2 is empty. A typical example we can consider the experiment consisting of tossing 2 distinct dies, red and green. The sample space consists of the set $\{(1,1), (1,2),\ldots,(6,6)\}$ with 36 elements, where (i,j) denotes the outcome ζ where the red die shows the number i and the green die shows the number j. We may describe the event E as those outcomes for which the sum of the dies is bigger than 10. Then E consists of the subset: $\{(5,6), (6,5), (6,6)\}$ of the sample space S, and if we toss (5,6) we say that E has occurred. The events of tossing an even number and an odd number are mutually exclusive.Probability of an event

The probability of an event is defined axiomatically. But we start by giving a relative frequency interpretation as follows. Suppose that an experiment with sample space S is repeated under exactly the same conditions n times. For an event E, we define n(E) to be the number of times the event E occurs.

Then P(E), the probability of the event E, can be defined as the limiting frequency of E

$$P(E) = \lim_{n\to\infty} \frac{n(E)}{n} \tag{2.1}$$

To give the axiomatic definition, we shall assume that for each event E we can assign a real number P(E), known as the probability of E, satisfying the following three axioms.

i) $0 \le P(E) \le 1$ (2.2)

ii) $P(S) = 1$ (2.3)

iii) For any sequence of mutually exclusive events E_1, E_2,...(that is, events for which

$E_i \cap E_j = \emptyset$ when $i \neq j$),

$$P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i) \tag{2.4}$$

Random variable

Let us consider an experiment with sample space S and outcomes ζ. To every ζ we can assign a real number $x(\zeta)$ according to some rule. This defines a function **x**, called "random variable" whose domain is the space S and whose range is R. This function can be continuous or discrete. In the example above, if the dies are fair, the random variable **x** assigns the value 1/36 to each outcome ζ. Here our random variable is discrete.

The random variables are chosen as to satisfy the following properties.

i. The set $\{\mathbf{x} \le x\}$ is an event for any real number x.

ii. The probability of the events $\{\mathbf{x} = +\infty\}$ and $\{\mathbf{x} = -\infty\}$ equals zero:

$$P\{\mathbf{x} = +\infty\} = P\{\mathbf{x} = -\infty\} = 0 \tag{2.5}$$

Probability distribution function

Given a real number x, the set $\{\mathbf{x} \le x\}$, consisting of all outcomes ζ such that $\mathbf{x}(\zeta) \le x$, is an event, by (i) above. The probability of this event $P\{\mathbf{x} \le x\}$ is a number depending on x. The function defined by

$$F_x(x) = P\{\mathbf{x} \le x\} \qquad (2.6)$$

is called the probability distribution function, or just the distribution function of the random variable. Here the distribution function will be denoted by F(x) instead of $F_x(x)$. The distribution function have the following properties.

a) $F(-\infty) = 0$ and $F(+\infty) = 1$ (2.7)

b) It is a no decreasing function of x: $F(x_1) \le F(x_2)$ for $x_1 \le x_2$

c) It is continuous from the right: $F(x^+) = F(x)$

The joint distribution function

Given two random variables **x** and **y** defined on the same sample space, with distribution functions $F_x(x)$ and $F_y(y)$, the sets $\{\mathbf{x} \le x\}$and $\{\mathbf{y} \le y\}$ are events with respective probabilities $P\{\mathbf{x} \le x\}=F_x(x)$ and $P\{\mathbf{y} \le y\}=F_y(y)$. The cartesian product of the sets $\{\mathbf{x} \le x\}\{\mathbf{y} \le y\}= \{\mathbf{x} \le x, \mathbf{y} \le y\}$, representing all outcomes ζ such that $\mathbf{x}(\zeta)\le x$ and $\mathbf{y}(\zeta) \le y$, is also an event. The probability of this event is a function of x and y, known as the joint distribution function of the random variables x and y denoted by $F_{xy}(x,y)$, i.e.

$$F_{xy}(x,y) = P\{\mathbf{x} \le x, \mathbf{y} \le y\} \qquad (2.8)$$

We shall use joint distribution functions only to define autocorrelation functions and power spectral densities. We shall give certain properties of the joint distribution functions below for completeness.

Although the joint distribution function $F_{xy}(x,y)$ is related to the distribution functions $F_x(x)$ and $F_y(y)$ of the random variables x and y respectevly, it cannot be determined solely from these functions. However knowing that $\{\mathbf{x} \leq \infty\}$ and $\{\mathbf{y} \leq \infty\}$ are the certain events, we can conclude that :

$$\{\mathbf{x} \leq x, \mathbf{y} \leq \infty\} = \{\mathbf{x} \leq x\}, \; \{\mathbf{x} \leq \infty, \mathbf{y} \leq y\} = \{\mathbf{y} \leq y\} \tag{2.9}$$

Finally, as $\{\mathbf{x} \leq \infty, \mathbf{y} \leq \infty\}$ is also the certain event,

$$F_{xy}(\infty, \infty) = 1 \tag{2.10}$$

On the other hand the events, $\{\mathbf{x} = -\infty, \mathbf{y} \leq y\}$ and $\{\mathbf{x} \leq x, \mathbf{y} = -\infty\}$ having zero probability, it is clear that

$$F_{xy}(-\infty, y) = 0 \text{ and } F_{xy}(x, -\infty) = 0 \tag{2.11}$$

Let suppose that $x_1 \leq x_2$, it is seen that;

$$\{\mathbf{x} \leq x_2, \mathbf{y} \leq y\} = \{\mathbf{x} \leq x_1, \mathbf{y} \leq y\} + \{x_1 < \mathbf{x} \leq x_2, \mathbf{y} \leq y\} \tag{2.12}$$

Therefore, $P\{\mathbf{x} \leq x_2, \mathbf{y} \leq y\} = P\{\mathbf{x} \leq x_1, \mathbf{y} \leq y\} + P\{x_1 < \mathbf{x} \leq x_2, \mathbf{y} \leq y\}$ (2.13)

From the latter inequality we conclude that:

$$P\{x_1 < \mathbf{x} \leq x_2, \mathbf{y} \leq y\} = F(x_2,y) - F(x_1,y) \geq 0 \tag{2.14}$$

Probability density function

The derivative, **if it exists,**

$$f(x) = \frac{dF(x)}{dx} \tag{2.15}$$

of the distribution function F(x) is called the probability density function of the random variable **x**. The density function is also known as frequency function.

The joint density function of two random variables **x** and **y**, assuming that the joint distribution function F(x,y) has partial derivatives of order up to two, is the second derivative of F(x,y) denoted by:

$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y} \tag{2.16}$$

Histograms

In many cases a graphic presentation of a frequency table gives concise and clear information about a frequency distribution. There are three types of graphic presentation: the histogram, the frequency polygone and the frequency curve. The sample space can be a discrete or a continuous set. If the sample space is a continuum, it can be divided into cells or intervals, and the number of occurences in each interval is called the frequency in this interval. The histogram plot is represented by bars. When the midpoints of the bars of a histogram are linked, we obtain a frequency polygon. The frequency curve is a smoothed frequency polygon. A normal probability density function is a symmetric distribution about the mean , with a frequency curve that is bell-shaped. **(Yamane, 1967).**

In practice, the division of the sample space into appropriate intervals is improtant for obtaining "smoother" histograms.

Median values of a random variable

The median of a random variable, **if it exists**, is the middle value of a frequency distribution such that the probabilities of the variable taking a value below or above it are equal. In a discrete distribution the median is the middle term if the number of terms is odd, or if the number of terms is even, the average of two middle terms, when the terms are classified in an ascending order **Borowski and Borwein (1991).** For example; the median of the set A=(3 7 9 11 19) is 9 and the median of the set B=(6 7 8 10 12 17) is (10+8)/2 = 9. Notice that the set A has 5 terms (odd) and the set B has 6 terms (even).

Expected Value of a random variable

The expected value or the mean of a random variable **x**, denoted by *E(x)* or η, is the integral, **if it exists,**

$$E(x) = \int_{-\infty}^{\infty} x f(x) dx \tag{2.17}$$

where $f(x)$ is the density of **x**. If **x** is of discrete type, taking the values x_n with probability p_n, then

$$E\{x\} = \sum_n x_n P\{x = x_n\} = \sum_n x_n p_n \tag{2.18}$$

Let us assume that in an n-time repetition of an experiment, the outcomes $\zeta_1, \zeta_2, \ldots, \zeta_n$ have been observed. For each outcome, the random variable **x** takes a numerical value such as $\mathbf{x}(\zeta_1)$, $\mathbf{x}(\zeta_2), \ldots, \mathbf{x}(\zeta_n)$. If n is sufficiently large, then the average of these numbers is approximately equal to the expected value of **x**:

$$E(x) \approx \frac{x(\varsigma_1) + x(\varsigma_2) + \ldots + x(\varsigma_n)}{n} \tag{2.19}$$

Remark: Notice that if the distribution function of a random variable is symmetric its median equals its expected value or its mean.

Two random variables $\mathbf{x_1}$ and $\mathbf{x_2}$ are said to be :

i) Independent if, $f(x_1,x_2) = f(x_1)\, f(x_2)$ (2.20)

ii) Uncorrelated if, $E\{x_1,x_2\} = E\{x_1\}\, E\{x_2\}$ (2.21)

iii) Orthogonal if, $E\{x_1,x_2\} = 0$ (2.22)

Remark: The independence of two events E1 and E2 is defined as $P(E1 \cap E2) = P(E1)P(E2)$. Then i) follows from the definitions of the distribution and density functions **(Papoulis, 1965).**

Variance

The mean η of a random variable determines the place of the center of gravity of f(x). Another important parameter is its variance or dispersion σ^2, defined by

$$\sigma^2 = E\{(x-\eta)^2\} = \int_{-\infty}^{\infty} (x-\eta)^2 f(x)dx \tag{2.23}$$

this quantity is equal the moment of inertia of the probability masses and gives some idea about their concentration near η. Its positive square root σ is called standard deviation. If **x** is of discrete type, then

$$\sigma^2 = \sum_n (x_n - \eta)^2 P\{x = x_n\} \tag{2.24}$$

We note that

$$\sigma^2 = E\{x^2 - 2x\eta + \eta^2\} = E\{x^2\} - 2\eta E\{x\} + \eta^2 = E\{x^2\} - \eta^2 \tag{2.25}$$

Which gives the important relationship

$$\sigma^2 = E\{x^2\} - (E\{x\})^2 \tag{2.26}$$

Higher order moments of a random variable

A more complete specification of the statistics of **x** is possible if one knows its moments m_n defined by

$$m_k = E\{x^k\} = \int_{-\infty}^{\infty} x^k f(x)dx \tag{2.27}$$

Clearly, $m_0=1$ *and* $m_1=\eta=E\{x\}$

The central moments of a random variable are

$$\mu_k = E\{(x-\eta)^k\} = \int_{-\infty}^{\infty}(x-\eta)^k f(x)dx \tag{2.28}$$

Note that $\mu_0=1$ $\mu_1=0$ $\mu_2=\sigma^2$

We can write the moments in terms of the central moments

$$m_k = E\{x^k\} = E\{[(x-\eta)+\eta]^k\} = E\left\{\sum_{r=0}^{k}\binom{k}{r}\eta^r(x-\eta)^{k-r}\right\} \tag{2.29}$$

$$m_k = \sum_{r=0}^{k}\binom{k}{r}\eta^r \mu_{k-r} \tag{2.30}$$

The joint moments m_{kr} of two random variables **x** and **y** are defined by

$m_{kr} = E\{x^k y^r\}$ which gives the following equality:

$$m_{kr} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^k y^r f(x,y)dxdy \tag{2.31}$$

Here the sum k + r = n is called order of the moments. The first order moments $m_{10} = \eta_x$ and $m_{01} = \eta_y$ are the expected values of **x** and **y** respectively. The point with coordinates η_x, η_y is the center of gravity of the probability masses. The joint central moments μ_{kr} are defined by the following formula,

$$\mu_{kr} = E\{(x-\eta_x)^k(y-\eta_y)^r\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\eta_x)^k(y-\eta_y)^r f(x,y)dxdy \tag{2.32}$$

which gives the variances of **x** and **y** $\mu_{20} = \sigma_x^2$ and $\mu_{02} = \sigma_y^2$ respectively.

The second central moment $\mu_{11} = E\{(x-\eta_x)(y-\eta_y)\}$ is called covariance of **x** and **y.**

$-\infty < \mu_{11} < \infty$ and its dimension is greater than 1.

The ratio

$$r = \frac{E\{(x-\eta_x)(y-\eta_y)\}}{\sqrt{E\{(x-\eta_x)^2\}E\{(y-\eta_y)^2\}}} = \frac{\mu_{11}}{\sigma_x \sigma_y} \tag{2.33}$$

is called correlation coefficient of **x** and **y**. $-1 \le r \le 1$ and it has no dimension.

Upper Decile

Any one of the nine values of a random variable that divide its distribution into ten equal parts, so that the probability of a variable having a value between one decile and the next is 1/10, is called decile. The cumulative relative frequency of the n^{th} decile is %10n. The ninth decile is the value below which %90 of the population lie **Borowski and Borwein, (1991).** In Chapter 6.2 we will use upper deciles for linear modeling which coincides with the ninth decile defined above.

l_2 norm

The norm is the length of a vector **v** in a finite dimensional vector space expressed as the square root of the sum of the squares of its components with an orthogonal basis, when these are discretes. In this work l_2 norm of the error is used in Chapter 4.3 to compute the error of Gaussian approximation.

2.2 Stochastic processes

Basic definitions

We consider an experiment specified by its outcomes ζ forming the sample space S, by events E which are certain subsets of the sample space S, and by the probabilities of these events. Assigning a time function to every outcome ζ by means of a certain rule, we obtain a family of functions for each ζ. This family is called a stochastic process. Thus a stochastic process is a function of two variables t and ζ whose domain is the set of real numbers and the sample space S, respectively. For a given t, **x**(t) is a random variable. The distribution function of $\mathbf{x}(t)$ denoted by $F(x;t)$ satisfiying the following equality is called first order distribution of the process $\mathbf{x}(t)$:

$$F(x,t) = P\{\mathbf{x}(t) \le x\} \tag{2.34}$$

The corresponding density function is:

$$f(x,t) = \frac{\partial F(x,t)}{\partial x} \tag{2.35}$$

Given two time instances t_1 and t_2, the joint distribution function of $\mathbf{x}(t_1)$ and $\mathbf{x}(t_2)$ denoted by $F(x_1, x_2;t_1, t_2)$ called also second order distribution of the process $\mathbf{x}(t)$ is:

$$F(x_1, x_2; t_1 t_2) = P\{x(t_1) \le x_1, x(t_2) \le x_2\} \tag{2.36}$$

The density function of two random variables $\mathbf{x}(t_1)$ and $\mathbf{x}(t_2)$ is given by

$$f(x_1, x_2; t_1, t_2) = \frac{\partial^2 F(x_1, x_2; t_1, t_2)}{\partial_{x_1} \partial_{x_2}} \tag{2.37}$$

The n'th order joint distribution function $F_{\underline{x}}(.,.)$ is defined as:

$$F_{\underline{x}}(x_1 ... x_n, t_1 .. t_n) = P(\underline{x}(t_1) \le x_1 .... \underline{x}(t_2) \le x_2) \tag{2.38}$$

and the n'th order joint density function $f(.,.)$ is defined as:

$$f_x(x_1 ... x_n; t_1 .. t_n) = \frac{\partial^n F_{\underline{x}}(x_1, ..., x_n; t_1, .. t_n)}{\partial_{x_1} ... \partial_{x_n}} \tag{2.39}$$

The mean $\eta(t)$ of a process $\mathbf{x}(t)$ is the expected value of the random variable $\mathbf{x}(t)$.

$$\eta(t) = E\{x(t)\} = \int_{-\infty}^{\infty} x f(x;t) dx \tag{2.40}$$

The autocorrelation $\mathbf{R_{xx}(t_1,t_2)}$ of a process $\mathbf{x}(t)$ is the joint moment of the random variables $\mathbf{x}(t_1)$ and $\mathbf{x}(t_2)$:

$$R_{xx}(t_1,t_2) = E\{x(t_1)x(t_2)\} = \int_{-\infty}^{\infty} x_1 x_2 f(x_1,x_2;t_1,t_2)dx_1 dx_2 \tag{2.41}$$

The autocorrelation function is a nonnegative definite function and this is a defining property. For real process the autocorrelation function is symmetric.

The autocovariance of $\mathbf{x(t)}$ is the covariance of the random variables $\mathbf{x}(t_1)$ and $\mathbf{x}(t_2)$

$$C_{xx}(t_1 t_2) = E\{[x(t_1)-\eta(t_1)][x(t_2)-\eta(t_2)]\} \tag{2.42}$$

From which we conclude that : $C_{xx}(t_1,t_2) = R_{xx}(t_1,t_2) - \eta(t_1)\eta(t_2)$. (2.43)

A real stochastic process $\mathbf{x}(t)$ is statistically determined if its nth order distribution functions are known.

Two processes $\mathbf{x}(t)$ and $\mathbf{y}(t)$ are said to be uncorrelated if, for any t_1 and t_2, we have $R_{xy}(t_1 t_2) = \eta_x(t_1)\,\eta_y(t_2)$ that is, if $C_{xy}(t_1,t_2) = 0$, they are called orhtogonal if $R_{xy}(t_1,t_2) = 0$. We say that the processes $\mathbf{x}(t)$ and $\mathbf{y}(t)$ are independent if the group $\mathbf{x}(t_1),\ldots,\mathbf{x}(t_n)$ is independent of the group $\mathbf{y}(t_1'),\ldots,\mathbf{y}(t_m')$ for any $t_1,\ldots,t_n,t_1',\ldots t_m'$.

Kolmogorov theorem states that the family of finite dimensional distribution functions satisfying the conditions of :

i)symmetry

$$F_x(x_1,\ldots x_n;t_1,\ldots t_n) = F_x(x_i,\ldots x_l;t_i,\ldots t_l) \tag{2.44}$$

ii)and consistency

$$F_x(x_1,\ldots,x_m,\infty\ldots\infty;t_1,\ldots,t_m,\ldots t_n) = F_x(x_1,\ldots,x_m;t_1,\ldots,t_m) \tag{2.45}$$

represents a stochastic process, and to each stochastic process we can associate a family of finite dimensional distribution functions.

A stationary process is a process whose joint distribution or density function is invariant in time. We say that a stochastic process **x**(t) is stationary in the strict sense if its statistics are not affected by a shift in the time origin. This means that the two processes **x**(t) and **x**(t+ε) have the same statistics for any ε.

A stochastic process is called strictly stationary of order n if for any t and h:

$$F_{\underline{x}}(x_1,...,x_n;t_1,...,t_n) = F_{\underline{x}}(x_1,...,x_n;t_1+h,...,t_n+h) \tag{2.46}$$

and if the above equality holds for any n.

It is said that a process **x**(t) is stationary of order k if its nth-order density function satisfies the following equality:

$$f(x_1,...,x_n;t_1,...t_n) = f(x_1,...,x_n;t_1+\varepsilon,...,t_n+\varepsilon) \tag{2.47}$$

not for any n, but only for $n \leq k$.

A stochastic process is called wide sense stationary if:

i) $E\{|\underline{x}(t)|^2\} < \infty$ for $\forall$ t (2.48)

ii) $E\{\underline{x}(t)\} = \text{constant}$ (2.49)

iii) $R_{\underline{xx}}(t_1,t_2) = R_{\underline{xx}}(t_1-t_2)$ (2.50)

Strict sense stationarity of order two implies wide sense stationarity.

Two processes $\underline{x}(t)$ and $\underline{y}(t)$ are said to be jointly stationary in the wide sense if:

$$R_{\underline{xy}}(t_1,t_2) = E\{\underline{x}(t_1)\underline{y}(t_2)\} = R_{\underline{xy}}(t_1-t_2) \tag{2.51}$$

For wide sense stationary processes, the autocorrelation function $R_{\underline{xx}}(t_1,t_2)$ is a function of (t_1-t_2) and the power spectral density can be defined as the Fourrier transform of the autocorrelation, and is denoted as $S_{\underline{xx}}(\omega)$. For real $\underline{x}(t)$, $R_{\underline{xx}}(t)$ is an even function.

A stochastic process is called respectively independent, uncorrelated and orthogonal if the pair $(\underline{x}_i,\underline{x}_j)$ of random variables with $\underline{x}_i = \underline{x}(t_i)$ and $\underline{x}_j = \underline{x}(t_j)$ are respectively independent, uncorrelated and orthogonal for arbitrary i and j.

Convergence

It is known that a sequence of numbers x_n tends to a limit x if, given $\varepsilon > 0$, we can find an integer n_o such that $|x_n-x| < \varepsilon$ for every $n > n_o$. But the convergence of a sequence $x_1,\ldots,x_n,\ldots$ of random variables is not defined as well. Here for each experimental outcome ζ we have a sequence of numbers $x_1(\zeta), x_2(\zeta),\ldots,x_n(\zeta),\ldots$, so $x_1,\ldots,x_n,\ldots$ represent a family of sequences. If every sequence converge to its corresponding outcome ζ, we can say that the family of sequences $x_1,\ldots,x_n,\ldots$ converges everywhere. But this is almost impossible. So we look for weaker convergence properties.

It is said that the sequence x_n converges to x with probability 1 if the set of outcomes ζ such that: $\lim \mathbf{x}_n(\zeta) = \mathbf{x}(\zeta)$ for $n \to \infty$, has probability equal to 1. This is written as:

$$P\{\mathbf{x}_n \to x\} = 1 \text{ for } n \to \infty \tag{2.52}$$

The sequence $\mathbf{x}_n$ tends to $\mathbf{x}$ in the mean square sense if :

$$E\{|\mathbf{x}_n-\mathbf{x}|^2\} \to 0 \quad \text{for } n \to \infty \tag{2.53}$$

The integral of a stochastic process $\mathbf{x(t)}$, denoted by $\mathbf{s}(\zeta)$ if it exists in the Riemann sense for every function $\mathbf{x}(t, \zeta)$ of the process is:

$$s = \int_a^b x(t)dt \tag{2.54}$$

When this integral does not exist for every ζ we can define $s(\zeta)$ in another way such as, if

$$\lim_{\Delta t_i \to 0} E\left\{\left[s - \sum_{i=1}^{n} x(t_i)\Delta t_i\right]^2\right\} = 0 \tag{2.55}$$

then s is defined as a mean square limit of a sum.

The stochastic process' integral

$$s = \int_a^b x(t)dt \tag{2.56}$$

being a limit of a sum, it can be written as $E\{a_1x_1 + \dots + a_nx_n\}$.

Using expected value properties we can conclude that :

$$E\{a_1x_1 + \dots + a_nx_n\} = a_1E\{x_1\} + \dots + a_nE\{x_n\} \quad \text{and} \tag{2.57}$$

$$E\{s\} = \int_a^b E\{x(t)\}dt = \int_a^b \eta(t)dt \tag{2.58}$$

The square of s can be written as a double integral:

$$s^2 = \int_a^b x(t_1)dt_1 \int_a^b x(t_2)dt_2 = \int_a^b\int_a^b x(t_1)x(t_2)dt_1dt_2 \tag{2.59}$$

The expected value $E\{s^2\}$ and the variance σ^2 of the stochastic integral s^2 can be written respectively as:

$$E\{s^2\} = \int_a^b\int_a^b E\{x(t_1)x(t_2)\}dt_1dt_2 = \int_a^b\int_a^b R(t_1,t_2)dt_1dt_2 \tag{2.60}$$

$$\sigma_s^2 = \int_a^b\int_a^b [R(t_1,t_2) - \eta(t_1)\eta(t_2)]dt_1dt_2 = \int_a^b\int_a^b C(t_1,t_2)dt_1dt_2 \tag{2.61}$$

If we apply the above equalities to a stationary process **w**(t) we obtain:

$$s = \frac{1}{2T} \int_{-T}^{T} w(t)dt \tag{2.62}$$

Here the expected value or ensemble average of **s** being E{**s**}= η its variance is:

$$\sigma_s^2 = \frac{1}{4T^2} \int_{-T}^{T} \int_{-T}^{T} C(t_1 - t_2)dt_1 dt_2 \tag{2.63}$$

However a stochastic process is time dependent, certain expected values can not be calculated as usual. Therefore we introduce the relationship of time and ensemble average. If the expected value of a stochastic process is time independent we will say that the process is at least mean value stationary. When a process is mean value stationary, then a time average **x̲** of the process may be defined as

$$\underline{x} = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} x(t)dt \tag{2.64}$$

if the limit average **x̲** of a mean value stationary process is equal to the constant ensemble average denoted by η_x.

White noise

Many random processes occuring in nature are approximately gaussian and approximately stationary and have a power spectrum which is approximately flat up to frequencies far higher than the maximum frequency at which a system is capable of significant response **Sage and Melsa (1971).** This concept will use in Section 4.3.

CHAPTER 3. DETERMINISTIC ASPECTS OF THE DATA

In this chapter we consider the ionospheric critical frequency foF2 process sampled at each station as a deterministic time function and determine the periodicities in the data. We repeat this procedure for each station hence we obtain the variation of various properties with geographic location.

In Section 3.1, we give a detailed description of the foF2 data. Then in Section 3.2 we investigate the deterministic components of the data using spectral analysis and finally in Section 3.3 we review a model based on parabolic fits and trigonometric expansions using least squares approximation to the data.

3.1 Description of the data

What is the ionosphere?

The ionosphere is the nearest part of the earth's atmosphere, to the earth. As it contains a large amount of ions and electrons, it effects the propagation of the radio waves. The ionospheric layer lie approximately from 75 km to 1-2 thousands km above the Earth. It is divided into 4 regions called D, E, F1, and F2 respectively from the lowest one to the highest. The D-Region is the region between about 75 and 95 km above the Earth in which the (relatively weak) ionization is mainly responsible for absorption of high-frequency radio waves. The E-Region is the one between about 95 and 150 km above the Earth that marks the height of the regular daytime E-layer. The ionization which is much more effective in the F2 layer (which lies above 400 km) disturbs the HF radio waves propagation.

What is the importance of the estimation of the ionospheric critical frequency?

The propagation of the electromagnetic waves used in radio, radar and navigation systems depends on the ionization rate of ionospheric layers. Radio waves below a certain frequency called "critical frequency" are reflected by the ionosphere and can be used for long-range communication. If the critical frequency is high, the available frequency band is larger. Thus the determination of the critical frequencies is important for the planning of radio communication and navigation systems. Hourly values of foF2 are recorded by vertical ionosondes, and distributed over the world. As the critical frequency has a strong dependency on the hour of the day, it is customary to work with the hourly values. Predictions of the monthly median values of foF2 for each hour are the popular and the determination of the reliability ranges below and above these predicted monthly medians have both theoretical and practical importance.

How data is processed?

In the framework of European Union Action COST 251, the critical frequency of the ionospheric layer F2 data, taken from different ionosonde stations in Europe in different years, are gathered in a CD-Rom. The available data consist of foF2 values taken from 48 stations between 1958-2000, which covers a period of 43 years. Eliminating those stations with less regular data coverage, the study is based on 20 stations between 1958-1998 (41 years). The geographical locations of these stations are shown in **Figure 3.1.1** below.

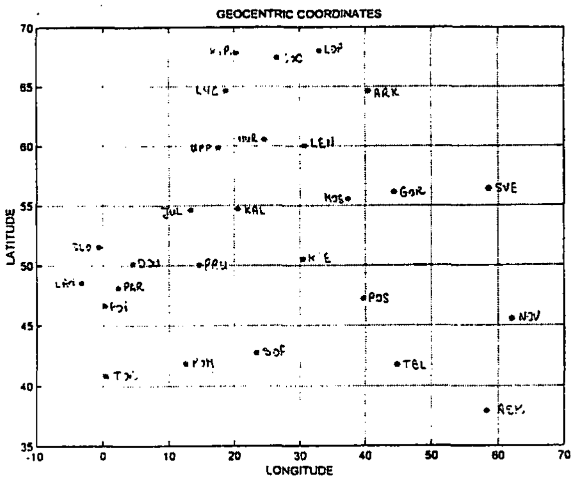

Figure 3.1.1. Representation of the geographic coordinates of 20 stations.

Since the original data foF2 has a latitude dependency, the 20 stations are arranged in 3 groups according to their latitude (from East to West). The first group, representing high latitudes, includes the stations Lycksele,Kiruna, Sodankyla, Arkhangelsk, Uppsala and Leningrad which lie nearly above 60N line. The second, mid latitude group includes the stations in the 45N-57N band, i.e. Slough, Juliusruh, Kaliningrad, Moscow, Sverdlovsk, Lannion, Poitiers, Dourbes, Pruhonice, Kiev, Novokazalinsk. The stations Tortosa, Rome and Sofia in the 40N-45N band, are arranged in a third low latitude group. **Table 3.1.1** given below is the representation of the coordinates and data coverage for these stations.

Table 3.1.1. List of stations. X means reliable data is available for that year. P stands for data from Paris-Scaley station with geographic coordinates 481N0023E, which stopped operating after 1970.

					1960										1970										1980										1990								
Code	Location	Station	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8
	Group I																																										
Code	Location	Station	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8
164	646N0188E	LYCKSELE	X					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X			X	X	X	X	X		X
167	678N0204E	KIRUNA	X	X			X		X			X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X							X	X	X	X	X	X	X
166	674N0266E	SODANKYLA	X	X	X	X	X	X	X		X	X		X	X	X	X		X			X		X			X	X			X		X	X	X	X	X	X	X	X	X	X	
163	646N0405E	ARKHANGEL												X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X						
158	598N0176E	UPPSALA	X				X	X	X	X			X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X		X	X	X		X	X	
160	600N0307E	LENINGRAD			X	X	X		X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X			X	X		X	X	X		
	Group II																																										
Code	Location	Station	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8
051	515N3594E	SLOUGH	X	X	X	X		X		X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X			
055	546N0134E	JULIUSRUH	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X		X	
154	547N0206E	KALININGRA								X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X						
155	555N0373E	MOSCOW	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X		X	X	X		
256	564N0586E	SVERDLOVSK	X	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X				
047	485N3567E	LANNION											P	P	P	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X					
046	466N0003E	POITIERS		X	X				X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X		
049	501N0046E	DOURBES	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X		X	X	X	X	X	X	X		
052	500N0146E	PRUHONICE		X		X			X	X	X	X	X	X			X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X					X	X			
151	505N0305E	KIEV							X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X			X							
246	455N0621E	NOVOKAZAL								X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X										
	Group III																																										
Code	Location	Station	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8
040	408N0005E	TORTOSA	X		X		X	X		X		X	X	X					X	X	X	X														X	X	X	X	X			
041	418N0125E	ROME	X	X	X	X	X		X	X	X	X	X	X	X	X					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
143	427N0234E	SOFIA							X	X		X	X	X	X	X	X	X	X		X					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X			X

The data from each station consist of hourly values presented in the form of 365x24 or 366x24 matrices and monthly medians of each hour, in the form of 12x24 matrices. One of the critical problems in data processing is filling missing values. This problem is partially solved by eliminating those stations and years for which data were not available or too irregularly available. As a first step the monthly median data for each year is collected and the ones for which the total number of missing values amounted to more than a month, is disregarded. Then the missing values, which were not too often, have been filled by a weighted average of the nearest neighbors. To simplify the data processing, the last day of the leap years has been disregarded. For computation purpose, cubic spline interpolation is used on 12x24 matrices of monthly medians to generate 365x24 matrices of interpolated monthly median values. This enables to replace the missing hourly values for a given day, by the corresponding value of the interpolated monthly medians. The ultimate goal of the work being the estimation of the ionospheric critical frequency foF2, the deviations of foF2 from monthly medians for each hour, denoted by

$$\Delta foF2 = foF2 - (foF2)_{median} \qquad (3.1)$$

can be considered as a first approximation to the original data foF2 and tabulated values for these medians provide a good guideline for frequency planning. Here, the spectral analysis of both the original data foF2 and the deviations from monthly medians ΔfoF2 is studied. The investigation of the statistical properties is done only for the deviations from monthly medians ΔfoF2.

To which physical facts is it related?

Ionosonde records of foF2 have been obtained since nearly 1938 at various stations over Europe. Hourly values are recorded and distributed by data banks. In very coarse terms, the variations of foF2 follow the solar sunspot activities with a period of 11 years very closely. The solar activity is quantified commonly by the "Solar Index", R_{12}, which is derived from the daily sunspot numbers. The sunspot numbers provide a well established index of solar activity, the 11-year cycles of which are accompanied by variations of the critical frequencies of the ionospheric layers. The R_{12} index is a twelve-month smoothed relative sunspot number. In **Figure 3.1.2** we see the monthly median values of foF2 for the Slough ionosonde station between 1968-1988, together with the variation of the 12-month smoothed sunspot number R_{12} **(Baykal, 1998).** A plot of foF2 versus R_{12} given in **Figure 3.1.3** shows that the dependency is nearly linear **(Baykal, 1998)**. The determination of this dependency is a crucial step in modeling.

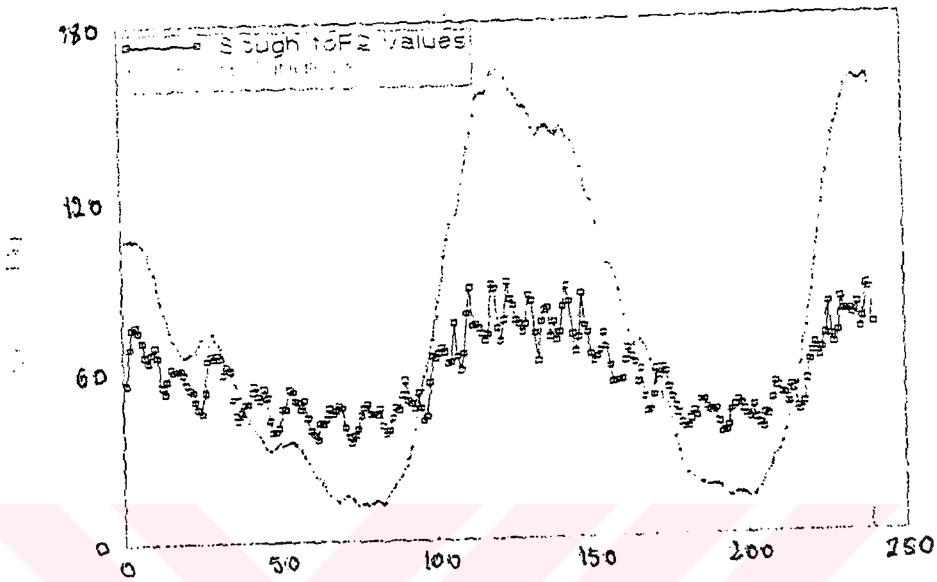

Figure 3.1.2. foF2 original data of station Slough(515N3594E) for 20 years (1970-1989), together with variations of R12 (Baykal, 1998).

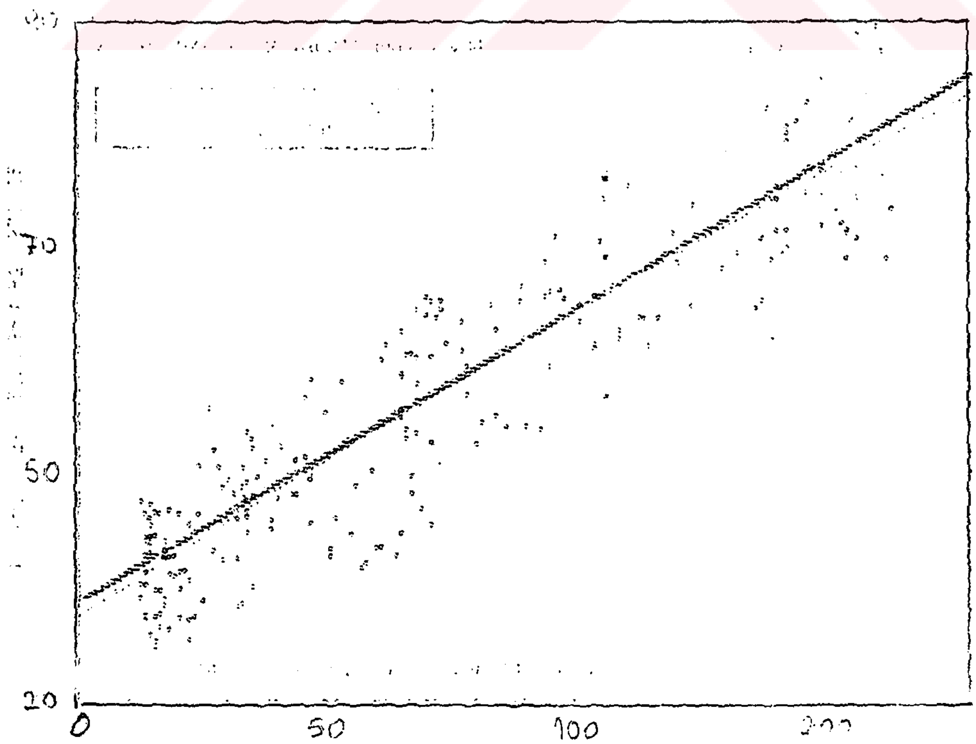

Figure 3.1.3. A plot of foF2 versus R_{12} (Baykal, 1998).

At the scale of a year, the median values of foF2 have a seasonal variation, which can be seen in Figure 3.1.2. Hourly values of foF2 for a year are shown in **Figures 3.1.4a,b** for Uppsala station, for low and high solar activity.

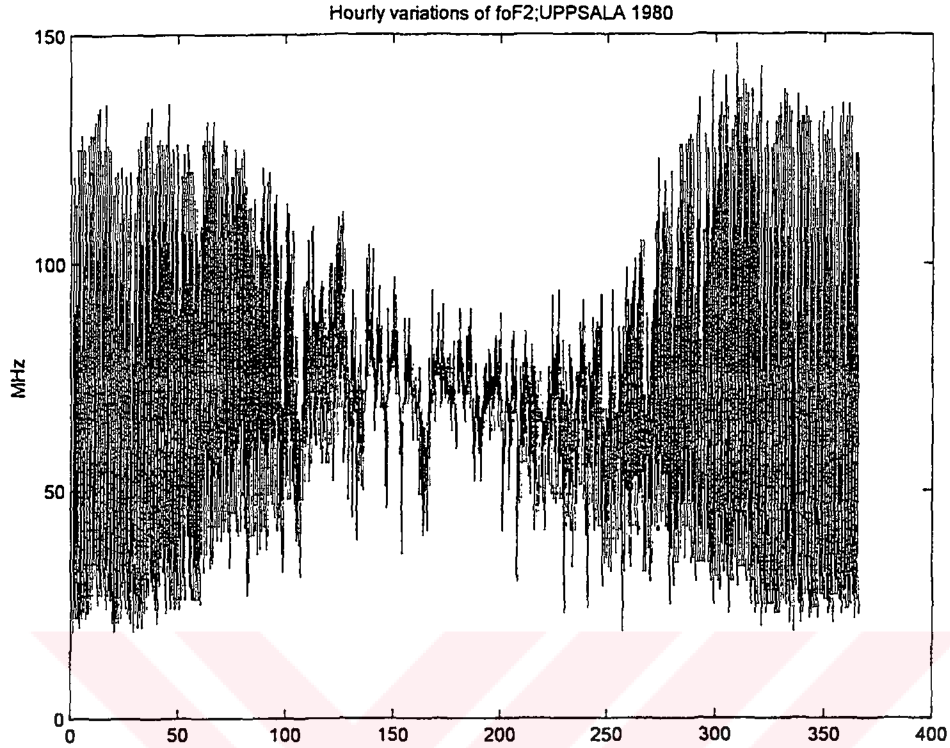

Figure 3.1.4a. foF2 original data of station UPPSALA(598N0176E) for 1980.

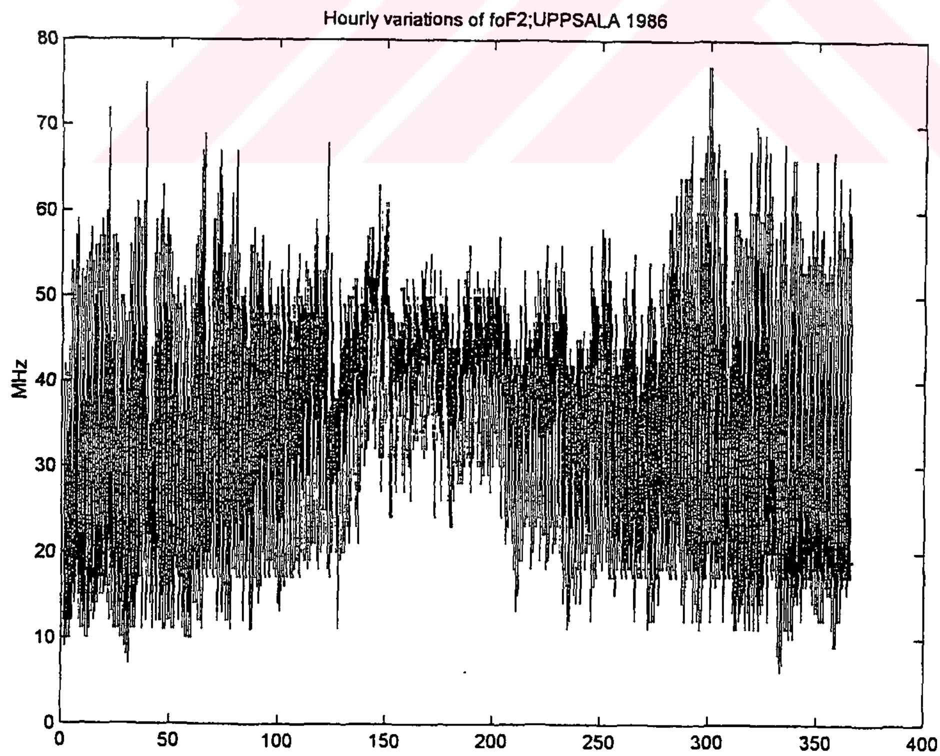

Figure 3.1.4b. foF2 original data of station UPPSALA(598N0176E) for 1986.

In addition to these regular variations, foF2 is also affected by magnetic storms. A typical storm time data is shown in **Figure 3.1.5 Davis, et al., (1996).**

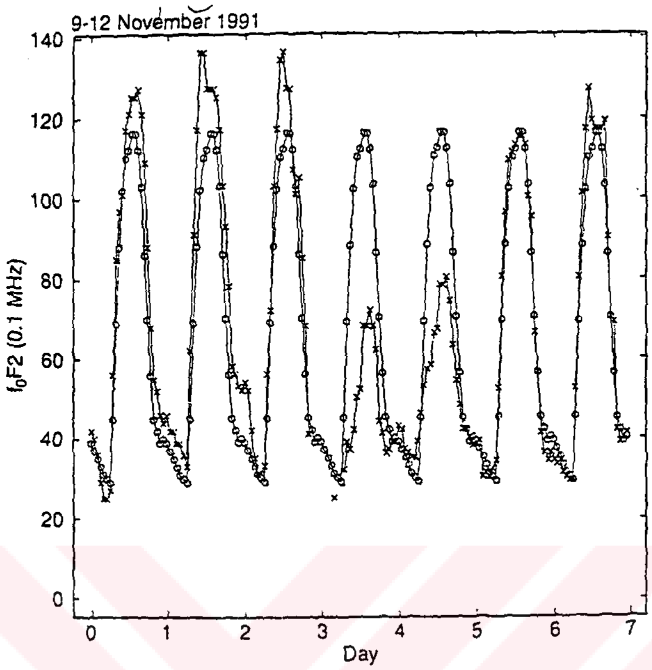

Figure 3.1.5. A typical storm-time behaviour of F2 layer critical frequency foF2 observed at Slough. The crosses are the hourly values observed on 6-12 November 1991. The open circles show the diurnal variation of the monthly median values (Davis et al., 1996)

To summarize, a simple study of the graphs shows that the critical frequency foF2 has a R_{12} dependency, seasonal dependency besides a dependence on longitude and latitude. Longitudinal dependency is eliminated by using L.T. (Local Time) instead of U.T. (Universal Time). On the other hand it is known that the magnetic storms, which occurs randomly, disturb the critical frequency foF2 propagation. All these observations and facts led to consider the critical frequency foF2 as a deterministic and stochastic process.

3.2. Spectral analysis

Theory

The Fourier transform of a function is a useful tool for determining the periodicities in the data. Also as it will be discussed in Section 4.1, the Fourier transform of a stationary stochastic process is a good approximation for the power spectral density of the process. We will describe below the Fourier transform of a function f(t) and practical problems associated with numerical computations.

Let f(t) be a function of time. The Fourier transform of f(t) is defined as the integral

$$F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt, \tag{3.2}$$

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(w) e^{i\omega t} d\omega \tag{3.3}$$

In the expression above f(t) is a continuous function defined for all t. In practice we work with a finite and sampled portion of the data. Thus we are lead to work with discrete signals of finite duration. When dealing with discrete signals we use the Discrete Fourier Transform (DFT) defined by

$$F(k) = \sum_{n=0}^{N-1} f(n) e^{-i\frac{2\pi}{N}nk} \tag{3.4}$$

When working with data sampled at say Δt intervals, the highest frequency that can be observed in the DFT spectrum corresponds to a period of 2Δt. For example, if the data is sampled at 1 hour intervals, then we can only observe variations with periods 2 hour or higher.

The Fast Fourier Transform (FFT) is an effective software routine for the computation of the Discrete Fourier Transform. The computation of the Fourier transform of time series is nowadays available in most numerical computation packages. We shall discuss below practical problems associated with the use of standard FFT routines.

Let $\{f_k\}$ be an N point sequence representing equispaced samples of a function f(t). The FFT of this series is a complex N point sequence. For the purposes of the determining periodicities in the data, we work with the absolute values of these numbers, denoted as a sequence of real numbers $\{F_k\}$. The sequence $\{F_k\}$ is symmetrical about its midvalue, hence only the first (N/2+1) terms of the sequence $\{F_k\}$ caries information. The first term of the sequence represents the mean value of the sequence $\{f_k\}$, and in practice one eliminates the mean value prior to computing FFT. In the plots of the absolute value of the FFT, the units on the horizontal axis are customarily called "bins", and the vertical axis is dimensionless. The conversion of bins to periods is done as follows.

Let N be the total number of data points sampled at Δt intervals. A peak in the absolute value of FFT at the k'th bin corresponds to a variation with period T given by

$$T=N/(k-1) \tag{3.5}$$

in units of Δt. For example, if 1 year of data are sampled daily, $\Delta t=1$ day, N=365, and a peak occurring at the k=13 corresponds to a periodicity of 365/12= 30.41 days. Conversely, if we search for a periodicity of 27 days, we should look at the nearest integers to 365/27+1=14.51. We also note that harmonics of a given periodicity appear at equidistant intervals. For example, if one year data are sampled hourly, N=365x24=8760, $\Delta t=1$ hour, and the main harmonic of the daily variation i.e. 24 hour appears at k_1= 8760/24+1=366, while the second harmonic of period 12 hours appear at k_2=8760/12+1=731, the third harmonic of period 8 hours at k_3=8760/8+1=1096, i.e. with 365 bin intervals.

In practice, the FFT spectra may have a number of unwanted peaks due to windowing effects arising from finite observation times and other numerical processing errors. Various averaging techniques have to be used in order to distinguish between weak deterministic periodicities embedded in noise and spurious peaks. In addition, it has been observed that variations with periods T such that N/T is an integer, the peaks in the FFT are quite sharp. Thus running the FFT with variable data lengths allows better resolution at certain frequencies.

Results

As an initial step in determining periodicities in the data we have computed the power spectra of foF2 and ΔfoF2 for various combinations of stations, years and hours. In all cases it is found that the dominant periodicities are the harmonics of 1 year and the first and second harmonics of 27 days. In the spectra of *foF2* the harmonics of the annual variation give rise to sharp peaks while the 27 days variation shows as a diffuse peak. The same periodicities are seen in the spectra of ΔfoF2, however the amplitude of the 27 days variation is now comparable with the amplitude of annual harmonics.

In **Figure 3.2.1** the frequency spectra for *foF2* and ΔfoF2 for Slough station is shown, to display the amplitudes of the periodicities of the order of months in *foF2* and in ΔfoF2.

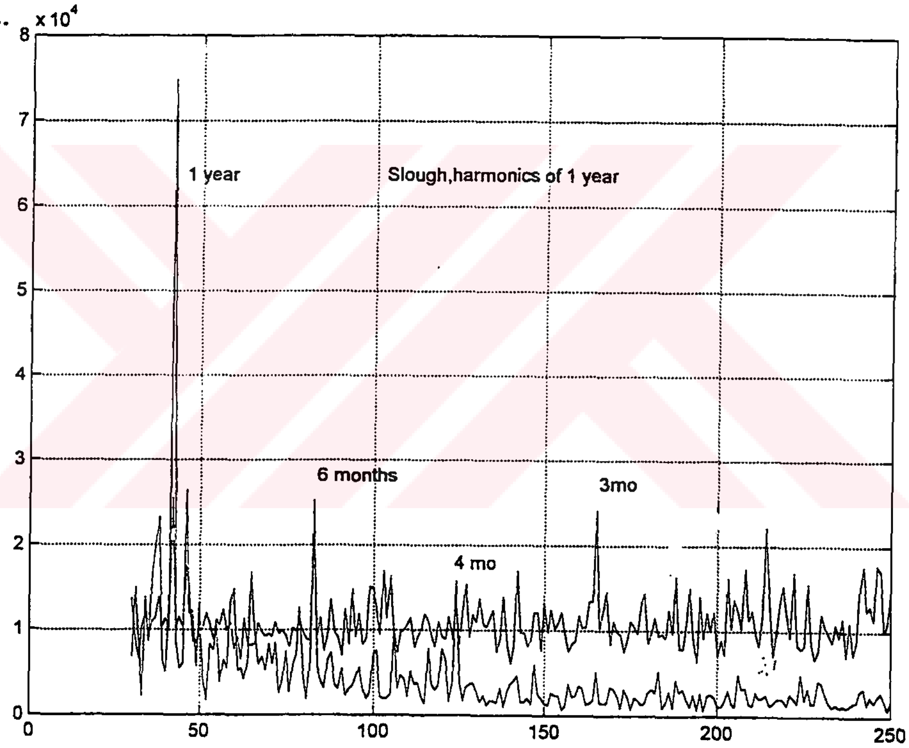

Figure 3.2.1. Lower portion of the spectrum of Slough station for all years.

Notice that in Figure 6, the upper graph displaying the spectrum of ΔfoF2 is scaled by a factor of 10. In the spectrum of *foF2* the peak corresponding to 1 year is dominant, the periodicities of 6 months and 4 months are visible. On the other hand, for ΔfoF2 the 1 year and 6 months periodicities are of the same order, 4 months periodicity is visible, in addition, there is a distinct peak corresponding to the periodicity of 3 months, which is absent in the *foF2* spectrum. The power spectra for different hours are more or less similar; the main observation is that the 27 days periodicity is seen only at day light hours. As an example, the spectra for Lycksele for 12h and 24h are compared in **Figure3.2.2.**

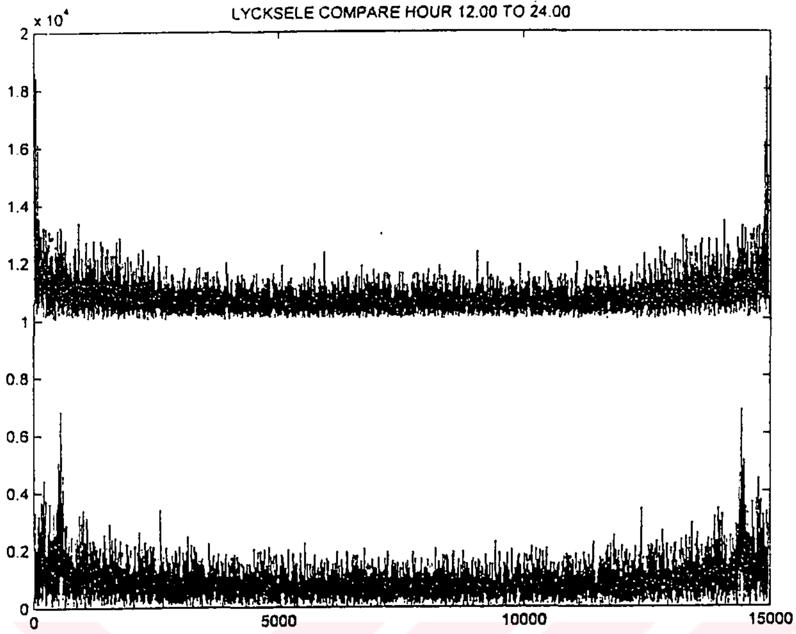

Figure 3.2.2.The power spectrum of Lyckseles' original data for all years.

Here the spectra are obtained using 41-year data, and the 27 days periodicity appears as most dominant peak. In **Figure 3.2.3a** the positive and negative deviations for Slough for all hours are arranged as a consecutive time series, in order to display the periodicity of 24 hours and harmonics.

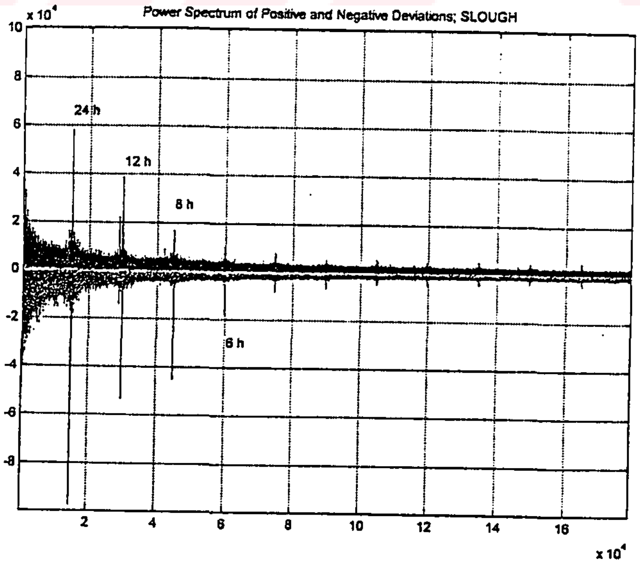

Figure 3.2.3a. A portion of the power spectrum of positive and negative deviations of Slough station for all years showing the periodicity of 24 hours and harmonics.

The periodicity seen just before the 12-hour periodicity is the second harmonic of the lunar rotation period, L_2. The spectrum of positive deviations is shown with positive amplitude and the spectrum of negative deviations is shown with negative amplitude. The zoomed graph, **Figure 3.2.3b** allows us to see the 1 year and 27 days periodicities.

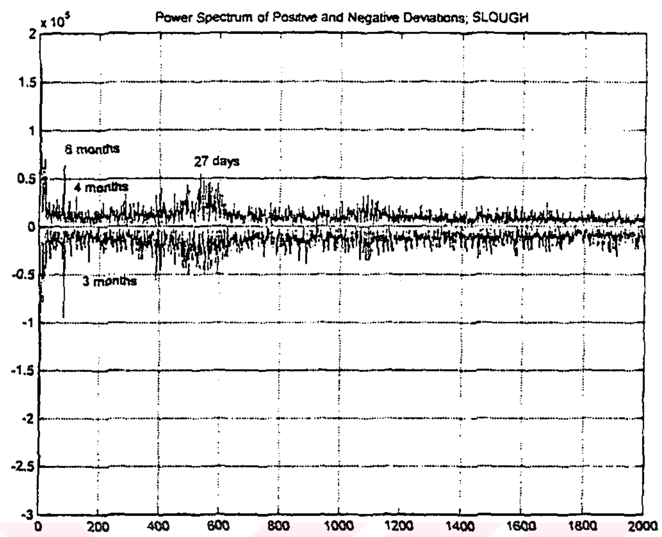

Figure 3.2.3b. Portion of the power spectrum of positive and negative deviations of Slough station.

Notice that the periodicity of 6 months has higher power for negative deviations. In **Figure 3.2.3c**, in the spectra for Slough, the 4 months periodicity can also be seen at both graphs, but the 3 months periodicity is clearly visible in negative deviations while it is absent in positive deviations. The appearance of the 3 months periodicity in negative deviations occurs for certain other stations at various latitudes, but not as sharply as in Slough. Thus further investigation is needed in order to decide on the existence of 3 months periodicity in negative deviations. This observation for Slough could not be generalized. The different character of the seasonal dependencies in positive and negative deviations can also be seen in the time domain graphs given in **Figure 3.2.4a** for Slough.

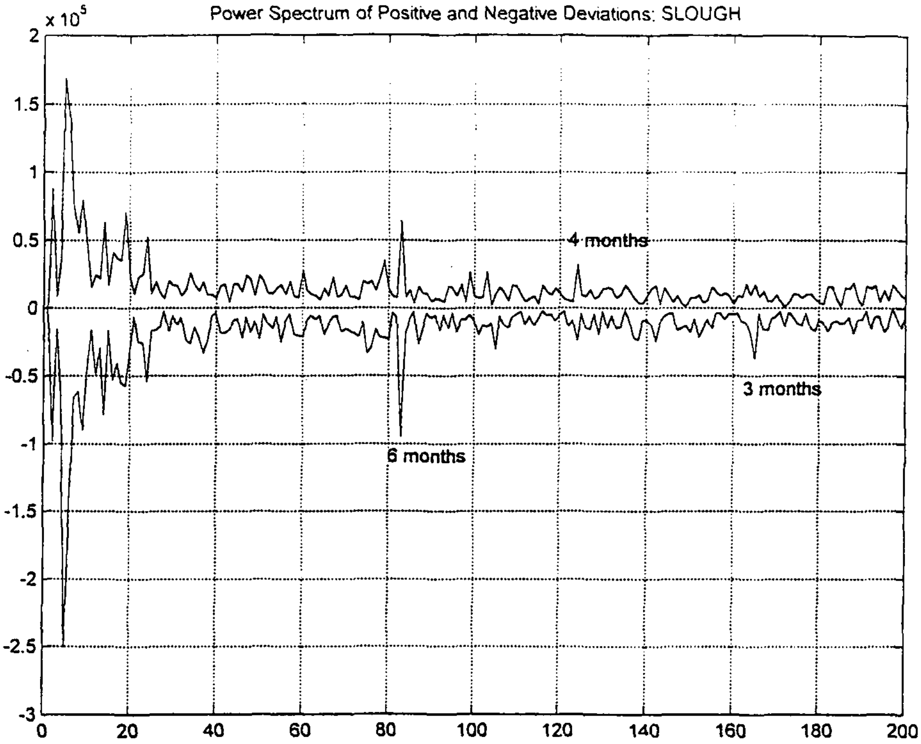

Figure 3.2.3c. Portion of the power spectrum of positive and negative deviations of Slough station for all years showing the seasonal variations.

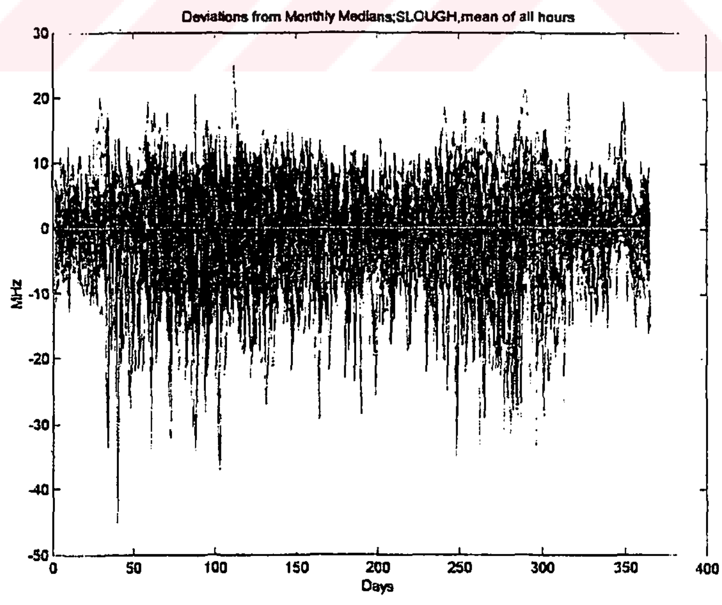

Figure 3.2.4. Time domain plots of positive and negative deviations for Slough.

3.3 Modeling monthly median foF2

Models for foF2

The long-term trend in ionosonde data of mid-latitude was investigated **(Bremer, 1992).** The linear regression analysis of monthly medians of foF2 from Juliusruh station using more than 30 years of observations has been made. Solar sunspot number R and geomagnetic activity index A_p has been used as parameter. Also another explanation of foF2 monthly medians has been made by using R_{12} solar index, which is a smoothed (averaged) sunspot number. The parabolic regression between foF2 monthly median values from Lannion station and R_{12} has been considered for each month of the year and for each hour of a day. This method finds applications of a ionospheric model for a long-term and short-term ionospheric prediction. In some models spherical harmonics has been used to explain foF2 monthly medians.

Most researchers now accept a parabolic dependency of foF2 on R_{12}. The dependencies on other solar indices may give more successful results, but the availability of an index is an important parameter in deciding to use it in a practical model. Thus, we prefer to work with models depending on R_{12} only. In modeling monthly medians, a trigonometric expansion in terms of the harmonics of one year (up to 5th harmonic) is also adopted **Bilge and Tulunay (1998).** The coefficients of these periodicities do depend however on R_{12}. Thus one works with a trigonometric expansion linearly modulated by R_{12}. Such a model is developed in **Baykal (1998)** and **Bilge and Tulunay (2000).** The determination of the best fitting trigonometric expansion is based on the so-called least squares approximation that will be discussed below.

Least Square Approximation in terms of arbitrary functions

When we use Least square approximation, we omit all stochastic assumptions and treat the estimation problem solely as a deterministic optimization problem.

The line that minimizes the sum of the squares of the derivations is chosen, to fit a line to no collinear points. The first assumption is that a linear equation will fit the data points. But furthermore it is noticed that the use of matrix notation satisfies every function (trigonometric, quadratic etc.). The final assumption is:

$$Y = AX \tag{3.6}$$

Here **Y** is a vector of which components are the data points. **A** is a matrix where the components are the chosen functions, and **X** is also a vector of which components are unknown. If all these equations are used correctly, errors may tend to average out. Since no one **x** can satisfy all the simultaneous equations, it is inappropriate to write the equality

$$Y = AX \tag{3.7}$$

Rather, an error vector **e** is introduced, such as: $e = Y - AX$ (3.8)

The least square approach yields the one **x** which minimizes the sum of the squares of the e_i components. That is, **x** is chosen to minimize.

$$\begin{aligned} \|e\|^2 &= e^T e = (Y - AX)^t (Y - AX) \\ \|e\|^2 &= Y^t Y - (AX)^t Y - Y^t AX + (AX)^t AX \\ \|e\|^2 &= \sum_{i=1}^{n} Y_i^2 - 2 \sum_{i=1, j=1}^{m,n} Y_i A_{ij} X_j + \sum_{i=1, j=1, k=1}^{n,m,h} (A_{ij} X_j)^T (A_{ik} X_k) \end{aligned} \tag{3.9}$$

Let's define

$$\|e\|^2 = E \tag{3.10}$$

$$E=\begin{pmatrix}Y_1^2 & Y_2^2 & \cdots & Y_n^2\end{pmatrix}\begin{pmatrix}1\\1\\\vdots\\1\end{pmatrix}-2\begin{pmatrix}Y_1 & \cdots & Y_n\end{pmatrix}\begin{pmatrix}a_{11} & \cdots & a_{1m}\\ \vdots & \ddots & \vdots\\ a_{n1} & \cdots & a_{nm}\end{pmatrix}\begin{pmatrix}X_1\\ \vdots\\ X_m\end{pmatrix}+\begin{pmatrix}X_1 & \cdots & X_m\end{pmatrix}\begin{pmatrix}a_{11} & \cdots & a_{n1}\\ \vdots & \ddots & \vdots\\ a_{1m} & \cdots & a_{nm}\end{pmatrix}\begin{pmatrix}a_{11} & \cdots & a_{1h}\\ \vdots & \ddots & \vdots\\ a_{n1} & \cdots & a_{nh}\end{pmatrix}\begin{pmatrix}X_1\\ \vdots\\ X_h\end{pmatrix}$$

(3.11)

$$\frac{\partial E}{\partial X_{io}}=-2\begin{pmatrix}Y_1 & \cdots & Y_n\end{pmatrix}\begin{pmatrix}a_{1_{io}}\\ \vdots\\ a_{n_{io}}\end{pmatrix}+\begin{pmatrix}a_{1_{io}} & \cdots & a_{n_{io}}\end{pmatrix}\begin{pmatrix}a_{11} & \cdots & a_{1h}\\ \vdots & \ddots & \vdots\\ a_{n1} & \cdots & a_{nh}\end{pmatrix}\begin{pmatrix}X_1\\ \vdots\\ X_h\end{pmatrix}+\begin{pmatrix}X_1 & \cdots & X_m\end{pmatrix}\begin{pmatrix}a_{11} & \cdots & a_{1h}\\ \vdots & \ddots & \vdots\\ a_{1n} & \cdots & a_{nh}\end{pmatrix}=0$$

(3.12)

Let's define

$$\begin{pmatrix}a_{1_{io}}\\ \vdots\\ a_{n_{io}}\end{pmatrix}=\sum_{i=1}^{n}A_{i_{io}} \qquad \text{and} \qquad \begin{pmatrix}a_{1_{io}} & \cdots & a_{n_{io}}\end{pmatrix}=\sum_{i=1}^{n}\left(A_{i_{io}}\right)^T \tag{3.13}$$

$$\frac{\partial E}{\partial X_{io}}=-2\sum_{i=1}^{n}Y_iA_{i_{io}}+\sum_{i=1,k=1}^{n,h}\left(A_{i_{io}}\right)^T A_{ik}X_k+\sum_{i=1,j=1}^{n,m}\left(A_{ij}X_j\right)^T A_{i_{io}} \tag{3.14}$$

We know that

$$\sum_{i=1,k=1}^{n,h}\left(A_{i_{io}}\right)^T A_{ik}X_k=\sum_{i=1,j=1}^{n,m}\left(A_{ij}X_j\right)^T A_{i_{io}} \tag{3.15}$$

So we have

$$\frac{\partial E}{\partial X_{io}}=-2\sum_{i=1}^{n}Y_iA_{i_{io}}+2\sum_{i=1,j=1}^{n,m}\left(A_{i_{io}}\right)^T A_{ij}X_j=0 \tag{3.16}$$

$$\frac{\partial E}{\partial X_{io}} = -2\begin{pmatrix} Y_1 & \cdots & Y_n \end{pmatrix}\begin{pmatrix} a_{1_{io}} \\ \vdots \\ a_{n_{io}} \end{pmatrix} + 2\begin{pmatrix} a_{1_{io}} & \cdots & a_{n_{io}} \end{pmatrix}\begin{pmatrix} a_{11} & \cdots & a_{1h} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nh} \end{pmatrix}\begin{pmatrix} X_1 \\ \vdots \\ X_h \end{pmatrix} = 0$$

(3.17)

Therefore we obtain

$$\Rightarrow \begin{pmatrix} a_{1_{io}} & \cdots & a_{n_{io}} \end{pmatrix}\begin{pmatrix} a_{11} & \cdots & a_{1h} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nh} \end{pmatrix}\begin{pmatrix} X_1 \\ \vdots \\ X_h \end{pmatrix} = \begin{pmatrix} Y_1 & \cdots & Y_n \end{pmatrix}\begin{pmatrix} a_{1_{io}} \\ \vdots \\ a_{n_{io}} \end{pmatrix} \quad (3.18)$$

$$\Rightarrow (A_{io})^T AX = Y^T A_{io} = (A_{io})^T Y \quad (3.19)$$

$$\Rightarrow X = \left(A_{io}^T A\right)^{-1} A_{io}^T Y \quad (3.20)$$

Since $Y^{\sim} = AX$ (3.21)

We finally obtain

$$Y^{\sim} = A(A_{io}^T A)^{-1} A_{io}^T Y \quad (3.22)$$

Results

The model developed in **Bilge and Tulunay (2000)** is used for preliminary investigations. It has been observed that the monthly medians for each hour can be predicted within of a 3-4% in the l_2 norm. These figures are quite satisfactory for prediction purposes. In the following chapters, we shall work with the deviations from monthly medians.

CHAPTER 4. STOCHASTIC ASPECTS OF THE DATA

4.1 Correlation and power spectrum

Correlation and Power spectrum of stationary processes

It is known that the second order moment of a stationary process **x**(t) is its autocorrelation denoted by R(τ) which satisfies:

$$R(\tau) = E\{x(t+\tau)x(t)\} = R_x(\tau) = R_{xx}(\tau) \tag{4.1}$$

The above equality implies that $R(0) = E\{|x(t)|^2\} \geq 0$ (4.2)

If **x**(t) is real, then R(τ) is real and even. Hence $R(-\tau) = R(\tau)$ (4.3)

Here our study is based on real **x**(t). The joint second moment

$$R_{xy}(\tau) = E\{x(t+\tau)y(t)\} = R_{yx}(-\tau) \tag{4.4}$$

of two jointly stationary processes is their cross-correlation. From the above definitions we can define the autocorrelation of the sum z(t) = x(t) + y(t) as

$$R_{zz}(\tau) = R_{xx}(\tau) + R_{yy}(\tau) + R_{xy}(\tau) + R_{yx}(\tau) \tag{4.5}$$

The autocorrelation of the product **x**(t)**y**(t) cannot be expressed in terms of second order moments of the given process with the exception that **x**(t) and **y**(t) are independent. In this case the random variables **x**(t + τ) and **x**(t) are independent of **y**(t + τ) and **y**(t). Hence

$$E\{x(t+\tau)y(t+\tau)x(t)y(t)\} = E\{x(t+\tau)x(t)\}E\{y(t+\tau)y(t)\} \tag{4.6}$$

Therefore

$$R_{ww}(\tau) = R_{xx}(\tau)R_{yy}(\tau) \quad \text{where } w(t) = x(t)y(t) \tag{4.7}$$

Having a real **x**(t) and knowing that $R(0) \geq 0$ we conclude that:

$$E\{[x(t+\tau) \pm x(t)]^2\} = 2[R(0) \pm R(\tau)] \tag{4.8}$$

The left-hand side of the above equality being nonnegative; we obtain

$R(0) \pm R(\tau) > 0$ which implies $-R(0) \leq R(\tau) \leq R(0)$. Thus we can say that $R(\tau)$ is maximum at the origin: $|R(\tau)| \leq R(0)$.

If we have two real processes **x**(t) and **y**(t) with a real constant **a**, we have :

$$E\{[x(t+\tau) + ay(t)]^2\} = R_{xx}(0) + 2aR_{xy}(\tau) + a^2R_{yy}(0) \tag{4.9}$$

This one being nonnegative for any **a**, its discriminant is nonpositive. Therefore

$$R_{xy}^2(\tau) \leq R_{xx}(0)R_{yy}(0) \tag{4.10}$$

Since the geometric mean of two numbers does not exceed their arithmetic mean, we also have: $2|R_{xy}(\tau)| \leq R_{xx}(0) + R_{yy}(0)$ (4.11)

Power spectrum

The power spectrum $S(\omega)$ or $S_x(\omega)$ or $S_{xx}(\omega)$ of a process **x**(t) is the Fourrier transform of its autocorrelation:

$$S(\omega) = \int_{-\infty}^{\infty} e^{-i\omega t} R(\tau)d\tau \tag{4.12}$$

where i is the imaginary i unit. $R(\tau)$ being an even function we conclude that $S(\omega)$ is a real function. The Fourier inversion formula allows us to compute $R(\tau)$ in terms of $S(\omega)$:

$$R(\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} S(\omega)e^{j\omega\tau}d\omega \tag{4.13}$$

With $\tau = 0$, the above yields

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} S(\omega)d\omega = R(0) = E\{|x(t)|^2\} \geq 0 \tag{4.14}$$

Thus, the total area of $S(\omega)/2\pi$ is nonnegative and equals the "average power" of the process $\mathbf{x}(t)$. If the process $\mathbf{x}(t)$ is real, then $R(\tau)$ is real and even. Therefore $S(\omega)$ is also even: $S(-\omega) = S(\omega)$. In this case we have:

$$S(\omega) = \int_{-\infty}^{\infty} R(\tau)\cos\omega\tau d\tau \tag{4.15}$$

$$R(\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} S(\omega)\cos\omega\tau d\omega \tag{4.16}$$

The power spectrum $S(\omega)$ of a process $\mathbf{x}(t)$ can be expressed in terms of its second-order density $f(x_1,x_2;\tau)$. Therefore, introducing the Fourrier transform of $f(x_1,x_2;\tau)$ with respect to τ

$$G(x_1,x_2;\omega) = \int_{-\infty}^{\infty} f(x_1,x_2;\tau)e^{-j\omega\tau}d\tau \tag{4.17}$$

and knowing that:

$$R(\tau) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1x_2 f(x_1x_2;\tau)dx_1dx_2 \tag{4.18}$$

we obtain

$$S(\omega) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1x_2 \int_{-\infty}^{\infty} f(x_1,x_2;\tau)e^{-j\omega\tau}d\tau dx_1dx_2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1x_2 G(x_1,x_2;\omega)dx_1dx_2 \tag{4.19}$$

4.2 Gaussian process

Gaussian distributions

It is known that **Papoulis (1965)** given any function G(x) such that:

$$G(-\infty) = 0, \quad G(+\infty) = 1 \text{ and} \tag{4.20}$$

$$\lim_{x \to \infty} G(x) = 1 \tag{4.21}$$

$G(x_1) \le G(x_2)$ if $x_1<x_2$ and $G(x^+) = G(x)$, we can find an experiment E and a random variable **x** defined on E such that its distribution function F(x) equals the given function G(x). Furthermore, we can also determine a random variable having as density a given function g(x), provided $g(x) \ge 0$

$$\int_{-\infty}^{+\infty} g(x)dx = 1 \quad \text{and} \quad G(x) = \int_{-\infty}^{x} g(y)dy \tag{4.22}$$

Among the distribution functions the most well known and widely used one is the Gaussian or "normal" distribution. We say that a random variable is normally distributed if its density function is a Gaussian curve such as

$$f(x) = Ae^{-\alpha x^2}, \alpha > 0 \tag{4.23}$$

Since

$$\int_{-\infty}^{\infty} e^{-\alpha x^2} dx = \sqrt{\frac{\pi}{\alpha}} \tag{4.24}$$

The probability density function of the Gaussian distribution is:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/2\sigma^2} \qquad -\infty < x < \infty, \quad \sigma > 0 \tag{4.25}$$

Two random variables **x** and **y** are said to be jointly normal if their density is:

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}}e^{-\frac{1}{2(1-r^2)}\left[\frac{(x-\eta_1)^2}{\sigma_1^2}-\frac{2r(x-\eta_1)(y-\eta_2)}{\sigma_1\sigma_2}+\frac{(y-\eta_2)^2}{\sigma_2^2}\right]} \tag{4.26}$$

Thus the joint density function of two jointly normal random variables **x** and **y** with zero mean $E\{x\} = \eta_x = E\{y\} = \eta_y = 0$ is:

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-r^2}}e^{-\frac{1}{2(1-r^2)}\left(\frac{x^2}{\sigma_1^2}-\frac{2rxy}{\sigma_1\sigma_2}+\frac{y^2}{\sigma_2^2}\right)} \tag{4.27}$$

Where, $E\{\mathbf{x}^2\} = \sigma_1^2$ and $E\{\mathbf{y}^2\} = \sigma_2^2$ and the parameter **r** is the correlation coefficient. It is known that **Papoulis (1965)** two jointly random variables **x** and **y** being uncorrelated, their covariance and correlation coefficient are zero. And if the correlation coefficient of two jointly random variables is zero, they are independent.

Properties of Gaussian processes

i) From the latter statement we have one of the properties of the Gaussian processes which is; if two jointly normal random variables **x** and **y** are uncorrelated, then they are independent.

We call a stationary process x(t) in the strict sense, a process that its statistics are not affected by a shift in the time origin.

This means that the two processes x(t) and $x(t + \varepsilon)$ have the same statistics for any ε. And we call a stationary process x(t) in the wide sense, a process that its expected value is a constant and its autocorrelation depends only on $t_1 - t_2$:

$E\{x(t)\} = \eta = \text{constant}$ $E\{x(t+\tau)x(t)\} = R(\tau)$, (explained in details in section 2.2)

Clearly, if x(t) is stationary of order two, then it is stationary in the wide sense. Wide sense stationarity involves only first and second order moments.

ii) Therefore if a process x(t) is normal and stationary in the wide sense, then it is stationary also in the strict sense. This follows from another property of the normal distribution mentioned below.

iii) The statistics of a normal process are uniquely determined in terms of its mean and autocorrelation.

iv) Let **z** = a**x** +b**y** and **w** = c**x** + d**y** be the linear transformations of two jointly normal random variables **x** and **y**, then **z** and **w** are also jointly normal. The converse is true for if **x** and **y** are independent.

Chebychev's inequality

Let η and σ^2 be the expected value and the variance of the random variable **x**, and ε a positive number **Cerit C., (2000).** We maintain that

$$P(|X-\eta| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2} \quad \text{or} \quad P(|X-\eta| \prec \varepsilon) \geq 1-\frac{\sigma^2}{\varepsilon^2} \tag{4.28}$$

The Law of Large Numbers

Let ε be an experiment and let A be an event associated with ε. Consider n independent repetitions of ε, let n(A) be the number of times A occurs among the n repetitions in a binomial distribution function, let f (A)= n(A)/n be the relative frequency and let P(A) = p be the occurrence probability of A. Knowing that the expected value of a binomially distributed random variable E(n(A)) = np and its variance V(n(A)) = np(1-p), for every positive number ε, we have:

$$P[|f(A)-p| \geq \varepsilon] \leq \frac{p(1-p)}{n\varepsilon^2} \qquad \text{(where 1-p = q)} \tag{4.29}$$

or, equivalently,

$$P[|f(A)-p| < \varepsilon] \geq 1-\frac{p(1-p)}{n\varepsilon^2} \tag{4.30}$$

Then for sufficiently large n we obtain:

$$\lim_{n\to\infty} P\left[\left|f(A)-p\right|<\varepsilon\right]=1 \qquad \lim_{n\to\infty} P\left[\left|f(A)-p\right|\geq\varepsilon\right]=0 \tag{4.31}$$

Thus f(A) tends to p in probability.

Characteristic functions

The characteristic function of a random variable **x** is the Fourier transform of its density function f(x) (with a reversal in sign).

$$\phi(\omega)=E\{e^{i\omega x}\} \tag{4.32}$$

This is the expected value of the complex function $e^{i\omega x}=\cos wx+i\sin\omega x$ of **x**, and it is given by the integral, if **x** is continuous,

$$\phi(\omega)=\int_{-\infty}^{\infty} e^{iwx} f(x)dx \tag{4.33}$$

and if **x** is of discrete type, taking the values $\mathbf{x_k}$,

$$\phi(\omega)=\sum_k e^{i\omega x_k} P\{x=x_k\} \tag{4.34}$$

The central-limit theorem

Consider a sequence $x_1,\ldots,x_n,\ldots$ of independent random variables with the same distribution function. Let η and σ^2 be the mean and variance of the random variable **x** and $\sigma^2\neq 0$. The random variable

$$S_n^{*}=\frac{S_n-n\eta}{\sigma\sqrt{n}} \tag{4.35}$$

has a normal distribution function **Cerit, C.,(2000).**

Proof: Here our aim is to show that for large **n** the characteristic function $\phi_{S_n^*}$ of S_n^* tends to the characteristic function of a standard normal random variable, $e^{-\frac{1}{2}t^2}$.

$$\phi_{S_n^*}(t) = \phi_{S_n}\left(\frac{t}{\sigma\sqrt{n}}\right)\exp\left(\frac{-in\eta t}{\sigma\sqrt{n}}\right) \tag{4.36}$$

$$\phi_{S_n^*}(t) = \left[\phi_{x_1}\left(\frac{t}{\sigma\sqrt{n}}\right)\right]^n \exp\left(\frac{-in\eta t}{\sigma\sqrt{n}}\right) \tag{4.37}$$

$$\phi_{S_n^*}(t) = \exp\left[n\ln\phi_{x_1}\left(\frac{t}{\sigma\sqrt{n}}\right) - i\eta n\left(\frac{t}{\sigma\sqrt{n}}\right)\right] \tag{4.38}$$

Hence we have to prove that :

$$\lim_{n\to\infty} n\left[\ln\phi_{x_1}\left(\frac{t}{\sigma\sqrt{n}}\right) - i\eta\left(\frac{t}{\sigma\sqrt{n}}\right)\right] = -\frac{t^2}{2} \tag{4.39}$$

i) for t = 0 both size of the equality vanishes and the equality holds.

ii) for $t \neq 0$, $$\frac{t^2}{\sigma^2}\lim_{n\to\infty}\frac{\ln\phi_{x_1}\left(\frac{t}{\sigma\sqrt{n}}\right) - i\eta\left(\frac{t}{\sigma\sqrt{n}}\right)}{\left(\frac{t}{\sigma\sqrt{n}}\right)^2} = -\frac{t^2}{2} \tag{4.40}$$

So we have $$\lim_{n\to\infty}\phi_{S_n^*}(t) = e^{-t^2/2}, \qquad -\infty \prec x \prec \infty \tag{4.41}$$

Where $e^{-t^2/2}$ is the characteristic function of the standard normal distribution function. Therefore the distribution function of the random variable

$$S_n^* = \frac{S_n - n\eta}{\sigma\sqrt{n}} \tag{4.42}$$

is a standard normal distribution.

The properties of Gaussian processes and the central-limit theorem, which proves that many of the distribution functions can be expressed in terms of the normal distribution function, allow us to investigate, in our study, the resemblance to the normal distribution function.

4.3 Statistical aspects of the data

Applications to foF2 testing statistical properties gaussian test

The first step in studying variability is to see to what extend these variations are "random". It is mentioned in the Preliminaries, that given a large number of observations, the histogram plot gives a good estimate of the probability density function. Most estimation techniques are designed to be applied to Gaussian white noise. Thus it is of interest to determine to what extend the deviations from monthly medians are Gaussian. For this purpose the histograms of the deviations from monthly medians for various stations and hours are obtained. In all cases the deviations were almost zero mean (all means were less than 1MHz), thus the means and medians do not differ much. We used least squares fit to find the standard deviation of the best Gaussian approximation and the l_2 norm of the error. The results are listed in **Table4.3.1**.

Table 4.3.1. Comparison between the histogram of the deviations from monthly medians and gaussian function with zero mean and appropriate sigma.

STATIONS	YEAR	SIGMA	ERROR
GIBILMANA	1979	0.737	0.2053
LISBONNE	1990	0.788	0.1249
TORTOSA	1965	0.407	0.8120
ROME	1970	0.661	0.1180
EL ARENOSILLO	1976	0.435	0.1282
ROME	1981	0.742	0.1660
GARCHY	1970	0.550	0.1523
POITIER	1968	0.563	0.1523
POITIER	1976	0.454	0.0960
POITIER	1983	0.660	0.1287
LANNION	1973	0.539	0.1633
FREIBURG	1958	0.750	0.1449
FREIBURG	1969	0.675	0.2210
DOURBES	1960	0.890	0.1148
DOURBES	1986	0.400	0.203
SLOUGH	1972	0.622	0.1505
SLOUGH	1992	0.680	0.1741
SOFIA	1983	0.688	0.1200
SIMFEROPOL	1958	0.667	0.2658
GORKY	1970	0.628	0.1500
UPPSALA	1963	0.339	0.2320
MOSCOW	1996	0.415	0.3250
KALININGRAG	1991	0.970	0.1390
KIEV	1968	0.562	0.1420
SVERDLOVSK	1976	0.413	0.2100

It is noted that however there is a systematic bias between the actual histograms and the approximating Gaussians, as seen in **Figure 4.3.1a-d.**

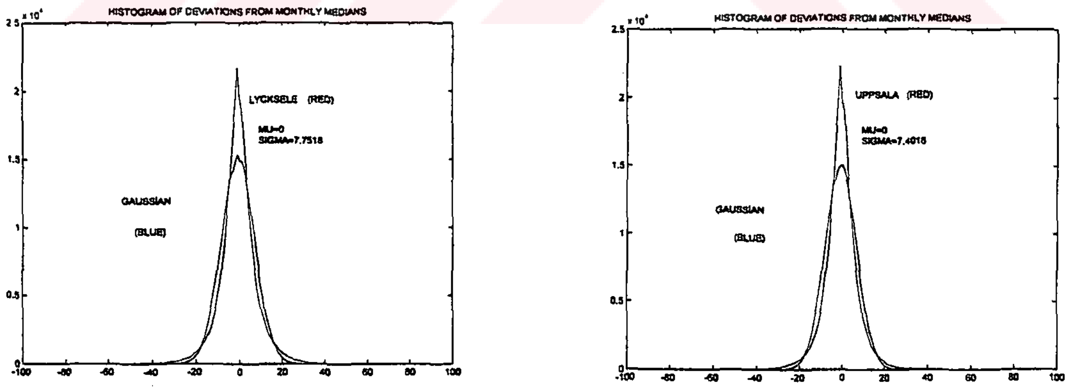

Figure 4.3.1a-b. Comparison between the histogram of the deviations from monthly medians and gaussian function with zero mean and appropriate sigma for stations a) Lycksele (over 60N), b) Uppsala (55N-60N) as high latitude stations.

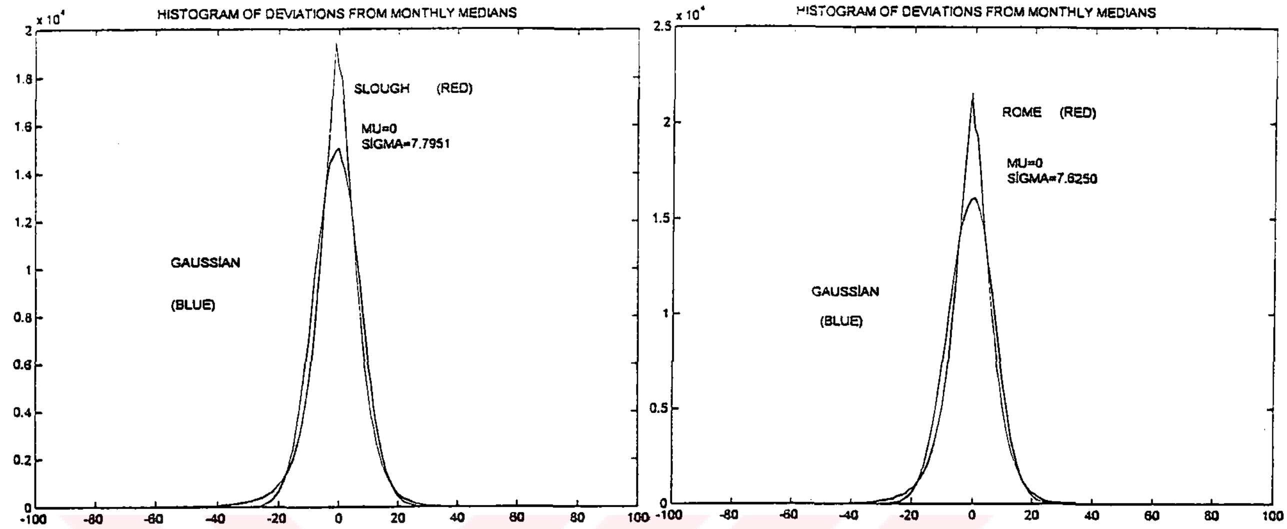

Figure4.3.1c-d.Comparison between the histogram of the deviations from monthly medians and gaussian function with zero mean and appropriate sigma for stations c) Slough(51N-55N) as mid latitude d)Rome(40N-45N) as low latitude.

Namely for certain hours, positive deviations were more probable than the approximating Gaussian while for certain hours negative deviations were more probable (not shown). Thus the hypothesis of considering the deviations form monthly medians as a single random variable is abandoned and these deviations are studied more closely in the time domain. Nevertheless as the overall errors range in %10-%15 interval, the Gaussian approximation is still a reasonable working hypothesis and the list of standard deviations given in Table 4.3.1 is expected to be useful in other applications (for example filtering applications). To illustrate the techniques used in the process it is recommended to present in details the properties of the data in details. Certain qualitative features of the data are as follows. Recall that a simple observation of the graph of ΔfoF2 shows that, extreme variations are rare, and there is a modulation by R_{12}. Thus the data is clipped to [-100,100].

A closer study using histograms shows that the variations ΔfoF2 range in [-50 to 50] with a probability of 0.9993, in [-10,10] with a probability of 0.8285, and in [-1,1] with a probability of 0.1273. The absolute value of the mean for various hours and various years is always less than 1. The shape of these variations can be compared to a normal distribution with the given mean and appropriate standard deviation.

The standard deviation leading to minimal mean square error is obtained by a one-dimensional optimization. Although there is not shown here in figures, it is observed that (i) for high R_{12}, the probability density for positive deviation Δf^{+} lie above the normal distribution, (ii) for low R_{12}, the probability density for negative deviation Δf^{-} lie below the normal distribution. The shape of the probability distribution function for each hour is also different. The peaks are narrower at nighttime, and wider at the afternoon and evening hours. The modulation by R_{12} in the time domain graphs shows itself in the Fourier spectrum, as dominant low frequency components.

CHAPTER 5. PARAMETER ESTIMATION BY MEANS OF ITS PHYSICAL DEPENDENCY

5.1. QUALITATIVE OBSERVATIONS

As a first step in the study of the time variations of the deviations from monthly medians, the deviations from monthly medians for 06h, 12h, 18h and 24h UT, for 6 stations in the geographic longitude band 10E-20E is shown in **Figure 5.1.1a-f.** A simple observation of these graphs shows that for all stations the deviations have larger amplitude at 12UT, the positive and negative deviations have different character and there is a clear modulation by R_{12}.

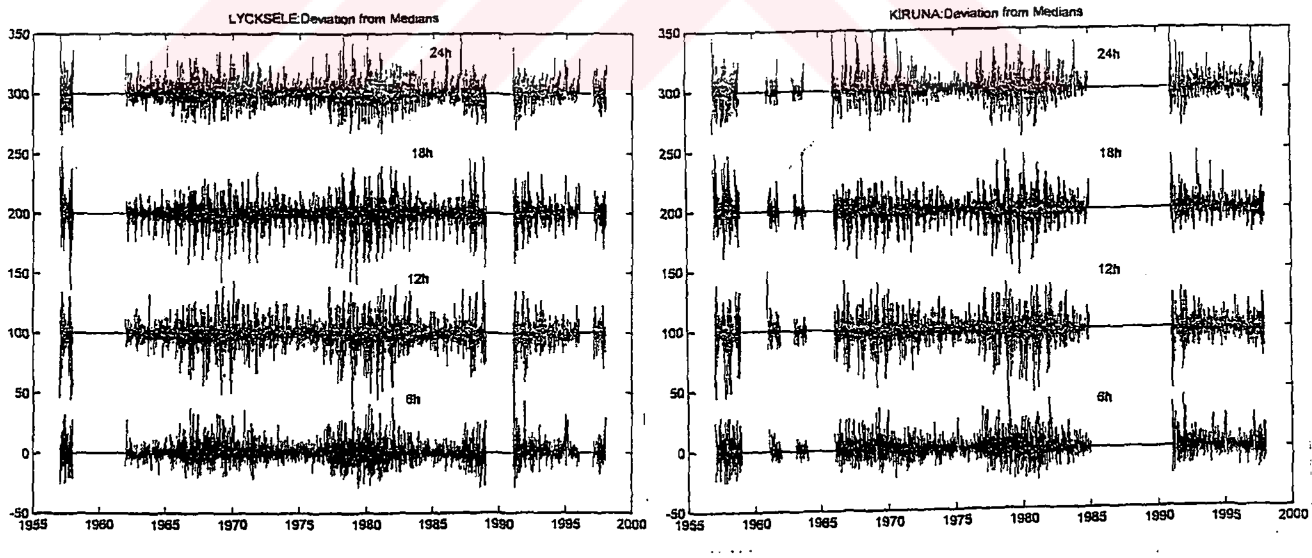

Figure 5.1.1a-b. Deviations from monthly medians for 06h, 12h, 18h, 24h U.T. for stations a) Kiruna b) Lycksele as high latitude (over 60N) in geographic longitude band 10E-20E in time domain.

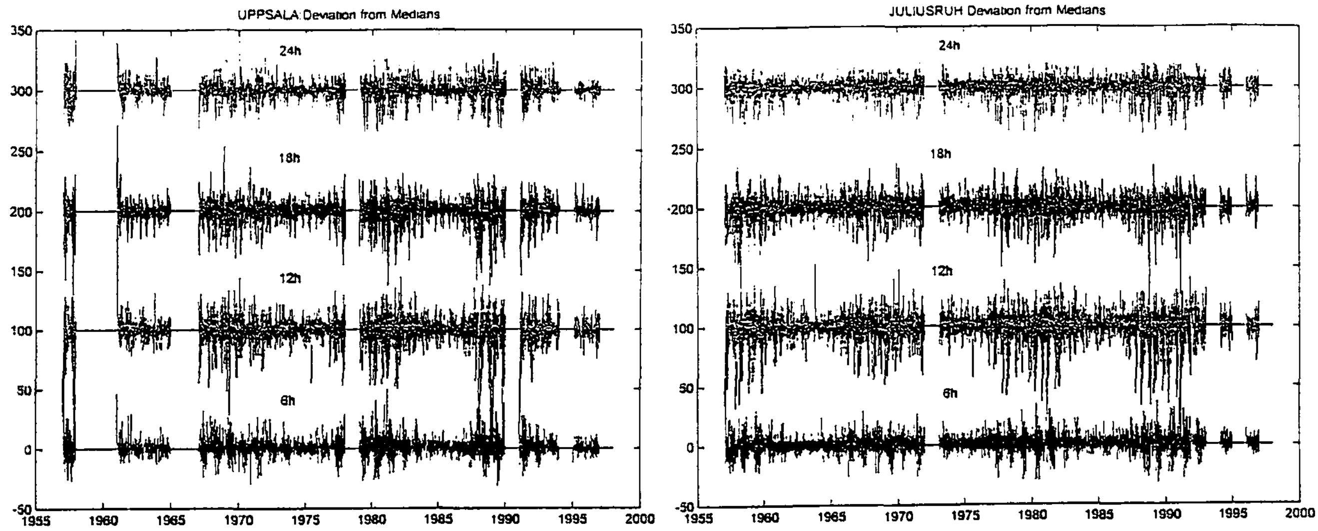

Figure 5.1.1c-d. Deviations from monthly medians for 06h, 12h, 18h, 24h U.T. for stations a)Uppsala b)Juliusruh on the latitude band (50N-60N) in geographic longitude band 10E-20E in time domain.

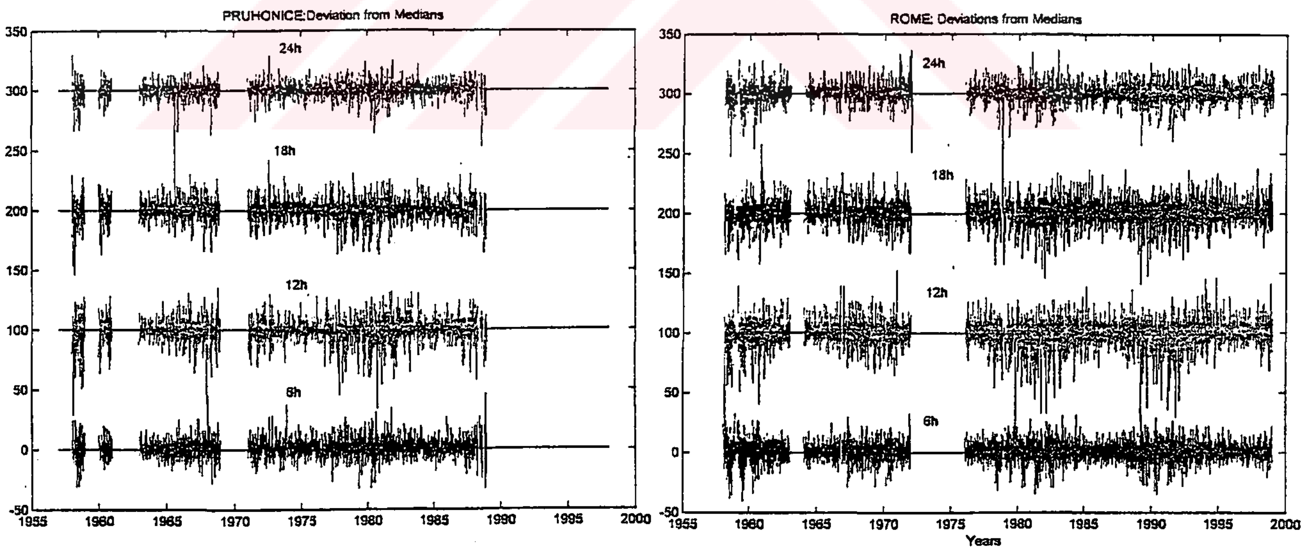

Figure 5.1.1e-f. Deviations from monthly medians for 06h, 12h, 18h, 24h U.T. for stations a) Pruhonice, b) Rome as low latitude (40N-50N) in geographic longitude band 10E-20E in time domain.

5.2 Physical dependency of the parameter

R_{12} dependency:

A closer look to the negative deviations of all stations for all years indicates that they are modulated by R_{12}. At high latitude, for Lycksele and Kiruna, both positive and negative deviations at all hours have modulation by R_{12}. For Uppsala only at daylight hours positive deviations have modulation by R_{12}. At mid latitude, for Juliusruh, there is a slight modulation of positive deviations by R_{12} at 06UT. At low latitude, for Pruhonice and Rome, positive deviations have practically no modulation by R_{12}. From these observations it is concluded that negative deviations are modulated by R_{12} at all latitudes, but positive deviations have a R_{12} dependency only at high latitudes.

Day and night difference:

In the histograms of Δf for each hour it is seen that, for daylight hours, there is an asymmetry between positive and negative deviations, negative deviations being more probable, while the histogram for 24UT is almost symmetric. This fact can be seen in **Figure5.2.1** on the histograms of Δf for Rome, for 24UT and 9UT. It is also noticed that, the histograms are wider for daylight hours and this is consistent with the observation that 27 days variation disappears at night time (see Figure 3.2.2).

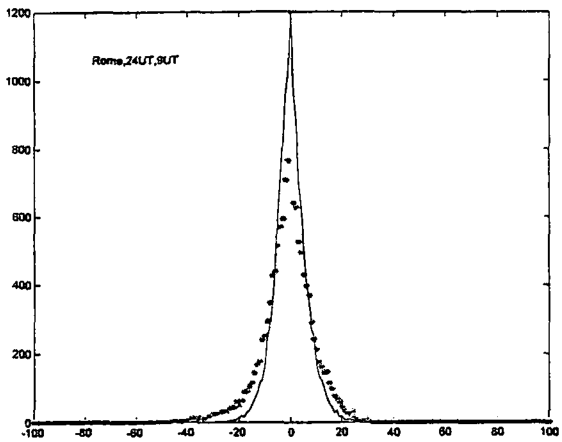

Figure 5.2.1. Histogram plots of the deviations from monthly medians for Rome station. Comparison of 24UT(-) and 9UT(*).

Latitude and longitude dependency:

The time domain plots given in **Figure 5.2.2a-b** for the daily deviations from monthly medians show that, at night hours, the negative deviations are below the positive deviations, while at day time the roles are reversed. The local time dependency can also be seen as a shift of the crossing points as moving longitudinally. Similar graphs are obtained for each latitude group in Table 1, where the similarities of the shapes of the curves inside each group justify the subdivisions with respect to latitude.

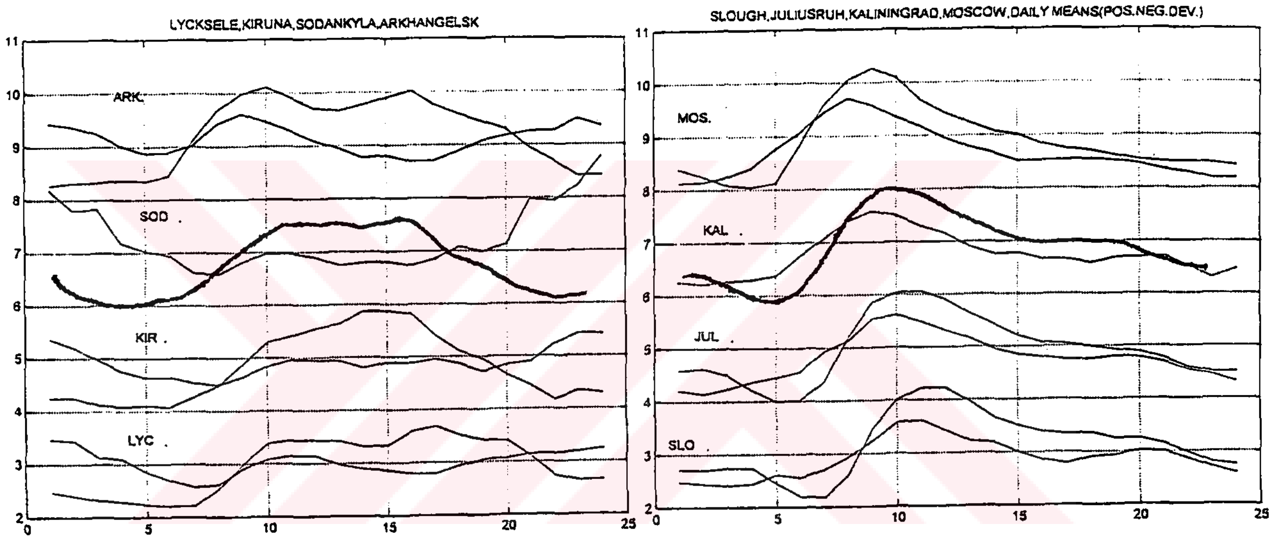

Figure 5.2.2a-b.The difference between the positive and negative deviations for stations a)Lycksele, Kiruna, Sodankyla and Arkhangelsk as high latitude. b)Slough, Juliusruh, Kaliningrad, Moscow as mid latitude.

Seasonal dependency:

Working with 1year samples, only qualitative remarks can be made for seasonal dependency. The time domain plots of ΔfoF2 given in **Figure 5.2.3.a-b** shows that at high latitudes, for example Uppsala, both positive and negative deviations have seasonal dependency while at lower latitudes, for example at Rome, only negative deviations have seasonal dependency. The amplitudes are higher during equinoxes. This is consistent with the spectral analysis results given in Figure 3.2.3c, where it can be seen that the periodicity of 6 months has higher power for negative deviations.

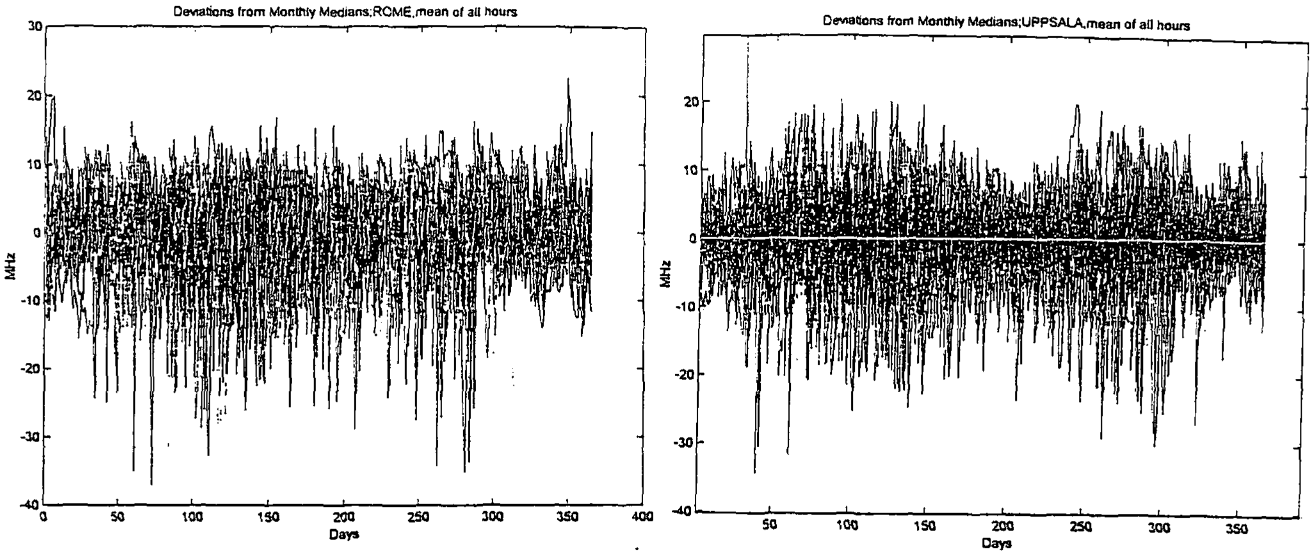

Figure 5.2.3a-b. Time domain plots of positive and negative deviations for stations a)Rome as low latitude, b) Uppsala as high latitude.

CHAPTER 6. MODELING POSITIVE AND NEGATIVE VARIATIONS SEPARATELY

The aim of the work is to give a convenient variability band to the variations of ΔfoF2, (we shall use $\Delta f^{\pm}$ instead of ΔfoF2). For this purpose we use upper deciles, whose definition is given in Chapter 2, to find the upper and lower bounds of this variability band.

We recall that we work with 365 days data samples of $\Delta f^{\pm}$ for each available stations, year and hour. For each sample we compute the %90 confidence interval, i.e. the value $(\Delta f^{\pm})_c$ such that $(\Delta f^{\pm})$ is less than $(\Delta f^{\pm})_c$ with a probability of 0.9, for the given station, hour and year. Thus we otain $(\Delta f^{\pm})_c$ values as a function

$$(\Delta f^{\pm})_c(\textbf{universal time, station, year}) \tag{6.1}$$

Recall that the longitude dependency can be taken into account by using local time, hence $(\Delta f^{\pm})_c$ is as well parameterized by the local time and latitude instead of universal time and station. Further we can replace parameters by year with a parameterized by R_{12}. Hence we have

$$(\Delta f^{\pm})_c = (\Delta f^{\pm})_c\ (\textbf{local time, latitude, } R_{12}) \tag{6.2}$$

At the first step, we aim to model the R_{12} dependency. We have observed that these upper deciles for positive deviations are virtually independent of R_{12}, and we decided to investigate only the negative deviations. As an example, in **Figure 6.1**, it is shown $(\Delta f^{-})_c$ for Sverdlovsk, for each hour as a function of R_{12}.

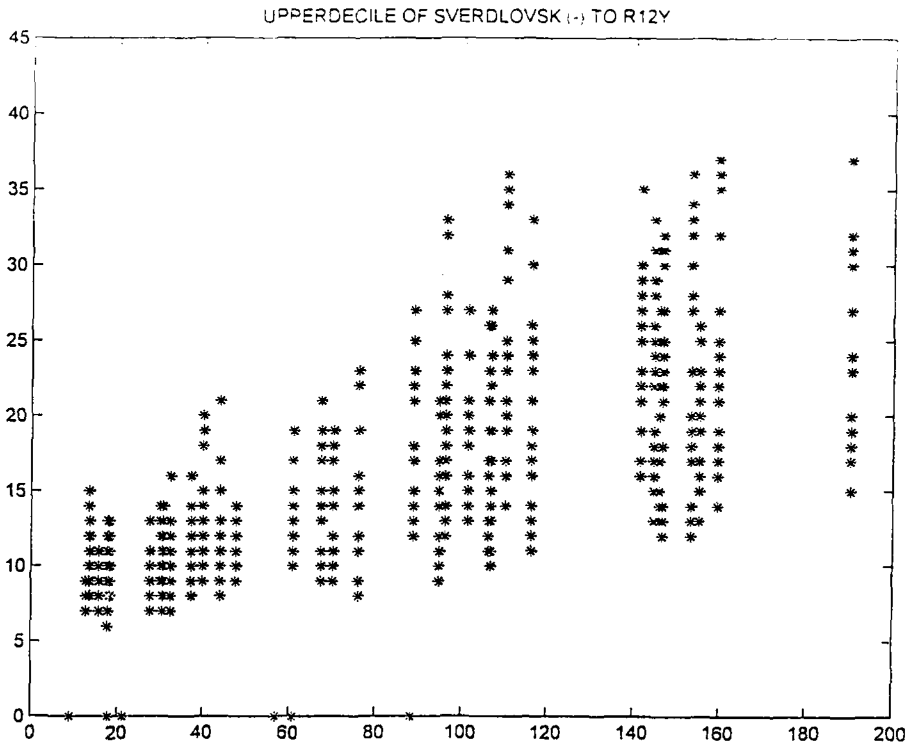

Figure 6.1. Upper deciles of negative deviations for Sverdlovsk station, with respect to R_{12} for all hours.

The data is quite scattered but there is a linear trend. Then after finding the slope and the intercept of the best line fit for each station and each hour, the required linear model shown below is obtained. Here the station or latitude dependency is parameterized as **θ**.

$$(\Delta f^-)_c(\text{local time}, \theta)=A(\text{local time}, \theta)R_{12}+B(\text{local time}, \theta) \qquad (6.3)$$

At the second stage we aim to model the latitude dependency. For this, we compute(the value $(\Delta f^-)_m$ by putting in the formula (6.3) $R_{12} = (R_{12})_{median}$. Hence we obtain values

$$(\Delta f^-)_m = (\Delta f^-)_m (\text{Local time, latitude}) \qquad (6.4)$$

Therefore we obtain models depending on 2 constants A and B for 18 stations and each hour. We compute 18*24*2 constants as shown in Table A1 in Appendix. The comparison of the analogues of Figure 6.2.1 for different stations suggestes that the values of the upper deciles tend to increase with the latitude, but there is a decrease after 60°N which is in agreement with previous work **(Kouris et al., 1999) (Kouris et al., 2000).**

In order to quantify these results, the computation is done after the conversion from Universal Time to Local Time and by inserting the median value of R_{12} in the formula below.

$$(\Delta f^{-})_m(\text{station,local time})=A(\text{station,local time})(R_{12})_{median}+B(\text{station,local time}) \quad (6.5)$$

Thus, a representative value of (Δf^{-}) for a given station and local time is obtained. In **Figure 6.2a-d**, the plot of $(\Delta f^{-})_m$ versus local time, for different stations grouped with respect to their latitudes are shown.

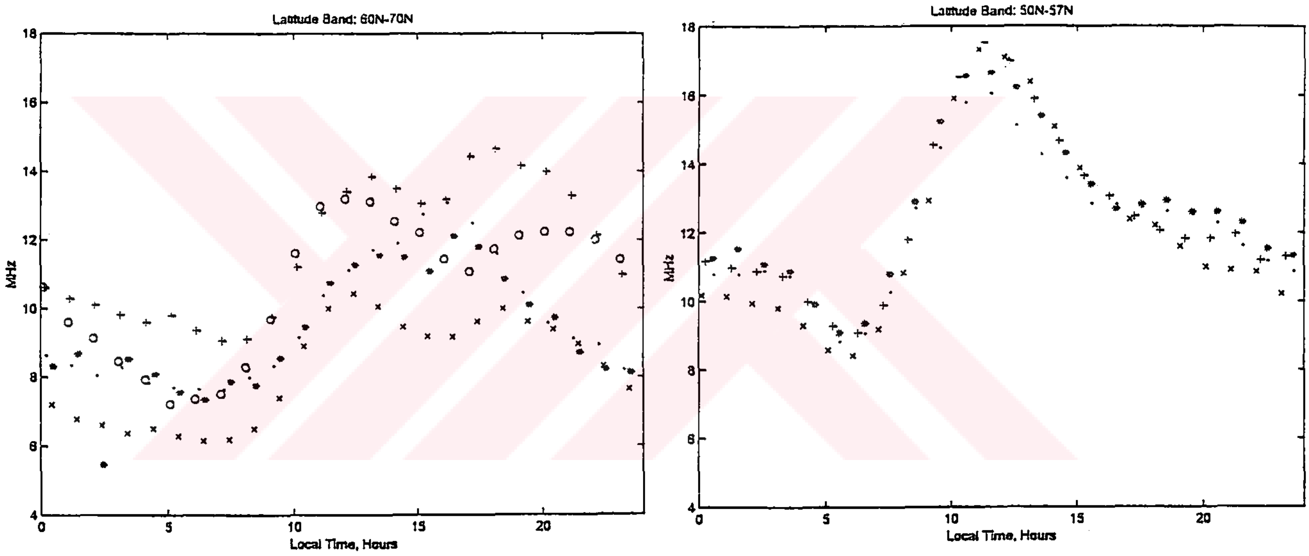

Figure 6.2a-b. Representative values of the upper deciles computed with the median value of R_{12} and plotted with respect to local time for the stations on latitude band a) 60N-70N b) 50N-57N.

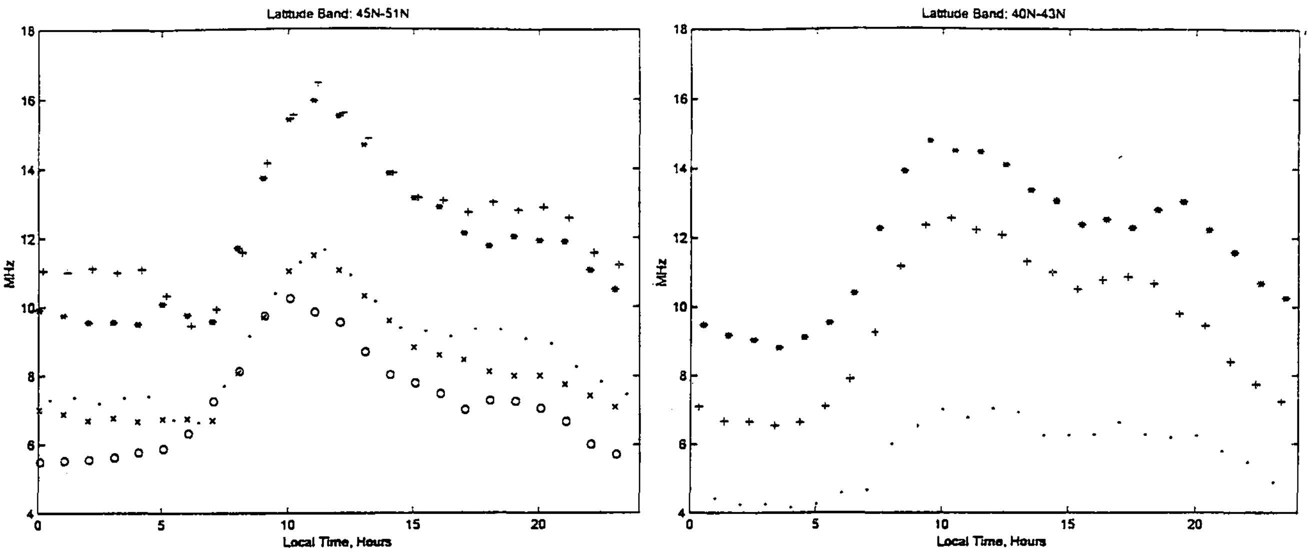

Figure 6.2.c-d. Representative values of the upper deciles computed with the median value of R_{12} and plotted with respect to local time for the stations on the latitude band c) 45N-51N d) 40N-43N.

It can be seen that the graphs for stations between 60°N-70°N are less coherent, the maxima is located at 12-14LT, and the amplitudes are less thatn 14 MHz. For the stations in the 50°N-57°N band, the variations are extremely coherent, there is a distinct peak at 12LT, and the amplitudes reach 18 MHz. Down to 45°N-51°N band, all graphs have a peak again around noon, but their amplitudes seem to have a shift, their respective maxima ranging from 10 to 16 MHz. In the lowest latitude band, 40°N-43°N, the peaks are located around 10-11LT, the maxima range from 7 to 14 MHz, but actually the lowest value of 7MHz, for Tortosa is questionable, because the data for Tortosa has large gaps.

CHAPTER 7. RESULTS AND DISCUSSION

In this study we applied statistical tools to model and estimate parameters of a process with deterministic and stochastic components. The process that we worked with was the ionospheric critical frequency foF2.

As a first step, for modeling, we considered the data as deterministic, at the scale of years and months and determined its periodicities using Fourier Analysis. Then, we studied the difference of foF2 and the monthly medians for each hour (ΔfoF2) as a stochastic process. We observed that the ΔfoF2 process considered as a time series sampled each hour still has deterministic components again at the timescales of years, months and hours. Thus we considered it as a stochastic process whose statistical properties change in time, but the change in these properties are tied to deterministic effects. Even that this observations lead to consider the process as a nonstationary process, we preferred to consider it as stationary because in previous works it is shown that models based on data from time intervals over which the process is nearly stationary are more successful. Thus we split the data to years and considered the data from each station, each year and each hour as stochastic process sampled daily and studied its power spectrum. We concluded that they are nearly white noise with some low frequency power.

As a second step we considered each set of 365 daily values as samples of random variables, and determined to what extend these random variables are Gaussian. We studied the histogram plots of the deviations from monthly medians for each station and each year. We concluded that they could be considered as Gaussian with a range of %10-%20 error. On the other hand we observed an asymmetry on the histogram plots. Based on this observation we studied the positive and negative deviations separately for each station and each year. This application on a large scale of data is a novelty in the literature.

We have shown that longitude dependency can be incorporated in Local Time and concluded that our main parameters are latitude, R_{12} and local time. We have studied local time dependency qualitatively and shown that the highest amplitude is at 12 am and graphed the 27 days periodicity.

As a third step we studied qualitatively the dependency of the statistical properties on physical parameters and concluded that the dependency of the negative deviations from monthly medians on the sunspot numbers R_{12} is the most crucial effects. We made a qualitative observation for the positive deviations and obtained a quantitative model for the upper deciles of the negative deviations from monthly medians in terms of the sunspot numbers. And we concluded that the upper deciles tend to increase with the latitude, but there is a decrease after 60°N.

It is known that geomagnetic storms result in a depression of foF2 and affect negative deviations. So the sensitivity of the negative deviations to R_{12} is not unexpected. Our observation that $\Delta f^{\pm}$ are different lead us to consider quiet and disturb days separately. This is also important for estimation, because after the onset of the storm, better estimation can be obtained if this information is also incorporated into the techniques.

In future works we aim to make better estimations of foF2 applying these results to different estimation techniques.

REFERENCES

Baykal, Ş.A., 1998. A model study of ionospheric Critical Frequencies for telecommunication purposes, M.Sc. Thesis, Middle East Technical University, Ankara.

Bilge, A.H., Tulunay, Y., 1998. Semi empirical single station modeling of foF2 variations: spectral analysis, Poster presented at the 23rd General Assembly of European Geophysical Society, Nice.

Bilge, A.H., Tulunay, Y., 2000. A novel on-line method for single station prediction and forecasting of ionospheric critical frequency foF2 1-hour ahead, Geophysical Research Letters, 27, 1383-1386.

Cerit, C., Yüksel, M., 2000. Olasılık, Beta Yayınevi, İstanbul.

Davis, C.J., Wild, M.N., Lockwood, M., Tulunay, Y. K., 1997. Ionospheric and geomagnetic responses to changes in IMF B_z: a superposed epoch study, Annales Geofisicae 15, 217-230.

Borowski, E.J. and Borwein J.M., 1991. Mathematics Dictionary, Harper Collins Publishers, New York.

Bremer J., 1992. Some Additional Results of Long-Term Trends In Vertical-Incidence Ionosonde Data, Journal of Atmospheric and Terrestrial Physics, 54, 11/12, 1505-1511.

Kouris, S.S., Fotiadis, D.N., Zolesi, B., 1999. Specifications of the F-Region variations for quiet and disturbed conditions, Phys. Chem. Earth(C), 24, 4, 321-327.

Kouris, S.S., Fotiadis, D.N., Hanbaba, R., 2000. On the day to dat variations of the MUF over Europe, Phys. Chem. Earth(C), 25, 4, 319-325.

Papoulis, A.,1965. Probability, Random Variables, and Stochastic Processe, McGraw-Hill Book Company, Tokyo.

Sage, A.P. and Melsa, J.L., 1971. Estimation Theory with Applications to Communications and Control, McGraw-Hill Book Company, New York.

Yamane, T., 1967. Statistics an Introductory Analysis, Harper International Edition, Tokyo.

APPENDIX A

Table A.1. Constants for polynomial fit to R_{12} of negative deviations from monthly medians of 18 stations, all years and each hour. U.T. (Universal time), L.T. (Local time).

$\Delta f = A^* R12 + B$; POLYFIT CONSTANTS

STATIONS	U.T. Hour. Minute	L.T. Hour. Minute	A CONSTANTS	B CONSTANTS
LYCKSELE	01.00	02.15	0.0371	7.7031
"	02.00	03.15	0.0381	7.3581
"	03.00	04.15	0.0381	7.1377
"	04.00	05.15	0.0355	7.5026
"	05.00	06.15	0.0347	7.1183
"	06.00	07.15	0.0305	7.0833
"	07.00	08.15	0.0382	6.6429
"	08.00	09.15	0.0463	6.7613
"	09.00	10.15	0.0589	7.3960
"	10.00	11.15	0.0618	8.7956
"	11.00	12.15	0.0718	8.7693
"	12.00	13.15	0.0715	9.2299
"	13.00	14.15	0.0696	9.0044
"	14.00	15.15	0.0766	8.1286
"	15.00	16.15	0.0792	8.0848
"	16.00	17.15	0.0979	8.1106
"	17.00	18.15	0.0901	8.8267
"	18.00	19.15	0.0586	10.3731
"	19.00	20.15	0.0525	10.6018
"	20.00	21.15	0.0416	10.5939
"	21.00	22.15	0.0257	10.4638
"	22.00	23.15	0.0314	8.9468
"	23.00	00.15	0.0344	8.3886
"	24.00	01.15	0.0394	7.7495

STATIONS	U.T. Hour. Minute	L.T. Hour. Minute	A CONSTANTS	B CONSTANTS
KIRUNA (-)	01.00	02.22	0.0453	5.1260
"	02.00	03.22	0.0473	5.2252
"	03.00	04.22	0.0479	4.7382
"	04.00	05.22	0.0497	4.4836
"	05.00	06.22	0.0462	4.6674
"	06.00	07.22	0.0307	5.6581
"	07.00	08.22	0.0391	5.4594
"	08.00	09.22	0.0453	5.3929
"	09.00	10.22	0.0558	5.5489
"	10.00	11.22	0.0725	5.6999
"	11.00	12.22	0.0801	5.9461
"	12.00	13.22	0.0844	6.2655
"	13.00	14.22	0.0951	5.7739
"	14.00	15.22	0.1267	4.5856
"	15.00	16.22	0.1340	4.4395
"	16.00	17.22	0.1057	5.6804
"	17.00	18.22	0.0845	6.1331
"	18.00	19.22	0.0576	6.7356
"	19.00	20.22	0.0464	6.5837
"	20.00	21.22	0.0434	6.3131
"	21.00	22.22	0.0377	6.4895
"	22.00	23.22	0.0416	5.5173
"	23.00	00.22	0.0493	5.4460
"	24.00	01.22	0.0480	5.2435
SODANKYLA	01.00	02.46	0.0301	9.5086
"	02.00	03.46	0.0728	3.8231
"	03.00	04.46	0.0673	3.7370
"	04.00	05.46	0.0681	3.1658
"	05.00	06.46	0.0609	3.3963
"	06.00	07.46	0.0539	4.3726
"	07.00	08.46	0.0599	3.8871
"	08.00	09.46	0.0705	3.9920
"	09.00	10.46	0.0895	3.6702
"	10.00	11.46	0.0991	4.3384
"	11.00	12.46	0.1101	4.1686
"	12.00	13.46	0.1023	4.9528
"	13.00	14.46	0.1132	4.2065
"	14.00	15.46	0.1151	3.6517
"	15.00	16.46	0.1224	4.2144
"	16.00	17.46	0.1228	3.8744
"	17.00	18.46	0.0995	4.4297

STATIONS	U.T. Hour. Minute	L.T. Hour. Minute	A CONSTANTS	B CONSTANTS
SODANKYLA	18.00	19.46	0.0847	4.6340
"	19.00	20.46	0.0668	5.4040
"	20.00	21.46	0.0656	4.4688
"	21.00	22.46	0.0710	3.6372
"	22.00	23.46	0.0638	3.9911
"	23.00	00.46	0.0661	4.0406
"	24.00	01.46	0.0662	4.4016
ARKHANGELSK	01.00	03.42	0.0556	2.7646
"	02.00	04.42	0.0597	2.6332
"	03.00	05.42	0.0557	2.6843
"	04.00	06.42	0.0552	2.5691
"	05.00	07.42	0.0511	2.8742
"	06.00	08.42	0.0549	2.9331
"	07.00	09.42	0.0659	3.1340
"	08.00	10.42	0.0778	3.8726
"	09.00	11.42	0.0894	4.2188
"	10.00	12.42	0.1001	3.9733
"	11.00	13.42	0.0917	4.1427
"	12.00	14.42	0.0888	3.7267
"	13.00	15.42	0.0952	3.0337
"	14.00	16.42	0.1022	2.5744
"	15.00	17.42	0.1100	2.5158
"	16.00	18.42	0.1176	2.3959
"	17.00	19.42	0.0998	3.1665
"	18.00	20.42	0.0752	4.5255
"	19.00	21.42	0.0707	4.3878
"	20.00	22.42	0.0633	4.2231
"	21.00	23.42	0.0572	3.9404
"	22.00	00.42	0.0562	3.5715
"	23.00	01.42	0.0538	3.3066
"	24.00	02.42	0.0559	2.9862
UPPSALA	01.00	02.10	0.0386	6.6393
"	02.00	03.10	0.0419	5.7393
"	03.00	04.10	0.0376	5.4949
"	04.00	05.10	0.0267	5.4831
"	05.00	06.10	0.0214	5.9711
"	06.00	07.10	0.0228	6.0180
"	07.00	08.10	0.0357	5.9720
"	08.00	09.10	0.0427	6.9264
"	09.00	10.10	0.0546	8.0788

STATIONS	U.T. Hour. Minute	L.T. Hour. Minute	A CONSTANTS	B CONSTANTS
UPPSALA	10.00	11.10	0.0641	8.8259
"	11.00	12.10	0.0692	8.7187
"	12.00	13.10	0.0662	8.8422
"	13.00	14.10	0.0613	8.5908
"	14.00	15.10	0.0629	8.1581
"	15.00	16.10	0.0614	7.4840
"	16.00	17.10	0.0604	7.1709
"	17.00	18.10	0.0546	8.2038
"	18.00	19.10	0.0655	7.8940
"	19.00	20.10	0.0526	8.8355
"	20.00	21.10	0.0426	9.4706
"	21.00	22.10	0.0353	9.7112
"	22.00	23.10	0.0337	9.2407
"	23.00	00.10	0.0386	8.1279
"	24.00	01.10	0.0386	7.1044
SLOUGH	01.00	00.58	0.0820	5.4863
"	02.00	01.58	0.0881	5.0896
"	03.00	02.58	0.0899	5.0808
"	04.00	03.58	0.0907	4.8503
"	05.00	04.58	0.0864	4.3177
"	06.00	05.58	0.0758	3.9191
"	07.00	06.58	0.0803	3.8605
"	08.00	07.58	0.0991	3.8485
"	09.00	08.58	0.1235	4.7197
"	10.00	09.58	0.1513	4.6887
"	11.00	10.58	0.1683	4.9078
"	12.00	11.58	0.1739	4.8162
"	13.00	12.58	0.1638	4.5306
"	14.00	13.58	0.1526	4.4196
"	15.00	14.58	0.1345	4.8949
"	16.00	15.58	0.1242	4.8152
"	17.00	16.58	0.1228	4.8941
"	18.00	17.58	0.1264	4.5616
"	19.00	18.58	0.1155	5.1592
"	20.00	19.58	0.1058	5.6982
"	21.00	20.58	0.0910	6.3909
"	22.00	21.58	0.0770	6.6339
"	23.00	22.58	0.0742	6.3566
"	24.00	23.58	0.0790	5.7554

STATIONS	U.T. Hour. Minute	L.T. Hour. Minute	A CONSTANTS	B CONSTANTS
JULIUSRUH	01.00	01.54	0.0648	7.3363
"	02.00	02.54	0.0618	7.0863
"	03.00	03.54	0.0637	6.7523
"	04.00	04.54	0.0612	5.9409
"	05.00	05.54	0.0622	5.0615
"	06.00	06.54	0.0735	4.6009
"	07.00	07.54	0.0856	5.2487
"	08.00	08.54	0.1096	5.8129
"	09.00	09.54	0.1328	6.6515
"	10.00	10.54	0.1546	6.5712
"	11.00	11.54	0.1488	7.0399
"	12.00	12.54	0.1587	5.9762
"	13.00	13.54	0.1445	6.0622
"	14.00	14.54	0.1291	5.9704
"	15.00	15.54	0.1206	5.6155
"	16.00	16.54	0.1074	5.7508
"	17.00	17.54	0.1093	5.7577
"	18.00	18.54	0.1051	6.1387
"	19.00	19.54	0.0898	6.7942
"	20.00	20.54	0.0748	7.7739
"	21.00	21.54	0.0620	8.2954
"	22.00	22.54	0.0533	8.0802
"	23.00	23.54	0.0564	7.6829
"	24.00	00.54	0.0651	7.0641
KALININGRAD	01.00	02.22	0.0465	4.6000
"	02.00	03.22	0.0462	4.4751
"	03.00	04.22	0.0424	4.0447
"	04.00	05.22	0.0419	3.5907
"	05.00	06.22	0.0382	3.7175
"	06.00	07.22	0.0489	3.8885
"	07.00	08.22	0.0573	4.5677
"	08.00	09.22	0.0731	5.3088
"	09.00	10.22	0.0849	5.9632
"	10.00	11.22	0.0943	5.8131
"	11.00	12.22	0.0878	5.9402
"	12.00	13.22	0.0849	5.3742

STATIONS	U.T. Hour. Minute	L.T. Hour. Minute	A CONSTANTS	B CONSTANTS
KALININGRAD	13.00	14.22	0.0819	5.0366
"	14.00	15.22	0.0751	4.7998
"	15.00	16.22	0.0725	4.6016
"	16.00	17.22	0.0684	4.5826
"	17.00	18.22	0.0610	4.9032
"	18.00	19.22	0.0618	5.0870
"	19.00	20.22	0.0506	5.5683
"	20.00	21.22	0.0432	5.9363
"	21.00	22.22	0.0385	5.7949
"	22.00	23.22	0.0386	5.5948
"	23.00	00.22	0.0416	5.2288
"	24.00	01.22	0.0509	4.4960
MOSCOW	01.00	03.29	0.0469	7.6914
"	02.00	04.29	0.0421	7.2626
"	03.00	05.29	0.0459	6.3029
"	04.00	06.29	0.0523	5.6885
"	05.00	07.29	0.0639	5.7432
"	06.00	08.29	0.0890	6.0287
"	07.00	09.29	0.1073	7.6152
"	08.00	10.29	0.1306	8.0827
"	09.00	11.29	0.1419	8.3495
"	10.00	12.29	0.1423	7.7880
"	11.00	13.29	0.1248	7.8236
"	12.00	14.29	0.1189	6.9829
"	13.00	15.29	0.1031	6.9978
"	14.00	16.29	0.0999	6.6225
"	15.00	17.29	0.0953	6.3243
"	16.00	18.29	0.1001	5.5840
"	17.00	19.29	0.0933	5.7883
"	18.00	20.29	0.0755	6.9404
"	19.00	21.29	0.0622	7.9404
"	20.00	22.29	0.0413	8.5174
"	21.00	23.29	0.0341	9.0937
"	22.00	00.29	0.0412	8.4967
"	23.00	01.29	0.0461	7.9902
"	24.00	02.29	0.0473	7.8097

STATIONS	U.T. Hour. Minute	L.T. Hour. Minute	A CONSTANTS	B CONSTANTS
SVERDLOVSK	01.00	05.10	0.0556	4.9830
"	02.00	06.10	0.0586	4.6191
"	03.00	07.10	0.0715	4.5531
"	04.00	08.10	0.0932	4.7907
"	05.00	09.10	0.1202	5.1584
"	06.00	10.10	0.1563	5.8101
"	07.00	11.10	0.1737	6.1008
"	08.00	12.10	0.1695	6.1395
"	09.00	13.10	0.1667	5.6130
"	10.00	14.10	0.1526	5.2307
"	11.00	15.10	0.1333	5.2705
"	12.00	16.10	0.1221	4.9500
"	13.00	17.10	0.1231	4.4351
"	14.00	18.10	0.1293	3.8624
"	15.00	19.10	0.1246	3.5419
"	16.00	20.10	0.1149	3.5722
"	17.00	21.10	0.0944	4.8042
"	18.00	22.10	0.0806	5.6434
"	19.00	23.10	0.0640	6.0758
"	20.00	00.10	0.0551	6.6060
"	21.00	01.10	0.0630	6.0726
"	22.00	02.10	0.0616	5.9587
"	23.00	03.10	0.0606	5.8829
"	24.00	04.10	0.0577	5.5578
LANNION	01.00	00.47	0.0557	3.6827
"	02.00	01.47	0.0528	3.9414
"	03.00	02.47	0.0547	3.6582
"	04.00	03.47	0.0561	3.7227
"	05.00	04.47	0.0539	3.9115
"	06.00	05.47	0.0533	3.2691
"	07.00	06.47	0.0576	2.9091
"	08.00	07.47	0.0710	3.1026
"	09.00	08.47	0.0868	3.5291
"	10.00	09.47	0.1080	3.3919
"	11.00	10.47	0.1163	3.7644
"	12.00	11.47	0.1169	4.0836

STATIONS	U.T. Hour. Minute	L.T. Hour. Minute	A CONSTANTS	B CONSTANTS
LANNION	13.00	12.47	0.1050	4.1223
"	14.00	13.47	0.0992	3.7218
"	15.00	14.47	0.0866	3.7599
"	16.00	15.47	0.0798	4.1123
"	17.00	16.47	0.0751	4.2659
"	18.00	17.47	0.0752	4.4733
"	19.00	18.47	0.0640	5.1707
"	20.00	19.47	0.0610	5.1229
"	21.00	20.47	0.0572	5.2282
"	22.00	21.47	0.0451	5.3353
"	23.00	22.47	0.0420	5.1194
"	24.00	23.47	0.0450	4.5634
POITIERS	01.00	01.01	0.0503	6.4881
"	02.00	02.01	0.0560	5.9246
"	03.00	03.01	0.0596	5.6876
"	04.00	04.01	0.0588	5.6744
"	05.00	05.01	0.0566	6.3957
"	06.00	06.01	0.0524	6.3411
"	07.00	07.01	0.0615	5.5729
"	08.00	08.01	0.0775	6.6697
"	09.00	09.01	0.1019	7.1104
"	10.00	10.01	0.1373	6.5200
"	11.00	11.01	0.1318	7.4344
"	12.00	12.01	0.1256	7.3982
"	13.00	13.01	0.1161	7.1671
"	14.00	14.01	0.1100	6.7375
"	15.00	15.01	0.1048	6.3588
"	16.00	16.01	0.0975	6.5747
"	17.00	17.01	0.0865	6.5298
"	18.00	18.01	0.0710	7.1705
"	19.00	19.01	0.0585	8.2608
"	20.00	20.01	0.0487	8.7818
"	21.00	21.01	0.0408	9.2613
"	22.00	22.01	0.0382	8.5950
"	23.00	23.01	0.0398	7.9202
"	24.00	00.01	0.0458	6.9449

STATIONS	U.T. Hour. Minute	L.T. Hour. Minute	A CONSTANTS	B CONSTANTS
DOURBES	01.00	01.18	0.0285	9.1562
"	02.00	02.18	0.0324	9.0197
"	03.00	03.18	0.0312	8.9612
"	04.00	04.18	0.0311	9.0440
"	05.00	05.18	0.0283	8.4666
"	06.00	06.18	0.0316	7.3736
"	07.00	07.18	0.0428	7.1402
"	08.00	08.18	0.0529	8.1273
"	09.00	09.18	0.0743	9.3229
"	10.00	10.18	0.0892	9.7680
"	11.00	11.18	0.0952	10.3018
"	12.00	12.18	0.0850	10.1017
"	13.00	13.18	0.0832	9.4719
"	14.00	14.18	0.0819	8.5690
"	15.00	15.18	0.0696	8.6427
"	16.00	16.18	0.0645	8.8885
"	17.00	17.18	0.0618	8.7259
"	18.00	18.18	0.0540	9.5403
"	19.00	19.18	0.0470	9.7576
"	20.00	20.18	0.0333	10.7334
"	21.00	21.18	0.0267	10.8532
"	22.00	22.18	0.0171	10.4595
"	23.00	23.18	0.0208	9.8655
"	24.00	00.18	0.0283	9.2242
KIEV	01.00	03.02	0.0166	5.6867
"	02.00	04.02	0.0174	5.5287
"	03.00	05.02	0.0166	5.6628
"	04.00	06.02	0.0211	5.3819
"	05.00	07.02	0.0276	4.9074
"	06.00	08.02	0.0379	5.6202
"	07.00	09.02	0.0506	6.3997
"	08.00	10.02	0.0636	6.9091
"	09.00	11.02	0.0615	7.4999
"	10.00	12.02	0.0599	7.1804
"	11.00	13.02	0.0542	6.7885
"	12.00	14.02	0.0480	6.4676

STATIONS	U.T. Hour. Minute	L.T. Hour. Minute	A CONSTANTS	B CONSTANTS
KIEV	13.00	15.02	0.0470	5.7552
"	14.00	16.02	0.0468	5.5556
"	15.00	17.02	0.0461	5.4596
"	16.00	18.02	0.0400	5.5102
"	17.00	19.02	0.0363	5.6384
"	18.00	20.02	0.0293	6.0997
"	19.00	21.02	0.0240	6.1971
"	20.00	22.02	0.0137	6.5355
"	21.00	23.02	0.0115	6.3546
"	22.00	00.02	0.0140	6.0972
"	23.00	01.02	0.0178	5.7246
"	24.00	02.02	0.0172	5.5719
NOVOKAZALINSK	01.00	05.08	0.0041	5.6005
"	02.00	06.08	0.0066	5.8825
"	03.00	07.08	0.0149	6.2754
"	04.00	08.08	0.0298	6.1826
"	05.00	09.08	0.0370	7.3192
"	06.00	10.08	0.0365	7.8636
"	07.00	11.08	0.0296	7.9111
"	08.00	12.08	0.0246	7.9292
"	09.00	13.08	0.0198	7.3833
"	10.00	14.08	0.0166	6.9333
"	11.00	15.08	0.0096	7.1287
"	12.00	16.08	0.0107	6.7533
"	13.00	17.08	0.0151	6.0157
"	14.00	18.08	0.0182	6.0846
"	15.00	19.08	0.0150	6.2716
"	16.00	20.08	0.0104	6.3606
"	17.00	21.08	0.0090	6.0715
"	18.00	22.08	0.0044	5.7014
"	19.00	23.08	0.0001	5.6975
"	20.00	00.08	-0.0019	5.6023
"	21.00	01.08	-0.0028	5.6925
"	22.00	02.08	-0.0008	5.5922
"	23.00	03.08	-0.0018	5.7183
"	24.00	04.08	0.0002	5.7390

STATIONS	U.T. Hour. Minute	L.T. Hour. Minute	A CONSTANTS	B CONSTANTS
TORTOSA	01.00	01.02	0.0116	3.6609
"	02.00	02.02	0.0143	3.3181
"	03.00	03.02	0.0141	3.3334
"	04.00	04.02	0.0162	3.1091
"	05.00	05.02	0.0139	3.3508
"	06.00	06.02	0.0126	3.7615
"	07.00	07.02	0.0126	3.8317
"	08.00	08.02	0.0193	4.7321
"	09.00	09.02	0.0283	4.6829
"	10.00	10.02	0.0259	5.3274
"	11.00	11.02	0.0276	4.9847
"	12.00	12.02	0.0240	5.4859
"	13.00	13.02	0.0204	5.6038
"	14.00	14.02	0.0275	4.4751
"	15.00	15.02	0.0238	4.7197
"	16.00	16.02	0.0106	5.5925
"	17.00	17.02	0.0096	6.0034
"	18.00	18.02	0.0109	5.5671
"	19.00	19.02	0.0045	5.8896
"	20.00	20.02	-0.0014	6.3238
"	21.00	21.02	0.0015	5.6678
"	22.00	22.02	-0.0011	5.5217
"	23.00	23.02	0.0017	4.7545
"	24.00	00.02	0.0032	4.4988
ROME	01.00	01.50	0.0611	5.2173
"	02.00	02.50	0.0671	4.6989
"	03.00	03.50	0.0656	4.5687
"	04.00	04.50	0.0645	4.9370
"	05.00	05.50	0.0627	5.4655
"	06.00	06.50	0.0875	4.7199
"	07.00	07.50	0.1127	4.9536
"	08.00	08.50	0.1340	5.2192
"	09.00	09.50	0.1506	5.0105
"	10.00	10.50	0.1372	5.5955
"	11.00	11.50	0.1285	6.1386
"	12.00	12.50	0.1199	6.3271

STATIONS	U.T. Hour. Minute	L.T. Hour. Minute	A CONSTANTS	B CONSTANTS
ROME	13.00	13.50	0.1135	6.0379
"	14.00	14.50	0.1146	5.6695
"	15.00	15.50	0.1002	5.9173
"	16.00	16.50	0.0969	6.2988
"	17.00	17.50	0.0867	6.7081
"	18.00	18.50	0.0793	7.7127
"	19.00	19.50	0.0686	8.6371
"	20.00	20.50	0.0617	8.2642
"	21.00	21.50	0.0574	7.8983
"	22.00	22.50	0.0584	6.9252
"	23.00	23.50	0.0616	6.2754
"	24.00	00.50	0.0617	5.4904
SOFIA	01.00	02.34	0.0122	5.8639
"	02.00	03.34	0.0130	5.7047
"	03.00	04.34	0.0094	6.0225
"	04.00	05.34	0.0102	6.4239
"	05.00	06.34	0.0198	6.5971
"	06.00	07.34	0.0277	7.4347
"	07.00	08.34	0.0444	8.2672
"	08.00	09.34	0.0570	8.6340
"	09.00	10.34	0.0515	9.1873
"	10.00	11.34	0.0408	9.5536
"	11.00	12.34	0.0359	9.7370
"	12.00	13.34	0.0355	9.0130
"	13.00	14.34	0.0417	8.3186
"	14.00	15.34	0.0397	7.9485
"	15.00	16.34	0.0494	7.6090
"	16.00	17.34	0.0552	7.3277
"	17.00	18.34	0.0415	8.0159
"	18.00	19.34	0.0261	8.1170
"	19.00	20.34	0.0160	8.4184
"	20.00	21.34	0.0101	7.7523
"	21.00	22.34	0.0060	7.3444
"	22.00	23.34	0.0077	6.7272
"	23.00	00.34	0.0089	6.5189
"	24.00	01.34	0.0135	5.7956

CURRICULUM VITAE

Eti Mizrahi was born on September 06,1956 in Istanbul. She graduated with a B.Sc. in Mathematical Engineering from Istanbul Technical University in 1999. The same year she began her M.S. studies in applied mathematics at the same university.
She is also a research assistant at ITU in the Department of Mathematics since November, 1999.