CRRAO Advanced Institute of Mathematics, Statistics and Computer Science (AIMSCS)

Research Report

Author (s): **T. Subba Rao, Gy. Terdik**

Title of the Report: **A space-time covariance function for spatio-temporal random processes and spatio-temporal prediction (kriging)**

Research Report No.: **RR2014-07**

Date: **March 19, 2014**

Prof. C R Rao Road, University of Hyderabad Campus,
Gachibowli, Hyderabad-500046, INDIA.
www.crraoaimscs.org

A space-time covariance function for spatio-temporal random processes and spatio-temporal prediction (kriging)

T. Subba Rao
University of Manchester,U.K. and
C. R. Rao AIMSCS, University of Hyderabad Campus, India

Gy. Terdik
University of Debrecen, Hungary

March, 2014

(Dedicated to the memory of Prof. M. B. Priestley, a great friend and a collegue, who passed away on 15th June 2013.)

Abstract

We consider a stationary spatio-temporal random process $\{Z(\mathbf{s},t) : (\mathbf{s},t) : \mathbf{s} \in \mathbb{R}^d,\ t \in \mathbb{Z}\}$ and assume that we have a sample $\{Z(\mathbf{s}_i,t);\ \mathbf{s}_i = 1,2,...m;\ t = 1,\ldots,n\}$ from $\{Z(\mathbf{s},t)\}$. By defining a sequence of discrete Fourier transforms at canonical frequencies at each location $\mathbf{s}_i$, $(i = 1,2,3...m)$, and using these complex valued random variables as observed sample, we obtain expressions for the spatio-temporal covariance functions and the spectral density functions of the spatio-temporal random processes. These spectra correspond to non separable class of random processes. The spatio-temporal covariance functions, obtained here are functions of the spatial distances and the temporal frequency and are similar to Matern class of covariance functions. These are in terms of modified Bessel functions of the second kind, and the parameters are in terms of the second order spectral density functions of the random proces and the spatial distances. We consider the estimation of the parameters of the covariance function and also briefly mention their asymptotic properties. The estimation of the entire data at a known location $\mathbf{s}_0$, $\{Z(\mathbf{s}_0,t).;t = 1,2,3...n\}$ and also the estimation of $Z(\mathbf{s}_0, n+v)$, for $v > 0$ given the above sample is also considered. The predictors are obtained using the vectors of

Discrete Fourier Transforms.The methods are illustrated with real and simulated data.

Keywords: Discrete Fourier Transforms, Covariance functions, Spectral density functions, Space-Time Processses, Prediction(kriging) Laplacian operators, Frequency Variogram, Long memory processes, Whittle likelihood.

1 Introduction and Summary

In recent years it has become necessary to develop statistical methods for analyzing data coming from diverse areas such as, environment, marine biology, agriculture, finance etc. The data which comes from these areas, are usually, functions of both spatial coordinates and temporal coordinates. Any statistical analysis developed must take into account both spatial dependence, temporal dependence and their interaction, if any. There is a vast literature on spatial analysis, (see [Cre93], for example) but not to the same extent in the case of spatio-temporal data. An addition of temporal dimension, which cannot be imbedded into spatial dimension, results in several problems, such as in spatio-temporal kriging, construction of finite parameter models for the data etc. One of the important problems often encountered and considered to be extremely important is the spatio-temporal prediction, commonly known as spatio-temporal kriging in the literature. The object in kriging, in the present context, is to predict the data at a known location where time series data is not observed, given the time series data at other locations. If one restricts to the construction of linear predictors as a linear combination of the entire observed data (the dimensions of which will be extremely large because of the number of spatial locations and number of time points) the optimal linear predictor will be a function of the covariance functions of the process which are functions of the space and time, and also a function of the data at the location $\mathbf{s}_0$ which is not observed. Besides, as pointed out by Cressie and Wikle [CW11](see chapter 6, p. 323-324), the problem is also related to ordering. Finding a suitable spatio-temporal covariance function, similar to Matern class, which is positive definite has become a challenge. However, [CH99], [Gne02], [DR07], [Ste05], [CG11] and [Ma02] and several authors in recent years (see [CW11] for details) have defined interesting classes of covariance functions which are positive definite. The estimation of the parameters of these functions have been discussed by [CH99] and [Gne02] and others, and recently by Subba Rao et al [SR13] proposed a frequency domain approach based on frequency domain version of the variogram for the estimation of the parameters which does not involve the inversion of matrices. The methodologies proposed by several authors for the spatio-temporal prediction which are in the time domain depend not only on the knowledge of the covariance function, but also on the inversion of large dimensional covariance matrices, the dimensions of which will increase as the number of observations over time and also the number of locations increase. In our present paper, based on the covariance functions of the Finite Fourier Transforms of the data we propose a method for the estimation of the entire data set at a known location $\mathbf{s}_0$ and also we consider the prediction of the future value at $(n+v, v>0)$, i.e. estimation of $Z(\mathbf{s}_0, n+v)$.

In section 1, the notation and the spectral representation of the spatio-temporal random processes are introduced . The spectral representations of the spatio-temporal processes and the properties of the discrete Fourier transforms are discussed in section 2. Expressions for the spatio-temporal covariances and spectral density functions when the random processes satisfy parametric models are derived in section 3 The estimation of the parameters of the spatio-temporal covariance functions are considered in section 4. The prediction of the entire data set at a known location given the data in the neighborhood using the Fourier transforms is discussed in section 5. In section 6, simulation of the data with a known covariance function and the estimation of the parameters of the covariance function is considered. Also the spatio-temporal prediction is considered in section 6. The methods are illustrated with real data in section 6.

1.1 Notation and Preliminaries.

Let $Z(\mathbf{s},t)$, where $\left\{(\mathbf{s},t) : \mathbf{s} \in \mathbb{R}^d, t \in \mathbb{Z}\right\}$, denote the spatio-temporal random process. We assume that the random process is spatially and temporally second order stationary, i.e.

$$\begin{aligned} E\left[Z(\mathbf{s},t)\right] &= \mu \\ Var\left[Z(\mathbf{s},t)\right] &= \sigma_Z^2 < \infty \\ Cov\left[Z(\mathbf{s},t), Z(\mathbf{s}+\mathbf{h},t+u)\right] &= c(\mathbf{h},u), \quad \mathbf{h} \in \mathbb{R}^d,\ u \in \mathbb{Z}. \end{aligned}$$

We note that $c(\mathbf{h},0)$ and $c(\mathbf{0},u)$ correspond to the purely spatial and purely temporal covariances of the process respectively. A further common stronger assumption that is often made is that the process is not only spatially stationary but also it is isotropic. The assumption of isotropy is a stronger assumption on the process The process is said to be isotropic if

$$c(\mathbf{h},u) = c(\|\mathbf{h}\|, u), \quad \|\mathbf{h}\| \in \mathbb{R}^1,\ u \in \mathbb{Z}$$

where $\|\mathbf{h}\|$ is the Euclidean distance. Without loss of generality, we set the mean μ equal to zero. As in the case of spatial processes, one can define the spatio-temporal variogram for $\{Z(\mathbf{h},t)\}$. It is defined as

$$2\gamma(\mathbf{h},u) = Var\left[Z(\mathbf{s}+\mathbf{h},t+u) - Z(\mathbf{s},t)\right]. \tag{1}$$

If the random process $\{Z(\mathbf{s},t)\}$ is spatially and temporally stationary, then we can rewrite the above as

$$2\gamma(\mathbf{h},u) = 2\left[c(\mathbf{0},0) - c(\mathbf{h},u)\right], \tag{2}$$

and for an isotropic process, $\gamma(\mathbf{h},u) = \gamma(\|\mathbf{h}\|,u)$. We note that $\gamma(\mathbf{h},u)$ is defined as the semi-variogram. Recently, [SRDB13] proposed a frequency domain version of the variogram which is used later for the estimation of the parameters of space-time covariance function.

In view of our assumption that the zero mean random process $\{Z(\mathbf{s},t)\}$ is second order spatially and temporally stationary,we have the spectral representation

$$Z(\mathbf{s},t) = \int_{-\infty}^{\infty}\int_{-\pi}^{\pi} e^{i(\mathbf{s}\cdot\underline{\lambda}+t\mu)} dZ_z(\underline{\lambda},\mu), \tag{3}$$

where $\mathbf{s}\cdot\underline{\lambda} = \sum_{i=1}^{d}\mathbf{s}_i\lambda_i$ and $\int_{-\infty}^{\infty}$ represents $d-$fold multiple integral. We note that $Z_z(\underline{\lambda},\mu)$ is a zero mean complex valued random process with orthogonal increments, with

$$\begin{aligned} E\left[dZ_z(\underline{\lambda},\mu)\right] &= 0 \\ E\left|dZ_z(\underline{\lambda},\mu)\right|^2 &= dF_z(\underline{\lambda},\mu), \end{aligned} \tag{4}$$

where $F_z(\underline{\lambda},\mu)$ is a non-decreasing function. If we assume further that $F(\underline{\lambda},\mu)$ is differentiable in all its $(d+1)$ arguments $\underline{\lambda}$ and μ, then $dF(\underline{\lambda},\mu) = f(\underline{\lambda},\mu)\, d\underline{\lambda}\, d\mu$. Here $f(\underline{\lambda},\mu)$ which is strictly positive and real valued, is defined as the spatio-temporal spectral density function of the random process $\{Z(\mathbf{s},t)\}$, and $-\infty < \lambda_1,\lambda_2,\ldots,\lambda_d < \infty, \quad -\pi \le \mu \le \pi$. In view of the orthogonality of the function $Z_z(\underline{\lambda},\mu)$, we can show that the positive definite covariance function $c(\mathbf{h},u)$ has the spectral representation

$$c(\mathbf{h},u) = \int_{-\infty}^{\infty}\int_{-\pi}^{\pi} e^{i(\mathbf{h}\cdot\underline{\lambda}+u\mu)} f(\underline{\lambda},\mu)\, d\underline{\lambda}\, d\mu \tag{5}$$

and by Fourier inversion, we have

$$f\left(\underset{\sim}{\lambda},\mu\right)=\frac{1}{\left(2\pi\right)^{d+1}}\sum_{u}\int\limits_{-\infty}^{\infty}e^{-i\left(\mathbf{h}\cdot\underset{\sim}{\lambda}+u\mu\right)}c\left(\mathbf{h},u\right)\,d\mathbf{h}, \tag{6}$$

where $d\mathbf{h}=\prod_{i=1}^{d}dh_i$. We further note that if the process is fully symmetric then $c\left(\mathbf{h},u\right)=c\left(-\mathbf{h},-u\right)$ and $f\left(\underset{\sim}{\lambda},\mu\right)=f\left(-\underset{\sim}{\lambda},-\mu\right)$ and $f\left(\underset{\sim}{\lambda},\mu\right)>0$ for all $\underset{\sim}{\lambda}$ and μ. Here $\underset{\sim}{\lambda}$ is the frequency associated with spatial coordinates (usually called wave number) and μ is the temporal frequency

2 Discrete Fourier Transforms and their properties

Let us assume that we have a sample $\{Z\left(\mathbf{s}_i,t\right); i=1,2,...m;\ t=1,\ldots,n\}$ from the zero mean spatio-temporal stationary process $\{Z\left(\mathbf{s},t\right)\}$. We now consider the time series data $\{Z\left(\mathbf{s}_i,t\right);\ t=1,\ldots,n\}$ at the location $\mathbf{s}_i$, and define the discrete Fourier transform

$$J_{\mathbf{s}_i}\left(\omega_k\right)=\frac{1}{\sqrt{2\pi n}}\sum_{t=1}^{n}Z\left(\mathbf{s}_i,t\right)e^{-it\omega_k}, \tag{7}$$

where $\omega_k=\frac{2\pi k}{n}$, $k=0,\pm1,\ldots,\pm\left[\frac{n}{2}\right]$. Define the corresponding second order periodogram by

$$I_{\mathbf{s}_i}\left(\omega_k\right)=\left|J_{\mathbf{s}_i}\left(\omega_k\right)\right|^2.$$

It is well known that the periodogram is asymptotically an unbiased estimator of the second order spectral density function, but it is not mean square consistent, and it is also well known that (see for example Priestley,1981[Pri81])

$$\begin{aligned}E\left(J_{\mathbf{s}_i}\left(\omega_k\right)\right)&=0\\ Var\left(J_{\mathbf{s}_i}\left(\omega_k\right)\right)&=E\left(I_{\mathbf{s}_i}\left(\omega_k\right)\right)\simeq g_{\mathbf{s}_i}\left(\omega_k\right).\end{aligned} \tag{8}$$

In view of our assumption of spatial and temporal stationarity the second order spectral density function of the process, $\{Z\left(\mathbf{s}_i,t\right)\}$ is same for all locations and hence we have,

$$g_{\mathbf{s}_i}\left(\omega_k\right)=g\left(\omega_k\right),\text{for all } i$$

where $g_{\mathbf{s}_i}\left(\omega_k\right)$ is the second order spectral density function of the spatial process $\{Z\left(\mathbf{s},t\right)\}$ .We assume that the second order spectral density function is a function of some p_1 parameters, say $\underline{\vartheta}_1$. From now onwards, we denote this spectral density function by $g\left(\omega_k, \underline{\vartheta}_1\right)$. In view of our assumption that $\{Z\left(\mathbf{s}_i, t\right)\}$ is temporally second order stationary, it can be shown, that for large n,

$$Cov\left(J_{\mathbf{s}_i}\left(\omega_k\right), J_{\mathbf{s}_i}\left(\omega_{k'}\right)\right) \simeq 0, \quad k \neq k'.$$

(see [Bri01], [Pri81], [DSR11]). If the random process $\{Z\left(\mathbf{s}_i, t\right)\}$ is Gaussian, then the complex valued random variables $\left\{J_{\mathbf{s}_i}\left(\omega_k\right);\ k = 0, 1, \ldots, \left[\frac{n}{2}\right]\right\}$ will be asymptotically independent, and will be complex Gaussian, each $J_{\mathbf{s}_i}\left(\omega_k\right)$ will be distributed with mean zero and variance proportional to $f_{\mathbf{s}_i}\left(\omega_k, \underline{\vartheta}_1\right)$, which is equal to $g\left(\omega_k, \underline{\vartheta}_1\right)$ in view of our assumption of spatial stationarity. Let us now evaluate the covariance function between the complex Fourier transforms $J_{\mathbf{s}_i}\left(\omega_k\right)$ and $J_{\mathbf{s}_j}\left(\omega_k\right)$. For large n, we can show ([Pri81])

$$Cov\left(J_{\mathbf{s}_i}\left(\omega_k\right), J_{\mathbf{s}_j}\left(\omega_k\right)\right) = E\left[I_{\mathbf{s}_i,\mathbf{s}_j}\left(\omega_k\right)\right] \simeq \frac{1}{2\pi}\sum_{n=-\infty}^{\infty} c\left(\mathbf{s}_i - \mathbf{s}_j, n\right) e^{-in\omega_k}. \tag{9}$$

We note that $I_{\mathbf{s}_i,\mathbf{s}_j}\left(\omega_k\right)$ is the cross periodogram between the spatial processes $\{Z\left(\mathbf{s}_i, t\right)\}$ and $\{Z\left(\mathbf{s}_j, t\right)\}$, and unlike the second order periodogram defined earlier which is always real valued, the cross periodogram is usually a complex valued function. and is aymptotically an unbiased estimator of the cross spectral density function given by

$$f_{\left(\mathbf{s}_i - \mathbf{s}_j\right)}\left(\omega\right) = \frac{1}{2\pi}\sum_{n=-\infty}^{\infty} c\left(\mathbf{s}_i - \mathbf{s}_j, n\right) e^{-in\omega}, \quad |\omega| \leq \pi \tag{10}$$

is also usually a complex valued function. However, under spatial isotropy and temporal stationarity assumptions, we have

$$c\left(\mathbf{s}_i - \mathbf{s}_j, n\right) = c\left(\left\|\mathbf{s}_i - \mathbf{s}_j\right\|, n\right) = c\left(\left\|\mathbf{s}_i - \mathbf{s}_j\right\|, -n\right).$$

and which in turn implies that the cross spectrum can be written as

$$f\left(\left\|\mathbf{s}_i - \mathbf{s}_j\right\|, \omega\right) = C(\| \mathbf{h} \|, \omega) = \frac{1}{2\pi}\sum_{n=-\infty}^{\infty} c(\left\|\mathbf{s}_i - \mathbf{s}_j\right\|, n)e^{-in\omega}, \quad |\omega| \leq \pi \tag{11}$$

which is symmetric over the frequency ω and is strictly positive and real valued. The above is a function of the Euclidean distance $\| \mathbf{h} \| = \|\mathbf{s}_i - \mathbf{s}_j\|$ and the temporal frequency ω. We use this function later when we consider prediction of the data at a known location. To obtain the spatio-temporal spectral density function of the random process defined earlier from the above function, we need to take Fourier transforms of the above over the Euclidean distance $\|\mathbf{s}_i - \mathbf{s}_j\|$. We will obtain expressions for the spatio-temporal density functions of the processes when they satisfy specific parametric models.

We now obtain an analytic expression for $f\left(\|\mathbf{s}_i - \mathbf{s}_j\|, \omega\right)$ under the assumption that random process satisfies a finite parameter model. (see [Whi53], [Whi54]). The expression derived will be similar to Matern's class of functions, but the parameters are functions of the second order spectral density of the random process. Later we use this covariance function for prediction. which is one of our main objects in this paper.

2.1 Fourier transform and Spectral Representation

In order to achieve the above objectives, we need a spectral representation for the discrete Fourier transform and this will be considered in the following section.

Consider the discrete Fourier transform (7),

$$J_{\mathbf{s}}(\omega) = \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^{n} Z(\mathbf{s}, t)\, e^{-it\omega}, \quad |\omega| \leq \pi \tag{12}$$

for any fixed location $\mathbf{s} \in \mathbb{R}^d$. Now substitute the spectral representation (3) for $Z(\mathbf{s}, t)$ in (12), and after some simplification, we obtain

$$J_{\mathbf{s}}(\omega) = \int \int e^{i\underline{s}\,\underline{\lambda}} \left[e^{i(n+1)\frac{\varphi}{2}} F_n^{\frac{1}{2}}(\varphi) \right] dZ_z(\underline{\lambda}, \mu) \tag{13}$$

where $\varphi = \mu - \omega$, $\int$ is a d dimensional multiple integral, (see Priestley, 1981.[Pri81], page 419.) and in obtaining the above, we used the fact that.

$$\sum_{t=1}^{n} e^{it\varphi} = e^{i(n+1)\frac{\varphi}{2}} \left[\frac{\sin n\frac{\varphi}{2}}{\sin\frac{\varphi}{2}} \right],$$

and the Fejér kernel $F_n(\varphi)$ is given by

$$F_n(\varphi) = \frac{1}{2\pi n} \frac{\sin^2 n\frac{\varphi}{2}}{\sin^2 \frac{\varphi}{2}}.$$

Hence

$$\sum_{t=1}^{n} e^{it\varphi} = e^{i(n+1)\frac{\varphi}{2}} \sqrt{2\pi n} F_n^{\frac{1}{2}}(\varphi).$$

It is well known that the Fejér kernel behaves like a Dirac Delta function as $n \to \infty$ and as $\varphi \to 0$, $F_n(\varphi) = 0(n)$. As pointed out by Priestley (1981.[Pri81], p. 419), $F_n^{\frac{1}{2}}(\varphi)$ does not strictly tend to a Dirac Delta δ - function as $n \to \infty$, nevertheless, behaves in a similar manner to a δ function. In particular as $n \to \infty$ and for all $\varphi \neq 0$, $F_n^{\frac{1}{2}}(\varphi) \to 0$, and as $\varphi \to 0$, $F_n^{\frac{1}{2}}(\varphi) \to \sqrt{\frac{n}{2\pi}}$. Therefore, as $n \to \infty$, $F_n^{\frac{1}{2}}(\varphi)$ vanishes everywhere except at the origin. In view of this, for large n, we can approximately write (13) as

$$J_{\mathbf{s}}(\omega) \simeq \int e^{i\underline{s}\,\underline{\lambda}} \sqrt{\frac{n}{2\pi}} dZ_z(\underline{\lambda}, \omega), \tag{14}$$

We note that the above integral is over the wave number space λ only. We use (14) later in our derivation of an expression for the spatio-temporal covariance function. In the following section we define a model similar to the models defined by Whittle(1953,1954)[Whi53], [Whi54], Jones and Zhang(1997)[JZ97], but we use the Laplacian operators on the complex valued random variables (Discrete Fourier transforms) to obtain expressions for the spatio-temporal spectral density functions and covariances which are functions of the spatial distances and the temporal frequencies. In their derivation, Jones and Zhang(1997)[JZ97] use a first order time derivative to accommodate the temporal dynamics, but in our derivation we use the frequency response function of the process to account for the temporal linear dependence in the time series.

3 Model and Derivation of the covariance function

We assume that corresponding to the spatio-temporal random process $\left\{Z(\mathbf{s},t); \mathbf{s} \in \mathbb{R}^d, t \in \mathbb{Z}\right\}$, we have a spatio-temporal random process $\left\{e(\mathbf{s},t); \mathbf{s} \in \mathbb{R}^d, t \in \mathbb{Z}\right\}$ which is like a white noise process in space and time. Similar assumptions are often made in spatial spatio-temporal analysis (for example see [CW11], [Gne02], [She11]) We assume that the random process $\{e(\mathbf{s},t)\}$ satisfies the following stationarity conditions

$$\begin{aligned} E(e(\mathbf{s},t)) &= 0 \\ Var(e(\mathbf{s},t)) &= \sigma_e^2, \text{ does not depend on } \mathbf{s} \text{ or } t. \\ Cov\left(e(\mathbf{s},t), e\left(\mathbf{s}',t'\right)\right) &= \sigma_e^2 I\left(\mathbf{s},\mathbf{s}'\right) I\left(t,t'\right), \end{aligned}$$

where

$$
\begin{aligned}
I(\mathbf{s},\mathbf{s}') &= \begin{cases} 1 \text{ if } \mathbf{s}=\mathbf{s}' \\ 0 \text{ otherwise} \end{cases} \\
I(t,t') &= \begin{cases} 1 \text{ if } t=t' \\ 0 \text{ otherwise.} \end{cases}
\end{aligned}
$$

As before, we assume that we have a sample $\{e(\mathbf{s}_i,t)\,;\ i=1,\ldots,m;\ t=1,\ldots,n\}$ corresponding to the observable spatio-temporal data $\{Z(\mathbf{s}_i,t)\,;i=1,\ldots,m;t=1,\ldots,n\}$. We define the discrete Fourier transform of these white noise processes

$$
J_{\mathbf{s},e}(\omega)=\frac{1}{\sqrt{2\pi n}}\sum_{t=1}^{n}e(\mathbf{s},t)\,e^{-it\omega}, \tag{15}
$$

where we assume, that the stationary process $e(\mathbf{s},t)$ has the spectral representation

$$
e(\mathbf{s},t)=\int\int e^{i(\underline{s}.\underline{\lambda}+t\mu)}dZ_e(\underline{\lambda},\mu)\,, \tag{16}
$$

where the orthogonal random process $Z_e(\underline{\lambda},\mu)$ satisfies

$$
\begin{aligned}
E\left[dZ_e(\underline{\lambda},\mu)\right] &= 0 \\
E\left|dZ_e(\underline{\lambda},\mu)\right|^2 &= \frac{\sigma_e^2}{(2\pi)^{d+1}}d\underline{\lambda}d\mu.
\end{aligned}
$$

Substituting (16) in (15) and proceeding as before, we obtain, for a fixed temporal frequency ω,

$$
J_{\mathbf{s},e}(\omega)\simeq\int e^{i\mathbf{s}\cdot\underline{\lambda}}\left[\sqrt{\frac{n}{2\pi}}\right]dZ_e(\underline{\lambda},\omega)\,. \tag{17}
$$

where $\int$ is a multiple integral. It is important to note that the integration is over only a wave number space.

For convenience of exposition, we consider the case $d=2.$, and later we will generalize this to any d. We define the Laplacian operator on the complex valued random variables, $J_{\mathbf{s}}(\omega)$ and $J_{\mathbf{s},e}(\omega)$ (here $\mathbf{s}=(s_1,s_2)$). Let $\upsilon>0$, and define the model

$$
\left[\frac{\partial^2}{\partial s_1^2}+\frac{\partial^2}{\partial s_2^2}-\left|c(\omega)\right|^2\right]^{\nu}J_s(\omega)=J_{s,e}(\omega)\,. \tag{18}
$$

where $J_s(\omega)$ and $J_{s,e}(\omega)$ are given by (14) and (17) respectively. We will see the significance of the frequency dependent function $c(\omega)$ in the above equation when we specialize the case $\nu = 1$. Now substitute the representations (14) and (17) in (18) and taking the operators inside the integrands and equating the integrands both sides of the equations (because of the uniqueness of the Fourier transforms this is valid), we obtain

$$\left(-\lambda_1^2 - \lambda_2^2 - |c(\omega)|^2\right)^{\nu} dZ_z(\underline{\lambda}, \omega) = dZ_e(\underline{\lambda}, \omega), \tag{19}$$

where $\underline{\lambda} = (\lambda_1, \lambda_2)$. Taking the modulus square, and taking expectations both sides of the modulus squares we obtain the spatio-temporal spectral density function. of the spatio-temporal process $Z(s,t)$ satisfying the above model (18) and it is given by

$$f_z(\underline{\lambda}, \omega) = \frac{\sigma_e^2}{(2\pi)^2 \left(\lambda_1^2 + \lambda_2^2 + |c(\omega)|^2\right)^{2\nu}} \tag{20}$$

Now we use the result for inverse transforms given in Whittle (1954.equation (65)) (note that we are considering the Isotropic processes) to obtain the covariance function. We have,

$$\frac{1}{4\pi^2} \int\int \frac{e^{i(x\omega_1 + y\omega_2)}}{\left(\omega_1^2 + \omega_2^2 + \alpha^2\right)^{\mu+1}} d\omega_1 d\omega_2 = \frac{1}{2\pi}\left(\frac{r}{2\alpha}\right)^{\mu} \frac{K_\mu(\alpha r)}{\Gamma(\mu+1)},$$

where $r = \left(x^2 + y^2\right)^{\frac{1}{2}}$, $K_\mu(x)$ is the modified Bessel function of the second kind of order μ, Using the above result to obtain the inverse transform of (20) ,we obtain

$$\begin{aligned} \frac{\sigma_e^2}{(2\pi)^2} \int\int \frac{e^{i(h_1\lambda_1 + h_2\lambda_2)}}{\left(\lambda_1^2 + \lambda_2^2 + |c(\omega)|^2\right)^{2\nu}} \quad & d\lambda_1 d\lambda_2 \\ &= \frac{\sigma_e^2}{2\pi}\left(\frac{\|\mathbf{h}\|}{2|c(\omega)|}\right)^{2\nu-1} \frac{K_{2\nu-1}(|c(\omega)|\,\|\mathbf{h}\|)}{\Gamma(2\nu)} \end{aligned} \tag{21}$$

where $\|\mathbf{h}\| = \left(h_1^2 + h_2^2\right)^{\frac{1}{2}}$. We note that ,under isotropy assumption, (21) is the covariance between the discrete Fourier Transforms $J_{\mathbf{s}}(\omega)$ and $J_{\mathbf{s+h}}(\omega)$, where $\mathbf{h} = (h_1, h_2)$. This covariance function is a function of the distance$\|\mathbf{h}\|$ and the temporal frequency ω. Hence, we have

$$\begin{aligned} Cov\left(J_{\mathbf{s}}(\omega), J_{\mathbf{s+h}}(\omega)\right) &= C(\|\mathbf{h}\|, \omega) \\ &= \frac{\sigma_e^2}{2\pi}\left(\frac{\|\mathbf{h}\|}{2|c(\omega)|}\right)^{2\nu-1} \frac{K_{2\nu-1}(|c(\omega)|\,\|\mathbf{h}\|)}{\Gamma(2\nu)} \end{aligned} \tag{22}$$

To see the significance of inclusion of $|C(\omega)|$ in the model (18), we study the limiting behavior of (22) as $||h|| \to 0$. We have noted earlier that $Var\left(J_s\left(\omega\right)\right)$ is proportional to the spectral density function $g\left(\omega\right)$ of the random process for all s. So it is interesting to examine the behavior of $C\left(\|\mathbf{h}\|,\omega\right)$ when $\|\mathbf{h}\| \to 0$, as the limit must tend to the second order spectral density function $g\left(\omega\right)$ of the spatio-temporal process defined earlier.

It is well known that, for all $\nu > 0$,

$$\lim_{x\to 0} \frac{x^{\nu} K_{\nu}\left(x\right)}{2^{\nu-1}\Gamma\left(\nu\right)} = 1. \tag{23}$$

Therefore, if we take the limit of $C\left(\|\mathbf{h}\|,\omega\right)$ given by (22) as $\|\mathbf{h}\| \to 0$, we get (using (23)),

$$C\left(0,\omega\right) = \frac{\sigma_e^2}{2\left(\left|c\left(\omega\right)\right|^2\right)^{2\nu-1}\left(2\nu-1\right)} = g\left(\omega\right). \tag{24}$$

From (22) and (24), we obtain the correlation coefficient

$$\begin{aligned} \rho\left(\|\mathbf{h}\|,\omega\right) &= \frac{C\left(\|\mathbf{h}\|,\omega\right)}{C\left(0,\omega\right)} = \\ &= \frac{\left(\|\mathbf{h}\|\left|c\left(\omega\right)\right|\right)^{2\nu-1}}{2^{2\nu-2}\Gamma\left(2\nu-1\right)} K_{2\nu-1}\left(\left|c\left(\omega\right)\right|\|\mathbf{h}\|\right). \end{aligned} \tag{25}$$

From (24), we observe that the function $\left|c\left(\omega\right)\right|$, which we used in defining the model (18) is in fact related to the second order spectral density function.

Consider the case of general d. Let $\rho = \|\underline{\lambda}\|$. We have ,

$$\begin{aligned} C\left(\|\mathbf{h}\|,\omega\right) &= \frac{\sigma_e^2}{(2\pi)^d}\int_{\mathbb{R}^d} \frac{e^{-i\mathbf{h}\cdot\underline{\lambda}}}{\left(\|\underline{\lambda}\|^2 + \left|c\left(\omega\right)\right|^2\right)^{2\nu}} d\underline{\lambda} \\ &= \frac{\sigma_e^2}{(2\pi)^d}\int_0^{\infty} \frac{\rho^{d-1}}{\left(\rho^2 + \left|c\left(\omega\right)\right|^2\right)^{2\nu}} \int_{\mathbb{S}_{d-1}} e^{-i\rho\|\mathbf{h}\|\cos\alpha} d\Omega d\rho \end{aligned}$$

where $\mathbb{S}_{d-1}$ is the unit sphere in $\mathbb{R}^d$ and Ω is Lebesque element of surface area on $\mathbb{S}_{d-1}$. Further we know

$$\int_{\mathbb{S}_{d-1}} e^{-i\rho\|\mathbf{h}\|\cos\alpha} d\Omega = (2\pi)^{\frac{d}{2}} \left(\rho\|\mathbf{h}\|\right)^{-\frac{d}{2}+1} \mathcal{J}_{\frac{d}{2}-1}\left(\rho\|\mathbf{h}\|\right),$$

where $\mathcal{J}_{\frac{d}{2}-1}$ denotes the Bessel function of the first kind, see [SW71], p.176. Now we use Hankel-Nicholson Type Integral, see [AS92], 11.4.44, if $d < 4\nu + 3$, then

$$\int_0^\infty \frac{\mathcal{J}_{\frac{d}{2}-1}(r\rho)}{\left(\rho^2 + |c(\omega)|^2\right)^{2\nu}} \rho^{\frac{d}{2}} d\rho = \frac{r^{2\nu-1} |c(\omega)|^{\frac{d}{2}-2\nu}}{2^{2\nu-1}\Gamma(2\nu)} K_{\frac{d}{2}-2\nu}(r|c(\omega)|).$$

Using the above integrals and noting $K_{\frac{d}{2}-2\nu} = K_{2\nu-\frac{d}{2}}$, we obtain for all d the covariance function

$$C(\|\mathbf{h}\|, \omega) = \frac{\sigma_e^2}{(2\pi)^{\frac{d}{2}} 2^{2\nu-1}\Gamma(2\nu)} \left(\frac{\|\mathbf{h}\|}{|c(\omega)|}\right)^{2\nu-\frac{d}{2}} K_{2\nu-\frac{d}{2}}(\|\mathbf{h}\| |c(\omega)|),$$

and the correlation function is given by

$$\rho(\|\mathbf{h}\|, \omega) = \frac{(\|\mathbf{h}\| |c(\omega)|)^{2\nu-\frac{d}{2}}}{2^{2\nu-\frac{d}{2}-1}\Gamma\left(2\nu-\frac{d}{2}\right)} K_{2\nu-\frac{d}{2}}(\|\mathbf{h}\| |c(\omega)|),$$

because

$$C(0, \omega) = \frac{\sigma_e^2}{(2\pi)^{\frac{d}{2}} 2^{\frac{d}{2}} \left(|c(\omega)|^2\right)^{2\nu-\frac{d}{2}}} \frac{\Gamma\left(2\nu-\frac{d}{2}\right)}{\Gamma(2\nu)} = g(\omega).$$

3.1 Special case:

To understand the significance of the equation (24), we consider the special case $\nu = 1$. By substituting $\nu = 1$ in (25), we obtain

$$\rho(\|\mathbf{h}\|, \omega) = (\|\mathbf{h}\| |c(\omega)|) K_1(\|\mathbf{h}\| |c(\omega)|),$$

a well known form except that the argument of the Bessel function is in terms of the function of $c(\omega)$,a function of the temporal frequency. From (24), we get when $\nu = 1$

$$C(0, \omega) = \frac{\sigma_e^2}{2|c(\omega)|^2} = g(\omega) > 0,$$

which implies that $|c(\omega)|^2$ is proportional to $g^{-1}(\omega)$, which is defined as the inverse second order spectral density function of the process. Let us assume that $g^{-1}(\omega)$ is absolutely integrable, then $g^{-1}(\omega)$ can be expanded in Fourier series,

$$g^{-1}(\omega) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} ci(k) \cos k\omega, \quad |\omega| \leq \pi.$$

where we used the fact that $g^{-1}(\omega) = g^{-1}(-\omega)$. The functions $\{ci(k)\}$ are known as inverse autocovariance functions, and are usually used to estimate the order of linear time series models. For example, if the series $\{Z(\mathbf{s}, t)\}$ satisfies for each s an autoregressive model of order p say, then it can easily be shown that $ci(k) = 0$ for all $k > p$, hence can be used to determine the order of the model. In other words, the covariance function $C(\|\mathbf{h}\|, \omega)$ which is in terms of the modified Bessel function is related to the spatial distance $\|\mathbf{h}\|$, and also a function which is related to the temporal dependence.

Now it is interesting to examine the case when $g(\omega)$ is independent of the frequency, which implies that the time series at each location is a white noise process. This assumption in turn implies that $|c(\omega)|^2$ is a constant, say α^2. Substituting this in (20) we see that the spatio-temporal spectral density function is proportional to only a positive definite function of the wave number λ.

3.2 Long Memory Process

If the spatio-temporal processes exhibits the long memory property, it is possible to accommodate this property in our definition of the model (18) by choosing appropriately the function $|c(\omega)|$ as follows.

Suppose the process $\{Z(\mathbf{s}, t)\}$, for each s, exhibits the long memory property (see Beran et al (2013) []) and we assume it can be modelled by a Fractional Autoregressive Moving Average Model (FARIMA) (p, d, q) of the form

$$\phi_p(B)(1-B)^d Z(\mathbf{s}, t) = \psi_q(B) e(\mathbf{s}, t)$$

where $-1/2 < d < 1/2$ and the process $\{e(\mathbf{s}, t)\}$ is a Gausian white noise in space and time. Then, it is well known that the spectral density function of the stationary process $\{Z(\mathbf{s}, t)\}$ is given by

$$g(\omega) = \frac{\sigma_e^2}{2\pi} \left| \frac{\psi_q(e^{-i\omega})}{\phi_p(e^{-i\omega})} \right|^2 |1 - e^{-i\omega}|^{-2d}, \ |\omega| \leq \pi$$

If we choose $g(\omega)$ as above and since $|c(\omega)|$ and $g(\omega)$ are related as in equation (24), by using the function $c(\omega)$ thus obtained, we get a spatio-temporal process with long memory property. The long memory parameter d can be estimated using the criterion to be defined in later sections. We hope to consider the estimation and their properties in detail in later publications.

4 Estimation of the parameters of the covariance function.

Let $C(\|\mathbf{h}\|, \omega)$, given by (22), be a function of the parameter vector $\underline{\vartheta}$, and now onwards we denote this function by $C(\|\mathbf{h}\|, \omega;\ \underline{\vartheta})$, and similarly we write the correlation function given by (25) as $\rho(\|\mathbf{h}\|, \omega;\ \underline{\vartheta})$. Our object is to estimate $\underline{\vartheta}$. We note that ω is the temporal frequency, $\|\mathbf{h}\|$ is the spatial Euclidean distance. To estimate the parameters $\underline{\vartheta}$, we use the frequency domain method recently proposed by Subba Rao [SRDB13] based on frequency variogram defined. We briefly summarize the procedure of [SRDB13]. We now define a new spatio temporal random process from $\{Z(s,t)\}$.

$$Y_{ij}(t) = Z(\mathbf{s}_i, t) - Z(\mathbf{s}_j, t), \quad \text{for each } \ t = 1, 2, \ldots, n$$

and for all locations $\mathbf{s}_i$, $\mathbf{s}_j$ where $\mathbf{s}_i$ and $\mathbf{s}_j$ $(i \neq j)$are the pairs that belong to the set $N(\mathbf{h}_l) = \{\mathbf{s}_i, \mathbf{s}_j;\ \|\mathbf{s}_i - \mathbf{s}_j\| = \|\mathbf{h}_l\|,\ l = 1, 2, \ldots, L\}$. Define the Finite Fourier transform (F.T.) of the new time series $\{Y_{ij}(t);\ i \neq j\}$ at the Fourier frequencies $\omega_k = \frac{2\pi k}{n}, k = 0, 1, \ldots, \left[\frac{n}{2}\right]$,

$$J_{\mathbf{s}_i,\mathbf{s}_j}(\omega_k) = \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^{n} Y_{ij}(t) e^{-it\omega_k} = J_{\mathbf{s}_i}(\omega_k) - J_{\mathbf{s}_j}(\omega_k), \qquad (26)$$

where

$$J_{\mathbf{s}_i}(\omega_k) = \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^{n} Z(\mathbf{s}_i, t) e^{-it\omega_k},\ (i = 1, 2, \ldots, m)\,.\text{for all } i$$

Let $I_{\mathbf{s}_i,\mathbf{s}_j}(\omega_k)$ be the second order periodogram of the time series $\{Y_{ij}(t)\}$ given by

$$I_{\mathbf{s}_i,\mathbf{s}_j}(\omega_k) = \left|J_{\mathbf{s}_i,\mathbf{s}_j}(\omega_k)\right|^2 = \frac{1}{2\pi} \sum_{u=-(n-1)}^{n-1} \hat{c}_{y,ij}(u) e^{-iu\omega_k},$$

where

$$\hat{c}_{y,ij} = \frac{1}{n} \sum_{t=1}^{n-|u|} \left(Y_{ij}(t+u) - \overline{Y}_{ij}\right)\left(Y_{ij}(t) - \overline{Y}_{ij}\right),\ |u| \leq n-1$$

is the sample autocovariance of lag u of the time series $\{Y_{ij}(t)\}$, and $\overline{Y}_{ij} = \frac{1}{n}\sum_{t=1}^{n} Y_{ij}(t)$. From (26), we obtain

$$E\left[I_{\mathbf{s}_i,\mathbf{s}_j}(\omega_k)\right] = E\left[I_{\mathbf{s}_i}(\omega_k)\right] + E\left[I_{\mathbf{s}_j}(\omega_k)\right] - 2\ Real\ E\left[I_{\mathbf{s}_i\mathbf{s}_j}(\omega_k)\right], \quad (27)$$

where $I_{\mathbf{s}_i\mathbf{s}_j}(\omega_k)$ is the cross periodogram between the processes $\{Z(\mathbf{s}_i, t)\}$ and $\{Z(\mathbf{s}_j, t)\}$. For large n, we can show for an isotropic process, the expectation (27) is

$$g_{\mathbf{s}_i,\mathbf{s}_j}(\omega_k; \underline{\vartheta}) = g_{\|\mathbf{h}\|}(\omega_k; \underline{\vartheta}) = 2\left[C(0, \omega_k;\ \underline{\vartheta}) - C(\|\mathbf{h}\|, \omega_k;\ \underline{\vartheta})\right], \qquad (28)$$

where $g_{\|\mathbf{h}\|}(\omega_k; \underline{\vartheta})$ is the spectral density function of the stationary process $\{Y_{ij}(t)\}$, $C(0, \omega_k;\ \underline{\vartheta}) = g(\omega_k, \vartheta)$ is the spectral density of the process $\{Z(\mathbf{s}_i, t)\}$ for all i, and $C(\|\mathbf{h}\|, \omega_k;\ \underline{\vartheta})$ is the spatio temporal covariance between $J_{\mathbf{s}_i}(\omega_k)$ and $J_{\mathbf{s}_j}(\omega_k)$ an expression of which we obtained earlier (22). [SRDB13] $g_{\mathbf{s}_i,\mathbf{s}_j}(\omega_k; \underline{\vartheta})$ as the frequency domain version of the variogram. It has similar properties as in the case of time domain as can be shown below.

4.1 Frequency Variogram, measurement errors and Nugget Effect

Consider the variogram equation (28). We would expect, the function $g_{\|\mathbf{h}\|}(\omega; \underline{\vartheta})$ if plotted against the distance $\| \mathbf{h} \|$(for all ω), will pass through the origin(i.e. as $\| \mathbf{h} \|$ tends to zero). Let us assume that there are measurement errors in the observed data, i.e. we observe $\widetilde{Z}(\mathbf{s}, t)$ instead of $Z(\mathbf{s}, t)$, where for each $\mathbf{s}$ and t, $\widetilde{Z}(\mathbf{s}, t) = Z(\mathbf{s}, t) + \eta(\mathbf{s}, t)$. We assume that the random errors $\eta(\mathbf{s}, t)$ are independent of $Z(\mathbf{s}, t)$, and is a white noise process in both $\mathbf{s}$ and t,. Also assume that it has zero mean and variance equal to σ_η^2. Then we can easily show that the function $g_{\|\mathbf{h}\|}(\omega; \underline{\vartheta})$ instead of passing through the origin will have a jump of magnitude proportional to σ_η^2 which will be the second order spectral density function of the white noise process $\eta(s, t)$. This is called Nugget effect in the context of spatial analysis.

Now for the estimation of the parameter vector $\underline{\vartheta}$ we proceed as in [SRDB13]. We consider the complex valued random vector,

$$\underline{J}'_{\|\mathbf{h}\|}(\omega) = \left[J_{\mathbf{s}_i,\mathbf{s}_j}(\omega_1), J_{\mathbf{s}_i,\mathbf{s}_j}(\omega_2), \ldots, J_{\mathbf{s}_i,\mathbf{s}_j}(\omega_M)\right]$$

which is distributed asymptotically as normal with mean zero and with variance covariance matrix with diagonal elements $\left[g_{\|\mathbf{h}\|}(\omega_1), g_{\|\mathbf{h}\|}(\omega_2),\right.$

$\ldots, g_{\|\mathbf{h}\|}(\omega_M)\big]$. We note that because of asymptotic independence of Fourier transforms at Fourier frequencies defined, the off diagonal elements of the variance covariance matrix are zero. The minus log likelihood function can be shown to be proportional to

$$Q_{n,N(\mathbf{h})}(\underline{\vartheta}) = \frac{1}{|N(\mathbf{h})|} \sum_{(\mathbf{s}_i,\mathbf{s}_j)\in N(\mathbf{h})} \sum_{k=1}^{M} \left[\ln g_{\mathbf{s}_i,\mathbf{s}_j}(\omega_k;\underline{\vartheta}) + \frac{I_{\mathbf{s}_i,\mathbf{s}_j}(\omega_k)}{g_{\mathbf{s}_i,\mathbf{s}_j}(\omega_k;\underline{\vartheta})} \right]. \tag{29}$$

where $g_{\mathbf{s}_i,\mathbf{s}_j}(\omega_k;\underline{\vartheta})$ is the variogram given by (28) and $I_{\mathbf{s}_i,\mathbf{s}_j}(\omega_k)$ is the periodogram of the $\{Y_{ij}(t)\}$. Here $N(\mathbf{h})$ is the collection of all distinct pairs $\mathbf{s}_i$ and $\mathbf{s}_j$ such that $N(\mathbf{h}) = \{(\mathbf{s}_i, \mathbf{s}_j);\ \|\mathbf{s}_i - \mathbf{s}_j\| = N(\mathbf{h})\}$. The above criterion (29) is defined only for one distance $\|\mathbf{h}\|$. Suppose we now define L spatial distances from the above data, we can define an over all criterion

$$Q_n(\vartheta) = \frac{1}{L} \sum_{l=1}^{L} Q_{n,N(\mathbf{h}_l)}(\underline{\vartheta}) \tag{30}$$

and minimize (30) with respect to $\underline{\vartheta}$. In defining the above, we have given equal weights to all the distances. The asymptotic normality of the estimator $\underline{\vartheta}$ obtained by minimizing (30) has been proved in Theorem 2 of the paper of [SRDB13]. It has been shown, that under certain conditions,

$$\sqrt{n}\left(\underline{\vartheta}_n - \vartheta_0\right) \xrightarrow{D} N\left(\underline{0}, \nabla^2 Q_n^{-1}(\underline{\vartheta}_0)\, V \nabla^2 Q_n^{-1}(\underline{\vartheta}_0)\right),$$

where $V = \lim_{n\to\infty} var\left[\frac{1}{\sqrt{n}} \nabla Q_n(\vartheta_0)\right]$, $\nabla Q_n(\vartheta_0)$ is a vector of first order partial derivatives, $\nabla^2 Q_n(\vartheta_0)$ is the matrix of second order partial derivatives. In view of the relation (28) and because (29) and (30) are in terms of the frequency variogram $g_{\mathbf{s}_i,\mathbf{s}_j}(\omega_k;\vartheta)$, we can rewrite the above expression in (29) in terms of $\rho(\|\mathbf{h}\|, \omega;\ \vartheta)$ as well. We note

$$g_{\mathbf{s}_i,\mathbf{s}_j}(\omega_k;\vartheta) = 2C(0,\omega_k;\ \vartheta)\left[1 - \rho(\|\mathbf{h}\|, \omega;\ \vartheta)\right], \tag{31}$$

where $|\rho(\|\mathbf{h}\|, \omega_k;\ \vartheta)| \le 1$. This correlation coefficient is the coherency coefficient defined earlier by Subba Rao et al [SRDB13]. Therefore, we can rewrite (29) as

$$\begin{aligned} Q_{n,N(\mathbf{h})}(\underline{\vartheta}) \quad = \quad & \frac{1}{|N(\mathbf{h})|} \sum_{(\mathbf{s}_i,\mathbf{s}_j)\in N(\mathbf{h})} \sum_{k=1}^{M} \Big[\ln 2 + \ln C(0,\omega_k;\ \vartheta) \\ & + \ln\left[1 - \rho(\|\mathbf{h}\|, \omega_k;\ \vartheta)\right] + \frac{I_{\mathbf{s}_i,\mathbf{s}_j}(\omega_k)}{2C(0,\omega_k;\ \vartheta)} \left[1 - \rho(\|\mathbf{h}\|, \omega_k;\ \vartheta)\right]^{-1} \Big]. \end{aligned}$$

Since $|\rho(\|\mathbf{h}\|, \omega_k;\ \underline{\vartheta})| < 1$, as an approximation we may consider alternatively minimizing, $\overline{Q}_{n,N(\mathbf{h})}(\underline{\vartheta})$, given by

$$\begin{aligned}\overline{Q}_{n,N(\mathbf{h})}(\underline{\vartheta}) &\simeq \frac{1}{|N(\mathbf{h})|} \sum_{(\mathbf{s}_i,\mathbf{s}_j)\in N(\mathbf{h})} \sum_{k=1}^{M} \Big[\ln C(0,\omega_k;\underline{\vartheta}) \\ &\quad + \rho(\|\mathbf{h}\|, \omega_k; \vartheta) + \frac{I_{\mathbf{s}_i,\mathbf{s}_j}(\omega_k)}{2C(0,\omega_k;\ \vartheta)}\rho(\|\mathbf{h}\|, \omega_k; \vartheta)\Big].\end{aligned}$$

5 Spatio-temporal prediction

Our object in this section is to estimate $\{Z(\mathbf{s},t);\ t=1,2,\ldots,n\}$ at the location $\mathbf{s}_0$ given the $m-$time series $\{Z(\mathbf{s}_i,t);\ i=1,2,\ldots,m; t=1,2,\ldots,n\}$ from the spatio-temporal stationary and isotropic process $\{Z(\mathbf{s},t)\}$. In other words, we are estimating the entire data set at the location $\mathbf{s}_0$ over the same period . Using the estimated set of observations at the location $\mathbf{s}_0$, we will also obtain optimal linear predictors of the future values, following a methodology similar to Box and Jenkins [BJ76].As in the case of the observed data $\{Z(\mathbf{s}_i,t)\}$, we define the discrete Fourier transform of $\{Z(\mathbf{s}_0,t)\}$ the data of which is not available, by

$$J_{\mathbf{s}_0}(\omega) = \frac{1}{\sqrt{(2\pi n)}} \sum_{t=1}^{n} Z(\mathbf{s}_0,t)\, e^{-it\omega}, \tag{32}$$

and by inversion, we have

$$Z(\mathbf{s}_0,t) = \sqrt{\frac{n}{2\pi}} \int_{-\pi}^{\pi} J_{\mathbf{s}_0}(\omega)\, e^{it\omega} d\omega. \tag{33}$$

In other words given $\{J_{\mathbf{s}_0}(\omega),\ \text{for all } -\pi \le \omega \le \pi\}$, we can uniquely recover the time series sequence $\{Z(\mathbf{s}_0,t);\ t=1,\ldots,n\}$. In view of this observation, we consider here the estimation of Fourier Transform $J_{\mathbf{s}_0}(\omega)$ for all ω, and from these complex valued observations, we can estimate $\{Z(\mathbf{s}_0,t)\}$ using the above equation(33). Consider the vector of the discrete Fourier transforms at a single frequency ω,

$$\underline{J}'_m(\omega) = [J_{\mathbf{s}_1}(\omega), J_{\mathbf{s}_2}(\omega), \ldots, J_{\mathbf{s}_m}(\omega)].$$

We note

$$\begin{aligned} E[\underline{J}_m(\omega)] &= 0 \\ E[\underline{J}_m(\omega)\underline{J}^*_m(\omega)] &= \underline{F}_m(\omega), \end{aligned} \tag{34}$$

where the square matrix $F_m(\omega) = (C(\|\mathbf{s}_i - \mathbf{s}_j\|, \omega)\,;\ i,j = 1,2,\ldots,m)$, and each element of $C(\|\mathbf{s}_i - \mathbf{s}_j\|, \omega)$ is given by (22). The complex random vector $\underline{J}_m(\omega)$ has a multivariate complex Gaussian distribution with mean zero and variance covariance matrix $\underline{F}_m(\omega)$, and the matrix is real and symmetric. Consider now the $(m+1)$ dimensional complex valued random vector,

$$\underline{J}'_{m+1}(\omega) = \left[J_0(\omega), \underline{J}'_m(\omega)\right],$$

which has zero mean, and variance covariance matrix

$$\begin{aligned} E\left[\underline{J}_{m+1}(\omega)\,\underline{J}^*_{m+1}(\omega)\right] &= \begin{bmatrix} C_0(0,\omega) & E\left(J_0(\omega)\,\underline{J}^{*\prime}_m(\omega)\right) \\ E\left(\underline{J}_m(\omega)\,J^*_0(\omega)\right) & E\left(\underline{J}_m(\omega)\,\underline{J}^*_m(\omega)\right) \end{bmatrix} \\ &= \begin{bmatrix} C_0(0,\omega) & \underline{G}'_0(\omega) \\ \underline{G}_0(\omega) & \underline{F}_m(\omega) \end{bmatrix}, \end{aligned}$$

where $C_0(0,\omega) = E\left(J_0(\omega)\,J^*_0(\omega)\right) = C(0,\omega)$, which is the second order spectral density function of the spatial process and it is given by (24), and the row vector $\underline{G}'_0(\omega)$ is given by

$$\begin{aligned} \underline{G}'_0(\omega) &= E\left[J_0(\omega)\,J^{*\prime}_m(\omega)\right] \\ &= \left[C(\|\mathbf{s}_0 - \mathbf{s}_1\|, \omega), C(\|\mathbf{s}_0 - \mathbf{s}_2\|, \omega), \ldots, C(\|\mathbf{s}_0 - \mathbf{s}_m\|, \omega)\right] \end{aligned}$$

and $\underline{F}_m(\omega)$ is defined above.. Therefore, the optimal linear least squares prediction of $J_0(\omega)$ given the vector $\underline{J}_m(\omega)$, is given by the conditional expectation

$$E\left[J_0(\omega)\,|\underline{J}_m(\omega)\right] = \underline{G}'_0(\omega)\,\underline{F}^{-1}_m(\omega)\,\underline{J}_m(\omega) \tag{35}$$

and the minimum mean squared error is given by

$$\sigma^2_m(\omega) = C(0,\omega) - \underline{G}'_0(\omega)\,\underline{F}^{-1}_m(\omega)\,\underline{G}_0(\omega)\,. \tag{36}$$

It is interesting and important to note from the equations (35 and 36) that the evaluation of the conditional expectation and the minimum mean square error involves only inversion of $m \times m$ dimensional matrices, unlike in the case of the time domain approach for prediction where one needs to invert $mn \times mn$ dimensional matrices. In many real data analysis the number of time points n will be very large. and m can be large too . Besides, there is no ordering problem involved here (see [CW11] p. 324). Once we have an expression for covariance function $C(\|\mathbf{h}\|, \omega)$, all the elements of the column vector $G_0(\omega)$ and the elements of $F_m(\omega)$ are known. By substituting the relevant expressions, we can evaluate (35) and (36). Usually , the covariance functions will have parameters which need to be estimated and this was considered earlier.

In obtaining the above, we assumed that the mean of the random process i.e. $E(Z(\mathbf{s},t)) = 0$,which is known as Ordinary Kriging in the literature. If $E(Z(\mathbf{s},t)) = \mu(\mathbf{s},t) \neq 0$, the usual approach is to model it in terms of covariates, and consider the estimation of the parameters. Once the parameters are estimated, one can use the estimated mean function in defining the Fourier Transforms., and consider the estimation and prediction as before. The estimation and the properties of the estimators will be considered in later publications.

. Let us now denote the estimate of the conditional expectation (35) by $\hat{J}_0(\omega)$, and therefore the estimate of $J_0(\omega)$ is given by

$$\hat{J}_0(\omega) = \underline{\hat{G}}_0'(\omega)\,\underline{\hat{F}}_m^{-1}(\omega)\,J_m(\omega)\,. \tag{37}$$

We can now estimate the time series by inverting using the equation (33). As pointed out earlier,in evaluating the above expression we need the spatio-temporal covariances, and the estimates of the parameters of the covariance functions which were considered earlier.. These involve unknown parameters, and can be estimated using methods described earlier. We use these estimated covariances to calculate the above Fourier transform. Using the above estimated Fourier Transform, we can now estimate the entire time series $\{Z(\mathbf{s}_0,t);\ t = 1,2,\ldots,n\}$.,.Using this as our data, we can find optimal linear predictors of $Z(\mathbf{s}_0, n+v)$ for all $v >$ .0.The methodology (Box and Jenkins, 1976) is well known, and hence details are omitted. In the parametric approach, a suitable linear time series model is assumed and fitted and, then the fitted model is used for the prediction of the future values. Since we have already computed Fourier transforms, it is convenient to describe the frequency domain approach for the estimation, which is based on Whittle likelihood approximation.

Briefly we describe Whittle likelihood approach. Let us assume that the second order stationary time series $\{Z(\mathbf{s}_0,t)\}$ satisfies a linear time series model, and let $g_0\left(\omega, \underline{\psi}\right)$ $|\omega| \leq \pi$ denote the second order spectral density function of the process and let the parameter vector be denoted by a q dimensional vector $\underline{\psi}' := \left(\psi_1, \psi_2, \ldots, \psi_q\right)$. We have the Fourier Transforms $J_0(\omega)$ for all ω The parameter vector $\underline{\psi}$ can now be estimated by minimizing the approximate negative log likelihood function,.

$$\int \left[\ln g_0\left(\omega, \underline{\psi}\right) + \frac{|J_0(\omega)|^2}{g_0\left(\omega, \underline{\psi}\right)}\right] d\omega$$

with respect to $\underline{\psi}$. The asymptotic sampling properties of the estimator ψ are now well established. and hence will not be repeated here. Having fitted a linear model to the data, it is now possible to obtain the optimal forecasts of the future values. Since the methodology is well known we omit the details.

6 Simulations and Real Data Analysis:

In the following we briefly indicate the procedure to generate a stationary spatio-temporal random process$\{Z(\mathbf{s}_i, t)\,;\ i = 1, 2, \ldots, m; t = 1, 2, \ldots, n\}$ at locations $\{\mathbf{s}_{i\}}$ with a given covariance spatio-temporal covariance function of the form given by (22). In order to generate the data , we simulate the Discrete Fourier Transforms $\{\mathbf{J}(\omega_k),\ k = 0, 1, \ldots, n-1\}$,and then by inversion we obtain the data. We briefly outline the steps.

6.1 Simulation:

1. The locations $\mathbf{s}'_i;\ i = 1, 2, \ldots, m$ are chosen randomly from the unit square and scaled by the $\max(\|\mathbf{s}'_i\|)$, thus obtaining the scaled locations $\mathbf{s}_i = \mathbf{s}'_i / \max(\|\mathbf{s}'_i\|)$ with Euclidean distances $d_{i,j} = \|\mathbf{s}_i - \mathbf{s}_i\|$.We , note that $\max(\|\mathbf{s}_i\|) = 1$.Here for our illustration purposes we have chosen d=2 and the number of locations $m = 9$. We estimated the data at the location 10.

2. The number of observations n is even, $n = 2^{11}$, and we generated a series of independent complex Gaussian zero mean random vectors $\mathbf{U}_k$ of order $m \times 1$, $k = 0, 1, \ldots, n/2$, , such that $Var(\mathbf{U}_k) = I$ an identity matrix. We note that $\mathbf{U}_0$ and $\mathbf{U}_{n/2}$ are real valued random vectors ..

3. We assume the spatio-temporal process $Z(s, t)$, for each $\mathbf{s}$,satisfies an ARMA(2,1) model and has the second order spectrum given by $g(\omega) = (\sigma^2/2\pi)\left|\vartheta\left(e^{-2i\pi\omega}\right)/\varphi\left(e^{-2i\pi\omega}\right)\right|^2$. The roots of the polynomials $\varphi(z)$ and $\vartheta(z)$ are greater than one in modulus. We have chosen the polynomials of the form

$$\begin{aligned}\varphi(z) &= 1 + 4/17z + 4/17z^2, \\ \vartheta(z) &= 1 - 2/3z.\end{aligned}$$

4. Let $\omega_k = 2\pi k/n,\ k = 0, \ldots, n/2,\ \sigma = 2$, and let $C(0, \omega_k) = g(\omega_k)$,

where (the equations (24) and (25) with $[\nu = 1]$)

$$c(\omega_k) = \sqrt{\frac{1}{2g(\omega_k)}},$$

and

$$\rho(d_{i,j}, \omega_k) = (d_{i,j} |c(\omega_k)|) K_1 (d_{i,j} |c(\omega_k)|),$$

$i, j = 1, 2, \ldots, m$.

5. Generate a series of covariance matrices $\mathbb{C}_k$, $k = 0, 1, \ldots, n/2$, with order $m \times m$ and with entries $c_{i,j}(k) = C(0, \omega_k) \rho(d_{i,j}, \omega_k)$, (note $d_{i,j} = \|\mathbf{s}_i - \mathbf{s}_i\|$). Let the vectors $\mathbf{J}(\omega_k) = [J_{\mathbf{s}_i}(\omega_k)]_{i=1}^m = \sqrt{\mathbb{C}_k}\mathbf{U}_k$. The variance covariance matrix of $J(\omega_k) = \mathbb{C}_k$. For indices $k = n/2 + 1, \ldots, n-1$ we put $\mathbf{J}(\omega_k) = \overline{\mathbf{J}}\left(\omega_{n/2-k}\right)$.
6. From equation (33) we note that the inverse Fourier transform of $\mathbf{J}(\omega_k)$ gives us the spatio-temporal data $Z(\mathbf{s}_i, t)$.

6.2 Estimation of the parameters and Prediction of $\widehat{Z}(\mathbf{s}_0, t)$

We briefly describe the steps required to estimate the parameters ϑ and also the steps required for prediction. We now assume that we have the spatio-temporal data $\{Z(\mathbf{s}_i, t);\ i = 1, 2, \ldots, m; t = 1, 2, \ldots, n\}$.

1. 1. Let $Y_{ij}(t) = Z(s_i, t) - Z(s_j,t)$ as defined in section 4. We compute the Discrete Fourier Transform of the differenced series $Y_{ij}(t)$ which is the difference of the individual series $\{Z(s_i, t)\}$and $\{Z(s_j, t)\}$. The parameters of the correlation function is now estimated by minimizing the criterion (30) .
2. We choose a new location $\mathbf{s}_0$ randomly and rescale the locations if it is necessary.
3. The parameters of the polynomials $\varphi(z)$, $\vartheta(z)$ and variance σ^2 are estimated in the case of our simulated data(see section 6.1) , and they are given by $\widehat{\sigma} = 2.0744$,

$$\begin{aligned} \widehat{\varphi}(z) &= 1 + 0.2404z + 0.2356z^2, \\ \widehat{\vartheta}(z) &= 1 - 0.6406z. \end{aligned}$$

These estimates are very close to the true values .

	Budapest	Debrecen	Szeged	Szombathely
Budapest	0	195.5222	161.6584	182.4531
Debrecen	195.5222	0	180.6993	377.2415
Szeged	161.6584	180.6993	0	287.0954
Szombathely	182.4531	377.2415	287.0954	0

Table 1: Distances Matrixt

4. The matrix $\mathbb{C}_k$ above is calculated using the estimated parameters and the inverse $\underline{\hat{F}}_m^{-1}(\omega_k)$, is evaluated for each ω_k . We have

$$\underline{G}_0'(\omega_k) = \widehat{C}(0, \omega_k)\left[\widehat{\rho}(d_{0,1}, \omega_k), \widehat{\rho}(d_{0,2}, \omega_k), \ldots, \widehat{\rho}(d_{0,m}, \omega_k)\right],$$

   where $d_{0,j} = \|\mathbf{s}_0 - \mathbf{s}_1\|$.

5. The Fourier transform of the predicted series $\widehat{Z}(\mathbf{s}_0, t)$ is given by $\hat{J}_0(\omega_k) = \underline{\hat{G}}_0'(\omega_k)\,\underline{\hat{F}}_m^{-1}(\omega_k)\,\mathbf{J}(\omega_k)$. Now the inverse Fourier transform of $\hat{J}_0(\omega_k)$ is evaluated using the equation (33) (considering the discrete sum) gives the predicted time series $\widehat{Z}(\mathbf{s}_0, t)$ for all t. The Figure 1 shows the simulated data according to the model of the Section 6.1. and its prediction.

6. In Figure 2, the log of the true spectrum spectrum $g(\omega)$, the log of the estimated $\widehat{g}(\omega)$spectrum and log of the periodogram $\hat{J}_0(\omega_k)$, are plotted , and the smoothing of this function gives an estimate of the spectral density function. We note that the estimated periodogram is very close to the true underlying spectrum indicating the prediction methodology does give good results.

6.3 Real Data

We consider Hungarian Meteorological Service (OMSZ) Climate time Series (http://met.hu/eghajlat/magyarorszag_eghajlata/) 50 years(1951-2000) monthly averages ($n = 600$) at four locations, namely, Budapest, Debrecen, Szeged and Szombathely. In Table 1, the distances (in Km) between these locations are given. The time series plots are given in Figure .

Because of presence of 12 months seasonality, we fitted a harmonic regression of the following form, and the estimates of the coefficients are summarized in Table 2.

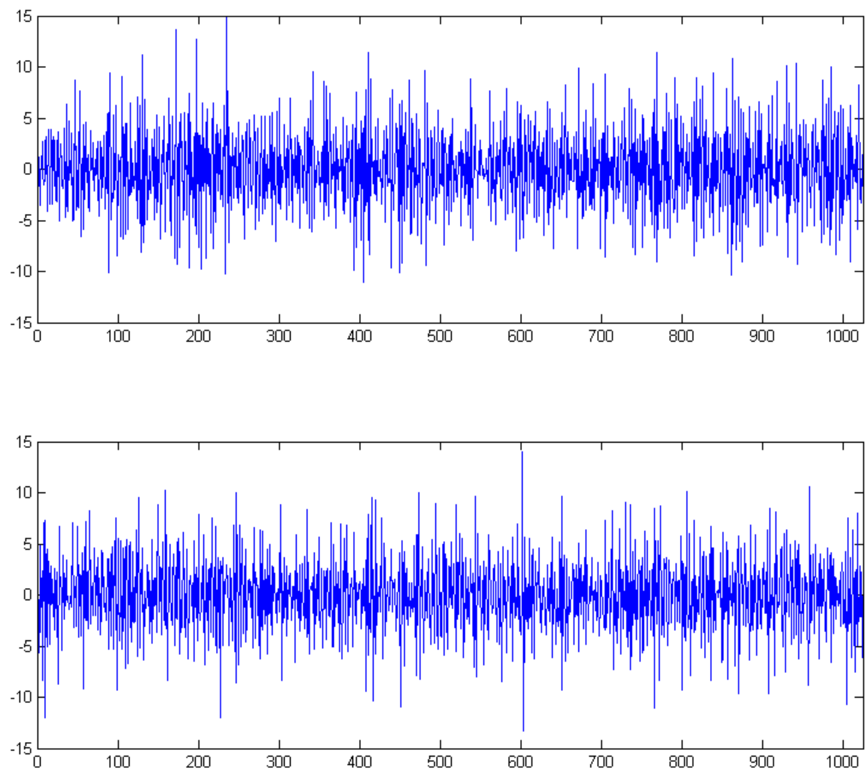

Figure 1: Simulated data (above) and its prediction

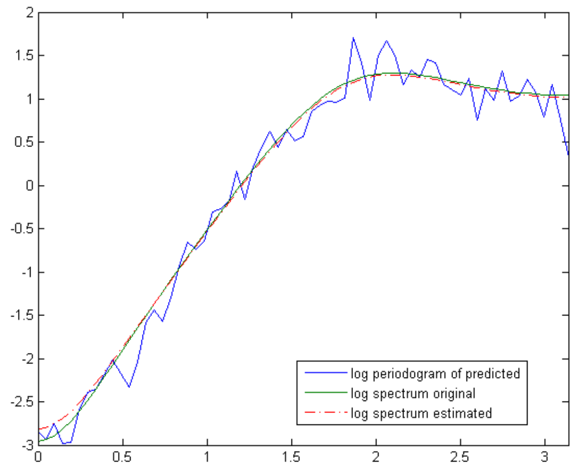

Figure 2: Log true spectrum and Log estimated spectrum (simulated)

Coeff	Coeff Budapest	Coeff Debrecen	Coeff Szeged	Coeff Szombathely
β_0	11.23400000	9.95600000	10.47483333	9.30866667
β_1	-9.39699315	-9.81061040	-9.68566405	-9.02813423
β_2	-5.41628301	-5.63468533	-5.68347870	-5.36693388

Table 2: Seasonal Parameter Estimates

$$X_t = \beta_0 + \beta_1 \cos\left(\frac{2\pi}{12}t\right) + \beta_2 \sin\left(\frac{2\pi}{12}t\right) + e_t,$$

It has been found that an ARMA (1,1) of the form $(1+\phi_1 B)Z(\mathbf{s},t) = (1+\theta_1 B)e(\mathbf{s},t)$ with var $e(\mathbf{s},t) = \sigma_e^2$ will fit well to the residuals thus obtained for the de-seasonalised data. The estimates of the parameters of these models for each location is summarized in the following Table 3.

The final estimates of the ARMA(1,1) model obtained by minimizing (30) are given by $\hat{\sigma}_e = 1.8501$, $\hat{\phi}_1 = -0.0296$, $\hat{\theta}_1 = 0.2231$. The estimated log spectrum from the original data and the log periodogram computed from the estimated data for each of the four cities is given in Figure 4. It is interesting to see the close behavior of the estimated log periodogram to the estimated log spectrum. It has to be noted that the log periodogram needs smoothing to obtain a consistent estimate.

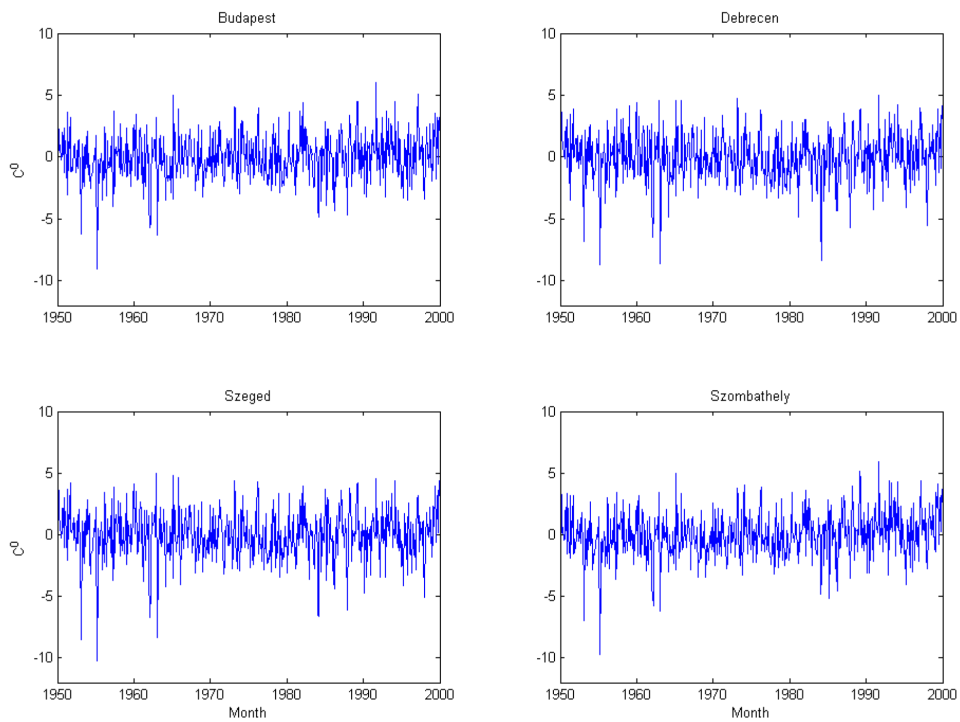

Figure 3: Time series plots of city temperatures

	$\hat{\sigma}_e^2$	$\hat{\phi}_1$	$\hat{\theta}_1$
Budapest	3.288912	0.137256	0.022879
Debrecen	3.768498	0.147554	0.042078
Szeged	3.795002	0.123754	0.086649
Szombathely	3.186594	0.547764	−0.382492

Table 3: Estimates of ARMA Parameters

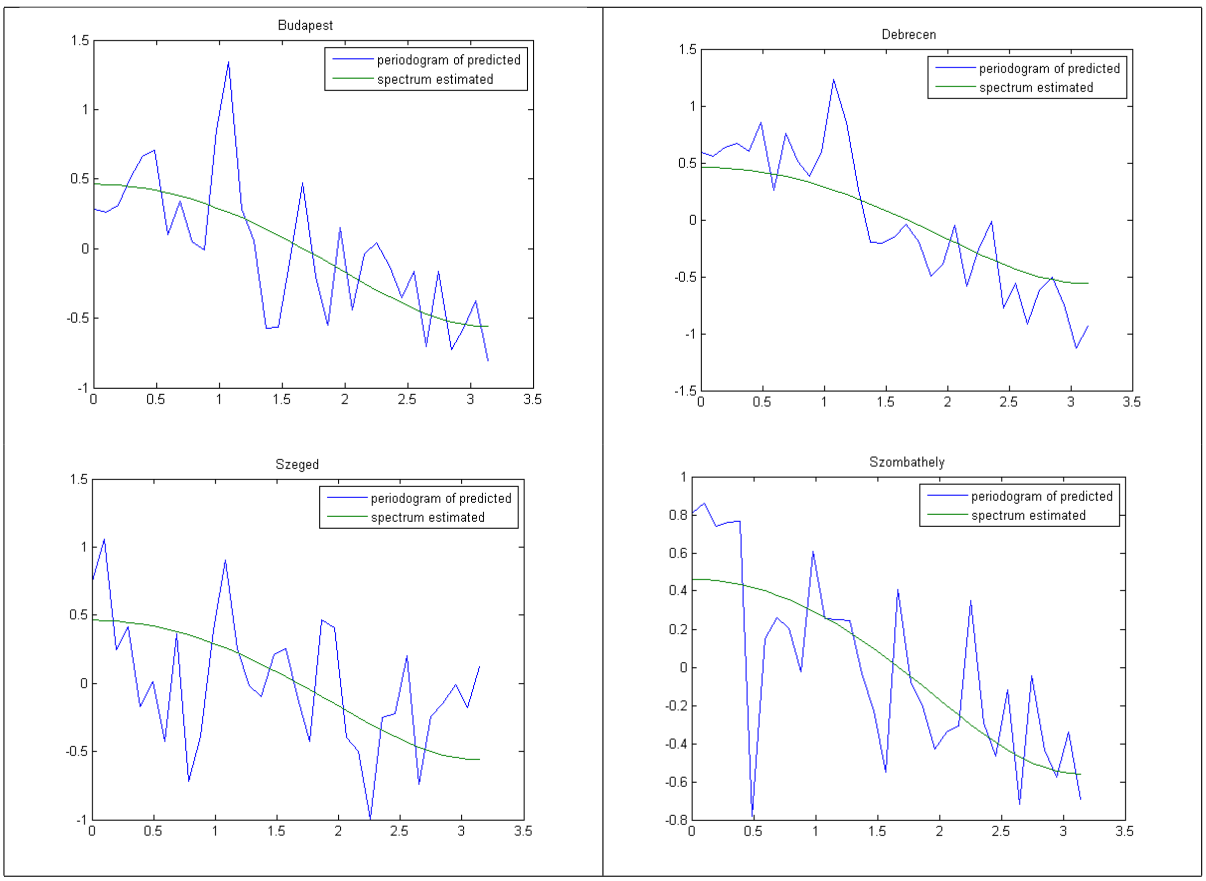

Figure 4: Log Periodogram vs Log Estimated Spectrum

Acknowledgement 1 *The publication was supported by the TÁMOP-4.2.2.C-11/1/KONV-2012-0001 project. The project has been supported by the European Union, co-financed by the European Social Fund. The visit of Subba Rao to the CRRAO AIMSCS was supported by a grant from the Department of Science and Technology,Government of India ,grant number SR/S4/516/07.*

References

[AS92] M. Abramowitz and I. A. Stegun. "Handbook of mathematical functions with formulas, graphs, and mathematical tables". Dover Publications Inc., New York (1992). Reprint of the 1972 edition.

[BJ76] G. E. P. Box and G. M. Jenkins. "Time Series Analysis, Forecasting and Control". Holden-Day, San Francisco (1976).

[Bri01] D. R. Brillinger. "Time Series; Data Analysis and Theory". Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA (2001). Reprint of the 1981 edition.

[CG11] Peter F Craigmile and Peter Guttorp. Space-time modelling of trends in temperature series. *Journal of Time Series Analysis* **32**(4), 378–395 (2011).

[CH99] Noel Cressie and Hsin-Cheng Huang. Classes of nonseparable, spatio-temporal stationary covariance functions. *Journal of the American Statistical Association* **94**(448), 1330–1339 (1999).

[Cre93] N. Cressie. "Statistics for Spatial Data". Wiley Series in Probability & Mathematical Statistics (1993).

[CW11] N. Cressie and C. K. Wikle. "Statistics for Spatio-Temporal Data". Wiley Series in Probability and Statistics (2011).

[DR07] Peter J Diggle and Paulo J Ribeiro. "Model-based geostatistics". Springer (2007).

[DSR11] Yogesh Dwivedi and Suhasini Subba Rao. A test for second-order stationarity of a time series based on the discrete fourier transform. *Journal of Time Series Analysis* **32**(1), 68–91 (2011).

[Gne02] T. Gneiting. Nonseparable, stationary covariance functions for space–time data. *Journal of the American Statistical Association* **97**(458), 590–600 (2002).

[JZ97] Richard H Jones and Yiming Zhang. Models for continuous stationary space-time processes. In "Modelling longitudinal and spatially correlated data", pages 289–298. Springer (1997).

[Ma02] C. Ma. Spatio-temporal covariance functions generated by mixtures. *Mathematical geology* **34**(8), 965–975 (2002).

[Pri81] M. B. Priestley. "Spectral Analysis and Time Series". Academic Press, New York (1981).

[She11] M. Sherman. "Spatial Statistics and Spatio-Temporal Data: Covariance Functions and Directional Properties (Wiley Series in Probability and Statistics)". John Wiley & Sons, Ltd (2011).

[SR13] T. Subba Rao. Time domain and frequency domain analysis of nonlinear astronomical time series. Technical Report, University of Manchester Institute of Science and Technology (2013).

[SRDB13] T. Subba Rao, S. Das, and G. Boshnakov. A frequency domain approach for the estimation of parameters of variograms of spatio-temporal stationary random processes. Technical Report, School of Mathematics. The University of Manchester, Submitted for publication. (2013).

[Ste05] Michael L Stein. Space–time covariance functions. *Journal of the American Statistical Association* **100**(469), 310–321 (2005).

[SW71] E. M. Stein and G. Weiss. "Introduction to Fourier analysis on Euclidean spaces". Princeton University Press, Princeton, N.J. (1971). Princeton Mathematical Series, No. 32.

[Whi53] P. Whittle. The analysis of multiple stationary time series. *J. Roy. Statist. Soc. Ser. B.* **15**, 125–139 (1953).

[Whi54] P Whittle. On stationary processes in the plane. *Biometrika* **41**(3/4), 434–449 (1954).