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

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

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 isotropic. The assumption of isotropy is a stronger assumption. The process is said to be isotropic if

$$c(\mathbf{h}, u) = c(\|\mathbf{h}\|, u), \quad \|\mathbf{h}\| \geq 0,\ u \in \mathbb{Z},$$

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

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

If the random process $\{Y_t(\mathbf{s})\}$ 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.

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

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

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 $Z_Y(\underline{\lambda}, \omega)$ is a zero mean complex valued random process with orthogonal increments with

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

where $dF_Y(\underline{\lambda}, \omega)$ is a spectral measure. If we assume further that $dF(\underline{\lambda}, \omega)$ is absolutely continuous with respect to In view of the orthogonality of the function $Z_Y(\underline{\lambda}, \omega)$, we can show that the positive definite covariance function $c(\mathbf{h}, u)$ has the representation

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

and by Fourier inversion, we have

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

where $d\mathbf{h}=\prod\limits_{i=1}^{d}dh_i$. From equation (5) we obtain

$$\begin{aligned}
c\left(0,u\right) &= \int\limits_{-\infty}^{\infty}\int\limits_{-\pi}^{\pi}e^{iu\omega}f\left(\underline{\lambda},\omega\right)\,d\underline{\lambda}\,d\omega=\int\limits_{-\pi}^{\pi}e^{iu\omega}g_0\left(\omega\right)d\omega, \\
c\left(\left\|\mathbf{h}\right\|,u\right) &= \int\limits_{-\pi}^{\pi}e^{iu\omega}\int\limits_{-\infty}^{\infty}e^{i\mathbf{h}\cdot\underline{\lambda}}f\left(\underline{\lambda},\omega\right)\,d\underline{\lambda}\,d\omega \\
&= \int\limits_{-\pi}^{\pi}e^{iu\omega}g_{\left\|\mathbf{h}\right\|}\left(\omega\right)d\omega
\end{aligned}$$

DFT, Discrete Fourier Transform

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

Let

$$\begin{aligned}
g_{\left\|\mathbf{h}\right\|}\left(\omega\right) &= Cov\left(J_{\mathbf{s}}\left(\omega\right),J_{\mathbf{s}+\mathbf{h}}\left(\omega\right)\right) \\
& \quad \int\limits_{-\infty}^{\infty}e^{i\left(\mathbf{h}\cdot\underline{\lambda}+u\omega\right)}f\left(\underline{\lambda},\omega\right)\,d\underline{\lambda}\,d\omega
\end{aligned}$$

be the covariance/spectrum function and, for example, let $g_{\left\|\mathbf{h}\right\|}\left(\omega\right)$ be of the form given by (**??**). Assume the function $g_{\left\|\mathbf{h}\right\|}\left(\omega\right)$ is characterized by the parameter vector ϑ. For convenience, we denote this covariance function by $g_{\left\|\mathbf{h}\right\|}(\omega,\vartheta)$.

Let us consider the case $\nu=1$. Then from the equation (**??**) we have

$$g_{\left\|\mathbf{h}\right\|}\left(\omega\right)=\frac{\sigma_e^2}{\left(2\pi\right)^2}\left(\frac{\left\|\mathbf{h}\right\|}{2\left|c\left(\omega\right)\right|}\right)K_1\left(\left|c\left(\omega\right)\right|\left\|\mathbf{h}\right\|\right), \tag{8}$$

and from the equation (**??**) we have

$$g_0\left(\omega\right)=\frac{\sigma_e^2}{2\left(2\pi\right)^2\left|c\left(\omega\right)\right|^2}, \tag{9}$$

which implies that $\left|c\left(\omega\right)\right|^2$ is proportional to $g_0^{-1}\left(\omega\right)$, which is defined as the inverse second order spectral density function of the process.

we can also obtain an expression for the auto-correlation function. We have the auto-correlation function when $d = 2$, and for all $\upsilon > 0$,

$$\begin{aligned}\rho\left(\|\mathbf{h}\|,\omega\right) &= \frac{g_{\|\mathbf{h}\|}\left(\omega\right)}{g_0\left(\omega\right)} \\ &= \frac{\left(\|\mathbf{h}\|\left|c\left(\omega\right)\right|\right)^{2\nu-1}}{2^{2\nu-2}\Gamma\left(2\nu-1\right)}K_{2\nu-1}\left(\left|c\left(\omega\right)\right|\|\mathbf{h}\|\right). \end{aligned} \tag{10}$$

we can also obtain an expression for the auto-correlation function. We have the auto-correlation function when $d = 2$, and for all $\upsilon > 0$,

$$\begin{aligned}\rho\left(\|\mathbf{h}\|,\omega\right) &= \frac{g_{\|\mathbf{h}\|}\left(\omega\right)}{g_0\left(\omega\right)} \\ &= \frac{\left(\|\mathbf{h}\|\left|c\left(\omega\right)\right|\right)^{2\nu-1}}{2^{2\nu-2}\Gamma\left(2\nu-1\right)}K_{2\nu-1}\left(\left|c\left(\omega\right)\right|\|\mathbf{h}\|\right). \end{aligned} \tag{11}$$

It is interesting to note that $\rho\left(\|\mathbf{h}\|,\omega\right)$ is in fact the coherency coefficient between two Discrete Fourier Transforms separated by the spatial distance $\|\mathbf{h}\|$ at the frequency ω.

1 Spatio-temporal Prediction

Our object in this section is to obtain an optimal predictor for $\{Y_t\left(\mathbf{s}\right); t = 1, 2, \ldots, n\}$ at the location $\mathbf{s}_0$ given the m spatial time series $\{Y_t\left(\mathbf{s}_i\right) | i = 1, 2, \ldots, m; t = 1, 2, \ldots, n\}$ from a spatio-temporal stationary, isotropic process $\{Y_t\left(\mathbf{s}\right)\}$. In other words, we are predicting the entire data set at the location $\mathbf{s}_0$. Using the predicted data at the location $\mathbf{s}_0$, we can obtain the optimal predictors for the future values$\{Y_t\left(\mathbf{s}\right); t = n + \nu, \nu \geqslant 0\}$ at the location $\mathbf{s}_0$. As in the case of the observed data $\{Y_t\left(\mathbf{s}_i\right)\}$, we define the discrete Fourier transform $J_{\mathbf{s}_0}\left(\omega\right)$ of $\{Y_t\left(\mathbf{s}_0\right)\}$,and predict the Fourier transform $J_{\mathbf{s}_0}\left(\omega\right)$ for all ω. Using the inverse Fourier Transform, we compute the predicted values of $Y_t\left(\mathbf{s}_0\right)$ for all $1 \leq t \leq n$. We pointed out earlier that there is a one to one correspondence between the discrete Fourier Transforms and the data. We have shown that if

$$J_{\mathbf{s}_0}\left(\omega\right) = \frac{1}{\sqrt{(2\pi n)}}\sum_{t=1}^{n} Y_t\left(\mathbf{s}_0\right)e^{-it\omega}, \tag{12}$$

then we have

$$Y_t\left(\mathbf{s}_0\right) = \sqrt{\frac{n}{2\pi}}\int_{-\pi}^{\pi} J_{\mathbf{s}_0}\left(\omega\right)e^{it\omega}d\omega. \tag{13}$$

Consider the vector of the discrete Fourier transforms obtained from all m locations at the frequency ω,

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

We note that

$$E\left[\underline{J}_m(\omega)\right] = 0,$$
$$E\left[\underline{J}_m(\omega)\,\underline{J}_m^*(\omega)\right] = F_m(\omega)\,, \tag{14}$$

wherethe real,$m \times m$ dimensional symmetric, positive definite square matrix $F_m(\omega) = \left(g_{\|s_i - s_j\|}(\omega)\,; i, j = 1, 2, \ldots, m\right)$, and each element $g_{\|s_i - s_j\|}(\omega)$ of the matrix $F_m(\omega)$ is given by (**??**). The complex random vector $\underline{J}_m(\omega)$ has a multivariate complex Gaussian distribution with mean zero and variance covariance matrix $F_m(\omega)$. Consider now the $(m+1)$ dimensional complex valued random vector,

$$\underline{J}'_{m+1}(\omega) = \left[J_{\mathbf{s}_o}(\omega)\,, \underline{J}'_m(\omega)\right].$$

It can be shown that the mean of the vector is zero, and the variance covariance matrix is given by

$$\begin{aligned} E\left[\underline{J}_{m+1}(\omega)\,\underline{J}^*_{m+1}(\omega)\right] &= \begin{bmatrix} E(J_{s_o}(\omega)\,J^*_{s_o}(\omega)) & E\left(J_{s_0}(\omega)\,\underline{J}^{*\prime}_m(\omega)\right) \\ E\left(\underline{J}_m(\omega)\,J^*_{s_0}(\omega)\right) & E\left(\underline{J}_m(\omega)\,\underline{J}^*_m(\omega)\right) \end{bmatrix} \\ &= \begin{bmatrix} g_0(\omega) & \underline{G}'_0(\omega) \\ \underline{G}_0(\omega) & F_m(\omega) \end{bmatrix}, \end{aligned}$$

where $g_0(\omega)$ is the second order spectral density function of the spatial process $\{Y_t(\mathbf{s}_0)\}$ and the row vector $\underline{G}'_0(\omega)$ is given by

$$\begin{aligned} \underline{G}'_0(\omega) &= E\left[J_{\mathbf{s}_o}(\omega)\,J^{*\prime}_m(\omega)\right] \\ &= [g_{\|s_0 - s_1\|}(\omega)\,, g_{\|s_0 - s_2\|}(\omega)\,, \ldots, g_{\|s_0 - s_n\|}(\omega)], \end{aligned}$$

and $F_m(\omega)$ is defined above. Therefore, the optimal linear least squares predictor of $J_{s_0}(\omega)$ given the vector $\underline{J}_m(\omega)$, is given by the conditional expectation

$$E\left[J_{\mathbf{s}_o}(\omega)\,|\,\underline{J}_m(\omega)\right] = \underline{G}'_0(\omega)\,F_m^{-1}(\omega)\,\underline{J}_m(\omega)\,, \tag{15}$$

and the minimum mean square prediction error is given by

$$\sigma_m^2(\omega) = g_0(\omega) - \underline{G}'_0(\omega)\,F_m^{-1}(\omega)\,\underline{G}_0(\omega)\,. \tag{16}$$

To predict the data $Y_t(\mathbf{s}_0)$ for all t, we use the inverse transform (13). In computing the predictor of $J_{\mathbf{s}_0}(\omega)$ using the expression (15) one usually replaces the elements of the matrices $\underline{G}_0(\omega)$ and $F_m(\omega)$ by their corresponding estimates and obtain

$$\widehat{Y}_t(\mathbf{s}_0) = \sqrt{\frac{n}{2\pi}} \int_{-\pi}^{\pi} e^{it\omega}\,\widehat{\underline{G}}'_0(\omega)\,\widehat{F}_m^{-1}(\omega)\,\underline{J}_m(\omega)\,d\omega. \tag{17}$$

It is interesting to compare the formulae (15) and (16) with the corresponding expressions obtained using the time domain approach (see [CW11], equations 6.49 and 6.50; [BCG14], equations 11.21 and 11.22). They are similar, but it is important to note that for evaluating the expressions given in [CW11], [BCG14], one requires the inversion of matrices of order $mn \times mn$, where as the evaluation of (ref:eq.5.3) given above only requires the inversion of $m \times m$ order matrices.

We note $E(Y_t(\mathbf{s}_0)) = 0$ and $Var(Y_t(\mathbf{s}_0)) \simeq \int_{-\pi}^{\pi} G_0^{'}(\omega) F_m^{-1}(\omega) G_0(\omega) d\omega$. We can show by an application of Parseval's Theorem

$$E\left(Y_t(\mathbf{s}_0) - \widehat{Y}_t(\mathbf{s}_0)\right)^2 = E\left|\mathcal{F}^{-1}\left(J_{\mathbf{s}_o}(\omega) - \widehat{J}_{\mathbf{s}_o}(\omega)\right)\right|^2$$
$$= \int_{-\pi}^{\pi} \sigma_m^2(\omega) d\omega. \tag{18}$$

In practice, the above integrals are approximated by finite sums of the form

$$\widehat{Y}_t(\mathbf{s}_0) = \sqrt{\frac{2\pi}{n}} \sum_{j=0}^{n-1} e^{it\omega_j} \underline{\widehat{G}}_0^{'}(\omega_j) \widehat{F}_m^{-1}(\omega_j) \underline{J}_m(\omega_j). \tag{19}$$

for all $t = 1, 2, \ldots n$,where the estimates $\underline{\widehat{G}}_0(\omega_j)$ and $\widehat{F}_m(\omega)$ are substituted for $\underline{G}_0(\omega)$ and $F_m(\omega)$ respectively. As noted earlier, the vector $\underline{G}_0(\omega)$ and the matrix $F_m(\omega)$ have covariance functions $g_{\|s_i - s_j\|}(\omega)$ as their elements.

1.1 Estimation of the Parameters of the Covariance function $g_{\|\mathbf{h}\|}(\omega, \vartheta)$ by Frequency Variogram (FV) Method.

Here we discuss briefly the FV methodology,The method here is based on the differences of DFTs.

We define a new spatio-temporal random process based on differences of the observed process $\{Y_t(\mathbf{s})\}$. Calculate the differences

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

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 of the new time series $\{X_{ij}(t) \,|\, i \neq j\}$ at the frequencies $\omega_k = 2\pi s_k/n, k = 1, \ldots, K$, where $s_k, k = 1, 2, \ldots K$ are set of integers and $\omega_j \pm \omega_k \neq 0 (\text{mod}\, 2\pi)$ for $1 \leq j < k \leq K$ (see [Bri01], [GKS12]). Let

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

Let $I_{X_{ij}}(\omega_k)$ be the second order periodogram of the time series $\{X_{ij}(t)\}$ given by

$$I_{X_{ij}}(\omega_k) = \left|J_{X_{ij}}(\omega_k)\right|^2. \tag{20}$$

Let $G_{X_{ij}}(\omega_k(n), \vartheta) = E(I_{X_{ij}}(\omega_k)$. The function $G_{X_{ij}}(\omega_k, \vartheta)$ is defined as the Frequency Variogram by [SRDB14]. It is very similar to the classical definition of spatio-temporal variogram $2\gamma(\mathbf{h},u)$ defined in section 2 (set $u=0$) in (1). The usefulness of FV as a measure of dissimilarity between two spatial processes and its further properties were discussed in a recent paper by the authors ([ST16])

From (20), we obtain

$$\begin{aligned} E\left[I_{X_{ij}}(\omega_k)\right] &= G_{X_{ij}}(\omega_k,\vartheta) \\ &= E\left[I_{\mathbf{s}_i}(\omega_k)\right] + E\left[I_{\mathbf{s}_j}(\omega_k)\right] - 2\,\mathrm{Re}\,E\left[I_{\mathbf{s}_i\mathbf{s}_j}(\omega_k)\right], \end{aligned} \tag{21}$$

where $I_{\mathbf{s}_i\mathbf{s}_j}(\omega_k(n))$ is the cross periodogram between the processes $\{Y_t(\mathbf{s}_i)\}$ and $\{Y_t(\mathbf{s}_j)\}$ and $I_{\mathbf{s}_i}(\omega_k)$ is the periodogram of the series $Y_t(s_i)$. For large n, it can be shown that for a stationary process $E\left[I_{\mathbf{s}_i}(\omega_k)\right] = E\left[I_{\mathbf{s}_j}(\omega_k)\right] = g_0(\omega_k;\underline{\vartheta})$ and for a stationary and an isotropic process $E\left[I_{\mathbf{s}_i\mathbf{s}_j}(\omega_k)\right] = g_{\|s_i-s_j\|}(\omega_k;\underline{\vartheta})$ which is real. Therefore, the expectation of (21) is given by

$$G_{(s_i,s_j)}(\omega_k;\underline{\vartheta}) = 2\left[g_0(\omega_k;\underline{\vartheta}) - g_{\|s_i-s_j\|}(\omega_k;\underline{\vartheta})\right], \tag{22}$$

It is interesting to compare $G_{(s_i,s_j)}(\omega_k;\underline{\vartheta})$ with spatio-temporal variogram $2\gamma(\mathbf{h},u)$ given by equation (2). The similarity between these two functions shows that one can use the Frequency variogram which is a frequency domain version of spatio-temporal variogram for estimating the effective range $||\mathbf{h}||$**,** and also the parameters etc.

Now for the estimation of the parameter vector $\underline{\vartheta}$ we proceed as in [SRDB14]. Consider the K-dimensional complex valued random vector,

$$\underline{F}_{\|s_i-s_j\|}(\omega) = \left[J_{X_{ij}}(\omega_1), J_{X_{ij}}(\omega_2), \ldots, J_{X_{ij}}(\omega_K)\right],$$

which is distributed asymptotically as complex normal with mean zero and with variance covariance matrix with diagonal elements $[g_{\|\mathbf{h}\|}(\omega_1,\vartheta), g_{\|\mathbf{h}\|}(\omega_2,\vartheta), \ldots, g_{\|\mathbf{h}\|}(\omega_K,\vartheta)]$,where $||h|| = ||s_i - s_j||$.We note that because of asymptotic independence of Fourier transforms at distinct frequencies chosen above, the off diagonal elements of the variance covariance matrix of the complex Gaussian random vector $\underline{\chi}_{\|\mathbf{h}\|}(\omega)$ are zero. Therefore, the minus of log likelihood function can be shown to be proportional to

$$Q_{n,N(\mathbf{h})}(\underline{\vartheta}) = \frac{1}{|N(\mathbf{h})|}\sum_{(\mathbf{s}_i,\mathbf{s}_j)\in N(\mathbf{h})}\sum_{k=1}^{K}\left[\ln G_{(\mathbf{s}_i,\mathbf{s}_j)}(\omega_k;\underline{\vartheta}) + \frac{I_{Xij}(\omega_k(n))}{G_{(\mathbf{s}_i,\mathbf{s}_j)}(\omega_k;\underline{\vartheta})}\right]. \tag{23}$$

Here $|N(\mathbf{h})|$ is the total number of all distinct pairs $\mathbf{s}_i$ and $\mathbf{s}_j$ such that $N(\mathbf{h}) = \{(\mathbf{s}_i,\mathbf{s}_j) \mid \|\mathbf{s}_i - \mathbf{s}_j\| = ||\mathbf{h}||\}$. The above criterion (23) is defined only for one distance $\|\mathbf{h}\|$. Suppose we now define L spatial distances from the observed data. We can now define an over all criterion for the minimization

$$Q_n(\vartheta) = \frac{1}{L}\sum_{l=1}^{L} Q_{n,N(\mathbf{h}_l)}(\underline{\vartheta}), \tag{24}$$

We minimize (24) with respect to $\underline{\vartheta}$ (for details refer to [SRDB14]). The asymptotic normality of the estimator $\underline{\vartheta}$ obtained by minimizing (24) has been proved in Theorem 2 of the paper of [SRDB14]. To avoid repetition, we refer to their paper for details. We state the asymptotic distribution of the estimates. It has been shown in [SRDB14] that under certain conditions, and for large n,

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

where $V = \lim_{n\to\infty} Var\left[\frac{1}{\sqrt{n}}\nabla Q_n\left(\vartheta_0\right)\right]$, $\nabla Q_n\left(\vartheta_0\right)$ is a vector of first order partial derivatives, $\nabla^2 Q_n\left(\vartheta_0\right)$ is a matrix of second order partial derivatives.

2 Algorithm

p=1 (16 előfordulás 24-ből)
q=0 (12/24) vagy q=1(11/24)
P=3 (18/24)
Q=0 (24/24)
d=0 (24/24)
D=1 (24/24)

Polynomials, AR, MA moving average polynomial is

$$\phi\left(z\right) = 1 - \varphi_1 z - \varphi_2 z^2 - \cdots - \varphi_p z^p,$$
$$\theta\left(z\right) = 1 + \vartheta_1 z + \vartheta_2 z^2 + \cdots + \vartheta_q z^q,$$

where z complex, $|z| = 1$ is the unit circle, i.e. $z = e^{i\omega}$.
no common factors of $\varphi\left(z\right)$ and $\vartheta\left(z\right)$

$$\Phi\left(z\right) = 1 - \Phi_1 u - \Phi_2 u^2 - \cdots - \Phi_P u^P,$$
$$\Theta\left(z\right) = 1 + \Theta_1 u + \Theta_2 u^2 + \ldots + \Theta_Q u^Q,$$

$u = z^S = e^{iS\omega}$
no common factors of $\Phi\left(z\right)$ and $\Theta\left(z\right)$
We are given a time series$\widetilde{Y}_t\left(\mathbf{s}_i\right)$ for each location $\mathbf{s}_i$.

1. First step, for each time series$\widetilde{Y}_t\left(\mathbf{s}_i\right)$ according to location $\mathbf{s}_i, i = 1, 2, \ldots, m$, use SARIMA model and estimate the order $\left(p, d, q, P, D, Q, S\right)_i$. Find the common order: $\left(p, d, q, P, D, Q, S\right)$.
   In our case $\left(p = 1, d = 0, q = 0, P = 3, D = 1, Q = 0\right)$

2. Apply the operator $\left(1 - B\right)^d \left(1 - B^s\right)^D$ on $\widetilde{Y}_t\left(\mathbf{s}_i\right)$, i.e.

   $$\left(1 - B\right)^d \left(1 - B^s\right)^D \widetilde{Y}_t\left(\mathbf{s}_i\right) = Y_t\left(\mathbf{s}_i\right),$$

   in case a trend/drift would be included, substruct the mean.
   In our case, use seasonal difference $D = 1$, $lag = 12$, diff(x, lag $= 12$), keep the first values for later inversion. Note the the drift will be eliminated since the later difference.

3. Use the above (p, q, P, Q, S), estimate the *common* parameters by Whittle likelihood.
   In our case, $(p, q, P, Q, S) = (p = 1, q = 0, P = 3, Q = 0)$

   (a) Calculate the distances between locations $\mathbf{s}_i$, $\mathbf{s}_j$ where $i < j$, they are denoted as length of differences $\|\mathbf{h}_{ij}\| = \|\mathbf{s}_i - \mathbf{s}_j\|$.

   (b) Calculate the differences

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

   and for all locations $\mathbf{s}_i$, $\mathbf{s}_j$ where $i < j$. Collect pairs $\mathbf{s}_i$ and $\mathbf{s}_j$, $(i \neq j)$ that belong to the set $N(\mathbf{h}_l) = \{\mathbf{s}_i, \mathbf{s}_j | \|\mathbf{s}_i - \mathbf{s}_j\| = \|\mathbf{h}_l\|, l = 1, 2, \ldots, L\}$. Note, the size of the matrix $[X_{ij}(t)]$, where $t = 1, 2, \ldots, n$, and $i < j$, is $n \times (m-1) m/2$.
   In our case each each location forms an individual set.

   (c) Finite Fourier transform of $\{X_{ij}(t) \,|\, i < j\}$ at the frequencies $\omega_k = 2\pi k/n_0$, $k = 0, 1, \ldots, n_0 - 1$, (see [Bri01], [GKS12]):

$$J_{X_{ij}}\omega_k) = \frac{1}{\sqrt{(2\pi n)}} \sum_{t=1}^{n} X_{ij}(t)\, e^{-it\omega_k}.$$

   Calculate the smoothed periodogram. Let $I_{X_{ij}}(\omega_k)$ be the second order periodogram of the time series $\{X_{ij}(t)\}$ given by

$$I_{X_{ij}}(\omega_k) = \left|J_{X_{ij}}(\omega_k)\right|^2. \tag{25}$$

   In our case

   (d) For given SARMA (p, q, P, Q, S) i.e. given $(\varphi_1, \varphi_2, \ldots \varphi_p)$, $(\vartheta_1, \vartheta_2, \ldots \vartheta_q)$, $(\Phi_1, \Phi_2, \ldots \Phi_P)$, and $(\Theta_1, \Theta_2, \ldots \Theta_Q)$

$$\phi(B)\, \Phi(B^s)\, Y_t = \theta(B)\, \Theta(B^s)\, W_t,$$

   Calculate $|c(\omega)|^2$

$$|c(\omega)|^2 = \left|\frac{\theta(z)\, \Theta(z^s)}{\phi(z)\, \Phi(z^s)}\right|^2$$

   $z = e^{i\omega}$, at the same frequencies as the periodogram.

   (e) Calculate $g_0(\omega_k; \underline{\vartheta})$ and $g_{\|\mathbf{s}_i - \mathbf{s}_j\|}(\omega_k; \underline{\vartheta})$ and

$$\begin{aligned} G_{(s_i, s_j)}(\omega_k; \underline{\vartheta}) &= 2\left[g_0(\omega_k; \underline{\vartheta}) - g_{\|\mathbf{s}_i - \mathbf{s}_j\|}(\omega_k; \underline{\vartheta})\right] \\ &= 2\left[\frac{\sigma_e^2}{2(2\pi)^2 |c(\omega)|^2} - \frac{\sigma_e^2}{(2\pi)^2}\left(\frac{\|\mathbf{h}_{ij}\|}{2|c(\omega)|}\right) K_1(|c(\omega)|\, \|\mathbf{h}_{ij}\|)\right] \\ &= \frac{\sigma_e^2}{(2\pi)^2 |c(\omega)|^2}\left[1 - \|\mathbf{h}_{ij}\|\, |c(\omega)|\, K_1(|c(\omega)|\, \|\mathbf{h}_{ij}\|)\right] \end{aligned}$$

   where $\mathbf{h}_{ij} = \mathbf{s}_i - \mathbf{s}_j$

(f) The exponent of the Whittle likelihod

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

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

Whittle estimator for ϑ

$$\arg\min Q_n\left(\vartheta\right).$$

4. Prediction

References

[BCG14] S. Banerjee, B. P Carlin, and A. E Gelfand, *Hierarchical modeling and analysis for spatial data*, Crc Press, 2014.

[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.

[CW11] N. Cressie and C. K. Wikle, *Statistics for spatio-temporal data*, Wiley Series in Probability and Statistics, 2011.

[GKS12] L. Giraitis, H. L. Koul, and D. Surgailis, *Large sample inference for long memory processes*, vol. 201/ 2, World Scientific, 2012.

[SRDB14] T. Subba Rao, S. Das, and G. Boshnakov, *A frequency domain approach for the estimation of parameters of spatio-temporal random processes*, Journal of Time Series Analysis **35** (2014), 357–377.

[ST16] T. Subba Rao and Gy. Terdik, *On the frequency variogram and on frequency domain methods for the analysis of spatio-temporal data*, ArXiv e-prints, 1610.05891 [math.ST] (2016), (submited for publication).